Money (Not) to Burn: Payments for Ecosystem Services to Reduce Crop Residue Burning

By

B. Kelsey Jack, University of California at Santa Barbara
Seema Jayachandran, Princeton University
Namrata Kala, MIT Sloan School of Management
Rohini Pande, Yale University

Griswold Center for Economic Policy Studies
Working Paper No. 307, March 2023

Money (Not) to Burn: Payments for Ecosystem Services to Reduce Crop Residue Burning*

B. Kelsey Jack†
Seema Jayachandran‡
Namrata Kala§
Rohini Pande¶

March 9, 2023

Abstract

Particulate matter significantly reduces life expectancy in India. We use a randomized controlled trial in the state of Punjab to evaluate the effectiveness of conditional cash transfers (also known as payments for ecosystem services, or PES) in reducing crop residue burning, which is a major contributor to the region's poor air quality. Credit constraints and distrust may make farmers less likely to comply with standard PES contracts, which only pay the participant after verification of compliance. We randomize paying a portion of the money upfront and unconditionally. Despite receiving a lower reward for compliance, farmers offered partial upfront payment are 8-12 percentage points more likely to comply than are farmers offered the standard contract. Burning measures based on satellite imagery indicate that PES with upfront payments significantly reduced burning, while standard PES payments were inframarginal. We also show that PES with an upfront component is a cost-effective way to improve India's air quality.

*We thank Patrick Behrer, Xavier Gine, Michael Greenstone, Rema Hanna, Chris Knittel, Michael Kremer, Jeff Pagel, Imran Rasul, Nick Ryan and many seminar audiences for helpful comments. We also thank Marshall Burke, Ben Moscona and especially Kendra Walker for guidance and support with the remote sensing data; Juliana Ariza Sanchez, Nils Enevoldsen, Valentina Goetz, Rohit James Joseph, Sahiba Lal, Aditya Madhusudan, and Ayushi Srivastava for excellent research assistance; and J-PAL South Asia for field support. We thank J-PAL and CEGA's Agricultural Technology Adoption Initiative (ATAI), J-PAL South Asia at IFMR's Cash Transfers for Child Health Initiative (CaTCH), and Jody and John Arnhold for funding. This RCT was registered in the American Economic Association Registry under number AEARCTR-0004508.

†UC Santa Barbara
‡Princeton University
§MIT Sloan School of Management
¶Yale University

1 Introduction

Poor air quality is a leading preventable cause of death and morbidity worldwide (Fuller et al., 2022). In North India, air pollution reduces life expectancy among the region's half a billion residents by up to 9 years, one of the largest health burdens from pollution in the world (Lee and Greenstone, 2021). The use of fires to clear agricultural land is a major source of particulate matter in India, as in other low- and middle-income countries. Despite a clear economic case for reducing this pollution, as well as efforts to prohibit and fine those who produce it, agricultural pollution in North India has increased over the last few decades (Appendix Figure A.1). Arguably, existing policies have failed to account for the incentives of two groups of actors: local officials' incentives to enforce penalties when the costs and benefits of polluting activities are in different political jurisdictions, and farmers' incentives to protect the environment given that the costs of pollution are largely borne by others.[1] In this paper, we ask whether a policy that explicitly considers these incentives can reduce pollution.

Every winter, farmers in North India burn rice stalk (residue) to clear fields. Smoke blankets the region and spreads downwind to major population centers such as New Delhi, affecting the health of millions of people (Cusworth et al., 2018).[2] A recent study estimates that crop residue burning caused 86,000 premature deaths in India in 2018, with over half of the deaths due to paddy (i.e., rice) residue burning in the state of Punjab (Lan et al., 2022). Roughly 4 million acres of paddy residue were burned in Punjab in 2018.[3] Using \$700,000 as the value of a statistical life (Majumder et al., 2018), the average mortality costs per acre of burning are \$8,000.[4] As a comparison, paddy production generates about \$500 per acre in

[1] Dipoppa and Gulzar (2022) provide evidence of this inter-jurisdictional externality, which undermines state and local actors' incentives to enforce fines and bans. Farmers' political clout and equity concerns around penalizing small farmers undermine enforcement even further.

[2] Beyond India, pollution from crop residue burning contributes to the pollution burden in China (Chen et al., 2017), Southeast Asia (Oanh et al., 2018), and Africa (Cassou, 2018).

[3] There were 6 million acres of Kharif season non-basmati paddy production in Punjab in 2018, according to Agricultural and Processed Food Products Export Development Authority (2018) (using official Government of India statistics) and the proportion burned was 66% for Punjab in 2018, according to Kumar et al. (2019).

[4] Throughout, we use 2018 real values and an exchange rate of ₹79 to \$1. Besides the mortality costs we estimate in this calculation, there are also climate change, labor productivity, and morbidity costs. Murali et al. (2010), estimates that residue burning in India emits around 4 million metric tons of carbon dioxide equivalents each year, or one percent of annual global greenhouse gas emissions (NASA, 2017; Jain et al.,

revenue, and, of course, even less in profits.[5] This suggests that eliminating residue burning would almost certainly increase societal welfare.

Eliminating crop residue burning does not require an end to rice production. There are ways to remove and manage residue without burning, and the government subsidizes certain farm equipment that can accomplish this. In addition, since 2015, farmers in North India have faced bans on burning residue. These policies, however, have been ineffective. In this paper, we investigate the feasibility of Payments for Ecosystem Services (PES) contracts, which pay farmers for not burning crop residue. By conditioning cash transfers on avoiding this environmentally harmful behavior, PES programs raise the private cost of environmental degradation for farmers. Functioning as a carrot rather than a stick, PES avoids the political unpalatability of enforcing fines when the benefits of doing so mainly accrue to other constituencies. In addition, compared to equipment (input) subsidies, PES offers more flexibility to farmers by rewarding them for the socially desired action regardless of how they achieve it.

However, contextual and institutional factors may limit the efficacy of PES — and, more broadly, conditional cash transfers — in low- and middle-income countries. PES participants must undertake a costly action to comply with the contract and receive payment. Farmers may not comply if they do not trust that the conditional payment will be made, or if they lack the cash on hand to pay for alternatives to burning before receiving the PES payment, limiting PES efficacy.

PES contracts that offer partial payment in advance may help with trust and liquidity. An upfront payment can send a costly signal, increasing trust that the subsequent conditional payment will occur. It can also alleviate liquidity constraints when farmers need to spend money on alternatives to burning. However, recouping the upfront payment if the participant fails to comply is often infeasible or undesirable in low-income settings. Practically, upfront payments must then be unconditional, potentially undermining their usefulness for at least

2014; Sarkar et al., 2018; Shindell et al., 2012).

[5] District-level data from ICRISAT show that in 2017, rice yields in Punjab were on average 17 quintals per acre, with a maximum yield of 20.4 quintals per acre (Rao et al., 2012). At a minimum support price of ₹2,000 per acre (Sharma, 2022), this implies a revenue of ₹35,000-41,000 per acre ($440 to 518). In our sample in Punjab, farmers reported profits of around $300 per acre. For consumer welfare to outweigh the mortality costs of burning, willingness-to-pay for rice would have to exceed 15 times the market price.

two reasons. First, for a given total payment, offering some of it upfront and unconditionally lowers farmers' marginal incentive to comply because the conditional payment is smaller. Second, given a certain level of effectiveness, upfront payments reduce cost-effectiveness due to payments to non-compliant farmers. Hence, the net effect of upfront and unconditional payments on compliance and cost-effectiveness is ambiguous.

Motivated by these observations, we conducted a randomized controlled trial in 171 Punjabi villages during the 2019 rice growing season to compare the efficacy of standard PES and partial upfront PES. We compare three farmer groups: those who were not offered a contract (control), those who were offered a contract with payment conditional on verification that the farmer did not burn (standard PES), and those who were offered a contract with a partial upfront payment that was explicitly unconditional on compliance, with the remainder conditionally paid after verification (upfront PES).[6]

Contract take-up was comparable and relatively high across treatment arms (72%). Our main finding is that, despite lower conditional payments, upfront PES led to 10 percentage points higher contract compliance than standard PES; this represents a doubling of the compliance rate. Remote sensing estimates of burning are consistent with the contract compliance results, showing a roughly 10 percentage point lower rate of burning among farmers offered upfront PES versus standard PES. The remote sensing measure also reveals that standard PES had no effect on burning when compared to the control group. This indicates that standard PES payments were inframarginal, i.e., paid to farmers who would not have burned even without PES. The upfront PES effect size corresponds to a 50-80% higher rate of not-burning than in the standard PES arm or control group. Consistent with the remote sensing results, in our endline survey, farmers in the upfront treatment arm were 9.5 percentage points more likely than those in the control group to report using balers (equipment used to bundle residue that has been removed from the field). Notably, balers are not among the "crop residue management" (CRM) farm equipment that is subsidized by the government. Meanwhile, farmers in the standard PES treatment arm reported no increase in the use of balers or other CRM equipment.

[6]Our design has subtreatments that vary specifics of the standard or upfront PES contract, but our primary empirical specification pools across subtreatments, as described in our pre-analysis plan.

Why did partial upfront contracts outperform standard PES contracts? An analysis of heterogeneous treatment effects using (pre-specified) baseline measures of generalized liquidity constraints and distrust is uninformative. Recognizing that generalized measures are imperfect proxies for specific PES-related beliefs and constraints, we also examine farmer responses to endline survey questions about the role of cash constraints and trust in determining their PES program response. Farmers assigned to the upfront PES treatment have 6.8% higher trust that contract payments will be made than those assigned to standard PES.[7] Around 70% of farmers say cash on hand affected their CRM decisions, suggesting that this was an important overall constraint, but responses did not differ by treatment.

To compare PES costs to the benefits of reduced residue burning, we first calculate the cost per additional acre not burned for the two treatments using our remote sensing based outcome. Standard PES has no statistically significant impact on burning, and, reflecting this, we observe noisily estimated positive costs that are sensitive to measurement choices. Upfront PES, on the other hand, reduced burning, and the cost per additional acre not burned under this arm ranges from ₹2,700 to ₹4,050 (or \$34 to \$51). Despite the fact that a substantial portion of the farmers paid upfront burned anyway, and that some compliers were inframarginal, the estimated cost of the program is drastically lower than our rough per-acre mortality benefit estimates (\$8,000). Alternatively, we can compare PES for crop residue burning to pollution abatement opportunities in other sectors, specifically India's electricity sector. Again, the difference is striking: the cost per life saved from installing scrubbers in coal-fired power plants is roughly \$400,000 (Cropper et al., 2019) versus \$3,000 to \$4,400 for upfront PES.

The paper is related to several literatures. First, multiple papers highlight the importance of contract design for agricultural outcomes, such as microfinance and crop insurance (Carter et al., 2017). Salient design features include timing of when the insurance or loan contract is offered (Burke et al., 2019; Fink et al., 2020; Casaburi and Willis, 2018), and how different contract features are bundled (Giné and Yang, 2009). We highlight the importance of contract design in environmental programs.

[7] If PES contracts were implemented as policy, repeated implementation might eventually address distrust in payments, making standard PES more cost effective over time.

Second, conditional cash transfers are widely used to incentivize specific outcomes, and recent research tests the importance of conditionality for take up and outcomes (Baird et al., 2011; Akresh et al., 2013; Attanasio et al., 2015; Aker and Jack, 2021). Instead of focusing on payment conditionality, we focus on payment timing and demonstrate that partial upfront payment during enrollment can improve compliance with a conditional contract.[8]

Third, while PES is widely used and studied, causal estimates of its impact, particularly comparisons of alternative PES program designs, are scarce. Exceptions include Jayachandran et al. (2017), who find that a PES program successfully reduced deforestation in Uganda, and Jack (2013) and Oliva et al. (2020), who show that program design affects the cost effectiveness of afforestation contracts. While PES programs have been used to address environmental externalities associated with land use, such as deforestation, they have rarely been applied to particulate matter, which is one of the most important environmental challenges worldwide. Important exceptions include Edwards et al. (2020), who find no impact of a bundled intervention of community-level training including payments to village governments for forest fire prevention, and Kramer and Ceballos (2018) who find that conditioning agricultural insurance payouts on not-burning in India lowers burning, but also insurance coverage. We contribute the first evaluation of the impact of alternative PES contract structures on efficacy and cost-effectiveness in crop residue burning.

Finally, a growing literature investigates the causes and consequences of crop residue burning (Behrer, 2019; Garg et al., 2021; Pullabhotla, 2018; Graff Zivin et al., 2020; Rangel and Vogl, 2019; He et al., 2020; Dipoppa and Gulzar, 2022; Nian, 2023). A smaller number of papers test the efficacy of burning-reduction policies; for instance, Edwards et al. (2020) and Kramer and Ceballos (2018) both vary incentives for burning in the context of bundled interventions. In contrast, we focus on contract design. We leverage novel data from a new remote sensing model to measure crop residue burning (see Walker et al. (2022) for further detail), which is better suited to detecting plot-level changes in burning than existing measures of active fires, burn scars or smoke (Balboni et al., 2021; Behrer, 2019; Edwards et al., 2020; Rangel and Vogl, 2019; Jayachandran, 2009).[9]

[8]Evidence from other settings supports the notion that the timing of transfers can affect program outcomes, because of liquidity (Coffman et al., 2019) or other channels.

[9]Existing remote sensing measures of fires have historically relied on high frequency but low spatial

2 Background and Study Design

2.1 Crop residue burning and policy responses

About 80% of the planted area of Punjab (known as the“granary of India”) is cultivated using an annual paddy rice-wheat dual crop system. Rice is farmed during Kharif, from June to October, and wheat during Rabi, from November to April. The introduction of mechanized harvesting in the 1980s resulted in widespread adoption of the dual crop system, but also created a need to manage waste on the field: mechanized rice harvesting leaves 8 to 12 inches of stalk, representing over 2.5 tons of residue per acre of paddy rice (Jain et al., 2014). In a 2019 listing survey with farming cooperative members in our two study districts, *all* farmers reported using a combination of mechanized harvesting techniques that necessitate residue removal.

Controlled burning has gradually emerged as the primary method used by Punjabi farmers to manage crop residue. Using satellite imagery from 2000 to 2018, Appendix Figure A.1 shows a rising incidence of fires across Punjab, as well as in our two study districts of Bathinda and Faridkot (Distributed Active Archive Center, 2018). Burning is prevalent despite reasonable levels of knowledge of CRM; in our baseline survey, nearly 60% of farmers reported awareness of and some use of CRM alternatives to burning. While cost is the primary reason for farmers' preference for burning, they also reported that renting CRM equipment from a cooperative or custom hiring center would take around a week during peak season, potentially delaying Rabi sowing. Furthermore, the majority of farmers who were familiar with CRM alternatives expressed concern about the perceived negative effects on agricultural productivity.[10]

Recognizing the negative pollution externalities from residue burning, the Indian Penal Code makes residue burning a punishable offense (Section 188). A 2015 court judgement

resolution sensors, such as MODIS (e.g., Vadrevu et al. (2011); Liu et al. (2019)). Newer sensors offer higher spatial resolution but less frequent observation, which potentially miss some crop residue fires due to their short visual signature (e.g., Badarinath et al. (2006); Singh et al. (2021)). Planet Labs high-resolution temporal and spatial imagery helps us overcome this tradeoff. See Walker et al. (2022) for additional details.

[10] Agronomic evidence suggests that CRM may indeed lower productivity in the short run by immobilizing available nitrogen in the soil. In the long run, however, incorporating residue into the soil may increase organic content and available nutrients, improving productivity (Sarkar et al., 2018).

banned residue burning and directed North Indian state governments to fine farmers who did so, with fines ranging from ₹2,500 to ₹15,000 depending on the farmers' landholdings (Bhuvaneshwari et al., 2019; National Green Tribunal, 2015).[11] Penalty-based policies have had limited impact due to agricultural lobby opposition and insufficient enforcement. A second set of policies subsidizes *in-situ* CRM alternatives (*in-situ* techniques either pull up and mulch the stalk or leave it standing). In 2017, the central government announced a two-year $144.3 million program for Punjab, Haryana, and Uttar Pradesh to subsidize equipment such as a new seeding machine known as the Happy Seeder, which sows directly through the paddy residue, and the Super Straw Management System, which chops residue. The subsidy program did not cover *ex-situ* equipment that removes the residue from the field, such as using equipment (or labor) to pull up and gather the straw, then using another piece of equipment (a baler) to bundle it into bales for industrial use.[12]

Even with the government subsidy scheme in place, the cost of CRM equipment rental remains high for farmers. Farmers typically rent the equipment from hiring centers or agricultural cooperatives, which qualify for an 80% discount on equipment purchases.[13] According to our baseline survey data, the median rental cost for a Happy Seeder was ₹1,250 per acre, and the total cost of *in-situ* residue management was about ₹3,000 per acre. Farmers may still prefer unsubsidized *ex-situ* methods because they are cheaper (e.g., renting a baler is around ₹1,000 per acre, even without a subsidy in place) or because they are perceived as less damaging to yields than *in-situ* methods.

To summarize, the policies that have been implemented to address residue burning have been largely ineffective to date. This motivates our focus on a more flexible alternative that aligns incentives: subsidizing not-burning through payments for ecosystem services (PES).

[11]The National Green Tribunal ruling applied to Rajasthan, Uttar Pradesh, Haryana, and Punjab.

[12]Baled straw is often used as an industrial heat source, displacing fossil fuel combustion. Nian (2023) provides evidence from China that the emergence of biomass power plants reduced agricultural burning in the immediate vicinity.

[13]Individuals can purchase equipment at a 50% subsidy, after which a Happy Seeder cost about ₹40,000 ($500), but few small- and medium-sized farmers own them.

2.2 PES contracts

We experimentally compare the effectiveness of alternative PES contract designs for reducing crop residue burning to a status quo control group. The treatment arms are summarized in Figure 1.

Our first treatment arm implements a standard PES contract, with payments contingent on not burning and paid after compliance is verified. (We discuss how contracts are monitored and enforced later in this section.) Based on discussions with the Punjab government and their view of what payment level was scalable, our base contract sets payment at ₹800 per acre. To change behavior, the payment needs to cover the marginal cost of CRM relative to burning, not the full cost of CRM. Burning may take time or require some material inputs, and some farmers may perceive it as detrimental to soil fertility. The variability in whether farmers burned in the status quo (8-20% did not burn in our control group, according to our remote sensing estimates) indicates heterogeneity in farmers' reservation price to switch away from burning; some farmers have a negative reservation price, but the majority have a positive reservation price, which likely varies considerably. While one would not expect an ₹800 per acre incentive to compel all farmers to avoid burning, the subset with a positive but modest cost to avoid burning should be on the margin of changing their behavior. To assess the importance of payment levels, we introduce a variant on the standard PES contract that pays twice as much as the base contract, or ₹1600 per acre not burned.[14]

The standard PES contract pays out only after the desired behavior has occurred (i.e., not burning). Farmers may not undertake a costly action in anticipation of a future payment if they do not trust the principal to follow through with the payment or if they lack cash on hand to carry out the action required for compliance. Both distrust and liquidity constraints are plausible barriers to compliance in our study setting. At baseline, only 13% of farmers said they had complete trust in the government, with even lower levels of trust in NGOs, about 7%. Less than half of the sample had ₹5,000 in savings, and the majority said that

[14]In November 2019 (after our intervention), the federal court ordered the Punjab government to pay farmers ₹100 per quintal of paddy, or about ₹2,500 per acre, conditional on (self-reported) not burning. Concerns regarding farmer's self-declarations that they did not burn led to the program's suspension a few weeks later. In our endline survey, about 30% of the sample was aware of the government program, and most of them reported learning about the program after having begun residue management. 80% of those who were aware of the program did not anticipate timely payments.

it would be somewhat difficult or difficult to access a ₹5,000 loan. (A typical farmer in our sample has 5 acres of paddy production, so renting a baler would cost about ₹5,000.) Our second treatment arm was designed to address these constraints by offering partial payment upfront, with the remainder paid conditional on contract outcomes. The sum of the (potential) upfront and ex post payments was held fixed at ₹800 to match our base contract. We evaluate two variants of the upfront PES treatment, which vary the share paid upfront (25% or 50%).

While upfront payments could, in principle, be recouped from a participant who does not comply, imposing such conditionality is challenging in practice (e.g., because participants are poor). Our contract explicitly made the upfront payment unconditional. Because the upfront component reduces the conditional payment amount — the farmer's incentive to comply — this feature could reduce compliance. In addition, because unconditional payments will be made to some farmers who then do not comply, the upfront payment could reduce PES cost-effectiveness, even if it increases compliance.

Farmers' participation in all treatments was voluntary, and non-compliance with the PES contract was only 'penalized' through non-payment. There are two reasons that payments made to farmers are not a good proxy for the program's impacts. First, some people who complied with the contract might have undertaken the desired activity even in the absence of the contract, i.e., the payments are inframarginal to their non-burning. Second, because there is no penalty for non-compliance, many farmers may enroll and later renege, increasing implementation costs particularly when some payment is offered upfront. Our study design allows us to evaluate these concerns by measuring "additionality" (i.e., whether contracts reduce burning relative to the status quo) as well as evaluating the cost-effectiveness of alternative PES contracts.

2.3 Sample

We chose Bathinda and Faridkot as our two study districts in Punjab because both have high rates of burning and little activity to encourage CRM adoption by other organizations.[15] We

[15]According to the Visible Infrared Imaging Radiometer Suite (VIIRS) satellite data (Distributed Active Archive Center, 2018), Faridkot had about 1.9 fires per square km in 2018, ranking sixth highest in the

limit our sample frame within these districts to villages with functioning farmer cooperatives, and within these villages, to farmers who were cooperative members (as of August 2019). From among the 393 villages in our sampling frame, we chose the 300 with the most members for screening.

Initial screening of farmers for eligibility was done by phone in fall 2019. A farmer was eligible if he or she grew between 2 and 12 acres of paddy, planned to harvest after the second week in October and plant the Rabi crop, and employed farm equipment indicative of burning (namely, used a reaper or did not use a chopper) in the previous year. The last criterion means that our sample has higher burning than average; by minimizing inframarginal payments through sample design, we increased statistical power to detect changes. We chose 176 villages with at least 18 eligible households for an in-person baseline survey.[16] Villages with fewer than six completed baseline surveys were eliminated, resulting in a final sample of 171 villages and 1,668 respondents.

To assess the representativeness of our sample, in early 2020 we conducted a census in four study villages. 70% of farmers reported being cooperative members, which was one inclusion criterion for the study sample. Appendix Table A.1 presents summary statistics for paddy farmers from the census (column 1), for cooperative members from the census (column 2), for study-eligible farmers (column 3), and for farmers enrolled in the study (column 4). We also report tests of whether cooperative members are different from census farmers. These two groups seems largely similar across a range of variables, including agricultural experience, area cultivated, knowledge of and prior experience with CRM techniques, distrust, and awareness of and application to government PES programs.[17] Second, we report tests of whether cooperative farmers differ from farmers enrolled in the study.[18] Since we only enrolled farmers cultivating 12 acres or less (a criterion that 80% of census farmers met),

state, while Bathinda had 1.7 fires per sq km, ranking ninth. We present total fires from this data source in Appendix Figure A.1.

[16]The baseline prioritized villages with the largest number of eligible households based on listing data. Baseline data collection stopped once the target sample of 176 villages was reached. Within a village, enumerators moved down a randomly ordered eligible list until 16 surveys were completed or the list was exhausted.

[17]We discuss measurement of these variables in detail in Section 3.3.

[18]Note that we only have 38 observations for the number of farmers enrolled in the study, since this table only includes farmers from the four census villages.

study farmers cultivated less land than census farmers (5.3 acres versus 7.7 acres), but are otherwise similar across the broad range of characteristics we measured.

2.4 Randomization and contract implementation

We randomly assigned villages to one of four treatment arms (two standard PES subtreatments and two upfront PES subtreatments) or the control group. Figure 1 summarizes the sample size per arm. Randomization occurred while the listing and baseline surveys were ongoing.[19] Eligible farmers in treatment villages received a follow-up visit and were offered a PES contract, on average 5.6 days after their baseline survey. If they could not be located or did not have a bank account, they were not offered the contract.[20]

Farmers received a basic program description as part of the contract offer (see Appendix A.2). The contract (offered by J-PAL) corresponding to their treatment arm was then shown to interested farmers, and the terms were read aloud by the enumerator.[21] Farmers who agreed to the contract terms received an information handout outlining the terms and conditions, as well as procedures for monitoring and verifying burning outcomes. During the enrollment visit, enumerators recorded whether the farmer was offered the PES contract and, if not, why (e.g., the household could not be found or did not have a bank account), and whether the farmer took up the contract.[22] A farmer's entire paddy acreage (as physically measured by the surveyors at baseline) was automatically enrolled if the farmer took up the contract. The maximum amount a farmer could be paid under the contract was ₹16,000 for farmers in the ₹1,600 per acre arm and ₹8,000 in the other treatment arms.[23] Farmers in the upfront PES arm received the upfront portion of their payment via direct bank deposit 2-3 days after enrollment.

[19] Initial assignment was stratified on district, below versus above median number of eligible households based on phone listing survey, baseline survey completion in the village, and listing survey completion in the village. We include a fifth strata of 15 villages that were randomly added to the different arms to supplement the sample size. These were villages where the listing survey included between 6 and 17 households. The order of listing was randomized, and the baseline and treatment implementation followed this order. Baseline surveyors did not know a village's treatment status.

[20] Contract payments were electronic, with bank account information only collected from treatment farmers.

[21] A sample contract for the ₹800 per acre arm with no upfront payments is included in Appendix A.3.

[22] Only 6 farmers were screened out because they did not have a bank account.

[23] Farmers eligible for the study cultivated 12 or fewer acres of paddy. For farmers with more than 10 acres of paddy, about 7% of the sample, the value of the contract per acre is constrained by the cap.

Contract monitoring and enforcement required verifying that none of the paddy plots were burned. To detect whether burning had occurred, project enumerators had to visit the plot during the short window between when the farmer removed the residue, whether by CRM or burning, and when the farmer tilled the soil for sowing the Rabi crop. Prior to the completion of residue management, a monitoring visit could not rule out future burning. After tilling, visual signs of burning become much less obvious. Because monitoring was required during this farmer-specific window of a few days, the farmer was responsible for contacting J-PAL after managing residue and at least four days before sowing the Rabi crop. Up to two monitoring visits could be requested if, for example, different plots had different sowing schedules. When farmers called in, appointments for monitoring were made, and the map of all study plots collected at baseline was given to the monitor.[24]

Visual inspection at the right moment can reveal quite clearly if a field was burned. Enumerators were required by the monitoring protocol to observe and record multiple observations on each plot, including burned straw or residue, grey or black ash on the soil, burned root residue, burned grass or weeds on the plot boundaries, and burned leaves or tree branches on the plot boundaries. They were told to inspect the plot perimeter and walk onto each plot and inspect the soil. Farmers were not immediately informed of the monitoring results after it was completed. Monitoring staff did not know specifically how the different data they recorded would be aggregated to determine contract compliance. The observations on different fields were combined into a single farmer-level compliance metric — any burning was a contract violation. Payments to farmers who had complied were made on a rolling basis, approximately 2-3 days after monitoring.

3 Data

Figure 2 presents the agricultural season for the paddy-wheat cropping cycle in Punjab, as well as the data collection timeline. Our data collection is focused on measuring impacts on

[24]Placing the onus for monitoring on the farmer may lower compliance. However, it increases monitoring accuracy since the precise timing for verifying absence of burning varies from mid-October to late November. Farmers who had not yet requested monitoring were contacted in the third week of October with a reminder that it was the farmer's responsibility to request monitoring.

farmer behavior. Because only a small proportion of farmers in each village participated in our study (on average, 9.8 farmers per village), we do not seek to detect treatment effects on village-level outcomes (such as air quality).

3.1 Survey and contract compliance data

Baseline survey We conducted baseline surveys with the identified person on the cooperative member list in October 2019. This person was in charge of household agricultural decisions in more than 90% of households. We collected data on demographics, agricultural production, income and credit constraints, trust in organizations, and barriers to CRM. Farmers listed all plots where paddy was grown, and accompanied the enumerator to each plot to collect geocoded perimeter measurements. The plot boundaries form the basis for monitoring treatment farmers' contract compliance and for linking satellite imagery to farmers for both the treatment and control groups.[25]

Endline survey We conducted a phone-based endline survey in June 2020, following the completion of the Rabi harvest.[26] The survey gathered information about self-reported field burning and CRM techniques, as well as perceptions about limited cash on hand as a factor in their CRM decision and trust in the payment (for the treatment groups only). We also gathered information on agricultural production and income over the previous two seasons (Kharif 2019 and Rabi 2020). Section 4.1 reports attrition between the baseline and endline surveys and balance tests.

Contract take-up and monitoring data At the time of contract offer, we record farmer enrollment decisions and non-enrollment reasons. We observe monitoring results for those who did request it, including the distinct plot-level signs of burning described above. These were combined to produce a binary burning outcome: A farmer complied with the contract

[25]At each plot, the enumerator walked the perimeter of the field using a mapping application on a GPS-enabled tablet. We drop 47 fields (1.6 percent) where either the field ID was missing from the geospatial data or two fields completely overlapped. Because the plot was measured before treatment was assigned, any measurement error in plot inventories or field perimeters should be orthogonal to treatment.

[26]The contract rollout and monitoring were completed prior to the COVID-19 pandemic. However, the in-person endline survey, scheduled for April 2020, had to be switched to a phone-based survey.

if none of his or her plots show signs of burning. A total of 187 farmers and 319 plots were monitored.

Spot check data We conducted spot checks on one randomly selected field for 50% of farmers in each sample village. The purpose of these data was to serve as training inputs to the remote sensing model described below. The spot check protocol was similar to that for monitoring with a few exceptions. Most importantly, unlike the monitoring visits that occurred when the farmer informed us they had removed their residue but not yet tilled the soil, spot checks could not be synchronized with the farmer's specific harvest and CRM timing.[27] Overall, we conducted 720 spot checks in November 2019.

3.2 Remote sensing measures

Satellite data The data on take-up and monitoring are informative of how many farmers signed up for the contract and were verified as not burning through monitoring, but not the rate of burning among non-monitored farmers. Some farmers in the control group may not have burned, and some farmers in the treatment group who did not enroll or did not call for monitoring may also not have burned. Thus, our analysis also uses a measure of burning constructed from high-resolution satellite imagery (see Appendix A.4 and Walker et al. (2022) for additional detail on the data, data processing and machine learning model). We use data from two sources, PlanetScope and Sentinel-2.[28] PlanetScope data is higher frequency (roughly every 2-3 days) and is less likely to miss a burning event. This feature is important because burned fields become observationally similar to unburned fields once the soil is tilled for Rabi planting. Appendix Figure A.2 depicts this in Sentinel-2 imagery for one study field. The absence of a persistent visual signature of burning is particularly problematic in PlanetScope's visual and near-infrared bands. Sentinel-2 data are collected at

[27] Other differences were that only one randomly selected plot was visited per farmer and, because the farmer did not accompany the enumerator, all observation occurred outside of the field. We do not use spot check data as a primary outcome because to have sufficient statistical power, we would have had to conduct repeated visits to address the challenge of observing the field after harvest but before tilling. Repeated monitoring might have directly affected behavior.

[28] PlanetScope has a 3 meter resolution and Sentinel 20 meter. Other commonly used sensors such as MODIS and VIIRS are available at lower resolution only, 375 meters to 1 kilometer, and therefore inappropriate for the small fields we study.

a lower frequency (every week to 10 days) but, unlike PlanetScope, they provide information in the mid-infrared range, which helps separate burned from unburned fields. Thus, the two sensors complement one another.

Data processing and classification We train a random forest (RF) model with labels (burn or no-burn) from the monitoring and spot check data, using pixel-level data from the two sensors. Our labeled set of 681 fields includes burn labels (positives) from the spot check and monitoring data and no-burn labels (negatives) only from the monitoring data.[29] The RF model outputs a pixel-level continuous prediction score ranging from 0 to 1, which represents the proportion of decision trees that the model classified as burned. We avoid over-fitting by holding out each field (consisting of many pixels) from the training set, obtaining a prediction for each held-out pixel. To aggregate pixel-level data to a field-level burning outcome, we average the predicted score across pixels in a field (omitting perimeter pixels) and then choose a threshold above which a field is classified as burned. To generate binary burn estimates for the entire data set, we apply the mean of the trained RF models (recall we have one trained model for each of the fields in the training set, given the procedure to avoid over-fitting) to unlabeled pixels, aggregate to the field-level, and apply the classification thresholds as described in the next paragraph. For consistency with other contract outcomes (a higher value is an environmental improvement), we invert the burning classification when estimating treatment effects. Hence, this farmer-level outcome equals zero if a farmer is predicted to have burned any of his fields.

We present two alternative thresholds: (1) *balanced accuracy* equalizes prediction accuracy for burn and no-burn labels (training data), or equivalently equalizes type I and type II errors or sensitivity and specificity, and (2) *max accuracy* maximizes overall model accuracy. Because burning is more common in our data than not-burning, the max accuracy measure improves overall accuracy at the expense of accuracy in predicting not-burned, i.e., by relaxing the threshold to classify fields as burned. This, however, also increases the likelihood of false positives (cases of non-burning that are misclassified as burned). The balanced accuracy measure avoids this problem by balancing the rate of false positives and false negatives,

[29] We exclude negative labels collected during spot checks because they indicate no burning in the days immediately preceding the spot check visit but provide no information about burning outside that window.

but at some (relatively minor) cost to overall accuracy (by classifying fewer fields as burned, which increases the scope for false negatives). The tradeoff between these two measures affects the control group mean burning rate: predicted burning is higher in the max accuracy measure than the balanced accuracy measure by construction. These differences only change treatment effects if the fields in the part of the distribution between the two thresholds are disproportionately from the treatment or control groups. We find similar treatment effects across the two measures.[30]

The RF model is trained using negative labels that are only available for the treatment group (and positive labels for both treatment and control groups). This could introduce bias into the classification if the spectral signature of not-burning is affected by treatment; if not-burning looks the same in treatment and control, this is not an issue. As one check on potential bias, Appendix Figure A.3 shows the distributions of the continuous random forest model output using fields not in the training set. Both for fields classified as burned and those classified as not-burned, the distributions are similar in the treatment and control groups. Formal statistical tests for equal distributions, conditional on classified burning status, confirm that there is no statistical difference.

The balanced accuracy model has an overall accuracy of 78 percent, while the max accuracy model is necessarily better at 82 percent (see Appendix Table A.2). The difference is due to 425 fields or 203 farmers (12% of the sample) being classified as not-burned under the balanced accuracy threshold and burned under the max accuracy threshold. 11% of the control group, 10% of the standard PES treatment and 15% of the upfront PES treatment fall between the thresholds, i.e., moving from balanced accuracy to max accuracy re-assigns their not-burned outcome from 1 to 0.

3.3 Burning and farmer heterogeneity measures

Burning measures We use multiple data sources to measure farmer residue burning, and consider ancillary outcomes that are likely related to burning. To begin, we use contract compliance data to create an indicator for whether a farmer requested monitoring and was

[30]Given that the farmer-level outcome depends on whether a farmer burns any of his fields, differences between the measures may be more pronounced than at the field level.

found not to have burned during the monitoring visit. This outcome allows us to compare the treatment effects of standard PES versus upfront PES. However, because we do not measure this outcome in the control group, we cannot use it to compare impacts relative to counterfactual burning.

Second, we use remote sensing measures that construct a predicted not-burning measure for all farmers. We present two versions of the remote sensing model, as described in the previous sub-section, which rely on the same underlying imagery and machine learning models but differ in the threshold used to determine a field-level binary burned/not-burned outcome.

Finally, we rely on endline survey data for measures of CRM practices among treatment and control farmers, which include the primary alternatives to burning: balers (*ex-situ*) and Happy Seeders (*in-situ*).

Farmer heterogeneity measures Following our pre-analysis plan, we use baseline data to create a distrust index and a financial constraints index, which are the two barriers that motivated us to test the upfront PES contract. The survey questions used to create all heterogeneity indices and their aggregation into binary variables are discussed in Appendix A.5. Because our study population had no experience with PES programs, and over 40% had no experience with CRM equipment, the survey questions focused on farmer trust in general and overall financial constraints related to agriculture. Both baseline indices thus capture generalized beliefs rather than specific beliefs related to choices around residue burning. For instance, when measuring financial constraints, we asked farmers if they could access savings or loans for agricultural equipment for two different amounts (₹5,000 and ₹10,000). Trust was measured using questions that asked farmers their level of trust in people in general, and specific groups (the Punjab Government, their village Panchayat, cooperative society, NGOs, and financial institutions like banks and insurance companies). In contrast, the endline survey directly asked treatment group farmers about trust in the PES program, and financial constraints specific to CRM.

We also construct indices of CRM equipment access, which may affect overall PES program take-up or compliance (we did not hypothesize they would differentially affect compliance with standard versus upfront PES). We construct three sub-indices to measure different access barriers: informational barriers (i.e., awareness of alternatives to burning and knowledge of where to rent them), CRM equipment access barriers (including delays), and negative beliefs about the impact of CRM equipment on soil health and agricultural yield.

4 Results

4.1 Estimation strategy

We estimate the following equation:

$$y_{ij} = \alpha + \beta StandardPES_j + \gamma UpfrontPES_j + \psi X_j + \epsilon_{ij}$$

where y_{ij} denotes an outcome for farmer i in village j, and $StandardPES_j$ and $UpfrontPES_j$ are binary variables that take the value 1 if village j is assigned to the standard PES treatment and the upfront PES treatment, respectively. X_j are strata fixed effects. Following our pre-analysis plan, and to increase statistical power, we pool treatment variants (different payment levels in $StandardPES_j$ and different proportions paid upfront in $UpfrontPES_j$). Standard errors are clustered at the village level. β is the effect of being assigned to the standard PES treatment, and γ is the effect of assignment to the upfront PES treatment (each relative to the control group).

Table 1 presents summary statistics and four types of balance tests: for the treatment (pooling all treatment arms) versus control group, for standard PES (pooling ₹800 and ₹1600 per acre) versus the control group, upfront PES (pooling 25% upfront and 50% upfront groups) versus the control group, and for upfront PES versus standard PES. In each comparison, self-reported burning in 2018 is balanced. We observe slight imbalance in measures of land size and CRM indices; the p-value of the joint F-test is 0.48 between treatment and control.

Table 1 also provides more details on our respondents. They have significant farming experience and produce on about five acres of land, on average. The majority of household income comes from agriculture, with a control group mean agricultural profit of ₹114,000 (the median is ₹58,000).[31] At baseline, 68 percent of farmers self-reported burning their fields in the previous year, though the actual value is likely higher given our screening criteria. Less than half of the sample, 48%, had previously signed a written contract.

We have no attrition for our main outcomes of PES compliance and remotely-sensed burning. For the endline survey, Appendix Table A.3 shows a reasonably high response rate for a phone survey (greater than 80%) but with some differential attrition; the treatment group was about 5 percentage points less likely to respond than the control. However, there is no difference in attrition between the standard and upfront PES groups.[32] Following Lee (2009), we report both the point estimates and upper and lower bounds on our estimates for regressions using endline survey outcomes for treatment and control groups.[33] We also test for whether there was differential attrition across treatment groups by baseline characteristics or by burning choice during the experiment. The results, presented in Appendix Table A.4, do not indicate that this is the case.[34]

4.2 Treatment effects

A. Did farmers take up the PES contract offers?

In Table 2, we examine differences across treatment arms in whether the surveyor found and offered the farmer the PES contract, as well as whether the farmer was eligible for the offer (i.e., he had a bank account and had not yet harvested his paddy). These outcomes are zero by construction in the control group, and they are determined for treatment farmers before they learned the details of the contracts. Farmers were found in treatment arms at a similar rate: 87.9% in the upfront PES treatments versus 88.5% in the standard PES treatment

[31] The baseline agricultural profit measure helps interpret the magnitude of the contract payments; the median base contract payment of ₹800 per acre equals ₹4,000 per farmer (since the average farmer cultivates 5 acres), or about 7% of median agricultural profit.

[32] 8% of the sample had invalid numbers, while 4.8% could not be contacted in eight attempts. The remaining attrition is due to: 1% of the sample didn't answer the phone, 0.12% had the wrong number, 0.59% refused to talk prior to the consent, 1.2% didn't consent to the survey, and 0.96% didn't complete the survey.

[33] Lee bounds entail estimating treatment effects after dropping (in turn) the observations with the lowest and highest outcomes in the less-attrited group, so that the proportion of included observations is equal across treatment and control. We trim observations from the control group within strata and include strata fixed effects in the estimation because our randomization is stratified.

[34] These results are estimated using a single regression that regresses attrition on all baseline characteristics, treatment status and each of the interactions between the two.

(column 1). The same is true for eligibility: 83.3% of farmers were eligible in the upfront PES treatment and 84.4% in the standard PES treatment (column 2).

Next we consider program take-up, which is again zero by construction for control farmers. Take-up (i.e., whether a farmer signed the PES contract) was high, with around 72% of farmers in both treatment arms signing the contract.[35] Conditional on being found and being eligible, the probability that a farmer took up the PES program is very high (around 87%). This is consistent with the option value associated with the contract being high in both treatments (since farmers who choose to burn forgo the conditional payments but face no penalties). The fact that contract take-up is not higher when the PES contract includes upfront payment is surprising, but could stem from the fact that farmers were not given the upfront cash right away. They were asked for their bank account information, and funds were transferred 2-3 days later. Trust barriers might have similarly deterred take up in the two treatment arms; farmers may have objected to sharing bank account information or granting access to their plots for monitoring.

B. Did farmers offered PES contracts reduce their crop residue burning?

Our first measure of not-burning is contract compliance, i.e., whether the farmer called for monitoring and in the monitoring visit was found in compliance with the contract (no fields were burned). Table 3, column 1 shows that 8.5% of farmers complied with the contract in the standard PES group. Compliance in the upfront arm is 10 percentage points higher, at 18%, and equality between the two groups can be rejected with $p < 0.01$.[36] Thus, despite similar contract take-up rates in the standard and upfront PES treatments, upfront payments make farmers twice as likely to comply with the contract.

Next, we consider our two remote sensing based measures of whether a farmer burned any fields.[37] Columns 2 and 3 of Table 3 present results for the balanced accuracy and max accuracy classifications, respectively. As with contract compliance, upfront PES outperforms

[35] As shown in Appendix Table A.5, the take-up rate was similar across all sub-treatments.

[36] Treatment effects on the outcome of whether the farmer called for monitoring (regardless of compliance) follow a similar pattern – 11% of farmers in standard PES called and there is a 10 percentage points higher likelihood of calling for monitoring in the upfront arm.

[37] About 18% of the sample had at least one field predicted to be burned and at least one field predicted not burned. We present plot-level treatment effects in Appendix Table A.6.

standard PES, with treatment effects that are similar across the two alternative remote sensing measures. Relative to the control group, upfront PES increases not-burning by 8-11.5 percentage points. In contrast, the effects of standard PES relative to the control group is indistinguishable from zero.[38] The latter result implies that standard PES payments were inframarginal. The farmers who complied with the standard PES contract and received payment would not have burned their fields even in the absence of the program.

The two remote sensing models differ in their estimate of the not-burning rate in the control group, which is 20% in the balanced accuracy classification and 9% in the max accuracy classification. Appendix Figure A.4 shows the control group mean and estimated treatment effects as the threshold is varied from a lower bound that classifies all true not-burned as not-burned to an upper bound that classifies all true burned as burned. The treatment effects are less sensitive to the choice of threshold (middle and bottom panel) than is the control group mean (top panel). The variability in the estimated control group mean does not affect the conclusion that the upfront PES arm was effective at lowering burning, but it does bear on whether the results imply that some farmers left money on the table. For example, with the higher control group mean estimated in the balanced accuracy model, the implied rate of not-burning in the upfront villages is 31%, implying that some farmers in the upfront arm who did not burn also did not call for the conditional payment.[39] The max accuracy results, on the other hand, imply that all farmers who did not burn received the conditional payment, so farmers did not leave money on the table.

Appendix Table A.5 disaggregates the effects by the four treatment arms. The point estimates for all outcomes other than contract take-up are larger for the upfront PES arms. Paying more in the standard PES variants (₹1,600 versus ₹800) increased the point estimates on compliance and the remote sensing based measure, though their difference is statistically indistinguishable from zero. For both measures, compliance with the ₹1,600 arm is lower than with either upfront arm. This is striking given that the upfront contract only pays ₹800 per acre. Across the two upfront arms (25% upfront versus 50% upfront), it is theoretically

[38] Treatment effects estimated on our subsample for which we have spot check data are similar in magnitude (see Appendix Table A.7), but with considerably lower statistical power and a (mechanically) higher control group mean given that not all burns were detected through a one-time surprise field visit.

[39] In the endline survey, some farmers reported losing the number to call for monitoring or forgetting when they were supposed to call.

ambiguous which should perform better, as increasing the fraction paid upfront meant a lower reward for compliance; we find similar effects for the two variants.[40]

In Table 4, we analyze treatment effects on the main *ex-situ* and *in-situ* alternatives to burning: baler and Happy Seeder use, respectively. Increased use of balers – by 10 percentage points – can explain all of the reduced burning achieved through upfront PES; every farmer who switched away from burning because of the upfront PES program seems to have switched to baling their straw. There are no detectable changes in Happy Seeder use.

Consistent with the null effect of standard PES on our remote sensing based burning measure, we see no difference in CRM usage in the standard PES arm relative to the control group, suggesting that standard PES had no impact on residue management behavior. In the second panel, we present Lee bounds to account for differential attrition by treatment status in the endline survey (from which we derive these results). Consistent with the point estimates, we only see a statistically significant increase in baler use for the upfront PES group (the lower and upper bound point estimates are 8.8 and 14.5 percentage points respectively).

Two important factors cited by farmers as arguments in favor of residue burning are reducing delays between harvesting and sowing the next crop and agricultural yields. In our setting, the significant upfront payouts in the upfront PES treatment may also have directly affected agricultural outcomes. Appendix Table A.8 presents treatment effects on crop yields for rice and the Rabi crop (which is wheat for 99.6% of the sample), and on delays in sowing the Rabi crop. Neither treatment impacted these outcomes.

4.3 Why did upfront payments increase farmer PES compliance?

Upfront PES reduces burning more than standard PES. We now examine why offering partial payment upfront boosted contract effectiveness.

Panel A of Table 5 presents heterogeneous treatment effects by pre-specified baseline indices: a distrust index and a financial constraints index.[41] These indices capture generalized

[40] The results across the two upfront arms diverge more for the max accuracy remote sensing based outcome than for either the compliance or balanced accuracy outcomes, though differences in the coefficients fall well within the standard errors around all of the estimates.

[41] We also use a generalized random forest model to identify baseline variables that explain heterogeneity

distrust and financial constraints faced by farmers, rather than CRM or PES-specific constraints. In both cases, we use a binary measure for being above the index median, or 'highly constrained'. Columns 1 and 2 present results using contract compliance as the outcome, and columns 3 through 6 using remote sensing measures of burning as the outcome. We do not find any differential impact of the upfront program by either heterogeneity index on any of the outcomes.[42]

In addition, in the endline survey, we asked a subset of farmers (63% of the sample) in the standard and upfront arms about distrust and financial constraints related to CRM decisions and the PES program.[43] To gauge trust in the PES program, we asked: "When I was deciding how to manage my residue during Kharif 2019, I was sure that J-PAL would return and pay me the amount owed to me if I did not burn." We create a binary variable that equals 1 if farmers answered yes (low constraint), and 0 otherwise. To test for cash constraints in CRM decisions, we asked farmers how much they agreed with the statement "When I was deciding how to manage my residue during Kharif 2019, shortage of cash limited what methods I could choose for residue management."[44] We create a binary variable that equals 1 if they answered "Shortage of cash was not at all important", "not very important" or "somewhat important" and 0 otherwise.[45]

Panel B of Table 5 presents results using these two endline variables as outcomes.[46]

in the treatment effect of the ₹1600 standard PES treatment relative to the control. The four most important variables determined from this procedure all belong to our trust index. However, when used in a second step to test heterogeneity in the treatment effect of the upfront PES arms relative to the ₹800 standard PES treatment, they are inconsistently signed and imprecisely estimated.

[42]Like all heterogeneity on observables, this null result could mean that the generalized measures are not relevant or that our baseline measures lack sufficient sample variability to pick up an effect. Other mechanisms may also be relevant, such as reciprocity. We are guided by our pre-analysis plan, but note that reciprocity seems unlikely as a mechanism given the high cost of contract compliance and the lack of differential endline survey participation by treatment arm.

[43]We asked these questions of only a subset of farmers because we reduced the survey duration partway through the sample due to respondent fatigue with a lengthy phone survey.

[44]Only contract signers were asked about trust, but both were asked about cash shortages. To keep the sample consistent, we only study contract signers here, but results on monetary constraints are similar if we include all respondents.

[45]The other options were that shortage of cash was very important or extremely important in determining CRM decisions. Alternative classification which classifies cash shortage as being unimportant if the respondent answered "Shortage of cash was not at all important" or "not very important" lead to qualitatively similar results.

[46]Note that since there is no differential attrition in the standard versus upfront arms for the endline survey, Lee bounds are not required.

Farmers in the upfront arm are nearly 7 percentage points more likely than farmers in the standard PES arm to say that they trusted that the conditional payment would be made if they complied. Recall that the difference in compliance between the two contract types is about 10 percentage points, so the increase in trust represents an economically significant effect.[47] We interpret this result as suggestive evidence that instilling trust was a mechanism by which upfront PES was more effective than standard PES.

In contrast, farmers in the upfront arm were not more likely to report cash shortages being unimportant for CRM decision-making. However, this does not necessarily imply that upfront payments do not affect the ability to pay for CRM equipment. The upfront amount (which could cover less than half the cost of baler rental) may have been too low to meaningfully ease this constraint for many farmers.

Appendix Table A.9 presents additional pre-specified heterogeneous treatment effects that pertain to the overall effect of PES (rather than upfront versus standard PES). These include information constraints about CRM equipment, barriers to accessing CRM equipment, and beliefs about the negative impact of CRM use on yields, pest prevalence, or soil health. These heterogeneity results are estimated without the control group, where the outcome of contract compliance is zero by construction. The correlation between the heterogeneity variable in the pooled treatment group and compliance can thus be interpreted as the heterogeneous treatment effect. Farmers with higher information constraints and more negative beliefs about burning alternatives are less likely to comply with PES, pooling all of the treatment arms.

We also test whether program take-up is differential by any of these five pre-specified indices (liquidity, distrust, information about CRM alternatives, CRM access constraints, and negative beliefs about burning alternatives). These results are presented in Appendix Table A.10 and do not indicate that the program had differential take-up by any of these indices.

[47] Note that the higher reported trust in the payment among farmers in the upfront PES treatment is not limited to the compliers in the upfront PES arms; levels of trust do not vary with compliance in this arm. In contrast, in the standard PES arms, farmers who complied – and received payment – were around 10 percentage points more likely to report a high level of trust in the payment.

5 Program Costs and Benefits

We present a cost effectiveness comparison across treatments in Table 6. Column 1 shows the treatment effects for farmer-level payments (zero by construction in the control group).[48] The results show that upfront PES pays out more, both because more people comply and because farmers who sign up yet burn still receive upfront payments. Columns 2 and 4 reproduce the treatment effect on the binary variable of not-burning shown earlier, using the balanced and max accuracy models, respectively. Note the slightly different results across the columns, which we return to below.

Column 3 shows the cost per additional unburned acre based on the balanced accuracy model, calculated by dividing the estimates in column 1 by the estimates in column 2 for each treatment. (Figure A.5 shows the results by treatment arm.) The estimated effect on burning of standard PES is very close to zero, resulting in a high cost per unburned acre (₹13,441). Upfront PES is more cost effective than standard PES, according to the point estimates, despite the higher payment per acre; the cost per averted acre of burning is ₹2,695 ($34). However, we cannot reject equality across the arms due to the very imprecise estimate for standard PES's cost per unburned acre (which stems from the standard error for standard PES in column 2 being much larger than the estimated coefficient).

For the max accuracy, the tradeoff between treatment effects and cost effectiveness is more apparent (column 5). For this measure, the effect of standard PES on burning (column 4) is small and statistically imprecise, but considerably larger than with the balanced accuracy measure. The treatment effect on not-burning from upfront PES is also smaller using the max accuracy classification than the balanced accuracy classification. Thus, upfront PES is "only" four times as effective as standard PES. Because upfront PES entails three times as large a payment per acre as standard PES, the cost-effectiveness gap between the contract types is not as stark: Upfront PES has a cost of ₹4050 ($51) per unburned acre, compared to ₹5157 ($64) for standard PES (column 5).

Next, we compare these costs to the benefits of reducing crop burning. Lan et al. (2022)

[48]We exclude the cost of monitoring contract compliance because it is sensitive to program scale and specific protocols around monitoring. Monitoring costs for our program were roughly ₹300/field, or less than ₹100/acre, which is small in comparison to the PES payments.

combine satellite data on fire intensity with air transport models to estimate that 86,000 premature deaths were caused by crop residue burning in India in 2018, 53.5% of which can be attributed to Kharif burning in Punjab. Estimates of the Value of a Statistical Life (VSL) for India range from $700,000 (Majumder et al., 2018) to 5.6 million (Madheswaran, 2007). Using the lower bound of this range as a conservative estimate, this implies $32 billion of annual damages. Kumar et al. (2019) estimates that 66%, or around 4 million acres, of Punjab's (non-basmati) paddy was burned in the Kharif season of 2018, which implies damages of $8,000 per acre (₹632,000). These mortality damages are 150 to 230 times the per acre cost of reducing burning through PES with upfront payment.[49]

To put this cost per life saved in perspective, we compare it to other pollution abatement options. Coal-fired power plants are a major source of particulate emissions in India, and installing SO_2 scrubbers lowers emissions of fine particulate matter. According to the Global Burden of Disease 2018, thermal power claims 84,650 lives in India each year. Installing scrubbers avoids 72% of deaths in the first year (Cropper et al., 2019), which would save 60,948 lives. India has 204 GW of coal capacity (Ministry of Power, 2022); on average, installing scrubbers in a 500 MW plant would save 150 lives per year, and costs about $1.2 billion over 20 years, or 60 million per year annualized. Assuming no discounting and constant benefits for the lifetime of scrubbers, the cost per life saved is around $400,000. For comparison, the cost per life saved from our most cost effective treatment is $2,930.[50] Thus, while several abatement opportunities have benefits that outweigh costs in India, residue burning reduction has the potential to save lives at a much lower cost.

6 Conclusion

Identifying low-cost pollution abatement opportunities allows for more efficient use of limited environmental protection budgets. Even if promising opportunities are identified, market imperfections and weak institutions can undermine the efficacy of policies aimed at these

[49] We use estimates of percentage of paddy burned from Kumar et al. (2019), who use satellite-based active fire products to quantify total rice area burned in Punjab in the Kharif season of 2018.

[50] Using above assumptions, crop residue burning causes 0.02 deaths per acre annually. Reducing one acre of burning costs ₹2695/acre, or $34/acre, which implies $2,930 per life saved.

abatement opportunities.

We show that crop residue burning, which has significant environmental and human health costs, can be reduced through well-designed PES payments. In particular, program design that takes institutional constraints and farmer concerns into account can significantly improve efficacy. Providing a portion of the contract payment upfront results in larger reductions in burning than providing the entire payment after participants have completed costly behavior change. Despite higher "wasted" payments (to farmers who continue to burn), PES with upfront payments is cost effective, resulting in burning reductions that provide benefits far in excess of their cost.

Scaling up PES contracts to avoid burning poses some challenges. First, the monitoring protocols that we implemented were not designed for scale. Viable approaches to large-scale monitoring, such as remote sensing, are likely to increase contract risk because of measurement error (though different forms of contractual risks may exist with scaled-up in-person monitoring as well). Second, equipment to manage crop residue is still scarce, and PES at scale will increase demand for this equipment, raising the cost of not burning. For example, if all farmers began renting balers, the rental price of balers would rise unless the supply of rental equipment in a village is elastic.

There are also reasons for optimism. The roll-out of a large-scale PES program would create incentives for innovation in better CRM equipment or rental market efficiency. Dynamic incentives — in the form of tying future eligibility to verified not burning — could reduce the cost of providing upfront payments. The need for upfront payments also might become less important as trust in being paid grows. Finally, PES programs are appealing because they can be implemented by organizations that want to reduce fires but lack the authority to levy fines. The enormity of the environmental damage caused by crop residue burning in India justifies such an investment and also highlights the need for further research to find viable solutions to this problem.

References

Agricultural and Processed Food Products Export Development Authority (2018). Basmati acreage yield estimation 2018. *Basmati Export Development Foundation*.

Aker, J. C. and K. Jack (2021). Harvesting the rain: The adoption of environmental technologies in the Sahel. Working Paper 29518, National Bureau of Economic Research.

Akresh, R., D. De Walque, and H. Kazianga (2013). Cash transfers and child schooling: Evidence from a randomized evaluation of the role of conditionality. Policy Research Working Paper 6340, The World Bank.

Attanasio, O. P., V. Oppedisano, and M. Vera-Hernández (2015). Should cash transfers be conditional? Conditionality, preventive care, and health outcomes. *American Economic Journal: Applied Economics 7*(2), 35–52.

Badarinath, K. V. S., T. R. K. Chand, and V. K. Prasad (2006). Agriculture crop residue burning in the Indo-Gangetic plains: A study using IRS-P6 AWiFS satellite data. *Current Science 91*(8), 1085–1089.

Baird, S., C. McIntosh, and B. Özler (2011). Cash or condition? Evidence from a cash transfer experiment. *The Quarterly Journal of Economics 126*(4), 1709–1753.

Balboni, C., R. Burgess, A. Heil, J. Old, and B. A. Olken (2021). Cycles of fire? Politics and forest burning in Indonesia. *AEA Papers and Proceedings 111*, 415–19.

Behrer, A. P. (2019). Earth, wind and fire: The impact of anti-poverty efforts on Indian agriculture and air pollution. Working Paper.

Bhuvaneshwari, S., H. Hettiarachchi, and J. N. Meegoda (2019). Crop residue burning in India: Policy challenges and potential solutions. *International Journal of Environmental Research and Public Health 16*(5), 832.

Burke, M., L. F. Bergquist, and E. Miguel (2019). Sell low and buy high: Arbitrage and local price effects in Kenyan markets. *The Quarterly Journal of Economics 134*(2), 785–842.

Carter, M., A. de Janvry, E. Sadoulet, and A. Sarris (2017). Index insurance for developing country agriculture: A reassessment. *Annual Review of Resource Economics 9*, 421–438.

Casaburi, L. and J. Willis (2018). Time versus state in insurance: Experimental evidence from contract farming in Kenya. *American Economic Review 108*(12), 3778–3813.

Cassou, E. (2018). Agricultural pollution: Field burning. *World Bank: Washington, DC, USA*.

Chen, J., C. Li, Z. Ristovski, A. Milic, Y. Gu, M. S. Islam, S. Wang, J. Hao, H. Zhang, C. He, et al. (2017). A review of biomass burning: Emissions and impacts on air quality, health and climate in China. *Science of the Total Environment 579*, 1000–1034.

Coffman, L. C., J. J. Conlon, C. R. Featherstone, and J. B. Kessler (2019). Liquidity affects job choice: Evidence from Teach for America. *The Quarterly Journal of Economics 134*(4), 2203–2236.

Cropper, M. L., S. Guttikunda, P. Jawahar, Z. Lazri, K. Malik, X.-P. Song, and X. Yao (2019). Applying benefit-cost analysis to air pollution control in the Indian power sector. *Journal of Benefit-Cost Analysis 10*(S1), 185–205.

Cusworth, D. H., L. J. Mickley, M. P. Sulprizio, T. Liu, M. E. Marlier, R. S. DeFries, S. K. Guttikunda, and P. Gupta (2018). Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India. *Environmental Research Letters 13*(4), 044018.

Dipoppa, G. and S. Gulzar (2022). Fires: The political economy and health impacts of crop burning in South Asia. *Working Paper*.

Distributed Active Archive Center (2018). MODIS and VIIRS land products global subsetting and visualization tool. Oak Ridge National Laboratory.

Edwards, R. B., W. P. Falcon, G. Hadiwidjaja, M. M. Higgins, R. L. Naylor, and S. Sumarto (2020). Fight fire with finance: A randomized field experiment to curtail land-clearing fire in Indonesia. Working Paper 55-e - 2020, Tim Nasional Percepatan Penanggulangan Kemiskinan.

Fink, G., B. K. Jack, and F. Masiye (2020). Seasonal liquidity, rural labor markets, and agricultural production. *American Economic Review 110*(11), 3351–92.

Fuller, R., P. J. Landrigan, K. Balakrishnan, G. Bathan, S. Bose-O'Reilly, M. Brauer, J. Caravanos, T. Chiles, A. Cohen, L. Corra, et al. (2022). Pollution and health: a progress update. *The Lancet Planetary Health*.

Garg, T., M. Jagnani, and H. K. Pullabhotla (2021). Agricultural labor exits increase crop fires. Working Paper.

Giné, X. and D. Yang (2009). Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of Development Economics 89*(1), 1–11.

Graff Zivin, J., T. Liu, Y. Song, Q. Tang, and P. Zhang (2020). The unintended impacts of agricultural fires: Human capital in China. *Journal of Development Economics 147*, 102560.

He, G., T. Liu, and M. Zhou (2020). Straw burning, PM2.5, and death: Evidence from China. *Journal of Development Economics 145*, 102468.

Jack, B. K. (2013). Private information and the allocation of land use subsidies in Malawi. *American Economic Journal: Applied Economics 5*(3), 113–35.

Jain, N., A. Bhatia, H. Pathak, et al. (2014). Emission of air pollutants from crop residue burning in India. *Aerosol and Air Quality Research 14*(1), 422–430.

Jayachandran, S. (2009). Air quality and early-life mortality evidence from Indonesia's wildfires. *Journal of Human Resources 44*(4), 916–954.

Jayachandran, S., J. De Laat, E. F. Lambin, C. Y. Stanton, R. Audy, and N. E. Thomas (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science 357*(6348), 267–273.

Kramer, B. and F. Ceballos (2018). Enhancing adaptive capacity through climate-smart insurance: Theory and evidence from India. Working paper, International Food Policy Research Institute.

Kumar, P., S. Rajpoot, V. Jain, S. Saxena, S. Ray, et al. (2019). Monitoring of rice crop in punjab and haryana with respect to residue burning. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 42*, 31–36.

Lan, R., S. D. Eastham, T. Liu, L. K. Norford, and S. R. Barrett (2022). Air quality impacts of crop residue burning in India and mitigation alternatives. *Nature Communications 13*(1), 6537.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies 76*(3), 1071–1102.

Lee, K. and M. Greenstone (2021). Air Quality Life Index annual update. Technical report, Energy Policy Institute at the University of Chicago.

Liu, T., M. E. Marlier, A. Karambelas, M. Jain, S. Singh, M. K. Singh, R. Gautam, and R. S. DeFries (2019). Missing emissions from post-monsoon agricultural fires in northwestern India: Regional limitations of MODIS burned area and active fire products. *Environmental Research Communications 1*(1), 011007.

Madheswaran, S. (2007). Measuring the value of statistical life: estimating compensating wage differentials among workers in India. *Social indicators research 84*, 83–96.

Majumder, A., S. Madheswaran, et al. (2018). *Value of statistical life in India: A hedonic wage approach*. Institute for Social and Economic Change.

Ministry of Power (2022). Annual report 2021–22. Technical report, Government of India.

Murali, S., R. Shrivastava, and M. Saxena (2010). Green house gas emissions from open field burning of agricultural residues in India. *Journal of Environmental Science & Engineering 52*(4), 277–284.

NASA (2017). Global fire emissions indicators, country-level tabular data: 1997–2015. Technical report, Center for International Earth Science Information Network (CIESIN), Columbia University.

National Green Tribunal (2015). Vikrant Kumar Tongad v. Environment Pollution Authority & Ors. Judgement.

Nian, Y. (2023). Incentives, penalties, and rural air pollution: Evidence from satellite data. *Journal of Development Economics*, 103049.

Oanh, N. T. K., D. A. Permadi, P. K. Hopke, K. R. Smith, N. P. Dong, and A. N. Dang (2018). Annual emissions of air toxics emitted from crop residue open burning in Southeast Asia over the period of 2010–2015. *Atmospheric Environment 187*, 163–173.

Oliva, P., B. K. Jack, S. Bell, E. Mettetal, and C. Severen (2020). Technology adoption under uncertainty: Take-up and subsequent investment in Zambia. *Review of Economics and Statistics 102*(3), 617–632.

Pullabhotla, H. (2018). Fires, wind, and smoke: Air pollution and infant mortality. Job Market Paper.

Rangel, M. A. and T. S. Vogl (2019). Agricultural fires and health at birth. *Review of Economics and Statistics 101*(4), 616–630.

Rao, P. P., L. Pandey, E. Jagadeesh, U. K. Deb, R. Jain, and K. Basu (2012). Meso-level database coverage and insights village dynamics in south asia.

Sarkar, S., R. P. Singh, and A. Chauhan (2018). Increasing health threat to greater parts of India due to crop residue burning. *The Lancet Planetary Health 2*(8), e327–e328.

Sharma, H. (2022). 14 of 15 States had Asked for Higher MSP for Paddy Than What Centre Announced. *Indian Express*.

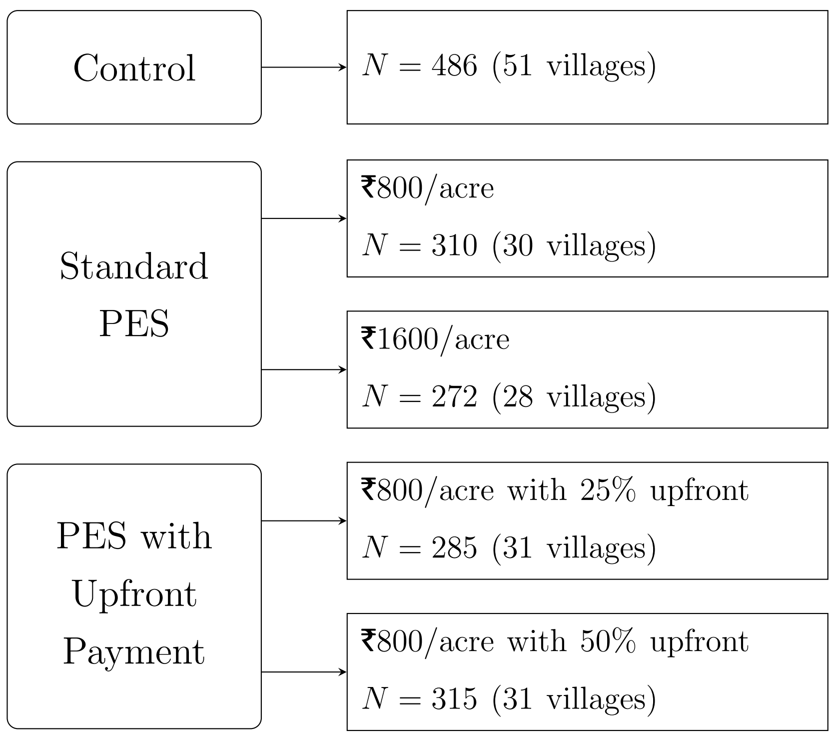
Shindell, D., J. C. Kuylenstierna, E. Vignati, R. van Dingenen, M. Amann, Z. Klimont, S. C. Anenberg, N. Muller, G. Janssens-Maenhout, F. Raes, et al. (2012). Simultaneously mitigating near-term climate change and improving human health and food security. *Science 335*(6065), 183–189.

Singh, D., N. Kundu, and S. Ghosh (2021). Mapping rice residues burning and generated pollutants using Sentinel-2 data over northern part of India. *Remote Sensing Applications: Society and Environment 22*, 1.

Vadrevu, K. P., E. Ellicott, K. V. S. Badarinath, and E. Vermote (2011). MODIS derived fire characteristics and aerosol optical depth variations during the agricultural residue burning season, north India. *Environmental Pollution 159*(6), 1560–1569.

Walker, K., B. Moscona-Remnitz, K. Jack, S. Jayachandran, N. Kala, R. Pande, J. Xue, and M. Burke (2022). Detecting crop burning in India using satellite data. Working Paper.

Figure 1: Experimental Design



Note: Treatments assigned at the village level. See text for additional detail.

Figure 2: Timeline

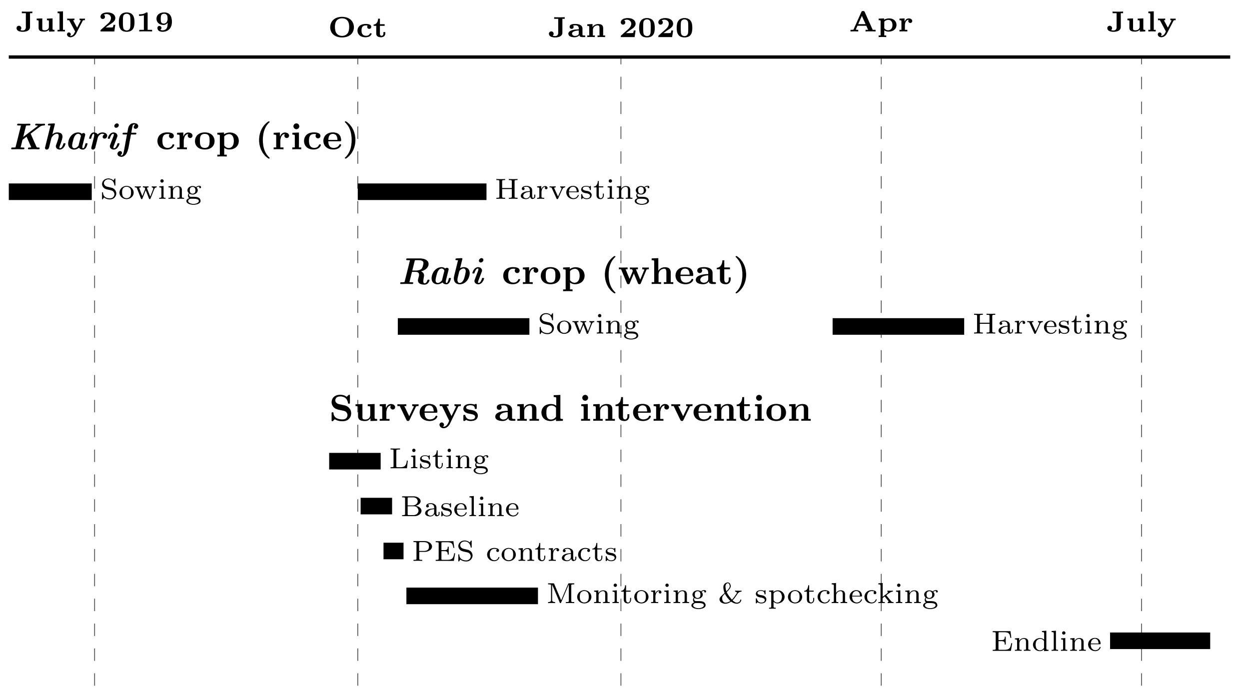


Table 1: Summary Statistics and Balance

	N	Control		Treatment vs Control	Standard vs Control	Upfront vs Control	Upfront vs Standard
		Mean	SD				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Demographics							
Age (years)	1668	48.675	12.732	-0.158	-0.448	0.124	0.610
				(0.751)	(0.816)	(0.825)	(0.653)
Total experience in agriculture (years)	1668	28.055	13.144	-0.184	-0.452	0.078	0.522
				(0.788)	(0.860)	(0.874)	(0.732)
Highest educational class passed	1658	7.213	4.197	-0.147	-0.200	-0.096	0.114
				(0.228)	(0.265)	(0.270)	(0.283)
1(Ever signed a written contract)	1440	0.483	0.500	-0.048	-0.049	-0.048	0.001
				(0.039)	(0.044)	(0.044)	(0.041)
Panel B: Income							
Total income	1602	125.694	172.588	-4.655	1.060	-10.190	-10.352
				(11.386)	(12.886)	(13.798)	(14.493)
Non-agricultural income	1455	18.076	66.407	-1.393	-2.545	-0.277	3.150
				(4.563)	(5.136)	(6.630)	(8.084)
Total agricultural profit	1521	114.177	155.748	-2.759	4.674	-9.905	-14.483
				(11.246)	(12.700)	(12.865)	(12.426)
Total area of land in acres (measured)	1668	4.986	2.816	0.327	0.350	0.304	-0.049
				(0.173)*	(0.203)*	(0.188)	(0.181)
Paddy production in 1000kg	1513	13.250	9.593	0.684	1.069	0.308	-0.768
				(0.625)	(0.736)	(0.700)	(0.722)
Panel C: Heterogeneity variables							
Liquidity constraints index	1668	0.504	0.500	0.011	0.020	0.003	-0.012
				(0.039)	(0.042)	(0.044)	(0.039)
Distrust index	1655	0.476	0.500	0.043	0.058	0.029	-0.032
				(0.035)	(0.039)	(0.040)	(0.034)
CRM information constraints index	1676	0.444	0.497	0.026	0.062	-0.009	-0.076
				(0.034)	(0.040)	(0.039)	(0.041)*
CRM access constraints index	1651	0.445	0.497	0.076	0.078	0.073	-0.009
				(0.033)**	(0.036)**	(0.038)*	(0.034)
CRM negative beliefs index	1676	0.500	0.501	0.069	0.095	0.043	-0.060
				(0.035)**	(0.039)**	(0.038)	(0.032)*
Panel D: Burning							
1(Burned paddy residue in 2018)	1576	0.684	0.465	0.025	0.044	0.006	-0.037
				(0.034)	(0.037)	(0.039)	(0.035)
P-value of joint F-test				0.475	0.482	0.693	0.808

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors reported in parentheses. Columns 2 and 3 show the summary statistics for the control group in the baseline. Column 3 shows the coefficient from regressing the baseline variable on a treatment dummy (taking value 1 if the respondent was in any treatment group). Columns 5 and 6 are the coefficients from regressing the baseline variable on two treatment dummies (taking the value 1 if the respondent was in the standard PES treatment, and taking value 1 if the respondent was in upfront PES treatment, respectively). Column 7 shows the coefficient from regressing the baseline variable on a dummy taking value 1 if the respondent was in the upfront payment treatment. Regressions in columns 4, 5 and 6 use the control group as the comparison group. Regressions in column 7 uses the standard PES group as the comparison group. All regressions cluster standard errors at the village level and include strata fixed effects. All income variables relate to income derived in the past 12 months and are measured in ₹1000. All index variables in Panel C are binary. There are 1,668 observations in the baseline; 486 observations for the control group and 1,182 observations for the treatment groups.

Table 2: Contract Eligibility and Take-Up

	Farmer Found (1)	Farmer Eligible (2)	Contract Take-Up (3)
Standard PES	0.885	0.844	0.726
	(0.014)***	(0.019)***	(0.023)***
Upfront PES	0.879	0.833	0.718
	(0.011)***	(0.014)***	(0.022)***
p-val: Standard PES = Upfront PES	0.720	0.627	0.798
Control mean	0.000	0.000	0.000
Standard PES mean	0.887	0.852	0.741
Upfront PES mean	0.880	0.837	0.728
N	1668	1668	1668

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects included. "Farmer Found" is a dummy variable taking the value 1 if the respondent was available during the intervention. "Farmer Eligible" is a dummy variable taking the value 1 if the respondent was available during the intervention and had a bank account. "Contract Take-Up" is a dummy variable taking the value 1 if the respondent signed a contract to participate in the PES program.

Table 3: Contract Compliance and Not Burning

	Compiled with Contract (1)	Not Burned (Balanced Accuracy) (2)	Not Burned (Maximum Accuracy) (3)
Standard PES	0.085	0.008	0.020
	(0.015)***	(0.042)	(0.030)
Upfront PES	0.183	0.115	0.077
	(0.020)***	(0.042)***	(0.032)**
p-val: Standard PES = Upfront PES	0.000	0.008	0.071
Control mean	0.000	0.202	0.091
Standard PES mean	0.084	0.198	0.098
Upfront PES mean	0.185	0.313	0.161
N	1668	1664	1664

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects included. "Complied with Contract" is a dummy variable taking the value 1 if the respondent called to request monitoring of his plots, and monitoring led to the conclusion that the respondent complied with the contract i.e. did not burn his paddy residue. "Not Burned" takes value 1 if the farmer did not burn any of his plots, and 0 otherwise. Column 2 classifies burning to balance type I and type II errors; column 3 classifies burning to maximize overall model accuracy.

Table 4: Crop Residue Management Methods

	Baler (1)	Seeder (2)
Standard PES	-0.010	-0.020
	(0.037)	(0.023)
Upfront PES	0.096	0.013
	(0.039)**	(0.026)
p-val: Standard PES = Upfront PES	0.014	0.157
Control mean	0.199	0.102
Standard PES mean	0.171	0.087
Upfront PES mean	0.295	0.112
N	1387	1387
Lee Bounds		
Standard PES		
Lower bound	-0.025	-0.028
	(0.027)	(0.021)
Upper bound	0.036	0.047
	(0.025)	(0.017)***
Upfront PES		
Lower bound	0.088	0.006
	(0.030)***	(0.022)
Upper bound	0.145	0.068
	(0.028)***	(0.019)***

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. "Baler" is a dummy variable taking value 1 if the farmer reported in the endline that he used a baler. "Seeder" is a dummy variable taking value 1 if the farmer reported in the endline that he used a Happy Seeder or a Super Happy Seeder. The second panel shows the Lee bounds for the treatment effects by treatment group. For details on the Lee bounds, see footnote 33.

Table 5: Liquidity and Distrust as Moderators of the Treatment Effects

Panel A: Liquidity Constraints and Distrust

Outcome variable:	Complied with Contract		Burned (Balanced Accuracy)		Burned (Maximum Accuracy)	
Type of constraint:	Distrust (1)	Liquidity (2)	Distrust (3)	Liquidity (4)	Distrust (5)	Liquidity (6)
Upfront PES	0.114	0.088	0.031	0.030	0.054	0.051
	(0.030)***	(0.029)***	(0.032)	(0.033)	(0.029)*	(0.029)*
Highly constrained	0.030	0.010	0.021	0.007	0.003	0.014
	(0.024)	(0.022)	(0.029)	(0.035)	(0.025)	(0.030)
Upfront PES $\times$ Highly constrained	-0.032	0.018	-0.037	-0.032	-0.025	-0.022
	(0.036)	(0.038)	(0.039)	(0.048)	(0.039)	(0.042)
Standard PES mean	0.083	0.084	0.167	0.167	0.104	0.105
Upfront PES mean	0.185	0.185	0.174	0.174	0.143	0.142
N	1172	1182	1168	1178	1168	1178

Panel B: Trust in Payment and Importance of Cash Shortage

Outcome variable:	Trusted Payment (1)	Cash Shortage Not Important (2)
Upfront PES	0.068**	0.038
	(0.028)	(0.043)
Standard PES mean	0.854	0.441
N	580	584

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. Panel A: "Type of constraint" is the heterogeneity variable which is indicated in the second row of the table. "Liquidity" is an index indicating liquidity constraints, including constrained access to cash and loans. "Distrust" is an index indicating the farmer's distrust in categories of people and organizations. All indices are binary and take value 1 if the farmer's constraints are larger than or equal to the median. The outcome variable is indicated in the top row: "Complied with Contract" is a dummy variable taking the value 1 if the monitoring led to the conclusion that the respondent complied with the contract i.e. did not burn his paddy residue. The comparison group is the standard PES group. Panel B: "Trusted Payment" takes value 1 if the respondent trusted that the payment by J-PAL will be made if they did not burn their paddy residue, and 0 otherwise. "Cash Shortage not Important" takes the value 1 if the respondent declared that cash shortage was not an important factor when deciding which crop residue management method to use, and 0 otherwise. These outcome variables are from the endline survey. The control group is excluded from the sample. The comparison group is the standard PES treatment group. Only those who signed a contract i.e. took up the treatment are included in the sample.

Table 6: Cost Effectiveness

		Balanced Accuracy		Maximum Accuracy	
	Amount Paid per Acre (1)	Not Burned (2)	Cost per Unburned Acre (3)	Not Burned (4)	Cost per Unburned Acre (5)
Standard PES	105.6	0.008	13440.5	0.020	5156.5
	(21.7)***	(0.042)	(71152.0)	(0.030)	(7156.0)
Upfront PES	310.5	0.115	2695.0	0.077	4051.3
	(15.4)***	(0.042)***	(948.7)***	(0.032)**	(1595.0)**
p-val: Standard PES = Upfront PES	0.000	0.008	0.879	0.071	0.864
N	1667	1664		1664	

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. "Amount Paid per Acre" is the per acre payment in ₹ that the farmer received. This includes the amount paid upfront for those in the upfront PES treatment, plus the amount paid conditional on compliance for those in the Upfront and standard PES treatment. "Not Burned" takes the value 1 if the farmer did not burn any of their plots, and 0 otherwise, and matches the estimates in Table 3. "Cost per Unburned Acre" is the "Amount Paid per Acre" divided by the "Cost per Unburned Acre" (column 1 divided by column 2 or column 4). Standard errors in columns 3 and 5 are calculated using the delta method.

Appendices

A.1 Appendix Tables and Figures

Table A.1: Comparison of Study Sample, Cooperative Listing, and Census Sample

	Census	Cooperative Members	Study Eligible	Study Enrolled	Diff Coop. - Census	Diff Coop. - Study
	(1)	(2)	(3)	(4)	(5)	(6)
Age (years)	46.79	46.99	46.53	48.34	0.20	1.35
	(14.31)	(14.31)	(14.53)	(12.82)	[0.85]	[0.54]
Total experience in agriculture (years)	25.66	26.29	25.05	28.92	0.63	2.63
	(14.58)	(14.72)	(14.25)	(13.39)	[0.54]	[0.26]
Total area of paddy land in acres (reported)	7.71	7.99	5.54	5.26	0.28	-2.73***
	(8.03)	(7.69)	(2.93)	(2.47)	[0.62]	[0.00]
1(Knowledge of CRM techniques)	0.87	0.89	0.86	0.79	0.02	-0.10
	(0.33)	(0.32)	(0.35)	(0.41)	[0.50]	[0.16]
1(Tried a CRM technique (oth. th. burning))	0.90	0.90	0.85	0.74	-0.00	-0.16*
	(0.30)	(0.31)	(0.36)	(0.45)	[0.89]	[0.04]
Distrust index excluding distrust in family (continuous)	-0.01	-0.32	-0.03	0.80	-0.31	1.12
	(3.49)	(3.43)	(3.57)	(3.92)	[0.20]	[0.10]
1(Aware of government PES program)	0.37	0.36	0.38	0.31	-0.01	-0.06
	(0.48)	(0.48)	(0.49)	(0.47)	[0.80]	[0.57]
1(Applied to government PES program 2019)	0.19	0.18	0.16	0.19	-0.01	0.01
	(0.40)	(0.39)	(0.37)	(0.40)	[0.68]	[0.91]
Observations	479	339	190	38		

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard deviations reported in parentheses and standard errors reported in brackets. Column 1 includes the sample of respondents in the census survey; column 2 includes the subgroup of participants in the census survey who are part of the local farmers' Cooperative Society; column 3 restricts the census sample to those respondents who would have been eligible for the baseline survey of the RCT; column 4 includes the sample of census respondents in the RCT. Columns 1 to 4 are the means in the samples, and columns 5 and 6 are the differences between the means.

Table A.2: Remote Sensing Model Accuracy in Holdout Sample

	Not Burned	
	Balanced Accuracy (1)	Maximum Accuracy (2)
Mean accuracy	0.78	0.82
False burn	0.08	0.13
False no burn	0.14	0.05
True burn	0.51	0.59
True no burn	0.27	0.22
No burn accuracy	0.76	0.63
Burn accuracy	0.79	0.92

Note: Accuracy statistics for remote sensing measures of burning, using different classifications thresholds. The true/false burn/no-burn rows show counts of the number of fields in each category. $N = 681$.

Table A.3: Attrition from the Endline Survey

	Attrition (1)
Standard PES	0.051
	(0.025)**
Upfront PES	0.044
	(0.025)*
p-val: Standard PES = Upfront PES	0.786
Control group mean	0.130
Standard PES group mean	0.187
Upfront PES group mean	0.182
N	1668

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. The outcome variable is a dummy that takes the value 1 if the respondent attrited from the endline.

Table A.4: Heterogeneity of Attrition from the Endline Survey (one regression)

	Age	Agric. Exp.	Educ.	Ever Signed Contract	Income	Non-Agric. Income	Agric. Revenue	Land Area	Paddy Prod.	Financial Const.	Distrust	Info. Const.	Access Const.	Neg. Beliefs	Burned Paddy Residue in 2018	Not Burned (Balanced)	Not Burned (Maximum)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Standard	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096
	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)
Upfront	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033
	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)	(0.197)
Het. Var.	-0.001	0.000	0.004	0.037	0.001	0.007	-0.001	0.003	0.004	0.054	-0.002	0.031	0.007	-0.043	-0.021	-0.064	0.212
	(0.003)	(0.003)	(0.004)	(0.036)	(0.001)	(0.003)	(0.002)*	(0.009)	(0.003)	(0.042)	(0.035)	(0.031)	(0.043)	(0.038)	(0.054)	(0.042)	(0.092)
Standard x Het. Var.	0.003	-0.002	-0.011	0.022	0.004	-0.013	-0.000	-0.012	-0.005	0.008	0.060	0.002	-0.049	0.061	0.011	-0.027	-0.127
	(0.005)	(0.004)	(0.007)	(0.055)	(0.002)	(0.004)**	(0.002)***	(0.013)	(0.004)	(0.056)	(0.055)	(0.046)	(0.06)	(0.063)	(0.072)	(0.07)	(0.126)
Upfront x Het. Var.	0.001	0.001	0.003	-0.013	-0.000	-0.008	0.001	-0.003	-0.002	0.024	-0.058	0.019	-0.032	0.031	0.060	0.071	-0.149
	(0.004)	(0.004)	(0.006)	(0.053)	(0.002)	(0.004)	(0.002)*	(0.012)	(0.004)	(0.058)	(0.049)	(0.05)	(0.058)	(0.054)	(0.071)	(0.072)	(0.121)
P-value test Standard x Het. Var. = Upfront x Het. Var.	0.705	0.576	0.066	0.545	0.096	0.060	0.205	0.465	0.401	0.767	0.033	0.736	0.758	0.641	0.457	0.226	0.852
Control group mean	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124
Standard group mean	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192
Upfront group mean	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180
N	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. "Standard" is the standard PES group and "Upfront" is the upfront payment PES group. The outcome variable is a dummy that takes the value 1 if the respondent attritted from the endline. Het. Var. is the heterogeneity variable. The column titles indicate of which heterogeneity variable the coefficients are displayed in the column. "Agric. Exp." refers to the total experience in agriculture (years). "Educ." refers to the highest educational class passed. "Ever Signed Contract" is a dummy taking value 1 if the farmer ever signed a written contract before and 0 otherwise. "Income" refers to the total income in ₹'s in the past 12 months. "Non-Agric. Income" refers to non-agricultural income in ₹in the past 12 months. "Agric. Revenue" refers to the total revenue from agriculture in 1000 ₹. "Land Area" refers to the total area of land in acres. "Paddy Prod." refers to the paddy production in 1000kg. "Financial Const." refers to a financial constraints index. "Distrust" refers to an index indicating the farmer's distrust in categories of people and organizations. "Info. Const." refers to a CRM information constraints index. "Access Const." refers to a CRM access constraints index. "Neg. Beliefs" refers to a CRM negative beliefs index. "Not Burned (Balanced)" refers to the remote sensing measure of not-burning using the balanced accuracy threshold, and "Not Burned (Maximum)" refers to the remote sensing measure of not-burning using the max accuracy threshold.

Table A.5: Treatment Effects Disaggregated by Subtreatment

	Contract Take-Up	Complied with contract	Not Burned: Balanced Accuracy	Not Burned: Max Accuracy
	(1)	(2)	(3)	(4)
800/acre	0.743	0.068	-0.028	0.007
	(0.030)***	(0.016)***	(0.046)	(0.033)
1600/acre	0.707	0.104	0.048	0.036
	(0.036)***	(0.025)***	(0.053)	(0.038)
800/acre with 25% Upfront	0.737	0.177	0.114	0.056
	(0.030)***	(0.029)***	(0.048)**	(0.032)*
800/acre with 50% Upfront	0.702	0.189	0.115	0.094
	(0.029)***	(0.029)***	(0.053)**	(0.046)**
p-val: 800/acre = 1600/acre	0.441	0.219	0.167	0.461
p-val: 800/acre = 800/acre with 25% Upfront	0.885	0.001	0.006	0.170
p-val: 800/acre = 800/acre with 50% Upfront	0.310	0.000	0.008	0.059
p-val: 1600/acre = 800/acre with 25% Upfront	0.510	0.051	0.249	0.617
Control mean	0	0	0	0
N	1668	1668	1664	1664

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects included. "Contract Take-Up" is a dummy variable taking the value 1 if the respondent signed a contract to participate in the PES program. "Complied with Contract" is a dummy variable taking the value 1 if the monitoring led to the conclusion that the respondent complied with the contract i.e. did not burn any fields. "Not Burned" takes value 1 if the farmer did not burn any of his plots, and 0 otherwise, using the balanced accuracy which classifies burning to balance type I and type II errors. "800/acre" is the standard PES group receiving ₹800 per acre conditional on not burning. "1600/acre" is the standard PES group receiving ₹1600 per acre conditional on not burning. "800/acre with 25% Upfront" is the upfront PES group receiving ₹800 per acre of which they receive 25% unconditionally upfront and 75% conditional on not burning. "800/acre with 50% Upfront" is the upfront PES group receiving ₹800 per acre of which they receive 50% unconditionally upfront and 50% conditional on not burning.

Table A.6: Treatment Effects on Remote Sensing Not-Burning Measure (Plot Level)

	Not Burned	
	Balanced Accuracy (1)	Maximum Accuracy (2)
Standard PES	0.003	0.022
	(0.046)	(0.036)
Upfront PES	0.123	0.101
	(0.045)***	(0.036)***
p-val: Standard PES = Upfront PES	0.005	0.023
Control group mean	0.296	0.150
Standard PES mean	0.284	0.154
Upfront PES mean	0.432	0.265
N	2875	2875

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. Plot level regressions are weighted by the inverse of the number of plots the farmer has. "Not Burned" takes value 1 if the farmer did not burn any of his plots, and 0 otherwise. Column 1 classifies burning to balance type I and type II errors; column 2 classifies burning to maximize overall model accuracy.

Table A.7: Treatment Effects Based on Spot Checks

	Not Burned (1)
Standard PES	0.014
	(0.077)
Upfront PES	0.105
	(0.073)
p-val: Standard PES = Upfront PES	0.233
Control mean	0.371
Standard PES mean	0.364
Upfront PES mean	0.456
N	715

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. "Not Burned" takes value 1 if the farmer did not burn any of his plots according to the spot checks, and 0 otherwise.

Table A.8: Effects on Agricultural Yield and Sowing Delays

	Paddy Yield (1)	Wheat Yield (2)	Days (3)
Standard PES	-0.026	-0.013	-0.217
	(0.039)	(0.015)	(0.643)
Upfront PES	-0.066	0.008	-0.120
	(0.045)	(0.015)	(0.627)
p-val: Standard PES = Upfront PES	0.356	0.151	0.881
Control mean	1.249	0.745	18.364
Standard PES mean	1.237	0.736	17.943
Upfront PES mean	1.194	0.756	18.380
N	1367	1378	1386
Lee Bounds			
Standard PES			
Lower bound	-0.103	-0.035	-1.032
	(0.037)***	(0.013)***	(0.552)*
Upper bound	0.037	0.008	0.320
	(0.037)	(0.015)	(0.524)
Upfront payment PES			
Lower bound	-0.113	-0.013	-0.937
	(0.037)***	(0.010)	(0.532)*
Upper bound	0.005	0.027	0.425
	(0.037)	(0.012)**	(0.505)

Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. "Paddy Yield" is a variable indicating the amount of paddy produced in Kharif 2019 (log of 1000 kg per acre). "Wheat Yield" is a variable indicating the amount of wheat produced in Rabi 2020 (log of 1000 kg per acre). "Days" is a variable indicating the number of days after the paddy harvest that passed before the farmer started sowing the Rabi crop. The second panel shows the Lee bounds for the treatment effects by treatment group. For details on the Lee bounds, see footnote 33.

Table A.9: Heterogeneity of Pooled Treatment Effects on Contract Compliance by CRM Equipment Constraints

Outcome variable:	Complied with Contract		
Type of constraint:	Information Constraints	Access Constraints	Negative Beliefs about Burning Alternatives
	(1)	(2)	(3)
Highly constrained	-0.063	-0.004	-0.040
	(0.019)***	(0.024)	(0.020)**
Pooled PES mean	0.135	0.136	0.135
N	1182	1168	1182

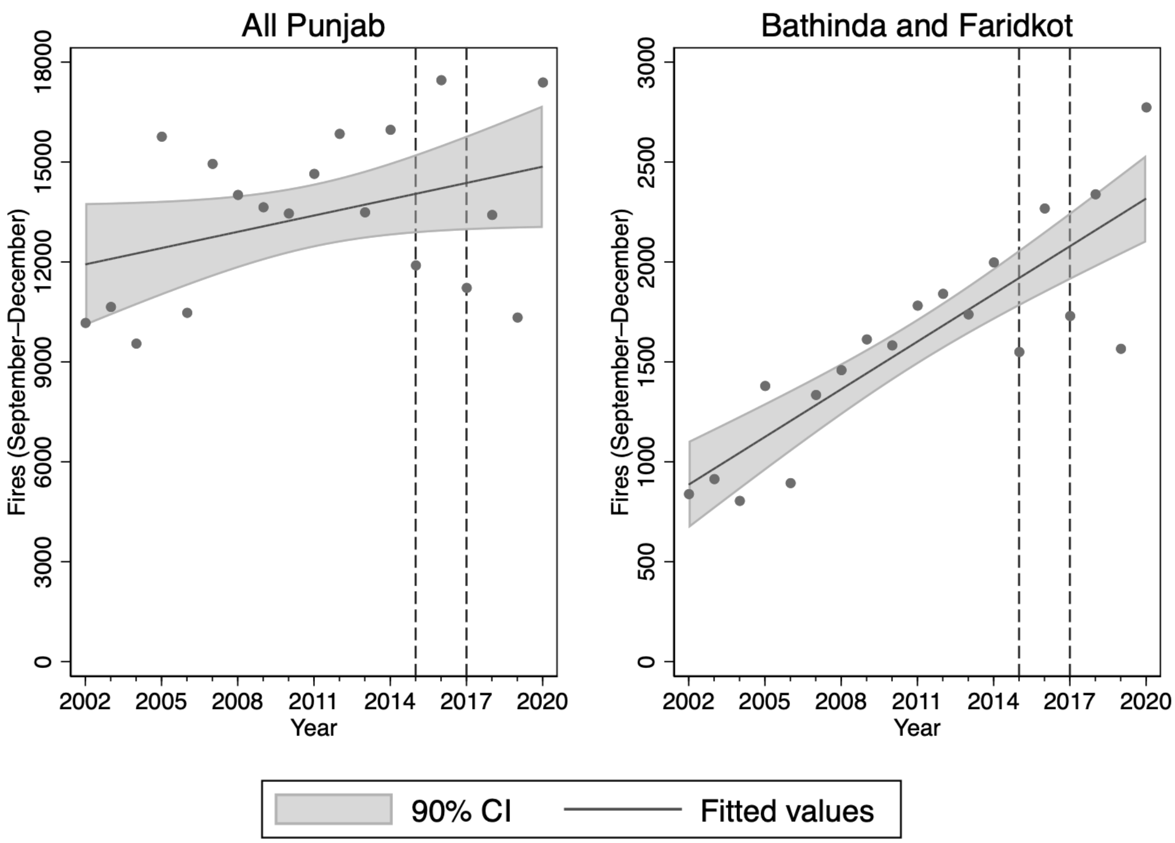
Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. The row labled "Type of constraint" indicates the heterogeneity variable: "Information Constraints" is an index indicating the farmer's lack of knowledge about CRM equipment. "Access Constraints" is an index indicating the farmer's difficulties in accessing CRM equipment. "Negative Beliefs about Burning Alternatives" is an index indicating the strength of the farmer's negative beliefs about the impact of CRM equipment on soil health and yield as compared to burning. All indices are binary and take value 1 if the farmer's constraints are larger than or equal to the median. The outcome variable is indicated in the top row: "Complied with Contract" is a dummy variable taking the value 1 if the monitoring led to the conclusion that the respondent complied with the contract i.e. did not burn his paddy residue.

Table A.10: Heterogeneity of Treatment Effects on Contract Take-Up

Outcome variable:	Program Take-Up				
Type of constraint:	Liquidity	Distrust	Information	Access	Negative Beliefs about Alternatives
	(1)	(2)	(3)	(4)	(5)
Upfront PES	0.004	-0.018	0.011	-0.014	-0.019
	(0.040)	(0.041)	(0.041)	(0.042)	(0.045)
Highly constrained	0.020	0.024	-0.057	-0.020	-0.041
	(0.036)	(0.042)	(0.047)	(0.030)	(0.046)
Upfront PES × Highly constrained	-0.015	0.035	-0.044	0.015	0.023
	(0.051)	(0.059)	(0.058)	(0.048)	(0.061)
Pooled PES mean	0.734	0.735	0.734	0.735	0.734
N	1182	1172	1182	1168	1182

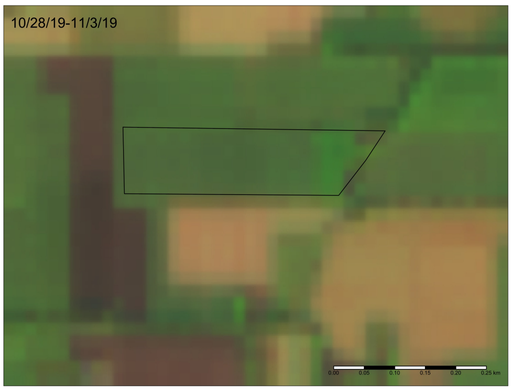
Note: ***(**)(*) indicates significance at the 1%(5%)(10%) level. Standard errors in parentheses clustered at the village level. Strata fixed effects are included. The row labled "Type of constraint" indicates the heterogeneity variable: "Liquidity" is an index indicating liquidity constraints, including constrained access to cash and loans. "Distrust" is an index indicating the farmer's distrust in categories of people and organizations. "Information" is an index indicating the farmer's lack of knowledge about CRM equipment. "Access" is an index indicating the farmer's difficulties in accessing CRM equipment. "Negative Beliefs about Alternatives" is an index indicating the strength of the farmers negative beliefs about the impact of CRM equipment on soil health and yield as compared to burning. All indices are binary and take value 1 if the farmer's constraints or strength of beliefs is larger than or equal to the median. The outcome variable is indicated in the top row of the table: "Program Take-Up" is a dummy variable taking the value 1 if the respondent signed a contract to participate in the PES program. The comparison group is the standard PES group.

Figure A.1: Time Trends in Fires Based on MODIS Satellite Data

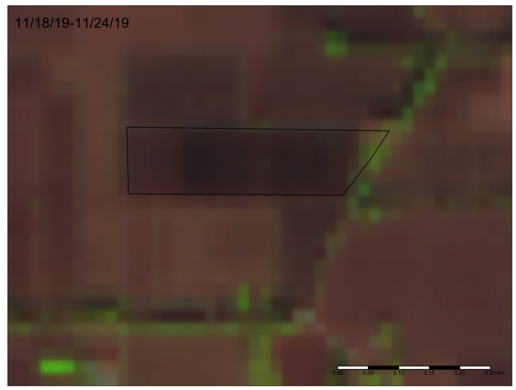


Note: Fire counts from September to December by year, based on MODIS imagery. Left panel shows the state of Punjab; right panel shows study districts. The line at 2015 indicates the introduction of the burning ban; the line at 2017 indicates the introduction of the two-year CRM subsidy.

Figure A.2: Visual Signs of Burning in Imagery: Example



Around harvest time

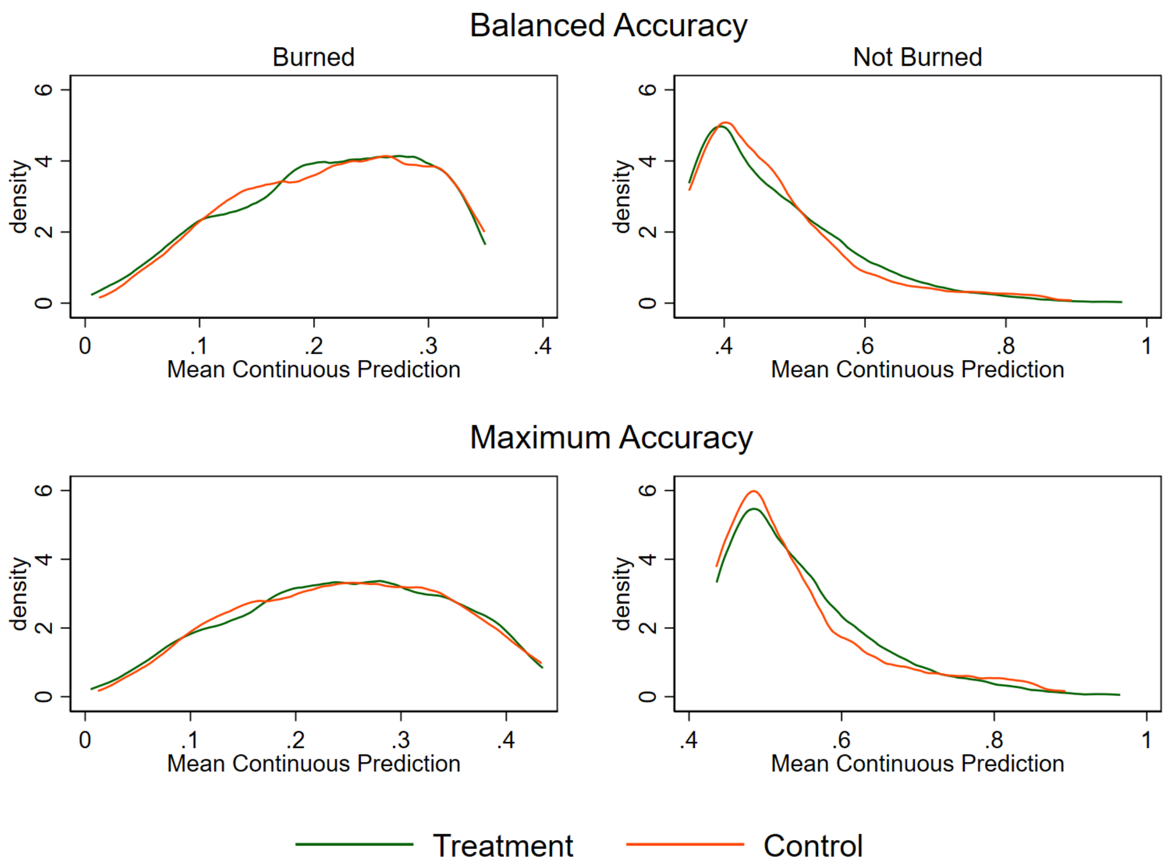


Just after burning

One week after burning

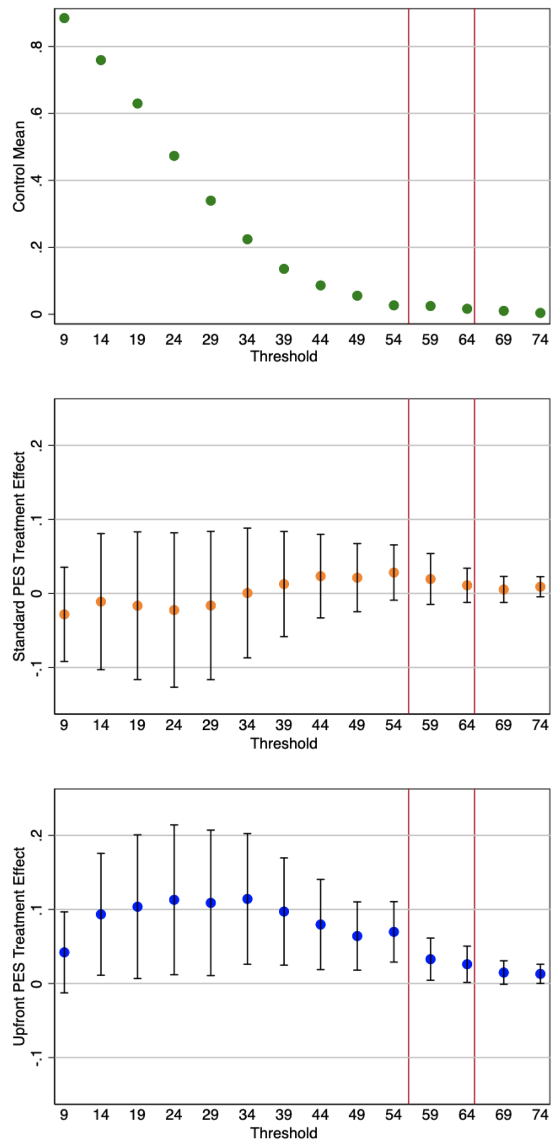
Note: Imagery from Sentinel-2 showing a study field.

Figure A.3: Distribution of Random Forest Predictions by Treatment



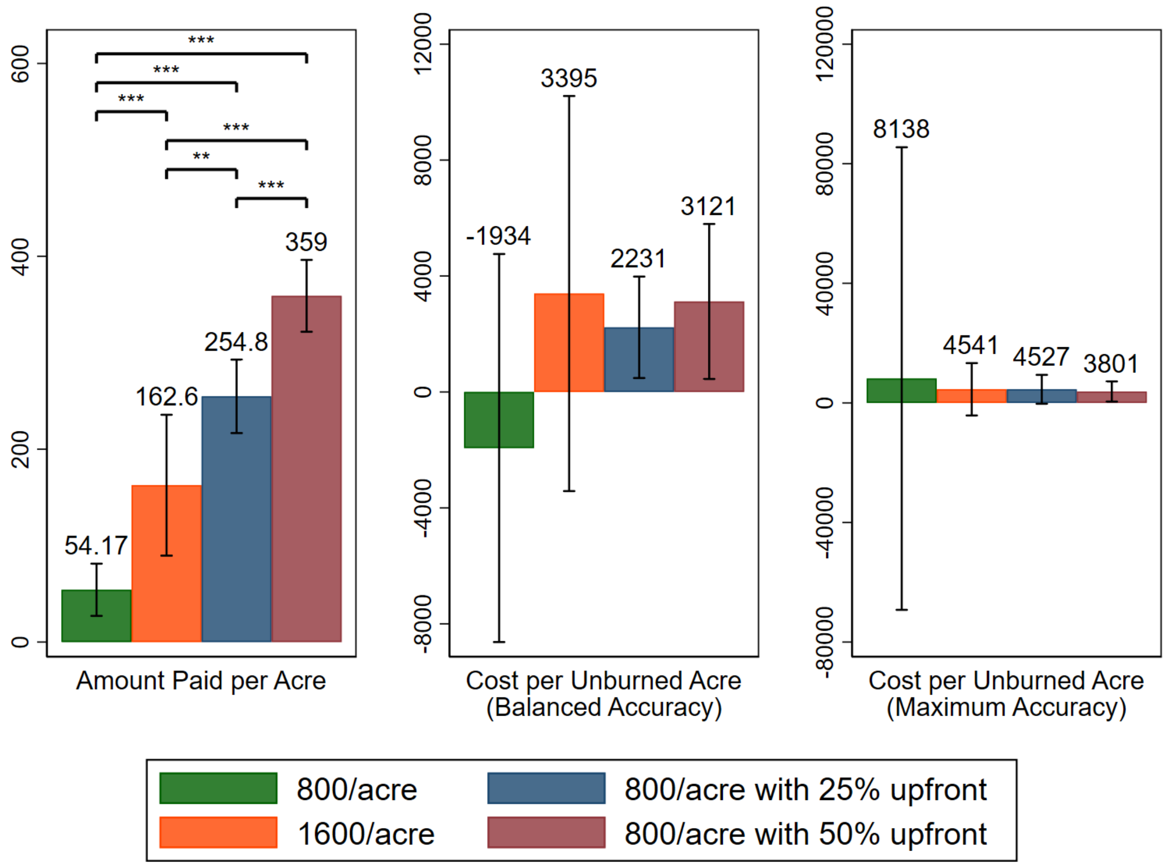
Note: The left panel shows the distribution of the continuous remote sensing measure of not-burning for the plots classified as having been burned. The right panel shows the distribution of the same measure for plots classified as not having been burned. The classification uses the balanced accuracy, and restricts to plots not included in the training data. The continuous remote sensing measure ranges from 0 to 1, where higher values mean that it is more likely that the plot has not been burned.

Figure A.4: Robustness of Treatment Effects on Not Burning



Note: The graphs show the control mean and treatment effects for binary remote sensing measures of not-burning based on different classification thresholds. The classification thresholds are indicated on the x-axis. The binary remote sensing measures of not-burning take value 1 if the farmer did not burn any of his plots, and 0 otherwise. The first graph shows the mean of the remote sensing measure of not-burning in the control group. The second graph shows the treatment effects on not-burning in the Standard PES treatment arm. The third graph shows the treatment effects on not-burning in the Upfront PES treatment arm. The two red lines in the second and third graph indicate the thresholds used for the remote sensing measures of not-burning with balanced accuracy which classifies burning to balance type I and type II errors, and with maximum accuracy which classifies burning to maximize overall model accuracy. The remote sensing measure of not-burning with balanced accuracy uses a threshold of 56 and the one with maximum accuracy uses a threshold of 65.

Figure A.5: Cost Effectiveness by Subtreatment



Note: Coefficients on top of bars. ***(**)(*) indicates significant difference between coefficients at the 1%(5%)(10%) level. "Amount Paid per Acre" is the payment in ₹the farmer received (it is the total amount for those who were monitored and complied (i.e. did not burn), it is the upfront amount for any respondents in the treatment groups with upfront component who participated but did not call for monitoring/burned, and it is 0 for respondents in treatment groups without upfront component who did not call for monitoring/burned, for respondents in any treatment groups who did not participate, as well as for respondents in the control group). "Cost per Unburned Acre" the amount paid per acre divided by the treatment effect.

A.2 Intervention Script

Program description

Our organization is working on agricultural and environmental issues and we want to help farmers manage paddy stubble after the paddy harvest this season. I am here to share details of a program that we are introducing to some farmers in this village during the paddy crop season in the month of October and November 2019.

To encourage farmers to manage paddy stubble in an environmentally-friendly manner, we will offer you an agreement that will pay you if you do not burn your paddy field(s) this season. We will compensate you at a rate of [treatment rate] per acre (up to a max. of [treat rate x 100]). You may use any alternate methods of managing the residue. Other than burning the stubble, we do not place any condition on what this method should be.

This monetary compensation will only be given to you if a monitor, during the months of October and November, assesses that that your paddy field has not been burned. If you are interested in participating, I will explain the terms and conditions of the agreement to you that will help you decide whether you want to enroll in the programme or not. If you are uncertain about signing the agreement because you are unsure whether you would be fulfilling the conditions of the contract, let me remind you that there is no harm in participating in the programme. If you burn, you will not be penalized in any way by us. If you do not burn, you will be given the reward. By signing you are only giving yourself a chance to win money.

If you would like someone in the house to help you make a decision and listen to the details of the programme, please feel free to invite them now. Please remember that whoever signs the agreement must have a bank account to enable payments at a later date.

Information handout

This document provides details on some of the items in the agreement and is to help the enrollee farmer with complying with the terms and conditions of the agreement.

Monitoring visits

1. The enrollee farmer is expected to initiate monitoring for all plots, with a maximum of two requests to J-PAL. All plots must be covered through these two requests.
2. Each request will result in up to two visits by J-PAL monitors. The second visit will only be performed if J-PAL determines that it is necessary to assess burning.
3. In addition to the requested and scheduled visits, J-PAL can also make unannounced visits to the plots for checks.

When to call for monitoring: The enrollee farmer should call once all the pre-sowing work related to stubble has been completed on the plots covered under the request. This means all activities related to stubble like removing or processing of stubble must have been completed and no further managing of stubble is required before sowing. In general, requests should be made at least four days before sowing. The request can occur if any of the following applies:

1. After the straw and stubble have been completely removed from the plot but no later than 4 days before sowing.
2. After the straw has been rolled into bales/bundles but no later than four days before sowing.
3. After the straw/stubble has been mixed or blended into the soil but no later than four days before sowing.
4. If using the Happy seeder or mulcher: once sowing preparation is complete but no later than four days before sowing. In these cases, a second monitoring visit will be made post sowing.

Remember, up to two requests can be initiated. If some plots are ready, call to schedule the first monitoring, keeping in mind that any plots not covered under the first request have to be monitored as part of the second request. If all plots are ready for monitoring, they can be inspected in a single visit.

Phone numbers for calling: xxx, xxx, xxx

What counts as burning? The agreement requires that farmers do not burn any of their plots. This will be broken if any of the following (or any other form of burning) are detected by the monitor. The farmer will not be eligible for payment if any of the following is detected during monitoring.

- Burning of the upper layer of loose straw left behind by the harvester.
- Burning of the standing stubble.
- Burning of straw collected in one part of the plot.
- Burning as mentioned above on any of the plots.

Important

1. J-PAL SA is not related to any government in any manner. The failure of the enrollee farmer to meet any term or condition in the agreement will not attract any penalty or fine, and no legal action will be taken. This is clearly stated in the agreement. We are only trying to find if this a good way to help the farmers with resolving the residue issue. We cannot impose any fines or penalties since we are not related to government.
2. The only consequence of not fulfilling any of the term or condition in the agreement will be that farmer will become ineligible for payment of amount as mentioned in the agreement.
3. In case the farmer does not request monitoring as specified above, J-PAL will not be liable to pay any amount as mentioned in the agreement. Decision on payment to be made will only be taken once all the plots have been fully monitored.
4. If after the first monitoring visit and after analyzing the observations recorded, the J-PAL SA team ascertains that burning happened in even one of the plots, no further monitoring visit will be conducted. In this case, the farmer will be ineligible to receive the payment.
5. At the time of the monitoring visit, we may also request you for bank account details. The bank account transfer is the fastest and easiest way to transfer the amount. After

the monitoring has been completed for all the plots and it is assessed that burning has not happened on any plots, the payment will be made directly into the account.

6. The enrollee farmer should keep the agreement and information handout safe for use later. The ID and phone numbers given on them are to be used for calling.

A.3 Sample Contract

${village_id}

${a_hhid}

${resp_id}

Contract for Incentive Program Offering Payment for No-Burning on Paddy Plots

This Agreement is executed on ___________________[Insert date]

by and between ${resp_name},

residing at

____[Insert Enrollee Address]

AND

Abdul Latif Jameel Poverty Action Lab South Asia at the Institute for Financial Management and Research, which is registered under Society Registration Act 1860 (hereinafter referred to as "J-PAL SA"), located at Buhari Towers, 2nd Floor, 4, Moors Road, Chennai 600006

Background

J-PAL SA proposes to partner with [${resp_name}] (hereinafter referred to as "Enrollee") with the following summary of responsibilities.

Based on the field measurement completed in a previous visit, (s)he cultivates _#ACRE____ acres of paddy currently.

Summary of responsibilities

J-PAL SA

1. Visit Enrollee's paddy plots, which were mapped during the survey visit to the Enrollee that was already conducted, to assess whether burning occurred. This monitoring visit will take place once Enrollee informs J-PAL SA by phone, as described below. J-PAL will visit the plots to assess whether they have been burned within **3** days of being called by the Enrollee. 'Monitoring will be available only beginning on 15 October 2019 or today (whichever date is later). Enrollees that call to be monitored before this date cannot be monitored by J-PAL South Asia and therefore are not eligible for payment.
2. The J-PAL SA team will determine if the field has been burned based on the observations made by the monitor during their visit. The process of inspection is summarized below:
    a) The J-PAL SA monitor will visit all the paddy plots as measured during a previous visit.
    b) The monitor will physically inspect each plot for visual cues and record the observations. Based on the recorded observations during the visit, the J-PAL SA team will determine whether the field was burned or not.
3. If the paddy plots do not appear to be burned, as assessed by the J-PAL SA team, then J-PAL SA will provide Enrollee with an amount such that the total payment amount for not burning is Rs **800** per acre of enrolled land. The maximum overall payment is Rs **8000**. The payment amount for the Enrollee is Rs ${pes_amount}.

Enrollee

1. Enrollee confirms, by signing this agreement, that the paddy plots mapped during the survey visit represent all of his/her paddy plots. All paddy plots cultivated in the 2019 Kharif season must be enrolled.
2. After harvesting paddy and managing and processing stubble, and at least 4 days before sowing wheat or any other rabi crop, Enrollee is required to call J-PAL SA at the numbers provided on the information handout between the hours of 9:00 am and 5:00 pm, on any date between 15 October and 30 November 2019 to indicate that the fields are ready to be monitored. We will not be able to monitor before the above mentioned date and farmers requesting for monitoring to be conducted before the 15th October will not be eligible.
3. The Enrollee may request up to two monitoring visits to cover all paddy plots, for example, for some plots that are ready for monitoring early and others that are ready late. Each plot will be monitored up to two times.
4. The Enrollee will also allow additional, unscheduled monitoring to occur at any point in time.

5. If it is assessed by the J-PAL SA team that the field is not burnt, the Enrollee will receive a payment amount as indicated above. For the enrollee to be eligible for payment, no burning should have taken place on any of the plots.
6. The assessment of whether a field is burnt or not is not dependent on whether the field was burnt deliberately or accidentally, or by the Enrollee or someone else.

Payment and contract

1. J-PAL SA shall not be obligated to pay the Enrollee any amount in excess of what is mentioned above.
2. By signing this agreement, the Enrollee acknowledges that J-PAL SA reserves the right to rescind the payment of the aforementioned amount if the Enrollee fails to fulfil any of the responsibilities designated to him/her under "Summary of Responsibilities" and/or breach of the terms of this agreement in any manner or extent.
3. There will be no legal implications for the Enrollee for the breach of the agreement. J-PAL SA will not take any legal action against the Enrollee if one or more responsibilities remain unfulfilled under the agreement.

ACCEPTED BY:	ACCEPTED BY:
J-PAL SA	Enrollee
Signature	Signature
Name	Name
Location	Location
Date	Date

A.4 Remote Sensing Model

This section provides additional detail on the construction of our remote sensing based outcome. For a complete description, please see Walker et al. (2022).

Model background: The goal of the model is to detect whether a plot in our sample was burned at any point during the burn season (from October 10 to December 15, 2019) based on satellite imagery. While burn scars are obvious if the plot is observed by satellite soon after burning, this signal erodes quickly with time. With a temporal resolution of about two days, PlanetScope imagery can often capture burned plots within this critical window. However, clouds and other abnormalities result in a maximum gap between any two images of 8 days, on average across plots in our sample. While Sentinel-2 imagery has a coarser temporal resolution of about eight days, it provides mid- and short-wave infrared (SWIR) bands that are able to detect signals of burning for a longer window post-burn. By combining observations from both sensors, we built a Random Forest (RF) model with an overall accuracy of 82% in detecting burning in smallholder rice plots.

Other studies have relied on burn detection based on active fires, using, for example, data from the Visible Infrared Imaging Radiometer Suite (VIIRS). The sensor has a spatial resolution of 375m, resulting in pixels that are around 140,000 m^2. A typical plot in our sample is around 10,000 m^2, and only a small share of farmers in a village are enrolled in the study, so existing active fire products are poorly suited to our measurement goals.

An overview of image processing for both types of satellite is as follows:

Imagery and image processing overview:

- PlanetScope: Four-band harmonized surface reflectance product from PlanetLabs
  - Resolution: Spatial: 3m, Temporal: 2.2-day on average (30-40 images per pixel)
  - Spectral bands: blue, green, red, Near Infrared (NIR)
  - Clouds: only included images with <10% cloud cover. Remaining clouds were masked using the unusable data masks (UDM2) provided with the imagery.

- Pre-processing: atmospheric correction based on the 6SV2.1 radiative transfer code already applied to product. Harmonized product also incorporates data from Sentinel-2 to normalize the spectral response functions between sensors.

- Sentinel-2: Level-1C products from USGS, converted to surface reflectance
  - Resolution: Spatial: 10m for visible and NIR bands, 20m for shortwave infrared (SWIR) bands. Temporal: 7-8 days on average
  - Spectral bands: Blue, Green, Red, NIR, SWIR1, SWIR2
  - Clouds: Cloudless layers from Google Earth Engine with cloud probabilities $\leq .5$ cloud were used as initial masks, then inspected and expanded manually to remove remaining cloud shadows.
  - Pre-processing: Geometric and radiometric corrections applied as Level-1C product, converted to bottom-of-the atmosphere reflectance with SNAP toolkit.

Feature creation and selection: As model inputs, we used individual bands and derived indices aimed at reducing noise and amplifying the portion of the spectrum most associated with burning. These indices were taken from the literature on burn mapping with a focus on char detection rather than vegetation change, as our primary separation task is between bare soil (harvested and often tilled plots) and charred soil (burnt plots). For PlanetScope images, we used the Bare Soil Index (BSoI), which uses all four bands, the Char Index (CI), which uses all visible bands, and the Burn Area Index (BAI), Simple Ratio (SR) and NDVI, which use the red and NIR bands. For Sentinel-2 images, we also included several bands using one or both SWIR bands including the Burn Scar Index (BSI), Mid-Infrared Bispectral Index (MIRIBI), and two variations of the Normalized Burn Index (NBR and NBR2). See Walker et al. (2022) for background and equations.

We stacked all images that overlap with any of the study participants' rice plots into a time-series and created pixel-level features based on statistics from each band and index across time. Statistics included min, max, median, and outer percentiles. An additional temporal differencing measure (Vdiff) was calculated for each band and index with the goal of capturing the moment the pixel changed from unburned to burned. This Vdiff measure was calculated based on the largest drop (or spike) in the sequence of values (V) for V_{t+1} - V_t. We used SequentialFeatureSelector in the sklearn toolkit in Python to reduce the feature space to an optimal number of features (around 30) prior to the final analysis. Retained features are presented in Walker et al. (2022).

Recognizing that pixels along the edge of a plot likely present differently due to the mixture of plot/non-plot classes and different burn patterns at edges, we flagged border pixels. These pixels were observed to have low importance in the construction of the RF model and were thus dropped from our analysis.

Model training and assessment: Training data consists of 441 burned and 240 unburned labels collected on the ground from participant farmers in 2019. Unburned labels come from plots where participants invited a monitor to visit to confirm that the stubble was managed without burning. Burned labels come from observations during unannounced spot checks of participant plots.

We used pixel-level features from the 681 labeled plots to train a RF model to provide burn predictions. Although data was retained at the pixel level, full plots were held out from the training data for use in optimization and accuracy assessment. Plot-level holdouts were necessary because pixels within the same plot have highly correlated features; if some pixels within a plot were used for training while others were used for testing, overfitting of the model and overestimation of accuracy would occur. A single plot was held out each time while a RF model was generated with the remaining 680 plots. This process was repeated 680 times in a Leave-One-Out Cross-Validation (LOOCV) format. Model accuracy was assessed based on the prediction score for each plot in the run where it was left out of model training.

To convert from pixel to plot-level predictions, we aggregated on the plot-level mean of the continuous RF output (we also tried the median and various percentiles and found the mean to perform best). We then used two approaches to set the classification thresholds based on this mean score, with plots exceeding the threshold classified as burned. First, we maximized overall accuracy ("max accuracy") by iterating over each threshold percentile and selecting the threshold with the highest accuracy for the full labeled set of plots. Alternatively, to balance accuracy across burned and not-burned labels ("balanced accuracy"), we iterated the burn accuracy and the no-burn accuracy over each threshold percentile, interpolated these accuracies into smooth functions, and selected the percentile threshold with the greatest accuracy for the mean at the point of intersection (where burned accuracy equals unburned accuracy). We tested using Cohen's Kappa for threshold optimization, which measures how a classifier compares when evaluated against a random classifier. In this case, maximizing kappa resulted in the same threshold selection as the max accuracy approach for all versions of our model.

Following plot-level aggregation, our best RF model achieves 82 percent overall accuracy, with 91 percent accuracy in detecting burned plots but only 63 percent accuracy in detecting unburned plots (details in Walker et al. (2022)). When the burned/unburned errors are balanced with our balanced accuracy procedure, the overall accuracy is reduced to 78 percent.

A.5 Survey Questions used in Constructing Indices for Heterogeneous Treatment Effects

This section details the (pre-specified) survey questions used in constructing the indices for heterogeneous treatment effects.

Financial constraints

1. If you needed to spend ₹5000 for agricultural equipment, would you have savings to draw on?

2. If you needed to spend ₹10,000 for agricultural equipment, would you have savings to draw on?

3. If you needed to spend ₹5000 for agricultural equipment, how easy would it be for you to get a loan for that amount?

4. If you needed to spend ₹10,000 for agricultural equipment, how easy would it be for you to get a loan for that amount?

These (standardized) variables are used to create an index, which is used to create a binary variable split at the median to denote high financial constraints.

Distrust

1. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

2. I'd like to ask you how much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all?

    - People in your neighborhood?
    - Strangers?

3. Even if you have had very little or no contact with these following institutions, please base your answer on your general impression of these institutions.

   - The Punjab Government?
   - The village Panchayat?
   - The cooperative society?
   - Non-governmental organizations (NGOs)?
   - Financial Institutions like Banks/Insurance Companies?

CRM access barriers indices

We construct three indices to measure different aspects of CRM equipment access barriers. The first measures information constraints, the second access barriers, and the third beliefs about how CRM equipment impacts agriculture relative to burning. All questions except the first are asked about the CRM equipment farmers reported being familiar with.

Information Constraints

1. Do you know about any crop residue management techniques to manage paddy stubble?
2. Where can you rent it (CRM equipment) from?

Access Barriers

1. Do you own [CRM equipment] as an individual or member of a CHC or Coop?
2. Is using [CRM equipment] more expensive or less expensive than burning paddy stubble?
3. In days, how long would it take you to access crop residue management equipment for managing paddy stubble at harvest time this year?
4. Including all costs, how much would the equipment cost per acre (in Rs.)?
5. How many days would it take to manage paddy stubble using this equipment?

Negative Beliefs About CRM Equipment

- Is using [CRM equipment] better for long-term soil health or worse for soil health than burning paddy stubble?
- Does using [CRM equipment] help yield of rabi season or hurt yield of rabi season compared to burning paddy stubble?