Essays on Development Economics

Nandita Krishnaswamy

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

ABSTRACT

Essays on Development Economics

Nandita Krishnaswamy

This dissertation consists of three empirical essays on agricultural incentives, risk, and rural labor markets.

Chapter 1 empirically estimates the effect of agricultural price support policies on crop choice and input (mis-)allocation, with important implications for spillover effects to other sectors. Agricultural price support policies are a popular way to alleviate the risk inherent in volatile prices, but, at the same time, may distort input allocation responses to agricultural productivity shocks across multiple sectors. This could reduce productivity in the agricultural sector in developing countries. I empirically test for misallocation in the Indian agricultural setting, with national price supports for rice and wheat. I first motivate the setting using a two-sector, two-factor general equilibrium model and derive comparative statics. I then use annual variation in the level of the national price supports for rice and wheat relative to market prices, together with exogenous changes in district-level agricultural productivity through weather shocks, in a differences-in-differences framework. I derive causal effects of the price supports on production patterns, labor allocation, wages, and output across sectors. I find that rice area cultivated, rice area as a share of total area planted, rice yields, and rice production all increase, suggesting an increase in input intensity (inputs per unit area) dedicated to rice. Wheat shows a similar increase in input intensity. The key input response is a reallocation of contract labor from the non-agricultural sector during peak cultivation periods, which results in an increase in wages in equilibrium in the non-agricultural sector (especially in response to price supports for the labor-intensive crop, rice, of 23%). The reallocation of labor reduces agricultural productivity by 82% of a standard deviation, and simultaneously reduces gross output in non-agricultural firms by 2.6% of a standard deviation. I also find that rice- and wheat-producing households do not smooth consumption more effectively in response to productivity

shocks in the presence of price supports.

Chapter 2 (with Emily Breza and Supreet Kaur) demonstrates the influence of collective action - specifically, through social sanctions imposed by informal labor unions - on labor supply in rural labor markets. A long tradition of work in social science posits that social norms affect labor market behavior. We use a field experiment to test whether community-wide norms against accepting wage cuts distort workers' labor supply during periods of unemployment. We undertake our test in informal spot markets for casual daily labor in India. We partner with 183 existing employers, who offer jobs to 502 randomly-selected laborers in their respective local labor markets. The job offers vary: (i) the wage level and (ii) the extent to which the offer is observable to other workers. On average, 26% of workers accept a job if it is offered at the prevailing wage, with no distinguishable differences by observability. In contrast, observability strongly mediates labor supply below the prevailing wage: while 18% accept work at a wage cut in private, this plummets to 4% when wage cuts are offered in public. The consequences of this behavior are substantial: workers are giving up 38% of average weekly earnings in order to avoid being seen as breaking the community norm. In a supplementary exercise, we document that workers are willing to pay to punish anonymous laborers who have accepted a wage cut. Costly punishment occurs both for workers in one's own village, and for workers in distant other labor markets—suggesting the internalization of norms in moral terms. Our findings support the presumption that collusive norms can develop even in the absence of formal labor institutions and can play a role in constraining labor supply behavior at economically meaningful magnitudes.

Chapter 3 investigates how households use engagement in criminal activity to smooth consumption in the face of agricultural risk. About 400,000 barrels of oil are stolen per day in the Niger Delta region. Much of this oil is stolen by militia groups with the help of local youth (who have the requisite knowledge about the terrain and placement of the pipelines). I use exogenous variation in households' access to oil pipelines, together with local shocks to agricultural productivity (both self-reported and due to variation in rainfall) to show that a proxy for theft from oil pipelines increases in the vicinity of households located close to pipelines that suffer unanticipated crop losses. This coincides with non-food expenditure-smoothing for these households (relative to households that are far from pipelines). Finally, I look at heterogeneity by household characteristics to identify households that are more likely to be affected by agricultural shocks or more likely to be targets for militia recruitment - households with young unemployed men and young men who are not in school, and households that lack financial infrastructure in their vicinity (which I take to be a proxy for a household's ability to access credit when faced with economic shocks). The findings from this paper suggest that there is potential for large spillover savings - in terms of reducing theft of oil from pipelines - for any policy that provides credit or other kinds of risk-mitigation mechanisms to households.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I am deeply grateful to my advisors and mentors, Eric Verhoogen, Emily Breza, and Supreet Kaur. Eric has challenged me to think broadly while keeping sight of the importance of methodological rigor, and I greatly appreciate his detailed feedback on many versions of this work. Emily and Supreet have provided mentorship from very early stages in my doctoral journey, and, with that, have taught me so much of what I know about good scholarship. I am very thankful for the opportunities they have given me to engage in fieldwork, and the patience they have demonstrated every step of the way.

The greater Columbia development community has been an incredible source of ideas and support. I have benefited so much from conversations with Jonas Hjort, Suresh Naidu, and Jack Willis. I thank Douglas Almond, Sakai Ando, Belinda Archibong, Ashna Arora, Cynthia Balloch, Iain Bamford, Michael Best, Varanya Chaubey, Anthony D'Agostino, Mark Dean, Golvine de Rochambeau, Tong Geng, Siddharth George, Francois Gerard, Siddharth Hari, Jonas Hjort, Adam Kapor, Pinar Keskin, Wojciech Kopczuk, Patrick McEwan, Suresh Naidu, Hailey Nguyen, Suanna Oh, Kiki Pop-Eleches, Qiuying Qu, Laura Randall, Daniel Rappoport, Saravana Ravindran, Debraj Ray, Bernard Salanie, Tobias Salz, Wolfram Schlenker, Stephanie Schmitt-Grohe, Yogita Shamdasani, Kartini Shastry, Jeff Shrader, Divya Singh, Rodrigo Soares, Matthieu Teachout, Danna Thomas, Lin Tian, Martin Uribe, Akila Weerapana, Scott Weiner, Jing Zhou, Jon Zytnick, and many others for valuable comments on parts of this work. I can only hope that I will give as much back to the community as it has provided me.

Our field team in Odisha, in India, has made all our fieldwork possible. I am grateful to Mohar Dey, Piyush Tank, Arnesh Chowdhury, Sayan Kundu, Shreoshee Mukherjee, Medha Aurora, Prasanta Nayak, Barada Behera, Santosh Kumar Sahoo, Ravi Kumar Rao, Saroj Kumar, and Deepak Kumar Nayak for their incredible hard work in running an efficient field team.

My friends from within and outside the program have been a constant source of support. Danna, Jing, and Lin, I am so glad that all of you were with me for each step we took towards earning our PhDs. From first-year study groups to post-defense celebrations, I have enjoyed every moment. To so many other friends from home and from Wellesley, and particularly Divya and Pooja, you always remind me that the work that I am doing matters, and not to lose sight of the world beyond Columbia.

My family, near and far, all figure in some way in the final product of this endeavor. My grandmother, Rajam, is perhaps the most important influence in my life and career. Kavita, Karuna, and Krishna have been my home away from home for the past six years, and I am so grateful for that. Smita and Gautam have always been a source of excellent advice or hilarious distraction whenever I have needed it. Above all, this dissertation wouldn't exist without the love and support of my parents, Radha and Mahesh. I owe them absolutely everything. Finally, my husband Mark has been a source of boundless love, laughter, and inspiration every single day. Thank you all!

Chapter 1. At What Price? Price Supports, Agricultural Productivity, and Misallocation

1.1 Introduction

Agricultural productivity in developing countries is low[1], and the productivity gap across sectors is large[2]. Simultaneously, farmers are unable to completely smooth consumption in response to shocks[3]. In response, a number of countries have adopted price support policies for various crops, in an effort to help farmers hedge against these risks[4]. However, prices on the open market, absent other frictions, are a mechanism for allocating inputs efficiently within the agricultural sector, and across sectors. We lack causal estimates of the effect of price supports on 1. distortions to farmers' production and input decisions, 2. total factor productivity in the agricultural sector, and 3. wages and output in the non-agricultural sector.

In this paper, I empirically study the extent to which price supports contribute to low agricultural productivity, and the productivity gap across sectors. I focus on the Public Distribution System in India, one of the largest such programs in the world. I look at the implications of price supports for farmers' crop choices, agricultural input selections, and decisions about non-agricultural work. I also study the resulting equilibrium effect on wages and output in both sectors. To do this, I interact weather-driven variation across space and time in local agricultural productivity (and therefore in local market prices) with changes in the level of the national-level price support in a differences-in-differences framework. I build a two-sector model of allocation decisions for capital and labor across the agricultural and non-agricultural sectors, with and without price supports. The model

[1]Kuznets (1971), Gollin et al. (2002), Caselli (2005), Restuccia et al. (2008), Chanda & Dalgaard (2008), Vollrath (2009), Lagakos & Waugh (2010), Gollin et al. (2011), Herrendorf & Schoellman (2011)

[2]Gollin, Lagakos, and Waugh (2011) estimate this to be 3.63 in the case of India

[3]Morduch 1995, Dercon 2002, Santangelo 2016

[4]Bangladesh, Brazil, Myanmar, Egypt, Indonesia, Mali, Pakistan, and Zambia, among other countries (World Bank Agricultural Distortions Database). The FAO finds that 27% of the 81 developing countries surveyed had price supports in place as of Jan 1, 2008.

describes the various channels through which prices mediate farmers' responses to agricultural productivity shocks and provides useful comparative statics.

There are two main reasons that India's price support policies are an effective context for testing the implications of such policies for farmers' decision-making. First, national price supports for rice and wheat[5] are announced in June at the beginning of each agricultural season, and are therefore known to farmers before planting. Second, there is variation between 1997 and 2012 - my time period of interest - in the extent to which the policy has kept up with local market prices, which provides important variation in the salience of the program to farmers[6][7].

First, I show that the support price is high in some years and low in others, relative to the entire predicted distribution of market prices for rice and wheat. This provides variation across the years in the probability that the price support will bind for a given district.

Second, each district's level of early-season rainfall serves as an exogenous, pre-planting, district-level shock to agricultural productivity. I verify that these local productivity shocks significantly affect the wholesale prices for rice and wheat that are eventually realized in the district; non-negative rainfall shocks (what I refer to in the paper as "good rain") lead to lower prices at harvest. So, it is clear that local market prices adjust in response to productivity shocks. There are two different distortions that price supports create in this environment; first, they allow farmers to sell output at a

[5]The Indian price supports are significant to farmers only for two staple crops, rice and wheat, in separate seasons. I discuss the implementation of these price supports in detail in Section 1.2.

[6]There are two, more minor, benefits to studying the Indian price support policy: First, this is a long-standing policy, with a single policy arm. The Indian government has provided price supports for staple crops since the 1970s, which reduces concerns that farmers are wary of the government reversing course on the price supports it announces at the time of planting, or that farmers need time to learn about the logistics of the policy. Second, the policy has shown little variation in the way that it is administered - eligibility criteria, key crops targeted, etc. - in the period I study.

[7]There are also advantages of assessing the impact of price supports in a developing country. Agricultural policies in developed countries (particularly in the US and across the EU), are often more nuanced than the Indian policy, and do not therefore provide an appropriate context for studying the direct influence of price supports on agriculture. They often involve a combination of income supports and quotas, do not apply in a blanket way to all farmers, and do not directly address price volatility. I also expect the responses of farmers to be very different in a context in which land-holdings tend to be smaller and more heavily focused on staple crops, farming is more labor-intensive, and farmers have less access to instruments such as futures contracts to address price volatility.

constant price (and not at the falling local market price) in response to positive productivity shocks, and second, in the case that they are set above a district's local market price, they provide an income shock that increases the marginal return to investing in agriculture relative to the non-agricultural sector. I consider both distortions together in this paper.

Importantly, both the level of the price support and the local early-season productivity shock are known to farmers before they make planting and input decisions.

To capture how responses to productivity shocks differ with and without price supports, I estimate the differential effect of "good rain", and therefore higher productivity, on various production metrics in years in which support prices are high relative to years in which they are low. Having determined that there is a positive effect on agricultural output and yield, I consider the effect of the policy on various inputs to agriculture, including labor, to identify the channels through which the production measures are affected. Third, I consider the effect of this input reallocation across sectors on productivity in the agricultural sector, and output in the non-agricultural sector. Finally, I study the differential effect of agricultural productivity shocks on household income for staple-producing households in high- and low- price support years as a measure of the income support provided by the policy.

There are four key results. First, the paper provides causal empirical evidence that price supports result in increased input intensity (amounts of input used per unit area) in the agricultural sector. I find that the Indian price support policy increases area, area share, yield, and production of rice. The increase in area and area share of rice suggest that farmers respond to the financial incentives of the price support by increasing the intensity of rice production. The increases in raw yield of rice further suggest increases in input intensity per hectare, beyond a simple reallocation of land towards a more input-intensive crop. I find a similar increase in yields and input intensity for wheat[8]. These production gains are restricted to districts that are relatively suitable for rice and

[8]There is no change in total area cultivated in the *Rabi* season- the main wheat-producing season - in response

wheat respectively.

The second key result is that these increases in area (for rice), yield and production (for both crops) coincide with a reallocation of labor from the non-agricultural to the agricultural sector, particularly during peak cultivation periods (when the marginal returns to investing labor in production are highest). I confirm that this reallocation is driven by contracted (short-term) employees rather than permanent employees of non-agricultural firms. For a sense of magnitude, among agricultural households, this is a decrease in days engaged in non-agricultural labor of 35% for rice and 19% for wheat. I find no effect of price supports on labor supply on the extensive margin, or other inputs. I turn to the model for the intuition behind these results. According to the two-sector, two-factor general equilibrium model I build, there are two competing effects of higher productivity in the agricultural sector on labor use in the *absence* of price supports: first, that a lower relative price for agricultural goods (and the resulting income effect) leads to increased demand in the non-agricultural sector and a reallocation of inputs away from agriculture[9], and second, that higher relative productivity in agriculture puts upward pressure on wages and results in a reallocation of labor into agriculture. Price supports partly negate the first channel, leaving the second to dominate.

Taken together, the results show crowding out effects in the non-agricultural sector as a result of the distortion in agricultural prices. I confirm this by analyzing output in the formal manufacturing sector, and find that it falls by 8.5% in years in which price supports are high, in response to positive productivity shocks in agriculture. The paper therefore provides initial evidence on the ability of price support policies to slow the growth of the (more productive) non-agricultural sector in a transition economy[10]. In addition, the loss in manufacturing output amounts to 0.83% of

---

to higher price supports, nor in the area or area share dedicated to wheat. However, even as area remains constant, production and yield both see significant increases, suggesting a similar increase in input intensity as for rice.

[9]Similar reasoning has been developed in models by Murphy et al. 1989, Kongsamut et al. 2001, and Gollin et al. 2002.

[10]The literature suggests that non-farm growth is key to increasing rural wages and reducing rates of poverty. In

India's GDP, which, when taken into account, effectively doubles the implicit cost of these price supports.

More broadly, these results can be extended to intuit the effect of increasing agricultural market integration (and therefore a single price across districts, in the extreme case) in developing economies on the sectoral allocation of inputs. In a world without market integration, increased productivity in the agricultural sector through a local rainfall shock reduces local market prices and strengthens the reallocation of inputs away from agriculture. With market integration, prices are inelastic to local productivity shocks, and farmers behave as they would when exposed to agricultural price supports.

Next, I ask whether the reallocation of labor into the non-agricultural sector can, in fact, reduce agricultural productivity. In accordance with the literature, I construct a Tornqvist-Theil index of agricultural TFP, aggregating across crops and across various inputs. I find a 0.82 standard deviation decrease in this measure of agricultural productivity in response to a positive agricultural productivity shock when price supports are high relative to when they are low. This is driven by the increase in labor use in the agricultural sector. This result, together with the crowding-out effects in the non-agricultural sector, suggests that not only do price supports policies hinder growth in the non-agricultural sector - they also have a negative impact on productivity *within* agriculture.

Finally, this paper also examines whether a price-support policy can provide income support in an environment in which prices run counter to productivity shocks and serve as an automatic stabilizer for income. Agricultural price support policies pay out when prices are low but production is high. In the case of India's price support program, I find that price supports do not improve consumption-smoothing in response to productivity shocks.

This paper contributes to the literature on the link between agricultural productivity and the growth

rural India, in particular, growth in the non-agricultural sector has been rapid, and has contributed more than double to rural growth than the use of agricultural technologies such as high-yielding varieties of seeds (Foster & Rosenzweig 2003).

of the non-agricultural sector (Bustos et al. 2012, Hornbeck & Keskin 2014, and many others). Studies based in India conclude that the factor bias of the productivity shock drives the direction of the effect on the non-agricultural sector[11]. Specifically, I examine the short-run effect of Hicks-neutral agricultural productivity shocks, driven by rainfall[12], on labor allocation and output in the non-agricultural sector, and then ask how these are affected by agricultural price supports. Studies on the effects of such rainfall shocks on the non-agricultural sector have identified two channels through which the sectors are related: (1) wages and (2) relative prices and demand (Lee 2014, Emerick 2016, Santangelo 2016). These studies find that the latter channel is stronger in the case of India, leading to labor movements out of agriculture in periods of good rainfall, which I confirm in this paper. In addition, as a contribution to this literature, this study is the first to separately identify the contribution of the producer price channel to this effect. I find that, in the presence of worker mobility[13], price supports simultaneously reduce wages for agricultural workers and increase the fraction of workers in agriculture, and reduce output and employment in the non-farm sector.

A second literature supports the idea that risk may have a significant impact on agricultural production. We know, for instance, that missing markets for insurance in many developing countries affect crop choice. Farmers continue to face shocks to output and prices, but lack access to financial instruments that could hedge against risk. Small farmers do not typically enter into futures contracts, and index insurance (that hedges against weather shocks) remains rare (Cole et al. 2009, Binswanger-Mkhize 2011)[14]. Their decisions about what to plant are therefore distorted by risk.

---

[11]Studies on the Green Revolution in India have found a negative relationship (over the long term) between labor-augmenting technological progress and output and labor allocation to the manufacturing sector in India (Foster & Rosenzweig 2004, Moscona 2017), while studies on short-term responses to rainfall shocks, assumed to be Hicks-neutral, have found the opposite (Emerick 2016, Santangelo 2016).

[12]There is an expansive related literature on the effect of rainfall shocks on agricultural inputs, including labor (Jayachandran 2006, Kaur 2017, and others), which suggests that rain is important for agricultural productivity.

[13]Prior work finds that, in the Indian context, there is a great deal of short-term movement of labor between sectors (Imbert & Papp 2015, Colmer 2017, and others). Workers are often engaged on a daily or weekly basis, and, even among those who are engaged primarily in agriculture, devote some time to non-agricultural activities.

[14]In the 2012-2013 agricultural cycle, 95% of rice- and wheat-producing households did not insure their staple crop.

There are two types of empirical work within this literature. First, farmers without access to insurance products tend to use production decisions to hedge against risk, even at the cost of expected income (Rosenzweig & Stark 1989, Fafchamps 1992, Morduch 1995, Dercon 1998, 2002, Dercon & Christiaensen 2011, Falco et al. 2014). Second, farmers diversify into more risky crops and invest more in inputs following the provision of various types of insurance (Karlan et al. 2014, Gehrke 2014, Cole et al. 2017), and large-scale government transfer programs (e.g. workfare programs, social transfers) (Bhargava 2015, Gehrke 2017)[15].

I contribute to this literature by examining the production and the labor allocation responses to a specific policy-driven reduction in *price* volatility in the agricultural sector. There are two ways in which this paper differs from the insurance and agricultural production literature. First, there is little existing evidence of the price support policy's effectiveness as an income support - this is because, unlike insurance, it pays out at times when lower prices might be offset by higher output. Second, experiments involving insurance tend to occur on a smaller scale. This price support policy covers all farmers in India, and my findings show that the aggregate effect (that cannot be studied through experiments) on labor allocation across sectors and non-agricultural outupt is large.

A third strand of literature deals with the direct effects of price volatility on farmers. Allen & Atkin (2016) find, for example, find that farmers shift towards less risky crops in the presence of increased income volatility (and decreased price volatility) in response to reduced trade costs[16]. This paper adds to this literature by using clear policy variation to assess the effect of price supports that are meant solely to alleviate price volatility, but which are themselves focused on staple (less risky) crops. This is in contrast to examining the impact of reducing price volatility through trade, in which there is wide-ranging impact on outcomes ranging from market access to input availability,

[15]In addition, these government policies have been shown to have labor market effects in similar contexts (Ardington et al. 2009, Basu et al. 2009, Azam 2012, Berg et al. 2013, Santangelo 2016), in particular by increasing non-agricultural wage rates.

[16]In the form of expansions in the highway network

rather than only on price volatility in a single sector.

A fourth strand of literature looks specifically at the effects of price supports in the agricultural sector, but does so by simulating an artificial price support as part of a structural model (Jonasson et al. 2014, Mariano & Giescke 2014). This paper adds to this literature by estimating the concrete effect of a particular price support policy, rather than making the various requisite assumptions for a structural estimation.

The paper proceeds as follows: Section 1.2 provides background on the agricultural sector and institutional details about the timing of the policy that drive my empirical strategy. Section 1.3 provides a two-sector model of input allocation and derives useful comparative statics. Section 1.4 presents my empirical strategy and validates that early-season rainfall affects realized market prices in the harvest period, which implies that it influences farmers' expectations of prices. Section 1.5 details how I aggregate information on prices, crops produced, area for each crop, production, yield, farmers' expenditure at harvest, and rainfall into a district-level panel for the time period 1997-2012. Section 1.6 presents results and a discussion of the broader implications of my findings. Section 1.7 provides a numeric estimate of the effects of the price support on agricultural productivity. Section 1.8 presents various robustness checks to validate my results, and discusses potential confounds. Section 1.9 concludes.

1.2 Background and Context

1.2.1 Agriculture in India

Indian agriculture is characterized by small-holder farmers (1-2.5 acres) who typically plant 1-2 crops each year[17]. The Green Revolution of the 1960s resulted in large increases in the use of high-yielding varieties and complementary inputs like fertilizer, even among small farmers.

[17] NSS rounds 55-68.

However, levels of technology investment remain low. Agricultural households commonly produce staples, and consume a significant proportion of their output[18]. They sell the rest of their produce either directly to wholesale markets *(mandis)* within their districts, or to middlemen who aggregate produce and sell it in the market.

There is strong evidence that local markets (at the level of the district, for example) are not well-integrated, because of which the effects of local weather shocks on prices are not completely arbitraged across districts[19].Transportation costs, the short shelf-life of most agricultural produce, and varied tastes for particular produce across states all result in large amounts of price variation between states, and even across districts within the same state[20][21]. Prices in the wholesale market are set using a system of first-price auctions and can, in some cases, involve brokers who facilitate sales.

Farmers and middlemen tend to transport their produce only to the nearest market, leading me to characterize them as price-takers from their own district's wholesale market in this context [22].

Price supports were introduced well before the time period over which I conduct my analyses, so I do not anticipate a "learning period" in my data in which farmers discover and begin utilizing the program. The Indian government has a long history of price support policies. Price supports for staple products began in 1972, as production boomed and prices began to fall. I only analyze the effects of the price support after 1997, when the consumption-side of the program underwent a

[18] NSS round 70

[19]I quantify the impact of local shocks on local market prices in Section 1.4.2.

[20]Within-year within-state standard deviation in wholesale prices averages Rs. 143 per 100 kgs of rice and Rs. 108 per 100 kgs of wheat.

[21]There are numerous regulatory barriers to inter-state movement of agricultural produce (Kohli & Smith 2003, Gulati 2012, Allen 2013) that also contribute to this price dispersion.

[22]Despite the lack of market integration described above, we can assume that there are at least some producers in any given district that are able to transport and sell their produce in a neighboring district. To the extent that this small group of farmers has the alternative option of selling in another district at a higher price than in their own (or the MSP), they are less likely to respond to the policy, and will dampen the magnitude of the effect that I find.

major overhaul that included, for the first time, targeted subsidies.

1.2.2 Setting and Implementing Minimum Support Prices

Support prices are announced at the beginning of each agricultural year, prior to planting, and paid at harvest. In June each year, at the time of the early annual monsoon, the Committee for Agricultural Costs and Prices (CACP) of the Central government announces a slate of national Minimum Support Prices (MSPs) for up to 25 crops[23]. However, government procurement at the MSP is a viable alternative only in the case of rice and wheat, which have both been procured at rates of higher than 15% since 1997[24] (Figure 1.1). As of the 2011-2012 season, procurement of rice and wheat stands at 40.2% and 39.7% of total production respectively.

Support prices are set independently for the two main cultivation seasons in the country, the Winter *Kharif* season (the main rice season[25]), and the Spring *Rabi* season (the main wheat season[26]), and paid out only the harvest period pertaining to that season[27]). The two seasons are distinct: either the rice support price is in effect, or the wheat support price is in effect, and not both[28]. At baseline,

[23]Data on Minimum Support Prices Recommended by CACP and Fixed by Government (Crop Year)

[24]In theory, farmers can sell any of 25 crops to various government *mandis* during the harvest. In practice, however, the price support policy focuses heavily on staples, particularly in the period between 1997-2012, the relevant period of study for this paper. In the case of pulses, for example, for which MSPs are regularly announced, under 1% of production is procured (Bhattacharya 2016). For cotton, a key cash crop, the proportion of procurement stands at a low 7% ("Cotton procurement at 2-2.5 million bales", Nov 14th 2015, Business Standard, and data from the Cotton Corporation of India http://cotcorp.gov.in/statistics.aspx).

[25]Planting in June-July, harvesting in Dec-Jan

[26]Planting in Nov-Dec, harvesting in Feb-Apr

[27]For example, the MSP for rice applies only between January and March for the main rice harvest from the *Kharif* season, while the MSP for wheat is effective in the *Rabi* season.

[28]Apart from the fact that the rice MSPs are only paid for harvesting in the *Kharif* season and the wheat MSPs are only paid for harvesting in the *Rabi* seasons, there are only 35 districts that cultivate both rice and wheat in the *Rabi* season (of which only 4 are significant wheat producers) and only 7 districts that produce both rice and wheat in the *Kharif* season (of which none are significant wheat producers). It is therefore unlikely that the support prices for rice and wheat intersect in decision-making within a season. However, farmers may certainly substitute production across seasons, since support prices are known before the earlier of the two seasons (the *Kharif* season) begins. I show this mitigating effect in Section 1.8.8.

we assume that farmers are aware of these prices before they make planting decisions (which occur 2-3 weeks after the main monsoon).

I consider the MSP-setting process to have elements of randomness from the perspective of the farmer, for four reasons, and these in turn validate the parallel trends assumption in the differences-in-differences framework I use. First, the precise algorithm that is used to set prices is not public knowledge, and certainly not known to the potential beneficiaries of the price supports [29]. Second, in addition to the information observed and taken into account by the national government, MSPs are set through a political process that introduces some randomness. There is a clear sense that political pressure sets ever-increasing MSPs[30]. Third, it is unlikely that there is meaningful district-specific information encoded into the national MSP announcement that was previously unknown to farmers in that district that could directly influence production decisions. Fourth, I verify that price supports do not correlate with various other metrics that are observable to farmers that may affect production: aggregate early-season rainfall (productivity) shocks[31] and monsoon forecasts

[29]While State governments provide recommendations to the CACP, the committee takes into account a wide range of information, including cost surveys from around the country and monsoon forecasts. The CACP describes the following considerations in setting these price supports (Terms of Reference, CACP 2009):

1. Cost of production, elicited through surveys, 2. Demand and supply, 3. Domestic and international price trends, 4. Inter-crop price parity, 5. Terms of trade between agriculture and non-agriculture, and 6. Likely implications of MSP on consumers of that product.

[30]MSPs have continued rising steeply in recent years and have never fallen in their entire history, even in periods in which world prices for rice and wheat are falling. I assume, therefore, that individual districts have no influence in setting the national MSP, once state-time trends are accounted for. State governments or the Central government sometimes announce surprise bonuses to the MSP, which are unknown to the farmer at the time of planting (and therefore do not factor into planting decisions).

[31]Since the monsoon begins in the earliest states in late May and announcement of price supports is made in June, it is possible that the aggregate of local-level rainfall shocks across the country is taken into account in setting the support price in the Kharif season, but I show that this is not the case. To do this, I test whether early-season rainfall across the country is predictive of the minimum support price (both in levels and first-differences) for rice and wheat, and find that it is not. Figure 1.9 indicates an increasing trend for real support prices over time for both rice and wheat, despite low (for example, 2012) and high (e.g. 2008) early-season rainfall realizations. Figure 1.10 shows that changes in support prices also show no consistent pattern in response to early-season rainfall. Second, even if there were such a pattern, it would not pose a threat to identification. I rely on local-level variation in early-season rainfall around the national average by including year fixed-effects in my specifications. This implies that changes in the national-level price support, even if based on some aggregate measure of early-season rainfall, are still random from the perspective of the individual farmer in a particular district. This does not pose a problem in the *Rabi* season since the announcement of support prices takes place in June, while early season *Rabi* rainfall only begins to be realized in

(Appendix Table 1.15).

The government serves as an alternative buyer for agricultural output at harvest, setting an effective (but not legislated) price support. At harvest-time, State and Central governments set up *mandis* in which any farmer can sell their harvest directly to government officials at the previously-announced MSP. At the time of the harvest, farmers observe realized prices in the wholesale market and make a decision about whether to sell their crops at the government *mandi* or at the wholesale market, taking into account transportation costs to both.

Since governments do not legislate a price floor, farmers often experience local market prices that are below the MSP in local wholesale markets in some years, but not in others. There are two main reasons I identify for continuing to observe prices below the MSP in some wholesale markets in some years: 1. Not all farmers are aware of the MSPs that have been set, and, as such, those producers do not consider the government price support policy in their planting or selling decisions[32] 2. Even among those who are aware of the MSP while making their production decisions, some might find that the additional transport cost required to take produce to the government *mandis* is too high, and therefore remain non-compliers[33]. It is this population of non-compliers and those with imperfect information who participate in the local wholesale market, in which I observe prices[34].

The policy's focus on rice and wheat is the result of the government's overall goal to procure staples and redistribute it at a single subsidized price to low-income households through a network

---

September. Nevertheless, I present evidence that support prices do not depend on early-season rainfall realizations in both seasons.

[32]Data from the 70th round of the NSS suggests that only 32% of rice-producing households accurately know the current MSP level for rice (39% for wheat). 12 % of rice-producing households and 16% of wheat-producing households reported sales to the government through the PDS system.

[33]Access to government mandis varies widely across districts and states, resulting in uneven access to price supports. I discuss this issue further in Section 1.8.6.

[34]There are, of course, operational constraints to accessing government *mandis* that extend beyond distance. These include the operational hours of *mandis*, potential bribes that need to be paid for the produce to be accepted, and overcrowded warehouses - all of which narrow the complier population and dampen the effect of the policy on producers. I discuss the implications of these in further detail in Section 1.8.

of close to 500,000 ration stores across the country[35]. This paper focuses only on production responses to the support price, and assumes that the consumption side of the program does not vary systematically with production-side factors in the period of study[36].

1.2.3 The Farmer's Timeline

I gather the details above into a timeline outlining the implementation of the policy for a representative state (Figures 1.2 and 1.3).

There are three key takeaways from the timing of implementation. First, the MSP is known (without uncertainty) when planting decisions are being made, and can influence planting and input decisions. Second, early-season monsoon rain is observed before planting occurs, and shocks to early-season rain reflect shocks to agricultural productivity. Third, farmers may form expectations of yield and market price based on monsoon rains, but these remain stochastic at the time of planting.

1.3 Two-Sector Framework

In this section, I present a two-sector model of allocation of capital and labor between agricultural and non-agricultural production. The model makes several simplifications to the context, but is used to provide useful comparative statics of farmers' responses to productivity shocks arising from local-level rainfall variation, both with and without price supports.

---

[35]Unlike the production-side price supports, which are available to all farmers, regardless of land-holding, the consumption-side subsidies are targeted toward poorer households. Rice and wheat are sold at a subsidized price (always below market retail prices) to people who hold Below-Poverty Line (BPL) cards (38% of the rural population), and at an even lower price to the ultra-poor. Consumer prices through the program are set at the national level also.

[36]The resale of rice and wheat procured by the government through ration stores may directly affect farmers' production choices, so I restrict my analyses to a time period (1997-2012) in which there are no major changes in administration and selection of beneficiaries by the government on the consumption side of the program. I discuss potential interactions between the production and consumption sides of the program in greater detail in Section 1.8.

I make the following simplifying assumptions in creating the framework. In Section 1.3.6, I discuss relaxing these assumptions.

1. That each district behaves like a small closed economy.

2. That within the agricultural sector, a single crop is produced with a single price, and that the price support (when I introduce it) applies to that one crop. This assumption allows me to focus on inter-sectoral labor shifts.

3. That realized prices are known with certainty immediately following the productivity shock - that is, that households observe early-season rainfall, and know the local market price for the agricultural good precisely.

4. That capital and labor are completely mobile across sectors.

1.3.1 Household Utility Maximization

I begin with a version of the framework without price supports. A representative household h earns income I from renting a stock of capital, K, and labor L, at rates r and w respectively. In turn, the household consumes two goods, an agricultural good and a manufacturing good. It maximizes a standard CES utility function[37] subject to a budget constraint (without credit). That is, the household chooses q_M and q_A to maximize:

$$U_h = [\alpha q_A^{\frac{\sigma-1}{\sigma}} + (1-\alpha) q_M^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \text{ s. t. } p_A qA + q_M = I,$$

where p_A is the price per unit of the agricultural output, and the price of the manufacturing good is normalized to 1.

Household optimization then satisfies the following conditions:

[37]I also show that Cobb-Douglas Stone-Geary preferences, also common to the literature, provide an even more stark version of the comparative statics derived here, due to stronger income effects arising from a subsistence constraint (derivation available upon request). Prior literature (Restuccia et al. 2008, Herrendorf 2013, Lee 2014) suggests that C-D Stone-Geary preferences better model the cross-country variation in responses to agricultural productivity shocks.

$$\frac{G\alpha q_A^{\frac{-1}{\sigma}}}{p_A} = G(1-\alpha)q_M^{\frac{-1}{\sigma}} \tag{1}$$

[38]

and

$$q_M = I - p_A q_A \tag{2}$$

I combine equations 1 and 2 to derive the optimal quantity consumed of the manufacturing good:

$$q_M = \frac{(1-\alpha)^{\sigma} I}{\alpha^{\sigma} p_A^{1-\sigma} + (1-\alpha)^{\sigma}} \tag{3}$$

[39]

[38]Where $G = [\alpha q_A^{\frac{\sigma-1}{\sigma}} + (1-\alpha)q_M^{\frac{\sigma-1}{\sigma}}]^{\frac{1}{\sigma-1}}$

[39]While we use equation 3 in deriving the general equilibrium in this model, an intuitive way to think about the equilibrium arises from examining quantity shares for the two goods:

$$\frac{q_M}{q_A} = \left(\frac{\frac{1-\alpha}{p_M}}{\frac{\alpha}{p_A}}\right)^{\sigma} \tag{4}$$

The representative household consumes according to the (price-weighted) ratios of the importance of each good in the utility function, downweighted by the substitutability between the goods. The higher the substitutability between the goods (higher σ), the closer the household gets to consuming only one good. If the goods are perfect complements, on the other hand, the goods will be consumed exactly in a 1:1 ratio.

1.3.2 Producers' Profit Maximization

In the firms' maximization problem, I make standard assumptions of perfect competition and profit maximization among producers in both sectors. Capital and labor are perfectly mobile across sectors and priced at r and w respectively.

Firms in both sectors possess a Cobb-Douglas production technology:

$$y_i = z_i K_i^{\beta_i} L_i^{1-\beta_i},$$

where the z's are industry-specific productivity factors, where the returns to capital in the non-agricultural production function are higher ($\beta_M > \beta_A$).

Firms in each sector $i = M, A$, facing input prices r, w, choose K_i, Li, to maximize profits:

$$\pi_i = p_i z_i K_i^{\beta_i} L_i^{1-\beta_i} - wL_i - rK_i, \tag{5}$$

where p_i represents the price of the output of sector i, and p_M, the price of manufacturing goods, is normalized to 1.

First-order conditions (FOC) from the firms' maximization problem give:

$$\beta_i p_i z_i (\frac{K_i}{L_i})^{\beta i-1} = r \tag{6}$$

and

$$(1-\beta_i)p_i z_i(\frac{K_i}{L_i})^{\beta_i} = w \tag{7}$$

for each of $i = M, A$.

The first-order conditions can be rearranged to express agricultural price as a function of inputs in manufacturing, and K_M as a function of L_M:

$$p_A = \frac{\beta_M z_M}{\beta_A z_A}(\frac{K_M}{L_M})^{\beta_M-1}(\frac{K-K_M}{L-L_M})^{1-\beta_A} \tag{8}$$

$$K_M = \frac{(1-\beta_A)\beta_M K L_M}{(1-\beta_M)\beta_A L + (\beta_M - \beta_A)L_M} \tag{9}$$

1.3.3 Equilibrium Without Price Supports

First, at equilibrium, the amount of manufacturing output must equal the manufacturing output consumed:

$$y_M = q_M = \frac{E_M}{p_M} = E_M \tag{10}$$

Second, the total capital and labor stock in the economy should be distributed among the sectors[40]:

$$K = K_M + K_A \tag{11}$$

$$L = L_M + L_A \tag{12}$$

Third, the total price of capital and labor used (in the firms' maximization problem) should equal total household income:

$$I = wL + rK \tag{13}$$

Taking FOC from both utility and profit maximization problems, and the equilibrium conditions detailed above, I express the optimal labor allocation to manufacturing, L_M, in an implicit function of the price for the agricultural good, p_A[41] :

$$L_M[\kappa_1 p_A^{1-\sigma} + \kappa_2] - \kappa_2 L = 0 \text{ [42]} \tag{14}$$

where

[40]With these preference structures and production technologies, it is clear that utility- and profit-maximization require the entire capital and labor stock in the economy to be utilized.

[41]A complete derivation is provided in the Model Appendix.

[42]Where $\kappa_1 = (\frac{\alpha}{1-\alpha})^\sigma$ and $\kappa_2 = \frac{1-\beta_M}{1-\beta_A}$.

$$p_A = \frac{\kappa_3 z_M}{z_A}[\frac{K}{(1-\beta_M)\beta_A L + (\beta_M - \beta_A)L_M}]^{\beta_M - \beta_A}\text{[43]} \tag{15}$$

At first glance, it is clear that p_A mediates the relationship between agricultural productivity, z_A, and labor in manufacturing, L_M, in equilibrium. I discuss this in further detail in Section 1.3.5.

1.3.4 Production, Consumption, and Equilibrium with Price Supports

I next turn to the case in which a price support is in effect for the agricultural good. That is, I assume the government purchases as much of the agricultural output as farmers want to sell at the support price p_S, and sells as much of the output as consumers demand at $p_C < p_S$[44]. In this case, the government also absorbs and over- or under-production in the agricultural sector, which implies that in a general equilibrium solution, local agricultural output need not equal consumption.

Demand for manufacturing goods now responds to consumer prices p_C:

$$q_M = \frac{(1-\alpha)^\sigma I}{\alpha^\sigma p_C^{1-\sigma} + (1-\alpha)^\sigma} \tag{16}$$

Firms' profit-maximization determines that the optimal ratio of capital to labor in both sectors is mediated by the agricultural producer price:

[43]Where $\kappa_3 = \frac{\beta_M[\beta_A(1-\beta_M)]^{1-\beta_A}}{\beta A[\beta_M(1-\beta_A)]^{1-\beta_M}}$.

[44]This can easily be extended to the small open economy case by setting $p_C = p_S$.

$$p_S = \frac{\beta_M z_M}{\beta_A z_A}(\frac{K_M}{L_M})^{\beta_M - 1}(\frac{K - K_M}{L - L_M})^{1-\beta_A} \tag{17}$$

I also note that the relationship between capital and labor in manufacturing remains unchanged in this context (equation 9). I therefore substitute equation 9 into equation 17, and obtain a relationship between agricultural productivity, labor in manufacturing, and the level of the price support:

$$L_M = \frac{(\frac{\kappa_4}{p_S z_A})^{\frac{1}{\beta_M - \beta_A}} - \kappa_5}{\beta_M - \beta_A} \tag{18}$$ [45]

In the case with price supports, the producer price in the agricultural sector, p_S, continues to figure in the relationship between non-agricultural labor allocation and agricultural productivity, but is a constant.

1.3.5 Comparative Statics

1.3.6 Assumptions

The framework outlined above is clearly a simplification of the Indian support price policy. The key differences between the model and the execution of the policy are as follows:

First and most importantly, Indian districts do not exist entirely in either the regime with, or without price supports. If realized prices are sufficiently high, price supports do not bind and we can expect that the district behaves according to the base case. If realized prices are low, then the district produces according to the binding support price. Adding a layer of complication to this

[45] Where $\kappa_4 = [(1-\beta_A)\beta_M]^{\beta_M - 1}[(1-\beta_M)\beta_A]^{1-\beta_A} K^{\beta_M - \beta_A}$ and $\kappa_5 = (1-\beta_M)\beta_A L$.

is the fact that farmers do not know, at the time of planting, whether the price support will bind. Farmers in each district can only estimate a probability that they will fall under one regime or another. Therefore, based on these probabilities, districts fall on a continuum between the two models.

In light of this, we expect a decrease in the amount of labor allocated to the non-agricultural sector in high price-support years for two reasons. First, the probability that the price support will bind at harvest for a given district is higher (therefore there is an increased chance of being in the price support case). Second, the level of the price support is higher relative to local prices, which we have shown to amplify the non-agricultural labor response. Combined, both these effects suggest that the differential response of non-agricultural labor allocation to an agricultural productivity shock in high- and low-price support years will be negative.

Second, capital and labor are not, in reality, perfectly mobile across sectors. Relaxing this assumption (in the extreme, this would mean that there are separate labor stocks for agriculture and manufacturing) should imply that competition for labor in the agricultural sector in response to positive productivity shocks drives up wages w_A, and demand, prices, and wages in the manufacturing sector, q_M and w_M. There should be smaller labor movements between sectors, both in the base case and with price supports.

Third, not every farmer is a complier - either because of lack of knowledge of the government program, or due to high transport costs to government depots. Because of this the general equilibrium effects will be significantly weaker when taken to the data.

Fourth, the model assumes that the consumption price, p_C, for agricultural goods, is exogenous. This closes off manufacturing demand responses to agricultural prices in the price support case (since demand relies entirely on p_C, which is exogenously set). However, in the Indian Public Distribution System, only a small selected fraction of the population can obtain (a quota of) rice at subsidized prices; the majority of consumers purchase rice on the open market. Open market rice prices, as I show in the next section, decrease in response to agricultural productivity shocks. We know from the model that income effects outweigh substitution effects and manufacturing demand should increase as a result. This, when taken to the data, will dampen the negative manufacturing labor response to agricultural productivity shocks that I derived in the model with price supports.

Fifth, farmers produce a wide variety of crops beyond the two for which price supports are significant, while the model assumes that agricultural output is a single crop. I show that in response to a positive productivity shock, farmers in fact substitute away from staples (perhaps due to utility from diet diversity, and the fulfillment of a caloric minimum intake), which the model cannot capture. This is negated in high price-support years. We may also be concerned that subsitution among crops that are heterogeneous in labor intensity entirely drives the labor market shifts I observe. However, I show increases in raw yield per hectare for staples, indicating an additional increase in input intensity for the crops under price supports - this means that the labor response does not arise simply from a shift to more labor-intensive crops that have price supports.

1.4 Empirical Strategy

Broadly, identification stems from the interaction between local weather-related productivity shocks, and the extent to which national price supports for rice and wheat keep up with local wholesale prices. Since districts face different weather shocks in different periods, this provides exogenous variation in agricultural productivity and therefore prices at the district-year level[46]. To separately identify the differential effect of productivity shocks on production with and without price supports,

[46]Weather shocks can take two forms: rainfall and temperature. Previous work confirms that temperature and rainfall are significant predictors of crop yields (Lobell et al. 2007, Schlenker et al. 2009). However, I avoid using temperature shocks due to their potential direct effect on workers' productivity in the non-agricultural sector (West 2003; Chen 2003; Chan 2009), in favor of focusing on rainfall shocks. Previous work (Dercon 2004, Miguel et al. 2004, Jayachandran 2006, Kaur 2017) has interpreted rainfall shocks as exogenous shifters of TFP.

I interact these with high and low support prices in a differences-in-differences framework.

1.4.1 Differences-in-differences Framework

I consider each district to be a distinct local market within which producers choose to sell either to the market or to the government at harvest-time. I assume that producers have the ability to sell their produce to any wholesale market in their district[47].

Districts vary across time and space in where their realized wholesale market price falls relative to the national MSP. Data from wholesale markets suggest that average district-level harvest-season wholesale prices for rice and wheat in the period 1997-2012 fall below the government's MSP for both rice and wheat for a significant proportion of districts[48].

I show that there is variation over time in whether the MSP is low or high relative to the entire distribution of realized local market prices, which does not necessarily follow any particular time-trend; there are early years with high MSPs relative to the distribution and later years with low MSPs relative to the distribution (Figure 1.5). This motivates my definition of the MSP as 'low' or 'high' in each year[49](Figure 1.6). I describe how I determine whether the price support in a given year is high or low in detail in Section 1.4.3.

I combine this with exogenous shocks to agricultural productivity derived from early-season rainfall to determine the how the price support policy affects production and input responses to productivity shocks. I describe how I define productivity shocks in detail in Section 1.4.2.

---

[47]This, as I have discussed in the previous section, is an approximation, given that producers who are close to district borders may well find that a wholesale market in a neighboring district is closer to them.

[48]That is, district-level wholesale prices at harvest are, in fact, *below* the MSP in approximately 39% of district-year observations for rice and 25% of observations for wheat. 90% of districts are below the MSP in one time period, but above it in another. Figure 1.4 shows the distribution of rice and wheat market prices relative to the MSP for all district-years.

[49]A continuous version of this variable, the percentile of the support price within the entire distribution of prices, provides similar results, but is more difficult to interpret directly. Results available upon request.

We can think of each district-year observation as falling into one of four categories:

(A) Low Rainfall, High MSP	(B) High Rainfall, High MSP
(C) Low Rainfall, Low MSP	(D) High Rainfall, Low MSP

In my reduced-form empirical strategy, the effect of the price support policy is reflected in the differential response to good rain (and therefore higher productivity) in low- and high-MSP years. I expect that farmers in districts that experience positive early-season rainfall shocks will anticipate higher agricultural productivity and lower prices (which I verify in detail in the next subsection). When support prices are high, the lower market prices do not factor into production decisions (which are driven by p_S, the level of the price support). Farmers also get an income boost, since the price support is higher than the local market price, exacerbating the effect. They are less likely to cut production in response to positive productivity shocks (good rainfall). As per the table above, that indicates that the difference in staple production in categories B-A will be significantly higher than D-C.

The parallel trends assumption assumes that any direct effect of agricultural productivity shocks on input allocation and crop mix that are not price-related are the same in high and low-MSP years (for example, early-season rainfall being a bigger boost for rice and wheat productivity than for other crops), except for the effect of price supports. Given that local productivity shocks are unlikely to influence whether price supports are high or low on a national level, the parallel trends assumption likely holds.

There are three main challenges that drive my choice of empirical strategy. First, I cannot use a cross-sectional comparison of districts with high and low prices relative to the support price to independently identify the effect of the price support policy. Districts in which realized wholesale market prices fall above the MSP are unobservably (to the econometrician) different from districts in which wholesale prices tend to be low. I therefore use a district-time panel of planting decisions between 1997 and 2012 and include district fixed-effects to compare the response of planting decisions to productivity shocks within the same district in years in which the national MSPs for rice and wheat are more salient to the farmer's decision-making (relatively higher) to years in which they are less so (relatively lower).

Second, using realized market prices at harvest to determine whether price supports are high or low in a given year results in reverse causality. Realized harvest wholesale market prices are, in equilibrium, determined both by planting decisions and market demand for each crop. They are also unknown to the farmer at the time that planting decisions are made. I therefore use price trends for each district to create a (parametric) prediction model for market prices (Section 1.4.3)[50]. It is important that this prediction be informative *before* planting decisions have been made. These predicted prices form the distribution of anticipated local market prices that determines whether the MSP is high or low in a given year.

Third, a direct comparison between districts with low and high market prices might not estimate the true effect of the support price policy. There are both income and insurance mechanisms at work - people could change planting decisions simply because anticipated income is higher from staple production under the program, or they could respond to the security of having a guaranteed price for rice and wheat, even if the probability of local market prices falling below the minimum support price is low[51]. Because of this, I choose to define all districts in a given year as affected by either a 'high' MSP or a 'low' MSP, and calculate average effects across all districts (both below and above the support price). I do test that the effect of the program is greater for districts in the lowest 30 percentiles of the price distribution in each year[52].

---

[50]Since each farmer is a price-taker, I abstract away from equilibrium effects in the prediction model.

[51]This would be even more significant if the ability to sell on the local market were limited through informal quotas or limited demand, leaving even farmers in high-price districts with no option other than to sell their remaining produce through the government program, or let it rot for no return.

[52]These districts' market prices typically always fall below the level of the price support in both low- and high-MSP years, and the model suggests that when the price support binds, the higher the level of the price support, the greater the response of farmers in that district. The results are provided in Appendix Tables A1.1 and A1.2.

1.4.2 Positive Productivity Shocks: Early-season Rainfall

If markets are sufficiently integrated that productivity shocks do not lead to price fluctuations, then price supports would not have a major role in mediating the allocative role of the price mechanism in this context. I test that local market prices do indeed respond to early season rainfall. Figure 1.7 provides a plot of price residuals (accounting for year and district fixed-effects) against deciles of early-season rainfall within a district. I find that highly negative shocks result in higher local wholesale prices for both rice and wheat.

Price responses to negative shocks are significant for both rice and wheat, as described in columns 1 and 2 of Tables 1.1 and 1.10.

In all main specifications, I define local-level shocks to prices as arising from *negative* deviations from the 40-year long-run average of early-season rainfall for the district, since positive deviations from the LR average are less informative about prices for both crops[53]. I define 'bad' rain - that is, rain that causes prices to increase - as any negative deviation of rainfall of more than 50% below the LR mean of early-season rainfall for the district[54], and confirm that negative shocks defined this way result in higher prices[55].

I also test that price responses to rainfall shocks are not significantly different in low- and high-MSP years, which means that farmers' local market option responds similarly to productivity shocks in

[53]Coefficients on positive deviations are small and insignificant in 1.1 and 1.10, and this pattern is observable in Figure 1.7 as well.

[54]I confirm graphically that prices are responsive to rainfall below this threshold in Figure 1.8. The relationship between prices and rainfall, initially steeply negative on the left of the rainfall distribution, and relatively flat beyond 50% below the LR mean level of rainfall, shows that prices are less responsive to rainfall when the price support is more likely to bind (that is, when the district experiences adequate rainfall), which suggests that the price support policy is indeed binding for some producers in some time periods with adequate rainfall.

[55] This includes 15.7% of observations for rice and 12.1% of observations for wheat. I also conduct a robustness check using a definition of bad rain common to the development literature: defining only the first quintile of observed deviations from the average for that district as a shock to prices. Using both definitions, prices are significantly higher in 'bad' rain years. While I focus mainly on the former in all main results, all results are robust to the latter specification. I present this in more detail in Section 1.8.

both high- and low-MSP years (Columns 4 and 7 of Tables 1.1 and 1.10).

Price responses are large relative to the residual variation in prices (after controlling for year and district fixed effects). A negative early-season rainfall shock causes prices to increase by Rs. 39 per quintal of rice (21% of residual standard deviation in rice prices) and Rs. 27 per quintal of wheat (28% of residual standard deviation in wheat prices). Early-season rainfall is, therefore, an important, and exogenous, shifter of agricultural productivity and therefore local market prices.

1.4.3 High and Low Price Supports: The Farmer's Prediction of Prices

There is an extensive literature that suggests that farmers adjust to information provided to them prior to the time of planting, based on anticipated profitability[56]. Here, I suggest that farmers use the information they have about productivity to make predictions about prices, and therefore about profitability of the their crop. I also assume that farmers' expectations of market prices given early-season rainfall are rational based on their past observations. I make price predictions in a parametric way, assuming that farmers have knowledge of past rainfall and prices[57], but limited recall. I use farmers' price predictions to classify support prices as high or low in a given year. A given district is 3.4pp, or 8.6% more likely to have a binding realized market price in a 'high' MSP year relative to a 'low' MSP year (Column 3 of Table 1.1).

To do this, I define farmers' information set in each time period, t, which includes msp_t and early-season rainfall w_t, and realized prices and early-season rainfall for the past α years. That is, they have observed the relationship between early-season rainfall and realized prices for the past α years, and use the parameters that define that relationship to predict this year's market price based on this year's early-season rainfall.

[56]Rosenzweig & Udry 2013, Kala 2015

[57]and, in some specifications, past MSP

I then use only the data contained in these information sets to make a prediction about this year's market prices during the harvest period for rice and wheat. Specifically, I use a district-specific quadratic function of early-season rainfall[58] , a district-specific time trend, a statextime trend, and district fixed effects, to predict market prices in t.

The empirical specification used in the prediction stage is as follows. I run the following specification using data from $t-1$ to $t-5$:

$$
\begin{aligned}
p^m_{dst} = \beta_0 + \beta_{1dst} EarlyRainfall_{dst} + \beta_{2dst} EarlyRainfall^2_{dst} \\
+ \beta_{3ds} \delta t_{dst} + \iota_{ds} + \epsilon_{dst}
\end{aligned}
$$

where p^m_{dst} is the local price in a given district d in state s in a given agricultural year and season t. The coefficients on $EarlyRainfall_{dst}$ describe a district-specific quadratic function of the relationship between early-season rainfall and local prices. I also include ι_{ds}, a district fixed effect. δ_t is a district-specific year trend, to account for districts being on different price trajectories over time. Xt_{st} are state-time trends. The error ϵ_{dst} is clustered at the district level.

In creating the specification in this way, I allow for farmers to use other aspects of the prices and data they have observed over the past five years (time trends, district fixed effects that capture the average prices of staples in their district over the five-year period, etc.) in their predictions.

I use the coefficients from the prediction specification to predict prices in time t. Using the predicted prices, I then calculate the percentile of price support in the predicted price distribution. I use median of this value to divide years into 'high' and 'low' MSP years.

[58]All rainfall terms are percent deviations from the 40-year long-run average of early-season rainfall taken from 1970 to 2015. Early-season rainfall enters as a quadratic function to allow both positive and negative deviations from the long-run mean to have an effect on prices.

The effect of price supports on production outcomes is robust to various alternatives to this type of prediction. In two alternate specifications (presented in Tables 1.17 and 1.18), I a) exclude early-season rainfall from the prediction, and b) include the level of the MSP in the prediction (so farmers take into account responses of harvest-season market prices to MSP announcements). I also run versions of the specification that vary the size of the information set, α, that the farmer considers in making his prediction (Table 1.19). I discuss these checks in detail in Section 1.8.3.

1.4.4 The Farmer's Decision Timeline

To make things more specific, the timeline of information and decision-making for the compliant farmer looks as follows:

1. The farmer's pre-planting information set $\mathscr{I}$ includes the realized wholesale market price in the district and information about early season rainfall for the past five years, together with the standard deviation of realized market prices around the prediction. Given the lack of empirical work on farmer decision-making and the extent of information considered in making a decision about this season's planting, this is simply a benchmark model, and I will later examine robustness to varying the size of information set.

2. Based on his information set, he creates a function that links early-season rainfall to realized wholesale market price within his district.

3. Before planting, the farmer observes the signal (early season rainfall), w_{dst} in district d, in state s, in time t.

4. Based on the weather shock, and the prediction model, he knows $p^{\hat{s}}_{dst}$, the expected wholesale market price for the staple crop, and the distribution of potential yields for various crops, $\hat{Y_{jdst}}$.

5. These are all stochastic because of a second, multiplicative weather shock, η_{dst}, which is

realized after planting and before the harvest, and affects the final distribution of the market price (but not planting decisions). The realized market price $p^s_{dst} = p^{\hat{s}}_{dst} * \eta_{dst}$, where η is centered around 1. That is, $E[\eta] = 1$.

6. At the same time, the government announces the national-level MSP for the year for the staple crop, msp_t.

7. Given his price prediction and knowledge of the MSP, together with the standard deviation of realized market prices around the predicted market price, the farmer knows the expected probability that realized market price will fall below the MSP. Since the realized market price p^s_{dst} is stochastic even after the initial realization of the weather shock, the distribution of expected prices gives farmers a probability, $\theta | w_{dst}$, that they will eventually sell their harvest at the msp_t.

8. Farmers use these expected probabilities that the market price will fall below the MSP (in which case they expect to sell their crop at the MSP), and predicted prices for rice and wheat, together with information encoded in early season rainfall about the year's relative prices, costs, potential yields, and revenues from various crops to select a portfolio of crops to plant.

9. Then, at harvest time, if the realized market price for the staple p^s_{dst} is higher than the msp_t, farmers sell their output at p^s_{dst}. If it is lower, they sell their output at msp_t.

1.4.5 Empirical Specifications

I implement the differences-in-differences strategy using the following empirical specification:

$$
\begin{aligned}
Y_{dst} = {} & \beta_0 + \beta_1 GoodRainfall_{dst} + \beta_2 GoodRainfall * HighMSP \\
& + \iota_{ds} + \delta_t + Xt_{st} + \epsilon_{dst}
\end{aligned}
$$

In this specification, Y_{dst} are the outcomes of interest in district d in state s in agricultural year t (June to June). These include total area cultivated, area cultivated of staples, area share of staples and other crops, yield, and production[59]. I include ι_{ds}, a district fixed-effect, to control for time-invariant district heterogeneity, such as suitability of the district to grow staples, terrain, how urban or rural a particular district is, average market prices in the district, and so on. δ_t is a year fixed-effect to isolate the effect of high MSP from other changes in production from one year to the next. Xt_{st} are state-time trends that aim to account for the potential influence of any particular state on support prices. For yield and production outcomes, I control for cubic polynomials of late-season rainfall which can have a direct effect on productivity after planting decisions are made[60] . Errors ϵ_{dst} are clustered at the district level.

1.5 Data

1.5.1 District-time Panel Data

District-time panel data comprise the main data in this paper. These types of data cover all sources of variation over time in prices and rainfall for the empirical analysis, as well as information on district-level planting patterns that change over time.

Data on cropping patterns are important for assessing the first-order responses to the price support policy. The government[61] collects information on area planted and quantity produced for various crops for each district in each season in each year for all districts in India -these are known as the

[59] I run specifications in levels rather than logs, to allow for switching into and away from producing staples. The data suggest that this pattern is fairly common. Of the districts covered, 74 rice-producing districts and 97 wheat-producing districts report zero production of the staple crop in the *Kharif* season for rice and in the *Rabi* season for wheat in at least one year, but not in all years.

[60]I also ensure that the inclusion of late-season rainfall controls does not drive the results. I present my main specifications without the rainfall controls in A1.5, and find that the sign, magnitude, and significance of the production results remains the same, barring some additional noise.

[61]Directorate of Economics and Statistics of the Ministry of Agriculture and Farmers' Welfare

Area Production Yield (or APY) data[62]. I derive area cultivated and raw yields (output per unit area) for each crop in each district-season-year from this dataset. For further analysis on changes in cropping patterns, I classify crops into four main categories: other staple crops, pulses, cash crops, and spices.

Rainfall data allow me to identify which districts face rainfall shocks that affect predicted market prices. I obtain monthly precipitation data at 0.5° resolution[63], which I aggregate to the district level[64]. I use total precipitation in the months of May and June for the *Kharif* season, and September, October, and November for the *Rabi* season, to define pre-planting shocks to agricultural productivity[65].

I use wholesale price data aggregated to the district-level as a measure of local market prices. I use these data, together with rainfall data, to predict harvest-time market prices for each district and create a distribution of anticipated prices for all districts. Daily wholesale price data are sparse in India, particularly in the period prior to 2005. I first compile all available price data for rice and wheat across markets in India reported by AGMARKNET (the number of markets and the number of districts covered varies over time, and currently stands at 3245 wholesale markets across the country)[66]. I average observed daily wholesale prices over the harvest period for each season[67]. I

[62] For a few states in a few years, missing APY data has to be supplemented with Land-Use Statistics Data instead, which does not provide production data. Minor crops that comprise less than 1% of the cultivated area in a district are excluded for a few state-years in the LUS data, due to the sheer number of crops.

[63] Climatic Research Unit of the University of East Anglia

[64] I calculate measures of monthly rainfall (in mm) at the district-level by superimposing these data on India's district boundaries and calculating means across all 0.5° cells that fall within each district.

[65] In order to define shocks to early-season rainfall more precisely, I calculate percent deviations of each district-year observations from the 40-year long-run district average of precipitation.

[66] Given the low coverage provided by AGMARKNET data, I supplement their wholesale price data using price data, where available, from the ICRISAT meso-level dataset, which covers all districts in 19 major states in India.

[67] January through March for the *Kharif* season and March through June for the *Rabi* season. Wholesale prices rarely move both above and below the price support in a single harvest period for a given district, and that there is little price-variation in markets during the season. This drives my decision to use mean wholesale prices for the entire harvest period for each district to construct my measure of local market prices.

convert prices into real terms using the World Bank's GDP deflator.

Input data are gathered from three rounds of the Agricultural Census Inputs Survey, and cover variable inputs by crop - use of high-yielding varieties, fertilizers, and proportion of area irrigated for rice and wheat.

I eliminate all districts that report no rice or wheat production in the relevant seasons in years of my data. My final sample comprises 5,113 (91% of area under rice production) district-year observations for rice, and 4,707 (94% of area under wheat production) district-year observations for wheat.

1.5.2 Repeated Cross-sectional Data

Households: The National Sample Survey (NSS) consumption/expenditure modules for rounds 55 to 68 are repeated cross-sectional household surveys that are representative at the district-level. I focus on households surveyed during the *Kharif* and *Rabi* harvest months. The surveys provide detailed information on per-capita household consumption at harvest, an estimate of the number of crops produced by each household during the period of this study, and whether the household produces rice or wheat[68].

In many analyses that use these data, I focus on households that consume rice (*Kharif* season) and wheat (*Rabi* season) out of home production, indicating that they are producers of staples[69].

Individuals: The NSS employment survey rounds 60-68 also provide weekly information on labor

[68]The data distinguish between home production and production from external sources. If the household reports consumption out of home produce of any crop, I assume that it produces that crop. This also includes products made from that crop - for example, I assume that if a household consumes wheat flour from home production, that it produces wheat.

[69]Some households might, particularly in response to the price supports, exclusively sell their staple produce to the market, and consume staples purchased from PDS or on the open market, in which case they would not be included in my sample. To the extent that my analysis excludes such households, my estimates are a lower bound on the effects of the price support on harvest-season expenditure by staple producers.

supply and wages throughout the agricultural cycle. I am able to distinguish between agricultural and non-agricultural time-use. I also calculate average daily wages from these data.

I focus on non-urban households in the agricultural sector surveyed during cultivation periods in the agricultural cycle, when labor supply is likely to be most responsive to incentives in the form of price supports. The survey provides a rich set of household- and individual-level control variables.

Firms: The Annual Survey of Industries data are repeated cross sections between 2002 and 2009 that survey all firms with above 100 employees, and a random sample of 1/3 of smaller firms, including both formal and informal firms. These data are able to validate my results on how labor utilization and output by non-agricultural firms respond to agricultural price supports. These data contain detailed firm-level information on characteristics, inputs, outputs, investment, capital, and employment.

I focus on non-urban firms. As a robustness check, I eliminate firms that use agricultural output as their inputs.

1.5.3 District-level Snapshot

Crop suitability measures[70] use soil, topographic, and climatic data to estimate suitability distributions (created as an index with a maximum value of 100 and a minimum value of 0) for a variety of crops for each 5 arc-minute grid cell. I obtain baseline suitability indices for 16 of the most prevalent crops in India, including rice and wheat. I aggregate suitability measures for each crop to the district-level[71], and create absolute and relative suitabiity measures[72].

[70]From the Global Agro-Ecological Zones database collected and disseminated by International Institute for Applied Systems Analysis (IIASA) and the Food and Agriculture Organization of the United Nations (FAO)

[71] I take an average of the Suitability Index across all 5-minute arc grid cells whose centroids lie within the district boundary.

[72]Details are provided in the Data Appendix.

Districts vary widely in their innate suitability for growing rice and wheat, but are all incentivized to grow staples through the government's national MSPs. These suitability indices can be considered a time-invariant baseline characteristic for each district. It is therefore informative to understand how growing patterns and crop yields in low-suitability districts change as a result of price supports, and how the gains from the program are distributed. The fact that low-suitability districts show no response to the program also serves as an additional test that the results derive directly from the program rather than some other random unobserved variation.

1.6 Results and Discussion

The set of results presented here are from the differences-in-differences framework. The preferred specification defines high and low support prices according to farmers' price predictions based on a five-year recall of local market prices and productivity shocks.

The first coefficient provided in each table (on the indicator for "Good Rain") describes the direct effect of a positive productivity shock on the outcome. The second coefficient, on the interaction term between good rainfall and high MSP, is our coefficient of interest. This estimates the differential production response to good rainfall (and therefore lower prices) in high MSP years relative to low MSP years. The last row of every table gives readers a sense of the magnitude of the effect: it shows the effect as a proportion of the mean of the dependent variable. I refer mostly to these magnitudes in the rest of this section.

1.6.1 Agricultural Production

The first set of results in Tables 1.2 and 1.11 cover five measures of agricultural production related to the staple crop: the area (hectares) planted with the staple crop in relevant season, the share of cultivated area devoted to the staple crop, total area cultivated across all crops (the extensive margin of production), and yield per unit area (tonnes per hectare) and total production of the staple.

Staple Area Planted and Area Share: The model suggests that the effect of good rain on staple production comprises two opposing effects - first, there could be an anticipated increase in rice yields that may lead farmers to increase area cultivated until the expected return from the marginal hectare planted is zero. However, the positive productivity shock also indicates lower prices for rice (as shown earlier in Table 1.1) , which puts pressure on farmers to decrease area cultivated, since both the average and marginal return to each unit of produce is now expected to be lower.

The coefficients on good rain in columns 1 and 2 of Table 1.2 suggest that, for rice, the second effect outweighs the first marginally: area share devoted to rice decreases in response to a positive productivity shock when price supports are low. From Column 3, the effect of good rainfall on total area cultivated (coefficient on "good rain") in low price support years is not statistically significantly different from zero, indicating that the two effects approximately offset each other on the extensive margin. For wheat (Table 1.11), the coefficient is positive and significant - good rainfall, in the absence of high MSPs, results in an overall increase in wheat area cultivated - the effect size is about 8% of the mean.

Now we turn to the interaction of productivity and price supports. Here, in the presence of price supports (i.e. when price supports are high and more likely to bind), the price that producers can expect for the staple crop remains stable in response to positive productivity shocks, rather than falling. This should lead to an unambiguously positive interaction effect, which is indeed the case. Columns 1 and 2 of Table 1.2 show that rice area planted increases 3.3% (3,342 hectares on average) in response to productivity shocks in years in which price supports are high relative to years in which they are low. Column 3 indicates no movement in the total area cultivated in the *Kharif* season. Taken together, these two results indicate that farmers shift land into the production of rice, but do not change the amount of land they cultivate on the extensive margin in response to price supports. I corroborate this by looking at area shares of rice (as a proportion of total area cultivated in the *Kharif* season) in column 2, which show an increase of 6.6%.

There are no shifts in area and area share of wheat in response to high support prices in the *Rabi* season. Columns 1 through 3 of Table 1.11 show no change in both land area devoted to wheat and total area cultivated in the *Rabi* season in response to good rain in high-MSP years relative to low-MSP years. However, results in the following sections show that the policy has bite among wheat producers too, even when area cultivated and area share remain unchanged.

Yield and Production of Staples: According to the model, farmers may respond to productivity shocks by devoting various kinds of capital or labor to agriculture. The direct effect of a positive productivity shock on raw yield and production within agriculture (that is, not controlling for inputs) are therefore determined in equilibrium according to the allocation of input across sectors. A positive rainfall shock serves as a Hicks-neutral boost to agricultural productivity and, therefore, yields. However, the model suggests that, in response to agricultural productivity shocks, more labor is allocated towards manufacturing (due to increased manufacturing demand), which decreases the amount of labor per unit area available to work in agriculture, leading to decreased yields for the staple crop. This labor movement out of agriculture in response to positive productivity shocks has also been observed empirically in the Indian context by Emerick (2016) and Santangelo (2016). In columns 4 and 5 of Tables 1.2 and 1.11, I show that the latter effect dominates, and that raw yields fall in reponse to positive productivity shocks. I note that while the yield result, measured in this way, is at first blush counter-intuitive with respect to prior work, my later results on the impact of good rainfall on agricultural productivity (controlling for all inputs including labor), are positive[73]
.

I now consider the interactive effect of productivity shocks and high price supports on yield and production of rice and wheat. According to the model, more labor is allocated to agriculture when price supports are binding. This should have an unambiguously positive effect on yields

[73]I present results for the productivity analysis in Section 1.7.

and production, and therefore a positive interaction effect between high price supports and positive shocks to agricultural productivity. I find an increase in raw yield of rice (production per hectare) of 7.2% when price supports are high (Column 4 of Table 1.2). Given the expansion in the area planted with rice, which we expect to be less productive land on average, the increase in yields suggests a reallocation of inputs towards rice as in the model, which I verify by looking at labor supply in agriculture in the next subsection. The expected return to investing in these inputs is now higher (due to stability of producer prices in the face of positive productivity shocks when price supports are high).

Coupled with the increased area planted with rice, the amount of rice produced (Column 4 of Table 1.2) increases by 8.5% in response to good rainfall in a high-support year relative to a low-support year, on average.

Wheat production also increases between 9.7% (Column 5 of Table 1.11). This is driven by a significant (and large) increase in raw yields of wheat of between 8% (Column 4 of Table 1.11).

Crop Mix: My model abstracts away from the crop mix decision of the farmer by considering a single agricultural output. However, Indian farmers often grow more than one crop, and the decision about how to allocate resources across crops is endogenous to the existence of price supports.

I find, in Table 1.3, that a positive productivity shock in low-suport years, leads to a shift in area share from rice production in the *Kharif* season to risker, higher-return crops like pulses and oilseeds. I interpret this as the direct effect of positive productivity shocks on crop choice. There are a number of potential microfoundations for this result: 1. Farmers might consider rice to be a giffen good in production: after reaching a basic amount of production and satisfying a basic caloric requirement (which happens more easily when there is a positive production shock), they might want to consume a more diverse diet. Second, farmers might be more willing to take on risk once their basic staple needs are met. Third, other crops may be more sensitive to early-season rainfall, both in productivity and prices.

I now move to the interactive effect of price supports and positive productivity shocks on crop mix responses. In the presence of high price supports, I find that, relative to low price-support years, there is a shift in crop mix towards the staple crop, rice. The overall effect on rice area shares in high price supports years (summing the direct and interaction terms in Column 2 of Table 1.2) is effectively zero: farmers do not shift away from rice in response to positive productivity shocks when price supports are high.

Crop Suitability and Responses to Price Supports: As described in Section 1.5, I calculate a measure of relative crop suitability that is akin to a measure of marginal cost of cultivating a unit of land with each of these staple crops *relative to planting other crops*. This is more suitable than a measure of absolute suitability given that landowners and making crop choices on the intensive (as well as extensive) margins. I use these measures for two purposes; first, to discern whether there is heterogeneity in gains from the program between low- and highly-suitable districts. In a policy that prioritizes staple crops over others, my results show that gains are distributed only among districts that are better able to switch into staple-production (Tables 1.4 and 1.12).

Second, I use this to verify that my estimates are driven by exposure to high price supports for rice and wheat specifically, rather than general (and universal) changes in agricultural production patterns that happen to correlate with high price supports.

1.6.2 Consumption

When there is a positive productivity shock, the shift of labor out of the agricultural sector and the resulting increase in agricultural wages, together with lower prices for agricultural produce, imply that the effects on household income are ambiguous. The most direct effect of the support policy on agricultural households is through monthly per-capita expenditure at harvest, when production and market prices are realized. Without a direct measure of income, this is the best proxy measure available. I first consider the direct effect of a positive productivity shock, which has two effects on consumption: it increases consumption through greater productivity, but gives farmers lower prices for their output. Overall, I find that households do not consume more in response to a positive productivity shock (Tables 1.5 and 1.13).

I then look at the effect of price supports on consumption responses to positive productivity shocks. Rice- and wheat-producing households surveyed at harvest both show no differential increase in monthly per-capita expenditure in response to good rainfall in high-MSP years, relative to low-MSP years (Tables 1.5 and 1.13). Agricultural households produce more price-supported output in high MSP years, but receive lower wages for the labor they sell to other producers (which I discuss in the next section). These opposing effects comprise the null result.

However, I find that rice- and wheat-producing households consume more in high price support years relative to low price support years, while agricultural households that do not produce these crops show no such increase. This negates the concern that the program is simply implemented badly, or that farmers face costs that are too high in accessing it.

1.6.3 Spillovers to the Non-Agricultural Sector

The increases in yield and production for both rice and wheat indicate a shift of inputs towards staple cultivation. Chief among these is labor. I find no increase in the use of irrigation, high-yielding varieties, or fertilizer in response to price supports for rice and wheat (Table 1.8)[74].

Labor Allocation Across Sectors: Prior literature suggests that Hicks-neutral or labor-saving productivity shocks can lead to industrial growth (Bustos et al. 2012, Emerick 2016, and Santangelo 2016) through reallocation of labor into the non-agricultural sector[75], in line with comparative statics

[74]This could be in response to low barriers to procuring heavily-subsidized inputs even in the absence of price supports

[75]The latter two papers explicitly document this pattern in response to weather-related positive productivity shocks in India.

from the model.

I consider the labor allocation of non-urban households in agriculture using the NSS employment surveys. I find, first, that my results hold true to previous work on the interlinkages between the agricultural and non-agricultural sectors - good rainfall results in an 8pp increase in an indicator for non-agricultural work for the week during the cultivation season (an approximately 35% increase for the typical agricultural household) (Column 6 of Table 1.6).

I then consider the interaction between the price support policy and labor responses to productivity shocks. I do find that price support policies completely negate the movement out of agriculture by providing incentives to allocate labor to agriculture. In Column 4 of Table 1.6 I use an alternative definition of labor allocation, the number of days worked in agriculture, and find a 9.78% increase (similar in magnitude to the indicator outcome). I also find that these labor movements occur on the intensive margin - overall labor supply (both as an indicator and in terms of number of days worked) shows no change (columns (1) and (2) of Table 1.6).

I note that this pattern emerges even in the case of wheat, a significantly less labor-intensive crop, though, as anticipated, magnitudes are lower (Table 1.14). I also find that, in the case of wheat, the pattern applies only to men and not to women. I present results here only for the men in the sample. This is, however, entirely unsurprising given that wheat production (to a much greater extent than rice production) tends to exclude women from the process, particularly during cultivation[76].

Occupational Choice: I then check which sectors see the largest decreases in labor allocation in response to productivity shocks in the presence of high price floors, and find that manufacturing and construction (key non-agricultural sectors in rural areas) are affected the most (Table 1.7).

This assuages any potential concern the labor movements are driven by fluctuations in forced

[76]Chen(1989) points out that women participate in tasks such as weeding, winnowing, drying, storage, and husking or milling, most of which are done at harvest-time. She also states that mechanization has displaced women from even these tasks, and that the shift to chemical fertilizers has shifted women away from a key cultivation-period task: manure-spreading.

entrepreneurship (which is often reported as non-agricultural employment when agricultural labor demand might be low).

Wages: Finally, I consider the effect on daily wages in both sectors. While the model assumes perfect labor mobility, any frictions in the labor market will cause wages to diverge across sectors. At baseline, wages in the non-agricultural sector are significantly higher than in the agricultural sector.

Coefficients on “good rain” in Columns 5 and 8 of Table 1.6 show that wages in the agricultural sector increase, and wages in the non-agricultural sector decrease in response to positve productivity shocks, as a natural extension of the labor market shifts I dicussed in the previous subsection. This indicates that falling prices in the agricultural sector and higher demand in the non-agricultural sector in response to positive productivity shocks encourage labor movements that lead to more equal wages across sectors.

I now turn to the differential effect of positive productivity shocks in high-support years relative to low-support years. Wages are higher in equilibrium in the non-agricultural sector during the cultivation season under higher price supports, corresponding with the negation of labor movements out of agriculture. This is especially true in response to the rice price supports, in which there is a 23% differential increase (Rs. 30.5) in non-agricultural daily wages in response to high price supports (Column 8 of Table 1.6).

A similar estimate for wheat is insignificant, though it moves in the right direction in conjunction with shifts in labor allocation (Column 5 of Table 1.14).

Labor Use from Industry Data: I use the Annual Survey of Industries (ASI) cross-section of firm-level data for rural firms to confirm that labor is reallocated from firms towards agriculture when rice support prices are high[77]. I find a fall in manufacturing worker-days for open[78] rural firms in response to agricultural productivity shocks in high price support years. This corresponds to the decrease in labor supply from the household surveys (Column 7 of Table 1.6). This strengthens the argument that price supports have direct effects on labor use in non-agricultural firms.

Interestingly, knowledge of worker-types in the data helps me identify that this effect is driven by a decrease in total worker-days of 17.8% for contract laborers, which is precisely the margin of adjustment for workers who divide their time across sectors. There is no such effect for permanent employees of these firms (Columns 10 and 11 of Table 1.6).

Output from Industry Data: Finally, I test the direct impact of price support policies on output in the non-agricultural sector using the ASI data. I find that gross output decreases by 2.6% of a standard deviation in high price support years in response to a positive agricultural productivity shock (Column 1 of Table 1.9). Value-added measures (Column 2) provide results of roughly similar magnitudes, though these estimates are noisier.

These results suggest that the crowding out of labor from non-agricultural allocation in response to high price supports has a concrete effect on production in the non-agricultural sector, at least in the short run.

1.7 Implications for Agricultural Productivity

Based on the results up to this point, I find that high price supports for rice and wheat crowd out labor allocation to, and output in the non-agricultural sector. However, the increase in labor usage in the agricultural sector may, in fact, be productivity-enhancing in that sector. To examine this, I

[77]Since the data are reported at an annual level, I cannot distinguish between the effects of rice and wheat price supports, and so I choose to focus on rice price supports.

[78]To the extent that firms shut down due to lack of access to labor, or higher wage rates, my estimates are a lower bound.

quantify the effect of increasing labor usage in the agricultural sector on agricultural productivity (in the form of agricultural TFP). I use an aggregate agricultural TFP measure common to the literature, the Tornqvist-Theil Index[79]:

$$ln(\frac{TFP_t}{TFP_{t-1}}) = \frac{1}{2}\sum_i (R_{it} + R_{it-1}) ln(\frac{Q_{it}}{Q_{it-1}}) - \frac{1}{2}\sum_j (S_{jt} + S_{jt-1}) ln(\frac{X_{jt}}{X_{jt-1}})$$

Where R_{it} refers to the revenue share of output i in time t and S_{jt} refers to the cost share of input j in time t.

This index allows us to not only capture changes in quantities produced of various crops and quantities used of various inputs, but also any associated changes in input and output prices. These price changes are particularly important since I have already shown that wages respond to the existence of high price supports in accordance with the movement of labor.

I calculate this index for each district for each year in my data for which the NSS modules are available (from which I extract the district-level prices of 14 different crops[80] that I use), and use the same differences-in-differences framework to examine the effect of the policy on agricultural productivity.

I find an increase in agricultural TFP in response to a positive agricultural productivity shock, as anticipated (the coefficient on good rain in column 3 in Table 1.9). Turning to the interaction term, I find that the increased use of labor in agriculture in response to the higher price support actually decreases agricultural productivity by 0.82 of a standard deviation, negating the positive productivity effect of the shock.

[79]Diewert 1976, Caves et al. 1982, Rosegrant & Evenson 1992, Murgai et al. 2001. This index provides an exact measure of technical change for linear homogeneous translog function that approximates - by a second-order Taylor polynomial - the Cobb-Douglas production function that I use in the model.

[80]Details about the crops used and data sources for this analysis are given in the data appendix A4.2.

However, there are clear data limitations that deem my estimate a lower bound, the key of which is the lack of annual data on district-level input quantities and prices. Instead, I use input quantities for a single year, in which the price support for rice is low, for all inputs other than labor. I use input prices at the state, rather than district, level to calculate cost shares of various inputs (other than labor). Then, to the extent that competition for inputs other than labor also increases simultaneously, pushing up their prices, or that quantities used of these inputs increase when price supports are high, my estimates do not take that into account, and are therefore a lower-bound estimate of the effect of price supports on agricultural productivity.

1.8 Potential Confounds and Robustness Checks

1.8.1 Defining High and Low Early-Season Rainfall

In the main specifications, I define rainfall shocks in the two seasons to be a greater-than-50% negative deviation from the 40-year average of early-season rainfall. I also use one alternate specification of good rainfall common in previous literature (Jayachandran 2006; Kaur 2017) that provides a weaker price differential between periods of ‘good’ and ‘bad’ rainfall[81], and is therefore more conservative. I define the first quintile of observed deviations from the average for that district as a shock to rainfall and therefore prices.

I present results using this alternate specification in Table 1.16. I find that results all follow with similar magnitudes as in my main rainfall specification, and remain significant.

[81] The average price differential between periods of good and bad rainfall is Rs.26 per 100 kg for rice, and Rs.16 per 100 kg for wheat, compared to the main definition of rainfall shocks, in which the differential was Rs. 39 per 100 kg for rice and Rs. 26 per 100kg for wheat.

1.8.2 Defining High and Low Early-Season Rainfall Without Predictions

1.8.3 Including Different Elements in the Farmer's Information Set At Planting

I test the robustness of the production responses to including different amounts of information in the farmer's information set at the time that planting decisions are made. Table 1.17 provides results for the five main agricultural production outcomes from three prediction methods (based on varying the elements in the farmer's information set) for the main rice production (*Kharif*) season. In the first column, support prices and rainfall do not figure into price prediction (so farmers base their price expectations merely on district-specific time-trends in prices). The second allows predictions to take into account the district-specific effect of rainfall on prices, and is the preferred specification in the main tables of the paper. The third specification accounts for both minimum support prices and rainfall in making price predictions. I show that these key effects of the policy generally move in the same direction and are of the same approximate magnitude for all three specifications. The same is true for wheat (Table 1.18).

I also test that the production results are robust to changing the number of years of information retained in the farmer's memory. I do this by testing a three-year and a seven-year recall period for rainfall and market prices for the farmer, and find the same increase in input intensity aross specifications (Table 1.19).

1.8.4 International Prices

If farmers plant rice and wheat for export, prices for staples in international markets could affect planting decisions. This could affect my empirical strategy if high MSPs correlate with higher prices on the world market. There are four reasons to think that this not the case. First, the trends in world prices and support prices (in real terms) do not coincide (Figure A1.1). Second, access to world markets should not change relative to local-level early-season rainfall shocks. Third, farmers very rarely sell directly for export[82]. Fourth, I check whether the effects of a price support for rice on agricultural production remain in the time period between October 2007/April 2008 and September 2011 (Sharma 2011)[83], when there was a rice export ban across the country, and find that they do (Figure A1.2 illustrates the effect of the ban and Table 1.20, Panel A provides results)[84]. I also analyze the effects of price supports on wheat cultivation in the period Feb 2007 to May 2010 [85], which coincided with a ban on wheat exports (Sharma 2011), and find similar overall patterns (Table 1.20, Panel B).

1.8.5 Consumption Side of the Program

Rice and wheat procured from the program are resold at subsidized rates to households below the poverty line. It is possible that years in which procurement is high due to relatively high support prices are a o years in which more is available at subsidized rates to households. Households then sell a larger proportion of the staples they produce (at the high government support price) and then purchase the max consumption quota (or the household's requirement, whichever amount is lower) through the program at subsidized rates. However, that effect functions at the post-production sale margin (the choice of whether to sell or to keep and consume). Any effect on the extensive margin of staple-production would be, if anything, negative.

1.8.6 Program Implementation

We might also be concerned about uneven program implementation across districts within the country. As mentioned in Section 1.2, there are multiple reasons for differential access to govern-

---

[82]NSS 70

[83]This corresponds to agricultural seasons 2008-2009 to 2011-2012.

[84]The overall direction and magnitude of the results remain the same, with similar significance.

[85]This corresponds to agricultural seasons 2007-2008 to 2009-2010.

ment purchases of foodgrains, the key of which are variation in the density of government depots (and transportation costs farmers face in taking their produce to the market), and anecdotal reports of delayed depot openings, delayed payments, and other operational constraints[86]. The empirical strategy I use assumes that harvest-season transportation costs and program implementation do not vary in a systematic way within a district with early-season rainfall. To the extent that these constraints are present but unresponsive to early-season rainfall, my estimates form lower-bound estimates for a well-implemented price-support policy.

To rule out the possibility that the program is simply implemented better (or solely) in districts that are relatively more suitable for rice and wheat, I check whether the lack of effect holds for districts that have low relative suitability for rice but high absolute suitability, and find that they do[87]. Districts that are low in relative suitability allocate, on average, 40% of their cultivated land area to rice production, and, on average, dedicate more land to rice production than their highly relatively suitability counterparts. Relative suitability is therefore unlikely to be a mechanism for selection in implementing the program.

1.8.7 Responses in Program Implementation to Planting Decisions

It is also possible that program implementation at harvest responds to planting decisions after early-season rainfall is observed. The government might choose to increase procurement when a higher amount of rice and wheat has been planted in a particular district, for example. If this is unanticipated on the part of the farmer, then it should not enter into consideration at the time of planting. If consumers are aware that this is the case, then they are indeed responding to increased access to the program, since they might be aware that their probability of being able to sell to the government, should they want to, is higher. If anything, this overcomes the access constraints

[86]For example, “Lackadaisical Govt Procurement Forces them to Sell Cheap”, *The Hindu*, 07/09/2017

[87]Results available upon request

mentioned in the paragraph above, getting us closer to an accurate estimate of the effect of a well-implemented program on farmer decision-making.

1.8.8 Spillover Effects of Rice Cultivation on Wheat Cultivation

Rice cultivation begins in June, immediately after the southwest monsoon, while wheat cultivation occurs primarily in the *Rabi* season, which occurs from October through April. Given the staggered timing of rice and wheat production, there is a concern that rice production (and the resulting increase in income due to high rice support prices) drives the increase in wheat yields. To understand the effect of wheat support prices on farmers separately from the effects of rice support prices, I focus on the effect of wheat support prices on production in four subsamples of my data:

1. Years in which rice price supports are low relative to the rice price distribution - that is, the percentile of the support price relative to the price distribution is lower than the median across all years
2. Districts that plant over half their cultivated area with wheat in the *Rabi* season, but less than half of their cultivated area with rice in the *Kharif* season
3. The intersection of subsamples 1 and 2
4. Top-ten wheat-producing states, excluding Rajasthan and Madhya Pradesh

I present results in Table 1.21. I find results that are entirely consistent with my basic results with regard to area cultivated with wheat (no effect), total area cultivated (no effect), and significant increases in wheat yield, and in wheat production. This indicates that the influence of support price policies on wheat production are not driven by spillover responses in the *Rabi* season from the rice-intensive *Kharif* season.

Finally, I check if high rice support prices mitigate, rather than drive, the increases in wheat yields and production, and find that they do. In the final four columns of Table 1.21, I present results from a falsification test: I consider districts that devote more than half their *Kharif* season area share to rice, and less than half their *Rabi* season area share to wheat - these are districts that might rely relatively heavily on rice price supports and not on wheat price supports. I find no effect of high wheat support prices on production in those districts, suggesting that rice price supports can, in some instances, dampen the effects of the wheat pricing policy.

1.8.9 Individual States' Influence on Results

I exclude Rajasthan and Madhya Pradesh from all wheat specifications, even though they are significant contributors to wheat production in India. This is because they are states that have also had a long history of significant state-level bonuses to the national-level support price policy (Rs. 100-150 above the MSP in the years in my sample). I verify that patterns of response to the policy are unique for these two particular states, but hold entirely well across the board when they are excluded (Appendix Table A1.4). Given the influence of the state government on production patterns in these two states, I achieve a more characteristic estimate of the policy response by using the other 31 states alone.

I test that results are not driven by state-level policies or trends that encourage particular patterns of production, with the exception of Rajasthan and Madhya Pradesh for wheat. Since rainfall is also likely to be serially correlated within a state, it is particularly important that these state policies do not respond to early-season rainfall or rainfall predictions for the state and, in turn, influence my results. I therefore create leave-one-out estimators for each state for both rice and wheat cultivation, which are presented in Appendix Tables A1.3 and A1.4. I find that my main results remain consistent in sign and magnitude and, in the vast majority of specifications, significant, across the state-level jackknife specifications.

1.9 Conclusion

This paper provides empirical evidence on producer responses to agricultural price supports, which can distort gains to farmers. I find that farmers respond in significant ways to price supports for two staples, rice and wheat, in the Indian Public Distribution System. Producers switch farmland into rice production, increasing output by 8% in response to good rainfall shocks in high-price-support years relative to years in which the price support is lower relative to the distribution of market prices. Wheat farmers, in contrast, do not change patterns of land cultivated, but similarly increase yield and total production significantly.

The production results, taken together, suggest that farmers are using more inputs per unit area cultivated for the two supported staple crops. Importantly, I find that the key source of increased agricultural yield is a reallocation of labor from the non-agricultural sector (particularly by contracted, rather than permanent, workers), resulting in a decrease in output in rural manufacturing and an increase in wages in the non-agricultural sector. This has, in line with other work by Gollin et al. (2013), Matsuyama (1992), Foster and Rosenzweig (2004, 2008), the potential to crowd-out growth driven by a more productive non-agricultural sector in favor of availing of these government incentives for agricultural production. The magnitudes of these effects are large: a measure of contract labor worker-days in manufacturing decreases by 17.8%, and gross output falls by 8.5%. In addition, when the loss in manufacturing output is taken into account, the implicit cost of the price support program doubles.

Simultaneously, the increased use of less-efficient quantities of labor in the agricultural sector results in a decrease in agricultural productivity of 0.82 standard deviations. Agricultural price supports therefore hinder the growth of the non-agricultural sector, while reducing productivity in the agricultural sector they are meant to support.

Finally, from a policy perspective, price supports can, and do, place heavy administrative burdens on governments. At the same time, countries like India continue to battle high rates of malnutrition, stunting, and seasonal hunger. Recent estimates suggest that up to 10% of rice sold through the Public Distribution System rots and goes to waste. It is plausible that open market sales of rice (perhaps coupled with heavy consumer subsidies), lead to lower production but more effective distribution of food to poor households.

Future work can consider the specific labor market implications and broader welfare effects of price supports relative to direct lump-sum payments to farmers, or payments to farmers when agricultural productivity is low (insurance). Future work will also consider how to balance price policy and procurement across the entire spectrum of crops for which the government announces support prices to prevent wastage while still offering farmers income support when necessary.

1.10 Figures

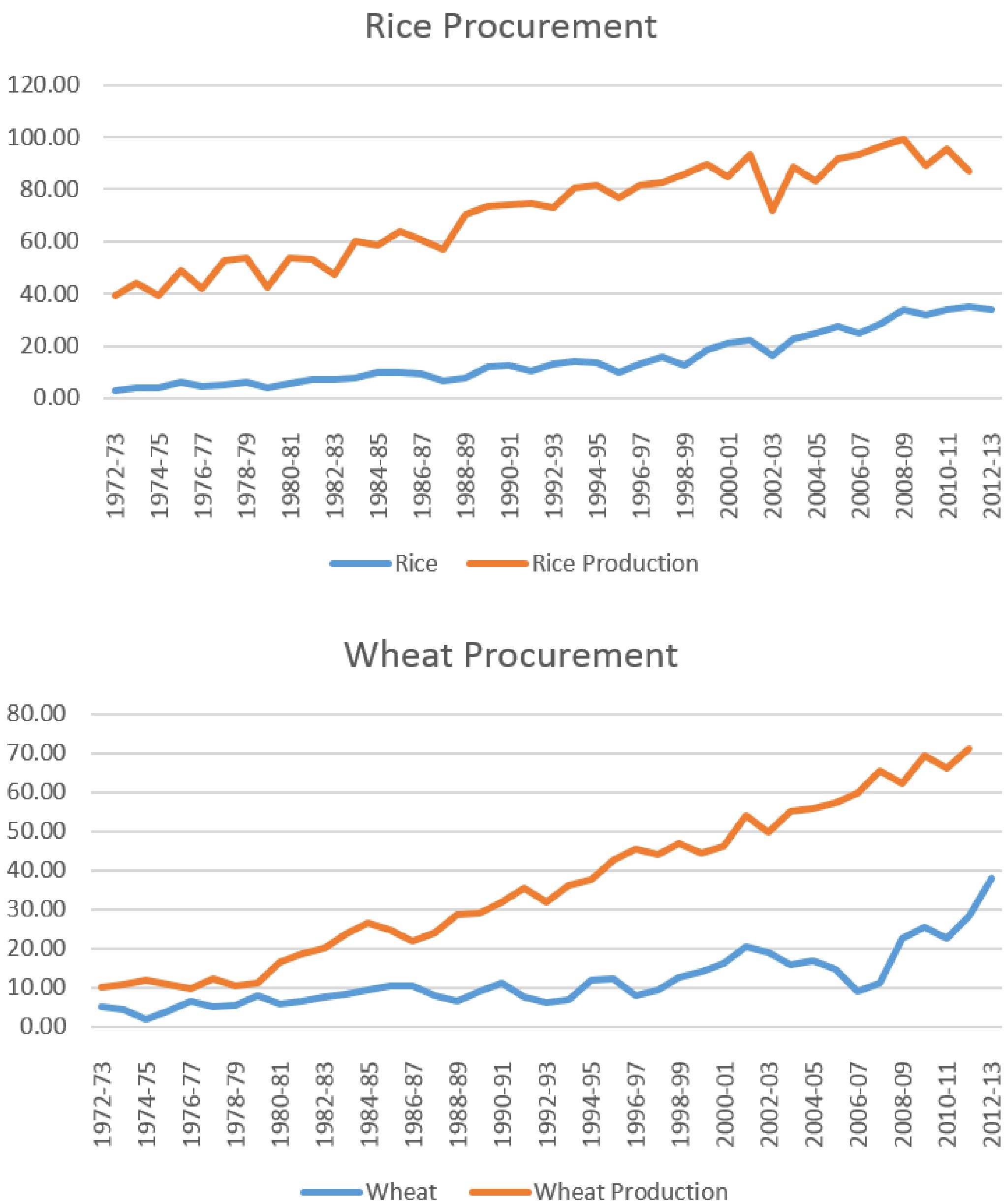


Figure 1.1: Amount of rice and wheat produced and procured (million tonnes)

Back to text

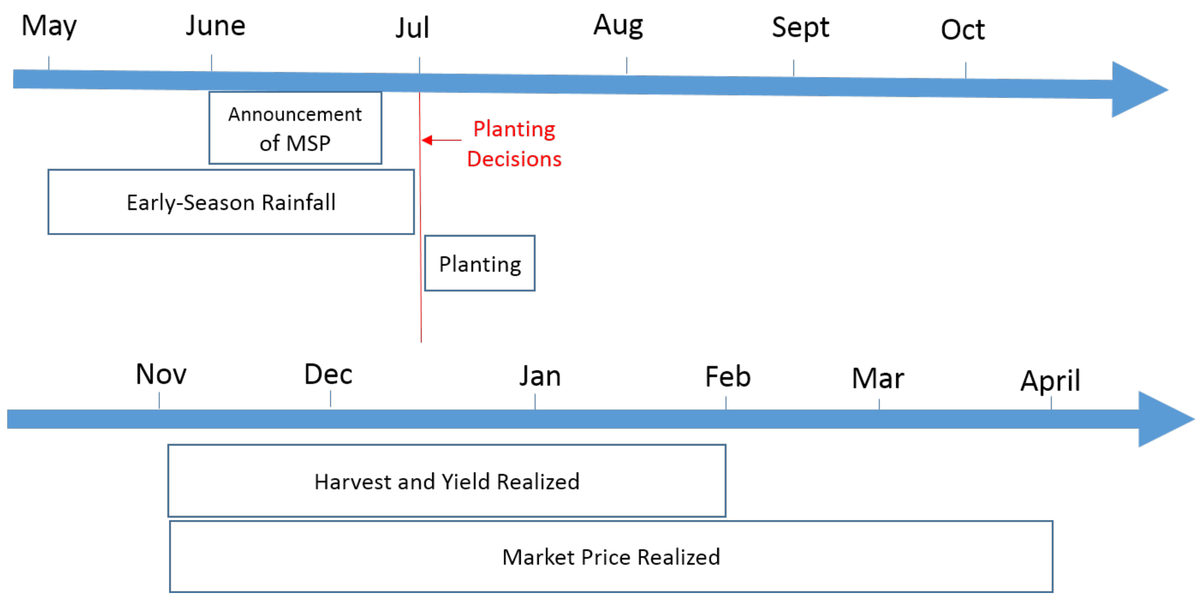


Figure 1.2: Timeline of events for the *Kharif* season.

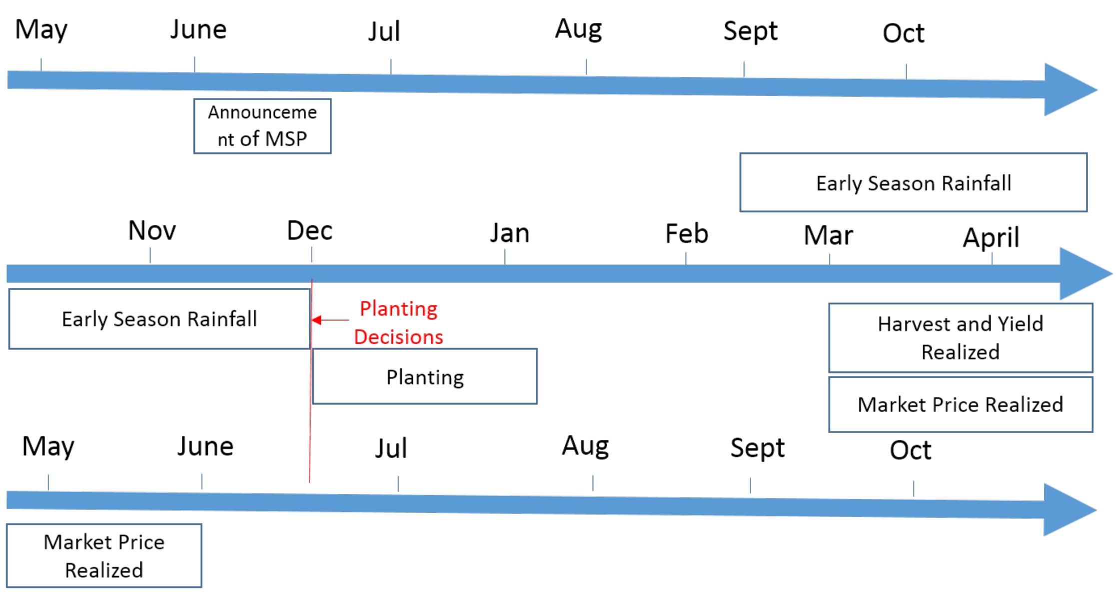


Figure 1.3: Timeline of events for the *Rabi* season.
Back to text

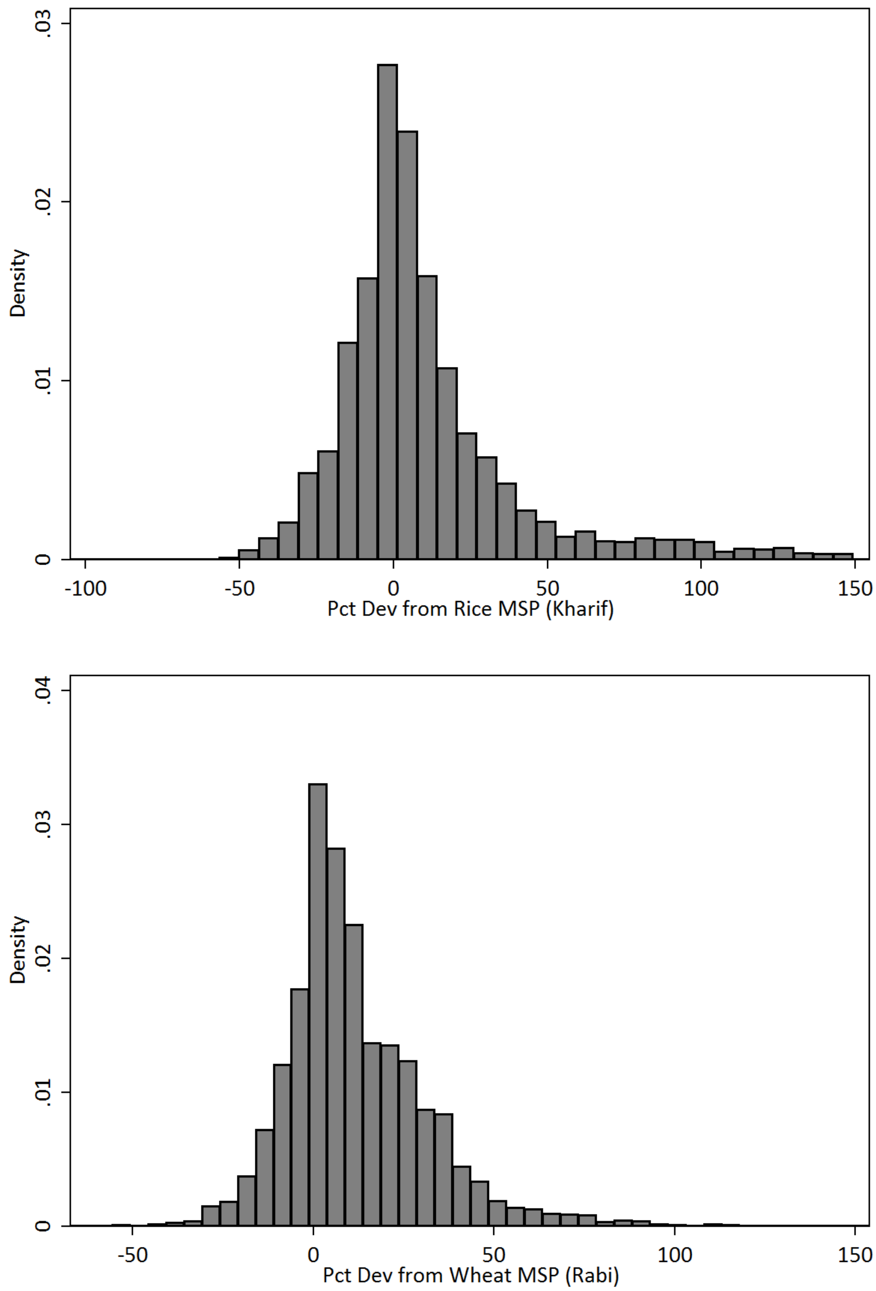
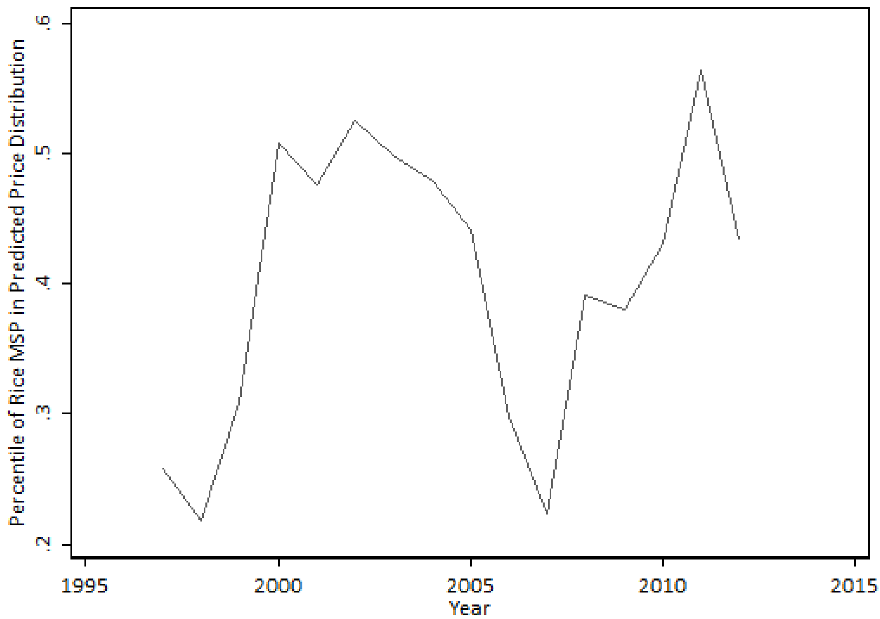


Figure 1.4: Distribution of percentage deviation of wholesale prices from MSP for all district-year observations between 1997 and 2012.

Back to text



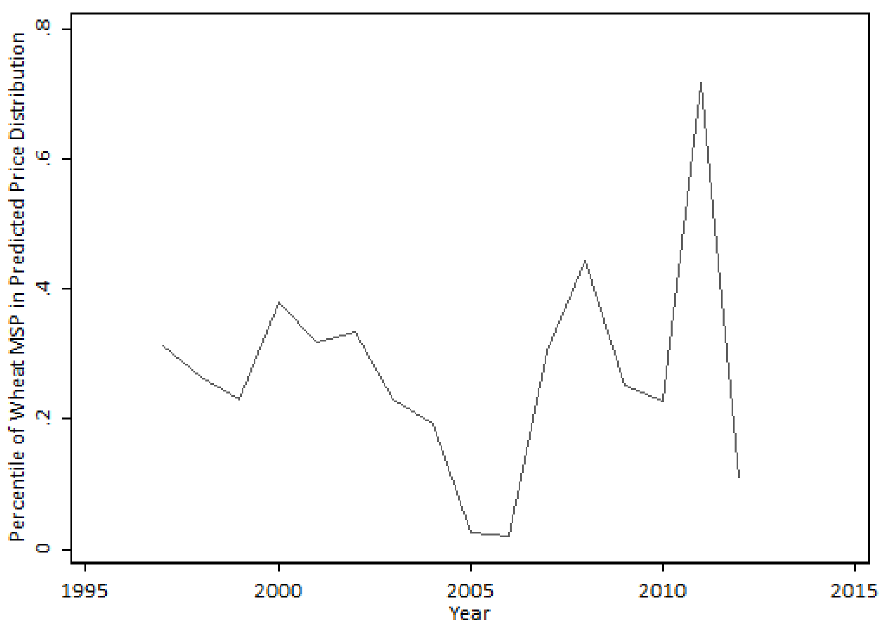


Figure 1.5: Percentile of MSPs in the price distribution for rice and wheat respectively.
Back to text

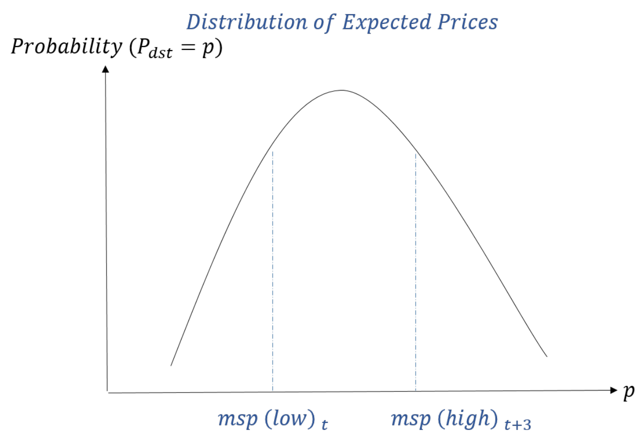


Figure 1.6: Illustration of variation in where MSP falls relative to the distribution of prices
Back to text

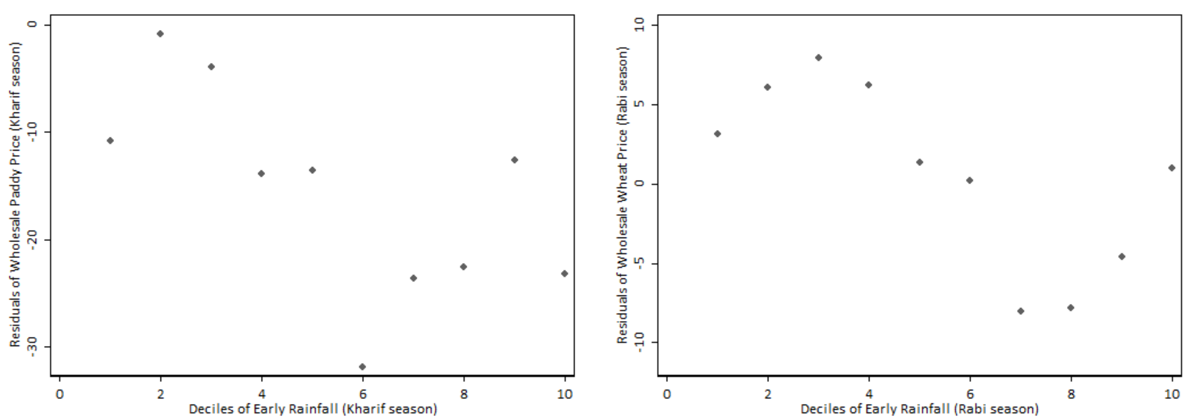


Figure 1.7: Mean binned residuals of wholesale price of Rice ((*Kharif* season) and Wheat (*Rabi* season) against deciles of early-season rainfall.
Back to text

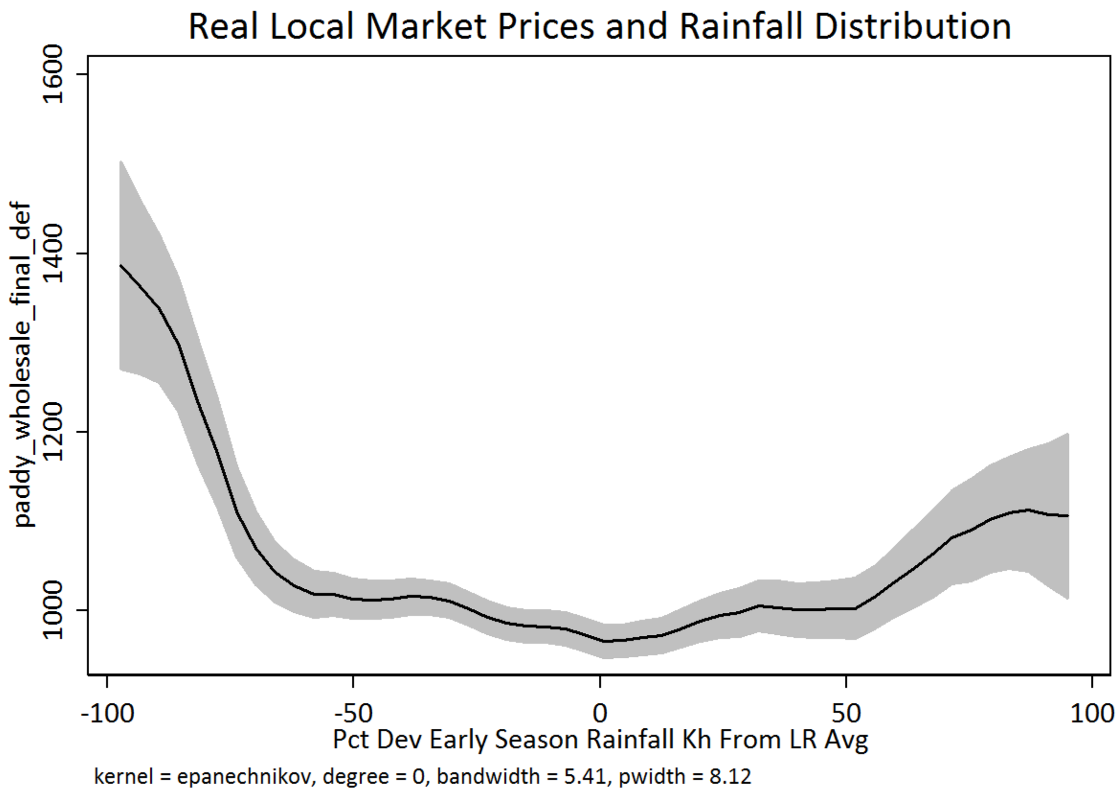


Figure 1.8: Local polynomial regressions of wholesale price of Rice ((*Kharif* season) against percent deviations from LR average of early-season rainfall.

Back to text

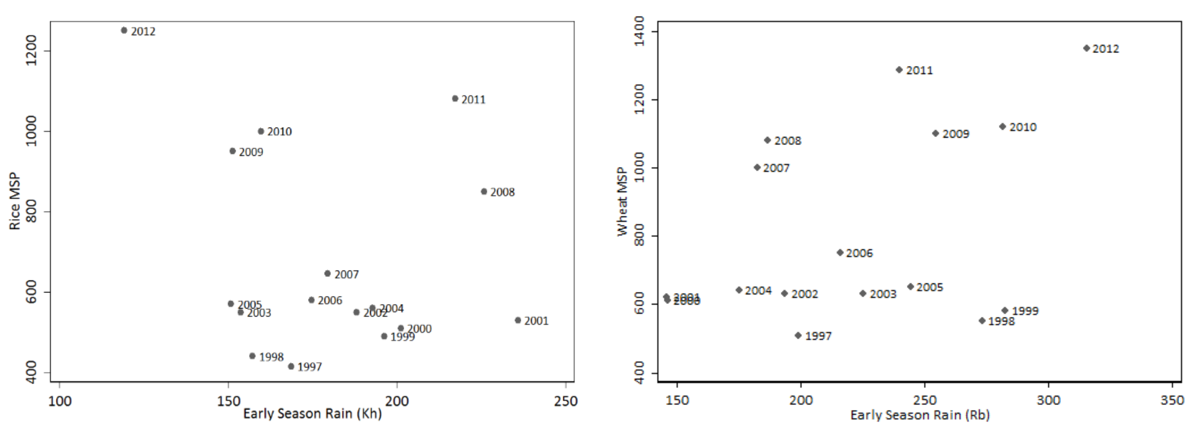


Figure 1.9: MSPs for rice (*Kharif* season) and wheat (*Rabi* season) plotted against early-season rainfall across the country between 1997 and 2012.

Note: Early-season rainfall is weighted by area cultivated in that season in a given district. Year t refers to the planting season t,t+1. For example, 2012 refers to the 2012-2013 season. Back to text

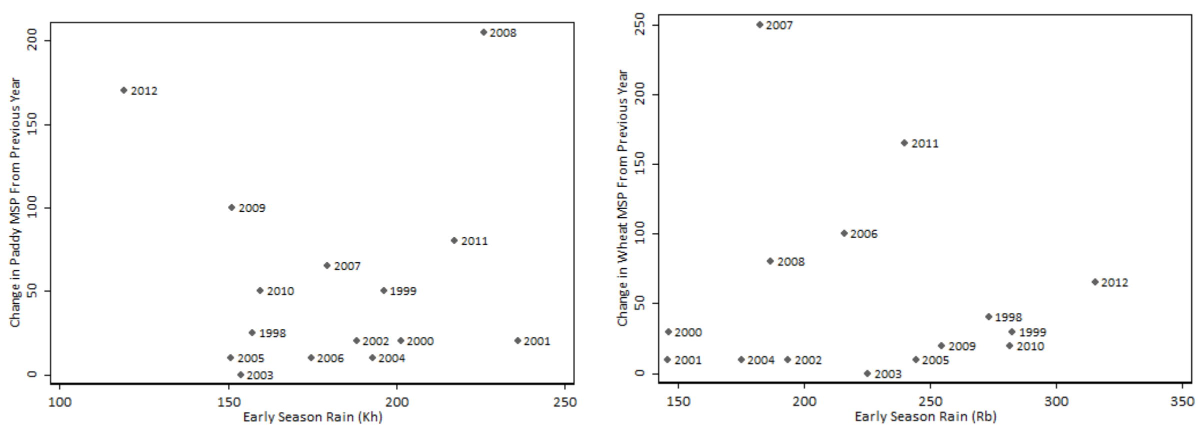


Figure 1.10: Change in MSPs for rice (*Kharif* season) and wheat (*Rabi* season) plotted against early-season rainfall across the country between 1997 and 2012.
Note: Early-season rainfall is weighted by area cultivated in that season in a given district. Year t refers to the planting season t,t+1. For example, 2012 refers to the 2012-2013 season. Back to text

1.11 Tables

Table 1.1: Response of Wholesale Rice Prices to Early Season Rainfall, *Kharif* season

	Continuous Definition of Rainfall				**Binary Definition of Rainfall**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SAMPLE	If Pct Dev Early Season Rainfall≥0	If Pct Dev Early Season Rainfall<0	Full Sample	Full Sample	Full Sample	Full Sample	Full Sample
VARIABLES	Wholesale Px At Harvest	Wholesale Px At Harvest	1(Binding)	1(Binding)	Wholesale Px At Harvest	Wholesale Px At Harvest	Wholesale Px At Harvest
Pct Dev Early Season Rainfall From LR Avg	0.269 (0.273)	-0.879*** (0.294)		0.000509* (.000299)			
1.(High Rice MSP)			0.0340** (0.0140)				
Pct Dev Early Season...Avg * 1.(High Rice MSP)				-0.000490 (0.000444)			
1.(Good Early Rain)					-39.53*** (11.69)		-38.62** (15.88)
1.(Above Lowest Quintile Early Season Rainfall)						-26.13*** (8.850)	
1.(High Rice MSP)* 1.(Good Early Rain)							-24.83 (15.12)
Observations	2,175	2,931	3,628	3,628	5,118	5,106	3,628
R-squared	0.705	0.789	0.486		0.722	0.722	0.745
Year FE	Yes	Yes	No	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses. Columns 1 and 2 show the response of rice wholesale prices to early-season productivity shocks for, respectively, the sample of districts with above-long-run-mean rainfall, and districts with below-long-run-mean rainfall. Columns 3 and 4 then provide a linear probability model that examines the first stage in the difference-in-difference strategy. Column 3 shows that the probability that the realized wholesale price binds (is lower than the support price) is higher in years defined as high price-support years. Column 3 does not include a year fixed effect, due to collinearity with the variable defining a given year as "high" or "low" price support. Column 4 shows that positive productivity shocks increase the probability that the support binds, in line with the decrease in price reflected in column 2, but that this is not significantly different in high and low support years. Then I move to my binary definition of productivity shocks, which I use in all tables that follow. Columns 5 and 6 show the responses of rice wholesale prices to two different (binary) definitions of positive productivity shocks. In Column 5, the productivity shock variable takes the value 1 if rainfall is above 50% below the long-run mean rainfall in that district. In Column 6, it takes the value 1 if rainfall is in the bottom quintile of the long-run rainfall distribution. Column 7 tests whether price responses to the rainfall shock defined in column 5 are significantly different in years defined as high and low price-support years, and provides support to Column 4 that this is not the case.

Back to text

Table 1.2: Agricultural Production, *Kharif* season

VARIABLES	(1) Rice Area Cultivated	(2) Rice Area Share	(3) Total Area Cultivated	(4) Rice Yield Per Unit Area	(5) Rice Production
1.(Good Early Rain)	-708.7	-0.0368***	4,219	-0.1184***	-20,615**
	(1,245)	(0.0109)	(2,967)	(0.0456)	(9,458)
1.(High Rice MSP)	3,342**	0.0397***	-2,715	0.1424***	18,682**
*1.(Good Early Rain)	(1,493)	(0.0135)	(3,441)	(0.0522)	(9,364)
Observations	3,608	3,608	3,608	3,608	3,608
R-squared	0.986	0.920	0.962	0.872	0.957
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0325	0.0657	-0.0126	0.0722	0.0853

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period.

Back to text

Table 1.3: Crop Mix, *Kharif* season

VARIABLES	(1) Rice	(2) Other Cereals	(3) Pulses	(4) Oilseeds/Cash Crops	(5) Spices
1.(Good Early Rain)	-0.0368***	-0.00567	0.0108**	0.0129**	-0.0002
	(0.0109)	(0.0045)	(0.0052)	(0.0058)	(0.0002)
1.(High Rice MSP)*1.(Good Early Rain)	0.0397***	0.0169**	-0.0083	-0.0253***	0.000625**
	(0.0135)	(0.0084)	(0.0064)	(0.0071)	(0.0003)
Observations	3,608	3,608	3,608	3,608	3,608
R-squared	0.920	0.927	0.803	0.894	0.828
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0657	0.112	-0.119	-0.154	0.270

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. The crops included in each category are detailed in the Data Appendix.

Back to text

Table 1.4: Low Vs. High Relative Suitability (Rice), *Kharif* season

VARIABLES	(1) Full Total Area Cultivated	(2) High Total Area Cultivated	(3) Low Total Area Cultivated	(4) Full Rice Area	(5) High Rice Area	(6) Low Rice Area	(7) Full Rice Prop of Area Cultivated	(8) High Rice Prop of Area Cultivated	(9) Low Rice Prop of Area Cultivated
1.(Good Early Rain)	4,219	3,756	5,556	-708.7	-2,141	224.5	-0.0368***	-0.0539***	-0.0272*
	(2,967)	(3,960)	(4,682)	(1,245)	(1,593)	(1,982)	(0.0109)	(0.0148)	(0.0155)
1.(High Rice MSP)*	-2,715	-8,388*	735.9	3,342**	4,851**	1,472	0.0397***	0.0765***	0.0119
1.(Good Rain Kh)	(3,441)	(4,983)	(5,001)	(1,493)	(2,192)	(2,209)	(0.0135)	(0.0219)	(0.0163)
Observations	3,608	1,919	1,689	3,608	1,919	1,689	3,608	1,919	1,689
R-squared	0.962	0.944	0.966	0.986	0.975	0.990	0.920	0.909	0.920
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0126	-0.0514	0.00270	0.0325	0.0530	0.0127	0.0657	0.109	0.0240

VARIABLES	(10) Full Rice Yield	(11) High Rice Yield	(12) Low Rice Yield	(13) Full Production of Rice	(14) High Production of Rice	(15) Low Production of Rice
1.(Good Early Rain)	-0.0796*	-0.0900	-0.0906	-13,893	-8,951	-20,347
	(0.0454)	(0.0600)	(0.0667)	(9,543)	(7,149)	(16,992)
1.(High Rice MSP)*	0.116**	0.174**	0.0907	14,919	16,935**	14,689
1.(Good Rain Kh)	(0.0534)	(0.0784)	(0.0745)	(9,626)	(8,355)	(16,638)
Observations	3,608	1,919	1,689	3,608	1,919	1,689
R-squared	0.872	0.878	0.875	0.956	0.964	0.952
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0593	0.0895	0.0465	0.0689	0.0856	0.0617

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

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. Yield and production specifications also include a cubic polynomial of monthly rainfall during the post-planting cultivation season. 'High' refers to districts with above-median suitability for rice, and 'low' to districts with below-median suitability for rice.

Back to text

Table 1.5: Monthly Per Capita Expenditure, *Kharif* season

	(1) Rice Households Trimmed	(2) Rice Households Winzorized	(3) Non-Rice Households Trimmed	(4) Non-Rice Households Winzorized
VARIABLES	Log(MPCE)	Log(MPCE)	Log(MPCE)	Log(MPCE)
1.(Good Early Rain)	0.00950	-0.00266	4.09e-05	-0.00313
	(0.0145)	(0.0148)	(0.0135)	(0.0140)
1.(High Rice MSP)	0.0589***	0.0518***	0.0205	0.0145
	(0.0167)	(0.0169)	(0.0164)	(0.0170)
1.(High Rice MSP)*	-0.0182	-0.0120	0.00863	0.0161
1.(Good Early Rain)	(0.0178)	(0.0181)	(0.0167)	(0.0177)
Constant	6.718***	6.823***	6.783***	6.840***
	(0.117)	(0.127)	(0.0790)	(0.0808)
Observations	37,652	38,034	71,321	72,037
R-squared	0.454	0.465	0.402	0.410
Early Rainfall in Prediction	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No
Rainfall	Yes	Yes	Yes	Yes
HH Char	Yes	Yes	Yes	Yes
Year FE	No	No	No	No
District FE	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district-year level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. I include only rural households that are surveyed in the harvest season and report consumption of home-produced rice (as rice households). Non-rice households are included if they produce at least one agricultural good. Household characteristics controlled for include religion, household type, household size, social group, and land possessed. All specifications also control for a cubic polynomial of monthly post-planting rainfall. Columns 1 and 3: I exclude the top 1% of per-capita expenditure observations. Columns 2 and 4: Per-capita expenditure is winzorized to the 99th percentile.

Back to text

Table 1.6: Labor Market Responses to Price Supports, *Kharif* season

	Overall		**Agriculture**			**Non-Agriculture**			**Non-Agriculture ASI Data**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	1.(Worked)	Worked Days	1.(Main Act Agri)	Agri Days (of past week)	Agri Wage	1.(Main Act Non-Agri)	Non-Agri Days (of past week)	NonAgri Wage	Total Days	Contract Days	Non-Contract Days
1.(Good Early Rain)	0.00692	0.167	-0.0641**	-0.448**	10.34**	0.0783***	0.548***	-25.07**	1,331	1,481	503.9
	(0.0164)	(0.114)	(0.0273)	(0.191)	(4.88)	(0.0248)	(0.173)	(12.51)	(1,870)	(2,413)	(1,226)
1.(High Rice MSP)*	0.00199	-0.114	0.0735**	0.512**	-9.91**	-0.0800***	-0.561***	30.56**	-4,560**	-5,617**	-1,609
1.(Good Early Rain)	(0.0190)	(0.132)	(0.0294)	(0.205)	(4.91)	(0.0265)	(0.185)	(13.90)	(2,155)	(2,738)	(1,205)
Observations	105,559	105,559	72,614	72,619	16,261	72,614	72,619	13,867	98,774	36,791	88,713
R-squared	0.326	0.517	0.251	0.250	0.514	0.228	0.228	0.442	0.191	0.166	0.188
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean			0.0982	0.0978	-0.206	-0.357	-0.358	0.230	-0.0940	-0.178	-0.0525

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. Columns 1-8 consider labor market outcomes from the NSS household survey of employment. I include only agricultural households in the rural sample that are surveyed in the kharif cultivation season. Columns (1) and (2) include all these individuals and columns (3)-(8) include only individuals who have indicated they have worked or searched for work in the past week. Standard individual and household controls - household size, household type, land possessed, social group, age, sex, education, religion and marital status- are included in all specifications. I also include cubic polynomials of monthly rainfall throughout the cultivation period. Wage regressions are restricted to individuals with non-zero wages. Columns (9)-(11) consider labor use in firms (both formal and informal) in the Annual Survey of Industries data. Column (9) considers total manufacturing days, while columns (10) and (11) consider permanent and contracted workers separately. The ASI analysis includes only open firms operating in the rural sector. It also includes a vector of firm-level controls, including industry, ownership, age, age squared, organization type, and number of plants.

Back to text

Table 1.7: Occupation Choice, *Kharif* season

VARIABLES	(1) Manufacturing 1.(Worked Manufacturing)	(2) Construction 1.(Worked Construction)
1.(Good Early Rain)	0.0191*	0.0128**
	(0.0104)	(0.0060)
1.(High Rice MSP)*1.(Good Early Rain)	-0.0228**	-0.0153**
	(0.0107)	(0.0064)
Observations	73,701	73,701
R-squared	0.381	0.051
Year FE	Yes	Yes
District FE	Yes	Yes
State x Time Trends	Yes	Yes
Proportion Mean	-0.300	-0.644

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. I include only agricultural households in the rural sample that are surveyed in the kharif cultivation season, and individuals who have indicated they have worked or searched for work in the past week. Standard individual and household controls - household size, household type, land possessed, social group, age, sex, education, religion and marital status- are included in all specifications. I also include cubic polynomials of monthly rainfall throughout the cultivation period.

Back to text

Table 1.8: Other Inputs, *Kharif* season

VARIABLES	(1) Proportion of Rice Area Irrigated	(2) Quant Fert Per Unit Area	(3) Proportion of Rice Area HYV
1.(Good Early Rain)	1.642*	-1.013	0.392
	(0.899)	(0.810)	(1.014)
Percentile of MSP*1.(Good Early Rain)	-0.0393*	0.0285	-0.0080
	(0.0223)	(0.0211)	(0.0249)
Observations	740	765	740
R-squared	0.904	0.959	0.928
Year FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes
Proportion Mean	-0.0543	0.0573	-0.0121

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. "Percentile of MSP" refers to the percentile of the support price in the predicted price distribution for the year.

Back to text

Table 1.9: Non-Agricultural Ouptut and Agricultural Productivity

	Non-Agricultural Output		**Agricultural Productivity**
	(1)	(2)	(3)
VARIABLES	Gross Output	Value-added	T-T Index
Good Early Rain	9.196e+07**	6.219e+07	0.6716***
	(4.462e+07)	(4.306e+07)	(0.2541)
1.(High Rice MSP)*1.(Good Early Rain)	-1.791e+08***	-6.250e+07	-0.9209***
	(6.174e+07)	(5.469e+07)	(0.2998)
Constant	2.775e+08	-2.058e+08	-1.550*
	(5.707e+08)	(3.951e+08)	(0.9295)
Observations	83,895	83,895	1,281
R-squared	0.067	0.038	0.376
Early Rainfall in Prediction	Yes	Yes	Yes
MSP in Prediction	No	No	No
Rainfall	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes
Firm Char	Yes	Yes	
Proportion SD	-0.0263	-0.0126	-0.822

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. Analysis on non-agricultural firms uses the ASI data, and includes only open firms operating in the rural sector. It also includes a vector of firm-level controls, including industry, ownership, age, age squared, organization type, and number of plants. Value-added is defined as total gross output minus total gross domestic inputs. In Column 3, I present results from my agricultural productivity analysis using the Tornqvist-Theil index. Details are provided in Appendix 4.2. All specifications include a cubic polynomial of rainfall in the cultivation period. In this table, I present results as a proportion of the standard deviation of the outcome variable, rather than its mean, due to the number of zero and negative observations in the value-added variable and in the Tornqvist-Theil index.

Back to text

Table 1.10: Response of Wheat Prices to Early Season Rainfall, *Rabi* season

	(1)	(2)	(3)	(4)	(5)
SAMPLE	If Pct Dev Early Season Rainfall≥0	If Pct Dev Early Season Rainfall<0	Full Sample	Full Sample	Full Sample
VARIABLES	Wholesale Price At Harvest	Wholesale Price At Harvest	Wholesale Price At Harvest	Wholesale Price At Harvest	Wholesale Price At Harvest
Pct Dev Early Season Rainfall From LR Avg	-0.0492	-0.321***			
	(0.0636)	(0.107)			
1.(Good Early Rain)			-25.68***		-39.37***
			(6.540)		(11.09)
1.(Above Lowest Quintile of Early Season Rainfall)				-15.78***	
				(4.564)	
1.(High Wheat MSP)* 1.(Good Early Rain)					10.88
					(18.05)
Constant	527.6***	525.6***	542.7***	531.0***	677.9***
	(12.62)	(13.43)	(8.743)	(8.033)	(16.54)
Observations	2,231	2,464	4,707	4,695	3,293
R-squared	0.921	0.901	0.904	0.904	0.897
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses.
Back to text

Table 1.11: Agricultural Production, *Rabi* season

VARIABLES	(1) Wheat Area Cultivated	(2) Wheat Area Share	(3) Total Area Cultivated	(4) Wheat Yield Per Unit Area	(5) Wheat Production
1.(Good Early Rain)	3,181**	-0.00806	6,591**	-0.0498	-4,009
	(1,228)	(0.00827)	(2,813)	(0.0289)	(4,936)
1.(High Rice MSP)*	354.5	0.0232*	-10,163	0.169***	25,138***
*1.(Good Early Rain)	(2,134)	(0.0137)	(5,389)	(0.0567)	(8,231)
Observations	2,598	2,598	2,598	2,598	2,598
R-squared	0.9364	0.875	0.960	0.936	0.988
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0800	0.0397	-0.0731	0.0800	0.0973

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero wheat production in the rabi season for all the years in the sample, and all districts in the states of Rajasthan and Madhya Pradesh. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period.

Back to text

Table 1.12: Low Vs. High Relative Suitability (Wheat)

VARIABLES	(1) Full Total Area Cultivated	(2) High Total Area Cultivated	(3) Low Total Area Cultivated	(4) Full Wheat Area	(5) High Wheat Area	(6) Low Wheat Area	(7) Full Wheat Prop of Area Cultivated	(8) High Wheat Prop of Area Cultivated	(9) Low Wheat Prop of Area Cultivated
1.(Good Early Rain)	6,690**	-4,138***	11,919***	3,865***	-1,117	7,110***	-0.0284**	-0.0503***	-0.00801
	(2,830)	(1,366)	(4,378)	(1,380)	(816.1)	(2,016)	(0.0111)	(0.0188)	(0.0127)
1.(High Wheat MSP)*	-6,675	7.352	-17,149	3,990	-392.0	3,176	0.0528***	0.0718*	0.0142
1.(Good Rain Rb)	(7,943)	(1,694)	(13,781)	(2,643)	(1,137)	(4,442)	(0.0192)	(0.0425)	(0.0247)
Observations	2,598	1,281	1,317	2,598	1,281	1,317	2,598	1,281	1,317
R-squared	0.958	0.982	0.946	0.988	0.992	0.986	0.809	0.777	0.837
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0480	8.25e-05	-0.0914	0.0478	-0.00578	0.0322	0.0903	0.113	0.0267
VARIABLES	(10) Full Wheat Yield	(11) High Wheat Yield	(12) Low Wheat Yield	(13) Full Production of Wheat	(14) High Production of Wheat	(15) Low Production of Wheat			
1.(Good Early Rain)	-0.0259	-0.0605	0.0126	9,536*	-1,938	18,505***			
	(0.0363)	(0.0755)	(0.0309)	(5,304)	(5,996)	(7,051)			
1.(High Wheat MSP)*	0.187***	0.228***	0.0973	10,703	1,329	3,367			
1.(Good Rain Rb)	(0.0552)	(0.0791)	(0.0725)	(10,695)	(8,759)	(17,628)			
Observations	2,598	1,281	1,317	2,598	1,281	1,317			
R-squared	0.907	0.873	0.945	0.982	0.989	0.979			
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes			
MSP in Prediction	No	No	No	No	No	No			
Year FE	Yes	Yes	Yes	Yes	Yes	Yes			
District FE	Yes	Yes	Yes	Yes	Yes	Yes			
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes			
Proportion Mean	0.0883	0.110	0.0450	0.0412	0.00618	0.0111			

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero wheat production in the Rabi season for all the years in the sample,and all districts in the states of Rajasthan and Madhya Pradesh. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period.'High' refers to districts with above-median suitability for wheat, and 'low' to districts with below-median suitability for wheat .

Back to text

Table 1.13: Monthly Per Capita Expenditure at Harvest, *Rabi* season

VARIABLES	(1) Wheat Households Winzorized	(2) Non-Wheat Households Winzorized
1.(Good Early Rain)	-65.86*	-10.59
	(34.79)	(40.61)
1.(High Wheat MSP)*1.(Good Early Rain)	72.17	-82.65
	(57.45)	(75.43)
Constant	909.6***	918.1***
	(111.3)	(173.1)
Observations	20,698	58,728
R-squared	0.533	0.256
Early Rainfall in Prediction	Yes	Yes
MSP in Prediction	No	No
Rainfall	Yes	Yes
HH Char	Yes	Yes
Year FE	Yes	Yes
District FE	Yes	Yes
State x Time Trends	Yes	Yes
Proportion Mean	0.0662	-0.0813

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero wheat production in the rabi season for all the years in the sample, and all households in the states of Rajasthan and Madhya Pradesh. I include only rural households that are surveyed in the harvest season and report consumption of home-produced wheat (as wheat households). Non-wheat households are included if they produce at least one agricultural good. Household characteristics controlled for include religion, household type, household size, social group, and land possessed. All specifications also control for a cubic polynomial of monthly post-planting rainfall. Per-capita expenditure is winzorized to the 99th percentile.

Back to text

Table 1.14: Labor Market Responses to Price Supports, *Rabi* season

	Agriculture			Non-Agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	1.(Main Activity Agri)	Agri Days	Agri Wage	1.(Main Activity Non-Agri)	Non-Agri Days	NonAgri
1.(Good Early Rain)	-0.00814	-0.0555	7.230	0.0101	0.0785	-8.643
	(0.0146)	(0.102)	(5.13)	(0.0141)	(0.0984)	(7.820)
1.(High Wheat MSP)*	0.0310	0.261*	-11.93	-0.0463**	-0.269**	6.095
1.(Good Early Rain)	(0.0202)	(0.141)	(8.18)	(0.0195)	(0.136)	(12.86)
Observations	40,453	40,455	8,681	40,453	40,455	10,771
R-squared	0.267	0.267	0.144	0.247	0.248	0.456
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0407	0.0488	-0.212	-0.224	-0.186	0.0446

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero wheat production in the rabi season for all the years in the sample, and all households in the states of Rajasthan and Madhya Pradesh. I include only agricultural households in the rural sample that are surveyed in the rabi cultivation season, and individuals who have indicated they have worked or searched for work in the past week. Standard individual and household controls - household size, household type, land possessed, social group, age, sex, education, religion and marital status- are included in all specifications. I also include cubic polynomials of monthly rainfall throughout the cultivation period.

Back to text

Table 1.15: Response of MSP to Various Potential Factors

PANEL A: Response of MSP to Early-Season Rainfall

VARIABLES	(1) Real Paddy MSP	(2) Change in Real Paddy MSP	(3) Real Wheat MSP	(4)) Change in Real Wheat MSP
Early Season Rain	0.261	0.564	0.512	-0.324
	(0.936)	(0.438)	(0.580)	(0.363)
Year Trend	10.81***	3.627	53.88***	7.598*
	(4.314)	(3.929)	(6.166)	(4.219)
Observations	16	15	16	15
R-squared	0.276	0.445	0.417	0.007

PANEL B: Response of MSP to Monsoon Forecasts

VARIABLES	(1) Real Paddy MSP	(2) Change in Real Paddy MSP	(3) Real Wheat MSP	(4)) Change in Real Wheat MSP
Num Days	0.261	0.564	0.512	-0.324
	(0.936)	(0.438)	(0.580)	(0.363)
Year Trend	10.81***	3.627	53.88***	7.598*
	(4.314)	(3.929)	(6.166)	(4.219)
Observations	16	15	16	15
R-squared	0.835	0.403	0.879	0.218

*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses. Early-season rainfall is weighted by area cultivated in that season in a given district. Monsoon forecasts in Panel B (num days) are defined as the number of days after the normal onset date that the monsoon is predicted to arrive (negative for early arrival).

Back to text

Table 1.16: Staple Production Response to Alternate Definition of Good Rain

PANEL A: Alternate Definition of Good Rain, *Kharif* season

VARIABLES	(1) Total Area Cultivated	(2) Rice Area	(3) Rice Proportion of Area Cultivated	(4) Rice Yield	(5) Production of Rice
1.(Above Lowest Quintile of Early Rainfall)	2,588	-400.8	-0.0304***	-0.0117	-6,336
	(2,564)	(1,244)	(0.00883)	(0.0406)	(7,702)
1.(Above lowest quintile...)* 1.(High Rice MSP)	-4,387	2,636**	0.0425***	0.0574	10,267
	(3,246)	(1,271)	(0.0114)	(0.0465)	(7,840)
Constant	351,784***	1,068	0.139***	1.448***	-78,719
	(24,173)	(14,747)	(0.0338)	(0.199)	(54,006)
Observations	3,597	3,597	3,597	3,597	3,597
R-squared	0.962	0.986	0.920	0.871	0.956
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0204	0.0257	0.0703	0.0295	0.0475

PANEL B: Alternate Definition of Good Rain, *Rabi* season

VARIABLES	(1) Total Area Cultivated	(2) Wheat Area	(3) Wheat Proportion of Area Cultivated	(4) Wheat Yield	(5) Production of Wheat
1.(Above Lowest Quintile of Early Season Rainfall)	4,394**	3,643***	0.00843	-0.00377	5,534
	(2,007)	(1,063)	(0.00727)	(0.0213)	(3,394)
1.(Above lowest quintile...)* 1.(High Wheat MSP)	-9,138**	565.1	0.0152	0.150***	17,909***
	(4,296)	(1,562)	(0.0117)	(0.0369)	(5,231)
Constant	58,602*	-11,108*	-0.300***	1.227***	33,031*
	(31,325)	(6,107)	(0.0860)	(0.292)	(17,889)
Observations	2,587	2,587	2,587	2,587	2,587
R-squared	0.960	0.989	0.875	0.935	0.987
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0656	0.00675	0.0260	0.0706	0.0688

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season (Panel A) or zero wheat production in the rabi season (Panel B) for all the years in the sample. Good rain is defined as rain in the 2nd through 5th quintiles of rainfall deviation from the long-run average. Panel B excludes all districts in the states of Madhya Pradesh and Rajasthan.

Back to text

Table 1.17: Rice Cultivation Responses using Various Predictions (*Kharif* season)

PANEL A VARIABLES	(1) Rice Area	(2) Rice Area	(3) Rice Area	(4) Total Area	(5) Total Area	(6) Total Area	(7) Rice Area Share	(8) Rice Area Share	(9) Rice Area Share
1.(Good Early Rain)	-650.3	-708.7	-2,559**	3,022	4,219	460.2	-0.0327***	-0.0368***	-0.0351***
	(1,403)	(1,245)	(1,172)	(3,215)	(2,967)	(2,728)	(0.0102)	(0.0109)	(0.0127)
1.(High Rice MSP)	4,099**	3,342**	6,002***	-789.7	-2,715	3,513	0.0412***	0.0397***	0.0335**
*1.(Good Early Rain)	(1,709)	(1,493)	(1,588)	(4,097)	(3,441)	(3,823)	(0.0155)	(0.0135)	(0.0149)
R-squared	0.986	0.986	0.986	0.962	0.962	0.962	0.920	0.920	0.920
Proportion Mean	0.0398	0.0325	0.0583	-0.00368	-0.0126	0.0164	0.0682	0.0657	0.0555
PANEL B VARIABLES	(10) Rice Yield	(11) Rice Yield	(12) Rice Yield	(13) Rice Production	(14) Rice Production	(15) Rice Production			
1.(Good Early Rain)	-0.199***	-0.180***	-0.202***	-29,099***	-26,733***	-28,565***			
	(0.0427)	(0.0489)	(0.0512)	(9,565)	(9,911)	(10,800)			
1.(High Rice MSP)*	0.219***	0.140***	0.158***	28,613***	18,446**	19,224*			
1.(Good Early Rain)	(0.0521)	(0.0527)	(0.0548)	(9,101)	(9,220)	(9,989)			
R-squared	0.8785	0.8778	0.8779	0.9573	0.9572	0.9572			
Proportion Mean	0.113	0.0718	0.0814	0.132	0.0853	0.0889			
Observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608	3,608	3,608
Early Rainfall in Prediction	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
MSP in Prediction	No	No	Yes	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period.

Back to text

Table 1.18: Wheat Cultivation Responses using Various Predictions (*Rabi* season)

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Wheat Area	Wheat Area	Wheat Area	Total Area	Total Area	Total Area	Wheat Area Share	Wheat Area Share	Wheat Area Share
1.(Good Early Rain)	3,578***	3,181**	2,518**	6,853**	6,591**	4,843**	-0.00355	-0.00806	-0.00725
	(1,316)	(1,228)	(1,164)	(3,041)	(2,813)	(2,228)	(0.00758)	(0.00827)	(0.00881)
1.(High Rice MSP)	-756.0	354.5	2,946	-7,762	-10,163	-4,092	0.00479	0.0232*	0.0216
*1.(Good Early Rain)	(1,685)	(2,134)	(2,983)	(5,389)	(7,361)	(6,622)	(0.0129)	(0.0137)	(0.0168)
R-squared	0.989	0.989	0.989	0.960	0.960	0.960	0.875	0.875	0.875
Proportion Mean	-0.00905	0.00425	0.0353	-0.0558	-0.0731	-0.0294	0.00818	0.0397	0.0370
PANEL B	(10)	(11)	(12)	(13)	(14)	(15)			
VARIABLES	Wheat Yield	Wheat Yield	Wheat Yield	Wheat Production	Wheat Production	Wheat Production			
1.(Good Early Rain)	-0.0704**	-0.0498	0-.0293	30.47	-4,009	-5,627			
	(0.0277)	(0.0289)	(0.0281)	(5,497)	(4,936)	(5,054)			
1.(High Rice MSP)*	0.167***	0.169***	0.110*	7,444	25,138***	33,091***			
1.(Good Early Rain)	(0.0560)	(0.0567)	(0.0607)	(7,357)	(8,231)	(9,063)			
R-squared	0.9364	0.9364	0.9367	0.9877	0.9878	0.9878			
Proportion Mean	0.0788	0.0800	0.0519	0.0288	0.0973	0.1281			
Observations	2,598	2,598	2,598	2,598	2,598	2,598	2,598	2,598	2,598
Early Rainfall in Prediction	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
MSP in Prediction	No	No	Yes	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero wheat production in the rabi season for all the years in the sample, and districts in the states of Rajasthan and Madhya Pradesh. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period.

Back to text

Table 1.19: Rice Yield Reponses by Size of the Farmers' Information Set, *Kharif* season

	(1) Three-Year Recall	(2) Five-Year Recall	(3) Seven-Year Recall
VARIABLES	Rice Yield	Rice Yield	Rice Yield
1.(Good Early Rain)	-0.1330***	-0.1184***	-0.1969***
	(0.0434)	(0.0456)	(0.0589)
1.(High Rice MSP)*1.(Good Early Rain)	0.0953*	0.1424***	0.1979***
	(0.0535)	(0.0522)	(0.0633)
Observations	4,180	3,608	2,904
R-squared	0.875	0.872	0.881
Year FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes
Proportion Mean	0.0486	0.0722	0.0983

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. Specifications include controls for a cubic polynomial of post-planting rainfall.

Back to text

Table 1.20: Staple Production During Export Bans

PANEL A: Rice Production During Export Ban (2008-2011), *Kharif* season

VARIABLES	(1) Total Area Cultivated	(2) Rice Area	(3) Rice Proportion of Area Cultivated	(4) Rice Yield	(5) Production of Rice
1.(Good Early Rain)	5,964	-3,527**	-0.0566***	-0.0540	-7,828
	(4,266)	(1,672)	(0.0181)	(0.0448)	(9,333)
1.(High Rice MSP)* 1.(Good Rain Kh)	-9,182	6,373**	0.0463	0.126	12,187
	(6,275)	(2,649)	(0.0286)	(0.0889)	(11,558)
Constant	94,524	-159,103***	0.216**	3.577***	-298,346*
	(60,033)	(40,774)	(0.0981)	(0.855)	(170,340)
Observations	1,353	1,353	1,353	1,353	1,353
R-squared	0.972	0.990	0.924	0.927	0.968
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0427	0.0623	0.0756	0.0640	0.0546

PANEL B: Wheat Production During Export Ban (2007-2009), *Rabi* season

VARIABLES	(1) Total Area Cultivated	(2) Wheat Area	(3) Wheat Proportion of Area Cultivated	(4) Wheat Yield	(5) Production of Wheat
1.(Good Early Rain)	8,656*	4,771	-0.0270	-0.0126	10,138
	(4,761)	(3,937)	(0.0183)	(0.0789)	(13,036)
1.(High Wheat MSP)* 1.(Good Early Rain)	2,003	199.4	-0.0121	0.259***	16,666
	(5,240)	(4,201)	(0.0315)	(0.0961)	(16,048)
Constant	78,994***	-1,657	0.0555**	1.083***	-11,094
	(5,655)	(4,319)	(0.0226)	(0.0863)	(14,213)
Observations	757	757	757	757	757
R-squared	0.987	0.995	0.901	0.935	0.994
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0154	0.00247	-0.0201	0.125	0.0662

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season (Panel A) or districts that have zero wheat production in the rabi season and all districts in the states of Rajasthan and Madhya Pradesh (Panel B) for all the years in the sample. Specifications for yield and production include a cubic polynomial of monthly post-planting precipitation during the cultivation period.

Back to text

Table 1.21: Tests on Various Subsamples of Wheat-Producing Districts

VARIABLES	(1) Test1a Wheat Area	(2) Test1b Total Area Cultivated	(3) Test1c Wheat Yield	(4) Test1d Production of Wheat	(5) Test2a Wheat Area	(6) Test2b Total Area Cultivated	(7) Test2c Wheat Yield	(8) Test2d Production of Wheat	(9) Test3a Wheat Area	(10) Test3b Total Area Cultivated	(11) Test3c Wheat Yield	(12) Test3d Production Wheat
1.(Good Early Rain)	576.2	1,021	-0.0793**	-8,774*	6,143**	2,290	-0.138***	5,743	1,392	-1,127	-0.133*	-18,195
	(994.6)	(2,465)	(0.0323)	(4,687)	(2,509)	(3,042)	(0.0499)	(8,616)	(3,455)	(5,322)	(0.0778)	(16,531)
1.(High Wheat MSP)*	3,148	7,036	0.240***	34,140***	-614.3	-1,425	0.361***	23,097	81.64	-4,269	0.307**	31,569
1.(Good Rain Rb)	(2,708)	(6,801)	(0.0548)	(11,590)	(3,274)	(4,488)	(0.0996)	(14,248)	(4,905)	(7,975)	(0.122)	(25,225)
Observations	1,403	1,403	1,403	1,403	522	522	522	522	279	279	279	279
R-squared	0.990	0.977	0.942	0.987	0.986	0.987	0.954	0.981	0.990	0.992	0.964	0.985
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

VARIABLES	(13) Test4a Wheat Area	(14) Test4b Total Area Cultivated	(15) Test4c Wheat Yield	(16) Test4d Production of Wheat	(17) Test5a Wheat Area	(18) Test5b Total Area Cultivated	(19) Test5c Wheat Yield	(20) Test5d Production of Wheat
1.(Good Early Rain)	3,986***	8,629**	-0.0201	5,147	-388.7	-5,729	0.0637	-1,202
	(1,497)	(3,460)	(0.0253)	(5,204)	(2,003)	(3,569)	(0.0741)	(9,699)
1.(High Wheat MSP)*	964.3	-10,677	0.225***	26,157***	-1,053	-1,593	-0.00319	10,848
1.(Good Rain Rb)	(2,529)	(8,792)	(0.0519)	(9,827)	(4,102)	(5,554)	(0.132)	(19,304)
Observations	1,773	1,773	1,773	1,773	658	658	658	658
R-squared	0.987	0.943	0.929	0.985	0.997	0.991	0.954	0.994
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero wheat production in the rabi season for all the years in the sample. Test 1: Restricting to years in which the rice price floor is low. Test 2: Restricting to districts in which the proportion of area cultivated with rice is less than 50% in the Kharif season and proportion of area cultivated with wheat in the rabi season is more than 50%. Test 3: Identical to test 2, but restricted to years in which the rice price floor is low. Test 4: Restricted to top 8 wheat producing states, excluding Madhya Pradesh and Rajasthan. Test 5: A falsification test that restricts analysis to years in which the paddy support price is high and to districts in which the proportion of area cultivated with rice in the Kharif season is greater than 50%. For test 5 alone, we observe no significant effect on wheat cultivation from the wheat support price, presumably due to a spillover effect from the Kharif season.

Back to text

Chapter 2. Scabs: Norm-driven Supression of Labor Supply

(with Emily Breza and Supreet Kaur)

2.1 Introduction

Traditional theories of the labor market presume that workers are atomistic individual agents, whose labor supply curve is pinned down solely by individual preferences and alternate sources of income. However, a long tradition of work in social science posits that worker behavior shows a tendency to deviate from this individualistic approach—for example, through phenomena such as coordinated restriction of output, strikes, and other collective behaviors. The labor economics literature has explored these outcomes within the context of formal unions (Farber, 1986). A broader set of literatures—spanning economics, political science, sociology, and psychology—suggests that such collective behaviors emerge even in the absence of any formal labor organization. This view has focused on the role of norms as a coordinating device that serves to constrain individual worker behavior.[88] For example, in *The Labor Market as a Social Institution*, Robert Solow argues that the social norms that arise in the workplace are inherent to what distinguishes labor markets from other commodity markets, and are important for understanding labor market equilibria.

This paper empirically examines community-wide norms against accepting wage cuts in a developing country context. Specifically, we test whether such implicit collusion distorts labor supply—preventing workers from accepting jobs at wages below the prevailing wage during times of unemployment. The setting for our study is informal markets for casual daily labor in India. Such markets are ubiquitous in poor countries, serving as the primary channel for hired employment for hundreds of million workers in India alone (National Sample Survey 2010). They are characterized

[88]We are agnostic about whether norms are the result of innate human preferences, for example for fairness, or are simply a way to describe a set of equilibrium strategies. A rich body of theoretical work models how, in the absence of any formal organization, collective behaviors can be sustained in equilibrium through sanctions or the internalization of norms into utility (Kandori, 1992; Ellison, 1994; MacLeod, 2007).

by a high degree of decentralization and informality: employment is bilaterally arranged between individual employers and laborers in spot contracts (typically lasting 1-3 days), with frequent resorting between employers and workers. Unions or other formal organizations are virtually nonexistent, and minimum wage laws are largely ignored (Rosenzweig 1980, 1988; Dreze and Mukherjee 1986). However, observational evidence points to a lack of wage flexibility in these markets—both across workers and across time (Dreze et al., 1986; Kaur, 2018; Breza et al., 2018).

Figure 2.1 provides initial suggestive support for the presence of norms against accepting wage cuts in this setting, using the approach developed in Kahneman et al. (1986). In a survey conducted with Indian agricultural laborers, over 80% of respondents said it was "unacceptable" or "very unacceptable" for an unemployed worker to offer to work at a rate below the prevailing wage (Panel A). In addition, about 80% of respondents stated that other workers in the village would become angry with a villager who accepted work below the prevailing wage (Panel B). While only speculative, these responses suggest that laborers view working for a lower wage, even when unemployed, as a violation which could result in sanctions.

In this paper, we use a field experiment to test whether these forces actually constrain workers' labor supply decisions. In doing so, we do not take a strong stance on why the norms arise. For example, they may stem from or be supported through primitives in utility. Alternately, they may simply describe an equilibrium strategy that enables decentralized workers to behave as a union (Kandori, 1992; Ellison, 1994). Moreover, these two views are not necessarily contradictory (e.g. MacLeod, 2007). Below, we discuss possible interpretations in light of our empirical results. However, our primary focus is on documenting the existence of these social forces and their effects on labor supply.

Specifically, we hypothesize that during times of unemployment, (at least some) workers find it privately optimal to take up jobs at wages lower than the prevailing wage, but are less likely to do so because this would be perceived as a norm violation, resulting in sanctions from co-villagers.

To test this hypothesis, we proceed in two steps. The main part of our study is a field experiment with existing employers in which we vary two aspects of job offers: (i) the level of the wage and (ii) the extent to which the job offer is observable to other workers. In addition, using a supplementary exercise, we document workers' willingness to destroy surplus in order to sanction those who have accepted wage cuts. We describe each of these components in turn below.

In this context, agricultural employers hire laborers in their village in one-day spot contracts to work on their land and perform a given cultivation activity (e.g. weeding). To implement the field experiment, we partner with such employers. We induce two types of variation during their hiring process. (i) First, the job is offered at a random wage level: at the prevailing wage, or 10% below the prevailing wage.[89] (ii) Second, we vary the extent to which the wage level is publicly observable: whether the job offer is made inside the worker's home or outside on the street where neighbors (who are typically other workers) can overhear the offer.

All offered jobs correspond to actual employment opportunities on the employer's land—so that our data reflects real employment decisions by workers. Treatment randomization is at the village level, so that all workers within a given village receive the same wageXobservability condition. In addition, the workers in our experiment (i.e. those who are offered jobs) are sampled randomly from the village population of laborers.[90] Aside from hiring, we are not involved in any other aspect of the employment relationship: employers supervise workers as usual, provide them food, etc. The experiment is conducted across 183 villages (i.e. labor markets) with 183 distinct partnering employers (one in each village), with jobs extended to 502 workers.

We predict that, when assured privacy, at least some unemployed workers will choose to accept jobs below the prevailing wage. However, if other villagers can observe their decision, this will dampen

[89]In this setting, there is a prevailing daily wage for each type of agricultural task. We provide direct evidence for this below.

[90]As is typical, hiring is done by employers who approach laborers at their homes to make job offers. We randomly select among workers who are home at the time of hiring. We provide several robustness checks to compare these workers with the overall population.

their willingness to accept a job at this lower wage. In contrast, we predict that observability will not decrease take-up of jobs at the prevailing wage—since taking up these jobs does not constitute a norm violation.

Our results are consistent with these predictions. At the prevailing wage, the average take-up rate of jobs is 26%, and we cannot reject that this take-up rate is the same regardless of whether the job offer is publicly observable by others.[91] In contrast, when a worker is offered a job below the prevailing wage, take-up depends crucially on whether his decision is publicly observable. When the lower wage is offered in private, take-up remains a robust 18%. However, this falls by 13.6 percentage points when low-wage offers are observable (significant at 1% level). When restricting the sample to workers who are in the agricultural labor market—defined as those who consider agricultural labor as their primary or secondary occupation—these results become even starker: the acceptance of wage cuts is 20.6% in private vs. 1.8% in public.

This distortion on individual labor supply is economically meaningful. The experiment was conducted during the lean season, when workers typically only find a few days of employment in a week. Consequently, passing up one day of work at a 10% wage cut is equivalent to foregoing 38% of average weekly earnings in our sample. This is a large magnitude, especially given that workers report skipping meals and struggling for cash during this time. Our experimental results suggest that, in our sample, 13.8% of workers (i.e. those who would have accepted the job in private but do not do so in public) choose to forego these earnings in order to avoid being seen as violating the village norm.

To provide positive evidence on the willingness to sanction those who violate labor supply norms, we use a supplementary exercise. In another set of villages—drawn from the same population as our study villages—we partner with employers to make private job offers to a random subset of

[91]Note that, even under the prevailing wage, we would not expect take-up to be 100%. Recall that we sampled randomly from the labor force in each village when making job offers. Workers may decline the job because they have another work activity already lined up, or because their reservation wage is higher than the prevailing wage.

laborers ("workers") at varying wage rates within each village. We then play a costly punishment game with another random subset of laborers ("players") in each village who were *not* offered jobs. Each player is paired with an anonymous worker and told that the worker is either in the player's own village, or in a village that is geographically far away. To implement the game, the player is told that his paired worker accepted a job at either (a) the prevailing wage or (b) 10% below the prevailing wage. The player can then give up some of his endowment to reduce the endowment of his paired worker.

As expected, we find that there is no punishment of workers who accept jobs at the prevailing wage. In contrast, when paired with a worker who accepted a wage cut, players punish the worker 37% of the time. When players do punish, the amount of money they deduct from those who violated the norm corresponds to 37.2% of average daily labor market earnings in our sample. In order to impose this punishment on their partner, the amount that players forego from their own endowment corresponds to 7.4% of typical daily earnings. Finally, we find that the desire to punish norm violations is not limited to workers in one's own village. Players also punish workers from distant villages who accepted a wage cut—even though that worker's action has no scope to affect the player's own labor market.

These results are consistent with the literature on social preferences, which indicates that individuals will be willing to destroy their own surplus to punish those who have engaged in norm violations (Charness and Rabin, 2002). Our findings are also consistent with contagious punishment models (Ellison, 1994), in which norms are an equilibrium strategy that is enforced through decentralized sanctions. We should note, however, that the willingness to punish those in other labor markets—where the deviating party's actions have no scope to affect one's own payoffs—is particularly consistent with villagers viewing norm violations in moral or general terms.

Our findings relate to the literature on wage adjustment and labor market distortions in poor countries. Early work in development economics focused heavily on the observation that wages in poor countries appear downwardly rigid, potentially contributing to high levels of involuntary unemployment (Lewis, 1954; Eckaus, 1955; Leibenstein, 1957). Recent empirical evidence documents that downward nominal wage rigidity continues to be relevant in village labor markets today, with consequences for unemployment levels (Kaur, 2018). The presence of rigidities in this setting has been a long-standing puzzle in the development literature. A substantial body of theoretical work has proposed various micro-foundations for rigid wages (Shapiro and Stiglitz, 1984; Dasgupta and Ray, 1986). However, many of these proposed micro-foundations, such as nutrition efficiency wages, have not withstood empirical scrutiny (Rosenzweig, 1988). To date, there is scant empirical evidence supporting any micro-foundation for why wage floors should arise in this setting. Osmani (1990) offers a model based on informal worker collusion that theoretically reconciles the different stylized facts about wage adjustment in this setting. Our study provides the first empirical test of this mechanism.

Our study also has bearing on the labor literature on formal and informal unions. While a large literature has sought to understand formal unions in developed countries (see Farber and Saks, 1980; Dickens et al., 2007), there has been less work on such forces in developing countries and also limited work in economics documenting the role of informal unions more broadly. Casual labor markets in poor countries display many of the same characteristics that are often rationalized by unions in developed markets: wage rigidity and wage compression (Kaur, 2018; Breza et al., 2016; Dreze et al., 1986). Consequently, documenting the presence of informal unions in our setting suggests that some of the considerations historically attached to formal unions may apply more broadly in the labor market.

Finally, our project relates to the literature on fairness norms and markets. Recent work in behavioral economics—building on psychology, sociology, and organizational behavior—has advanced the notion that fairness norms potentially affect wage and employment behavior (Akerlof and Yellen, 1990; Fehr et al., 2009; Kahneman et al., 1986; Card et al., 2012; Kaur, 2018; Breza et al., 2016).

This paper directly documents the presence of a community-wide norm against accepting wage cuts, and shows that this norm has a distortionary effect on labor supply in a high-stakes field setting: employment in real jobs. Our finding that workers are willing to give up a large amount of earnings in order to avoid being seen as violating the norm, coupled with costly punishment of deviators, supports the view that such norms have relevance for labor market behaviors.

While our design enables us to understand whether workers' labor supply is affected by collusive pressure, our evidence does not allow us to make predictions about what wage levels would exist in equilibrium in the absence of such pressure. Such predictions would require understanding the demand side of the market, which is outside the scope of our study. In addition, our findings will of course be specific to the five Indian districts (and 183 villages) in which our study was conducted. However, the features of our setting, such as wage rigidity and low employment rates, are mirrored across India and in other developing countries (Kaur, 2018; Beegle et al., 2015). Providing the first piece of evidence for a potential micro-foundation for wage rigidity in such a setting would advance the literature and suggest an exploration of this micro-foundation in other locations as well.

The paper proceeds as follows. In Section 2.2, we describe the setting, research hypotheses, and experimental design. We present the results of the main experiment in Section 2.3, and the costly punishment game in Section 2.4. Section 2.5 discusses potential threats to validity. Section 2.6 concludes.

2.2 Setting and Experimental Design

2.2.1 Overview of Setting

The study takes place rural Odisha, one of India's most underdeveloped states. Markets for casual daily wage labor are extremely active, and provide the primary source of wage labor earnings for residents in the area. A large proportion of workers in construction, unskilled manufacturing, and other factories are hired through these labor markets. In the data collected by Breza et al. (2016), there appears to be a wage floor. This floor coincides with the prevailing wage in agriculture. Concretely, denote as W the prevailing wage in agriculture in a village. When workers from that village work in the non-agricultural sector – largely in jobs that take place outside the village – the wage they earn is at or above W in 98% of cases. This is despite the fact that unemployment levels appear high. Employment rates (in terms of total worker-days across all sectors) are below 50%. A striking 80% of workers report being involuntarily unemployed at least one day in the past two weeks. Understanding the source of the wage floor inside the village therefore has potential bearing on understanding determinants of the wage in the labor market as whole.

2.2.2 Research Hypotheses

As before, we denote the prevailing village wage as W. If worker collusion contributes to downward wage rigidity at W, then we hypothesize that during times of high unemployment, (at least some) workers would find it privately optimal to take up jobs at wages lower than the prevailing wage, but do not do so because this would result in sanctions from co-villagers. Specifically, we predict:

H1.) The true private opportunity cost of working for a subset of individuals is less than W – i.e., workers will be privately willing to accept work at wages below W.

H2.) When other workers can observe an individual's job take-up decision, workers will be less likely to accept work below W.

H3.) Workers will sanction others who have accepted work below W.[92]

We construct a design to test these hypotheses, and rule out confounding factors.

[92]We are agnostic as to how the community sustains the costly punishment of deviations by workers. Punishment may be a community norm or may be enforced through mechanisms similar to Ellison (1994) and Kandori (1992).

2.2.3 Experimental Design

Our experiment takes place in five rural districts in Orissa, India. In each study village, we partner with a local agricultural employer (i.e. landowner in that village). We induce experimental variation in the wage rate offered by the employers, and in the observability of these offers to other laborers in the community. Partner employers typically hire daily-wage laborers for tasks like weeding and field preparation. Our experiment involves measuring the job take-up of each worker approached by the employer under the different treatment conditions below. Importantly, workers in the experiment make decisions about real jobs, working for an actual local employer who is typically familiar to the workers.

The research design requires that we have full experimental control over offered wages and the observability of those offers. In exchange for this control, we subsidize the cost of the labor for the employers. The size of the employer's contribution is the same regardless of the size of the wage offer. This allows us to make employment offers in some treatment arms, described below, without the employer knowing the wage.[93] Note that because we care about the labor supply side only, internal validity is not affected by the fact that the employer is being compensated for his cooperation (see Section VII).

Our full experimental design is presented in Figure 2.2. The core experiment follows a 2x2 design (treatment cells A-D). We incorporate 2 supplemental treatment cells (D and E) to allow for additional tests. Note that randomization is at the village (i.e. labor market) level, so that only one treatment cell will be implemented in any given village.

The first dimension of exogenous variation in the core 2x2 design sets the wage offer at either W or W-10%. The second dimension changes the observability of the wage offer. In the core design,

[93]Following the completion of work, wage payments are made to workers by members of the research staff. The employer pays his contribution directly to the research field team and never makes any payments directly to the workers. For ethical reasons, the employer in each village is aware of the possibility that we offer a wage below the prevailing wage.

there are two observability conditions:

i.) Fully Private: Employment offers are made in private, inside the worker's home. The employer does not enter the worker's home, and the research team never informs the employer of the wage.

ii.) Fully Public: Employment offers are made in public, on the street in front of the worker's home. The employer and any other passers-by can hear the terms of the wage offer.[94]

Treatment cells A and B give rise to basic tests of H1. First, the take-up rate in cell B measures whether there is any willingness to work below the prevailing wage when the offer is made in private. Under H1, take-up in cell B should be strictly positive. Second, a comparison of job take-up rates in treatment cells A versus B measures the fall in willingness to work attributable to a 10% lower wage, when job offers are made in private. Note that if workers do not believe that the information will be kept completely private, then this will result in a downwardly-biased estimate of an individual's true willingness to work below the prevailing wage, making it harder to validate H1. We return to this in our discussion of threats to validity in Section 2.5.

The difference in take-up rates between treatment cells B and D offers a basic test of hypothesis H2. This comparison identifies how much an individual's willingness to work below the prevailing wage falls when the take-up decision is made observable to the community.

By examining the impact of observability on take-up at the prevailing wage—i.e. the differences between cells A and C—we can validate whether observability itself has any impact on job take-up even when no community norm is being violated. We predict observability will have no impact in this case (i.e. A and C will have the same take-up rate). This helps us rule out a story where changes

[94]Villages in the study districts are typically quite compact, with small dwellings that share adjacent walls and no real doors. When an employer visits the home of a worker, it is not uncommon for curious neighbors to overhear the wage offer. It is important to note that in the study villages, laborers and large employers live in distinct neighborhoods. Thus in most cases, all passers-by who overhear the wage offers will be individuals whose primary occupation is also wage labor and not landowners (i.e. employers).

between B and D are not due to community pressure around wage cuts, but some other level shifter. Similarly, we can perform our test of H2 in a differences-in-differences framework by examining (D-B) - (C-A) — thereby partialling out any level shifters in from observability itself.

Note that in moving from fully private to fully public, there is a change both in whether community members at-large learn about the wage rate and also whether the employer himself learns about the wage rate. Consequently, one interpretation of any differential take-up between cells B and D could stem from a desire to avoid having the employer learn the worker's reservation wage, which may affect future bargaining dynamics with employers. This story would not necessarily rely on co-worker sanctions. To help distinguish between these two interpretations, we introduce a third source of variation in the observability of the wage offer.

iii.) Partially Private: Employment offers are made in private, inside the worker's home. However, the employer does enter the worker's home and overhears the wage offer.

The difference (B-F) captures the aversion to taking a wage below the prevailing rate in front of an employer, while the difference (F-D) captures the aversion to accepting a wage below the prevailing rate in front of other laborers. Of course, if an employer is aware of the wage, then information transmission through the village may lead workers to learn it as well. Consequently, to the extent that we observe a take-up difference between B and F, we cannot disentangle whether it results from employer knowledge only, or the indirect channel of employer knowledge spreading to workers. We acknowledge that, to the extent that the purpose of collusive behavior is to enforce a collective bargaining outcome with employers, the difference between D and F is not necessarily well defined. Regardless, we view this as a useful additional source of variation.

2.2.4 Context and Protocols

Context This experiment takes place in 183 villages in four districts of Odisha, India. Agricultural production in these districts focuses mainly on paddy, which is both seasonal and labor-intensive. Over 70% of survey respondents are primarily engaged in agriculture, with 53% listing daily-wage agricultural labor as their main occupation. 91% of all respondents engage in daily-wage agricultural labor. There is strong baseline evidence of wage rigidity and wage compression in this area. In the survey conducted by Kaur (2018) in this area of Odisha, 100% of laborers and employers reported that they could not recall a year when the prevailing nominal wage in the village was lower than the wage in the previous year. This is consistent with the distribution of wage changes across India as a whole (Figure 2.3).[95] In addition, the baseline survey evidence collected by Breza et al. (2016) indicates that there tends to be very little variation in wages within a village (Figure 2.4). Over 80% of agricultural workers in a village receive the modal village wage. This is consistent with the presence of a clear wage norm that can be easily followed by laborers when making labor supply decisions.

In our setting, the village constitutes a prominent boundary for the labor market. Agricultural employers hire daily-wage laborers solely from within or close to their village. For example, in our pilot surveys, laborers report that 70% of worker-days in agriculture involve work within the village, and 97% of agricultural work-days take place within 5 kilometers of the village. In addition, workers within a village (i.e. those whose primary source of earnings is wage labor) tend to live the same tightly packed area of the village (referred to as the "labor colony"). This is common in India as laborers within a village generally stem from low-caste groups, and live in designated areas. Such an environment may be expected to make worker collusion and sanctions easier. Indeed, prior work suggests that the presence of strong within-village ties, risk-sharing, and job search networks allow sanctions to have a significant clout against those who violate a village norm.[96]

[95]The figure shows some areas where nominal wage cuts may occur. This could be driven by measurement error and compositional changes (see Kaur (2018)). More generally, the occurrence of wage cuts is not inconsistent with the presence of wage rigidity.

[96]See, for example, Townsend (1994), Chandrasekhar, Kinnan, and Larreguy (2014), Karlan, Mobius, and Szeidl

We take advantage of a few distinct features of production in our study area. First, as is typical in subsistence farming, paddy production has lean periods in which employment declines (particularly between February and June, and again between September and November), which allows us to document that informal unions hold even during the eight months of the year when the opportunity cost of turning down a job is potentially high.[97] Second, the labor-intensive nature of paddy production in Odisha results in the ubiquitous use of casual daily-wage labor. Third, because of the uniformity of the crop produced in the region, we can work with local employers to offer nearly-identical agricultural jobs in different villages. We can work with everyone who participates in the agricultural daily-wage labor market, without selecting for people who have special skills or knowledge in a particular type of production or crop, which is helpful for the external validity of our results.

These features of village agricultural labor markets in Odisha (and, indeed, in many places elsewhere in India and the rest of the developing world)—siloed labor markets, a clear and consistent prevailing wage, and a relatively homogeneous skill and knowledge base among those who participate in the agricultural labor market—make this an effective context in which to study the micro-foundations of the wage rigidity and wage compression we observe.

Protocols

Village Selection We sample 183 villages in rural areas (i.e. at least 20 km from a town) across four districts in Odisha, India. We limited the sample to villages that have forty or more households

(2014), and Dhillon, Iverson, and Torsvik (2013).

[97]We pause the experiment during the four months of peak demand in the agricultural labor market. The labor market is most likely to clear at this time, and workers' alternative to taking up a job with us is taking a job with another employer at the prevailing wage with very high certainty. At these times of the year, a worker's reservation wage of employment is likely to be the prevailing wage, leading us to expect zero takeup at below the prevailing wage in both public and private. While conducting the experiment during these times would be fascinating, for cost and sample size reasons, we stick to the remaining eight months of the year for our study.

in the labor colony.[98]

Employer Selection Once a village has been selected, we conduct a preliminary visit in which we ask an informant to list 20 employers in the village and tell us how much land they own and cultivate. After using this information to understand the distribution of land size in the village, we then recruit a mid-sized employer willing to hire up to three workers to work on his land within the next week.

Employers are told that the wage rate offered for a job on their land may be above or below the prevailing wage. They are not given any information about the level of observability of the job offers. Their contribution to the wage that will be paid to each worker is always Rs. 100 (approximately US$.1.6) per worker hired. This ensures that their incentives to hire are not changed differentially across treatments, and enables us to keep wage offers blind to the employer in the fully private treatments. Our field staff accompanies the employer during hiring, as discussed below. After that point, we are not involved in the employment arrangement and the employer supervises the worker on his land as usual etc.

Treatment Assignment Once an employer has been selected in the village, we assign the village to one of six treatment cells, in accordance with the sampling weights assigned to each treatment.[99] Our unit of randomization is the village. Thus within a village, all workers are recruited under the same wage and observability condition.

[98]We use a floor on the size of the village to ensure heterogeneity across villages in the level of information spread, particularly in private treatments. In smaller villages, information may consistently be transmitted to all households.

[99]Specifically, we generated a random treatment assignment order in accordance with our desired sampling weights. Villages in the sample were then sequentially assigned to the next treatment assignment on this list as we moved through the study areas. In the second half of the sample, we also stratified by whether the labor colony population size in the village is above or below median for the block (a geographical subunit of a district). We did not perform this stratification for the first half of the sample due to an oversight.

Participant Selection and Hiring We select one informant from the labor colony of the village to create a comprehensive census of daily-wage agricultural workers. We then partner with the employer to offer jobs to two-three workers (depending on the task, which was specified before treatment assignment and based on the area). We approached a random subset of workers in the labor colony with job offers, moving on to the next randomly selection household in case the worker was not home.

In accordance with local practice, job offers are made two days in advance, and at dusk, when the majority of workers are home. In our survey of employers in the area, employers in 60% of villages typically hire two days in advance of when they would like to complete the work, and less than two days in advance in 85% of villages. The norm in every village is to hire less than four days before the day of work.

In all hiring, the employer informed the worker he wanted to hire laborers to work on his land. The employer then introduces one of our field staff, saying "this person is here with me, and will give you some more details". Across all treatments, the field staff person accompanying the employer then relays the wage level to the worker. This enables us to keep which information is being conveyed constant across all the observability treatments.

Hiring - Public observability Job offers are made outside the participant's home, which generally lead to others in the labor colony observing the job offer and take-up decision. Field staff do not interact with or provide the onlookers with any information directly. As we document below, on average there were 5 onlookers present when public hirings happened; these would typically have been other residents of the labor colony, i.e. laborers.

Hiring - Private observability Job offers are made in the participant's home. After his initial conversation with the worker, the employer wanders away with a staff member out of earshot, while a second staff member continues the conversation with the worker and informs him of the exact wage level.

Hiring - Partially private observability Job offers are made in the participant's home, but the employer remains present for the entire conversation.

Confirmation and Day of Work On the day before work is scheduled, the employer and field staff confirm the work with those who accepted the job. This is the common practice by employers in our study area. On the day of work, the employer meets the workers in his fields. The work itself proceeds as it would normally: the employer supervises the work, provides in-kind benefits like tea and lunch, without our staff present. Members of the research team do verify that the workers who agreed to the job actually work a full day. They also deliver the physical wage payments to workers at the end of the day across all treatments. This enables us to hold total wages confidential from the employer.

Surveys After the workday is complete, we conduct a variety of surveys to provide further support for our hypotheses, including surveys with the employer, approached workers, and randomly chosen workers from the labor colony who were not approached for job offers.

2.3 Results

2.3.1 Take-up of the Job

Panel A of Figure 2.6 presents the raw job take-up rates in the public and private treatments across wage offers made at the prevailing wage and at a 10% discount to the prevailing wage. In all the analysis, the outcome variable for job take-up equals 1 if the worker accepted the job: i.e. showed up to and completed the work. The figure shows that when the wage is set below the prevailing rate, take-up falls substantially when the job offers are made in public instead of private. However, when the wage is set at the prevailing rate, if anything, public offers lead to a weak increase in take-up rates. Table 2.2 presents the results in regression form across treatment cells. Cols. (1)-(2) reports OLS regression results for the full sample, where Fully private: Prevailing wage - 10% is the omitted category.

For job offers at a 10% wage cut, take-up falls by 13.6 percentage points when an offer is public versus fully private (Col. 2). The results are similar with and without controls, with p-values of this difference ranging from 0.019-0.032 in Cols. (1)-(2) (p-values reported at bottom of table).

In contrast, for jobs made at the prevailing wage, take-up rates are positive in sign and statistically indistinguishable under public and private. This is consistent with no role for observability when the social norm is not being violated (i.e. under the prevailing wage), but an important role for social observation when workers are contemplating whether they will take up jobs below the prevailing wage.

In addition, our design enables us to gain some suggestive evidence on whether the difference between Fully Private versus Public is driven by the presence of the employer rather than presence of other workers. Our results suggest that in the presence of the employer (under Partially Private (Employer)), take-up of low wage offers declines—the difference with Fully Private is negative but insignificant—but remains substantively higher than take-up under Public offers. In Col. (2), we can reject that (Partially Private W-10% = Public W-10%) at the 10% level (p-values reported at the bottom of the table). This test has limited statistical power due to the smaller sample size in the Partially Private treatment cells. As discussed above, it is also not perfectly interpretable as the incremental impact of co-worker observability, since the employer could also indirectly spread information to others in the village. However, these results support our presumption that pressure from other workers plays an important role in depressing labor supply below the prevailing wage.

Finally, note that the absolute magnitude of difference between Fully Private at the prevailing wage and Fully private at a 10% cut is 6-8 percentage points, but statistically not different than zero in any column. This is consistent with an underlying labor supply curve that is likely upward sloping, but not highly elastic around the prevailing wage. However, in the presence of social pressure (i.e. under the Public treatments), observed labor supply drops substantially below the prevailing wage—behavior that can reinforce a wage floor at the prevailing wage.

As explored in detail below, we hypothesize that the collective behavior of the informal union is stronger for "insiders" in the village union. We define an "insider" to be a worker who engages regularly in the agricultural wage market and show our main results for only this group of individuals. Panel B of Figure 2.6 shows the raw take-up rates for the subset of workers who self-identify in the endline survey as working in wage agricultural labor as a primary or secondary occupation. Interestingly, take-up rates in private look very similar to those in the full sample at both wage rates. However, we observe that when the below-prevailing wage offer is made public, take-up falls almost to zero. The Col (3) of Table 2.2 shows the results of the OLS regression specification, restricting the sample to this same group of "insiders." Indeed, the key treatment effect of public wage offers holding fixed the below-prevailing wage rate increases in magnitude to 24.6% (significant at the 1% level). Moreover, for this group of insiders, the treatment effect is much larger when the low wage offer is made in front of other workers in comparison to when the offer is made only in front of the employer (25.6% versus 7.58%). This difference is also significant at the 1% level. These results provide preliminary evidence that workers who are members of the informal union are especially concerned with violating the village wage norms in front of other workers.

A potential concern with our design is that when public offers begin in the village, information may spread about the wage rate at which offers are being made. This, in turn, could affect which workers are available to be approached by the employer for a job offer (for the subsequent second and third offers made in the village). Note that information spread affecting the take-up decision is not a problem by itself. However, it is important that the treatments do not induce a type of differential selection into receiving a job offer. To ensure that this is not a problem, in columns 1-4 of Appendix Table B.1, we restrict our analysis to the first household and first two households, respectively, approached in our randomized list. While results are noisier due to the reduced sample size, the results are qualitatively similar.

Recall that during hiring, if a household that was approached for a job was not home, the employer moved to another household—approaching no more than six households total (regardless of how many people were home). The process of approaching three households in each village was sufficiently quick that we do not think any aspect of the wage level would have led to differential door locks. However, as a robustness check against this concern, in columns 5-6 of Appendix Table B.1, we code any household that was not home as zero take-up. While this mechanically dampens the observed take-up levels across all treatments, and consequently predictably decreases statistical power, the results remain qualitatively similar. For job offers below the prevailing rate, the take-up difference between public and fully private offers is statistically significant at the 10% level across most specifications. In addition, for jobs at the prevailing wage, take-up levels are similar across the observability levels.

2.3.2 Heterogeneous Treatment Effects

We next turn to two tests for heterogeneous treatment effects. First, we ask whether the treatment effects are any different statistically between insiders in the informal union and outsiders. We hypothesize that insiders, who rely on casual agricultural labor markets as an important income source, have more to lose from violating the village wage norm in front of other workers. We look for support for this hypothesis in Table 2.3. Here we define outsider in two different ways using each worker's endline survey responses: first, as above, an outsider is any individual who does not participate in the agricultural labor market as a primary or secondary occupation; second, we define an outsider as an individual who participates in the non-agricultural labor market as a primary or secondary occupation. Because we undersampled the partially private treatment cells, we pool Fully Private and Partially Private together for each wage. Cols. (1)-(2) present the main treatment effects, but using the pooled specification. Col. (3) presents heterogeneous treatment effect estimates using the first outsider definition, and Col. (4) does the same for the second definition. We find that insiders respond to a public low-wage offer with an 18-22 percentage point decrease in job take-up rates (both significant at the 1% level). Moreover, this treatment effect is detectably smaller for outsiders, supporting our insider vs. outsider hypothesis.

Our view of the mechanism underlying our main results is that workers reduce their take-up of below-prevailing wage jobs when they worry that their decisions are observable to others in the village. Therefore, in villages where more individuals are likely to learn of worker take-up decisions, we hypothesize that the main treatment effects will be larger. Following the completion of hiring in all villages, we returned to all but one of them to capture village-level characteristics that were not recorded in our initial endline. In this survey exercise we asked approximately five randomly-chosen workers per village a series of questions, some of which pertained to information flow in the village. Two of the questions can help us to explore our hypothesis: first, we asked each individual the extent to which laborers learned about the wages at which others accept agricultural work; second, we asked how many others would find out if a worker accepted an agricultural job at below the prevailing wage. In each case, we aggregate responses at the village level and create an indicator for whether a village has below-median information flow. We predict that the magnitude of the treatment effect will be relatively smaller for these low-diffusiveness villages.

In Table 2.4 we explore heterogeneous treatment effects based on both measures of low information flow. Again, we augment the pooled average treatment effects specification. In Col. (1) we consider knowledge about others' wages, while in Col. (3), we use the measure pertaining to information spread about a wage norm violation. Both measures of low information flow deliver similar qualitative results. In both cases, publicizing a low wage offer in highly diffusive villages leads to an approximately 20 percentage point decline in take-up rates (significant at the 1% level). However, in low diffusiveness villages, this large treatment effect is almost completely offset, leading to no measurable differences in take-up rates between public and private low-wage offers. These findings are consistent with our proposed mechanism.

2.4 Costly Punishment - Lab Games Results

To provide positive evidence on the willingness sanction those who violate labor supply norms, we use a costly punishment game in a supplementary lab-in-the-field exercise.

In another set of 11 villages—drawn from the same population as our study villages—we again partner with employers to make job offers to a random subset of workers at varying wage rates within each village. These offers are always made in private. Each worker is first offered a job at 10% below the prevailing wage, and if he says no, is asked if he would be willing to work for the employer at the prevailing wage. By approaching 10-15 workers with job offers in each village (with the number of workers per village decided ex ante), we guarantee that in each village, at least some workers have accepted a wage cut.[100]

Specifically, we then recruit another (random) subset of laborers in each village who were *not* offered jobs. These other laborers, who we will refer to as "players", are the ones who participate in the costly punishment game. Each player is paired with an anonymous worker (the “partner”) who received a job offer. The player and his anonymous partner are both given an endowment of Rs. 100. The player can "punish" his partner, reducing his endowment, by giving up some of his own endowment. Specifically, for every Rs. 5 that is removed from the partner’s endowment, the

[100]Note that the costly punishment game is played in the evening after job offers are made, but before the day of employment occurs. After the game is played, we announce that those laborers who do get jobs will receive the full prevailing wage (regardless of their initial response at the time of the wage offer). This enables us to fully preserve the anonymity of workers’ take up decisions and prevent any sanctions outside the game.

player must give up Rs. 1 of his own endowment. To make visualization easy, we implement this by placing Rs. 100 in each of 2 trays, placed in front of the player. The player than removes money from his tray and his partner's tray, in accordance with the above proportion, until he is satisfied with the final allocations.

To test for costly punishment, we randomize two features of the partner's characteristics. First, we randomly vary whether the player is partnered with a worker in the player's own village, or is partnered with a worker in a village that is geographically far away. Note that in this latter case, the worker's job acceptance decision has no direct consequences for the player, since the partner's actions take place in a different labor market. Second, the player is told that his paired worker accepted a job at either (a) the prevailing wage or (b) 10% below the prevailing wage. The sample is weighted so that there is an equal number of observations in each of the 2x2=4 cells.

Furthermore, in order to obfuscate the reason for the exercise, we add in three "placebo" rounds of the game, which are played by the player before he receives one of the above conditions.[101] The player's payoff is determined by a random roll of the dice, in which one of his four rounds is implemented.

If accepting a wage cut violates the social norm, then the literature on social preferences indicates that individuals will be willing to destroy their own surplus to punish those who have engaged in norm violations. In contrast, we do not expect to see punishment among workers who accept work at the prevailing wage—providing a helpful benchmark.

Figure 2.7 shows the raw punishment frequencies in our different experimental treatments. As expected, the figure shows that there is virtually no punishment of workers who accept jobs at the prevailing wage.[102] In contrast, when paired with a worker who accepted a wage cut from their

[101]In each round, the player's paired partner is a different individual. In each of these earlier rounds, the paired worker undertakes a positive, negative, and neutral action, respectively: baking someone a cake, stealing someone's bike, and traveling to the city for work.

[102]Note that we use a comprehension check after training to verify that workers understand the rules of the game.

own labor market, players punish the worker 45.1% of the time.

In addition, we find that the desire to punish norm violations is not limited to actions in one's own village. Players also punish workers from distant villages in similar frequencies (45.0% of the time) who have accepted a wage cut—even though that worker's action has no scope to affect the player's own labor market.

Table 2.5 presents these results in regression form. Column 1 shows that, pooling across "partners" in the own and other labor market, the punishment probability increases by 44.1 percentage points when the "partner" accepts a wage lower than the prevailing level (statistically significant at the 1% level). Column 2 shows that this effect size is of very similar magnitude and is statistically indistinguishable when the "partner" lives in a different labor market versus the player's own labor market. Columns 3-4 show that these results are robust to village fixed effects and to considering only the first experimental round pertaining to the "partner's" labor supply decisions. Finally, Column 5 shows that "partners" who accept a job below the prevailing wage from the same labor market receive payoffs that are Rs. 13.3 smaller (on a base of Rs. 100).

When players do punish, the amount of money they deduct corresponds to 42.8% of average daily labor market earnings in our sample. In order to impose this punishment on their partner, the amount that players forego from their own endowment, conditional on punishment, corresponds to 8.6% of typical daily earnings.

These results are consistent with the literature on social preferences, which indicates that individuals will be willing to destroy their own surplus to punish those who have engaged in norm violations (Charness and Rabin, 2002). Our findings are also consistent with contagious punishment models (Ellison, 1994), in which norms are an equilibrium strategy that is enforced through decentralized sanctions. We should note, however, that the willingness to punish those in other labor markets—

The results presented here exclude workers who failed the objective comprehension check (i.e. could not answer a series of 5 questions about the rules of the game). This reduces measurement error in the results. Results are similar, but noisier, if we include workers who failed the comprehension check.

where the deviating party's actions have no scope for equilibrium effects on one's own payoffs—is particularly consistent with villagers viewing norm violations in moral or general terms.

2.5 Threats to Validity

2.5.1 Internal Validity

We discuss some potential confounds that could contaminate the interpretation of the results of the main take-up field experiment.

Information Spread One might worry that the private wage offers do not remain private given that we make job offers to multiple workers. This is a valid concern, a priori, for several reasons. First, if the number of employment offers is high relative to village size, then even in the private offer condition, the wage offers will essentially become public. This would likely bias take-up at W-10% in private toward zero. As mentioned above, we limit the number of job offers to a small number in each village. The fact that we observe robust take-up in the private wage cut treatment (as opposed to close to zero take-up in the public wage cut treatment) validates our premise that at least a portion of workers believed that confidentiality would be maintained in the private treatment. To the extent that workers did not believe their take-up decision would remain confidential, this suggests our take-up estimates are a lower bound.

Number of Onlookers Our design rests on the idea that the presence of onlookers during public job offers will affect take-up behavior, because it directly enables observability by other laborers in the labor colony. In addition, in Appendix Table B.3, we validate that the number of onlookers was similar under the Public treatments under the two different wage rates (prevailing wage and 10% cut).

Information About the Prevailing Wage One potential concern with our design is that the public treatments provide workers with information about the prevailing wage—e.g., through potential comments from onlookers. This information, in turn, could depress take-up of public jobs below the prevailing wage. This is not consistent with this setting: the prevailing wage is general knowledge, as validated in our endline survey (Figure 2.5). As further evidence in support of this idea, in Appendix Table B.4, we document that among workers who were approached for job offers, reports of the prevailing wage are not systematically different across treatment cells. Importantly, there is no evidence that knowledge of the prevailing wage is different among Public and Fully Private treatments.

Poverty Signaling It is possible that in some villages, only the poorest households might absolutely need to take a job below the prevailing wage. Thus, other households that might prefer to take such a job in private might worry that doing so in public might send a signal about their wealth to the community. If individuals experience disutility from being classified as very poor, such a mechanism could explain a fall in take-up at W-10% when the wage offer is public. The difference between cells A versus C, and G versus I, respectively, provides a possible suggestive test against such an explanation. If projecting status and wealth is desirable, workers should be marginally more likely to reject job offers in public versus private in these conditions as well, providing a helpful, albeit imperfect, placebo test. In addition, in endline surveys, the majority of workers state that accepting job offers below the prevailing wage would result in anger and sanctions from others—consistent with our hypothesized mechanism. Finally, our costly punishment game results provide positive support for sanctions. If accepting a wage cut is only costly because it is a sign of financial destitution, then it is unclear why workers would punish such individuals by taking money away from them.

Side Payments In our endline surveys we checked whether employers tried to compensate workers for the low offer wage by making side transfers, and do not find evidence for this. Furthermore, if such behavior were to exist, then it would most likely cause an increase in take-up across all W-10% treatment cells. Thus side payments can't rationalize our hypotheses.

2.5.2 External Validity

The magnitudes of our estimates are, of course, specific to labor markets in the study districts in Orissa, India, during the agricultural lean season. Our primary goal is to provide evidence that in our setting, villagers belong to informal unions and use social sanctions to enforce adherence to a village wage. While the shape and form of informal unions may vary across settings, we have reason to believe that the phenomenon of interest is not limited to rural Orissa. Several papers provide descriptive evidence consistent with worker co-ordination in setting and enforcing wages across rural labor markets in South Asia (Kaur, 2018; Osmani, 1990; Dreze et al., 1986). There is also reason to believe that similar informal institutions may exist in other settings. For example, Prothero (1912) points to a similar phenomenon in the early stages of industrialization in England. Further, the implications of informal unions – such as wage rigidity and low employment rates – are phenomena that are observed in many developing country contexts even outside of South Asia (Beegle et al., 2015).

It is also important to reiterate that in our study, we are interested in understanding the labor supply consequences of informal unions. Given that we subsidize employment, we cannot make any claims about the demand side of the market. We leave this to future work.

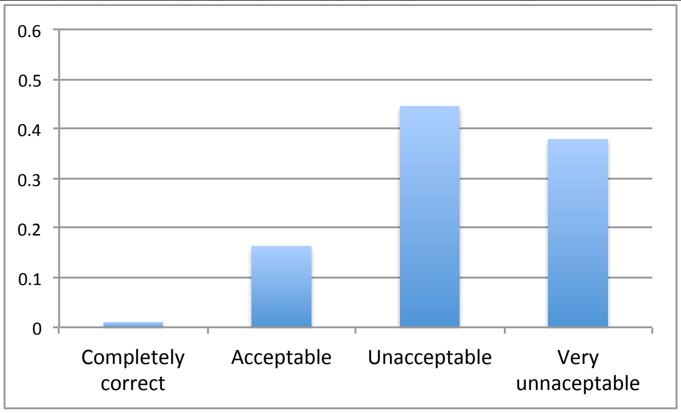
2.6 Conclusion

We find evidence that workers privately would like to supply labor below the prevailing wage, but do not do so when their take-up decision is publicly observable. This supports the hypothesis that collective pressure dampens labor supply below the prevailing wage, supporting the presence of wage floors in village labor markets. Our findings provide documentation of a way in which norms against accepting wage cuts distort labor supply behavior, with large impacts on the foregone earnings of unemployed workers.

Finding evidence that co-worker pressure dampens labor supply below the prevailing wage—even during times of high unemployment—provides impetus for exploring this mechanism in other settings. If this mechanism is indeed more generally applicable, then this can inform our understanding of the role of norms in shaping labor market outcomes, such as wage rigidity and wage compression.

2.7 Figures

[Acceptability of Taking a Wage Cut. Suppose it is the lean season. The prevailing wage is Rs. 200. To increase his chance of finding work, a laborer tells farmers that he would be willing to work any day that week at Rs. 180. Is the laborer's behavior acceptable?]



[Sanctions for Accepting Wage Cuts. If a laborer accepts work at a rate lower than the prevailing wage, how likely is it that the other laborers in the village become angry?]

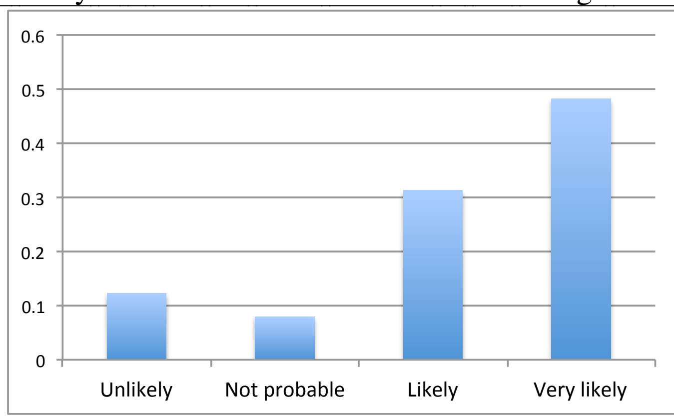


Figure 2.1: Survey Evidence

	Wage	
Observability of Wage	W	W-10%
Fully Private	A	B
Fully Public	C	D
Partially Private (Employer Observes Offer)	E	F

Figure 2.2: Experimental Design

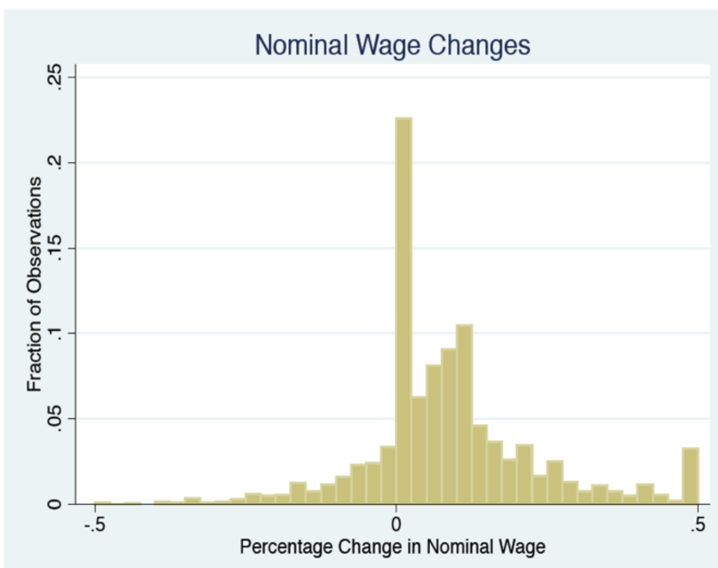


Figure 2.3: Nominal Wage Changes. Source: Kaur (2015), World Bank Climate & Agricultural Data (256 districts).

Figure 2.4: Distribution of Wages Inside the Village. Source: Breza, Kaur, and Shamdasani 2017.

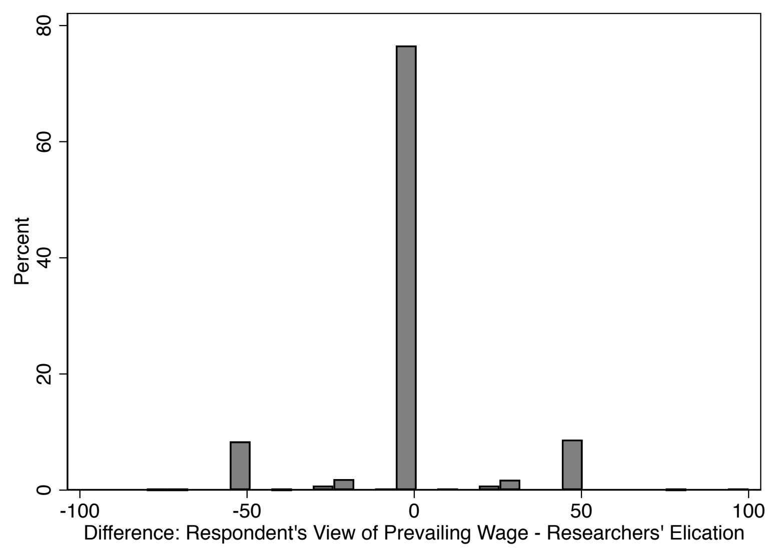
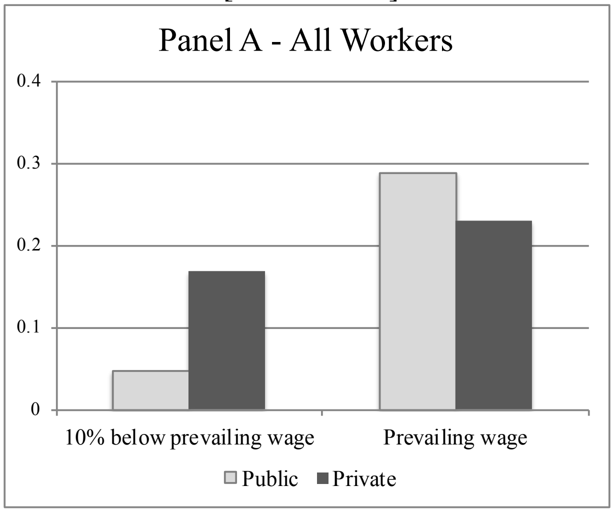


Figure 2.5: Control Group Reports of Prevailing Wage, Normalized by Informant Report.

[All Workers]



[Agricultural Workers Only (Insiders)]

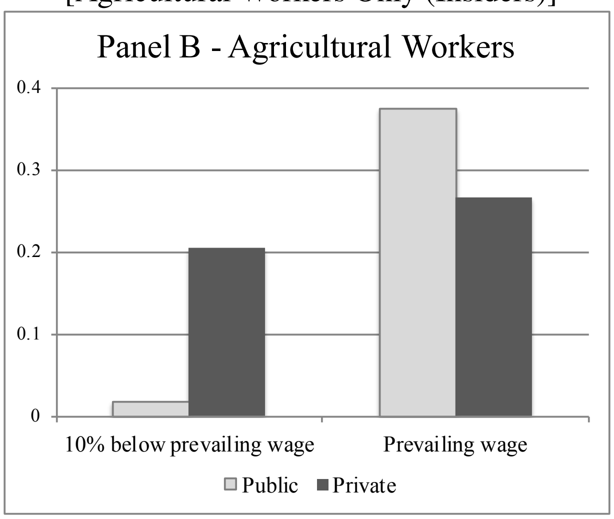


Figure 2.6: Job Take-Up by Treatment: Raw Data

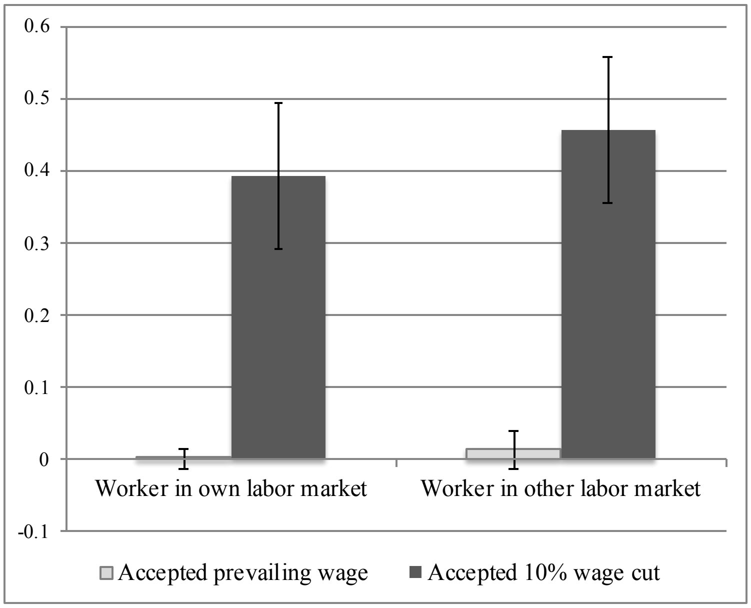


Figure 2.7: Laboratory Games: Fraction of Respondents Punishing

2.8 Tables

Table 2.1: Covariate Balance

VARIABLES	(1) Number HHs	(2) Field Prep	(3) Weeding	(4) Compost Crush.	(5) Age	(6) Sched Tribe	(7) Has Worked Empl.	(8) Empl. Influence	(9) Days Paid Wage 30	(10) Not Ag. Laborer
Public: Prevailing Wage - 10%	-2.227	0.0708	-0.0736	-0.0144	0.561	-0.0264	0.0897	-0.0419	0.984	0.0370
	(3.642)	(0.108)	(0.0798)	(0.0567)	(1.826)	(0.109)	(0.0887)	(0.183)	(1.063)	(0.0634)
Partially Private (Employer): Prevailing Wage - 10%	-1.826	0.0724	-0.108	-0.0628	1.940	-0.00439	0.191	-0.143	-0.912	0.0810
	(3.722)	(0.116)	(0.0745)	(0.0476)	(1.820)	(0.118)	(0.0911)	(0.226)	(0.995)	(0.0700)
Fully Private: Prevailing Wage	5.316	0.0653	-0.00345	-0.0601	1.285	0.0103	0.0511	0.315	1.379	-0.00601
	(4.146)	(0.124)	(0.0966)	(0.0486)	(2.224)	(0.126)	(0.0901)	(0.193)	(1.179)	(0.0604)
Public: Prevailing Wage	-2.477	0.0206	0.0425	-0.0443	-1.492	0.194	0.197	-0.0324	1.336	-0.0251
	(4.091)	(0.112)	(0.117)	(0.0565)	(2.116)	(0.132)	(0.105)	(0.191)	(1.058)	(0.0650)
Partially Private (Employer): Prevailing Wage	4.454	0.101	-0.0831	0.0170	3.619	-0.0167	-0.0205	-0.000504	1.149	0.0503
	(4.979)	(0.149)	(0.0981)	(0.0915)	(2.374)	(0.149)	(0.103)	(0.186)	(1.012)	(0.0856)
Observations	502	502	502	502	442	444	426	383	427	446
Task and Year x Month FE										
Sample	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main
Depvar Mean (Private: Prevailing Wage - 10%)	46.58	0.214	0.165	0.126	44.34	0.333	0.318	3.333	8.841	0.156
Test: Full Private W-10% = Public W-10%	0.542	0.513	0.358	0.799	0.759	0.809	0.313	0.819	0.356	0.560
Test: Full Private W = Public W	0.104	0.731	0.709	0.729	0.215	0.197	0.194	0.0817	0.974	0.768
Test: Full Private - Public, W-10% = Full Private - Public, W	0.355	0.495	0.416	0.678	0.249	0.242	0.693	0.260	0.538	0.536
Test: Partial Private W-10% = Public W-10%	0.920	0.989	0.579	0.283	0.355	0.845	0.305	0.652	0.0911	0.544
Test: Full Private W-10% = Full Private W	0.201	0.599	0.972	0.218	0.564	0.935	0.572	0.106	0.244	0.921

Notes: Column 1 reports the likelihood of successfully completing an endline survey with a salient respondent, by treatment. The outcome variable in column 2 is the number of control endline surveys conducted in the salient household's village. Column 3 reports the presence of a supplemental survey for the salient household's village. Note that we successfully completed supplemental surveyors in all but one village. The treatment of this failed village was Fully Private: Prevailing Wage. Thus, we restrict the sample to only the fully private treatment arms. In all columns, the omitted category is the Fully Private: Prevailing Wage - 10% treatment. Standard errors are clustered at the village level and are reported in parentheses. Observations are weighted by the number of salient individuals in each village.

Back to text

Table 2.2: Main Results

VARIABLES	(1) Accepted Offer	(2) Accepted Offer	(3) Accepted Offer
Public: Prevailing Wage - 10%	-0.122	-0.136	-0.246
	(0.0564)	(0.0573)	(0.0644)
Partially Private (Employer): Prevailing Wage - 10%	-0.0657	-0.0516	-0.0758
	(0.0611)	(0.0633)	(0.0788)
Fully Private: Prevailing Wage	0.0609	0.0791	0.0663
	(0.0703)	(0.0659)	(0.0819)
Public: Prevailing Wage	0.119	0.116	0.104
	(0.0808)	(0.0713)	(0.0856)
Partially Private (Employer): Prevailing Wage	0.0364	0.0690	0.0935
	(0.0775)	(0.0886)	(0.0992)
Observations	502	502	363
Task and Year x Month FE		✓	✓
Sample	Main	Main	Agricultural HHs
Depvar Mean (Private: Prevailing Wage - 10%)	0.175	0.175	0.211
Test: Full Private W-10% = Public W-10%	0.0316	0.0188	0.000181
Test: Full Private W = Public W	0.460	0.589	0.658
Test: Full Private - Public, W-10% = Full Private - Public, W	0.0629	0.0481	0.00858
Test: Partial Private W-10% = Public W-10%	0.143	0.0865	0.0107
Test: Full Private W-10% = Full Private W	0.387	0.232	0.419

Notes: In all specifications, the dependent variable is an indicator for whether the worker signed up for the job and showed up for work. In all columns, the omitted category is the Fully Private: Prevailing Wage - 10% treatment. Columns 1 and 2 include the full sample. Column 3 restricts the sample to workers who answered the endline questionnaire and who indicated that they engage in agricultural labor as a primary or secondary occupation. Observations are weighted by the number of salient individuals in each village. Standard errors are clustered at the village level and are reported in parentheses.

Back to text

Table 2.3: Heterogeneous Treatment Effects: Insiders vs. Outsiders

VARIABLES	(1) Accepted Offer	(2) Accepted Offer	(3) Accepted Offer	(4) Accepted Offer
Public: Prevailing Wage - 10%	-0.113	-0.142	-0.182	-0.216
	(0.0439)	(0.0487)	(0.0447)	(0.0539)
Fully or Partially Private: Prevailing Wage	0.100	0.0984	0.0590	0.0190
	(0.0478)	(0.0516)	(0.0595)	(0.101)
Public: Prevailing Wage	0.141	0.148	0.139	0.0839
	(0.0599)	(0.0655)	(0.0803)	(0.0988)
Public: Prevailing Wage - 10% x Outsider			0.297	0.167
			(0.103)	(0.0689)
Fully or Partially Private: Prevailing Wage x Outsider			-0.0629	0.0626
			(0.108)	(0.124)
Public: Prevailing Wage x Outsider			-0.105	0.0793
			(0.131)	(0.133)
Outsider			-0.0793	-0.0721
			(0.0503)	(0.0532)
Observations	502	446	446	446
Task and Year x Month FE	✓	✓	✓	✓
Sample	Main	Individual Endline	Individual Endline	Individual Endline
Depvar Mean (Omitted)	0.147	0.160	0.211	0.245
Outsider Definition			Not ag. laborer	Non-ag. laborer

Notes: This table presents heterogeneous treatment effects by insider versus outsider status using the responses of salient households in the worker endline. In all columns, we pool Fully Private: Prevailing Wage - 10% and Partially Private (Employer): Prevailing Wage - 10%. In columns 1 and 2, we present the pooled version of the main results for the full sample and endline survey sample, respectively. In column 3, outsider is defined as an individual who does not claim agricultural labor as a primary or secondary occupation. In column 4, outsider is defined as an individual who works in non-agricultural labor as a primary or secondary occupation. In all specifications, the dependent variable is an indicator for whether the worker signed up for the job and showed up for work. In columns 1 and 2, the omitted category is the pooled Fully or Partially Private: Prevailing Wage - 10% treatment. In columns 3 and 4, the omitted category is the Fully or Partially Private: Prevailing Wage - 10% treatment for insiders only. Observations are weighted by the number of salient individuals in each village. Standard errors are clustered at the village level and are reported in parentheses.
Back to text

Table 2.4: Heterogeneous Treatment Effects: Village Information Spread

VARIABLES	(1) Accepted Offer
Public: Prevailing Wage - 10%	-0.199
	(0.0677)
Fully or Partially Private: Prevailing Wage	0.0565
	(0.0750)
Public: Prevailing Wage	0.111
	(0.0972)
Public: Prevailing Wage - 10% x Low Wage Info Spread	0.170
	(0.0934)
Fully or Partially Private: Prevailing Wage x Low Wage Info Spread	0.0613
	(0.0996)
Public: Prevailing Wage x Low Wage Info Spread	0.0335
	(0.130)
Low Wage Info Spread	-0.0731
	(0.0671)
Observations	499
Task and Year x Month FE	✓
Depvar Mean (Omitted)	0.204

Notes: This table presents heterogeneous treatment effects by village-level diffusiveness, as measured in the mop-up survey. In column 1, the heterogeneous variable of interest is an indicator for below-median knowledge of the wages of others. In column 2, we use an indicator for below-median spread of information about other workers accepting a job below the prevailing wage. In all specifications, the dependent variable is an indicator for whether the worker signed up for the job and showed up for work. In this table, we pool Fully Private: Prevailing Wage - 10% and Partially Private (Employer): Prevailing Wage - 10%. In all columns, the omitted category is the pooled Fully or Partially Private: Prevailing Wage - 10% treatment. Observations are weighted by the number of salient individuals in each village. Standard errors are clustered at the village level and are reported in parentheses.

Back to text

Table 2.5: Costly Punishment Games: Pilot Results

VARIABLES	(1) Any Punishment	(2) Any Punishment	(3) Any Punishment	(4) Any Punishment	(5) Partner's Payoff
Partner Accepts a Job Below Prevailing Wage	0.420	0.393	0.393	0.436	-14.57
	(0.0447)	(0.0632)	(0.0647)	(0.103)	(4.425)
Partner Accepts a Job Below Prevailing Wage x Different Village		0.0494	0.0494	-0.00310	5.569
		(0.0894)	(0.0916)	(0.137)	(4.551)
Partner lives in Different Village		0.0143	0.0133	0.00737	-0.701
		(0.0143)	(0.0185)	(0.0294)	(1.259)
Observations	262	262	262	131	131
Village FE			✓	✓	✓
First Round Only				✓	✓
Depvar Mean: Partner Accepts Job at Prevailing Wage	0.00763	0.00763	0.00763	0	100

Notes: Each participant ("player") was anonymously paired with either another worker in his village or in a distant village, and given various scenarios about his paired worker. A player could take away money from his paired worker's endowment by giving up money from his own endowment. The table reports results under the 2 employment scenarios: (i) the worker accepted a job at the prevailing wage, or (ii) the worker accepted a job at a wage 10% below the prevailing wage. OLS regressions. The dependent variable in Cols. (1)-(4) is a dummy for whether the player punished the other worker at all; in Col. (5) it is the payoff of the anonymous partner (his initial endowment minus the amount deducted by the participant). Each player plays these two scenarios in random order; Cols. (4)-(5) report results only from the first of these two rounds. Standard errors clustered by player. N=131 participants (i.e. agricultural laborers) in 31 villages (villages are different from those in the main experimental sample).

Back to text

Chapter 3. Agricultural Risk and Motivation for Crime: Theft from Oil Pipelines in Nigeria

3.1 Introduction

Participation in crime in general, and in organized crime in particular, has been found to have various economic motives. In particular, in environments with incomplete markets for insurance against risk, agents may turn to criminal activity as a way to hedge against income shocks (Miguel, Satyanath, & Sergenti 2004, Hidalgo et al. 2010). However, past literature has relied on unanticipated economic shocks for identification, without being able to exogenously vary individuals' level of access to criminal employment[103]. The nature of an extremely profitable type of criminal activity in the Niger delta region - the theft of oil from transport pipelines - allows me to vary (exogenously) both the economic shock faced by households, and the extent of local labor demand for criminal activity.

Nigeria is Africa's largest (and the world's thirteenth-largest) producer of oil. Oil production is concentrated in the nine Southern Niger delta region states (Orogun, 2010). According to OPEC statistics, oil and gas production accounts for around a third of Nigeria's GDP. At the same time, the country loses at least 4% of daily production (and likely even more) to theft, largely spearheaded by militia groups. Militia groups in the Niger delta frequently recruit youth from communities around pipelines to provide localized knowledge regarding the precise location of pipelines and routes to access them without detection. Anecdotal evidence suggests that youth receive monetary compensation in return for their engagement in the theft of oil.

In this paper, I rely on households' differential probability of employment with militia groups using a proxy measure - the household's distance from the closest oil pipeline. In contrast with some of the past literature that addresses organized crime, this provides exogenous variation in access

[103]Some important exceptions are Di Tella & Schargrodsky 2004 and Dell 2015, who use innovative methods to predict the displacement of criminal activity in a network when particular edges of that network are shut down due to increased enforcement.

to criminal employment. I interact this with a measure of unanticipated shocks to agricultural productivity at the household level to investigate whether households use criminal activity as a form of insurance against agricultural risk.

There are four key results in this paper. First, I find that negative agricultural productivity shocks faced by a household lead to a significant increase in the number of incidents of theft from oil pipelines in the vicinity of the household- this corresponds to a doubling of the number of large-scale oil spill incidents due to theft when households report large unanticipated crop losses. Since this pattern holds true for both small- and large-scale theft, this would suggest the involvement of local militia groups, and the existence of opportunities for short-term engagement in oil theft from these groups.

Second, I verify this using data on the number of militia conflicts in the vicinity of households facing exogenous decreases in agricultural productivity. I find that the number of militia incidents increases to a greater extent relative to good agricultural years for households close to oil pipelines than for households located farther away. Militia groups are therefore more active in areas facing negative agricultural productivity shocks, and even more so in areas close to oil pipelines. However, this paper does not take a stand about whether this is a supply-driven or demand-driven effect in the labor market. It is possible that local labor supply to militia activity increases, or that militia groups respond by moving to areas that are facing poor agricultural outcomes, or both.

Third, I find that the households driving these results are the ones who are most vulnerable to agricultural shocks - households engaged in agriculture, households with unemployed youth or youth who do not report school attendance, and households without access to credit.

Fourth, I find that these households, relative to households farther from oil pipelines who do not have access to employment in criminal activity, show a greater degree of consumption-smoothing, particularly with regard to expenditure on non-food items. This confirms that households in areas in which oil pipeline theft is increasing in response to negative agricultural shocks show increases in income. This could be in the form of the direct sale of oil by the household, or it could be due to monetary transfers from militia groups to employees from local households.

The results have important implications for policies that target the placement of security resources and the dispensation of credit, if one the goal of these policies is to reduce economic losses due to theft of oil from pipelines. For example, the selection of households for government loans or subsidized agricultural insurance could include an assessment of the distance from oil pipelines or the amount of militia activity in the neighborhood. In comparison, longer-term labor market interventions such as vocational training may have less of an impact on households' short-term decision about engaging in criminal activity (Blattman & Ralston 2015).

Oil is an especially interesting resource to examine in this respect because a) tapping into this resource is dangerous (with a number of fires, explosions, and deaths resulting from sabotage) (Onuoha 2009), and b) exploiting this resource has demonstrably negative effects on the local environment, health, and agriculture (Salau 1993, Adeyemo 2002, Kadafa 2012). That is, the negative externalities from state-controlled production of oil, and from sabotaging this production, are particularly stark. Yet, as I will discuss in the next section, oil pipeline theft remains a common problem in Nigeria, and one that is crippling for the petroleum industry. The existence of such externalities suggests that understanding participation in the theft of oil from pipelines, and channeling of resources towards reducing its incidence could provide economic benefits beyond just the monetary value of the stolen oil.

This paper contributes to the existing literature on the relationship between economic shocks - either to the returns from criminal activity or to the returns from the next best alternative - and conflict, to the extent that conflict stems from the desire to appropriate another's resources. Existing literature examines the effect of shocks to commodity prices (Angrist & Kugler 2008, Dube & Vargas 2013, Blakeslee & Fishman 2014, Iyer & Topalova 2014, Dix-Carneiro et al. 2017) and productivity shocks (Collier & Hoeffler 1998, 2001, and 2002, Hidalgo et al. 2010, Miguel & Sathyanath 2011, Axbard 2016) on conflict, and find, largely, effects that operate in the anticipated directions. This paper confirms the opportunity cost channel for engaging in crime - the lower the returns from agriculture, the higher the labor supply to militia operations to steal oil (Becker 1968, Grossman 1991). However, existing work in this area uses either global shocks from world prices of various commodities, or localized rainfall shocks, which could affect (in an unobservable way) any variety of other factors that contribute to conflict. The unique structure of the oil pipeline network in Nigeria provides an additional layer of identification; some households, selected in an essentially random manner, have access to alternative employment with militia groups, while others do not. This paper is also able to provide suggestive evidence on precisely which types of households participate in this alternate employment and quantify the economic impact.

My reliance on exogenous variation in criminal access suggests that this paper is able to separate the rapacity effect of positive economic shocks on crime by focusing on theft from oil pipelines - that is, a theft from a third-party who is unaffected by local economic shocks. This is in contrast to previous literature, which has acknowledged the existence of both channels, and has focused on identifying the stronger channel (Freedman & Owens 2016) .

This work also fits in with a literature that studies recruitment into criminal groups and individuals' motivation to engage in crime. Employment in militia groups and gangs has been extensively studied in the sociological literature in particular, with mixed outcomes. Hill et al. 2001, Boas and Hatloy 2008 and Decker et al. 2008 find that poverty is a weak predictor of membership in organized crime groups, or, at the most, on par with numerous other explanations (social belonging, protection etc.). Finding that economic shocks in the agricultural sector are indeed causally linked to involvement in militia activity therefore has important implications for policies to minimize recruitment into militia groups and theft from oil pipelines in the context of the Niger delta.

The rest of the paper proceeds as follows: Section 3.2 provides background on oil pipeline sabotage in Nigeria; Section 3.3 details my empirical strategy; Section 3.4 provides an overview of the data; Section 3.5 lays out my hypotheses and estimating equations, Section 3.6 presents results, Section 3.7 presents various robustness checks, and Section 3.8 discusses and concludes.

3.2 Background

Conflict in Africa has been linked with poverty, weak institutions and economic dependence on natural resources (Elbadawi & Sambanis 2000). Many of these conflicts involve citizens working against governments who impose federal control of resources that have long been a source of people's livelihoods - as Olesegun Obasanje's government did in Nigeria. After 1999, when Nigeria made the transition from military to civilian rule, oil reserves became the *de facto* property of the federal government under the Nigerian constitution. Individual communities, therefore, have little say in the extent of extraction of oil from the delta.

3.2.1 Placement of Oil Pipelines

Oil pipelines in the Niger Delta region belong to either private petroleum companies or to the state-run oil corporation, the Nigerian National Petroleum Corporation (NNPC), which operates various subsidiaries. Such entities aim to minimize costs of construction of pipelines by constructing them on linear paths between oilfields.

The owners of proposed pipelines are required to sign Memoranda of Understanding with the local communities located close to the anticipated path of the pipelines. In these Memoranda, they outline ways to compensate communities for any potential damage resulting from pipeline construction and future leaks from the pipeline. Compensation is not linked to the current economic conditions in communities. However, in practice, MOUs are rarely enforced, and pipelines continue to laid in their planned locations without any kind of compensation to the communities affected[104].

[104]Any transfers made to households close to pipelines through these MOUs are captured in the direct effect of being close to pipelines, and should not affect the interactive effect of agricultural shocks and being located close to

3.2.2 Protecting Oil Pipelines

Both types of entities provide their own dedicated security for the pipelines and the oil and gas drilling stations. Local law enforcement is also mandated to prevent theft from oil pipelines. I do not directly observe the level of security provided to each segment of the pipeline or the communities around them. However, if additional security is provided to communities that are facing economic shocks, or communities that are close to pipelines, this should dampen any response in theft from oil pipelines when households close to pipelines are facing shocks to agricultural productivity.

3.2.3 Oil Pipeline Sabotage in Nigeria

The scale of oil pipeline sabotage in Nigeria is staggering. Over 100,000 barrels a day (valued at $1.6 billion per year) were stolen from pipelines and wellheads, and shipped to Eastern Europe, Singapore, Brazil, and elsewhere for refining (Ikelegbe 2005, Pagnamenta 2009). This corresponds to 4% of Nigeria's daily production. 2013 figures from the (former) governor of the central bank, Lamido Sanusi, put the overall shortfall due to theft and corruption, and the resulting stoppages in production, at $20 billion.

These large-scale operations are typically coordinated by militia groups in the delta, which have the expertise and equipment required to steal vast quantities of oil and finance their day-to-day operations. Anifowose et al. (2011) point out that "the technological sophistication required to untie a pipeline valve and subsequently use pumping machines to load products into waiting tanker(s) is likely beyond what the poor can afford." The level of theft from oil pipelines fairly closely tracks world oil prices, which drive the economic incentives of the militia groups (Figure 3.3).

The theft of large quantities of oil is economically motivated on the part of the militia, but could be politically motivated on the part of citizens who take part. These attacks are politically

---

oil pipelines.

motivated in three ways. First, the Niger delta region is owed a derivation of 13% of the revenues it produces from the federal government. However, the corruption of state officials and the low level of accountability has led to improper use of these funds, and lack of basic services. Second, producing oil in the region and transporting it to ports for export causes significant environmental harm to the surrounding areas - Inoni, Omotor, and Adun (2006) found that oil spills (partly due to sabotage) could lead to a loss of 1.3% of farm crop yield and a 5% reduction in farm income[105]. That Niger delta citizens would have to suffer all of the negative consequences of oil production with little of the gain exacerbates their displeasure with the federal government. Third, the history of Nigerian governance has led to the division of the country according to ethnic power blocs, leading to resentment from key minority groups. There is some anecdotal evidence that this history is being used as an incentive for minority citizens (especially significant minority groups like the Ijaw) to help militia groups sabotage the federal government through the theft of oil.

Future work will examine these political motives for participation in theft from oil pipelines. In this paper, however, I examine whether economic shocks induce local populations to engage in oil bunkering through participation in local militia activity, and whether this allows households to smooth consumption in the face of agricultural risk. I note that there is wide variation over time in the number of spills per month (Figure 3.1) and in the size of spills (Figure 3.2), suggesting that there is potential for engagement in criminal activity to be a short-term decision made independently in each time period (each agricultural cycle, for example).

3.2.4 Resources Used in Theft from Oil Pipelines

It is clear that pipeline sabotage is highly localized. Pipelines take up such a small portion of available land that the majority of the population in the region lives too far from pipelines to be actively involved in vandalism (Anifowose et al. 2012). I find in my household survey data that

[105] I further consider the implications of the direct impact of oil theft on crop yields in Section 3.7.

only 8.7% of households surveyed live within 15 kilometres of an oil pipeline. The level of local engagement necessary to carry out the large-scale theft of oil (known as bunkering) is high; "the ease with which vandals identify the pipelines buried in remote areas, sometimes six feet below the ground, shows some expert knowledge. The confidence with which they have been operating also reveal that they definitely have powerful godfathers (Adeniyi 2007)." According to Human Rights Watch, oil bunkering activities have been on the rise in the past decade because of lax security and political will to shut down bunkering routes, and more sophistication on the part of militia organizations.

Young, unemployed youth are typically the ones in these areas who, with the support of militia groups, engage in theft from oil pipelines. They are also the ones most likely to be affected by agricultural shocks, because of the labor-intensive, male-dominated nature of agriculture in the delta region (Ofuwoku & Chukwuji 2012). Lack of employment opportunities and poverty have proved to be correlates of theft from oil pipelines (Anifowose et al. 2012). However, previous work on this subject has considered only correlations between poverty and bunkering, and has not attempted to causally identify reasons for increased militia activity and bunkering in response to economic shocks. There is anecdotal evidence, however, that the militia groups compensate youth for their time and even provide public goods to the villages in the area. In this work, I causally identify incentives to join militia groups using various sources of identification that I detail in the next section.

3.3 Empirical Strategy

I use a classic differences-in-differences framework in my empirical strategy, with two key sources of variation that I exploit for identification.

3.3.1 Engagement in Crime

One of the limitations of studying employment in militia activity (or indeed, crime in general), is the difficulty inherent in measuing engagement in crime. To get around this issue, I use geographic variation in distance from the households to pipelines as a proxy for access to potential employment with militia groups. First, I argue that the precise locations of oil pipelines can be taken as exogenous to local economic conditions, particularly because they are laid in ways that will maximize network efficiency. If this is the case, communities that are closer to pipelines are comparable to those that are farther away, with the important exception that the households in those communities are more likely to have the opportunity to participate in theft from oil pipelines.

This stems from the fact that communities that have a higher density of pipelines closer to them are less likely to be chosen for their special characteristics, and more likely to be chosen for their geographic positioning. The generally linear pattern of the pipeline network corroborates this(Figure 3.4). In addition, pipeline segments are typically short, with a median length of 16 kilometres, and begin and end at oilfields.

3.3.2 Shocks to Agricultural Productivity

I interact households' distance to the nearest oil pipeline with farmer's self-reports on the gap between planned and realized crop yields, a measure of household-level shocks to agricultural productivity. These crop losses are measured as the percentage difference between the anticipated yield of the farmer at the time of planting and the actual yield he reports in the follow-up wave of the survey. If income from engaging in theft from oil pipelines is used to mitigate agricultural risk, then consumption should not fall, or should fall less, in response to negative agricultural shocks for households close to pipelines, relative to households farther away from pipelines.

In addition, this implies that there are more incidents of theft from pipelines in times of negative agricultural productivity shocks, given the increase in local labor supply to militia groups. I verify this by testing whether the number of reported incidents of oil pipeline sabotage are higher at spots along pipelines that are close to households reporting agricultural crop losses. More details on the data on oil pipeline sabotage are provided in Section 3.4.

I use a self-reported survey measure of crop losses to be able to identify household-level shocks to agricultural productivity. However, this variable presents a great deal of noise. In addition, to the extent that these losses partly result from fewer resources (material inputs or labor) given to agriculture by households who are employed in militia activity, they may over-report losses for households that are located close to pipelines. I therefore verify my results using community-level shocks to agricultural productivity that stem from insufficient or excessive rainfall, a common instrument in the literature.

3.4 Data

3.4.1 Oil Spills

No complete database exists of all incidents of pipeline vandalism in the Niger Delta region. I instead use a subset of pipeline sabotage events in which oil was spilled. Data on oil spills comes from the National Oil Spill Detection and Response Agency (NOSDRA), a regulatory body set up by the federal government to ensure an effective response to oil spills. Data are available on 9737 spills since 2001, together with data (for a subset of spills) on estimated spill quantity, precise location, and pipeline ownership. Using the latitude and longitude of each spill, I am able to pinpoint each spill relative to the households in the survey data. I also calculate the number of spills in a given radius from the household.

The estimated spill quantity includes both any amount that has been stolen and any amount that has simply spilled out of the pipeline in the process. Spills range from small amounts of sabotage by individual actors to large-scale theft of over a thousand barrels at a time. I assume that the former are related to individual households engaging in small-scale theft (which can also be used to mitigate income losses from agricultural shocks), while the latter stem from organized militia activity (requiring specialized equipment and knowledge). Both are potentially economically and politically motivated, so I incorporate both types of sabotage separately in my analyses.

There is the possible concern that a) the database does not include thefts that do not result in oil spills, and b) contains spills that are a result of mechanical failure due to accident or natural wear and tear to pipelines. I restrict all my analyses to the set of spills that have been labelled as 'sabotage', a broad categorization that requires that there was evidence of human intervention at the spill site. Assuming a negative relationship between the number of pipeline attacks carried out and the likelihood of a spill (as perpetrators hone their skills), the results I obtain are an underestimate of the response in number of spills following an economic shock.

3.4.2 Oil Pipelines

I have digitized oil and gas pipeline data from 2005 and 2015 from a detailed map of petroleum infrastructure in Nigeria titled the Petroleum Economist. The map provides high-resolution information on all oil and gas pipelines in Nigeria, as well as related features such as oil fields, refineries, depots, and tanker terminals.

For all the analyses in this paper, I use only the 2005 pipeline data, since my survey data cover a period up to 2013. There are two reasons for this. First, there is little variation in the pipeline network between the two years of the map data. Second, to the extent that new pipelines were constructed between 2005 and 2013 that affect households that I classify (using 2005 pipeline data) as far away from pipelines, the resulting expenditure effects I find are an underestimate.

3.4.3 Household Characteristics, Expenditure, and Crop Losses

Household-level data come from the General Household Surveys from 2006 to 2010, Demographic Health Surveys from 2008 and 2013, and LSMS surveys conducted between 2010 and 2013. The General Household Surveys are repeated cross-sections, and provide information on health, expenditures, and education. The LSMS survey is a household panel, and, crucially, provides precise GPS coordinates of the location of households (offset up to 5 km to preserve the privacy of the respondent[106]). I know, therefore, how close each household is to the nearest point on the oil pipeline network, and therefore how much access households are likely to have to militia activity. I view this variation in household distance from the pipeline network as essentially random (as explained in further detail in the previous section). I match household data with data on the number of oil spills in the month in which the survey is conducted.

In this paper, I use the LSMS data only, and focus specifically on expenditures in two broad categories – food and non-food expenditure. I consider expenditures in the month of the survey (after harvest-time), which I use as a proxy for household income.

In my analyses, I consider households with access to criminal employment to be those that are within 15 kilometers of an oil pipeline segment (Figure 3.7). I set 15 kilometers to be a reasonably conservative threshold because it is the typical distance that can be traversed by bicycle. As a robustness check, I group households according to whether they are less than 15km, 15-30km, 30-60km, and further than 60km from pipelines, to investigate how quickly the the effects die off with distance from the pipeline network.

The LSMS survey also provides information on shocks to agricultural productivity. Each of the two rounds of the survey consists of a pre-planting and a post-harvest wave. Each household is

[106]The offset leads to measurement error in the household's distance to the closest oil pipeline, which should dampen treatment effects that I find. Due to the methodology for the offset being entirely independent from the placement of oil pipelines, there should be no systematic difference in how houeholds closer than 15km and households farther away than 15km from pipelines are affected.

therefore surveyed twice. The first wave of the survey asks respondents about anticipated yields for the season ahead, while the second asks about realized yields. I take the difference between the two for each crop (as a percentage of output), and create an average across all crops weighted by the amount of land devoted to each crop. I use this as a measure of households' unanticipated negative shocks to agricultural productivity. This defines the shock for the entire agricultural cycle for each household. Should households adjust the level of inputs they put into agriculture to adjust for realizations of shocks during the course of the agricultural cycle, this variable measures the loss in output that farmers were unable to compensate for by adjusting inputs.

One possible issue is that militia activity may itself cause economic harm to communities near pipelines in unobservable ways – by making it difficult for citizens to move around and engage in productive activity, for example (Pinotti 2012). To the extent that militia activity itself depresses economic gains, households near oil pipelines should be worse off overall, and we would, in addition, expect the effects to be the same on households with and without young men.

A possible further confounding factor is whether the direct environmental effect of oil pipelines on crop yields affects household responses to economic shocks. I present two tests of whether yields differ in a systematic way between households close to and further from pipelines, and find that they do not[107]. I also exclude any households within 1km of a pipeline from all my analyses, to exclude the direct effect of soil contamination from oil spills on crop yields.

[107]Results are presented in 3.7.1.

3.4.4 Rainfall

I use data on rainfall shocks[108] as instruments for the level of alternate employment in agriculture in a given area. To the extent that rainfall alone is potentially correlated with propensity for militia activity[109], the interaction of this instrument with the distance of the household from the nearest point on the oil pipeline network provides a suitable independent measure for the differential response in crime rate and household consumption in periods of low agricultural productivity for households with access to oil pipelines.

3.4.5 Other Types of Conflict

To examine the impact of agricultural shocks on militia activity and other types of conflict, I turn to the ACLED database, which uses various media sources to collate information on the geolocations of various types of conflict occurring in Nigeria. The database also contains information on conflict intensity (proxied by number of casualties), and the general affliliations of the actors involved in the conflict. I narrow this database to 747 incidents of conflict in Nigeria in the period 2010-2014, and, within these, focus on conflicts that involve at least one militia group[110]. I calculate the number of such events that occur within 15km of each household in my data.

[108]Rainfall shocks are constructed as in Chapter 1, as a percentage deviation from the long-run average for the village. Unlike in Chapter 1, however, I take rainfall data for the entire agricultural cycle, rather than only in the early part of the season, since I consider measures of expenditure and of theft from oil pipelines in the post-harvest period only.

[109]This relationship is well documented in the previous literature.

[110]Within these, I also include conflicts involving unidentified armed groups.

3.4.6 Access to Financial Institutions

The Nigeria Cash-In-Cash-Out dataset of provides the precise geographic coordinates of close to 16,000 Cash-In-Cash-Out points, defined as bank branches, mobile money agents[111], savings and credit cooperatives, sand micro-finance institutions. I use the number of locations within a 10km radius of the household as a proxy, broadly, for credit access. The median household has two such locations within a 10km radius. However, the number of such locations close to households a) may correlate with other characteristics that have a direct impact on household income, and b) may not necessarily be a direct measure of the ability to mitigate risk, but could correlate with other measures of the ability of the household to insure itself against risk. To deal with this, I include the variable both directly and interacted with the interaction term in my main specification. I provide more details in Section 3.6.

3.5 Hypotheses and Empirical Specifications

Table 1 provides summary statistics for observable characteristics of households close to (within 15km) and further from (greater than 15km) from pipelines.

Defining Crop Losses: In all following specifications, I include crop shocks in two ways: a dummy variable losses of over 50% of anticipated yields, and a continuous variable for percentage of anticipated output lost. 30% of all households in my post-harvest sample report losing 50% or more of the output they had anticipated in the pre-planting survey.

Defining Distance to the Pipeline: I define a household as close to an oil pipeline as a binary variable, which is an indicator for being less than 15km away from a pipeline in all main specifications.

[111]Broadly, any location that has been registered by a mobile money operator to offer mobile money services.

As a robustness check, I repeat my analyses using distance as a categorical variable[112].

Hypotheses and Empirical Specifications: My hypotheses and the empirical specifications I implement to test them are as follows:

H1: I anticipate that there will be a higher number of spills in the month of the survey within a 15km radius of households facing larger proportions of crop loss. I also anticipate a higher number of large-scale spills (indicative of militia activity) within this 15km radius of households that lose a greater proportion of their crop. To test this I run the following specification:

$$\text{Num Spills in 15km Radius}_{jlt} = \beta_0 + \beta_1 1(\text{Pct Losses>50\%})_{jlt} + \beta_2 1.(\text{Dist to Pipeline <15km})_j$$
$$\beta_3\ 1.(\text{Pct Losses>50\%})*1.(\text{Dist to Pipeline <15km})_{jlt} + \gamma_l + \delta_t + \epsilon_{jlt}$$

where j is the household surveyed, l is the geographic unit (LGA) and t refers to the time period of the survey. I include LGA[113] and survey monthXyear fixed effects (for the post-planting survey, which corresponds with the timing of oil spills and measures of household expenditure) in all analyses. I also include a vector of household and village characteristics.

Test: $\beta_3 > 0$

H2 I predict that the increase in the number of oil spills will correspond to a greater increase in militia activity close to oil pipelines, and, in particular, a higher number of conflict events in the vicinity of oil pipelines. This is over and above any baseline increase in conflict events in response to negative agricultural shocks (as people whose labor is less productive in agriculture join militia groups in response).

H3a: I expect that households that are close to oil pipelines are able to smooth consumption

[112] Results are presented in Section 3.7.4.

[113] I do not include village fixed effects because it would absorb the distance variable when it is defined as a binary, which would prevent me from examining the direct impact of being close to an oil pipeline on a household.

expenditure using income from militia activity when income from their main activity (agriculture) is curtailed. Therefore, households less than 15 km from pipelines should show a smaller decrease in household expenditure in periods of agricultural shock than those farther away.

H3b: I anticipate that there will be an effect on household consumption specifically for households with young men self-reporting as unemployed, who are most likely to be recruited into militia groups[114]. The category of expenditure will be affected will depend on whether the households in my data are at the subsistence threshold.

H3c: I anticipate that there will be an effect on household consumption specifically for agricultural households, and not for households that do not cultivate crops and are therefore not affected by shocks to agricultural productivity.

H4: Households that are credit-constrained will be more likely to turn to theft from oil pipelines in response to negative shocks to agricultural productivity.

To estimate the effect of the shocks on expenditures for various types of households, I use the same specification as above, but with food, non-food, and total expenditures as the outcome variables for H3a, 3b, and H4. For H3c, I use rainfall shocks at the community-level rather than household-level shocks as defined by unanticipated lost output. This is due to the fact that rainfall allows me to define the *potential* shock to non-agricultural households (defined as those households that do not produce any crops), and show that they do not respond. To test this I run the following specification:

$$\begin{aligned} \text{Num Spills in 15km Radius}_{jlt} &= \beta_0 + \beta_1 1(\text{Rainfall Shock})_{jlt} + \beta_2 1.(\text{Dist to Pipeline} <15\text{km})_j \\ &\beta_3\ 1.(\text{Rainfall Shock})*1.(\text{Dist to Pipeline} <15\text{km})_{jlt} + \gamma_l + \delta_t + \epsilon_{jlt} \end{aligned}$$

As before, I test whether β_3 is positive.

[114]Gould et al. (2002) find causal links between labor market opportunities and crime for young men in the 1980s and 1990s in the US.

3.6 Results

In this section, I present results from my main empirical strategy, with a focus on the interaction between a household's unanticipated crop losses and being located close to an oil pipeline.

Oil Spills: *H1:* Table 3.1 presents results from the main specification, with the number of oil spills within a 15km radius of the household as the outcome of interest. I find an increase of 0.66 spills per month when percentage losses in anticipated yields are above 50% for households within 15km of pipelines. Further, I find an increase of 0.13 *large* spills of over 1000 barrels per month in these households (results presented in Table 3.2). This corresponds to an increase of 92.6% on the mean number of large spills in the vicinity of a household whose losses are less than 50% of anticipated output.

Furthermore, the overall increase in the number of spills in a 15 km radius of the household is entirely driven by households with greater than 50% loss of anticipated agricultural output. That is, households close to pipelines that do not face significant agricultural shocks do not experience a greater number of oil spills along the pipeline in their immediate vicinity. This suggests that only large shocks to agricultural productivity can trigger a shift of labor supply into criminal activity.

Militia Conflict Events: *H2:* Table 3.3 considers how the number of conflict events in which militia groups participate responds to agricultural shocks, particularly those close to oil pipelines. I find an increase of 5% in the number of conflict events involving militia groups in response to a negative agricultural shock faced by the household. However, this doubles when the household in question is less than 15km from an oil pipeline, indicating an even higher incidence of militia activity close to oil pipelines when communities face income shocks from unanticipated shocks to agricultural productivity and are likely to smooth income by engaging in criminal activity.

Expenditures: *H3a:* Table 3.4 examines the response in various kinds of expenditure to agricultural productivity shocks, proximity to oil pipelines, and the interaction of the two. In it, I present results for specifications that include percentage of crop lost as both a continuous variable (columns 1-3) and as a binary variable taking the value 1 if over 50% of anticipated yields are lost (columns 4-6). I describe in detail the results from the latter set of specifications, given that I do not expect participation in militia activity to be linear in the increase in crop loss. While the results in columns 1-3 have the sign I would expect, estimates are noisy. I therefore use the most reduced-form specification possible.

I look at overall expenditure, and food and non-food components separately. I find that non-food expenditures are the hardest hit in times of shock to agricultural output, for an average impact of about 28K Naira, as expected. However, households close to pipelines do not face decreases in non-food expenditure when they face agricultural shocks – in fact, they make up for this shortfall by almost 200% (about half of median non-food expenditure in the sample). It is unlikely that these respondents are subsistence farmers given the relatively large expenditures at baseline. Therefore, it follows that this must be due to an additional source of income in bad agricultural seasons for those close to pipelines. This, combined with the corresponding increase in oil spills, indicates that income from oil theft is likely compensating for the economic loss.

H3b: I use the same specification as above to look at heterogeneity in household expenditures by households close to oil pipelines with and without unemployed male youth, the target demographic for militia recruitment. I present results in Table 3.5. Column 1 presents effects on non-food expenditures for all households using my preferred specification (identical to Table 3.4). Columns 2 and 3 divide the sample into households with and without unemployed youth between the ages of 14 and 25 respectively. Columns 4 and 5 divide the sample into households reporting in the post-harvest survey that they had youth between the ages of 14 and 25 who did not attend school in the past year, and households that did not, respectively.

Both Columns 2 and 4 therefore select households with individuals who are, at least anecdotally, the target for militia recruitment. Among these households, those that are less than 15 km from a pipeline show much higher non-food expenditure at baseline when compared with households farther away, relative to households without these groups of individuals (in a comparison of the direct effect of the distance indicator in Column 2 with 3, or 4 with 5). This suggests that such households have an alternate source of income, which I attribute to a greater degree of involvement in theft from oil pipelines at baseline.

These households also show slightly larger - about 86.5K Naira - estimates of the interaction term in the main specification (albeit with some noise). That is, households close to pipelines with unemployed male youth facing large crop losses reduce non-food expenditures less than households close to pipelines without unemployed male youth facing crop losses. This supports the idea that economic gains from oil theft are making up for losses due to shocks to agricultural productivity.

H3c: Households engaging primarily in agriculture (that is, without alternate sources of income) should be affected to a greater degree by agricultural productivity shocks than households that already report income from other sources in the pre-planting period. I test this by, again, dividing the sample between agricultural and non-agricultural households, and find that agricultural households drive the patterns of consumption smoothing close to pipelines. In these specifications, I use the community-level measure of agricultural shock for each household - that is, the deviation from the long-run average level of rainfall during the year.

I present results in Table 3.7. Columns 1 and 2 examine the response of spill incidence to agricultural shocks and the binary distance variable for agricultural and non-agricultural households separately. The results suggest that only spills in the vicinity of agricultural households respond to agricultural shocks. Column 3 confirms that non-food expenditure does not respond to distance from oil pipelines, rainfall shocks, or the interaction of the two, suggesting that households that do not cultivate crops are not vulnerable to agricultural shocks, and do not turn to oil pipeline theft as an insurance mechanism.

H4: Access to credit should mediate households' response to agricultural shocks, and their resulting engagement in theft of oil from pipelines. I therefore divide my sample by the median number of Cash-In-Cash-Out points within 10km of the household, and run the main specification separately for each half of the distribution. I present results in Table 3.6.

The direct impact (for households which lose less than 50% of their anticipated yields) of being close to pipelines on the number of oil spill incidents is positive only in areas with a higher-than-median number of CICO points, which indicates that credit access and oil spills are positively correlated at baseline. This runs counter to the argument that militia activity may make it *more* difficult to keep financial access points active. I then turn to the interaction between crop loss and proximity to oil pipelines, and find that it is positive and signifcant only for households with low access to CICO points. This suggests that, as a source of protection against agricultural risk, credit access is able to mitigate households' participation in theft from oil pipelines.

However, this is merely suggestive evidence that credit access is an important mediator of households' engagement in criminal activity; since none of the variation in credit acess is quasi-random, I make no claims of causality in this particular analysis of household heterogeneity. This analysis is also silent on the specific mechanism for the link between credit access and agricultural risk - it is possible that households with greater credit access also have better access to insurance mechanisms of various kinds. It is also possible that credit itself serves as insurance when household consumption is squeezed due to agricultural losses.

3.7 Robustness Checks

I carry out a number of robustness checks of my main results.

3.7.1 Yield Responses to Environmental Impact of Oil Pipelines

To rule out a direct environmental effect of oil pipelines or additional spills on agricultural yields (which would affect the unanticipated nature of shocks to agricultural productivity), I use a householdXplotXtime level dataset. There is only one crop per 'plot'. I test whether yields for each crop differ for plots owned by households less than 15 km from the pipeline relative to plots farther away, and find that they do not. I present results in Table C.1. For this analysis, in addition to the controls in the main specification, I include crop fixed effects and plot characteristics.

3.7.2 Yield Responses to The Presence of Militia Groups

The presence of militia groups near households (which we might expect to particularly be those close to pipelines) alone may be sufficient to have either a positive (due to income transfers) or negative (due to destruction of surrounding infrastructure and general fear) impact on agricultural productivity, yields, or income. There are two reasons my empirical strategy still applies. First, I allow for the direct effect of being close to pipelines separately from the interactive effect of agricultural productivity shocks and access to pipelines (and the resulting criminal employment opportunities). Second, the interaction I specify may be affected in the following way: militia groups may be more likely to target communities facing agricultural shocks. If they do have a negative impact on these communities, then my estimates of the differential change in expenditure in response to agricultural productivity shocks will understate the transfers from militia groups to local households. Should the mere presence militia groups have a positive effect on household incomes, I would not expect to specifically find differential effects for households with unemployed young men or agricultural households, with the exception of the oil bunkering mechanism. This is further supported by the increase in spill incidents near those households.

3.7.3 Community-level Rainfall Shocks

I use an alternate measure, at the community-level, of shocks to agricultural productivity from rainfall. I use percentage deviation from the long-run average for each enumeration area, and proxy for unanticipated agricultural income losses using this measure. I find similar increases in theft from oil pipelines for households less than 15 kilometers from a pipeline, in response to rainfall shocks as measured by percentage deviation from the long-run mean level of rainfall for that community over the agricultural cycle.

I present results in Table C.3. I find positive coefficients on being closer to an oil pipeline, but the bulk of the increase in the number of spills is driven by spills in the vicinity of households facing rainfall shocks, as the interaction term suggests.

3.7.4 Distance Measurements

I classify households into four buckets - less than 15km, 15-30km, 30-60km, and further than 60km from pipelines (the omitted category), to examine how quickly the impact on theft from oil pipelines decreases with distance. I present results in Table C.2. The first column confirms that a) There are zero spills less than 15km from a household that is more than 15km from a pipeline (this is one piece of evidence to confirm both the accuracy of the pipeline digitization and the location-coding of the spills in the database).

Column 2 then looks at non-food expenditure for each category of distance interacted with percentage loss in anticipated crop output. I find, as expected, a direct negative effect of percentage losses over 50% on non-food expenditure for every distance category. However, only households less than 15km from pipelines are able to more than make up for this shortfall. This, again, suggests that proximity to pipelines provides an additional (and, in fact, more lucrative) source of income than agriculture.

3.8 Conclusion

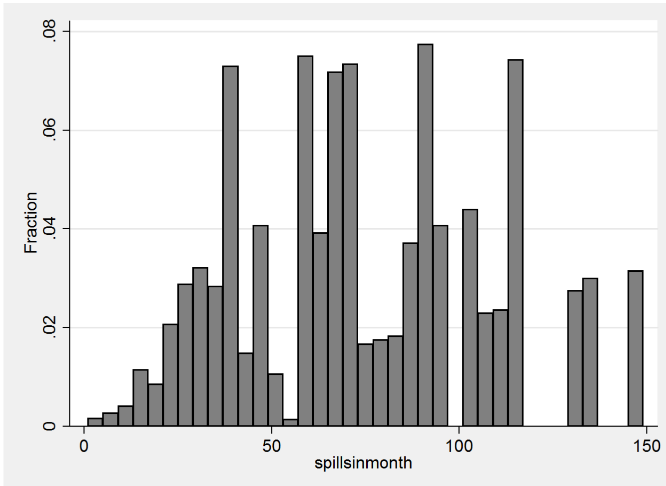
This paper investigates whether agricultural risk provides impetus for engaging in criminal activity - specifically, in the form of working with militia groups in the Niger delta to steal large quantities of oil from pipelines - and finds that it does. It also uses various types of household heterogeneity analyses to identify and confirm factors that select at-risk households - those with unemployed or non-school-attending youth, and those with lack of access to a robust financial structure (which I interpret as lack of access to credit).

From a policy perspective, a significant amount of money can be saved by reducing the incidence of theft from oil pipelines - that is a first-order priority for both the government and private oil companies. Importantly, security can be targeted to locations that are known to be facing local economic shocks. To the extent that access to insurance and credit can mitigate households' engagement in theft from oil pipelines, the risk factors I suggest are important for militia recruitment can be taken into account in providing such safety nets for farmers.

Future work will study the political motives for joining militia groups - whether ethnic fractionalization or resentment against the government contributes, and how the magnitudes of those effects compare to the economic opportunity cost motive. It is also important to study whether economic shocks, by encouraging households to turn to employment with militia groups, could lead to the formation of a criminal career - that is, whether these effects lead to a permanent increase in militia activity after the agricultural productivity shock has dissipated.

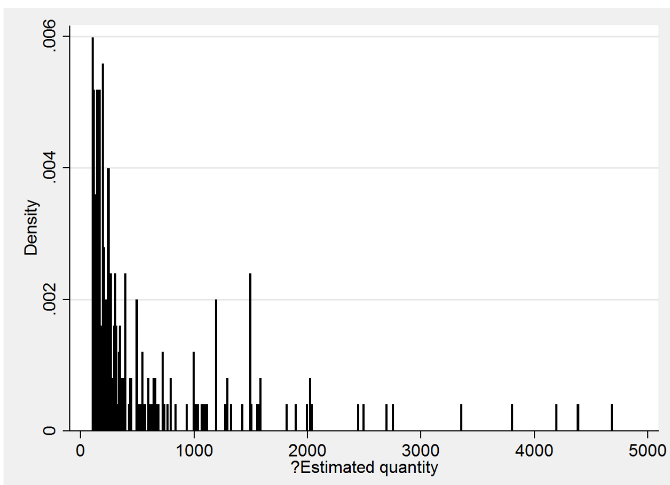
3.9 Figures

Figure 3.1: Distribution of Number of Spills



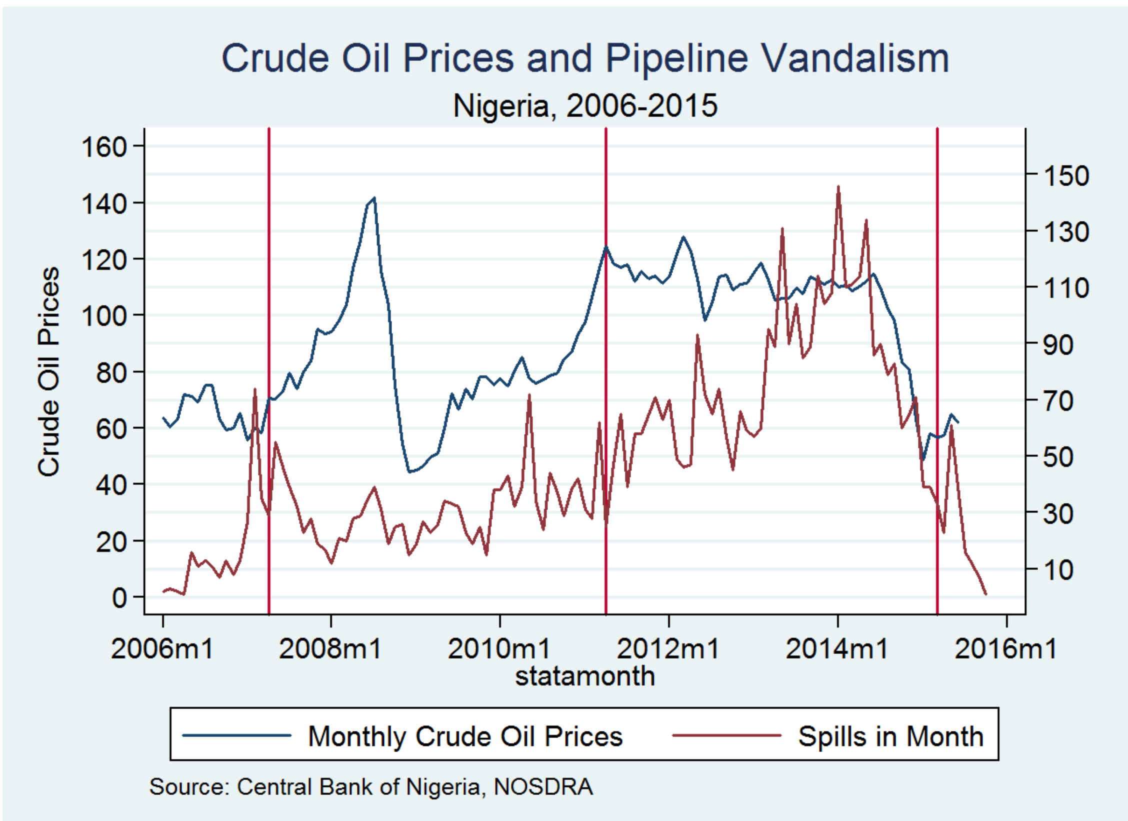
Back to text

Figure 3.2: Distribution of Spill Volume (Barrels)



Back to text

Figure 3.3: World Oil Prices and Oil Theft



Back to text

Figure 3.4: Oil Pipeline Network in Nigeria

Source: The Petroleum Economist, 2005

Back to text

Figure 3.5: Surveyed Households

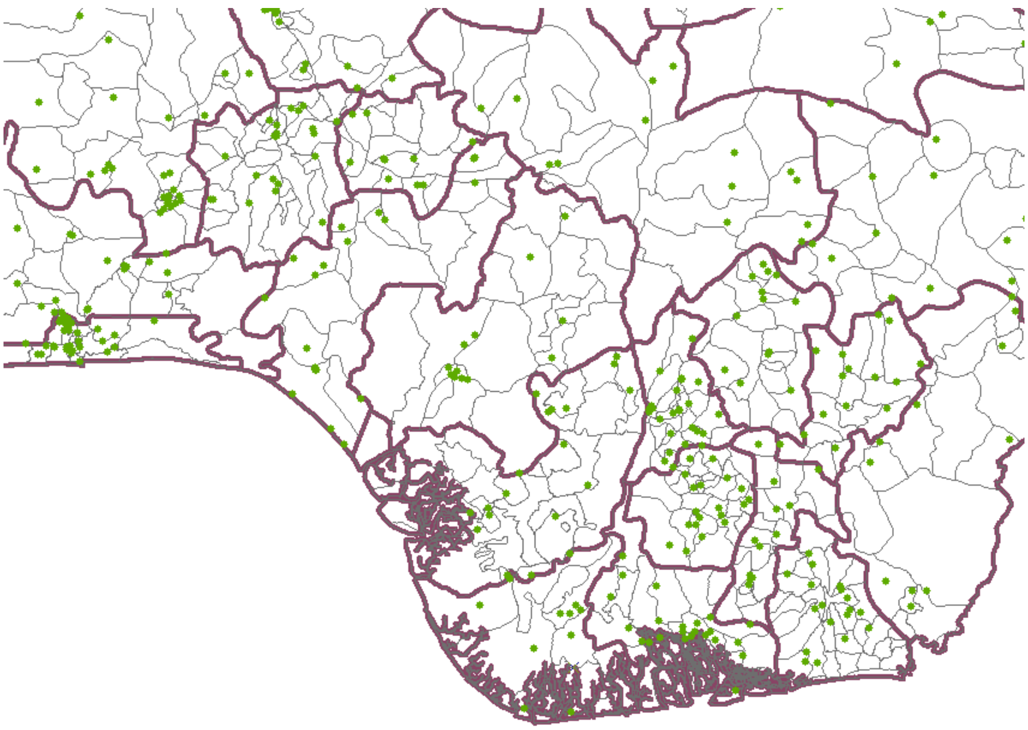


Figure 3.6: Empirical Method: Pipeline Network

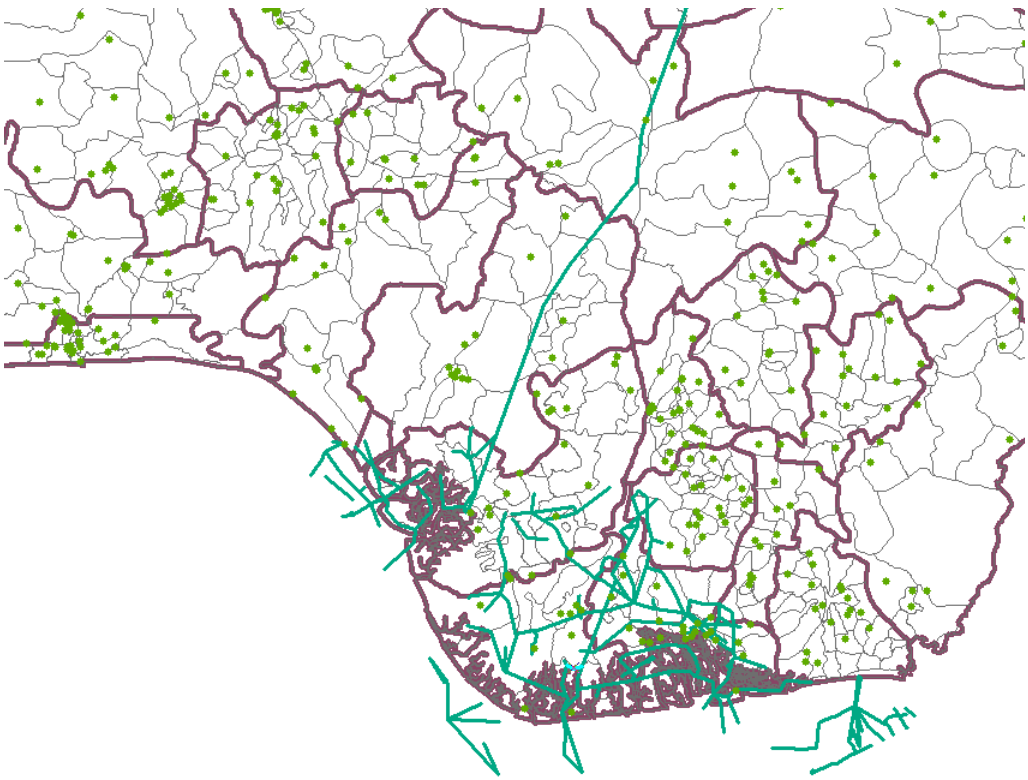
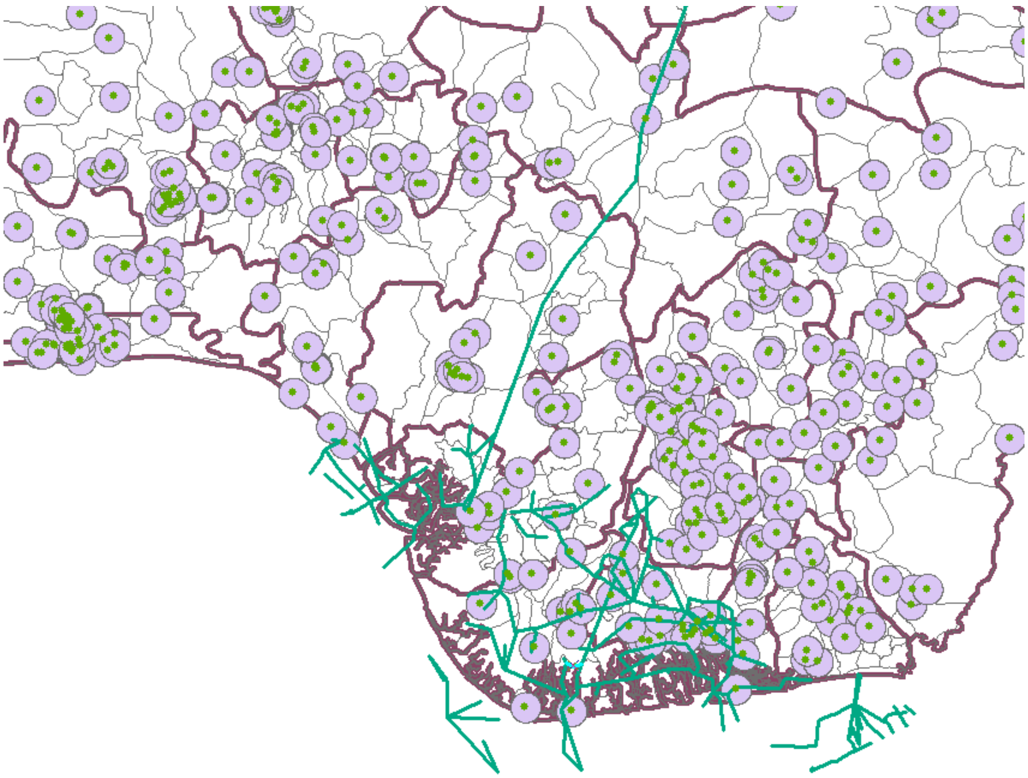


Figure 3.7: Empirical Method: Distance to Pipeline



Back to text

3.10 Tables

Table 3.1: Spills in Month- Response to Agricultural Productivity Shocks

VARIABLES	(1) Spills in Month	(2) Spills in Month	(3) Spills in Month	(4) Spills in Month
1(Dist<15km)	-0.0169	-0.277*	-0.0376	-0.286
	(0.149)	(0.167)	(0.190)	(0.179)
Pct Crop Lost	-5.92e-05	-0.000116		
	(0.000134)	(0.000110)		
1(Dist<15km)*Pct Loss)	0.00837**	0.00676*		
	(0.00392)	(0.00380)		
1(Pct Loss > 50%)			-0.00506	-0.00573
			(0.0150)	(0.0122)
1(Dist<15km)*1(Pct Loss>50)			0.767**	0.659*
			(0.385)	(0.385)
Observations	10,905	10,905	10,905	10,905
HH Char	Yes	Yes	Yes	Yes
Village Char	Yes	Yes	Yes	Yes
State FE	Yes		Yes	
LGA FE		Yes		Yes
Time FE	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline.

Back to text

Table 3.2: Large Spills in Month- Response to Agricultural Productivity Shocks

VARIABLES	(1) Spills in Month	(2) Spills in Month	(3) Spills in Month	(4) Spills in Month
1(Dist<15km)	-0.0795*	-0.0433	-0.106*	-0.0587
	(0.0432)	(0.184)	(0.0570)	(0.221)
Pct Crop Lost	1.53e-05	-1.48e-05		
	(0.000193)	(0.000180)		
1(Dist<15km)*Pct Loss	0.00232***	0.00111**		
	(0.000520)	(0.000479)		
1(Pct Loss > 50%)			0.00739	0.00533
			(0.0245)	(0.0230)
1(Dist<15km)*1(Pct Loss>50)			0.234***	0.132**
			(0.0563)	(0.0523)
	(0.0330)	(0.302)	(0.0386)	(0.344)
Observations	10,905	10,905	10,905	10,905
HH Char	Yes	Yes	Yes	Yes
Village Char	Yes	Yes	Yes	Yes
State FE	Yes		Yes	
State FE	Yes		Yes	
LGA FE		Yes		Yes
Time FE	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline.

Back to text

Table 3.3: Number of Violent Militia Events

VARIABLES	(1) Events in Year
1(Dist<15km)	0.105
	(0.653)
1(Pct Loss > 50%)	0.217*
	(0.117)
1(Dist<15km)*1(Pct Loss>50)	0.199*
	(0.102)
Observations	268
LGA FE	Yes
Time FE	Yes
Mean	4.184

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. The outcome variable is the number of conflict events from the ACLED database involving at least one milita group that occur within 15 km of the household during the agricultural cycle.

Back to text

Table 3.4: Expenditure Responses to Agricultural Shocks

VARIABLES	(1) Total Exp	(2) Total Food Exp	(3) Non-food Exp	(4) Total Exp	(5) Total Food Exp	(6) Non-Food Exp
Distance < 15km	45,905	38,962*	-39,791**	64,965**	63,692***	-34,268**
	(28,579)	(22,220)	(15,855)	(27,685)	(21,610)	(14,649)
Pct Crop Loss	-307.3	42.21	-344.5***			
	(187.5)	(179.1)	(45.12)			
Distance < 15*Pct Crop Loss	210.7	-183.4	399.3			
	(403.5)	(308.1)	(349.9)			
1(Pct Crop Loss > 50%)				-11,941	16,774	-28,274***
				(16,374)	(16,006)	(3,469)
Distance < 15*1(Pct Crop Loss > 50%)				30,298	-20,572	52,281*
				(37,349)	(30,369)	(31,252)
Observations	10,905	10,905	10,905	10,905	10,905	10,905
HH Char	Yes	Yes	Yes	Yes	Yes	Yes
Village Char	Yes	Yes	Yes	Yes	Yes	Yes
LGA FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. I exclude one observation with an outlying level of expenditure.

Back to text

Table 3.5: Heterogeneous Expenditure Responses to Agricultural Shocks

	(1) All Households Non-Food Exp	(2) Unemployed Youth Non-Food Exp	(3) No Unemployed Youth Non-Food Exp	(4) Youth Not in School Non-Food Exp	(5) No Youth Not in School Non-Food Exp
1(Dist<15km)	-34,268**	318,219***	-52,118***	194,867***	-32,382*
	(14,649)	(109,591)	(16,651)	(44,593)	(16,590)
1(Pct Loss > 50%)	-28,274***	-30,526**	-20,772*	-29,415**	-23,265**
	(3,469)	(12,005)	(11,251)	(13,914)	(11,712)
1(Dist<15km)*1(Pct Loss>50)	52,281*	88,705	21,638	84,715*	21,535
	(31,252)	(67,799)	(14,184)	(50,236)	(14,912)
Observations	10,905	580	8,883	1,092	8,371
HH Char	Yes	Yes	Yes	Yes	Yes
Village Char	Yes	Yes	Yes	Yes	Yes
LGA FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. I exclude one observation with an outlying level of expenditure. Column 1 includes all households in the main sample. Column 2 includes only households who report unemployed youth between the ages of 14 and 25 in the household, while Column 3 includes all other households. Column 4 includes only households who report that youth between the ages of 14 and 25 in the household have not attended school in the past year, while Column 5 includes all other households.

Back to text

Table 3.6: Heterogeneity By Credit Access

VARIABLES	(1) Low Access	(2) High Access
1(Dist<15km)	-0.111	0.431***
	(0.0851)	(0.151)
1(Pct Loss > 50%)	-0.00192	0.00748
	(0.0105)	(0.00855)
1(Dist<15km)*1(Pct Loss>50)	0.796**	-0.128
	(0.354)	(0.332)
Observations	6,053	4,811
R-squared	0.297	0.262
HH Char	Yes	Yes
Village Char	Yes	Yes
LGA FE	Yes	Yes
Time FE	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. Low Access is defined as households that have a lower than median number of CICO points wiithin a 10 km radius of the household.

Back to text

Table 3.7: Heterogeneity by Cultivating Household

VARIABLES	(1) Ag HH Spills in Month	(2) Non-Ag HH Spills in Month	(3) Non-Ag HH NonFood Expenditure
1(Dist<15km)	-0.145	0.00434	17,048
	(0.181)	(0.0210)	(68,445)
Village-Level Rainfall Shock	0.0589***	0.0274**	-5,624
	(0.0156)	(0.0121)	(4,124)
1.(Distance<15km)*1(Rainfall Shock)	1.969**	0.0622	-2,417
	(0.875)	(0.0546)	(48,395)
Observations	10,905	7,268	7,268
R-squared	0.270	0.376	0.350
HH Char	Yes	Yes	Yes
Village Char	Yes	Yes	Yes
LGA FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. A rainfall shock is defined as a positive or negative deviation from the village's LR average of greater than 20%. Column 1 restricts the sample to agricultural households only. Column 2 and 3 use the sample of households that do not cultivate any crops.

Back to text

References

Akerlof, G. A. and J. L. Yellen (1990): "The Fair Wage-Effort Hypothesis and Unemployment," The Quarterly Journal of Economics, 105, 255–283. 2.1

Ambrus, A., M. Mobius, and A. Szeidl (2014): "Consumption Risk-Sharing in Social Networks," The American Economic Review, 104, 149–182.

Angelucci, M., D. Karlan, and J. Zinman (2013): "Win some lose some? Evidence from a randomized microcredit program placement experiment by Compartamos Banco," NBER Working Paper 19119.

Angrist, J. D. and A. D. Kugler (2008): "Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia," The Review of Economics and Statistics, 90, 191–215.

Anifowose, B., D. M. Lawler, D. Van der Horst, and L. Chapman (2012): "Attacks on oil transport pipelines in Nigeria: A quantitative exploration and possible explanation of observed patterns," Applied Geography, 32, 636–651.

Ardington, C., A. Case, and V. Hosegood (2009): "Labor supply responses to large social transfers: Longitudinal evidence from South Africa," American Economic Journal: Applied Economics, 1, 22–48.

Atkin, D. (2013): "Trade, tastes, and nutrition in India," American Economic Review, 103, 1629–1663.

Attanasio, O., A. Kugler, and C. Meghir (2011): "Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial," American Economic Journal: Applied Economics, 188–220.

Augsburg, B., R. De Haas, H. Harmgart, and C. Meghir (2012): "Microfinance at the margin: Experimental evidence from Bosnia and Herzegovina," NBER Working Paper 18538.

Axbard, S. (2016): "Income Opportunities and Sea Piracy in Indonesia: Evidence from Satellite Data," American Economic Journal: Applied Economics, 8, 154–94.

Azam, M. (2011): "The impact of Indian job guarantee scheme on labor market outcomes: Evidence from a natural experiment," Unpublished Manuscript.

Bandiera, O., I. Barankay, and I. Rasul (2013): "Team incentives: Evidence from a firm level experiment," Journal of the European Economic Association, 11, 1079–1114.

Banerjee, A., E. Breza, E. Duflo, and C. Kinnan (2015a): "Do credit contstraints limit entrepreneurship? Heterogeneity in the returns to microfinance," Working Paper.

Banerjee, A., D. Karlan, and J. Zinman (2015b): "Six randomized evaluations of microcredit: Introduction and further steps," American Economic Journal: Applied Economics, 7, 1–21.

Bardhan, P. K. (1973): "Size, productivity, and returns to scale: An analysis of farm-level data in Indian agriculture," The Journal of Political Economy, 1370–1386.

Basu, A. K. (2013): "Impact of rural employment guarantee schemes on seasonal labor markets: Optimum compensation and workers' welfare," Journal of Economic Inequality, 1–34.

Beaman, L. and J. Magruder (2012): "Who gets the job referral? Evidence from a social networks experiment," The American Economic Review, 3574–3593.

Becker, G. S. (1968): "Crime and punishment: An economic approach," in The economic dimensions of crime, Springer, 13–68.

Beegle, K., E. Galasso, and J. Goldberg (2015): "Direct and Indirect Effects of Malawi's Public Works Program on Food Security," Working Paper. 2.1, 2.5.2

Behrman, J. R., A. D. Foster, and M. R. Rosenzweig (1997): "The dynamics of agricultural production and the calorie-income relationship: Evidence from Pakistan," Journal of Econometrics, 77, 187–207.

Benjamin, D. (1992): "Household composition, labor markets, and labor demand: Testing for separation in agricultural household models," Econometrica: Journal of the Econometric Society, 287–322.

Berg, E., S. Bhattacharyya, R. Durgam, and M. Ramachandra (2012): "Can rural public works affect agricultural wages? Evidence from India," .

Berndt, E. R. and L. R. Christensen (1973): "The translog function and the substitution of equipment, structures, and labor in U.S. manufacturing 1929-68," Journal of Econometrics, 1, 81 – 113.

Besley, T. and S. Coate (1995): "The design of income maintenance programmes," The Review of Economic Studies, 62, 187–221.

Bhargava, A. K. (2014): "The impact of India's rural employment guarantee on demand for agricultural technology," Unpublished Manuscript.

Binswanger-Mkhize, H. P. (2012): "Is There Too Much Hype about Index-based Agricultural Insurance?" The Journal of Development Studies, 48, 187–200.

Blakeslee, D. and R. Fishman (2014): "Weather shocks, crime, and agriculture: evidence from India," Social Science Research Network.

Blattman, C., N. Fiala, and S. Martinez (2013): "Generating skilled self-employment in developing countries: Experimental evidence from Uganda," Quarterly Journal of Economics.

Blattman, C. and L. Ralston (2015): "Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs," Unpublished Manuscript.

BøÅs, M. and A. Hatløy (2008): "'Getting in, getting out': militia membership and prospects for re-integration in post-war Liberia," The journal of modern African studies, 46, 33–55.

Bowles, S. and H. Gintis (2002): "Social Capital and Community Governance," The Economic Journal, 112, F419–F436.

Breza, E., S. Kaur, and Y. Shamdasani (2016): "The Morale Effects of Pay Inequality," NBER Working Paper 22491. 2.1, 2.1, 2.2.1, 2.2.4

——— (2018): "Testings for Labor Rationing: Revealed Preference Estimates from Demand Shock," Working. 2.1

Brooks, J., P. M. Filipski, E. Jonasson, and E. Taylor (2012): "Modeling the Welfare Implications of Agricultural Policies in Developing Countries," .

Burgess, R. and D. Donaldson (2017): "Railroads and the Demise of Famine in Colonial India," Unpublished Manuscript.

Bustos, P., B. Caprettini, and J. Ponticelli (2016): "Agricultural productivity and structural transformation: Evidence from Brazil," American Economic Review.

Card, D. (2011): "Origins of the unemployment rate: The lasting legacy of measurement without theory," The American Economic Review, 101, 552–557.

Card, D. and D. Hyslop (1996): "Does Inflation "Grease the Wheels of the Labor Market"?" National Bureau of Economic Research Working Paper Series, No. 5538.

Card, D., A. Mas, E. Moretti, and E. Saez (2012): "Inequality at work: The effect of peer salaries on job satisfaction," The American Economic Review, 102, 2981–3003. 2.1

Caselli, F. (2005): "Chapter 9: Accounting for Cross-Country Income Differences," Elsevier, vol. 1 of Handbook of Economic Growth, 679 – 741.

Caves, D. W., L. R. Christensen, and W. E. Diewert (1982): "The economic theory of index numbers and the measurement of input, output, and productivity," Econometrica: Journal of the Econometric Society, 1393–1414.

Chanda, A. and C.-J. Dalgaard (2008): "Dual economies and international total factor productivity differences: Channelling the impact from institutions, trade, and geography," Economica, 75, 629–661.

Chandrasekhar, A. G., C. Kinnan, and H. Larreguy (2014): "Social networks as contract enforcement: Evidence from a lab experiment in the field," NBER Working Paper 20259, national Bureau of Economic Resarch Working Paper w20259.

Charness, G. and M. Rabin (2002): "Understanding social preferences with simple tests," The Quarterly Journal of Economics, 117, 817–869. 2.1, 2.4

Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery (2013): "Barriers to Household Risk Management: Evidence from India," American Economic Journal: Applied Economics, 5, 104–135.

Cole, S., X. Giné, and J. Vickery (2017): "How does Risk Management Influence Production Decisions? Evidence from a Field Experiment," The Review of Financial Studies, 30, 1935–1970.

Collier, P. and A. Hoeffler (1998): "On economic causes of civil war," Oxford economic papers, 50, 563–573.

Colmer, J. (2016): "Weather, labour reallocation, and industrial production: Evidence from India," Work. Pap., London Sch. Econ., London.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013): "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment," The Quarterly Journal of Economics, 128, 531–580.

Dasgupta, P. and D. Ray (1986): "Inequality as a determinant of malnutrition and unemployment: Theory," The Economic Journal, 1011–1034. 2.1

De Janvry, A., M. Fafchamps, and E. Sadoulet (1991): "Peasant household behaviour with missing markets: Some paradoxes explained," The Economic Journal, 1400–1417.

De Mel, S., D. McKenzie, and C. Woodruff (2008): "Returns to capital in microenterprises: Evidence from a field experiment," The Quarterly Journal of Economics, 1329–1372.

Decker, S. H., C. M. Katz, and V. J. Webb (2008): "Understanding the black box of gang organization: Implications for involvement in violent crime, drug sales, and violent victimization," Crime & Delinquency, 54, 153–172.

Dell, M. (2015): "Trafficking networks and the Mexican drug war," American Economic Review, 105, 1738–79.

Demeke, M., G. Pangrazio, M. Maetz, et al. (2009): "Country responses to the food security crisis: Nature and preliminary implications of the policies pursued," .

Dercon, S. (2002): "Income risk, coping strategies, and safety nets," The World Bank Research Observer, 17, 141–166.

Dercon, S., T. Bold, and C. Calvo (2016): Social Protection for the Poor and Poorest: Concepts, Policies and Politics, Springer, chap. Insurance for the Poor?, 47–63.

Dercon, S. and L. Christiaensen (2011): "Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia," Journal of Development Economics, 96, 159–173.

Dhillon, A., V. Iversen, and G. Torsvik (2013): "Employee referral, social proximity and worker discipline: Theory and Evidence from India," CESifo Working Paper Series No. 4309.

Di Tella, R. and E. Schargrodsky (2004): "Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack," American Economic Review, 94, 115–133.

Dickens, W. T., L. Goette, E. L. Groshen, S. Holden, J. Messina, M. E. Schweitzer, J. Turunen, and M. E. Ward (2007): "How Wages Change: Micro Evidence from the International Wage Flexibility Project," The Journal of Economic Perspectives, 21, 195–214. 2.1

Diewert, W. E. (1976): "Exact and superlative index numbers," Journal of Econometrics, 4, 115–145.

Dix-Carneiro, R., R. R. Soares, and G. Ulyssea (2017): "Economic Shocks and Crime: Evidence from the Brazilian Trade Liberalization," Tech. rep., National Bureau of Economic Research.

Drexler, A., G. Fischer, and A. Schoar (2014): "Keeping it simple: Financial literacy and rules of thumb," American Economic Journal: Applied Economics, 6, 1–31.

Drèze, J., R. Khera, et al. (2013): "Rural poverty and the Public Distribution System," Economic and Political Weekly, 47.

Dreze, J., L. Leruth, and A. Mukherjee (1986): "Rural Labour Markets in India: Theories and Evidence," in 8th World Congress of the International Economic Association, New Delhi, December. 2.1, 2.1, 2.5.2

Dube, O. and J. F. Vargas (2013): "Commodity price shocks and civil conflict: Evidence from Colombia," The Review of Economic Studies, 80, 1384–1421.

Dupas, P. and J. Robinson (2015): "The Daily Grind: Cash Needs, Labor Supply and Self-Control," Working Paper.

Eckaus, R. S. (1955): "The factor proportions problem in underdeveloped areas," The American Economic Review, 539–565. 2.1

Elbadawi, E. and N. Sambanis (2000): "Why are there so many civil wars in Africa? Understanding and preventing violent conflict," Journal of African economies, 9, 244–269.

Ellison, G. (1994): "Cooperation in the prisoner's dilemma with anonymous random matching," The Review of Economic Studies, 61, 567–588. 88, 2.1, 2.1, 92, 2.4

Emerick, K. (2016): "Agricultural productivity and the sectoral reallocation of labor in rural India," Unpublished Manuscript.

Faber, B. (2014): "Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System," Review of Economic Studies, 81, 1046–1070.

Falco, S. D., F. Adinolfi, M. Bozzola, and F. Capitanio (2014): "Crop insurance as a Strategy for Adapting to Climate Change," Journal of Agricultural Economics, 65, 485–504.

FARBER, H. S. (1986): "The analysis of union behavior," Handbook of labor economics, 2, 1039–1089. 2.1

FARBER, H. S. AND D. H. SAKS (1980): "Why Workers Want Unions: The Role of Relative Wages and Job Characteristics," Journal of Political Economy, 88, 349–369. 2.1

FEHR, E., L. GOETTE, AND C. ZEHNDER (2009): "A Behavioral Account of the Labor Market: The Role of Fairness Concerns," Annual Review of Economics, 1, 355–384. 2.1

FEIGENBERG, B., E. FIELD, AND R. PANDE (2013): "The economic returns to social interaction: Experimental evidence from microfinance," The Review of Economic Studies, 80, 1459–1483.

FINK, G., B. K. JACK, AND F. MASIYE (2014): "Seasonal credit constraints and agricultural labor supply: Evidence from Zambia," NBER Working Paper.

FOSTER, A. D. AND M. R. ROSENZWEIG (2003): "Agricultural development, industrialization and rural inequality," Unpublished Manuscript.

——— (2004): "Agricultural productivity growth, rural economic diversity, and economic reforms: India, 1970–2000," Economic Development and Cultural Change, 52, 509–542.

——— (2011): "Are Indian farms too small? Mechanization, agency costs, and farm efficiency," Working Paper.

FREEDMAN, M. AND E. G. OWENS (2016): "Your friends and neighbors: Localized economic development and criminal activity," Review of Economics and Statistics, 98, 233–253.

GEHRKE, E. (2014): "An employment guarantee as risk insurance? Assessing the effects of the NREGS on agricultural production decisions," The World Bank Economic Review.

GEORGE, P. S. (1983): "Government Intervention in Foodgrain Markets: Procurement and Public Distribution of Foodgrains in India," Indian Institute of Management, Centre for Management in Agriculture, CMA Monographs, 100.

GINÉ, X., R. TOWNSEND, AND J. VICKERY (2009): "Patterns of Rainfall Insurance Participation in Rural India," World Bank Economic Review, 1–28.

GOLDBERG, J. (2016): "Kwacha gonna do? Experimental Evidence about labor supply in rural Malawi," American Economic Journal: Applied Economics, 8, 129–49.

GOLLIN, D., D. LAGAKOS, AND M. E. WAUGH (2013): "The agricultural productivity gap," The Quarterly Journal of Economics, 129, 939–993.

GOLLIN, D., S. PARENTE, AND R. ROGERSON (2002): "The role of agriculture in development," American Economic Review, 92, 160–164.

GOULD, E. D., B. A. WEINBERG, AND D. B. MUSTARD (2002): "Crime rates and local labor market opportunities in the United States: 1979–1997," Review of Economics and statistics, 84, 45–61.

GROSSMAN, H. I. (1991): "A general equilibrium model of insurrections," The American Economic Review, 912–921.

GUITERAS, R. P. AND B. K. JACK (2014): "Incentives, selection and productivity in labor markets: Evidence from rural Malawi," NBER Working Paper.

HECKMAN, J. J. (1993): "What has been learned about labor supply in the past twenty years?" The American Economic Review, 116–121.

HERRENDORF, B., R. ROGERSON, AND K. VALENTINYI (2013): "Two Perspectives on Preferences and Structural Transformation," American Economic Review, 103, 2752–89.

HERRENDORF, B. AND T. SCHOELLMAN (2015): "Why is measured productivity so low in agriculture?" Review of Economic Dynamics, 18, 1003–1022.

HIDALGO, F. D., S. NAIDU, S. NICHTER, AND N. RICHARDSON (2010): "Economic determinants of land invasions," The Review of Economics and Statistics, 92, 505–523.

HILL, K. G., C. LUI, AND J. D. HAWKINS (2001): Early precursors of gang membership: A study of Seattle youth, US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention Washington, DC.

HORNBECK, R. AND P. KESKIN (2014): "The Historically Evolving Impact of the Ogallala Aquifer: Agricultural Adaptation to Groundwater and Drought," American Economic Journal: Applied Economics, 6, 190–219.

IKELEGBE, A. (2005): "The economy of conflict in the oil rich Niger Delta region of Nigeria," Nordic Journal of African Studies, 14, 208–234.

IMBERT, C. AND J. PAPP (2015): "Labor market effects of social programs: Evidence from India's employment guarantee," American Economic Journal: Applied Economics, 7, 233–263.

IYER, L. AND P. TOPALOVA (2014): "Poverty and crime: evidence from rainfall and trade shocks in India," Unpublished Manuscript.

JACOBY, H. G. (2016): "Food prices, wages, and welfare in rural India," Economic Inquiry, 54, 159–176.

JAYACHANDRAN, S. (2006): "Selling labor low: Wage responses to productivity shocks in developing countries," Journal of Political Economy, 114, 538–575.

JESSOE, K., D. T. MANNING, AND J. E. TAYLOR (2017): "Climate Change and Labour Allocation in Rural Mexico: Evidence from Annual Fluctuations in Weather," The Economic Journal, n/a–n/a.

JHA, SHIKHA; RAMASWAMI, B. (2012): "The Percolation of public expenditure: Food subsidies and the Poor in India and the Philippines," in India Policy Forum 2011-12, ed. by B. . P. A. Shah, Shekar; Bosworth, chap. India Policy Forum 2011-12, 95–138.

Jha, R. K. V. B. M. A. S. (2005): "Market Integration in Wholesale Rice Markets in India," ASARC Working Papers 2005-03 2005., aSARC Working Papers 2005-03 2005.

Jolejole-Foreman, M. C., M. Mallory, and K. Baylis (2013): "Impact of Wheat and Rice Export Ban on Indian Market Integration," 2013 Annual Meeting, August 4-6, 2013, Washington, D.C. 150595, Agricultural and Applied Economics Association.

Jonasson, E., M. Filipski, J. Brooks, and J. E. Taylor (2014): "Modeling the welfare impacts of agricultural policies in developing countries," Journal of Policy Modeling, 36, 63–82.

Kadafa, A. A. (2012): "Environmental impacts of oil exploration and exploitation in the Niger Delta of Nigeria," Global Journal of Science Frontier Research Environment & Earth Sciences, 12, 19–28.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986): "Fairness as a constraint on profit seeking: Entitlements in the market," The American Economic review, 728–741. 2.1, 2.1

Kala, N. (2015): "Ambiguity aversion and learning in a changing world: The potential effects of climate change from Indian agriculture," Ph.D. thesis, Ph. D. Dissertation, Yale University.

Kandori, M. (1992): "Social norms and community enforcement," The Review of Economic Studies, 59, 63–80. 88, 2.1, 92

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

Kaur, S. (2018): "Nominal Wage Rigidity in Village Labor Markets," American Economic Re. 2.1, 2.1, 2.1, 2.1, 2.1, 2.2.4, 95, 2.5.2

Kaur, S., M. Kremer, and S. Mullainathan (2010): "Self-control and the development of work arrangements," The American Economic Review Papers and Proceedings, 100, 624–628.

——— (2015): "Self-control at work," Journal of Political Economy, 123, 1227–1277.

Keynes, J. M. (1936): General theory of employment, interest and money, Atlantic Publishers & Dist.

——— (1937): "The general theory of employment," The Quarterly Journal of Economics, 209–223.

Kochar, A. (1999): "Smoothing consumption by smoothing income: Hours-of-work responses to idiosyncratic agricultural shocks in rural India," Review of Economics and Statistics, 81, 50–61.

——— (2005): "Can Targeted Food Programs Improve Nutrition? An Empirical Analysis of India's Public Distribution System," Economic Development and Cultural Change, 54, 203–235.

Kohli, Renu;Smith, M. (2002): Developing Agricultural Trade: New Roles for Government in Poor Countries, Springer, chap. Country Studies of the Changing Role of Government in Agricultural Trade: India, 33–48.

Kongsamut, P., S. Rebelo, and D. Xie (2001): "Beyond balanced growth," The Review of Economic Studies, 68, 869–882.

Kozicka, M., R. Weber, and M. Kalkuhl (2016): "Public Distribution System in India - Leakage, Self-Selection and Targeting Errors," .

LaFave, D. and D. Thomas (2013): "Height and cognition at work: Labor market performance in a low income setting," Working Paper.

Lam, D. and S. Duryea (1999): "Effects of schooling on fertility, labor supply, and investments in children, with evidence from Brazil," Journal of Human Resources, 160–192.

Lee, H. (2014): "Industrial Output Fluctuations in Developing Countries: General Equilibrium Consequences of Agricultural Productivity Shocks," Unpublished Manuscript.

Leibenstein, H. (1957): Economic backwardness and economic growth : Studies in the theory of economic development., New York: Wiley. 2.1

Lewis, W. A. (1954): "Economic Development With Unlimited Supplies of Labour," The Manchester School, 22, 139–191. 2.1

MacLeod, W. B. (2007): "Can Contract Theory Explain Social Preferences?" The American Economic Review, 97, 187–192. 88, 2.1

Mariano, M. J. M. and J. A. Giesecke (2014): "The Macroeconomic and Food Security Implications of Price Interventions in the Philippine Rice Market," Economic Modelling, 37, 350–361.

McKenzie, D., C. Theoharides, and D. Yang (2014): "Distortions in the International Migrant Labor Market: Evidence from Filipino Migration and Wage Responses to Destination Country Economic Shocks," American Economic Journal: Applied Economics, 6, 49–75.

Miguel, E. and M. K. Gugerty (2005): "Ethnic diversity, social sanctions, and public goods in Kenya," Journal of Public Economics, 89, 2325–2368.

Miguel, E. and S. Satyanath (2011): "Re-examining economic shocks and civil conflict," American Economic Journal: Applied Economics, 3, 228–32.

Miguel, E., S. Satyanath, and E. Sergenti (2004): "Economic shocks and civil conflict: An instrumental variables approach," Journal of political Economy, 112, 725–753.

Mobarak, A. M. and M. Rosenzweig (2014): "Risk, insurance and wages in general equilibrium," NBER Working Paper 19811.

Morduch, J. (1995): "Income smoothing and consumption smoothing," The Journal of Economic Perspectives, 9, 103–114.

Moscona, J. (2017): "The Impact of India's Green Revolution: An Empirical Investigation of Modern Agricultural Development," Unpublished Manuscript.

Murgai, R., M. Ali, and D. Byerlee (2001): "Productivity Growth and Sustainability in Post-Green Revolution Agriculture: The Case of the Indian and Pakistan Punjabs," The World Bank Research Observer, 16, 199–218.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1989): "Income Distribution, Market Size and Industrialization," The Quarterly Journal of Economics, 104, 537–564.

Odjuvwuederhie Emmanuel, I., O. Doulason Gordon, and A. Felicia Nkem (2006): "The effect of oil spillage on crop yield and farm income in Delta State, Nigeria," Journal of Central European Agriculture, 7, 41–48.

Office, N. S. S. (2013): "Key Indicators of Situation of Agricultural Households in India," Tech. rep., Ministry of Statistics and Programme Implementation, Government of India.

Ofuoku, A. U. and C. O. Chukwuji (2012): "The impact of rural-urban migration on plantation agriculture in the Niger Delta region, Nigeria," Journal of Rural Social Sciences, 27, 137.

Onuoha, F. C. (2009): "Why the poor pay with their lives: oil pipeline vandalisation, fires and human security in Nigeria," Disasters, 33, 369–389.

Orogun, P. S. (2010): "Resource control, revenue allocation and petroleum politics in Nigeria: the Niger Delta question," GeoJournal, 75, 459–507.

Osmani, S. R. (1990): "Wage determination in rural labour markets," Journal of Development Economics, 34, 3–23. 2.1, 2.5.2

Prothero, R. E. (1912): English Farming, Past and Present, Longmans, Green. 2.5.2

Restuccia, D. and R. Rogerson (2013): "Misallocation and productivity," Review of Economic Dynamics, 16, 1–10.

Restuccia, D., D. T. Yang, and X. Zhu (2008): "Agriculture and aggregate productivity: A quantitative cross-country analysis," Journal of Monetary Economics, 55, 234 – 250.

Rosegrant, M. W. and R. E. Evenson (1992): "Agricultural Productivity and Sources of Growth in South Asia," American Journal of Agricultural Economics, 74, 757–761.

Rosenstein-Rodan, P. N. (1956): "Disguised unemployment and under-employment in agriculture," .

Rosenzweig, M. and C. R. Udry (2013): "Forecasting Profitability," Tech. rep., National Bureau of Economic Research.

Rosenzweig, M. R. (1980): "Neoclassical theory and the optimizing peasant: An econometric analysis of market family labor supply in a developing country," The Quarterly Journal of Economics, 31–55.

——— (1988): Labor markets in low-income countries, Elsevier, vol. 1, 713–762. 2.1

Salau, A. T. (1993): "Environmental crisis and development in Nigeria," Inaugural Lecture, University of Port Harcourt, Choba, Nigeria, 35–48.

Sanchez de la Sierra, R. (2017): "On the Origin of the State: Stationary Bandits and Taxation in Eastern Congo," Unpublished Manuscript.

Santangelo, G. (2016): "Firms and farms:The impact of agricultural productivity on the local Indian economy," Unpublished Manuscript.

Schiff, M. and A. Valdés (2002): "Agriculture and the macroeconomy, with emphasis on developing countries," in Agriculture and its External Linkages, Elsevier, vol. 2 of Handbook of Agricultural Economics, chap. 27, 1421 – 1454.

Schultz, T. W. (1964): Transforming traditional agriculture., Yale University Press.

Sen, A. K. (1967): "Surplus labour in India: A critique of Schultz's statistical test," The Economic Journal, 154–161.

Shapiro, C. and J. E. Stiglitz (1984): "Equilibrium unemployment as a worker discipline device," The American Economic Review, 433–444. 2.1

Sharma, R. (2011): "Food Export Restrictions: Review of the 2007-2010 Experience and Considerations for Disciplining Restrictive Measures," FAO Commodity And Trade Policy Research Working Paper No. 32.

Singh, I., L. Squire, J. Strauss, et al. (1986): Agricultural household models: Extensions, applications, and policy., Johns Hopkins University Press.

Tarozzi, A., J. Desai, and K. Johnson (2015): "The impacts of microcredit: Evidence from Ethiopia," American Economic Journal: Applied Economics, 7, 54–89.

Taylor, J. (2008): "Involuntary unemployment," The New Palgrave Dictionary of Economics, Second Edition.

Tobin, J. (1972): "Inflation and Unemployment," The American Economic Review, 62, 1–18.

Townsend, R. M. (1994): "Risk and Insurance in Village India," Econometrica, 62, 539–591.

Tripathi, A. K. (2014): Agricultural Prices and Production in Post-reform India, Routledge.

Udry, C. (1996): "Efficiency and market structure: Testing for profit maximization in African agriculture," Working Paper.

Vollrath, D. (2009): "How important are dual economy effects for aggregate productivity?" Journal of Development Economics, 88, 325 – 334.

Watts, M. (2007): "Petro-insurgency or criminal syndicate? Conflict & violence in the Niger Delta," Review of African Political Economy, 34, 637–660.

Wischnath, G. and H. Buhaug (2014): "Rice or riots: On food production and conflict severity across India," Political Geography, 43, 6–15.

Zimmermann, L. (2012): "Labor Market Impacts of a Large-Scale Public Works Program: Evidence from the Indian Employment Guarantee Scheme," IZA Discussion Paper No. 6858.

Appendix A1. Chapter 1 Supplementary Figures & Tables

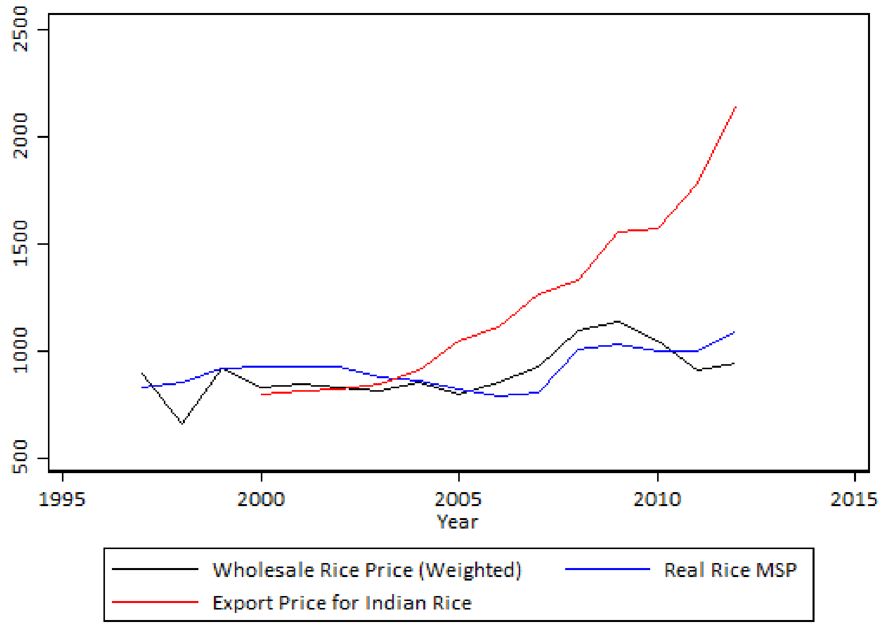


Figure A1.1: Variation over time in real MSP, weighted mean real price, and real export price of common rice from India

Back to text

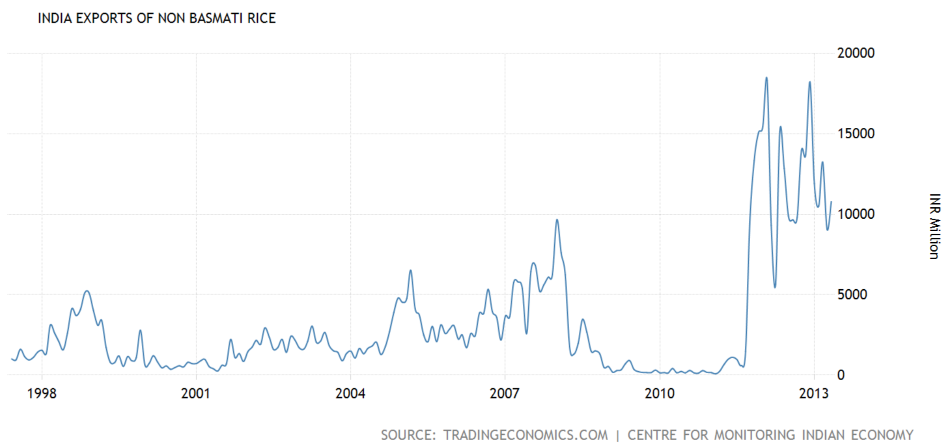


Figure A1.2: Amount of common rice exported from India
Back to text

Table A1.1: Low Vs. High Prices (Rice)

VARIABLES	(1) Full Total Area Cultivated	(2) High Total Area Cultivated	(3) Low Total Area Cultivated	(4) Full Rice Area	(5) High Rice Area	(6) Low Rice Area	(7) Full Rice Prop of Area Cultivated	(8) High Rice Prop of Area Cultivated	(9) Low Rice Prop of Area Cultivated
1.(Good Early Rain)	4,219	2,957	19,487**	-708.7	-763.9	2,354	-0.0368***	-0.0295**	-0.0963***
	(2,967)	(3,926)	(8,526)	(1,245)	(1,477)	(3,603)	(0.0109)	(0.0125)	(0.0313)
1.(High Rice MSP)*	-2,715	-4,804	-12,881	3,342**	1,237	4,663	0.0397***	0.0228	0.110***
1.(Good Early Rain)	(3,441)	(4,788)	(9,423)	(1,493)	(1,931)	(4,280)	(0.0135)	(0.0153)	(0.0409)
Observations	3,608	2,533	1,075	3,608	2,533	1,075	3,608	2,533	1,075
R-squared	0.962	0.967	0.967	0.986	0.990	0.977	0.920	0.940	0.923
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0126	-0.0215	-0.0665	0.0325	0.0117	0.0485	0.0657	0.0394	0.165

VARIABLES	(10) Full Rice Yield	(11) High Rice Yield	(12) Low Rice Yield	(13) Full Production of Rice	(14) High Production of Rice	(15) Low Production of Rice
1.(Good Early Rain)	-0.0796*	-0.0228	-0.434***	-13,893	-9,308	-38,878*
	(0.0454)	(0.0468)	(0.147)	(9,543)	(10,372)	(22,820)
1.(High Rice MSP)*	0.116**	0.0563	0.452***	14,919	6,442	44,561**
1.(Good Rain Kh)	(0.0534)	(0.0581)	(0.156)	(9,626)	(11,415)	(21,278)
Observations	3,608	2,533	1,075	3,608	2,533	1,075
R-squared	0.872	0.889	0.878	0.956	0.960	0.956
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0593	0.0274	0.266	0.0689	0.0276	0.253

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample. Yield and production specifications also include a cubic polynomial of monthly rainfall during the post-planting cultivation season.

Back to text

Table A1.2: Low Vs. High Prices (Wheat)

VARIABLES	(1) Full Total Area Cultivated	(2) High Total Area Cultivated	(3) Low Total Area Cultivated	(4) Full Wheat Area	(5) High Wheat Area	(6) Low Wheat Area	(7) Full Wheat Proportion of Area Cultivated	(8) High Wheat Proportion of Area Cultivated	(9) Low Wheat Proportion of Area Cultivated
Good Early Rain Rabi	6,690**	10,098***	-2,119	3,865***	4,597***	2,856	-0.0284**	-0.0301**	-0.0515
	(2,830)	(3,602)	(5,315)	(1,380)	(1,501)	(4,986)	(0.0111)	(0.0130)	(0.0481)
1.(High Wheat MSP)*	-6,675	-8,232	-13,104	3,990	3,570	-4,715	0.0528***	0.0482**	0.0782
1.(Good Rain Rb)	(7,943)	(9,413)	(15,426)	(2,643)	(3,099)	(5,349)	(0.0192)	(0.0211)	(0.0547)
Observations	2,598	1,857	741	2,598	1,857	741	2,598	1,857	741
R-squared	0.958	0.957	0.983	0.988	0.988	0.991	0.809	0.808	0.890
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	-0.0480	-0.0570	-0.104	0.0478	0.0438	-0.0534	0.0903	0.0851	0.124

VARIABLES	(10) Full Wheat Yield	(11) High Wheat Yield	(12) Low Wheat Yield	(13) Full Production of Wheat	(14) High Production of Wheat	(15) Low Production of Wheat
1.(Good Early Rain)	-0.0259	-0.00824	-0.0571	9,536*	8,976	36,322
	(0.0363)	(0.0412)	(0.145)	(5,304)	(5,627)	(31,003)
1.(High Wheat MSP)*	0.187***	0.150**	0.364	10,703	11,454	-52,249
1.(Good Rain Rb)	(0.0552)	(0.0622)	(0.250)	(10,695)	(12,144)	(43,061)
Observations	2,598	1,857	741	2,598	1,857	741
R-squared	0.907	0.916	0.928	0.982	0.985	0.985
Early Rainfall in Prediction	Yes	Yes	Yes	Yes	Yes	Yes
MSP in Prediction	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0883	0.0728	0.160	0.0412	0.0463	-0.181

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. ‘Low’ refers to districts in the bottom 30% of the price distribution in each given year. ‘High’ refers to the remaining 70% of districts. This analysis excludes all districts that have zero rice production in the kharif season or zero wheat production in the Rabi season for all the years in the sample,and all districts in the states of Rajasthan and Madhya Pradesh. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period.

Back to text

Table A1.3: Leave-one-out Jackknife Estimates by State (Rice), *Kharif* season

	Total Area Cultivated	Rice Area	Rice Proportion of Area Cultivated	Rice Yield	Production of Rice
1	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.737)	(1,493.106)**	(0.013)***	(0.053)**	(9,625.661)
2	-2,373.973	3,716.996	0.043	0.111	14,680.202
	(3,511.123)	(1,528.422)**	(0.014)***	(0.054)**	(9,447.069)
3	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.737)	(1,493.106)**	(0.013)***	(0.053)**	(9,625.661)
4	-2,640.916	3,604.428	0.042	0.143	17,357.793
	(3,564.549)	(1,562.396)**	(0.014)***	(0.054)***	(9,823.698)*
5	-4,420.329	2,667.692	0.043	0.090	11,893.432
	(3,608.341)	(1,546.972)*	(0.014)***	(0.055)	(9,986.738)
6	-2,667.590	3,356.243	0.039	0.116	15,013.360
	(3,441.381)	(1,494.136)**	(0.013)***	(0.053)**	(9,627.970)
7	-2,765.168	3,286.993	0.040	0.116	12,839.045
	(3,483.211)	(1,515.628)**	(0.014)***	(0.054)**	(9,386.131)
8	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.737)	(1,493.106)**	(0.013)***	(0.053)**	(9,625.661)
9	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.271)	(1,492.904)**	(0.013)***	(0.053)**	(9,624.358)
10	-4,312.272	3,441.975	0.040	0.118	15,550.353
	(3,251.550)	(1,536.286)**	(0.014)***	(0.056)**	(10,049.762)
11	-3,158.647	2,997.463	0.039	0.130	15,932.322
	(3,527.852)	(1,498.406)**	(0.014)***	(0.055)**	(9,982.898)
12	-2,866.078	3,533.896	0.054	0.080	15,479.447
	(3,466.593)	(1,510.242)**	(0.013)***	(0.051)	(9,681.397)
13	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.271)	(1,492.904)**	(0.013)***	(0.053)**	(9,624.358)
14	-4,406.079	2,023.837	0.014	0.142	17,584.005
	(3,495.002)	(1,405.444)	(0.012)	(0.055)***	(9,921.524)*
15	-3,771.047	3,246.985	0.043	0.124	14,690.060
	(3,543.703)	(1,548.162)**	(0.014)***	(0.055)**	(9,806.999)
16	-2,879.326	3,448.405	0.040	0.120	15,397.875
	(3,608.652)	(1,553.257)**	(0.014)***	(0.055)**	(10,009.681)
17	1,533.918	3,833.846	0.039	0.114	15,907.008
	(3,401.072)	(1,565.910)**	(0.014)***	(0.055)**	(10,242.516)
18	-2,283.067	3,570.960	0.043	0.115	15,190.658
	(3,476.787)	(1,553.892)**	(0.014)***	(0.055)**	(9,809.386)
19	-2,777.167	3,341.011	0.040	0.116	14,948.021
	(3,448.734)	(1,497.017)**	(0.013)***	(0.054)**	(9,661.002)
20	-2,726.673	3,346.391	0.040	0.116	14,940.083
	(3,442.149)	(1,493.469)**	(0.013)***	(0.053)**	(9,624.491)
21	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.737)	(1,493.106)**	(0.013)***	(0.053)**	(9,625.661)
22	-2,661.762	3,360.753	0.040	0.115	14,902.172
	(3,441.604)	(1,494.867)**	(0.013)***	(0.053)**	(9,628.713)
23	-2,068.204	4,344.941	0.042	0.048	9,949.660
	(3,552.382)	(1,513.693)***	(0.014)***	(0.051)	(9,688.847)
24	-2,734.469	3,363.746	0.040	0.116	14,973.620
	(3,445.911)	(1,495.201)**	(0.013)***	(0.053)**	(9,632.170)
25	-1,634.329	3,358.680	0.036	0.125	17,087.773
	(3,514.405)	(1,520.263)**	(0.014)***	(0.056)**	(9,986.824)*
26	-3,266.106	3,271.340	0.039	0.122	14,764.370
	(3,492.685)	(1,511.951)**	(0.014)***	(0.054)**	(9,862.856)
27	-2,715.026	3,342.424	0.040	0.116	14,918.688
	(3,440.737)	(1,493.106)**	(0.013)***	(0.053)**	(9,625.661)
28	-1,009.402	3,314.318	0.029	0.081	7,752.261
	(3,655.043)	(1,737.931)*	(0.011)***	(0.045)*	(7222.477)
29	-3,351.619	3,411.073	0.043	0.137	16,713.695
	(3,560.298)	(1,556.973)**	(0.014)***	(0.055)**	(10,060.946)*
30	-2,607.378	3,458.748	0.040	0.120	15,510.199
	(3,473.227)	(1,511.843)**	(0.014)***	(0.054)**	(9,749.161)
31	-203.589	5,223.755	0.046	0.214	32,410.098
	(3,947.901)	(1,721.362)***	(0.016)***	(0.073)***	(13,407.322)**
32	-2,887.329	3,309.690	0.040	0.118	14,844.629
	(3,506.642)	(1,510.259)**	(0.014)***	(0.055)**	(9,829.899)
33	-4,165.939	2,176.149	0.042	0.128	12,561.164
	(3,540.985)	(1,412.450)	(0.014)***	(0.055)**	(9,798.029)

Back to text height

Table A1.4: Leave-one-out Jackknife Estimates by State (Wheat), *Rabi* season

State Excluded	Total Area Cultivated	Wheat Area	Wheat Proportion of Area Cultivated	Wheat Yield	Production of Wheat
1	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
2	-10,159.133	352.431	0.023	0.190	21,095.094
	(7,360.220)	(2,132.011)	(0.014)*	(0.048)***	(8,048.350)***
3	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,357.308)	(2,132.308)	(0.014)*	(0.048)***	(8,049.235)***
4	-10,299.551	495.664	0.024	0.190	22,028.000
	(7,587.063)	(2,226.691)	(0.014)*	(0.049)***	(8,400.560)***
5	-12,276.823	-455.290	0.023	0.151	18,118.098
	(7,626.272)	(2,122.878)	(0.014)	(0.048)***	(8,193.088)**
6	-10,155.486	335.601	0.023	0.194	21,047.500
	(7,361.068)	(2,132.142)	(0.014)*	(0.048)***	(8,048.120)***
7	-9,727.381	664.879	0.026	0.197	22,097.885
	(7,486.715)	(2,154.998)	(0.014)*	(0.048)***	(8,184.793)***
8	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
9	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
10	-15,376.751	-1,112.667	0.014	0.106	12,845.268
	(8,350.440)*	(1,441.902)	(0.014)	(0.045)**	(5,558.498)**
11	-10,762.203	54.596	0.020	0.194	21,334.894
	(7,565.564)	(2,147.067)	(0.014)	(0.049)***	(7,957.791)***
12	-9,713.937	541.996	0.030	0.170	20,452.874
	(7,502.831)	(2,154.290)	(0.013)**	(0.048)***	(8,179.771)**
13	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
14	-10,631.872	190.622	0.020	0.202	22,307.833
	(7,587.330)	(2,191.757)	(0.012)	(0.048)***	(8,269.579)***
15	-9,875.742	511.950	0.025	0.190	21,860.378
	(7,545.507)	(2,215.730)	(0.014)*	(0.049)***	(8,397.622)***
16	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
18	2,639.960	937.144	0.020	0.220	23,376.513
	(4,099.020)	(2,513.737)	(0.016)	(0.052)***	(9,883.903)**
19	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
20	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
21	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
22	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,358.682)	(2,132.706)	(0.014)*	(0.048)***	(8,050.737)***
23	-10,986.271	317.249	0.012	0.226	21,804.949
	(7,654.842)	(2,183.843)	(0.013)	(0.047)***	(8,307.409)***
24	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
25	-10,657.993	-279.150	0.019	0.183	16,996.442
	(7,456.585)	(2,113.214)	(0.014)	(0.049)***	(7,804.578)**
27	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
28	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,360.100)	(2,133.117)	(0.014)*	(0.048)***	(8,052.288)***
29	-10,198.387	354.944	0.023	0.192	21,158.492
	(7,363.932)	(2,133.174)	(0.014)*	(0.048)***	(8,059.735)***
30	-10,162.863	354.489	0.023	0.191	21,109.362
	(7,361.473)	(2,133.515)	(0.014)*	(0.048)***	(8,053.790)***
31	-14,331.462	1,926.886	0.049	0.230	22,482.760
	(9,809.050)	(2,760.303)	(0.019)***	(0.063)***	(9,927.051)**
32	-11,165.413	439.902	0.025	0.202	23,651.840
	(7,933.434)	(2,335.142)	(0.015)*	(0.050)***	(8,683.103)***
33	-10,916.277	519.454	0.026	0.190	22,644.611
	(7,953.558)	(2,316.098)	(0.015)*	(0.052)***	(8,733.157)**

Back to text height

Table A1.5: Yield and Production, No Rain Controls, *Kharif* season

VARIABLES	(1) Yield	(2) Yield (Unconditional)	(3) Production of Rice Kh
1.(Good Early Rain)	-0.0869*	-0.0796*	-13,893
	(0.0455)	(0.0454)	(9,543)
1.(High Rice MSP)*1.(Good Rain Kh)	0.125**	0.116**	14,919
	(0.0536)	(0.0534)	(9,626)
Observations	3,561	3,608	3,608
R-squared	0.867	0.872	0.956
Early Rainfall in Prediction	Yes	Yes	Yes
MSP in Prediction	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes
Proportion Mean	0.0632	0.0593	0.0689

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season for all the years in the sample.

Table A1.6: Agricultural Production, *Kharif* season; No Predictions

VARIABLES	(1) Total Area	(2) Rice Area	(3) Rice Prop of Area Cult	(4) Rice Yield	(5) Rice Prod
1.(Good Early Rain)	-2,933	-995.3	-0.00539	-0.0959**	-18,738*
	(3,513)	(1,936)	(0.00913)	(0.0428)	(9,558)
1.(High Rice MSP)	8,212*	3,229	0.0127	0.0882*	13,817*
*1.(Good Early Rain)	(4,346)	(2,253)	(0.0121)	(0.0456)	(8,254)
Observations	3,608	3,608	3,608	3,608	3,608
R-squared	0.962	0.986	0.920	0.872	0.957
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
State x Time Trends	Yes	Yes	Yes	Yes	Yes
Proportion Mean	0.0383	0.0314	0.0210	0.0447	0.0639

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the district level. This analysis excludes all districts that have zero rice production in the kharif season or zero wheat production in the rabi season for all the years in the sample. Yield and production specifications also include cubic specifications of monthly rainfall in the post-planting period. High and low rainfall are defined with respected to the distribution of real realized wholesale prices for rice.

Appendix A2. Chapter 1 Model Appendix

A2.1 CES Equilibrium

A2.1.1 Without Price Supports

According to the framework, we have the following equations from the FOCs:

$$q_M = \frac{(1-\alpha)^\sigma I}{\alpha^\sigma p_A^{1-\sigma} + (1-\alpha)^\sigma} \tag{A2.19}$$

$$\beta_i p_i z_i (\frac{K_i}{L_i})^{\beta_i - 1} = r \tag{A2.20}$$

$$(1-\beta_i) p_i z_i (\frac{K_i}{L_i})^{\beta_i} = w \tag{A2.21}$$

$$p_A = \frac{\beta_M z_M}{\beta_A z_A} (\frac{K_M}{L_M})^{\beta_M - 1} (\frac{K - K_M}{L - L_M})^{1-\beta_A} \tag{A2.22}$$

$$y_M = q_M = \frac{E_M}{p_M} = E_M \tag{A2.23}$$

$$K = K_M + K_A \tag{A2.24}$$

$$L = L_M + L_A \tag{A2.25}$$

$$I = wL + rK \tag{A2.26}$$

First, I rewrite K_M in terms of L_M using A2.20, A2.21, A2.24, and A2.25:

$$\begin{aligned} K_M &= \frac{(1-\beta_A)\beta_M}{(1-\beta_M)\beta_A} \frac{K - K_M}{L - L_M} L_M &(27) \\ &= \frac{(1-\beta_A)\beta_M K L_M}{(1-\beta_M)\beta_A L + (\beta_M - \beta_A) L_M} &(28) \end{aligned}$$

Then, I use the market clearing condition from equation A2.23, together with A2.19, and A2.26 to express L_M in terms of p_A.

$$y_M = q_M \tag{29}$$
$$z_M K_M^{\beta_M} L_M^{1-\beta_M} = \frac{(1-\alpha)^\sigma}{\alpha^\sigma p_A^{1-\sigma} + (1-\alpha)^\sigma}(wL + rK) \tag{30}$$

I substitute in for w and r using A2.20 and A2.21:

$$z_M K_M^{\beta_M} L_M^{1-\beta_M} = [\frac{(1-\alpha)^\sigma}{\alpha^\sigma p_A^{1-\sigma} + (1-\alpha)^\sigma}]((1-\beta_M) z_M (\frac{K_M}{L_M})^{\beta_M} L + \beta_M z_M (\frac{K_M}{L_M})^{\beta_M - 1} K) \tag{31}$$

I divide through by $z_M(\frac{K_M}{L_M})^{\beta_M}$, then substitute in for K_M using 28, and rearrange, to get a implicit expression for L_M:

$$L_M[(\frac{\alpha}{1-\alpha})^\sigma p_A^{1-\sigma} + \frac{1-\beta_M}{1-\beta_A}] - \frac{1-\beta_M}{1-\beta_A} L = 0 \tag{32}$$

Finally, I use equation A2.22 and substitute in for K_M using 28, to obtain an expression for p_A in terms of L_M and constants:

$$p_A = \frac{\beta_M z_M}{\beta_A z_A}(\frac{K_M}{L_M})^{\beta_M - 1}(\frac{K - K_M}{L - L_M})^{1-\beta_A} \tag{33}$$
$$= \frac{\frac{\beta_M[\beta_A(1-\beta_M)]^{1-\beta_A}}{\beta A[\beta_M(1-\beta_A)]^{1-\beta_M}} z_M}{z_A}[\frac{K}{(1-\beta_M)\beta_A L + (\beta_M - \beta_A) L_M}]^{\beta_M - \beta_A} \tag{34}$$

2.1.2 With Price Supports

When price supports are set at the level p_S in the agricultural sector, with a consumer price for agricultural goods set at p_C, I start with a modified set of first-order conditions. I rewrite A2.22 as follows:

$$p_S = \frac{\beta_M z_M}{\beta_A z_A}(\frac{K_M}{L_M})^{\beta_M - 1}(\frac{K - K_M}{L - L_M})^{1-\beta_A} \tag{35}$$

Then, I substitute for K_M using 28, and rearrange to get:

$$p_S = \frac{\beta_M z_M}{\beta_A z_A} \frac{[(1-\beta_A)\beta_M]^{\beta_M - 1} K^{\beta_M - \beta_A}[(1-\beta_M)\beta_A]^{1-\beta_A}}{[(1-\beta_M)\beta_A L + (\beta_M - \beta_A)L_M]^{\beta_M - \beta_A}} \tag{36}$$

Rewriting in terms of L_M,

$$L_M = \frac{(\frac{[(1-\beta_A)\beta_M]^{\beta_M - 1}[(1-\beta_M)\beta_A]^{1-\beta_A} K^{\beta_M - \beta_A}}{p_S z_A})^{\frac{1}{\beta_M - \beta_A}} - (1-\beta_M)\beta_A L}{\beta_M - \beta_A} \tag{37}$$

2.2 CES Comparative Statics

Now, I derive various relevant comparative statics for the two different cases.

2.2.1 Without Price Supports

Labor in Manufacturing

I substitute 34 into 32 to get:

$$L_M[\kappa_1(\frac{\kappa_3 z_M}{z_A})^{1-\sigma}(\frac{K}{(1-\beta_M)\beta_A L + (\beta_M - \beta_A)L_M})^{(\beta_M - \beta_A)(1-\sigma)} + \kappa_2] - \kappa_2 L = 0 \tag{38}$$

$$\tag{39}$$

As long as $\sigma < 1$, an increase in z_A decreases the left-hand side of the equation. The effective size of the exponent on L_M is positive, so it is clear that L_M must increase in response to an increase in

z_A to keep the system in equilibrium.

$$\frac{\partial L_M}{\partial z_A} > 0$$

Relative Price of Agricultural Goods

From 34, and since $\frac{\partial L_M}{\partial z_A} > 0$, an increase in z_A results in a decrease in p_A.

$$\frac{\partial p_A}{\partial z_A} < 0$$

Production in Manufacturing

From A2.19, it is clear that q_M increases as p_A falls, which happens in response to an increase in z_A. Therefore,

$$\frac{\partial q_M}{\partial z_A} > 0$$

Labor in Agriculture

With a given total stock of labor, L,

$$\frac{\partial L_A}{\partial z_A} = -\frac{\partial L_M}{\partial z_A} < 0$$

Production in Agriculture

Agricultural production is influenced by two different pressures: 1. The posiive effect of increased agricultural productivity z_A, and 2. The resulting decrease in capital and labor allocated to agriculture in equilibrium.

Overall, using equations A2.19 and the budget constraint from the household problem, we have:

$$q_A = \frac{I\alpha^\sigma}{p_A^\sigma(\alpha^\sigma p_A^{(1-\sigma)} + (1-\alpha)^\sigma)} \tag{40}$$

Since we know agricultural prices fall with a positive productivity shock,

$$\frac{\partial q_A}{\partial z_A} > 0$$

However, I should note that the amount of rice produced in response to a positive productivity shock actually *falls* in my empirical analysis, due to substitution with other crops, which the model cannot capture.

2.2.2 With Price Supports

I now turn to the version of the model with price supports. The equilibrium is defined in equation 18:

$$L_M = \frac{(\frac{\kappa_4}{p_S z_A})^{\frac{1}{\beta_M - \beta_A}} - \kappa_5}{\beta_M - \beta_A}$$

Labor in Manufacturing

$$\frac{\partial L_M}{\partial z_A} = \frac{-1}{(\beta_M - \beta_A)^2}(\frac{\kappa_4}{p_S})^{\frac{1}{\beta_M - \beta_A}}(\frac{1}{z_A})^{\frac{1-\beta_M+\beta_A}{\beta_M-\beta_A}}\frac{1}{z_A^2} < 0 \tag{41}$$

Production in Manufacturing

From 28, K_M decreases when L_M falls, which happens in response to an increase in z_A. Therefore, overall manufacturing production, q_M, falls.

$$\frac{\partial q_M}{\partial z_A} < 0 \tag{42}$$

Labor in Agriculture

With a given total stock of labor, L

$$\frac{\partial L_A}{\partial z_A} = -\frac{\partial L_M}{\partial z_A} > 0$$

Production in Agriculture

With a given stock of capital and labor, and equations 41 and 28, we know that K_A and L_A both increase in response to a positive productivity shock z_A. Therefore,

$$\frac{\partial q_A}{\partial z_A} > 0 \tag{43}$$

Size of the effect relative to the level of the price support

The effect of the level of the price support on the size of the labor and production effects can be determined from equation 37.

$$\frac{\partial L_M}{\partial z_A \partial p_S} = \kappa_4^{\frac{1}{\beta_M - \beta_A}} \frac{1}{(\beta_M - \beta_A)^3} \left(\frac{1}{p_S z_A}\right)^{\frac{1-\beta_M+\beta_A}{\beta_M-\beta_A}} \frac{1}{z_A^2} \frac{1}{p_S^2} > 0 \tag{44}$$

This suggests that the negative effect of the productivity shock on labor allocated to manufacturing is larger when price supports are higher.

Appendix A3. Chapter 1 Data Appendix

A3.1 Crops by Category

Category	Crops
Other Cereals	Bajra, Barley, Jowar, Maize, Ragi, Wheat, Small Millets, Others
Pulses	Arhar/Tur, Beans, Blackgram, Cowpeas, Gram, Horsegram, Khesari, Masoor, Moth, Greengram, Urad, Other pulses
Cash Crops and Oilseeds	Arecanut, Cashewnut, Castorseed, Cotton, Coconut, Groundnut, Guarseed, Hemp, Jute, Linseed, Mesta, Nigerseed, Safflower, Sesamum, Sugarcane, Soybeans, Sunflower, Tobacco
Spices	Black pepper, Cardamom, Coriander, Dry Chillies, Dry Ginger, Garlic, Ginger, Turmeric

A3.2 Crops Used in the Productivity Calculation

Jowar, bajra, maize, barley, small millets, ragi, gram, arhar/tur, moong, masoor, urad, peas, groundnut, cotton.

Data Type	Data Source	Notes
Output		
District-level Production of Various Crops	APY data	

Data Type	Data Source	Notes
District-level crop prices	District-level averages of prices faced by households in that district for each crop in the NSS consumption/expenditure data	Rounds 60-68, Schedule 1.0. Soybean is not reported in NSS rounds 66 and 68. Cotton is not reported in NSS round 68. Revenue shares are adjusted accordingly. Where a crop's prices are not available for a particular district, I replace the missing data with the state-average price for the crop.
Inputs		
Land use in agriculture	APY data	
Quantities and price of labor	NSS Employment/Unemployment Surveys	Rounds 60-68, Schedule 10. To obtain correct labor cost share estimates, I aggregate up the NSS data using the multipliers provided.

Data Type	Data Source	Notes
Fertilizer Use (N, P, and K)	Agricultural Input Survey 2006-07	The survey does not provide season-wise input use, so, to the extent that the cost shares of other inputs are high relative to labor, the effect of the labor-use increase in high MSP years is understated, resulting in a lower-bound estimate of the fall in agricultural productivity.
Use of agricultural machinery	Agricultural Input Survey 2006-07	
Use of irrigation	Agricultural Input Survey 2006-07	
Use of high-yielding varieties	Agricultural Input Survey 2006-07	
Prices of non-labor inputs	Cost of Cultivation Survey 2006-07	Prices are aggregated to the state-level.

A3.3 Absolute and Relative Suitability

Absolute suitability for rice and wheat is simply the value of the Suitability Index for the district. Relative suitability is arguably a more important measure, since a district that is absolutely bad for staple production could still do relatively better by planting staples than by plantin g other crops (that have absolutely worse letterpaper suitability measures). It also captures the district's comparative (rather than absolute) advantage, a key determinant of potential gains from trade.

To calculate the relative suitability of crops, I weight the absolute suitability levels by the average share of the country's land area used in the production of that crop in the main growing season (to avoid placing too high a weight on the suitability of a district to grow more minor crops). I then calculate an overall index of suitability for the district:

$$CropIndex_{ds} = \sum_{j=1}^{16} AreaProp_j \, Suit_{js}$$

I then calculate the relative suitability of the staple crops by taking:

$$RelSuitRice_{ds} = \frac{SuitRice_{ds} - CropIndex_{ds}}{CropIndex_{ds}}$$

with an analogous measure for relative wheat suitability. I then run analyses separately for low- and high-suitability districts (dividing the sample by the median of the relative suitability measure).

Appendix B. Chapter 2 Supplementary Figures & Tables

Table B.1: Main Results: Sample Robustness

VARIABLES	(1) Accepted Offer	(2) Accepted Offer	(3) Accepted Offer
Public: Prevailing Wage - 10%	-0.126	-0.122	-0.0817
	(0.0820)	(0.0645)	(0.0474)
Partially Private (Employer): Prevailing Wage - 10%	0.0260	-0.0374	-0.0377
	(0.0911)	(0.0702)	(0.0493)
Fully Private: Prevailing Wage	0.0664	0.0788	0.0598
	(0.100)	(0.0754)	(0.0598)
Public: Prevailing Wage	0.136	0.0966	0.0793
	(0.102)	(0.0776)	(0.0514)
Partially Private (Employer): Prevailing Wage	0.126	0.137	0.0629
	(0.131)	(0.105)	(0.0746)
Observations	188	359	545
Sample Restriction	First HH	First Two HHs	Intended Sample
Task and Year x Month FE	✓	✓	✓
Depvar Mean (Private: Prevailing Wage - 10%)	0.158	0.173	0.213
Test: Full Private W-10% = Public W-10%	0.127	0.0611	0.0869
Test: Full Private W = Public W	0.506	0.824	0.725
Test: Full Private - Public, W-10% = Full Private - Public, W	0.139	0.171	0.170
Test: Partial Private W-10% = Public W-10%	0.0628	0.161	0.241
Test: Full Private W-10% = Full Private W	0.508	0.297	0.318

Notes: In columns 1-2, sample restricted to the first household approached in each village, and in columns 3-4, sample restricted to the first two households approached in each village. In columns 5-6, sample restricted to the intended salient households in the village, including households where no respondent was home. In these cases, we code the outcome variable "Accepted Job" as 0 (job refusal). In all specifications, the dependent variable is an indicator for whether the worker signed up for the job and showed up for work. In all columns, the omitted category is the Fully Private: Prevailing Wage - 10% treatment. Standard errors are clustered at the village level and are reported in parentheses. Observations are weighted by the number of salient individuals in each village.

Table B.2: Survey Attrition and Control Sample Composition

VARIABLES	(1) Has Endline Survey	(2) Num Control in Village
Public: Prevailing Wage - 10%	0.0342	0.407
	(0.0514)	(0.376)
Partially Private (Employer): Prevailing Wage - 10%	0.0124	0.104
	(0.0525)	(0.355)
Fully Private: Prevailing Wage	0.0383	-0.154
	(0.0525)	(0.397)
Public: Prevailing Wage	-0.0857	0.214
	(0.0662)	(0.433)
Partially Private (Employer): Prevailing Wage	0.0696	0.834
	(0.0554)	(0.486)
Observations	502	502
Task and Year x Month FE	✓	✓
Sample	Main	Main
Depvar Mean (Private: Prevailing Wage - 10%)	0.879	5.364

Notes: Column 1 reports the likelihood of successfully completing an endline survey with a salient respondent, by treatment. The outcome variable in column 2 is the number of control endline surveys conducted in the salient household's village. In all columns, the omitted category is the Fully Private: Prevailing Wage - 10% treatment. Standard errors are clustered at the village level and are reported in parentheses. Observations are weighted by the number of subjects in each village.

Table B.3: Number of Onlookers in the Public Treatments

VARIABLES	(1) Number of Onlookers	(2) Number of Onlookers
Public: Prevailing Wage - 10%	0.100	0.520
	(0.568)	(0.614)
	[0.860]	[0.400]
Observations	160	160
Task and Year x Month FE		✓
Depvar Mean (Public: Prevailing Wage)	5.294	5.294

Notes: Sample restricted to the Public Rs. 180 and Rs. 200 wage treatments. In all specifications, the dependent variable counts the number of onlookers present at the hiring. In all columns, the omitted category is the Public: Prevailing Wage treatment. Standard errors are clustered at the village level and are reported in parentheses. P-values are reported in brackets. Sample restricted to the salient households in the village.

Table B.4: Endline Reports of Village Prevailing Wage

VARIABLES	(1) 1(Agree)	(2) Difference	(3) Abs. Difference
Public: Prevailing Wage - 10%	0.0442	-1.126	-1.291
	(0.0713)	(3.246)	(3.051)
	0.536	0.729	0.673
Partially Private (Employer): Prevailing Wage - 10%	0.0333	-1.900	-1.266
	(0.0841)	(4.011)	(3.557)
	0.692	0.636	0.722
Fully Private: Prevailing Wage	0.123	-1.598	-2.557
	(0.0771)	(4.109)	(3.718)
	0.112	0.698	0.493
Public: Prevailing Wage	0.0579	2.640	-0.109
	(0.0856)	(4.505)	(4.084)
	0.499	0.559	0.979
Partially Private (Employer): Prevailing Wage	0.122	-0.675	-3.194
	(0.0918)	(6.082)	(4.805)
	0.185	0.912	0.507
Observations	431	431	431
Sample	Salient	Salient	Salient
Task and Year x Month FE	✓	✓	✓
Depvar Mean	0.800	5.650	8.875
Test: Full Private W-10% = Public W-10%	0.536	0.729	0.673
Test: Full Private W = Public W	0.399	0.369	0.561

Notes: Sample restricted to all salient households who responded to our endline survey. In column 1, the dependent variable is an indicator for whether the respondent reports the same prevailing wage at endline as the village informants reported prior to the intervention. In column 2, the dependent variable is the difference between the respondent's view of the prevailing wage and the informant's report. In column 3, the dependent variable is the absolute value of this difference. In all columns, the omitted category is the Fully Private: Prevailing Wage - 10% treatment. Standard errors are clustered at the village level and are reported in parentheses.

Appendix C. Chapter 3 Supplementary Figures & Tables

Table C.1: Yield vs. Distance from Pipeline

VARIABLES	(1) Yield	(2) Yield
Distance from Pipeline (km)	0.886	
	(1.288)	
1.(Distance < 15km)		6.092
		(4,213)
Pct Crop Loss	-0.870	-1.094
	(5.724)	(5.717)
Observations	20,443	20,443
Adjusted R-squared	0.102	0.102
Crop FE	Yes	Yes
Plot Char	Yes	Yes
Time FE	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. Yield is measured in kg per unit planted area. This analysis only contains plots that are in standardized units for yields (kg) and planted area (sq km).

Back to text

Table C.2: Results For Various Distance Categories

VARIABLES	(1) Spills in Month	(2) Non-Food Exp
1(Dist<15km)	-0.00232	4,685
	(0.0733)	(22,702)
1(Pct Loss > 50%)	-0.00462	-25,367**
	(0.0228)	(11,993)
1(Dist<15km)*1(Pct Loss>50)	0.797***	36,085***
	(0.0752)	(12,615)
1(15<Dist<=30km)	-0.0169	28,716*
	(0.0554)	(16,868)
1(15<Dist<=30km)*1(Pct Loss>50)	0.0687	-32,924***
	(0.0685)	(11,462)
1(30<Dist<=60km)	-0.00536	14,950
	(0.0448)	(13,603)
1(30<Dist<=60km)*1(Pct Loss>50)	0.00418	-22,202***
	(0.0490)	(8,168)
Observations	10,905	10,905
HH Char	Yes	Yes
Village Char	Yes	Yes
LGA FE	Yes	Yes
Time FE	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. I exclude one observation with an outlying level of expenditure.

Back to text

Table C.3: An Alternate Measure of Agricultural Productivity

VARIABLES	(1) Spills in Month
1(Dist<15km)	0.107*
	(.0611)
Village-Level Rainfall Shock	0.0593***
	(0.00847)
1.(Distance<15km)*1(Rainfall Shock)	1.618**
	(0.717)
Observations	18,135
R-squared	0.300
HH Char	Yes
Village Char	Yes
LGA FE	Yes
Time FE	Yes

*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses. Errors clustered at the village level. I leave out observations within a kilometre of the pipeline. A rainfall shock is defined as a positive or negative deviation from the village's LR average of greater than 20%.

Back to text