What happens when new technology is not good enough? Innovation frictions in developing countries' agriculture

Sergio Puerto, Cornell University, sap257@cornell.edu

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25, 2023

What happens when new technology is not good enough? Innovation frictions in developing countries' agriculture

Sergio Puerto

This is a preliminary draft. Please do not cite.

This version: May 17, 2023

Abstract

Innovation is crucial for farm productivity and economic growth in low-income countries, but poor farmers are still reluctant to adopt seemingly profitable new inputs and practices. A possible explanation is that the absence of mechanisms crowding out bad agricultural technology allows innovators and input suppliers to remain in low targeting efficiency equilibria, hindering technological change. To test this, I conducted a two-stage randomized control trial to estimate the effect of targeting frictions on farmers' adoption of improved seeds. The experiment recreated counterfactual scenarios for innovators' seed development decisions, in which some farmers received an offer to purchase the new variety they preferred the most, whereas others were offered a recommended but unknown new variety. Results indicate that preference-driven mismatch significantly reduces farmers' adoption of new technology. Take-up among farmers offered the recommended variety is 50% lower than targeted farmers. Further, I find no impact of learning-by-doing and biased performance beliefs on take-up, but rather that mismatch effects can be explained by technology's features and how innovators prioritize local adaptation efforts. As a result, technological improvements making agriculture more productive or climate-resilient may fail in the long run because today's innovation supply carrying those crucial traits is not good enough to induce adoption among poor farmers.

1 Introduction

A persistent obstacle to the development of low-income countries is the slow adoption of new agricultural technologies. Innovation is crucial for farm productivity and economic growth, but poor farmers are still reluctant to adopt seemingly profitable new inputs and practices (Fuglie et al., 2019). A vast literature has examined this issue, focusing mainly on demand-side factors limiting farmers' technology choices.[1] Relaxing those constraints has also been the focus of policy in the developing world for decades. In contrast, we know considerably less about the supply side of this technology adoption puzzle (Emerick et al., 2016; Suri, 2011), and while it is commonly presumed that modern technologies are indeed better for most farmers, several examples challenge this assumption.[2]

An under-appreciated aspect of this problem is that new technologies are not usually tailored to the needs of poor farmers. Although public research centers and state-owned enterprises are the main suppliers of agricultural technology in many low-income countries, they often lack the resources and incentives to match innovations with farmers' demand for new technology.[3] These constraints can limit the ability of agricultural research and development (R&D) to prioritize and transfer innovations. For example, country-level evidence suggests that ecological mismatch can induce transfers of inappropriate technology (Moscona and Sastry, 2021). However, there is no definitive explanation for why technological mismatch in agriculture persists over time. In this paper, I argue that the absence of mechanisms crowding out bad technology allows innovators and input suppliers to remain in low targeting efficiency equilibria. What happens then to technological change when innovations do not

[1]These factors include time and risk preferences (Liu, 2013; Duflo et al., 2011), learning failures (Hanna et al., 2014; Conley and Udry, 2010), and limited credit and insurance access (Karlan et al., 2014).

[2]Findings on heterogeneous returns on adoption and yield risk show that improved technology is not necessarily optimal under real-world conditions (Suri, 2011; Barrett et al., 2004). Lemon technologies such as low-quality fertilizer can lower farmers' returns (Bold et al., 2017), although accurate quality estimates are needed to prevent quality misattribution (Michelson et al., 2021). Also, yield improvements of new crop varieties are often overestimated in agronomic trials (Laajaj et al., 2020).

[3]Public spending accounts for three-quarters of global agricultural R&D investment, which is mostly driven by spending in China, India, and other middle-income countries (Beintema et al., 2020). At the same time, the inability of researchers to appropriate the returns of innovation through patents or plant breeder's property rights removes some of the incentives to develop more competitive technology (Fuglie et al., 2019).

reflect the preferences of those who must ultimately take up new technology?

To answer this question, I examined the development and adoption of improved crop varieties. It is common for plant-breeding programs to release a single new variety to supply a large and heterogeneous group of farmers. In developing countries where seed markets are often informal and incomplete, public releases are usually the only new crop varieties available to farmers. A key challenge, however, is that we only observe the varieties that breeders release, not the counterfactual varieties they may have developed had they been able to properly match farmers' preferences. Thus, I conducted a two-stage randomized control trial in Costa Rica to estimate the effect of targeting frictions on farmers' adoption of improved seeds.

In the first stage, a group of small and medium-scale farmers tested five new drought and heat resistant bean varieties in their fields. As a short-duration cash crop with stable demand, bean production is essential for farmers' food security and incomes, as they use it to finance the production of major crops (maize and rice). However, with no irrigation systems in place, poor farmers rely on rain to water their crops, making weather shocks a constant productivity risk. The agronomic trials allowed me to elicit farmers' preferences and estimate the new varieties' performance under farmers' heterogeneous conditions, which are usually private information unknown to breeders.

In stage two, I recreated counterfactual scenarios for breeders' release decisions such that each farmer faces his or her own seed market. Contrary to adoption studies that distribute new inputs for free, farmers in the intervention were offered to buy a fixed amount of seed of the new varieties tested in the trials. I randomized these offers so that targeted farmers were offered the variety they preferred the most based on their experience during the trials. In contrast, farmers in the mismatch treatment received a recommended but unknown new variety. The recommendation was made by a group of crop scientists from the national breeding program, following a process similar to the actual release of new varieties.[4] I

[4]The variety recommended to farmers was indeed released by Costa Rica's public breeding program in March 2023. See press release here (in Spanish).

compared the take-up rates and outcomes between these groups and with farmers that did not participate in the trials or received an offer.

The main results indicate that preference-driven mismatch significantly reduces farmers' adoption of new technology. Take-up among farmers who were offered the recommended variety is 50% lower compared to targeted farmers, who received offers that matched their stated preferences. This difference is entirely driven by the mismatch between farmers' preferences and the available supply of new varieties. Since participation in the agronomic trials (a proxy for learning-by-doing) had no impact on take-up, there is no evidence of farmers trading off first-hand experience in favor of expert advice (i.e., varietal recommendation). I also find no impact of yield expectations on adoption decisions. However, differences in take-up rates of the recommended variety across treatments suggests that the recommendation promoted purchases of a variety that may be sub-optimal for one in six adopters. Further, take-up in the mismatch group is comparable to the demand for certified seeds of current varieties. This result attenuates the inferred adoption under mismatch since these take-up levels are similar to the average seed replacement rate. These findings hold after experimentally controlling for well-known adoption frictions, such as liquidity constraints, heterogeneous transaction costs, quality misattribution, and limited-attention learning.

Moreover, trial results suggest that trial productivity alone does not predict adoption. Yield comparisons show that non-adopters did not under-perform in the trials relative to adopters. For some farmers, there are significant differences at the low end of the productivity distribution. No differences are observed in higher quantiles of the yield distribution in either region. These results follow a similar pattern to the returns on adoption reported by Suri (2011). Nevertheless, I find no consistent improvement across varieties when compared to the reference variety (a widespread variety released in 2003), even though breeders' experimental evidence in controlled environments suggests otherwise. So, the productivity gains some farmers experienced are either not linked to varietal improvement but likely caused by the superior quality of the trial seed or only occurring in the tails of the weather distribution, where improved resistance to extreme weather may increase yield stability but not yield potential.

A possible explanation for the mismatch effects is that heterogeneity across farmers impedes technological change. Suppose innovators cannot capture that heterogeneity, or there are no mechanisms motivating them to internalize it. Then, new technologies will only respond to the needs of a specific group of farmers. So, I study differences in research efforts and new varieties' features to estimate heterogeneous treatment effects. Results on research efforts indicate no mismatch effect on farmers located closer to the experimental station where the varieties were developed and tested, followed by increasingly negative mismatch effects on take-up for farther away farmers. Other estimates using weather variability suggest significant heterogeneous mismatch effects. Improved resistance to related to drought and extreme temperature are key expected traits of the new varieties. Results show that mismatch effects are greater among farmers who experienced drought and extreme heat events prior the intervention. Regression analysis using atmospheric data confirm these results for dry spells (the number of consecutive days with daily precipitation lower than 1mm) but show null results for the number of hot days (maximum daily temperature greater than 75 percentile average temperature).

In essence, this paper shows how detrimental innovation frictions can be for technological change in developing countries' agriculture. This paper contributes to the recent literature on supply-side constraints on technology diffusion in low-income countries (Dar et al., 2021; Emerick et al., 2016). Here I tested an alternative explanation, intrinsic to the technology and its development (Suri, 2011), and different to adoption frictions or cases when adverse selection crowds out adoption (e.g., bad quality inputs). Moreover, my results are closely related to the inappropriate technology hypothesis (Moscona and Sastry, 2021; Stewart, 1977). This paper provides direct evidence that technological mismatch negatively affects technology adoption. If little take-up occurs, there is also little benefit of reducing the technological mismatch, unless inappropriateness itself is the cause of limited adoption. Thus, the impacts of inappropriate technology transfers would be large when efforts to adapt these new technologies to local conditions are rare or unsuccessful, which primarily depends on local researchers' incentives and capabilities.

Furthermore, findings in this paper support the conclusion that imperfect targeting help to explain the technology adoption puzzle in agriculture. From an adopter's perspective, new technology delivering no significant benefit is rejected in favor of current alternatives. On the supply side, however, innovation is as good as it can be under the constraints innovators face, especially in the context of public agricultural R&D (e.g., competing interests, limited funding and expertise, no property rights of innovations). Since many of the world's innovators operate under these conditions, which limit their ability to internalize farmers heterogeneity, imperfect targeting contributes to the slow pace of technological change observed in developing countries' agriculture. As a result, technological improvements making agriculture more productive or climate-resilient may fail in the long run because today's innovation supply carrying those crucial traits is not good enough to induce high adoption among poor farmers.

The rest of this paper is organized as follows. Sections 2 describes the setting and research design. Section 3 summarize the data and presents the empirical estimation strategy. Section 4 report the experimental treatment effects on new varieties take-up. Section 5 summarize the potential mechanisms explaining the main results on adoption. Section 6 presents preliminary results on learning, expectations, and heterogeneous mismatch effects. Finally, section 7 includes a short discussion based on preliminary results.

2 Research Design

2.1 Setting

The sample consists of 800 farmers from two regions of Costa Rica.[5] I focused on these regions because together they account for the vast majority of beans produced in the country, each region has distinct agroecological and social conditions, and they host farmers for whom investments in new seeds is the most relevant. The northern region is located along the border with Nicaragua and includes the Huetar and Chorotega subregions (see figure 1). In the south, the Brunca region is located near the pacific ocean, stretching in along mountain range near the border with Panama.

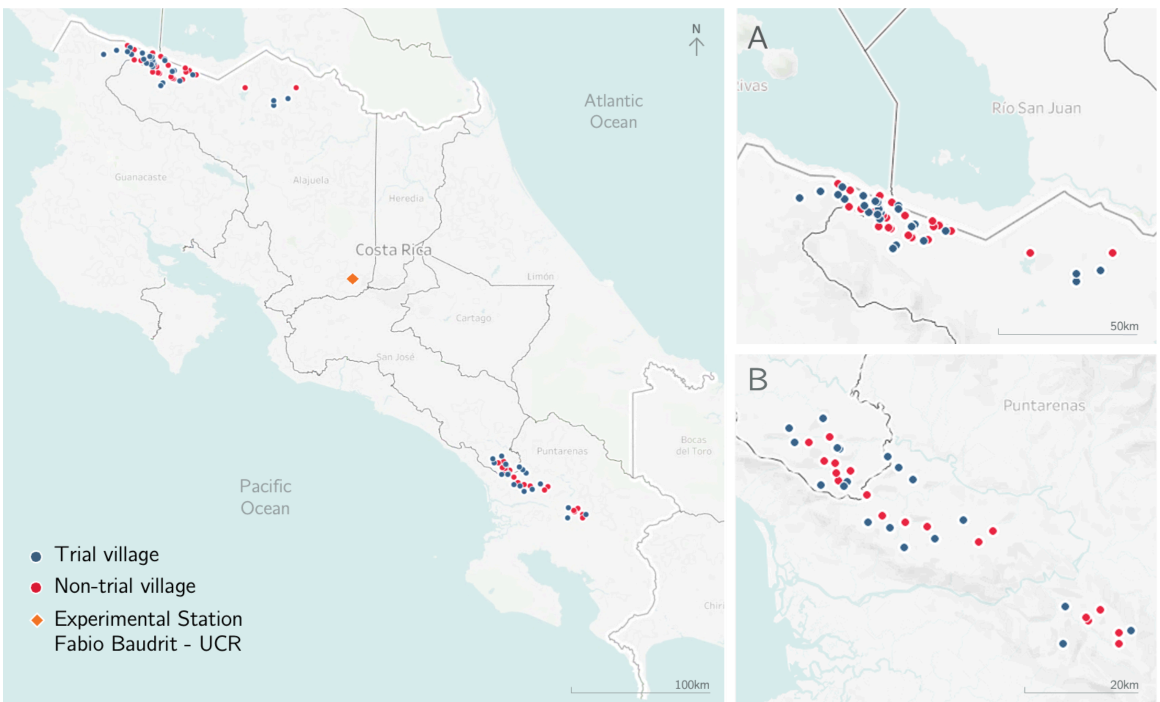


Figure 1: This map of Costa Rica shows the 117 villages sampled for the study, highlighting the selected villages for the agronomic trials. Panels A and B zoom in over the North and South regions, respectively. The diamond-shaped icon locates the experimental station in which the new varieties were developed and tested.

[5] Protocols for this study were reviewed by the Institutional Review Board for Human Participants at Cornell University (#2106010430). The randomized evaluation was included in the American Economic Association RCT registry in November 2021, prior to the main intervention (AEARCTR-0008452).

The main economic activity in these regions is agriculture. As a cash-crop with constant demand and a short farming cycle (about 75 days), bean production allow farmers to finance the production of maize and other minor crops for commercial purposes and self-consumption. In the south, farmers plant beans twice a year, following the dry (October to February) and wet seasons (May to August). The steep terrains of the south limit mechanization, forcing bean producers to rely heavily on farm labor. In comparison, the majority farmers in northern Costa Rica only plant once per year during the dry season. Farms in the north are located in flat terrains, which allows for mechanization and larger planting areas. Importantly, in the south have the support of a stronger network of agricultural associations and cooperatives offering commercialization services, credit, and mechanized processing, which are lacking for most farmers in the north.

2.2 Sampling

From these regions, I selected 117 villages using administrative records from the National Productive Council of Costa Rica (CNP henceforth for its name in Spanish). I sampled villages with at least six small or medium scale bean farmers (farm-size of less than 50 hectares) registered with the CNP.[6] By recommendation of the breeders, villages in indigenous communities were excluded to prevent the replacement the traditional bean varieties they grow. Other villages were ignored because of logistical reasons (places with few bean farmers, inaccessible locations or with insufficient extension support to conduct the study). Then, I drew a stratified random sample of farmers from the CNP registry using the selected villages as strata (between 6 and 7 farmers per village).

The intervention was implemented using a two-stage randomized control trial as described in figure 2. In the first stage, agronomic trials were conducted to test the performance of five red bean varieties on farmers' fields and to elicit farmers' preferences for new varieties.

[6]Most farmers in Costa Rica have incentives to register in the CNP. Being part of this registry allows farmers to access public assistance programs, including being able to sell produce to the public procurement program at a higher price than the market, receive extension services, subsidies, and free inputs, such as certified seed and fertilizers.

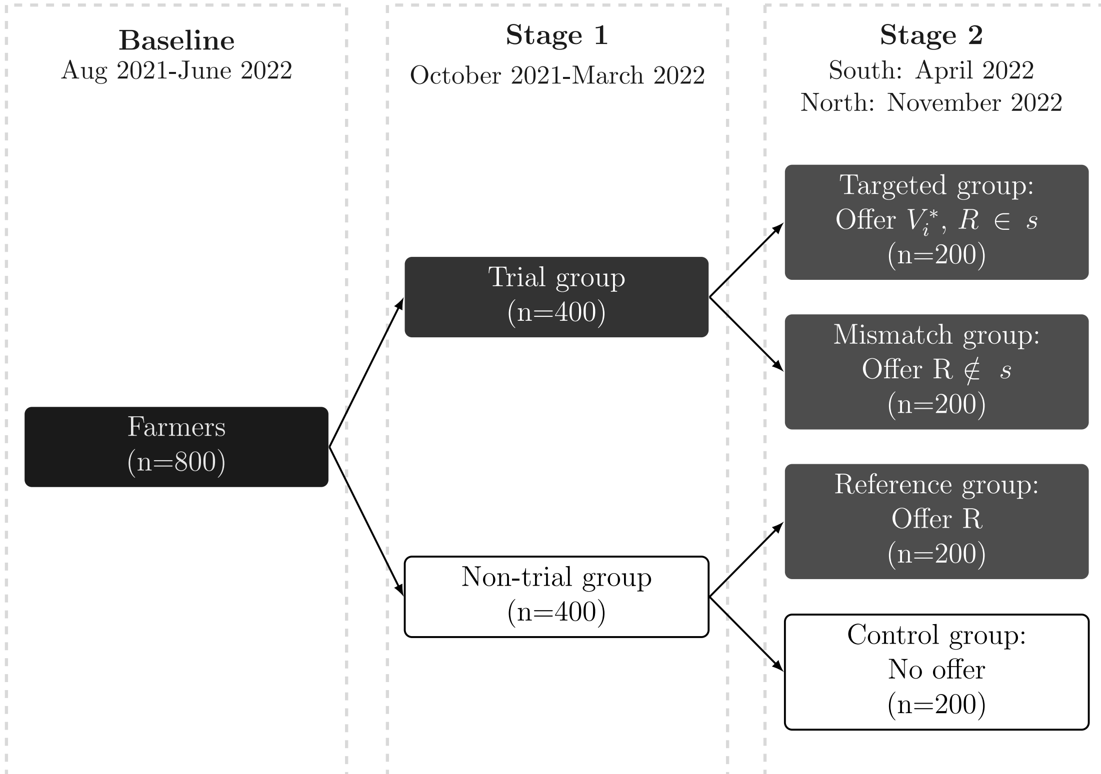


Figure 2: This figure describes the research stages, intervention timeline and the distribution of farmers in the experimental groups. The trial group refers to farmers assigned to agronomic trials. The targeted group refers to farmers who were offered the variety of their preferences (V_i^*) from the testing set of varieties (s). Farmers in the mismatch group were offered the variety recommended by the breeders (R), which was not part of their testing set. The reference group was part of the non-trial group but received an offer to buy R. The control group did not participate in the trials nor receive an offer of new varieties.

To control for the participation in the trials, the sample of farmers was divided into two experimental groups randomized at the village-level. In the second stage, farmers were offered one of the new varieties for purchase, conditional on participation in the trials and assignment into four experimental groups depending on the random distribution of varieties, and the breeders' recommendation. Each stage is explained in detail in the following sections.

The germplasm used in the intervention was the result of a five-year breeding cycle (at the time of on-farm trials) conducted by Costa Rica's national plant-breeding program. The selection process started with more than 100 genetic lines from the International Center for Tropical Agriculture in Colombia (CIAT for its name in Spanish), and produced five candidates for new red bean variety be released in Costa Rica. The main breeding objective of this process was to improve seeds' drought and heat resistance compared to Cabecar, a high-yield and widely adopted variety that was released in 2003, which was used as reference variety in the trials. This implies that farmers had no access to these varieties prior to the trial with very few exceptions.[7]

2.3 Farm trials

The Triadic Comparison of Technologies method (Tricot for short) was used to allocate varieties to farmers (five new and a reference variety).[8] Each farmer in the trial group received a set (s) of three new varieties of the five varieties at random to plant during the dry season of 2021. The information about the varieties in the testing sets was not revealed to farmers, so varieties were given letters as names, such that $s = \{A, B, C\}$. To limit spillovers about the varieties' performance, farmers were informed that each participant in the trial received a different set of varieties. By design, each one of the six varieties only appears in half of the testing sets used in the trials.

Farmers also participated in training sessions on how to plant and evaluate the new varieties. They were asked to apply the same overall management to the trial plots as their current bean plots, and to choose the location of the trials in their fields. However, farmers were told the number of plants, seeds per plant, and distance between plants to use. They were given 150 grams of each new variety to plant in a 5x5 meters plot per variety. A group of local collaborators were in charge of preparing and staking the trial plots for identification, support farmers in how to fill out a performance scorecard, and collect three waves of trial data (planting, mid-season, and post-harvest).

During the trials, farmers examined the varieties' relative performance by choosing the best and worst varieties in their set. This evaluation was structured using specific traits to prevent differential learning, as some farmers may ignore important aspects of technology

[7]Those who had access to the new germplasm were part of a small group of farmers in the south that allowed breeders to test the varieties as part of the selection process. In any case, farmers were not able to identify the specific varieties in the trials.

[8]See van Etten et al. (2019) for detailed information about the Tricot method.

(Hanna et al., 2014).[9] Participants answered which variety in their testing set, if any, they wanted to adopt in the next season. They also indicated their overall ranking of varieties, and compared of each new variety with their current variety. With the help of local collaborators, farmers weighted the bean output from each plot to measure the new varieties' yield.

2.4 New variety offer

In the second stage, I use the random allocation of new varieties to further divide the sample. Based on the trial's information and farmers' preferences, farmers were grouped into four experimental groups (three intervention groups and a control). The intervention groups received an offer to buy one of the new varieties tested in the trials and were presented with the same information about the trials' results. Each offer consisted of three kilograms of a single variety for a fixed price of 4800 colones (about $7.5 USD), matching the retail cost of a same amount of certified seed. Note, however, that the quality of the seed offered to farmers in the intervention (foundation seed) was superior to that of the commercial certified seed. In addition, this price does not take into account other costs (e.g., transportation). The offers were made at the farm gate to prevent that differences in transaction costs to influence uptake decisions, as reported by Suri (2011). Thus, field assistants visited each farmer in the intervention groups in their farms, or at their most convenient place for them to meet.

During the seed sale, flexible payment methods were allowed to prevent liquidity constraints limiting seed purchases (Karlan et al., 2014). Payment methods included cash, interest-free pay-later loans (up to two weeks), online payment using a smart-phone app (called SINPE). Farmers were also allowed to reschedule a visit to deliver the seeds and collect the payment personally, or through a third party (relatives, neighbors and local shop-

[9] The traits included in the performance scorecard are: plant structure, maturity, pest resistance, drought resistance, yield (referring to grain weight and the number of grains per pod), commercial value (grain size and color), taste, and cooking time. Some of these traits were recommended by the breeders and others were market-oriented attributes relevant for farm profitability. The Spanish version of the scorecard used by farmers is included in figure A1 in the appendix.

keepers).

The varieties offered farmers were determined by farmers' preferences and crop scientists' recommendations. The process of deciding the variety to recommend imitated real-life decisions breeders make when a new variety is released. Results from the agronomic trials were used to determine the best performing variety in each region, as well as yield differences between the new and the reference varieties. Breeders also used information from previous trials conducted in the public experimental stations and exhibition plots, and qualitative results from discussions with selected farmers' groups and association leaders.

Half of the farmers in the trial group were assigned to the targeted group (see figure 2 under stage 2). Farmers in this group were offered their preferred variety from the seeds tested in the trials (denoted V_i^* for each farmer i). First-hand experience with the new varieties in the trials may allow farmers to update their beliefs and reduce the uncertainty related to investments in new technology. The targeted group represents an ideal but unrealistic situation in which the supply of crop varieties matches exactly farmers' revealed preferences for new varieties.

The other half of trial participants formed the Mismatch group, who were offered the variety recommended by the breeders (R). Note that farmers in this group had no previous experience with the recommended variety, given that R was not in their testing set s. Since these farmers formed preferences for varieties other than R, this group captures the mismatch problem caused by a constrained seed-supply unable to match farmers' preferences. In this case, neither breeders or farmers knew the actual performance of R in farmers' fields. However, to maintain the same level of information across intervention groups, farmers were informed about the results of the trials at the moment the offer was made. This information includes the average performance in each trait of R, relative to the varieties in the trial.

The non-trial group was divided into the Reference and Control groups. The Reference group was treated exactly as the Mismatch group, so the only difference is that reference farmers did not evaluate the new varieties under evaluation. Thus, results for the conventional group allow me to identify the effect of the participation in the trials on farmers' take-up of the new varieties. This group is the closest to the reality of many farmers, as they usually have limited information and experience with new technology before deciding whether or not to adopt it, and they rely on performance information provided by breeders, input suppliers, and their peers.

Finally, the control group is comprised of sampled farmers who took part of the baseline survey but were not part of the trials and did not receive any offer of new varieties. Results from this group allows me to control for new seed, higher-quality purchases of current the varieties. Moreover, I use the control group's information to test whether the adopters (those who take-up of the new varieties) and non-adopters are different from an untreated group of farmers from the same population selected at random.

The end-line survey information was collected at the end of the following productive season after the offers were made. In the south, offers were made three weeks before the start of the dry season of 2022, and the end-line survey was conducted in August 2022. For the northern regions, offers spanned from October to December 2022, and the survey information was collected in January and February 2022.

3 Data and empirical Strategy

3.1 Administrative records

Data from the CNP was used to sample farmers for the study. Table A1 in the appendix compares the sample of farmers with the population of small- and medium-scale farmers in the CNP registry. I find no statistically significant differences for the relevant variables included in the registry, except for a slightly higher proportion of farmers sampled from the norther region, suggesting that both samples of farmers are comparable and supporting the external validity of my results.

3.2 Survey data

Survey data was collected before and after the intervention. In the southern region, survey data was collected in the second half of 2021 before the start of the rainy season. In the north baseline data was collected during the first semester of 2022. Using survey questionnaires, a team of local surveyors visited farmers to collect baseline information on individuals' characteristics, household composition, and farm management. The farm survey included plot-level questions on productivity, input use and farming practices. Table 1 reports mean values for relevant characteristics and compares trial and non-trials group. To test for sample balance between these groups I use two-tail difference in means tests. I find no significant difference for most variables except the education level (p-val=0.07) and farm area in hectares (p-val=0.04). Thus, I include these two variables as baseline controls in the estimation strategy.

Table 1: Sample balance: trial versus non-trial groups

	Trial (n=352)	Non-trial (n=392)	Difference p-value
Age (years)	46.47	47.93	0.15
Gender (male=1)	0.83	0.79	0.17
Education level (1-8)	2.07	1.94	0.07
Household income (usd/month)	317.58	298.83	0.59
Family size (# members)	3.51	3.48	0.79
Farming experience (years)	18.30	18.37	0.95
Bean yield (quintal/ha)	19.10	19.07	0.97
Planted area (ha)	5.43	5.47	0.97
Farm size (ha)	7.48	9.33	0.04
Bean plots (#)	1.74	1.70	0.55
Land renting (yes=1)	0.40	0.40	0.99
Input subsidy (yes=1)	0.22	0.19	0.39
Distance to exp. station (km)	146.24	146.37	0.91

Notes: This table compares the trial groups (Targeted and Mismatch) and non-trial groups (Reference and Control) at the baseline. Trial assignment was randomized at the village level. P-values are reported for differences in means tests.

The average farmer in the sample is male, middle-age, with elementary education, and part of a household making $5.34 ppp dollars per person per day (for reference, the World Bank estimates a the poverty line for Costa Rica of $6.85 in 2017 ppp dollars).[10] The average farm uses on average 5.4 hectares of land and has a productivity level is on par with national estimates of 18 to 20 quintals per hectare for small- and medium-scale farmers.

3.3 Trial data

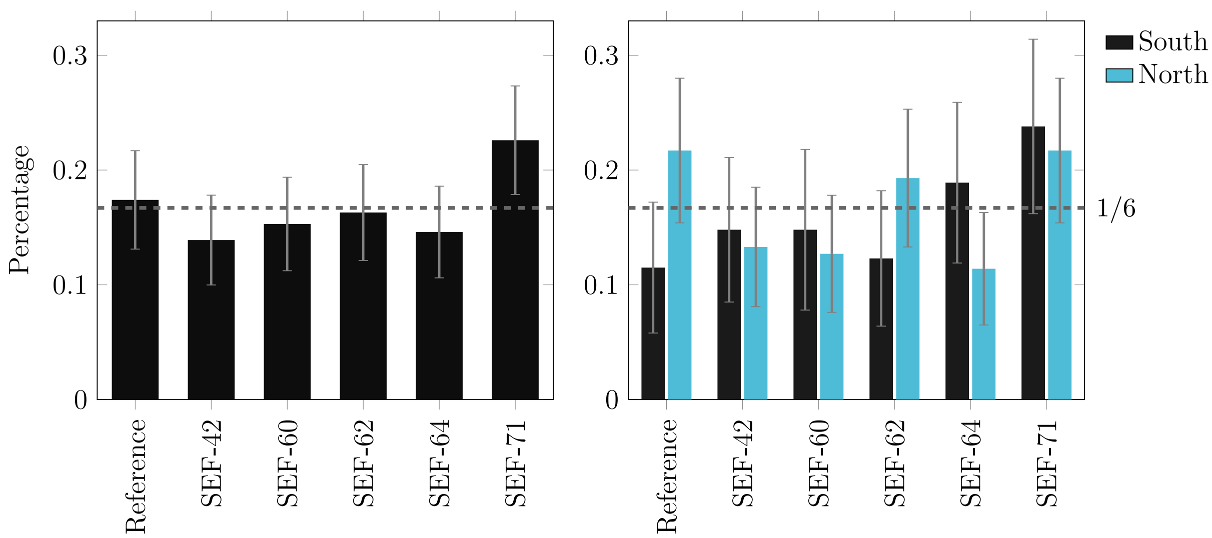
The data collection for the trials was divided in four short waves in which local collaborators visited or called over the phone farmers. In each wave, farmers evaluated traits related to the development of the plant and the bean output (see figure A1 in the appendix). In the first visit, 30 days after the estimated planting date, farmers were asked about plant structure. This visit was also used to confirm the actual planting date of the trial plots. 15 days later, farmers were asked about maturity (the time plants took to flower) and plants' resistance to pests and drought. During harvesting, approximately 70 days into the trial, farmers were asked to compare the varieties in terms of yield and commercial value. They also were instructed to cook and taste the varieties to estimate the cooking time and their preferences as end-consumers. Some days after the harvest when beans were dried and ready for sale, farmers were visited to evaluate the overall performance and their preferences. They were also asked to compare the new varieties with their current variety.

Two questions are used to determine farmers preferences. Farmers were asked to decide which varieties they would like to plant in the next season, as a measure of their stated preferences. They also decided what variety was the best variety overall from their testing set. Responses to these two questions coincide 92% of the time, but only a 60% of farmers chose a single variety as strictly preferred. Also, 2.4% of farmers indicated they preferred not to adopt any new variety. So, whenever more than one variety is chosen, the overall best variety is set as their most preferred variety.

Figure 3 reports the frequency in percentage terms that each variety was selected as

[10] This calculation was made using the 2021 purchasing power parity (ppp) conversion factor of 343.9 (The World Bank, 2023).

farmers preferences. In the cases that more than one variety was chosen, I assign the best overall variety as their preferred variety. The pooled trial data (panel A) shows that SEF-71 is selected more than the rest, with a percentage that is significantly higher than the uniform probability of selecting any variety at random (one out of six). This is not necessarily true when data is desegregated by region (panel B). Differences between regions show that the SEF-71 results are driven mostly by the south, where the reference variety (called Cabecar) under-performed compared to the new varieties. This is supported by qualitative observations in the field indicating an appetite for alternatives that are not as susceptible as Cabecar to winter pests and diseases. In the north, where farmers only plant during the summer season, the reference is the most preferred variety, tly more than the SEF-71. Regional results, however, indicate no significant differences when compared to a random choice as shown by the 95% confidence intervals.



(A) All trials

(B) By region

Figure 3: Figure reports farmers' stated preferences for the varieties in the trials. Each bar corresponds to the percentage of farmers who chose a particular variety as the one they wanted to plant in the next season. The corresponding vertical lines indicate the 95% confidence intervals. The horizontal dashed line shows the uniform probability of choosing one of the six varieties at random.

Yield comparisons indicate no significant improvement of the new varieties (see A2 in the appendix for details). The average yield for all trials was about two kilograms for a five squared meters plot. Consistent with farmer choices, the reference variety performed better in the north, and although there are no consistent differences between regions, most of the new varieties seem to be better adapted to the south. Also matching with farmers' preferences, SEF-71's yield was the highest in the south. In the north, SEF-60 exhibit the highest yield, despite it was one of the least preferred varieties, indicating that traits other than yield may be more important for farmers. I find no significant yield improvements by comparing the new and reference varieties, and the only significant differences are lower yields for the SEF-42 and SEF-64 varieties. Considering these varieties were developed to improve drought and high-heat resistance, both low-probability events, these averages could mask differences in performance due to extreme weather which take place at the tails of weather distribution.

3.4 Estimation strategy

What is the effect of imperfect targeting on the new variety take-up?

The main empirical approach is to use the treatment assignment to identify the effect of the intervention on adoption. The estimation is based on the reduced form regression model described by equation (1). The outcome *take-up* is an indicator variable that equals one when a farmer buys the new variety offered in the intervention, zero otherwise. Explanatory variables include indicator variables *R-group* that describes assignment of farmer i in village j to the treatment groups offered the recommended variety (mismatch and reference), *T-group* describes assignment into the trial (targeted and mismatch), δ_j capture village-specific fixed effects (strata used for randomization), and ϵ_{ij} is the error term, which is clustered at the village level (randomization level in the first stage of the intervention). Thus, information from the reference groups is used as the benchmark.

$$\textit{take-up}_{ij} = \alpha_0 + \alpha_1 \textit{R-group} + \alpha_2 \textit{T-group} + \delta_j + \epsilon_{ij} \tag{1}$$

Coefficient $\hat{\alpha}_1$ is captures the preferences mismatch effect on the adoption relative to the targeted group. Village fixed effects δ_j control for location-specific effects, including differentials in varietal adaptability. The trial participation effect is captured by $\hat{\alpha}_2$, which identifies the difference between the mismatch and reference groups. The main hypothesis is that the targeted farmers should exhibit the highest take-up rate of all experimental groups. Thus, the expected estimated effects are $\hat{\alpha}_1 < 0$, $\hat{\alpha}_2 > 0$, and $\hat{\alpha}_1 - \hat{\alpha}_2 \leq 0$.

What is the impact of adopting a new variety? I estimate the effect of taking-up a new variety on farming outcomes by comparing adopters versus non-adopters using a difference-in-differences approach. The estimating model is described by equation (2),where Y_{ijt} is the outcome variable in period t. $Post_t = 1$ identifies information post-intervention, and μ_{ijt} is the error term. Coefficient β_3 identifies an intention-to-treat (ITT) estimate of adoption on outcomes. Coefficient estimate β_1 captures adopters' group-specific effects, and β_2 identifies the time trend common to both groups.

$$Y_{ijt} = \beta_0 + \beta_1 take\text{-}up_{ij} + \beta_2 Post_t + \beta_3 take\text{-}up_{ij} * Post_t + \delta_j + \mu ij \tag{2}$$

I study three sets of potential outcomes. The new bean varieties are expected to deliver improvements in productivity and drought/heat resistance. Under normal conditions, the new varieties should provide a similar yield as the current varieties in the market, while they should perform better under heat and drought stress. First, this paper focuses on changes in average yield and yield variance as measures of productivity. I estimate these effects using plot-level data collected before and after the intervention. If the new varieties are indeed better, we should observe that adoption has some positive effect on yields, either by improving yield potential or by reducing yield losses. Given the trials results, I do not expect overall yield improvements from taking-up a new variety. However, the magnitude of these effects may vary with the specific varieties a farmer chooses to adopt.

Second, changes in quality traits will be estimated at the farm level. Quality improvements can be estimated using the adjustments to sale quantities made by buyers based on discarded beans due to quality issues. The new varieties are expected to maintain the appearance of beans desired by end-consumers (beans with a bright red color). Buyers discard darker, dull colored, and damaged beans, which are usually sold as second-class beans for processed foods. I use self-reported quality losses from survey data, which is corroborated using information from farmers associations that buy beans from farmers.

Third, I focus on the indirect effects of adoption on input use. Evidence in the literature suggests that adoption can crowd in complementary factor-deepening technologies (Emerick et al., 2016), which implies that adoption can positively affect capital investments related to machinery use and fertilizer use. If estimated effects of adoption are factor-deepening, comparable effects on on-farm and off-farm labor allocation can be expected.

Are farmers better off accepting the recommended offer?

An important question is whether there are heterogeneous adoption effects due to imperfect targeting. To answer this, I estimate the triple-differences model described in equation (4), where indicator variable *R-groups* identify farmers assigned to treatment groups $Mismatch$ and $Reference$, which received the offer to buy variety R. Coefficient estimate γ_6 identifies the effect on the outcomes of adopters who adopted the recommended variety. A negative and significant coefficient would indicate that farmers who can choose according to their preferences are better-off than those in the treatment group who adopted variety R.

$$\begin{aligned} Y_{ijt} =& \gamma_0 + \gamma_1 \textit{take-up}_{ij} + \gamma_2 \textit{R-groups}_{ij} + \psi_3 Post_t \\ &+ \gamma_4 \textit{take-up}_{ij} * Post_t + \gamma_5 \textit{R-groups}_{ij} * Post_j \\ &+ \gamma_6 \textit{take-up}_{ij} * \textit{R-groups}_{ij} * Post_t + \delta_j + \varepsilon_{ijt} \end{aligned} \tag{3}$$

Identification threats

A significant concern is the potential endogeneity of adoption decisions and outcomes.

As mentioned earlier, results indicate that adopters were more productive in the trials, particularly for below average yield values. To address this, I use the information of farmers in the control group. The objective is to test whether the group of adopters is different from a group of farmers from the same population selected at random. More importantly, it allows us to test whether groups exhibit the same time trend in the absence of the intervention. I estimate the same model presented in equation (2) using (i) a sub-sample including only the information of non-adopters and control (as the comparison group), and (ii) the full sample relative to the control group. These results would be indicative of any potential difference pre-treatment.

Another important identification challenge is farmers systematically using higher (or lower) quality plots to plant the new variety (Emerick et al., 2016). This can lead to biased estimates of the adoption impact of the new varieties. During baseline survey data collection, farmers were asked to map the plots in their farm and identify their perceived plot quality. I use use this information and plot-level productivity data to control for plot selection in the estimation of the adoption impact.

4 Treatment Effects

4.1 New varieties take-up

Figure 4 reports the take-up rates of new varieties across treatment groups. Overall, results show that 43% of farmers purchased the new variety when offered during the intervention. Also, there here are significant differences across treatment groups. Uptake among farmers in targeted group is about 0.56, which is 17 and 23 percent points higher than the mismatch and reference groups, respectively. On average, this translates into 50% percent higher adoption as a result of matching farmers' preferences with the appropriate new variety. I find virtually no difference in adoption between in adoption between mismatch and reference groups. Although results by regions show the same take-up pattern across treatment groups, farmers in the south seem to have take-up rates that match the seed replacement rate (purchases of certified seed of current varieties). This suggest that adoption may be driven, to some extent, by the need of higher quality seed and not the demand for new varieties. On the contrary, seed replacement in the northern region is significantly lower than uptake.

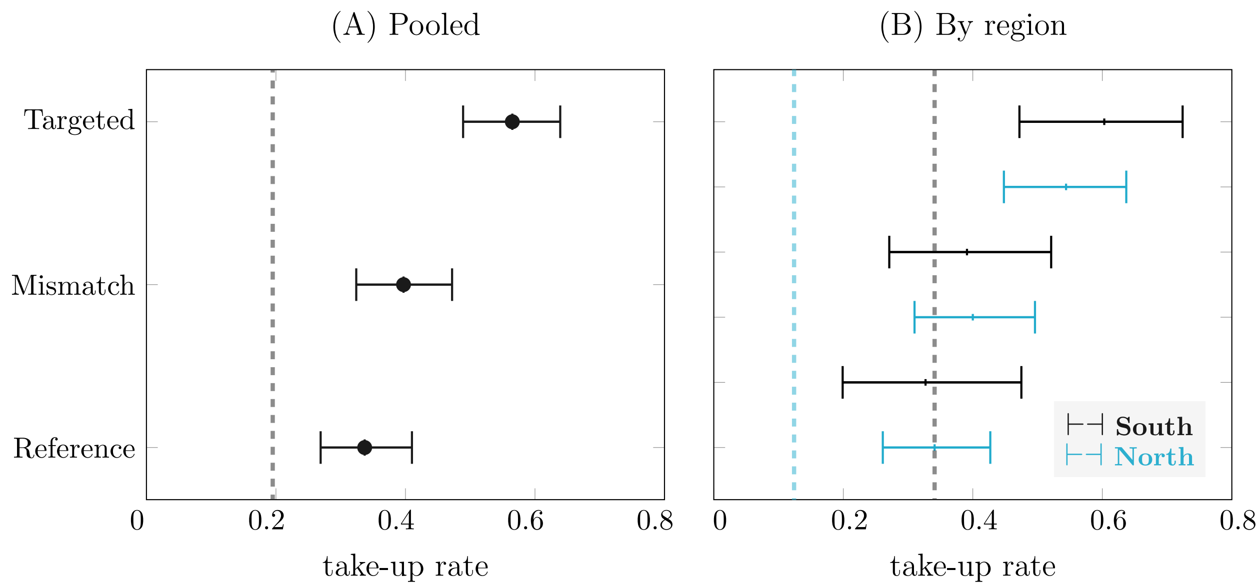


Figure 4: The figure reports the average take-up rate across intervention groups and the corresponding 95% confidence intervals. The left panel shows pooled results. The right panel shows results by region. Take-up is defined as a farmer purchasing the variety offered in the intervention. The vertical line shows the percentage of farmers who bought seeds of current varieties from the CNP or bean farmers associations during the season when the intervention occurred.

Uptake results also give us a sense of the mismatch magnitude. If take-up of the recommended variety is similar across treatment groups, we could conclude that only farmers who benefit from it are buying into the breeder's recommendation. While 39% farmers in the mismatch group purchased the recommended variety, only 22% of farmers in the targeted group did so too (a difference of 17 percent points). Note that this percentage matches the fraction of farmers who preferred the recommended variety in the trials (see SEF-71 in figure 3), showing that stated preferences are a good predictor of uptake. Given that the recommended variety was tested by all targeted farmers, and assuming that the randomization balanced out any other difference in preferences, these results suggest that the recommendation pushed purchases of a variety that may be sub-optimal, if not unfavorable, for one in six adopters.

Table 2: Treatment effects on take-up

	(1) take-up	(2) take-up	(3) take-up	(4) Uptake
Mismatch	-0.182***	-0.182***	-0.175***	-0.177***
	(0.058)	(0.058)	(0.059)	(0.058)
Trial participation	0.102	0.098	0.099	0.116
	(0.081)	(0.079)	(0.077)	(0.076)
Seed replacement		0.077		
		(0.063)		
Seed replacement x South			0.181**	0.178**
			(0.080)	(0.077)
Seed replacement x North			-0.011	-0.016
			(0.090)	(0.089)
Lost trial				-0.198*
				(0.100)
Constant	0.486***	0.474***	0.470***	0.478***
	(0.073)	(0.075)	(0.074)	(0.074)
Dependent variable mean	0.431	0.431	0.431	0.431
Village fixed effects	yes	yes	yes	yes
R-squared	0.200	0.201	0.204	0.213
Observations	540	540	540	540

Notes: This table reports coefficient estimates from linear probability models using Pr(take-up=1) as the dependent variable. Fixed effects at the village-level included, which was the stratification for the randomization of the agronomic trials. Robust standard errors clustered at the village level in parenthese. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 2 reports coefficient estimates of the reduced-form regression model described in equation (1). These linear regression models use take-up as the dependent variable and the reference group as the comparison category. Results confirm the previous findings on the significant differences across treatments, and show consistent mismatch effects for all specifications. When seed replacement is included in the model, its coefficient is not different than zero (column 2). Once this effect is disaggregated by region (column 3 and 4), seed replacement in south matches the take-up rate of about 18 percent points. As before, there is no significant association between seed replacement and take-up for northern farmers.

4.2 Adoption impact

[*Editorial note:* Endline data collection is finished in the south region but currently underway in the north region. Results on adoption effects are expected to be included by July, 2023.]

5 Conceptual framework

In the uptake results, I have documented three facts: (i) farmers for which new technology matched their stated preferences exhibit a markedly higher adoption rate. (ii) Testing the new technology through agronomic trials prior to the adoption decision had no positive impact on take-up. On the contrary, participating in the trial but failing to successfully test the technology reduces adoption. (iii) The demand for higher quality inputs is positively correlated with new technology adoption for some farmers but not all. Taken together, these facts indicate that the usual one-size-fits-all strategy that characterizes public breeding programs' releases leads to preference-driven mismatch that limits new technology diffusion.

In this section I summarize the potential mechanisms explaining the intervention's treatment effects on adoption. These ideas are based on a agricultural household model rationalizing farmers' response to frictions in the input supply, under a discrete-technology structure of production. For most part, I ignore adoption frictions from the demand side, considering that my intervention controls for well-known constraints by design. Instead, I focus on innovation frictions caused by costly heterogeneity across farmers' individual conditions.

The first mechanism relates to learning, expectations and the risk associated with new technology investments. Participation in the agronomic trials may allow farmers gather first-hand information that help to reduce the uncertainty related to new technology investments.

Being able to learn about the new variety performance by testing it (a proxy for learning-by-doing), prior the adoption decision should incentive farmers who observe net economic gains to adopt. This explanation requires that there is no performance misattribution, so that farmers correctly identify benefits and losses derived from the new technology (e.g., net productive improvements or lower risks).

The second mechanism is dis-economies of scope in the development of new technology. If the cost of adaptation is too high, in terms of R&D effort, innovators face dis-economies of scope that reduce the number of new technologies supplied to the market. As a result, innovators develop technologies that are only adapted to match the conditions of a the few areas where the marginal cost of adaptation is less or equal than the price of technology. Thus, take-up rates will depend on how similar farmer's conditions are relative to the areas for which the new variety was adapted.

The third mechanism relates to technology's features. If innovators cannot observe individual-level net returns on adoption, new technologies will likely cater to the needs of the average farmer. Under high heterogeneity across farmers, technology features (i.e., variety's traits) increase adoption among over-represented groups of farmers, but reduce take-up for the rest. For example, imagine that most farmers higher yield prefer over better taste. So, the development of new varieties will favor small yield gains over large improvements in taste. Thus, take-up depends on how successful innovators are at prioritizing technology features beneficial for the marginal adopter.

6 Adoption drivers

6.1 Learning and beliefs

Reduced-form results in table 2 show that trial participation has no impact on take-up. This could indicate that the agronomic trials were a short-lived experience with no effect on new varieties purchases. However, as reported in column 4, there is a negative effect (albeit significant at the 90% confidence level) of farmers losing the agronomic trial –roughly 17% of trials were lost due to extreme weather, plagues, or mismanagement. These mixed results indicate that even if participation in the trials does not incentivize farmers to take-up a new variety, they seem to respond to their individual experience during the trials. So, how does the trial experience, not just participation, influence farmers' adoption decision?

Here I explore two possible explanations: learning-by-doing and biased beliefs. First, trial data shows that non-adopters did not underperform in the trials relative to adopters. Figure 5 compares the distribution of trial yields between adopters and non-adopters (panel A), and by region (panels B and C). The kernel density of yield values also show that the average trial yield is not significantly higher for adopters (difference in means= 0.22 kg, p-value= 0.19). The standard deviation is about the same for both groups too (1.61 kg versus 1.50 kg). In the south, the cumulative distribution of non-adopters' yield is higher for values below the mean yield (1.95 kg). Little differences are present for the upper tail of the yield distribution, suggesting an asymmetric response to trials yield. Furthermore, the gap between distributions at zero values shows again the negative effect that lost trials had on adoption decisions in the north. Kolmogorov-Smirnov tests confirm that, for the whole sample, trial yield distributions of adopters and non-adopters are not different (p-val=0.219). In the south, the distributions are statistically different but only at a 90% significance level.

Taken together, these results suggest that trial productivity alone does not predict adoption. In the south, however, it is possible that the asymmetric response could be evidence of performance misattribution at the low-end of the productivity distribution. This explanation is unlikely because farmers could compare the trial plots with their regular fields. Farmers were asked to follow the same management for all plots, so that regular plots in the farm serve as controls for individual farmer's conditions. Plot selection for the trials was also controlled by design. Thus, farmers could infer that any difference between the trial and regular plots are caused by the new varieties, correcting for misattribution issues.

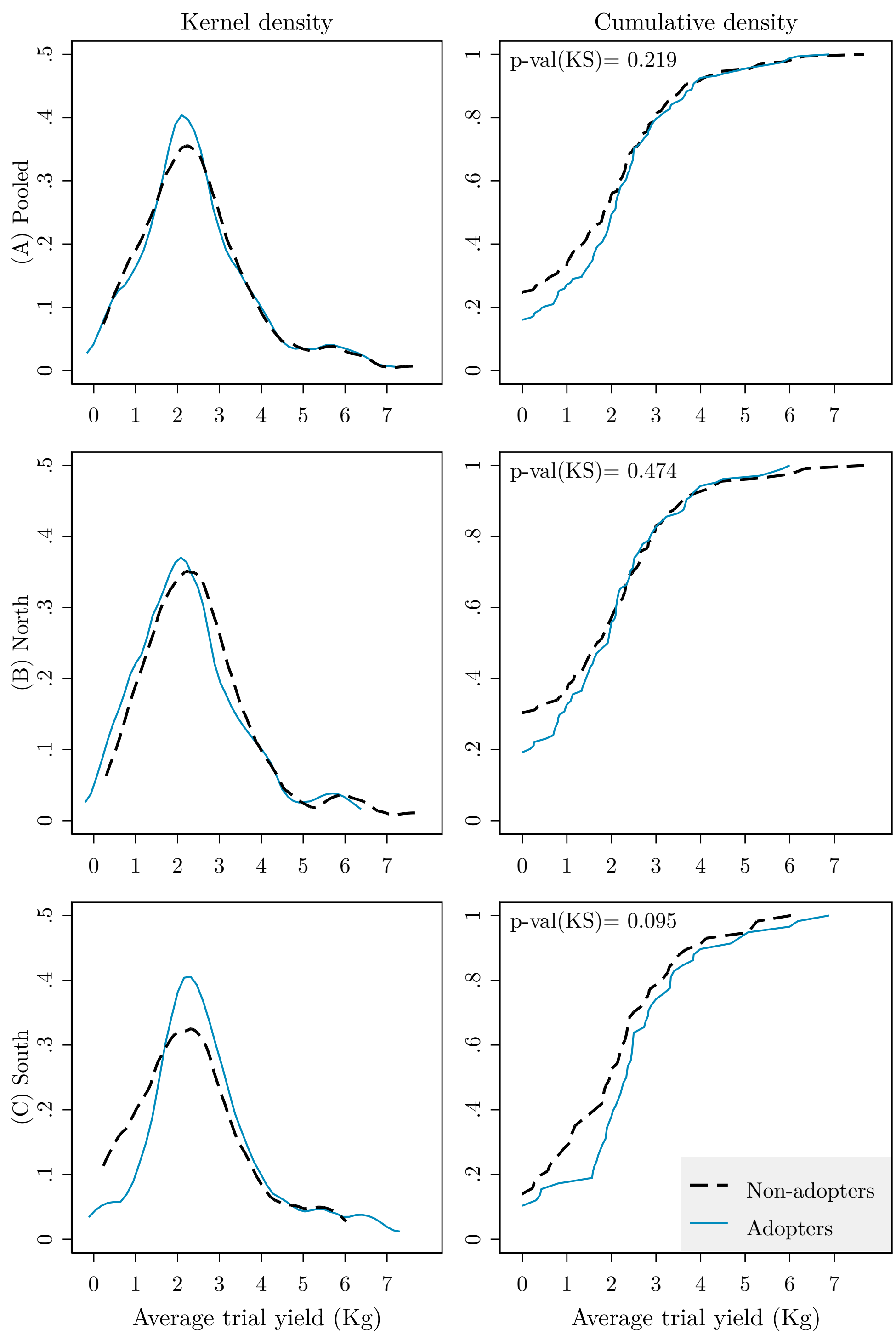


Figure 5: Graphs comparing the distribution of yield from the trials (kilograms per experimental plot) between farmers who purchased the variety offered in the intervention (Adopters) versus those who rejected the offer (Non-adopters). Left panels show the Cumulative distribution, and panels on the right the kernel density (Epanechnikov kernel and optimal bandwidth) of average yield of the trial plots. The p-value from the Kolmogorov-Smirnov tests on the equality of distribution is reported in the top-left corner of each cumulative distribution plot.

A second explanation is that farmer's expectations about new varieties performance affect how they respond to the trial results. If farmers believe that the new technology is only worth adopting for an improvement greater than a given threshold, small positive gains may not be enough to trigger adoption. This threshold may depend on the specific productive conditions and varietal preferences of each farmer. For example, low-productivity farmers may expect higher performance from an improved variety to compensate for other productive constraints. Other farmers may favor seeds with grain colors that are more attractive to consumers over significant but small yield improvements.

To test this explanation, I use baseline information to construct an expectation gap measure. At the baseline, farmers were asked to determine the yield per hectare that, on average, an improved variety is likely to produce. The expectation gap is then defined as the relative difference between the expected and current yields in percentage terms. The average expected yield is 10 quintals per hectare above the current yield, although a 14% of farmers believes a new variety would not improve productivity or even decrease yields. For reference, note that the relative price of certified seed is approximately two quintals per each quintal of commercial seed required to plant an hectare.[11] So, at the margin and ignoring transaction costs, a rational farmer should be indifferent between the seed they produce and save, and the new variety (priced as certified seed) for yield gains of two quintals.

Results in table 3 show no effect of the expectation gap alone (column 1). When interacted with trial participation (columns 2), the expectation gap shows the expected null effect for non trial farmers, and a negative but insignificant effect of -0.3% for farmers who participated in the trials. In column 3 and 4, I test whether the direct comparison between trial versus expected and baseline yields confirm these results. I find null effects for both models. This comparison is done using a binary variable for cases in which the expected (or baseline) yield is higher than the trial yield, zero otherwise, including for non-trial farmers. Other

[11] Prices for certified seeds are determined by the CNP. Formal farmers groups, such as cooperatives and associations, also produce certified seed, usually sold at the same price or slightly lower. The CNP and agricultural extension services offer seed credit scheemes in which a farmer is given a quintal of certified seed in exchange of two quintals of the commercial seed they produce.

similar results are also found comparing above and below the median expectation gap, or by restricting the sample to the trial groups (not reported).

Table 3: Expectations

	(1) take-up	(2) take-up	(3) take-up	(4) take-up
Mismatch	-0.183***	-0.182***	-0.184***	-0.184***
	(0.058)	(0.058)	(0.058)	(0.058)
Trial participation	0.119	0.131	0.132	0.110
	(0.079)	(0.089)	(0.086)	(0.098)
Expectations gap	-0.003	0.002		
	(0.008)	(0.016)		
Expectations gap x Trial		-0.006		
		(0.018)		
Baseline yield > Trial tield			-0.039	
			(0.050)	
Expected yield > Trial tield				0.008
				(0.073)
Constant	0.466***	0.458***	0.463***	0.465***
	(0.093)	(0.099)	(0.093)	(0.093)
Dependent variable mean	0.431	0.431	0.431	0.431
Village fixed effects	yes	yes	yes	yes
Baseline controls	yes	yes	yes	yes
Trial controls	yes	yes	yes	yes
R-squared	0.205	0.203	0.205	0.204
Observations	540	540	540	540

Notes: This table reports coefficient estimates from linear probability models using Pr(take-up=1) as the dependent variable. Expectation gap is measured as the relative difference between expected and current yields at the baseline. Baseline controls include education and farm size. Trial controls include seed replacement and lost trials. Robust standard errors clustered at the village level in parenthese. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

6.2 Heterogeneous mismatch effects

In this section, I investigate the factors driving the mismatch effect on adoption. As explained in section 3, high heterogeneity across farmers may impede the technological change by limiting research and development. If innovators are unable to capture that heterogeneity or there are no mechanisms motivating them to internalize it, new technologies will only respond to the preferences of a specific group of farmers. Under these constraints, how representative is that group of farmers, who they are, and how innovators target them would determine how successful the new technology is in terms of performance and adoption.

To test these hypotheses, I study the heterogeneous mismatch effects on take-up relative to innovators' work conditions. First, I use the travel time from each village to the experimental station where the new varieties where developed and tested.[12] This distance is an indicator of the relative effort innovators should invest to get to know farmers' local conditions. Given that the experimental station is located Costa Rica's central valley, on average 245 km away from farmers' location, the time traveled is also rough measure of the differences in conditions between farmers' locations and where innovators work.

Table 4 reports regression models estimating mismatch effect on take-up depending on how close farmers are to the experimental station. Column 1 reports a simple model estimating the pure mismatch for reference. Model in column 2 includes the travel time variable interacted with the indicator variable for the groups that received the recommended variety offer. Results for this models show a negative and significant mismatch effect that increases with travel time. Column 3 shows the heterogeneous mismatch effects over the distribution of travel times (quintiles). The overall pattern shows null mismatch effects on farmers closest to the the station (-3 percent points for the first quintile coefficient), followed increasingly negative effects for farmers farther away (up to -25 percent points for the farthest farmers).

[12] Researchers also conducted several test on farmers fields, most of them in the south (villages of Veracruz and Changuena). All development, however, was done in the experimental station Fabio Baudrit Moreno, part of the Universidad de Costa Rica in Alajuela, a city located 20km northwest from the capital San Jose (10.0073° N, 84.2659° W).

Table 4: Travel distance

	(1) take-up	(2) take-up	(3) take-up
Mismatch	-0.204***	0.077	-0.029
	(0.046)	(0.169)	(0.083)
	(0.050)	(0.236)	(0.094)
Mismatch x travel time		-0.062*	
		(0.036)	
		(0.049)	
Mismatch x 2nd quintile travel time			-0.183*
			(0.101)
			(0.080)
Mismatch x 3rd quintile travel time			-0.203*
			(0.107)
			(0.092)
Mismatch x 4th quintile travel time			-0.207**
			(0.102)
			(0.118)
Mismatch x top quintile travel time			-0.254**
			(0.102)
			(0.130)
Constant	0.517***	0.520***	0.520***
	(0.060)	(0.059)	(0.060)
Dependent variable mean	0.431	0.431	0.431
Region fixed effects	yes	yes	yes
Baseline controls	yes	yes	yes
Trial controls	yes	yes	yes
R-squared	0.047	0.053	0.057
Observations	540	540	540

Notes: This table reports coefficient estimates from linear probability models using Pr(take-up=1) as the dependent variable. Region fixed effects included in all specifications. Controls include baseline education level and farm size, and seed replacement and lost trials. Robust standard errors (SE) clustered at the village level in parentheses. The standard errors corrected for spatial correlation using a distance threshold of 10km are also reported below each SE estimate. Significance reported using robust SE: *** p<0.01, ** p<0.05, * p<0.1.

Overall, results support the idea that there is differential adoption relative to innovators locations as a proxy for research effort. A possible explanation is that farmers anticipate innovators being more successful in adapting the new variety to their local conditions if these conditions are similar to where innovators operate. Thus, farmers that observe a greater effort by plant breeding programs in their communities are more likely to accept the recommended offer, and that is why there are no significant differences relative to the targeted group. Outside the innovators' reach, or aside frequently visited communities, farmers may not trust the recommended offer because they expect it not to be a good match for their farms. On the contrary, if farmers being able to choose their most preferred variety based on their own experience, reputational risk are less relevant in adoption decisions.

However, how heterogeneity across farmers' conditions impact adoption? and more importantly, does matching help to improve take-up of the new varieties? To answer these questions, I test whether specific weather conditions explain differences between mismatch and targeted groups. To do so, I use precipitation and temperature data, as well as baseline information on weather shocks.

Note that the new varieties were developed to improve drought and extreme heat resistance compared to the current varieties in the market. During the intervention these varieties were marketed as such to all farmers. So, I focus on the differences between treatment groups of extreme temperature and low precipitation conditions. If researchers are successful at prioritizing traits demanded by farmers, we should expect no differences treatment groups because both, the recommended and the most preferred variety, should make the marginal adopter better off. On the contrary if drought events, for instance, are not relevant for farmers, we should expect their demand for the new varieties to be lower among the mismatch group relative to the targeted group, which is composed by farmers who can choose whatever variety they prefer, regardless of whether the new variety is (or expected to be) drought resistant or not.

Table 5 reports results from regression models estimating the effect of extreme weather

Table 5: Heterogenous effects of weather conditions

	(1) Uptake	(2) Uptake	(3) Uptake	(4) Uptake
Mismatch	-0.194***	-0.189***	0.523**	-0.245***
	(0.056)	(0.057)	(0.257)	(0.055)
	(0.052)	(0.049)	(0.205)	(0.082)
Mismatch x Flood event	-0.067			
	(0.181)			
	(0.139)			
Mismatch x Drought event		-0.253***		
		(0.128)		
		(0.059)		
Mismatch x Dry spell			-0.026***	
			(0.008)	
			(0.007)	
Mismatch x Hot days				0.001
				(0.001)
				(0.001)
Constant	0.546***	0.542***	0.508***	0.513***
	(0.062)	(0.063)	(0.060)	(0.060)
Dependent variable mean	0.431	0.431	0.431	0.431
Village fixed effects	yes	yes	no	no
Region fixed effects	no	no	yes	yes
Controls	yes	yes	yes	yes
R-squared	0.205	0.211	0.057	0.047
Observations	540	540	540	540

Notes: This table reports coefficient estimates from linear probability models using Pr(take-up=1) as the dependent variable. Village-level fixed effects included in models (1) and (2). Region fixed effects included in models (2) and (3) (north=1) . Controls include baseline education level and farm size, and seed replacement and lost trials. Robust standard errors clustered at the village level in parentheses. The standard errors (SE) corrected for spatial correlation using a distance threshold of 10km are also reported. Significance using corrected SE: *** p<0.01, ** p<0.05, * p<0.1.

variables. The objective of these regression is to determine if the mismatch effects change conditional on weather related events and conditions. Column 1 reports estimates for interaction term between the mismatch (i.e., indicator of assignment to groups that received the recommended variety) and an indicator variable that takes the value of one if the farmer experienced floods, cold weather or storms at the baseline. This model works as a sort of falsification test since the new varieties should offer no improved resilience to these events. As expected, I find no differential mismatch effect for flood events. Model in column 2 reproduces this model but using drought and extreme heat related events. Results show a stronger mismatch effect among farmers who experienced this type of extreme weather shocks prior to the intervention.

To confirm these results, models in columns 3 and 4 in table 5 use atmospheric information from the ERA5 Global Reanalysis database (Hersbach et al., 2020).[13] Column 3 shows the effects of dry spells during 2021, the year prior to the intervention. Dry spell is defined as the number of consecutive dry days (precipitation < 1 mm). Results show a negative and significant coefficient for the interaction term between mismatch and dry spell, which is consistent with the drought event results of column 2, meaning that mismatch effects are larger for who experienced a longer dry spell. Lastly, model in column 4 estimates the differential effect of hot days (daily max temperature $>$ 75th percentile average temperature). Column 4 results show an insignificant effect of hot days relative to the average mismatch effect.

7 Discussion

[*Editorial note:* Discussion and specific conclusions will be included when the data collection and analysis finalize. A full version of the paper expected to be completed by August, 2023.]

[13]The resolution of the ERA5 grid is 25x25 km at the Equator. For each farm location, precipitation and temperature were calculated as the daily average of the surrounding area (5 km radius).

References

Barrett, C. B., Moser, C. M., McHugh, O. V., and Barison, J. (2004). Better technology, better plots, or better farmers? identifying changes in productivity and risk among malagasy rice farmers. *American Journal of Agricultural Economics*, 86(4):869–888.

Beintema, N. M., Nin-Pratt, A., and Stads, G.-J. (2020). ASTI global update 2020. Technical report.

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

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

Dar, M. H., de Janvry, A., Emerick, K., Sadoulet, E., and Wiseman, E. (2021). Private input suppliers as information agents for technology adoption in agriculture. Working paper.

Duflo, E., Kremer, M., and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review*, 101(6):2350–2390.

Emerick, K., de Janvry, A., Sadoulet, E., and Dar, M. H. (2016). Technological innovations, downside risk, and the modernization of agriculture. *American Economic Review*, 106(6):1537–1561.

Fuglie, K., Gautam, M., Goyal, A., and Maloney, W. F. (2019). *Harvesting Prosperity: Technology and Productivity Growth in Agriculture*. Washington, DC: World Bank.

Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3):1311–1353.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Laajaj, R., Macours, K., Masso, C., Thuita, M., and Vanlauwe, B. (2020). Reconciling yield gains in agronomic trials with returns under african smallholder conditions. *Scientific Reports*, 10(1).

Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in china. *Review of Economics and Statistics*, 95(4):1386–1403.

Michelson, H., Fairbairn, A., Ellison, B., Maertens, A., and Manyong, V. (2021). Misperceived quality: Fertilizer in tanzania. *Journal of Development Economics*, 148:102579.

Moscona, J. and Sastry, K. (2021). Inappropriate technology: Evidence from global agriculture. *SSRN Electronic Journal*.

Stewart, F. (1977). *Technology and Underdevelopment*. Palgrave Macmillan UK.

Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica*, 79(1):159–209.

The World Bank (2023). Costa Rica, Country Overview. https://www.worldbank.org/en/country/costarica/overview. [Online; accessed 19-Dec-2022].

van Etten, J., de Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., Fadda, C., Gebrehawaryat, Y., van de Gevel, J., Gupta, A., Kiros, A. Y., Madriz, B., Mathur, P., Mengistu, D. K., Mercado, L., Mohammed, J. N., Paliwal, A., Pè, M. E., Quirós, C. F., Rosas, J. C., Sharma, N., Singh, S. S., Solanki, I. S., and Steinke, J. (2019). Crop variety management for climate adaptation supported by citizen science. *Proceedings of the National Academy of Sciences*, 116(10):4194–4199.

Appendix

Table A1: Sample comparison with CNP farmers population

Variable	Sample (n=800)	CNP registry (N=2959)	Difference p-value
Bean production (quintal)	91.40	91.36	0.99
Planted area (ha)	4.40	4.53	0.52
Yield (quintal/ha)	41.05	40.49	0.49
Seed quantity	75.81	77.17	0.77
Gender (female=1)	0.20	0.22	0.25
Associated (yes=1)	0.27	0.25	0.26
Region (north=1)	0.63	0.59	0.04

Notes: This table compares the study sample with the small- and medium-scale farmers registered in the National Productive Council of Costa Rica (CNP) for the 2020-2021 period. Differences are estimated using differences in means tests. Bean production and yield include red and black common beans. Areas only included those plots destined to bean production. A quintal of seed refers to 46 kilograms bags. Seed quantity is the ammount of certified seed used. Associated captures membership to any farmers group (associations and cooperatives).

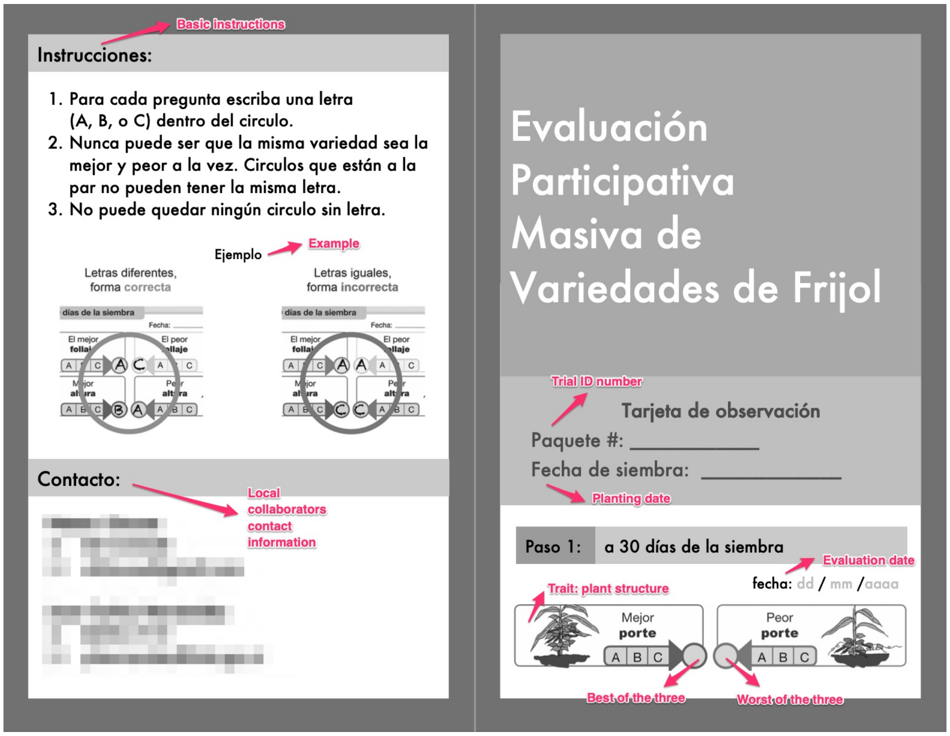
Table A2: Productivity comparison between reference and new varieties

	Yield (g)			
variety	mean	std. dev.	diff.	p-value
(A) All trials (N=400, lost=17.6%)				
Reference	2580	1543		
SEF-42	2237	1278	-343	0.012
SEF-60	2598	1651	18	0.586
SEF-62	2421	1596	-159	0.155
SEF-64	2296	1371	-284	0.081
SEF-71	2533	1470	-47	0.320
(B) South region (N=140, lost=5.4%)				
Reference	2538	1422		
SEF-42	2465	1370	-343	0.785
SEF-60	2365	1638	-173	0.557
SEF-62	2574	1276	37	0.898
SEF-64	2540	1276	2	0.992
SEF-71	2699	1404	161	0.545
(C) North region (N=260, lost=24.1%)				
Reference	2646	1680		
SEF-42	2061	1198	-585	0.020
SEF-60	2762	1646	116	0.148
SEF-62	2340	1393	-306	0.185
SEF-64	2139	1393	-507	0.056
SEF-71	2382	1490	-264	0.487

Notes: Table compares the average yield of each new variety in the agronomic trials versus the reference variety (Cabecar). Each variety only appeared in half of the testing sets that farmers received. Lost trials refer to attrition, non-compliance, and lost agronomic plots due to weather or biotic related causes. P-values are estimated using difference in means tests. Panel A reports all trials pooled. Panels B and C report results for the southern (Brunca) and northern (Chorotega and Huetar) regions, respectively.

Figure A1: Trials Scorecard

(a) Front



(b) Back