You Reap What (You Think) You Sow? Misperceptions of Improved Maize Varieties in Ethiopia

Paola Mallia *

January 23, 2026

Abstract

The adoption of improved seed varieties has the potential to substantially increase agricultural productivity. However, only 23.3 percent of farmers who grow an improved maize variety report doing so in Ethiopia. This paper provides causal evidence of the impact of misperceptions about improved maize varieties on farmers' production decisions, productivity, and profitability. I employ an Instrumental Variable approach, taking advantage of the roll-out of a governmental program that increases transparency in the seed sector. I find that farmers who correctly classify the improved maize variety they grow experience large increases in input usage (urea, NPS, labor), yields, and profits per hectare. However, there are no statistically significant changes in other agricultural practices. Using machine learning techniques, I develop a model of interpolation to predict objectively measured varietal identification from farmers' self-reported data, providing a proof-of-concept for scalable approaches to obtaining reliable measures of crop varieties and allowing for the extension of the analysis to a nationally representative sample.

*Paris School of Economics. I would like to thank the Standing Panel of Impact Assessment (SPIA) of the CGIAR, the Living Standard Measurement Study - ISA team, and the Agricultural Transformation Agency of Ethiopia. I thank Karen Macours, Sylvie Lambert, Tavneet Suri, Marcelo Gantier-Mita, and the participants of the seminars at PSE for their insightful comments. I thank ANR for the support received through the EUR grant ANR-17-EURE-0001. All remaining errors are my own. Contact: paola.mallia@psemail.eu

1 Introduction

It is widely accepted that agricultural productivity in sub-Saharan Africa has grown only modestly over the past six decades compared with other regions, such as Asia and Latin America. Given that roughly half of the population is employed in agriculture, this slow growth is often cited as a key reason why the region has struggled to achieve sustained economic development and why poverty remains pervasive (Suri & Udry, 2022). Persistently low productivity is frequently linked to the limited adoption of improved agricultural technologies. Despite the apparent profitability of many available innovations, adoption rates remain uneven, raising the question of why certain technologies fail to diffuse widely (Macours, 2019). A clearer understanding of the constraints to adoption, as well as the optimal allocation of complementary inputs, is therefore considered essential for explaining an important component of economic growth (de Janvry & Sadoulet, 2020). Recent assessments of technology use in sub-Saharan Africa suggest a mixed picture. On the one hand, average use of inorganic fertilizer and agrochemicals remains low. On the other hand, government efforts to promote modern inputs—particularly in maize production—appear to have achieved some success. Nevertheless, these inputs are rarely applied jointly at the plot level, despite substantial evidence of agronomic complementarities (Sheahan & Barrett, 2017).

This paper provides causal evidence of the effect of farmers' wrong beliefs of whether the maize variety used is improved on production decisions, productivity and profitability. For the purpose of this analysis, "improved varieties" refer to those containing CGIAR-related germplasm. CGIAR germplasm stems from breeding initiatives that deliberately focus on traits anticipated to yield high returns, indicating adaptation to the local context and incorporation of enhanced traits. Consistent with the behavioral economics literature (Angner & Loewenstein, 2007; Kahneman & Tversky, 1984; Tversky & Kahneman, 1973), I show that farmers act upon their beliefs about the improved varietal traits through the allocation of complementary inputs, which in turn affects yields and profits achieved. I rely on the Ethiopia Socioeconomic Survey collected in 2018/2019 (ESS4) to recover a measure of farmers' correct identification of maize improved varieties, that is whenever the variety grown on one plot is identified as improved both by the farmer and by the objective measure (DNA fingerprinting).[1] To isolate the behavioral adjustment of farmers' beliefs about the variety, I focus on the sample of farmers who grow an improved maize variety according to the objective measures, but who have either correct or incorrect beliefs about it (i.e. true positive and false negative, respectively).[2] I take advantage of administrative data

[1] DNA fingerprinting is commonly accepted as the gold-standard measure for crop varietal identification. A detailed description of the methodology can be found in Kosmowski et al. (2021)

[2] Farmers' self-reported identification of whether the cultivated variety is improved or not is gathered through the following question: *What type of Seed / was used for [CROP] on [FIELD]?* where the possible an-

on districts' exposure to a governmental program that favors transparency in the seed supply chain, the Direct Seed Marketing (DSM). The exogenous variation used in this paper is its interaction with households' proximity to the closest urban center, which captures within-district variation in exposure to the program. Longer exposure to the program facilitates the creation of an environment where farmers learn to recognize the type of variety grown and are able to experiment dynamically with agricultural investments. Correct classification leads farmers to adopt more fertilizers per hectare of land in the form of urea (164 kg/ha increase) and NPS (130 kg/ha increase), employ more hired labor at harvest and plant less seeds per hectare of land. I do not find evidence of any impact on agricultural practices such as, conservation agriculture, soil and water conservation, and agricultural water management practices. The gain in yields experienced by those who correctly classify is substantial, accounting for a 120 percent increase. Given that both true positives and false negatives cultivate an improved variety, the differences in yields can be attributed to the different production decisions. I do not find statistically significant increases in absolute profits, but instead large increases in profits per hectare (an increase of 253 USD per hectare of land). Additional results suggest that farmers, who correctly identify the improved variety grown on at least one of their maize plots, are not systematically reallocating resources to maize cultivation at the expense of other crops, but instead, they are crowding-in additional resources. Finally, applying machine learning techniques allows to predict the objective measure for varietal identification with 85 percent accuracy using farmers' self-reported data. This enables to extend the analysis to the nationally representative sample, thus addressing potential concerns linked to low statistical power and finding qualitatively similar results.

This study contributes to the literature in multiple ways. The literature on agricultural technology adoption has investigated different potential constraints that farmers may face[3], such as credit constraints, risk market inefficiencies (Karlan, Osei, Osei-Akoto, & Udry, 2014); input and output market inefficiencies (Ashour, Gilligan, Hoel, & Karachiwalla, 2019; Bold, Kaizzi, Svensson, & Yanagizawa-Drott, 2017; Michelson, Fairbairn, Ellison, Maertens, & Manyong, 2021; Shamdasani, 2021; Suri, 2011); behavioural constraints (Duflo, Kremer, & Robinson, 2011); and lack of knowledge about profitability and proper use (Foster & Rosenzweig, 2010). I provide a complementary explanation by relaxing the underlying assumption in the literature, which assumes that farmers possess complete awareness of which technology is being used. Not all agricultural technologies have traits that are easy to identify given farmers' knowledge. Improved seed varieties, in particular, tend to exhibit distinguishable traits only after germination, and even in that case, there may be differences in farmers' ability to interpret a noisy yield signal. Farmers may have wrong

swers are: *1. Traditional; 2. Improved/New; 3. Improved saved from the previous year; 4. Improved/Recycled.* Answers from 2 to 4 are coded as "Improved". (Section 4 - question 11 of Post-Planting survey of ESS4.)

[3]See Foster and Rosenzweig 2010, de Janvry, Sadoulet, and Suri 2017, Magruder 2018 for literature reviews.

beliefs when using or, alternatively, foregoing to use an improved variety. Most improved varieties perform well relative to non-improved varieties even under unfavorable production conditions, but they express their full yield potential only with favorable management (Gollin, Morris, & Byerlee, 2005). For this reason, the mere decision to adopt does not automatically imply achieving the full potential of one technology, as per agronomists' estimates. Yield increases predicted in agronomic trials do not immediately translate into the same gains when applied in the field, since farmers' behavioral adjustment to the innovation tends to lower the expected gains. Thus, correctly identifying the variety grown should, in principle, allow farmers to make the correct behavioral adjustment in complementary input allocation to achieve the yield gains promised by the technology. Whenever farmers misidentify the technology in use, observed input misallocation might reflect either their failure to adapt to the technology or their successful optimization of the objective function based on incorrect beliefs about the true technology. The findings of this study indicate that farmers achieve higher productivity levels when they hold accurate beliefs about the technology in use. While I refrain from drawing conclusions on the optimality of the new input allocation, considering the extremely low levels of adoption of inputs in the sample and the complementarity between inputs use and improved varieties, the additional amounts of usage deriving from correct classification can be considered a positive improvement with respect to low national averages.

Secondly, as the instrumental variable is identified by the cumulative length of exposure to a policy intervention promoting transparency in the seed system, it offers suggestive evidence of learning regarding agricultural technologies and the allocation of complementary inputs. This contribution is aligned with existing literature that acknowledges farmers' learning processes as a critical factor influencing decisions regarding technology selection and input allocation. In a learning-about-productivity model, individuals' beliefs about new technology evolve through experience. Farmers learn about the overall profitability of a new technology based on their experiences and compare it to the assumed profitability of the well-established existing technology (Besley & Case, 1993; Besley, Case, et al., 1994). However, uncertainty about the profitability of a new technology is not the sole challenge that farmers must overcome. Learning may also involve acquiring information on how to optimally manage the new technology (Foster & Rosenzweig, 1995). Agricultural research organizations and extension agents conduct controlled experiments on new technologies, allowing them to determine maximum possible yields and, in some cases, maximum profitability given specific price scenarios. However, they cannot always provide information on the best practices for achieving these yields considering the unique characteristics of individual farmers' plots (Laajaj, Macours, Masso, Thuita, & Vanlauwe, 2020). Consequently, farmers may need to experiment with a new crop on their land to determine the correct input applications. As farmers decipher the optimal management of a new technology, the

technology's profitability increases as knowledge accumulates.

Third, I contribute to the wider literature on measurement error in agricultural surveys that investigates the extent of nonclassical measurement error in land size, soil quality, yields, and crop variety identification (Abay, Abate, Barrett, & Bernard, 2019; Abay, Bevis, & Barrett, 2021; Carletto, Savastano, & Zezza, 2013; Gourlay, Kilic, & Lobell, 2019; Ilukor et al., 2017; Kosmowski et al., 2019; Wossen et al., 2019), by providing empirical evidence that the implications of measurement error can go beyond statistical inference. The availability of novel measurement methods in agricultural surveys has provided the opportunity to investigate whether farmers are able to identify correctly the crop variety grown on their plots as improved. A fast-growing literature documents that there are remarkable discrepancies between farmers' belief of the seed variety grown and objective measures of varietal identification. This evidence spans through different countries and crops: for example, in a sample of cassava growers in Colombia, Floro IV, Labarta, Becerra López-Lavalle, Martinez, and Ovalle 2018 find that farmers correctly identify the variety grown as improved only on 9.2 percent of their plots. Similar percentages are also reported in Maredia et al. 2016 for farmers growing cassava in Ghana (13 percent) and beans in Zambia (9 percent). Finally, Wossen et al. 2019 documents 43 percent of correct varietal identification for cassava farmers in Nigeria and Kosmowski, Alemu, Mallia, Stevenson, and Macours 2020 report that farmers correctly classify the variety grown on 23 percent of maize plots in Ethiopia[4]. Despite the importance of investments in complementary inputs to obtain benefits for improved varieties (Emerick, de Janvry, Sadoulet, & Dar, 2016), little attention has been devoted to investigating the effect of farmers' mistaken beliefs of the variety grown on decision-making processes and behaviors, which is the goal of this study.

Finally, I show the possibility of predicting objective measures from self-reported data with a high level of accuracy using machine learning techniques. Although validated gold standard measures are becoming increasingly available at low-cost thanks to technological advancement, national statistical agencies may be unable to easily scale at the national level the best survey methodologies. Nevertheless, accurate and timely crop production statistics are critical to adequate government policy responses and the availability of reliable measures is pivotal to establishing credible performance evaluation systems. Therefore, the application developed in this paper serves as a proof-of-concept to the use of self-reported data and objective measures collected on a regionally representative subsample to recover national figures of adoption.

The rest of the paper is organized as follows. Section 2 develops a conceptual framework on the role of misperception in farmers' input allocation decisions and economic outcomes. Section 3 provides an overview of government programs targeting the Ethiopian seed sector.

[4]In this paper I rely on the same data as in Kosmowski et al. 2020, available at https://www.openicpsr.org/openicpsr/project/124681/version/V9/view.

Section 4 describes the data and presents descriptive statistics. Section 5 outlines the empirical strategy. Section 6 presents the main results, which are discussed in Section 7. Section 8 provides empirical evidence supporting the validity of the exclusion restriction. Section 9 tests for misreporting. Section 10 develops a machine learning model to predict objective measures from self-reported data and extends the analysis to the nationally representative sample. Section 11 concludes.

2 Conceptual Framework

Improved seed varieties exhibit latent benefits that are difficult for farmers to observe prior to planting and may remain ambiguous even post-harvest. Yield-enhancing traits require complementary inputs to manifest fully, while tolerance traits (e.g., drought resistance) only become evident under specific environmental conditions. This informational opacity creates a fundamental challenge: farmers must commit to input investments before fully understanding the production technology at their disposal.

This section develops a conceptual framework to characterize how misperceptions about seed variety identity affect input allocation decisions and economic outcomes in smallholder agriculture. The framework emphasizes two interrelated learning processes: (*i*) learning to *recognize* which agricultural technology is being employed, and (*ii*) learning to *optimally manage* that technology conditional on correct identification. I demonstrate that misperception disrupts both processes, generating inefficiencies in resource allocation and reduced profitability. The framework generates testable predictions regarding input use, yields, and profitability that guide the subsequent empirical analysis.

2.1 The Smallholder Information Problem

Smallholder farmers in developing countries face a distinctive set of constraints that shape their returns to information. Unlike large-scale commercial agriculture, smallholders operate on very limited land areas, typically measured in fractions of hectares rather than tens or hundreds of hectares. This scale imposes two features that dominate economic decision-making.

First, absolute returns are constrained by scale. Even substantial percentage increases in per-hectare profitability translate into modest absolute gains when multiplied by small plot areas. This mechanical relationship between intensive and extensive margins has important implications: per-hectare measures can be informative about productivity improvements in smallholder contexts.

Second, upfront information is critical for optimizing complementary input decisions. Unlike other inputs that can be adjusted over the growing season (e.g., weeding intensity,

pest management), fertilizer application decisions are largely irreversible and must be made weeks or months before harvest reveals productivity outcomes. In the presence of uncertainty about seed variety identity, farmers face a compound decision problem: they must simultaneously choose input quantities and form beliefs about the underlying production function these inputs will augment.

This second feature distinguishes information constraints from credit constraints, though both may bind simultaneously. A farmer with access to sufficient capital to purchase fertilizer will nevertheless systematically under-invest if she fails to recognize that her seed variety exhibits strong input complementarities. Conversely, a farmer who correctly identifies her variety but faces binding credit constraints will optimize conditional on her budget but remain constrained at a second-best allocation.

2.2 Setup: Production, Costs, and Beliefs

I formalize the farmer's problem as a sequential decision under incomplete information. Consider a farmer who has acquired seed of unknown type prior to the planting season. Let $s \in \{I, T\}$ denote the true seed type, where I represents improved varieties (containing CGIAR-related germplasm) and T represents traditional or local varieties. The true variety type s has been determined at the time of seed acquisition but may not be perfectly observable to the farmer.

2.2.1 Timing

The production process unfolds over four periods:

Period 0: Farmer acquires seed through formal channels (agricultural cooperatives, government agencies) or informal channels (saved seed, informal markets).

Period 1: Farmer observes signals $\sigma \in \Sigma$ about seed type and forms posterior belief $\mu(\sigma) = \Pr(s = I \mid \sigma)$. Signals may include purchase source, price paid, physical characteristics, and extension advice.

Period 2: Based on belief μ, farmer chooses complementary input level $\theta \in \mathbb{R}_+$ (fertilizer, labor) to maximize expected profits.

Period 3: Harvest occurs, revealing output $q = f(s, \theta) + \varepsilon$ where ε is a productivity shock. Farmer observes realized profits and updates beliefs.

2.2.2 Production Technology

Output is produced according to:

$$q = f(s, \theta) + \varepsilon \tag{1}$$

where θ represents complementary inputs and $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is an i.i.d. productivity shock. I assume $f(\cdot, \cdot)$ is twice continuously differentiable with $f_\theta > 0$ and $f_{\theta\theta} < 0$.

The defining characteristic of improved varieties is stronger complementarity with inputs:

Assumption 1 (Input Complementarity)**.** *The marginal productivity of inputs is strictly higher for improved varieties:*

$$\frac{\partial f(I, \theta)}{\partial \theta} > \frac{\partial f(T, \theta)}{\partial \theta} \quad \textit{for all } \theta \in \mathbb{R}_+ \tag{2}$$

This captures the agronomic reality that improved varieties are bred specifically to respond to fertilizer and other inputs (Gollin et al., 2005).

2.2.3 Profit Maximization

Since the productivity shock ε is realized at harvest (Period 3) but farmers choose inputs θ in Period 2, farmers maximize expected profit. Given $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ with $\mathbb{E}[\varepsilon] = 0$, expected profits equal:

$$\mathbb{E}[\pi(s, \theta)] = p \cdot f(s, \theta) - c(\theta) - r_s \tag{3}$$

Under complete information, the farmer solves:

$$\max_\theta \ \mathbb{E}[\pi(s, \theta)] \tag{4}$$

The first-order condition characterizes optimal input use:

$$p \cdot \frac{\partial f(s, \theta^*)}{\partial \theta} = c'(\theta^*) \tag{5}$$

Proposition 1 (Optimal Input Intensity)**.** *Under complete information, profit-maximizing input use is strictly higher for improved varieties:* $\theta_I^* > \theta_T^*$.

Proof. By Assumption 1, $f_\theta(I, \theta) > f_\theta(T, \theta)$ for all θ. At the optimum for traditional varieties, $p \cdot f_\theta(T, \theta_T^*) = c'(\theta_T^*)$. Since improved varieties have higher marginal productivity at θ_T^*, we have $p \cdot f_\theta(I, \theta_T^*) > c'(\theta_T^*)$. Given $f_{\theta\theta} < 0$ and $c'' \geq 0$, the optimum for improved varieties must satisfy $\theta_I^* > \theta_T^*$. □

Henceforth, I use $\pi(s,\theta)$ as shorthand for expected profit, with the understanding that farmers optimize under uncertainty.

2.2.4 Incomplete Information and Farmer Beliefs

Farmers form beliefs $b \in \{I,T\}$ about seed type based on observable signals. I assume farmers choose inputs optimally conditional on beliefs:

- If $b = I$ (believes improved), applies $\theta_H = \theta_I^*$
- If $b = T$ (believes traditional), applies $\theta_L = \theta_T^*$

Cross-classifying true type (s) and beliefs (b) generates four possible states:

Table 1: Classification of Beliefs and True Seed Type

	Believes Improved	**Believes Traditional**
Actually Improved	True Positive (TP)	False Negative (FN)
Actually Traditional	False Positive (FP)	True Negative (TN)

2.3 Economic Consequences of Misperception

2.3.1 Efficient Outcomes

When beliefs align with reality (TP and TN), farmers apply appropriate input levels and achieve expected outcomes. Their realized profits approximate the complete-information optimum for their seed type.

2.3.2 False Negatives: Foregone Profits from Under-Investment

A farmer who plants improved variety $s = I$ but believes it is traditional ($b = T$) applies only θ_L inputs, achieving:

$$\pi_{FN}^* = p \cdot f(I,\theta_L) - c(\theta_L) - r_I \tag{6}$$

Had she correctly identified the variety and applied θ_H, she would have earned:

$$\pi_{TP}^* = p \cdot f(I,\theta_H) - c(\theta_H) - r_I \tag{7}$$

The profit loss from misperception is:

$$\Delta\pi_{FN} = \underbrace{p \cdot [f(I,\theta_H) - f(I,\theta_L)]}_{\text{Foregone revenue}} - \underbrace{[c(\theta_H) - c(\theta_L)]}_{\text{Avoided costs}} \tag{8}$$

This decomposes into:

1. **Foregone revenue:** Insufficient inputs fail to exploit the improved variety's yield potential

2. **Avoided costs:** False negatives save on input expenditures

Given the strong complementarity (Assumption 1), the marginal return to additional inputs is high for improved varieties, suggesting $\Delta\pi_{FN} > 0$.

Proposition 2 (Profit Loss from Misidentification)**.** *Under Assumption 1, false negatives earn strictly lower profits than true positives:* $\pi^*_{FN} < \pi^*_{TP}$.

Proof. Consider the profit function for a farmer growing improved variety $s = I$:

$$\pi(\theta) = p \cdot f(I, \theta) - c(\theta) - r_I$$

The second derivative is $\pi''(\theta) = p \cdot f_{\theta\theta}(I, \theta) - c''(\theta) < 0$, so the profit function is strictly concave in θ. By Proposition 1, θ^*_I is the unique profit-maximizing input level. True positives apply $\theta_H = \theta^*_I$ and achieve $\pi(\theta^*_I)$. False negatives apply $\theta_L = \theta^*_T < \theta^*_I$ and achieve $\pi(\theta_L)$. By strict concavity with unique maximum at θ^*_I, any $\theta \neq \theta^*_I$ yields strictly lower profits. Therefore $\pi(\theta_L) < \pi(\theta^*_I)$, i.e., $\pi^*_{FN} < \pi^*_{TP}$. □

2.3.3 Why False Negatives Persist

False positives are more likely to receive error signals: they expect high yields but observe disappointing outcomes, creating incentives to revise beliefs. In contrast, false negatives observe yields $q = f(I, \theta_L) + \varepsilon$ that may not obviously disappoint if they expected traditional-variety yields around $f(T, \theta_L)$. The productivity shock ε obscures whether disappointing realizations reflect bad luck or incorrect variety beliefs. Absent a salient counterfactual, farmers lack basis for questioning beliefs.

This asymmetry implies false negatives require *external shocks* to initiate belief revision—precisely the role played by seed market transparency improvements.

2.4 Intensive versus Extensive Margins in Smallholder Agriculture

Total household profits equal per-hectare profitability times cultivated area:

$$\Pi_{\text{household}} = \frac{\pi}{A} \times A_{\text{total}} \tag{9}$$

In smallholder contexts, farmers may be unable to expand A_{total} due to:

- **Land scarcity:** Limited availability of additional plots
- **Labor constraints:** Fixed household labor supply limits feasible scale
- **Credit constraints:** Scaling requires proportional increases in input purchases

This implies that large improvements in per-hectare profitability (π/A) may coexist with modest changes in absolute household profits ($\Pi_{\text{household}}$) when plot sizes are small and expansion is constrained. Per-hectare profitability thus becomes the appropriate welfare metric, as it captures intensive margin efficiency gains independently of extensive margin constraints.

2.5 Information Constraints versus Credit Constraints

A key empirical question is whether observed inefficiencies reflect information constraints or credit constraints. These generate different predictions:

Credit constraints predict reallocation: If farmers face fixed budgets, increased input use on correctly identified plots should come at the expense of other plots. Total input use remains constant across the farm.

Information constraints predict crowding-in: If information is the binding constraint, correct identification may trigger:

1. Updated beliefs about returns to agricultural investments generally
2. Mobilization of previously uninvested resources (savings, informal credit)
3. Improved access to formal input markets through seed purchase channels

This predicts increased total input use across all plots, not just reallocation.

2.6 Dynamic Learning and Seed Market Transparency

2.6.1 Bayesian Learning with Noisy Signals

Farmers observe signals σ about seed type and form posterior beliefs:

$$\mu(\sigma) = \Pr(s = I \mid \sigma) = \frac{\Pr(\sigma \mid s = I) \cdot \Pr(s = I)}{\Pr(\sigma)} \tag{10}$$

In opaque seed markets, signals are highly noisy due to counterfeiting, informal recycling, and uninformative physical characteristics. The likelihood ratio $\Pr(\sigma \mid s = I)/\Pr(\sigma \mid s = T)$ is close to unity, and posterior beliefs remain highly uncertain.

2.6.2 How Transparency Improves Signal Quality

Seed market transparency programs improve learning through:

1. **Direct observation:** Labeled, certified seed at retail points provides reliable information
2. **Reduced counterfeiting:** Shorter supply chains limit adulteration opportunities
3. **Enhanced social learning:** Common knowledge enables precise learning from neighbors

These mechanisms increase the informativeness of purchase source signals: $\Pr(\sigma = \text{formal channel} \mid s = I) \approx 1$.

2.6.3 Cumulative Learning

Learning is not instantaneous. Let μ_t denote a farmer's belief in period t. Belief evolution follows:

$$\mu_{t+1} = \mu_t + \lambda \cdot \mathbb{1}(\text{transparency exposure}) \cdot \eta_t \tag{11}$$

where $\lambda > 0$ is a learning rate and η_t captures period-specific shocks. Cumulative exposure facilitates:

- More experimentation cycles
- Stronger social learning networks
- Deeper market penetration

This implies current correct identification reflects accumulated past learning, not just contemporaneous effects.

2.6.4 Differential Effects by Remoteness

Transparency programs may generate larger marginal improvements for remote farmers who:

- Face higher baseline information barriers
- Have limited access to formal suppliers concentrated in urban centers
- Rely more heavily on informal dealers vulnerable to counterfeiting

This predicts the effect of transparency programs should be larger for remote households than for urban households with existing access to formal channels.

2.7 Testable Predictions

The framework generates predictions that guide the empirical analysis:

Table 2: Theoretical Predictions

Outcome	Prediction
Fertilizer use	$\theta_{TP} > \theta_{FN}$: Complementarity implies higher optimal input use for improved varieties (Proposition 1)
Labor allocation	Forward-looking farmers anticipating higher yields should pre-contract harvest labor; planting labor requirements may be inelastic
Yields	$q_{TP} > q_{FN}$: Higher input use combined with complementarity generates higher output
Profits per hectare	$(\pi/A)_{TP} > (\pi/A)_{FN}$ if foregone revenue exceeds avoided costs (Proposition 2)
Absolute household profits	Ambiguous: large per-hectare gains may translate to modest absolute gains if plot sizes are small and expansion is constrained
Input use on other plots	Crowding-in (total increase) if information constraints bind; reallocation (zero sum) if credit constraints bind
Correct identification	Increases with cumulative transparency exposure; effects larger for remote farmers facing higher baseline information barriers

The subsequent empirical analysis tests these predictions. Section 7 returns to the framework to interpret the empirical findings and assess economic magnitudes.

2.8 Scope and Limitations

This framework abstracts from several factors:

Risk: I assume risk neutrality. Risk aversion would generate additional wedges between observed and optimal input use, but conditional on risk-adjusted optima, misperception still causes deviations.

Heterogeneity: I treat farmers as ex-ante identical. Heterogeneity in ability, credit access, or learning capacity could generate differential treatment effects.

General equilibrium: I take prices as exogenous. At smallholder scales with incomplete adoption, general equilibrium effects are likely negligible.

Despite these simplifications, the framework provides tractable intuition for how information constraints shape agricultural decision-making and identifies the key empirical

patterns to investigate.

3 Seed supply and distribution in Ethiopia

The focus of this paper is on maize cultivation in Ethiopia. Maize stands out as the most prevalent cereal crop in the country, with approximately 9.8 million farmers engaged in its cultivation, as reported by the Central Statistical Agency in 2019. This crop is of paramount importance for food security and is well-suited to thrive in most of Ethiopia's main agroecological zones. Historically, the Ethiopian government held substantial control over the seed sector. The formal seed system operated under a highly centralized model, with the government assuming responsibility for seed production, distribution, and regulation. To address the persistent issues related to access to quality and transparent seed, the government took a significant step by launching the Direct Seed Marketing (DSM) program in 2011. DSM was designed as a mechanism through which both public and private seed producers directly supply adequate amounts of quality seeds to farmers in convenient places and in a timely fashion, selling the seed to the farmers directly at competitive prices. The goal of the program is to ensure transparent access to improved seeds of high quality by smallholder farmers (ATA, 2015). Importantly for this study, DSM is also intended to improve transparency and verifiability of varietal identity in seed transactions by reducing intermediaries and strengthening traceability. Under DSM, seed producers are allowed to sell seed directly to farmers, in contrast to the Conventional Seed Marketing (CSM) system in which seed passes from seed producers to regional Bureaus of Agriculture, then to woreda's Agricultural Offices and finally to Development Agents, cooperative unions, and primary cooperatives, who, in turn, sell the seed to farmers (see Figure A1). Through the DSM program, private seed firms are also involved in certified seed production, producing mainly hybrid maize seed (Alemu et al., 2019; Benson, Spielman, & Kasa, 2014). Private seed firms are able to sell their products directly to farmers through One Stop Shops which are located in the woreda capital or major urban centers or arrange deliveries to farmers cooperatives. The DSM program was initially piloted in the Amhara region in 2011, in Oromia and SNNP in 2012, and later scaled up in Tigray in 2013. The regional governments of Amhara, Oromia, and SNNP established their parastatal seed production units—Amhara Seed Enterprise (ASE), Oromia Seed Enterprise (OSE), and South Seed Enterprise (SSE), respectively—in 2009 and 2010 to provide seed for farmers in their regions. At the same time, a growing number of private seed producers and agricultural cooperatives emerged to supply maize seed primarily to the regional Bureaus of Agriculture (BoA) for distribution through their supply channels to farmers (Mekonen, Minot, Warner, & Abate, 2019). While the design did not embed specific targeting criteria to define the roll-out of the program across the country, it emerged from various official sources that regions were chosen based

on their seed production capacity and the importance of maize as a key commodity. The regional governments thus, targeted woredas based on their potential in maize production (Mekonen et al., 2019). In 2018, DSM was expanded to 228 woredas covering 63 percent of seed supply in the four major regions: 45 percent in Amhara and SNNPR, 56 percent in Tigray, and 74 percent in Oromia. In March 2018 the MoA enacted Seed Marketing Guidelines and a Certificate of Competence for the One-Stop-Shop Directives. In 2019, DSM coverage has increased to more than 290 woredas. Crops included under DSM include maize, teff, and wheat. Figure 1 is based on administrative data and shows the geographical coverage of DSM over time for maize in the sample relevant to this study.

For the purpose of this paper, the key feature of DSM is that it changes the *information environment* in which seed transactions occur. By allowing seed producers to sell directly to farmers through one-stop shops and cooperative channels, DSM reduces the number of intermediaries in the supply chain and strengthens the scope for transparent communication of varietal characteristics and seed quality. This is particularly relevant in settings where seed provenance is difficult to verify and informal channels are more exposed to mislabeling, counterfeit risk, and uncertainty about varietal identity.

This institutional setting motivates the instrumental variable strategy developed in Section 5, which exploits that longer DSM exposure is expected to improve the transparency and verifiability of varietal identity, particularly in more remote areas that rely more heavily on informal seed sourcing.

The Input Voucher Sales System (VSS, henceforth) is another initiative introduced in ATA's agenda by the Ministry of Agriculture with the overall aim to address difficulties that smallholder farmers face in accessing credit for agricultural inputs such as fertilizer, improved seeds, and labor-saving tools[5]. The program links local microfinance institutions to farmers to issue cash or credit vouchers that can be used to redeem inputs at nearby cooperative stores. As shown in Figure 2, by 2019 the VSS was rolled out in Tigray, Amhara, Oromia, and SNNP regions. Since the two programs were designed and implemented by the Ministry of Agriculture to address different but complementary issues of the Ethiopian agricultural sector, I control for the VSS presence when exploiting the length of exposure to the DSM program, even though by 2019 the program's coverage was virtually universal in our sample.[6]

[5] http://www.ata.gov.et/programs/highlighted-deliverables/input-voucher-sales-system-ivs-2/

[6] I also performed the analysis controlling for the presence of the Agricultural Commercialization Cluster and obtain very similar results which we report in Appendix 14.4.

4 Data and descriptive statistics

4.1 The Ethiopia Socioeconomic Survey

The analysis is informed by the Ethiopia Socioeconomic Survey (ESS), a regionally and nationally representative household survey collected by the Ethiopian Central Statistics Agency (CSA) with the support of the Living Standard Measurement Study (LSMS) of the World Bank and the Standing Panel on Impact Assessment (SPIA) of the CGIAR. The ESS is implemented every two years: the households in the survey were visited in 2011/2012 (ESS1), re-visited in 2013/2014 (ESS2), and then re-visited again in 2015/2016 (ESS3). For the purpose of this analysis, I mainly rely on the wave collected in 2018/19 (ESS4) and take advantage of the previous data collections to provide evidence in support of the validity of the identifying assumptions. Despite the name, the 2018/19 ESS (ESS4) is the first wave of a new panel and it covers all nine states and two cities, Addis Ababa and Dire Dawa. Unlike previous ESS waves, ESS4 is also representative of regions as well as rural and urban areas. In particular, the sample of ESS4 was extended to ensure the representativeness of regions that had previously been aggregated in an "Other region" category: Afar, Benishangul-Gumuz, Dire Dawa, Gambela, Harari, and Somali. The ESS entails comprehensive household, community, and agricultural surveys. Farmers' self-reported measures of agricultural inputs and outputs are complemented with the collection of objectively measured data of land size, harvest quantity, and crop variety for maize, obtained through GPS measurement, crop-cut, and DNA fingerprinting, respectively. However, the collection of samples for DNA fingerprinting was restricted to the subsample of maize-growing regions: Amhara, Dire Dawa, Harar, Oromia, Southern Nations, Nationalities, and Peoples Region (SNNPR), and Tigray. For each sampled EA, a random sample of a maximum of 10 fields for each crop plots belonging to 12 ESS households was selected to collect the DNA fingerprinting measures.[7] The data of the DNA fingerprinting sample is combined with the reference library of maize varieties in Ethiopia that was previously compiled under a DNA fingerprinting research project conducted by the International Maize and Wheat Improvement Center (CIMMYT) and the Ethiopian Institute of Agricultural Research (EIAR). The total sample of maize plots with DNA fingerprinting measures in ESS4 consists of 479 plots cultivated by 423 households.[8]

[7]The 12 ESS households are a subsample of 20 households selected for a wider survey (AgSS). On the randomly selected 10 plots per crop, priority was given to pure-stand over intercropped plot.

[8]The original sample of plots was meant to be 506 plot observations. However, due to technical issues during data collection the final sample that can be merged with the reference library and, household and community surveys is of 479 plot observations.

4.2 Descriptive statistics

Comparing farmers' self-reported variety and the objective measure of varietal identification for maize varieties allows to obtain a picture of misperception in our sample. Table 4 reports the share of plots by category. Notably, correct classification of improved maize varieties (true positives) occurs in only 23 percent of plots. Conversely, 37 percent of cases involve farmers growing improved varieties while believing they are planting non-improved ones (false negatives). False positives, where farmers incorrectly report planting improved varieties account for 22 percent of the sample and the remaining 17 percent exhibit accurate beliefs about growing non-improved varieties (true negatives).

Table 5 provides a comprehensive overview of key variables used in our analysis, spanning plot and household-level data. Farmers in our sample typically cultivate relatively small maize plots, averaging 0.13 hectares.[9] However, there is significant variability among different farmer categories, with false positives tending to have larger plots. Regarding fertilizer usage, it is noteworthy that farmers who believe they are growing non-improved varieties (TN and FP categories) tend to apply lower amounts of urea, DAP, and NPS per hectare. This observation suggests that farmers make agricultural decisions influenced by their beliefs rather than the true genetic variety, signaling awareness of the complementarities between improved varieties and fertilizers. In terms of labor allocation, assuming 1 day equates to 8 hours of work, farmers allocate, on average, approximately 191 days of family labor and only 21 days of hired labor per hectare during preparation and planting. Conversely, during the harvest, labor input decrease to an average of 84 days for family labor and 5 days for hired labor per hectare. Additionally, the table provides insights into various agricultural practices. Crop rotation with a legume, crop residue cover, minimum or zero tillage, terracing, and plowing along the plot contour are practiced on 27 to 36 percent of plots, separately. Soil erosion prevention measures are taken on 72 percent of the plots, and 64 percent of farmers employ Soil Water Conservation Practices. Notably, only 1 percent of the plots are irrigated. Furthermore, I look at variables related to government programs and household proximity to urban centers. The unconditional average length of exposure to the DSM program is 1.5 years, with program implementation covering 34.24 percent of the sampled woredas. I calculate household proximity to the closest woreda town, as the inverse of self-reported distance in kilometers, revealing an average distance of 21 kilometers.[10] As mentioned, the VSS initiative is implemented in nearly all woredas in the sample (99 percent). Plot suitability for maize cultivation averages 24 on a scale from 0 (less suitable)

[9] A pattern consistent with findings reported by (Desiere & Jolliffe, 2018).

[10] Proximity of household h to the closest woreda towns is defined as

$$Proximity_h = max(distance_h) - distance_h$$

where $distance_h$ is the distance of household h to the closest woreda town.

to 100 (highly suitable), with the maximum in the sample reaching 55. Farmers, on average, plant 93.5 kg of maize seed per hectare. Plot characteristics such as potential wetness, elevation, and slope exhibit limited variation from the Enumeration Area (EA) mean.

To measure yield, I take advantage of the harvest measure obtained through crop-cut (dry weighted crops) on 4x4 meters sub-plots. Whenever necessary, I adjust the plot area taking into account the percentage of land dedicated to maize in case of intercropping with other crops. Maize cultivation profits are calculated based on the monetary value of harvested quantities using market center prices. Costs considered encompass hired labor and input purchases, including transportation expenses. Notably, I do not include costs for items that were not purchased, such as family labor or recycled seed, due to data limitations. Finally, the table reveals that maize yields average almost 2.9 tons per hectare, while household profits from maize cultivation amount to approximately 34 USD per year (PPP adjusted conversion rate).

The descriptive patterns are consistent with the belief that cultivating an improved variety influences farmers' decisions regarding input usage, and this ultimately impacts yields and profits. Hence, arguably, the data in the table suggests that farmers tend to misperceive rather than deliberately misreport the variety they are cultivating: a hypothesis discussed more comprehensively in Section **??**. In addition, the number of differences in observable characteristics is likely to be correlated with differences in unobservable characteristics, which calls for an identification strategy that goes beyond selection on observables in order to establish causality.

4.3 Descriptive insights on misperception and the degradation hypothesis

While I may not be able to pinpoint the causes of misperception, the available data offer valuable descriptive insights. Table 6 outlines the characteristics of farmers across identification categories. Household-head characteristics are largely homogeneous across groups in terms of age and gender. There is a notable difference in education: farmers who correctly identify improved varieties (true positives) have nearly twice the years of schooling as false positives, though this difference is not statistically significant at conventional levels (p-value = 0.11). This suggests that education may play a role in variety recognition, though the evidence is not conclusive. Other observable farmer characteristics show limited variation across groups.

The second segment of Table 6 reports seed characteristics obtained from DNA fingerprinting. Several patterns stand out. False negatives are disproportionately associated with Open-Pollinated Varieties (OPVs): 82 percent of false negatives are OPVs, compared to only 18 percent among true positives. False negatives also report substantially less current-season

purchasing (21 percent versus 93 percent for true positives), which is suggestive of less frequent seed replacement and greater reliance on saved seed. Moreover, they rely more heavily on informal sources such as local markets (41 percent) and neighbors (13 percent). By contrast, true positives and false positives overwhelmingly report purchasing seeds in the observed season, consistent with hybrid cultivation, and obtain them through formal channels such as agricultural cooperatives, government agencies, and farmers' organizations. True positives also cultivate nearly all drought-tolerant and quality protein maize varieties in our sample, indicating that correct identification is strongly correlated with the uptake of recent breeding innovations.

These sourcing differences are also consistent with uncertainty and trust frictions in seed markets. Informal channels typically involve weaker traceability and less verifiable information about varietal identity, which can increase both misperception and farmers' uncertainty about whether a purchased seed is truly improved.

A key interpretive concern is that false negatives may not reflect misperception at all, but rather a biologically rational assessment of seed performance. In particular, a farmer could report that they are not growing an improved variety even when DNA fingerprinting classifies the seed as improved if the seed has been recycled and has lost the traits associated with improved performance. This mechanism is most plausible for hybrid maize: when farmers replant saved grain (e.g., F2 generations), hybrid vigor can be lost, so farmers may reasonably perceive the seed as "no longer improved." Since false negatives are much less likely to have purchased seed in the current season and rely more on informal sources, their self-reports could plausibly reflect this type of loss of improved traits through recycling rather than mistaken beliefs.

The descriptive patterns in Table 6, however, point in the opposite direction. First, false negatives maintain high genetic purity (93.4 percent), only 3.4 percentage points lower than true positives. If false negatives were primarily planting recycled hybrid seed with substantial loss of improved genetic content, we would expect markedly lower purity. The persistently high purity therefore suggests that the improved genetic material remains largely intact.

Second, the loss-of-traits explanation is difficult to reconcile with the variety composition of false negatives. False negatives are overwhelmingly OPV growers (82 percent), and improved OPVs are specifically bred to maintain performance across generations when saved and replanted (Pixley & Bänziger, 2004). Put differently, the biological argument that recycling causes improved traits to disappear applies primarily to hybrids, not to OPVs. The fact that false negatives are concentrated among OPVs therefore strengthens the case for misperception: many farmers appear to be cultivating improved OPVs but failing to recognize them as improved.

Taken together, these descriptive patterns suggest that false negatives are unlikely to

reflect a purely biological loss of improved traits through recycling, and are instead consistent with misperception—in particular, a failure to recognize improved OPVs as improved. Importantly, OPV cultivation should not be interpreted as a backward or inferior technology choice. Improved OPVs are bred with desirable traits and can be a profitable option for resource-constrained farmers, especially in settings where hybrid seed and fertilizer costs are high relative to grain prices. Because OPV seeds are cheaper and can be recycled with minimal yield penalties, the savings from adopting OPVs instead of hybrids can be reinvested in complementary inputs such as fertilizer, herbicide, or additional labor, potentially raising returns even when expected yields are lower than under hybrids.

In this sense, DSM may help correct misperceptions not only about hybrids, but also about the value of improved OPVs, thereby reducing the likelihood that farmers underinvest in complementary inputs or overlook varieties well-suited to their resource constraints and agroecological conditions. More broadly, DSM may affect behavior not only by changing farmers' information about varietal identity, but also by increasing the perceived reliability of seed claims and reducing quality uncertainty, which can matter for input choices if farmers are averse to such uncertainty.

5 Empirical strategy

The variable of interest for this analysis is the correct classification of improved maize varieties, that is, cases in which the crop grown on a plot is classified as improved according to both farmers' self-reports and DNA fingerprinting analysis. Throughout the paper, "improved" refers to maize varieties containing CGIAR-related germplasm as identified by DNA fingerprinting.[11]

The specific focus on farmers' correct classification of varieties containing CGIAR-related germplasm—rather than on non-improved varieties—is motivated by the fact that these varieties require specific management to exploit their yield-enhancing traits. In particular, agronomists highlight the key complementarity between improved seed adoption and fertilizer application. Moreover, tolerance traits to adverse climatic events or pests are becoming increasingly relevant, and correct knowledge is central to fully exploiting their potential and to complementing crop choices with appropriate climate-smart practices. Thus, I hypothesize that farmers' beliefs about the crop variety grown (improved vis-à-vis non-improved) shape complementary input allocation decisions and, conditional on the true variety planted, determine productivity outcomes.

[11]This implies that improved varieties outside the CGIAR germplasm pool would not be classified as improved by our objective measure. Our results are therefore interpreted as the consequences of correctly recognizing CGIAR-related improved varieties that require specific complementary management to realize their yield and profit potential.

I focus on farmers who cultivate an objectively improved maize variety according to DNA fingerprinting, comparing those who correctly report it as improved (true positives) to those who report it as non-improved despite cultivating an improved variety (false negatives). This comparison isolates the role of correct variety recognition for complementary input choices among farmers facing the same underlying seed technology.

Our main approach is to use correct identification of an improved maize variety to explain production decisions at the plot level. The main specification is therefore:

$$Y_{phw} = \beta_0 + \beta_1 CI_{phw} + \beta_2 P_{hw} + \beta_3 DSM_w + \beta_4 VSS_w + \beta_5 VSS_w * P_{hw} + \\ + \beta_6 Suitability_w + \beta_7 Suitability_w * P_{hw} + \alpha_r + \varepsilon_{phw} \quad (12)$$

where Y_{phw} is the input allocation observed on plot p of household h in woreda w, CI_{phw} is the correct identification of the improved variety on plot p of household h in woreda w, P_{hw} is the log transformation of the measure of proximity to the nearest woreda town (Km) of household h residing in woreda w, DSM_w is the number of years of exposure to the DSM program of woreda w, VSS_w is the presence of the VSS system in woreda w, $Suitability_w$ is a measure of maize suitability in woreda w and, α_r are regional fixed effects. Both VSS and Suitability are interacted with proximity to woreda towns for reasons that we will explain in greater detail in what follows. The error term ε_{phw} is clustered at the EA level which corresponds to the first level of sampling.

Thus, I define CI_{phw} as an indicator equal to one if the farmer reports cultivating an improved variety and DNA fingerprinting confirms that the plot is planted with an improved variety (true positive), and equal to zero if the farmer reports cultivating a non-improved variety while DNA fingerprinting indicates the plot is planted with an improved variety (false negative). The main hypothesis is that farmers growing improved varieties allocate complementary inputs differently when they are aware of the variety status, compared to cases in which they cultivate an improved variety but believe it to be non-improved. In addition, farmers who correctly identify the improved variety in the year of observation are also more likely to have held correct perceptions in the past, which increases opportunities to learn about optimal complementary input allocation over time and to act upon those beliefs.

In this setting, I focus on a broad definition of inputs, including land area, fertilizer application, labor activities, as well as plot characteristics such as elevation, slope, and potential wetness. Since farmers in our sample cultivate on average three maize plots, plot selection provides an additional margin through which farmers can optimize their objective function.

DSM is not interpreted as a treatment whose overall effects are of independent interest in this paper; rather, DSM exposure is used only to generate variation in farmers' correct classification of improved varieties.

Correct identification of adoption status is endogenous since farmers who misperceive adopting an improved variety differ from those who do not. These differences are evident in observed characteristics, which are likely correlated with unobserved characteristics that influence input allocation. Thus, the ordinary least squares (OLS) estimates of equation (12) would be biased. To address the endogeneity concern, I implement an Instrumental Variable (IV) approach.

Hence, I instrument correct identification by implementing the following first-stage:

$$CI_{phw} = \theta_0 + \theta_1 P_{hw} + \theta_2 DSM_w + \theta_3 P_{hw} * DSM_w + \theta_4 VSS_w + \theta_5 VSS_w * P_{hw} + \\ + \theta_6 Suitability_w + \theta_7 Suitability_w * P_{hw} + \alpha_r + \eta_{phw} \quad (13)$$

where $P_{hw} * DSM_w$ is the interaction between the number of years of exposure to DSM and proximity to woreda town and represents the instrumental variable, while the other variables are defined as above.

The instrument is motivated by the idea that DSM affects correct identification primarily by improving the transparency and verifiability of varietal identity in local seed markets. Proximity to woreda towns captures baseline access to formal outlets and information, whereas remote households rely more heavily on informal transactions where traceability is weaker. The interaction between DSM exposure and proximity therefore isolates the component of DSM-induced transparency that differentially affects remote households' ability to correctly recognize improved varieties.

As already mentioned, the DSM program allows both private and public seed suppliers to market seeds directly to farmers. Competition between different suppliers increases incentives for transparency regarding varietal characteristics in order to gain market share. Based on this, longer exposure to the program should facilitate the creation of an environment where farmers become more aware of the variety grown and are able to experiment dynamically with complementary agricultural investments. Therefore, the length of exposure to the DSM program is used in this context as a proxy for learning, both individually and socially, about the varieties available on the market, the trustworthy actors involved in the supply chain, and the reliability of varietal claims, which should ultimately increase farmers' ability to correctly recognize varietal identity, and thereby affect complementary input decisions.

At the same time, farmers who reside in areas that are closer to woreda towns should have easier access to formal channels of seed supply, *ceteris paribus*. This is because, under both the conventional seed marketing system and under DSM, the formal supply chain unwinds from higher to lower administrative levels where One-stop-shops and larger markets are available, without relying on informal dealers to perform the last mile in the distribution. Instead, farmers that reside in remote rural areas may find it too costly to travel to urban centers and thus end up relying on informal channels that are not subject to quality

controls and are more likely to engage in counterfeiting than officially recognized suppliers.

I hypothesize that DSM exposure improves information for all households, but that the marginal effect is larger in more remote areas where baseline traceability is weaker and informal sourcing is more prevalent. The interaction between DSM exposure and proximity to woreda towns is designed to capture precisely this differential improvement in the information environment.

From an identification point of view, given the non-random targeting of the DSM program, I cannot rely on exposure to DSM alone as a valid instrumental variable. Indeed, woredas that have accumulated greater exposure to DSM over the years are likely different from woredas that did not in terms of land suitability to maize cultivation and/or the presence of other initiatives promoted by the government (such as DSM), which affect input allocation. Therefore, I exploit the interaction between exposure to DSM and proximity to woreda towns to identify exogenous variation that affects input allocation only through correct classification, while controlling for other possible confounding factors.

For this reason, in both specifications, I include a measure of maize suitability and its interaction with proximity to woreda towns[12]. In addition, I control for the presence of other major agricultural programs in the woreda promoted by the Agricultural Transformation Agency (ATA) which possibly followed similar targeting criteria than DSM.[13] Controlling for some of the targeting criteria of DSM, that is VSS and maize suitability, conditioning on proximity to woreda town, relaxes concerns about a direct effect of the instrumental variable on input allocation, lending confidence to the validity of our identifying assumptions. I discuss in greater detail the validity of the exclusion restriction in Section 8.

6 Results

6.1 First-stage results

Table 7 presents the first-stage estimates for Equation 13. As anticipated, an increase in exposure to the DSM program raises the likelihood of correctly identifying improved maize varieties. Conversely, proximity to a woreda town does not exhibit a significant effect. The coefficient of the interaction term between proximity to a woreda town and years of exposure to DSM is negative. This finding suggests that as farmers become more accustomed to transparent supply channels, such as those facilitated by DSM, the influence of proximity to an urban center diminishes, and vice versa. Notably, the instrumental variable's coefficient is

[12]This measure was aggregated for the purpose of this analysis at the woreda level.

[13]In particular, I focus on the VSS described in Section 3. Using administrative data on the geographical coverage of the program allows us to take into account any difference across woredas derived from VSS's activities in terms of input use through access to credit.

substantial and statistically significant at the 1 percent confidence level. Our results indicate that our model accounts for approximately 27 percent of the variance in farmers' accurate classification of maize varieties. Moreover, the size of the F-statistic for the instrument is 14.97, indicating that our approach is unlikely to be afflicted by concerns regarding weak instrumental variables.

A valid concern is that the DSM program may, by merely increasing the availability of improved varieties, mechanically raise the probability of farmers believing that they are cultivating improved varieties without increasing transparency. To rule out this concern Table A1 shows that the instrumental variable is not predictive of farmers' reporting to grow an improved variety (column (1)), whether correctly or not; nor of adoption of an improved variety according to the objective measure (column (2)). Instead, the instrument is still relevant in predicting the correct identification of both improved and non-improved varieties, thus corroborating the transparency channel argument.

6.2 Inputs use and plot selection

Table 8 reports the OLS and Two-Stage-Least-Square (2SLS) estimates of Equation 12, in Panel A and B, respectively. Panel B shows that farmers who correctly classify the variety grown as improved allocate more inputs on their plots. Column (3) reports a large and positive effect of correct classification on the quantity of urea applied per hectare. On plots cultivated with a correctly classified improved variety, farmers apply 164 kg per hectare more than false negatives. The effect is statistically significant at the 10 percent level. While there is no detectable effect on the quantity of DAP used, the allocation of NPS is greater on correctly classified plots. In particular, the effect amounts to an increase of 129.7 kg per hectare where the mean for the plots without a correct classification is 13 kg per hectare[14]. Both DAP and NPS are blended fertilizers based on nitrogen and phosphorus, but the latter also contains sulfur. With the introduction of NPS in Ethiopia, farmers have progressively shifted from DAP to NPS (Balemi et al., 2019), which is consistent with the results.

Despite the increase in fertilizers use, I do not find any statistically significant effect on the measures of labor at planting, although the coefficients are positive and big in size. However, I find a positive and statistically significant increase in days of hired labor per hectare at harvest. The effect is 4.7 additional days of hired labor for plots where the variety is correctly identified, which corresponds to an 85 percent increase with respect to the (control) average. As in Beaman, Karlan, Thuysbaert, and Udry 2013, the adoption of improved seeds and/or fertilizers is not associated with a significant change in the amount of family labor applied, but increased expenses in hired labor. Indeed, the results are in line with the

[14]The effects on the extensive margin for urea and NPS use are reported in Table A2. The large impact reflects important changes at the extensive margin, with an increase of 18 percentage points and 17 percentage points, starting from 15 for urea and 9 for NPS.

complementarity between improved varieties and fertilizer use predicated by agronomists, and cultivating a yield-enhancing variety translates into a greater amount of work at harvest. Moreover, I test if farmers select different types of plots for the cultivation of improved seeds if they know they are improved. Plots where farmers are aware of growing improved varieties do not exhibit deviation from the community average in terms of potential wetness and slope however, farmers select plots with slightly lower elevation. Finally and importantly, the coefficient on seedling rate in column (11) is negative and big in size. This result can be interpreted as a signal that farmers who correctly classify are confident of the quality of seed used since it is not unusual to observe farmers increasing the quantity of seed planted in order to increase the probability of germination of the crop cultivated. I do not find any statistically significant effect on plot area under maize cultivation from correct classification, although the coefficient is positive and big in size.

Even though the survey instruments used in this analysis do not provide the exact timing of fertilizer application, the fact that the different nutrients are generally applied at different times in the growing cycle helps speculate on the investment decisions of farmers who correctly classify. In particular, phosphorous and potassium (NPS) contribute mostly to soil conditioning and roots development, thus it is usually applied at the time of planting. In contrast, nitrogen – in the form of urea – contributes to healthy plant and leaf development and therefore is mostly applied later in the season. The picture that emerges when considering the results on fertilizer applications and seedling rate, suggests that farmers who correctly classify improved varieties are confident of the quality of their seeds which makes them willing to make early investments. This is also in line with a cumulative effect of learning from experimentation over years, which again lends confidence to the story outlined in Section 2 (Conceptual Framework) and Section 6.1 (First-Stage results).

Comparing the coefficients reported in Panel A and Panel B of Table 8, I observe that the IV coefficients of the different outcomes of interest are larger than the OLS estimates most of the time. This discrepancy could be attributed to differences in how the IV and OLS coefficients place weights on the covariates (i.e. the covariate weight difference) and the endogeneity bias.[15] I perform the IV-OLS gap decomposition developed in Ishimaru 2022 to assess the contribution of these two terms to the overall gap. The results reported in Table A3 show that in the vast majority of the cases, most of the gap can be attributed to the endogeneity bias from omitted variables. Indeed, it is plausible to assume that farmers with greater (unobserved) ability use more inputs in comparison to farmers with low ability. However, since farmers with low ability are less likely to have access to improved varieties, they are consequentially unlikely to be affected by misperception, given that they only adopt non-improved varieties. The interaction between these two effects is consistent with

[15] Given that the treatment is binary, the difference attributable to weights on treatment levels is mechanically zero.

a negative omitted variable bias. Instead in the case of hired labor, while we still expect low-ability farmers to be less likely to misclassify an improved variety, this time high-ability farmers should hire less labor given a potentially higher marginal productivity of family labor. While recent studies take misclassification of improved varieties either as exogenous (Wossen, Abay, & Abdoulaye, 2022) or assume selection on observable characteristics (Bohr, Deisemann, Gollin, Kosmowski, & Lybbert, 2022), these results suggest that ignoring the endogeneity bias due to unobservables would lead to an underestimation of the true effect of correct identification on inputs application/[16]

6.3 Cultivation practices

Departing from input allocation, another interesting hypothesis to investigate relates to whether the correct classification of improved varieties grown encourages the adoption of other agricultural practices. Research over the last decade from CIMMYT shows that improved maize varieties are compatible with sustainable intensification practices. These promising findings have motivated efforts such as the Sustainable Intensification of Maize-Legume Cropping Systems for Food Security in Eastern and Southern Africa (SIMLESA) program introduced by CIMMYT in 2010 with the aim to develop sustainable intensification based on conservation agriculture (CA), among others. Indeed, empirical evidence on the joint impact of modern seed adoption and conservation agriculture practices shows promising results in increasing farmers' income and reducing production costs (Teklewold, Kassie, Shiferaw, & Köhlin, 2013). In particular, conservation agriculture (CA) is based on a set of principles for on-farm natural resource management i.e. minimal mechanical disturbance of the soil, permanent soil cover, and diversification of crops cultivated on any given plot. Farmers can follow a variety of practices to fulfill these principles. For example, minimum soil disturbance may involve zero or minimum tillage, ripping, basin planting, or other practices. Farmers may retain crop residues covering at least 30 percent of the soil or plant cover crops between seasons. Other practices worth looking at include Soil and Water Conservation (SWC) practices and Agricultural Water Management practices (AWM). SWC includes terracing, the construction of small walls along the contours of the land, contour plowing, and water catchments. Motorized and treadle pumps are instead part of AWM. Table 9 shows no impact of correct identification of improved varieties on the adoption of any of the practices under analysis. The lack of statistically significant results may be due to the low adoption rates for most of the practices, except soil erosion prevention and SWC, as also reported in Kosmowski et al. (2020). The results do not change either when using an aggregate index of all practices. However, the estimates of some of the practices

[16] Similar considerations apply to the results on agricultural productivity and profitability reported in Section 6.5

although not statistically significant are large in size. Indeed, correct classification seems to be positively correlated with rotation with a crop legume and zero tillage and negatively associated with most crop residue cover, minimum tillage, and Soil and Water Conservation practices.

6.4 Crowding-In versus Reallocation of Resources from Other Plots

After having shown that farmers allocate more inputs on plots where improved varieties are correctly identified, a natural next question is whether the additional resources come from reallocation from other household plots cultivated with crops other than maize. The conceptual framework (Section 2) generates competing predictions: if farmers face binding budget constraints, increased input use on correctly identified maize plots should reflect zero-sum reallocation across plots. In contrast, if information is the primary constraint, correct identification may trigger crowding-in through belief updating about agricultural returns more broadly or mobilization of previously uninvested resources.

To investigate this question, I estimate the 2SLS at the household level—instead of the plot level as before—where I aggregate the outcome measures across all farmers' plots. The correct identification variable now takes value one if farmers correctly identify the improved variety grown on at least one of their maize-cultivated plots. Thus, I compare aggregate input allocation on all maize plots versus input allocation on all plots under crops other than maize.

The random sampling framework for the collection of DNA fingerprinting does not ensure that the objective measure was collected for all maize plots of the household. This implies that it is not possible to establish the correct classification for each of the household's plots. Thus, in order to conduct this part of the analysis at the household level, I assume that if the variety was correctly classified on at least one of the household's plots, then it is likely that this is true also for the others. Indeed, it is arguably plausible that farmers used the same seed variety on all of their plots.[17]

Results are shown in Table 10. Panel A shows the first-stage results at the household level. The instrumental variable has a slightly stronger negative effect on correct classification than the plot level analysis and it is still highly statistically significant. The F-statistic is now above 21, lending confidence to the relevance of the instrumental variable.

Panel B reports the results of the 2SLS estimation on maize plots. It shows that the effects discussed in the previous section are also found when the estimation is done at the

[17] I use the 26 households where more than one plot was randomly selected to collect the DNA fingerprinting measure. Among these households, in 73 percent of cases, their misclassification status remained consistent across all plots, which suggests the assumption is reasonable. Because of the 26 households for whom more than one plot was randomly selected, the number of observations is a bit smaller for the plot-level analysis than for the household-level analysis.

household level. That is, households in which farmers correctly identify the maize variety grown on at least one of the plots use on average more NPS and hired labor per hectare on maize plots, in comparison to those who do not correctly classify. The effect on urea application is also positive though no longer statistically significant at the household level, likely reflecting reduced precision from aggregation.

Panel C shows what happens on cultivations of other crops when the correct identification condition on maize plots is met. As shown in Column (2), farmers who correctly classify use 78.8 kg/ha more urea on non-maize plots, an effect that is statistically significant at the 10 percent level. The coefficient on NPS application is also positive and statistically significant at the 5 percent level, with an increase of 88.0 kg/ha. In contrast, there is no statistically significant increase in labor use on non-maize plots, neither at planting nor at harvest.

The overall picture provided by these results is inconsistent with systematic reallocation of fixed resources from other crops to maize cultivation. Instead, the findings point towards crowding-in of additional resources at the household level. This pattern is consistent with information constraints—rather than binding credit constraints—being the primary barrier to input use, as formalized in the conceptual framework (Section 2.5).

Several mechanisms may explain why correct identification on maize plots is associated with increased input use on other crops. First, farmers who recognize improved varieties on maize plots and observe high returns to complementary inputs may update their beliefs about agricultural productivity more broadly, encouraging investment across all crops. Second, correct identification may increase farmers' willingness to mobilize previously uninvested resources—household savings, informal credit, or family labor—that they were reluctant to deploy under uncertainty. Third, farmers who purchase certified seed through formal DSM channels may simultaneously gain improved access to complementary inputs (fertilizer, extension services, credit) that benefit multiple crops, not just maize.

An alternative explanation for the effects observed in Panel C is directly linked to exposure to DSM. Although the instrumental variable relies on the length of exposure to the DSM program specifically for maize, the program was implemented with the exact same logic also for wheat and teff in the regions of Tigray, SNNP, and Oromia. Therefore, I cannot exclude that part of the observed effects for plots cultivated with crops other than maize may be driven by the direct effect of exposure to DSM whenever the program was implemented in the same woredas for multiple crops. While I cannot fully disentangle these mechanisms with available data, both interpretations support the conclusion that information and access constraints, not merely fixed budgets, limit input use.

The crowding-in pattern has implications for policy design. Interventions that improve seed market transparency do not merely reallocate fixed resources more efficiently—they can increase total agricultural investment by reducing uncertainty and improving market access. This suggests potential complementarities between seed sector reforms and broader

agricultural development efforts.

6.5 Yields and profits

The results discussed so far indicate that correctly classifying the variety grown on one plot as improved causes several changes in input allocation. Keeping in mind that both groups considered in this analysis cultivate improved varieties (true positives and false negatives), it is interesting to test whether changes in input allocation due to correct classification of varieties translate into higher productivity and profitability levels. Table 11 shows a noticeable increase in yields on plots where there is correct classification. The effect is statistically significant and represents a 122 percent increase from the mean of farmers who do not correctly classify the improved variety grown. This result highlights that utilizing more inputs as a result of correct beliefs leads to substantial productivity gains. The mere adoption of an improved variety does not automatically translate into higher yields unless it is accompanied by higher fertilizer use, thus the importance of correct beliefs on the variety adopted and the consequential behavioral adjustment to it. Indeed, the percentage increase in yield is consistent with the magnitude of change in fertilizer use. Although, I cannot detect any statistically significant effect of correct classification on household's absolute profits from maize cultivation in column (2), conditioning on the size of land cultivated the effect on profits is positive and big in magnitude (column (3)). Farmers who correctly classify the variety grown earn 253 USD more per hectare. The effect accounts for a 170 percent increase and is statistically significant (p-value$<$0.01).

7 Discussion of the main results

This section interprets the main findings in light of the conceptual framework, quantifies the associated economic magnitudes, and discusses their policy implications.

To assess the economic significance of misperception, I calibrate the profit decomposition implied by equation (8). Correct identification of improved seed varieties induces additional input expenditures of approximately 164 kg/ha of urea (USD 82/ha), 130 kg/ha of NPS fertilizer (USD 78/ha), and 4.7 days/ha of hired labor (USD 24/ha), for a total incremental cost of USD 184/ha. The estimated 120 percent yield increase relative to the false-negative baseline yield of 1,614 kg/ha corresponds to an additional 1,937 kg/ha of output. Valued at a maize price of USD 0.25/kg, this yields an additional USD 484/ha in revenue. Net profits therefore increase by approximately USD 300/ha, closely matching the direct empirical estimate of USD 253/ha reported in Table 8, Column 3.

The implied return on the incremental investment of USD 184/ha is approximately 163 percent, substantially exceeding prevailing annual interest rates in rural Ethiopia, which

range from 15 to 100 percent. This magnitude indicates that credit constraints alone are unlikely to explain the observed under-investment. Instead, false negatives appear to operate on a steep portion of the production function, where marginal returns to inputs are extremely high. This interpretation is consistent with the very low baseline input use observed in the data—for example, average urea application of only 19 kg/ha (Table 3), far below agronomic recommendations of 100–150 kg/ha for improved maize varieties.

Despite the large per-hectare profit gains, Table 11, Column 2 shows no statistically significant effect on total household profits on maize plots. This apparent discrepancy reflects scale constraints inherent in smallholder agriculture. With a mean plot size of 0.13 hectares, a gain of USD 253/ha translates into an average increase of only USD 33 per plot, which is modest in absolute terms and difficult to detect statistically. Farmers do not respond by expanding cultivated area, as most households already farm nearly all accessible land (mean total maize area is 0.14 hectares), face inelastic household labor supply—as evidenced by null effects on family labor use in Table 8—and encounter credit constraints that limit lumpy investments in land, even when information facilitates more intensive input use. Consequently, the estimated USD 253/ha gain reflects an efficiency improvement along the intensive margin, while the muted effects on aggregate household profits arise from binding extensive margin constraints. From a policy perspective, translating large per-hectare productivity gains into substantial welfare improvements requires complementary interventions that relax land and labor constraints.

Evidence on input use across crops further suggests that information rather than fixed budgets is the primary constraint. The crowding-in effects documented in Table 10, Panel C show that farmers who correctly identify improved maize varieties also significantly increase fertilizer use on other crops, applying an additional 79 kg/ha of urea and 88 kg/ha of NPS on non-maize plots. This pattern is inconsistent with simple reallocation within a fixed input budget. Instead, correct identification may update beliefs about agricultural returns more broadly, mobilize previously uncommitted resources, or improve access to inputs through formal seed and input markets. Alternatively, DSM exposure may directly affect multiple crops, given its simultaneous implementation for maize, wheat, and teff. Regardless of the precise channel, the evidence points to information and access constraints, rather than purely financial constraints, as key barriers to efficient input use.

Several additional patterns reinforce the conceptual framework. Labor responses are asymmetric over the production cycle: there is no effect on labor use at planting, but a large and statistically significant increase in hired labor at harvest (+4.7 days/ha), consistent with forward-looking optimization by farmers anticipating substantial yield gains. Correct identification also leads to lower seeding rates (Table 8, Column 11), consistent with reduced quality uncertainty. Seeds that are correctly identified are overwhelmingly sourced from formal channels (93 percent for true positives versus 21 percent for false negatives; Table 6),

allowing farmers to reduce insurance planting. The absence of effects on complementary agricultural practices (Table 7) suggests weak short-run complementarities with improved varieties or learning dynamics that extend beyond the duration of DSM exposure. Finally, the negative and statistically significant interaction between DSM exposure and proximity (-0.268, $p < 0.01$) in the first stage confirms that transparency improvements disproportionately benefit more remote farmers facing higher baseline information barriers, as predicted by the learning mechanism outlined in Section 2.6.

Several limitations are worth noting. The instrumental variables estimates identify a Local Average Treatment Effect for compliers induced by DSM exposure. The machine learning extension presented in Section 9 suggests that this LATE is broadly representative of the average treatment effect, although measurement error in predicted DNA-based classification limits precision. The framework abstracts from risk aversion; however, the large estimated returns (163 percent) indicate movement toward, rather than beyond, risk-adjusted optima. Prices are treated as exogenous, and general equilibrium effects are unlikely to be first-order concerns given low national adoption rates of improved maize varieties (23.3 percent) and the small scale of individual farms.

The results have clear policy implications. Investments in seed market transparency—through programs such as DSM that emphasize certification, labeling, and quality assurance—can generate substantial productivity gains on the order of USD 253/ha, corresponding to a 170 percent increase in profits. These interventions are likely to be most effective in remote rural areas, where information barriers are most severe. Moreover, correcting information frictions can amplify the returns to complementary agricultural investments, such as extension services and input subsidies, as evidenced by the observed crowding-in effects. Nonetheless, achieving large absolute welfare gains will require parallel efforts to address land and labor constraints that limit responses along the extensive margin.

Overall, the empirical evidence strongly supports the conceptual framework. Misperceptions about seed variety identity generate large inefficiencies, with false negatives foregoing returns of approximately 163 percent on incremental fertilizer investments. The close correspondence between theoretical predictions and empirical findings—across input use, labor allocation, yields, profitability, and spillover effects—suggests that the framework captures the key economic mechanisms at work. Although extensive margin constraints limit absolute household profit gains at current farm sizes, the intensive margin efficiency improvements are substantial and underscore the potential of information-based interventions to raise productivity in smallholder agriculture.

8 Exclusion restriction

The exclusion restriction implied by the instrumental variable approach is that, conditional on the controls included in the regression, the interaction between the duration of exposure to the DSM program and households' proximity to woreda towns affects production decisions only through its effect on farmers' correct identification of the improved variety grown on their plot.

A natural threat to this identifying assumption is that DSM was implemented earlier in some woredas rather than in others for unobserved reasons that may also affect production decisions differentially by proximity to woreda towns. For example, DSM could have expanded earlier in locations where input demand was increasing over time among households closer to woreda towns but not among those farther away. In that case, exposure to DSM would be directly associated with the outcome variables independently of the channel of correct classification of the seed variety. Given the non-random targeting of DSM discussed in Section 3, random assignment of the program to woredas can be ruled out.

Importantly, I do not observe consistent woreda-level panel measures of infrastructure investments, extension deployment, or other contemporaneous agricultural initiatives that could have evolved alongside DSM roll-out. As a result, I cannot directly control for all potential time-varying confounders. I therefore provide empirical evidence that speaks to the plausibility of the exclusion restriction by testing whether the instrument is associated with differential patterns in input allocation *prior* to ESS4. To do so, I employ two complementary strategies.

Placebo reduced-form tests in pre-ESS4 waves.
First, I perform a placebo test by estimating the reduced-form relationship using the waves collected prior to ESS4: ESS1 (2011/2012), ESS2 (2013/2014), and ESS3 (2015/2016). The reduced-form regression at the household level is:

$$\begin{aligned} Y_{hw} = \beta_0 + \beta_1 P_{hw} * DSM_w^{2019} + \beta_2 P_{hw} + \beta_3 DSM_w^{2019} + \beta_4 VSS_w + \\ + \beta_5 VSS_w * P_{hw} + \beta_6 Suitability_w + \beta_7 Suitability_w * P_{hw} + \alpha_r + \varepsilon_{phw} \end{aligned} \quad (14)$$

where the variables are defined as in equations (1) and (2), and DSM_w^{2019} denotes the length of exposure to DSM through 2019 for woreda w.

All the waves share almost identical questionnaires, with the exception that the quantity applied of NPS and seeds was not collected in ESS1 and ESS2. Moreover, although harvested quantities were meant to be collected in all waves, in practice most crops/fields do not have self-reported harvest information in ESS1.[18] In addition, the crop-cut measure, which I rely on for the main results on productivity and profitability, is only available in

[18] Source: ESS1 Basic Information Document.

ESS4 and for a very limited subsample in previous waves. The measurement error affecting self-reported harvested quantity (relative to crop cuts) makes these measures less comparable across waves. For this reason, I do not use yields or profits in the placebo test or in the differential-trend exercise below, and instead focus on input allocation outcomes that are consistently measured.

For comparability across waves, I restrict the samples to maize-cultivating households in the regions where the DNA fingerprint sample was collected for ESS4: Amhara, Dire Dawa, Harar, Oromia, SNNP, and Tigray. If the instrument were primarily capturing baseline differences between close-by and more remote households in woredas that later accumulated greater DSM exposure, it would tend to predict input use already in ESS1–ESS3.

Table 12 reports the placebo estimates.[19] Reassuringly, the instrumental variable (years of exposure to DSM through 2019 interacted with proximity to woreda towns) does not predict fertilizer use or hired labor in ESS1, ESS2, and ESS3, while it does in ESS4. While some coefficients are statistically significant in earlier waves, they do not concern the variables of primary interest (urea, NPS, and hired labor at harvest) or they have opposite signs. Estimates in ESS3 are the closest to ESS4, which is consistent with the smaller difference in cumulative DSM exposure between ESS3 and ESS4 and with the possibility that learning effects become more salient closer to the ESS4 period.

Results are consistent when using the balanced household panel sample reported in the Appendix (Table A5)[20], suggesting that this evidence is not driven by sample composition across waves.

Differential pre-trend tests.

Second, I test whether changes in input use across the pre-ESS4 panel waves are systematically related to the instrument. Using the panel component of the three waves prior to ESS4, I estimate the following household-level specification:

$$\Delta Y_{t,t-1} = \beta_0 + \beta_1 P_{hw} * DSM_{ESS4} + \varepsilon \tag{15}$$

where $\Delta Y_{t,t-1}$ represents the change in input use between ESS2 and ESS1, ESS3 and ESS2, and ESS3 and ESS1, and the independent variable is DSM exposure in ESS4 interacted with proximity to woreda town.

The results reported in Table 13 show that pre-trends in input allocation are not correlated

[19] To ensure comparability across waves, I use the sample of all households growing maize in the regions where DNA fingerprinting was collected. This is a larger sample for ESS4 than the one used in the main analysis since for the purpose of this exercise it is not necessary to restrict to observations where DNA fingerprinting was collected. I report the estimates of the reduced form regression for the DNA fingerprinting sample of ESS4 in Table A4 (Appendix). The weaker results in Panel A of Table 12 relative to Table A4 are due to sample composition: the larger ESS4 sample includes EAs in areas that are less important for maize production and thus were not sampled for DNA fingerprinting.

[20] The panel dimension mentioned refers to ESS1, ESS2, and ESS3, but not ESS4.

with the instrumental variable, with the only exception of family labor at harvest. Overall, these tests provide evidence against the presence of differential *pre-existing* trends correlated with the instrument. At the same time, they cannot definitively rule out all time-varying shocks that may have coincided with DSM expansion during the ESS4 period, which remains an important limitation when interpreting the causal estimates.

9 Misperception versus misreporting

One potential issue to consider is the possibility that the discrepancy between farmers' beliefs and the true genetics of the variety is due to misreporting rather than misperception. When misreporting occurs, farmers' underlying beliefs align with reality, as those who misreport their true beliefs are presumed to be aware of and make decisions based on these hidden beliefs rather than the measurements. Therefore, if misreporting is prevalent, farmers' self-reported varietal identification will not substantially influence their decision-making processes. In such cases, accurate inference would necessitate the use of genetic identification, which reflects farmers' latent perceptions.

A straightforward method to distinguish whether farmers' errors stem from misperception or misreporting is to examine input allocation. If input allocation is correlated with the actual genetic characteristics of the variety, it suggests misreporting rather than misperception. Conversely, if input allocation is associated with farmers' beliefs, then we are dealing with a situation of misperception. To provide evidence in support of whether farmers' beliefs are equal to the self-reported variety, I follow the approach proposed by Wossen et al., 2022. In particular, I estimate the following equation:

$$X_{ph} = \beta_0 + \beta_1 SR_{ph} + \beta_2 DNA_p + \varepsilon_{ph} \tag{16}$$

where X_{ph} is the input applied on plot p by household h, SR_{ph} is the self-reported variety identification for plot p by household h and equals to one if variety is reported as improved and zero otherwise, DNA_{ph} is the genetic identification on plot p and equals to one if the variety is improved and zero otherwise and ε_{ph} is the error term clustered at Enumeration Area (EA) level.

Table 14 reports the results of the estimation of equation (16) and the p-values of the Naive Wald statistics and the p-values of the Lochner-Moretti (LM) statistics. Both statistics allow to test for the presence of misclassification in *SR*, which is equivalent to testing for the exclusion restriction of *DNA*, but LM is also robust to mispecifications.

The results concerning the quantity of inputs, in particular fertilizers application, suggest that farmers' decisions are primarily associated with their self-reported measures rather than the actual genetic variety. The only exception to this pattern is the usage of DAP.

Nevertheless, when combining DAP and NPS (Column (2)), the outcomes still align with the idea of misperception rather than misreporting. This suggests that the choice between DAP and NPS may be influenced by other factors, such as access to one of these fertilizers, as the two are equivalent in terms of purpose. In contrast, I do not observe any labor-related choices correlated with either of these measures. Lastly, only in the case of seedling rate, I cannot reject the presence of misreporting, since the error may arise from both sources[21].

Therefore, these findings are consistent with those reported in Wossen et al., 2022 for Cassava in Nigeria, reinforcing our emphasis on misperception over misreporting in varietal identification and supporting the idea that wrong beliefs about the variety grown have consequences on agricultural decision-making.

10 Extension to the nationally representative sample

The analysis conducted so far suffers from the caveat of limited sample size, due to the fact that the DNA fingerprinting measure is available only for a subsample of farmers given the relatively high costs of collecting this type of data. For this reason, I use machine learning (ML) techniques to predict the objective measure of improved maize variety in the part of the sample where this measure was not collected. This allows to extend the analysis to the sample where the DNA fingerprinting measure was not collected, thus increasing the sample size. The contribution of this exercise is twofold: first, it is of interest to investigate whether the results discussed so far are robust when using the full nationally representative sample of ESS4. Second, given the difficulties in scaling the DNA fingerprinting data collection and analysis due to logistics and budgetary constraints, it is worth exploring other techniques that rely on farmers' self-reported information and could be used to interpolate more accurate national figures of adoption of improved varieties. Specifically, using the sample of 491 plot observations[22] where both the DNA fingerprinting measure and the self-reported measure of improved maize variety are available (the training set, henceforth), I predict the former measure on the remaining sample of 1490 plots under maize cultivation (i.e. the test set).
As recommended by the literature, the first step consists of selecting the best-performing model. In order to do so, I test different machine learning algorithms on the training set and cross-validate the models to assess their performance in terms of Minimum Squared Error (MSE), also known as risk. Finally, I choose the best-performing model and predict

[21]There is also a marginally significant coefficient on land area dedicated to maize cultivation, but it has the opposite sign as the coefficient for misreporting. Moreover, the result is driven by a few very large values for land area and upon excluding these outliers from the sample only the hypothesis of misperception remains supported.

[22]To train the algorithm I take advantage of the full sample available which also includes false positive and true negative. Moreover, I recover 12 additional observations which were excluded due to missing sampling weights.

the variable of interest on the test set (Athey & Imbens, 2019). Before performing this procedure, I exclude from the original training set, 100 observations that I use to estimate the accuracy of our preferred model. That is, I predict the variable of interest on "unseen" data (i.e. data not used to train the algorithm) and compare the results with the true value of the DNA fingerprinting. In this way, I can assess the percentage of correctly predicted data points using the models of choice. Table A6 provides a summary of the different samples used in the procedure just described.

I take advantage of the SuperLearner package available in R, which provides an easy and efficient way to perform the training and the prediction processes using a wide library of machine learning algorithms. The training procedure is fed with 120 covariates, which can be broadly categorized as follows:

- Self-reported improved seed variety measure.
- Seed characteristics: source, if re-cycled from previous seasons, if purchased and location of purchase.
- Plot characteristics: land tenure, self-reported plot type and fertility measures, distance to the household's dwelling;
- Household characteristics: household size, dependency ratio, credit and insurance access, household's head education;
- Climate, rainfall, soil fertility measures from satellite data matched with GPS plot location;
- Farmer's agricultural practices adopted on the plot.

I purposely do not include any of the outcome variables used so far in the analysis, such as input application, yields, profits, or any other covariate used in the first stage. This exclusion is essential as I replicate the same specifications using the predicted sample. Moreover, it is important to note that I also incorporate measures that are not self-reported and, thus, not known by the farmer. This step is taken to ensure that the algorithm does not merely mimic the data generation process of the farmer, and thus her mistakes.

For each of the machine learning algorithms, I use a 5-fold cross-validation, that is 5 randomly selected sub-samples of the training set, 4 used to train the ML algorithms independently and 1 to calculate the MSE. Given the binary nature of the prediction variable, I rely on the Xg.boost, Ipredbagg and Random Forest learners [23]. Table 15 reports the average MSE based on 5-fold cross-validation. The best-performing individual model

[23] Xg.boost, Random Forest and Ipredbagg are decision-tree-based ensemble machine learning algorithms that perform well in classification problems.

is Random Forest, with a risk coefficient (MSE) of 0.1336. However, as shown also in Figure A2, SuperLearner selects as the best performing model the ensemble (i.e. the linear combination) of all the three models. The ensemble method predicts the variable of interest as a weighted linear combination of the predictions of the single models, where weights are inversely proportional to the MSE of the single models. Overall, testing the model on unseen data allows to reach a level of accuracy of 85 percent. Taking into account the classification problem, the number of variables used and the potential noise of some of these variables, obtaining 85 percent of accuracy is quite high. This is particularly reassuring as it will affect the statistical power of the next steps. Moreover, given the prediction power obtained by relying on self-reported and GIS data, the results provide evidence that performing DNA fingerprinting on subsamples can be a reliable way to scale the objective measure at the national level.

In what follows I discuss the results of the first-stage and 2SLS estimations of correct identification, constructed using the predicted DNA fingerprinting measure on input usage. To assess the validity of the predictions and to disentangle the potential source of error, I replicate the analysis using the *predictions* of i) the original sample used so far [24]; ii) the sample of EAs of the original sample; iii) the sample of regions of the original sample; iv) the full national ESS4 sample. As in the previous analysis, I replicate the analysis on the sample true positive (correct identification) and false negative cases.

Table 16 shows the results of the first-stage regression. Estimates in column (1), almost perfectly mimic the first stage results of the original sample explored in Table 7. All the coefficients in the regression are of the same sign and magnitude as the core values, if not completely identical. In particular, the coefficient of the instrumental variable is still negative, highly statistically significant and slightly bigger in size. The F-statistic for the IV is 11.05. While the results just discussed are reassuring, they do not provide an exhaustive picture of the model's performance. This is because I am using the predictions obtained on the same sample that was used to train the algorithms. Thus, columns (2)-(4) depart from this sample and explore the validity of the predictions over "unseen" data. The estimates are overall robust in all the samples used, however, there is variation in the relevance of the IV in the different contexts. In column (2), I report the first-stage using the sample of households from the same EAs, but including also plots that were not randomly selected for the DNA fingerprinting measure: the coefficient of the IV is slightly smaller and statistically significant, which is also reflected in a smaller F-statistics. Estimates on the sample of households from the same regions of the DNA fingerprinting subsample and the full national sample are reported in columns (3) and (4), respectively. The value of the F-statistic for IV is 4.13, which suggests that the instrument is now less relevant than it was in the original

[24] In this Section, I refer to "original sample" as the sample of plot-level observations used in the main analysis containing both the self-reported and DNA fingerprinting data.

sample. There are at least two potential motivations for why this is the case. The woreda and regions included in columns (3) and (4) are notably either areas where maize is much less relevant or barely exposed to the DSM program. Indeed, DSM was not implemented in the regions of Beninshangul Gumuz, Gambela and Afar, which are the regions included in column (4) but that were not surveyed in the original sample.

Following the same logic, Table 17 shows the 2SLS results. As shown in Panel A, the results obtained from the predictions are almost identical to the original sample discussed in the previous sections for input application (Table 8). In particular, the effect of correct identification on quantity of urea and NPS are qualitatively the same in size in Panel A, than previous results and exhibit the same level of significance. Moving to Panel B, the effect of correctly identified varieties on the quantity of urea is larger in magnitude and highly statistically significant. The effect of correct classification on hired labor at harvest is still statically significant and close to the original estimate. In Panel C and D, most of the results remain consistent with the original results, with a few notable exceptions: indeed, the seedling rate and yields are no longer statistically significant. However, since I am by design adding woredas (and regions) where maize is much less important and where only a few woredas were exposed to the DSM, both the predictions and the first-stage become less relevant, thus affecting the results in Panel C and D.

Overall, the application conducted using machine learning techniques supports the robustness of the results and the discussed mechanisms behind them. Moreover, it provides empirical evidence of the possibility of interpolation of DNA fingerprint measures from self-reported data with only a 15 percent of error in predictions.

11 Concluding remarks

The analysis conducted in this paper provides causal evidence of the impact of correct classification of improved maize varieties on input allocation, agricultural practices, yields and profits. I find that farmers who correctly classify the variety grown as improved use more fertilizers and hired labor at harvest, while I do not find any detectable changes in agricultural practices in comparison to farmers who have wrong beliefs about the improved variety planted. Furthermore, I observe substantial productivity gains associated with accurate variety identification and profitability conditional on the size of the cultivated land. Given that the effect of correct classification is identified through the cumulative length of exposure to a program that favors transparency in the seed sector, I argue that the underlying mechanism is based on farmers dynamically learning about the optimal management of the agricultural technology used in the past seasons.

There are a few lessons that can be drawn from this analysis. First, misperception in which agricultural technology is used has testable implications for farmers' produc-

tion decisions. Wrong beliefs on the improved variety grown can be a factor influencing (mis)allocation of agricultural resources. Second, since the exogenous variation exploited in this analysis relies on the cumulative effect of reducing misperception, it provides suggestive evidence of the dynamic nature of farmers' learning. Thus, optimal management of agricultural technologies and complementary inputs should be conceived as the outcome of continuous experimentation where misperceptions may affect the learning process. Third, this study contributes to the research agenda on agricultural data measurement, by recognizing not only the contribution of objectively measured data but also, and foremost the importance of combining it with farmers' self-reported measures to understand possible misallocation. Fourth, although programs like DSM do not have the primary objective of reducing misperception, they do have this additional side effect. When evaluating the impact of these policies, it is advisable to consider these effects.

The quasi-experimental nature of the empirical strategy implies some caveats. First, the temporarily cross-section nature of the ESS4 does not allow for comparison within farms across years, which inhibits an in-depth investigation of the dynamic adoption process. Second, the estimated effect is a Local Average Treatment Effect that identifies the impact of correct classification only for the population of compliers. That is, the effect of correct classification for farmers who are exposed to the DSM program conditional on their distance to urban centers and end up identifying the variety grown as improved when this is the case.

While the literature has so far neglected the role of farmers' misperception in which agricultural technology is used, this paper provides empirical evidence of the large magnitude of the effects of correcting misperception. The interest in improved seed varieties and optimal input allocation is particularly relevant in light of climate change. The overall impacts of climate change on agriculture are expected to be negative due to climate variability and the frequency of extreme climatic events, such as droughts and flooding, that affect precipitation and, higher temperatures that impact yields in a negative way and favor the growth of weeds and the proliferation of crop pests (Nelson et al., 2009). Because agricultural production remains the main source of income for most rural communities in low and middle-income countries, adaptation of the agricultural sector to the adverse effects of climate change is imperative for protecting and improving the livelihoods of the poor and ensuring food security (Lemessa, Watebaji, & Yismaw, 2019). For this reason, national, regional and international plant breeding efforts involved multilocational trials with the goal to develop crop varieties that are resistant to climate-related phenomena and more efficient in their use of resources to reduce the impact on the agricultural ecosystem and the wider environment. Resistance to drought, salinity, flooding, diseases and pests are the most common climate-related traits for which crop varieties are bred (Lemessa et al., 2019). Farmers' accurate knowledge of the variety used is pivotal to taking advantage of the tolerance traits. For this reason, farmers' correct classification of improved varieties acquires new meaning, especially

in the case of varieties with tolerance traits, but more and foremost for the complementarity between those and climate-smart practices in the pursuit of climate adaptation.

References

Abay, K. A., Abate, G. T., Barrett, C. B., & Bernard, T. (2019, 6). Correlated non-classical measurement errors, 'second best' policy inference, and the inverse size-productivity relationship in agriculture. *Journal of Development Economics*, *139*, 171-184. doi: 10.1016/J.JDEVECO.2019.03.008

Abay, K. A., Bevis, L. E., & Barrett, C. B. (2021, 3). Measurement error mechanisms matter: Agricultural intensification with farmer misperceptions and misreporting. *American Journal of Agricultural Economics*, *103*, 498-522. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/ajae.12173https://onlinelibrary.wiley.com/doi/abs/10.1111/ajae.12173https://onlinelibrary.wiley.com/doi/10.1111/ajae.12173 doi: 10.1111/AJAE.12173

Alemu, D., Rashid, S., Tripp, R., Alemu, D., Rashid, S., & Tripp, R. (2019, 3). Seed system potential in ethiopia: constraints and opportunities for enhancing the seed sector. *Gates Open Research 2019 3:948*, *3*, 948. Retrieved from https://gatesopenresearch.org/documents/3-948 doi: 10.21955/GATESOPENRES.1115840.1

Angner, E., & Loewenstein, G. (2007). Behavioral economics. *Handbook of the philosophy of science: Philosophy of economic*, 641–690.

Ashour, M., Gilligan, D. O., Hoel, J. B., & Karachiwalla, N. I. (2019). Do beliefs about herbicide quality correspond with actual quality in local markets? evidence from uganda. *The Journal of Development Studies*, *55*(6), 1285–1306.

Ashraf, N., Giné, X., & Karlan, D. (2009, 11). Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in kenya. *American Journal of Agricultural Economics*, *91*, 973-990. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-8276.2009.01319.xhttps://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8276.2009.01319.xhttps://onlinelibrary.wiley.com/doi/10.1111/j.1467-8276.2009.01319.x doi: 10.1111/J.1467-8276.2009.01319.X

ATA. (2015). Seed system development strategy: Vision, systemic challenges, and prioritized interventions.

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685–725.

Balemi, T., Kebede, M., Abera, T., Hailu, G., Gurmu, G., & Getaneh, F. (2019, 10). Some maize agronomic practices in ethiopia: A review of research experiences and lessons from agronomic panel survey in oromia and amhara regions. *African Journal of Agricultural Research*, *14*, 1749-1763. Retrieved from https://academicjournals.org/journal/AJAR/article

-abstract/A3827DC62088 doi: 10.5897/AJAR2019.14338

Beaman, L., Karlan, D., Thuysbaert, B., & Udry, C. (2013, 5). Profitability of fertilizer: Experimental evidence from female rice farmers in mali. *American Economic Review*, *103*, 381-86. doi: 10.1257/AER.103.3.381

Benson, T., Spielman, D., & Kasa, L. (2014). Direct seed marketing program in ethiopia in 2013: An operational evaluation to guide seed-sector reform. *IFPRI Discussion Paper 1350. Washington, D.C.: International Food Policy Research Institute (IFPRI).*

Besley, T., & Case, A. (1993). Modeling technology adoption in developing countries. *The American economic review*, *83*(2), 396–402.

Besley, T., Case, A., et al. (1994). *Diffusion as a learning process: Evidence from hyv cotton* (Tech. Rep.).

Bohr, N., Deisemann, T., Gollin, D., Kosmowski, F., & Lybbert, T. J. (2022). The seeds of misallocation: Fertilizer use and maize varietal misidentification in ethiopia. *Available at SSRN: https://ssrn.com/abstract=4587151*. doi: http://dx.doi.org/10.2139/ssrn.4587151

Bold, T., Kaizzi, K. C., Svensson, J., & Yanagizawa-Drott, D. (2017, 8). Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in uganda. *The Quarterly Journal of Economics*, *132*, 1055-1100. Retrieved from https://academic.oup.com/qje/article/132/3/1055/3064350 doi: 10.1093/QJE/QJX009

Burke, M., Bergquist, L. F., & Miguel, E. (2019, 5). Sell low and buy high: Arbitrage and local price effects in kenyan markets. *The Quarterly Journal of Economics*, *134*, 785-842. Retrieved from https://academic.oup.com/qje/article/134/2/785/5266398 doi: 10.1093/QJE/QJY034

Carletto, C., Savastano, S., & Zezza, A. (2013, 7). Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. *Journal of Development Economics*, *103*, 254-261. doi: 10.1016/J.JDEVECO.2013.03.004

Cole, S., Giné, X., & Vickery, J. (2017, 6). How does risk management influence production decisions? evidence from a field experiment. *The Review of Financial Studies*, *30*, 1935-1970. Retrieved from https://academic.oup.com/rfs/article/30/6/1935/2433387 doi: 10.1093/RFS/HHW080

Cole, S. A., & Fernando, A. N. (2021, 2). 'mobile'izing agricultural advice technology adoption diffusion and sustainability. *The Economic Journal*, *131*, 192-219. Retrieved from https://academic.oup.com/ej/article/131/633/192/5867759 doi: 10.1093/EJ/UEAA084

CSA. (2019). *Sample survey area and production of major crops.*

de Janvry, A., & Sadoulet, E. (2020, 9). Using agriculture for development: Supply- and demand-side approaches. *World Development*, *133*, 105003. doi: 10.1016/

J.WORLDDEV.2020.105003

de Janvry, A., Sadoulet, E., & Suri, T. (2017). Field experiments in developing country agriculture. In *Handbook of economic field experiments* (Vol. 2, pp. 427–466). Elsevier.

Desiere, S., & Jolliffe, D. (2018). Land productivity and plot size: Is measurement error driving the inverse relationship? *Journal of Development Economics*, *130*, 84–98.

Duflo, E., Kremer, M., & Robinson, J. (2011, 10). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review*, *101*, 2350-90. doi: 10.1257/AER.101.6.2350

Emerick, K., de Janvry, A., Sadoulet, E., & Dar, M. H. (2016). Technological innovations, downside risk, and the modernization of agriculture. *American Economic Review*, *106*(6), 1537–61.

Evenson, R. E., & Gollin, D. (2003, 5). Assessing the impact of the green revolution, 1960 to 2000. *Science*, *300*, 758-762. Retrieved from https://www.science.org/doi/abs/10.1126/science.1078710 doi: 10.1126/SCIENCE.1078710/ASSET/4B81E052-332D-4728-BF5B-FA7743B5E8BC/ASSETS/GRAPHIC/SE1731473002.JPEG

Fink, G., Jack, B. K., & Masiye, F. (2020, 11). Seasonal liquidity, rural labor markets, and agricultural production. *American Economic Review*, *110*, 3351-92. doi: 10.1257/AER.20180607

Floro IV, V. O., Labarta, R. A., Becerra López-Lavalle, L. A., Martinez, J. M., & Ovalle, T. M. (2018). Household determinants of the adoption of improved cassava varieties using dna fingerprinting to identify varieties in farmer fields: a case study in colombia. *Journal of Agricultural Economics*, *69*(2), 518–536.

Foster, A. D., & Rosenzweig, M. R. (1995, 9). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, *103*, 1176-1209. Retrieved from https://www.journals.uchicago.edu/doi/abs/10.1086/601447 doi: 10.1086/601447

Foster, A. D., & Rosenzweig, M. R. (2010, 8). Microeconomics of technology adoption. *Annual Reviews Econonomics*, *2*, 395-424. Retrieved from https://www.annualreviews.org/doi/abs/10.1146/annurev.economics.102308.124433 doi: 10.1146/ANNUREV.ECONOMICS.102308.124433

Gollin, D., Hansen, C. W., & Wingender, A. M. (2021, 6). Two blades of grass: The impact of the green revolution. *Journal of Political Economy*, *129*, 2344-2384. Retrieved from https://www.journals.uchicago.edu/doi/abs/10.1086/714444 doi: 10.1086/714444

Gollin, D., Morris, M., & Byerlee, D. (2005). Technology adoption in intensive post-green

revolution systems. *American Journal of Agricultural Economics*, *87*(5), 1310-1316. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8276.2005.00824.x doi: https://doi.org/10.1111/j.1467-8276.2005.00824.x

Gourlay, S., Kilic, T., & Lobell, D. B. (2019, 11). A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale - productivity relationship in uganda. *Journal of Development Economics*, *141*, 102376. doi: 10.1016/J.JDEVECO.2019.102376

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014, 8). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, *129*, 1311-1353. Retrieved from https://academic.oup.com/qje/article/129/3/1311/1817927 doi: 10.1093/QJE/QJU015

Ilukor, J., Kilic, T., Stevenson, J., Gourlay, S., Kosmowski, F., Kilian, A., ... Asea, G. (2017). Blowing in the wind: the quest for accurate crop variety identification in field research, with an application to maize in uganda. In *Cgiar spia conference on the impact of agricultural research: Rigorous evidence for policy, nairobi, kenya.*

Ishimaru, S. (2022, 01). Empirical Decomposition of the IV–OLS Gap with Heterogeneous and Nonlinear Effects. *The Review of Economics and Statistics*, 1-45. Retrieved from https://doi.org/10.1162/rest_a_01169 doi: 10.1162/rest_a_01169

Kahneman, D., & Tversky, A. (1984, 4). Choices, values, and frames. *American Psychologist*, *39*, 341-350. doi: 10.1037/0003-066X.39.4.341

Karlan, D., Osei, R., Osei-Akoto, I., & Udry, C. (2014, 5). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, *129*, 597-652. Retrieved from https://academic.oup.com/qje/article/129/2/597/1867065 doi: 10.1093/QJE/QJU002

Kosmowski, F., Alemu, S., Mallia, P., Stevenson, J., & Macours, K. (2020). Shining a brighter light: Comprehensive evidence on adoption and diffusion of cgiar-related innovations in ethiopia. *SPIA Synthesis Report*.

Kosmowski, F., Ambel, A., Tsegay, A. H., Negawo, A. T., Carling, J., Kilian, A., & Agency, T. C. S. (2021). A large-scale dataset of barley, maize and sorghum variety identification using dna fingerprinting in ethiopia. *Data*, *6*(6). Retrieved from https://www.mdpi.com/2306-5729/6/6/58 doi: 10.3390/data6060058

Kosmowski, F., Aragaw, A., Kilian, A., Ambel, A., Ilukor, J., Yigezu, B., & Stevenson, J. (2019, 6). Varietal identification in household surveys: Results from three household-based methods against the benchmark of dna fingerprinting in southern ethiopia. *Experimental Agriculture*, *55*, 371-385. Retrieved from https://www.cambridge.org/core/journals/experimental-agriculture/article/varietal-identification-in

-household-surveys-results-from-three-householdbased-methods-against-the-benchmark-of-dna-fingerprinting-in-southern-ethiopia/C9F13CD010BEFDFF6AAE24CE8BBAF7CE doi: 10.1017/S0014479718000030

Laajaj, R., Macours, K., Masso, C., Thuita, M., & Vanlauwe, B. (2020, 8). Reconciling yield gains in agronomic trials with returns under african smallholder conditions. *Scientific Reports 2020 10:1*, *10*, 1-15. Retrieved from https://www.nature.com/articles/s41598-020-71155-y doi: 10.1038/s41598-020-71155-y

Lemessa, S. D., Watebaji, M. D., & Yismaw, M. A. (2019). Climate change adaptation strategies in response to food insecurity: The paradox of improved potato varieties adoption in eastern ethiopia. *Cogent Food & Agriculture*, *5*. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/23311932.2019.1640835 doi: 10.1080/23311932.2019.1640835

Macours, K. (2019, 10). Farmers' demand and the traits and diffusion of agricultural innovations in developing countries. *Annual Review of Resource Economics*, *11*, 483-499. Retrieved from https://www.annualreviews.org/doi/abs/10.1146/annurev-resource-100518-094045 doi: 10.1146/ANNUREV-RESOURCE-100518-094045

Magruder, J. R. (2018, 10). An assessment of experimental evidence on agricultural technology adoption in developing countries. *Annual Review of Resource Economics*, *10*, 299-316. Retrieved from https://www.annualreviews.org/doi/abs/10.1146/annurev-resource-100517-023202 doi: 10.1146/ANNUREV-RESOURCE-100517-023202

Maredia, M., Reyes, B., Manu-Aduening, J., Dankyi, A., Hamazakaza, P., Muimui, K., ... others (2016). Testing alternative methods of varietal identification using dna fingerprinting: Results of pilot studies in ghana and zambia. *Food Security International Development Working Papers 246950. Michigan State University, Department of Agricultural, Food, and Resource Economics*.

Mekonen, L. K., Minot, N., Warner, J., & Abate, G. T. (2019). *Performance of direct seed marketing pilot program in ethiopia: Lessons for scaling-up* (Vol. 132). Intl Food Policy Res Inst.

Michelson, H., Fairbairn, A., Ellison, B., Maertens, A., & Manyong, V. (2021, 1). Misperceived quality: Fertilizer in tanzania. *Journal of Development Economics*, *148*, 102579. doi: 10.1016/J.JDEVECO.2020.102579

Mobarak, A. M., & Rosenzweig, M. R. (2012). Selling formal insurance to the informally insured.

Nelson, G. C., Rosegrant, M. W., Koo, J., Robertson, R., Sulser, T., Zhu, T., ... others (2009). *Climate change: Impact on agriculture and costs of adaptation* (Vol. 21). Intl Food Policy Res Inst.

Pixley, K., & Bänziger, M. (2004). Open-pollinated maize varieties: A backward step or valuable option for farmers? *Seventh Eastern and Southern Africa Regional Maize Conference*. doi: http://dx.doi.org/10.2139/ssrn.4587151

Shamdasani, Y. (2021, 9). Rural road infrastructure & agricultural production: Evidence from india. *Journal of Development Economics*, *152*, 102686. doi: 10.1016/J.JDEVECO.2021.102686

Sheahan, M., & Barrett, C. B. (2017, 2). Ten striking facts about agricultural input use in sub-saharan africa. *Food Policy*, *67*, 12-25. doi: 10.1016/J.FOODPOL.2016.09.010

Suri, T. (2011, 1). Selection and comparative advantage in technology adoption. *Econometrica*, *79*, 159-209. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.3982/ECTA7749https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA7749https://onlinelibrary.wiley.com/doi/10.3982/ECTA7749 doi: 10.3982/ECTA7749

Suri, T., & Udry, C. (2022). Agricultural technology in africa. *Journal of Economic Perspectives*, *36*(1), 33–56.

Teklewold, H., Kassie, M., Shiferaw, B., & Köhlin, G. (2013, 9). Cropping system diversification, conservation tillage and modern seed adoption in ethiopia: Impacts on household income, agrochemical use and demand for labor. *Ecological Economics*, *93*, 85-93. doi: 10.1016/J.ECOLECON.2013.05.002

Tversky, A., & Kahneman, D. (1973, 9). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232. doi: 10.1016/0010-0285(73)90033-9

Wossen, T., Abay, K. A., & Abdoulaye, T. (2022). Misperceiving and misreporting input quality: Implications for input use and productivity. *Journal of Development Economics*, *157*, 102869. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304387822000414 doi: https://doi.org/10.1016/j.jdeveco.2022.102869

Wossen, T., Abdoulaye, T., Alene, A., Nguimkeu, P., Feleke, S., Rabbi, I. Y., ... Manyong, V. (2019, 1). Estimating the productivity impacts of technology adoption in the presence of misclassification. *American Journal of Agricultural Economics*, *101*, 1-16. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1093/ajae/aay017https://onlinelibrary.wiley.com/doi/abs/10.1093/ajae/aay017https://onlinelibrary.wiley.com/doi/10.1093/ajae/aay017 doi: 10.1093/AJAE/AAY017

Zabel, F., Putzenlechner, B., & Mauser, W. (2014). Global agricultural land resources–a high resolution suitability evaluation and its perspectives until 2100 under climate change conditions. *PloS one*, *9*(9), e107522.

12 Tables

Table 3: Comparison between farmers' beliefs and true genetics of the variety grown.

	Farmer's belief			
Genetics	Improved		Non-improved	
Improved	True (TP)	Positive	False (FN)	Negative
Non-improved	False (FP)	Positive	True (TN)	Negative

Table 4: Cross-tabulation of DNA fingerprinting results and farmers' beliefs

DNA fingerprinting	**Farmer's belief**		**Total**
	Improved	Not improved	
Improved	23.3%	37.3%	60.6%
Not improved	22.1%	17.3%	39.4%
Total	45.4%	54.6%	100%

Notes: Sample size: 491.

Table 5: Descriptive statistics

	TP		FP		FN		TN		Total	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Inputs application										
Maize plot area (HA) - GPS	0.14	95	0.21	81	0.09	219	0.10	84	0.13	479
Total household area under maize (HA)	0.15	78	0.23	72	0.10	201	0.12	72	0.14	423
Urea use on plot	0.69	95	0.76	81	0.15	219	0.23	84	0.41	479
Quantity of UREA (Kg/HA)	141.20	95	122.43	81	20.02	219	51.23	84	73.41	479
Use of DAP on plot	0.34	95	0.04	81	0.10	219	0.08	84	0.14	479
Quantity of DAP (Kg/HA)	66.21	95	5.66	81	13.13	219	12.53	84	23.13	479
Use of NPS on plot	0.52	95	0.64	81	0.09	219	0.10	84	0.30	479
Quantity of NPS (Kg/HA)	68.07	95	106.12	81	12.55	219	10.23	84	43.71	479
Pre-harvest family labor (days/HA)	163.92	95	178.72	81	198.32	219	222.16	84	190.96	479
Pre-harvest hired labor (days/HA)	22.48	95	16.53	81	25.60	219	14.04	84	20.97	479
Harvest family labor (days/HA)	61.89	93	82.36	71	94.09	206	96.43	82	84.77	452
Harvest hired labor (days/HA)	5.11	93	1.18	71	7.18	206	2.12	82	4.62	452
Total labor per plot (days/HA)	252.72	95	266.72	81	317.55	219	330.85	84	295.16	479
Seedling rate (Kg/HA)	36.87	95	45.48	81	81.44	219	245.39	84	93.45	479
Plot characteristics (deviation from EA average)										
Potential Wetness Index	0.01	94	0.06	80	-0.08	207	-0.09	83	-0.03	464
Elevation (m)	-0.02	94	0.02	80	0.02	207	-0.01	83	0.01	464
Slope (percent)	-0.12	94	0.03	80	-0.08	207	-0.04	83	-0.06	464
Agricultural Practices										
Plot is irrigated	0.02	95	0.00	81	0.01	219	0.03	84	0.01	479
Plot prevented from soil erosion	0.85	95	0.70	81	0.72	219	0.59	84	0.72	479
River dispersion	0.00	95	0.00	81	0.00	219	0.01	84	0.00	479
Motor pump used for irrigation	0.02	95	0.00	81	0.01	219	0.00	84	0.01	479
Crop rotation with a legume	0.46	95	0.26	81	0.35	219	0.33	84	0.35	479
Crop residue cover - visual aid	0.40	95	0.36	81	0.37	217	0.29	84	0.36	477
Minimum tillage	0.17	95	0.36	81	0.39	219	0.53	84	0.36	479
Zero tillage	0.16	95	0.32	81	0.22	219	0.45	84	0.27	479
Soil Water Conservation practices	0.79	95	0.60	81	0.65	219	0.46	84	0.64	479
Terracing	0.40	95	0.32	81	0.21	219	0.18	84	0.27	479
Water catchments	0.07	95	0.01	81	0.09	219	0.03	84	0.06	479
Afforestation	0.03	95	0.03	81	0.02	219	0.01	84	0.03	479
Plough along the contour	0.30	95	0.24	81	0.32	219	0.24	84	0.28	479
Conservation Agriculture - using Minimum tillage	0.00	95	0.00	81	0.01	219	0.00	84	0.00	479
Household distance and government programs										
Proximity to woreda town (KM)	70.54	95	76.48	81	67.20	219	69.90	84	70.34	479
No. years of DSM	2.00	95	2.41	81	0.77	219	1.42	84	1.50	479
Input Voucher System	1.00	95	0.98	81	0.99	219	0.94	84	0.98	479
Suitability Index	20.86	95	34.09	81	20.91	219	23.08	84	24.02	479
Household Productivity and Profitability										
Maize yields (Kg/Ha)	2,352.06	78	3,019.34	72	1,582.68	201	2,257.49	72	2,157.67	423
Household maize profits (USD)	30.85	78	71.70	72	19.77	201	27.44	72	34.19	423
Household maize profits per land cultivated (USD/HA)	178.17	78	261.38	72	156.47	201	-81.26	72	142.04	423

Table 6: Descriptive statistics by categories of misperception.

	TP		FP		FN		TN		Total		P-value					
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	TP-FP	TP-FN	TP-TN	FP-FN	FP-TN	FN-TN
Household-head characteristics																
Male	0.88	80	0.83	72	0.86	204	0.82	67	0.85	423	0.405	0.782	0.418	0.476	0.967	0.492
Age (years)	45.66	80	46.75	72	48.48	204	45.57	67	47.04	423	0.716	0.253	0.977	0.492	0.698	0.248
Years of education completed	2.32	80	1.41	72	1.58	204	2.07	67	1.78	423	0.114	0.158	0.698	0.684	0.258	0.358
DNA fingerprinting seed characteristics																
OPV variety	0.18	95	0.04	81	0.82	219	0.08	84	0.39	479	0.002	0.000	0.036	0.000	0.249	0.000
Purity percent	96.80	95	96.81	81	93.38	219	92.47	84	94.68	479	0.988	0.000	0.000	0.000	0.000	0.209
Drought tolerant variety	0.54	95	0.00	81	0.11	219	0.00	84	0.16	479	0.000	0.000	0.000	0.000	.	0.000
Quality Protein Maize (QPM)	0.02	95	0.00	81	0.03	219	0.00	84	0.01	479	0.311	0.617	0.311	0.046	.	0.046
Self-reported - Seed purchase sources																
Seed used was purchased	0.93	91	0.92	79	0.21	215	0.30	80	0.53	465	0.815	0.000	0.000	0.000	0.000	0.231
Agr. Coop.	0.30	76	0.43	64	0.04	39	0.07	17	0.28	196	0.169	0.001	0.018	0.000	0.001	0.693
Farmer-based organization	0.34	76	0.19	64	0.01	39	0.00	17	0.20	196	0.059	0.000	0.000	0.001	0.000	0.334
Gov. Agency	0.19	76	0.08	64	0.00	39	0.00	17	0.10	196	0.100	0.000	0.000	0.014	0.014	.
Saving-Credit coop.	0.00	76	0.10	64	0.00	39	0.00	17	0.04	196	0.007	.	.	0.007	0.008	.
Local Mkt	0.02	76	0.14	64	0.41	39	0.65	17	0.19	196	0.023	0.000	0.000	0.014	0.000	0.125
Main Mkt	0.03	76	0.01	64	0.37	39	0.19	17	0.09	196	0.503	0.000	0.122	0.000	0.087	0.193
Relative	0.01	76	0.00	64	0.04	39	0.00	17	0.01	196	0.321	0.373	0.324	0.201	.	0.207
Friend/Neighbor	0.00	76	0.02	64	0.13	39	0.08	17	0.04	196	0.316	0.059	0.300	0.129	0.455	0.653
NGO	0.11	76	0.02	64	0.00	39	0.00	17	0.05	196	0.052	0.008	0.009	0.162	0.164	.
Self-reported - Seed type																
Traditional	0.00	95	0.00	81	1.00	219	1.00	84	0.57	479	.	.	.	.	.	.
Improved new	0.88	95	0.90	81	0.00	219	0.00	84	0.38	479	0.735	0.000	0.000	0.000	0.000	.
Improved saved from previous year	0.06	95	0.04	81	0.00	219	0.00	84	0.02	479	0.527	0.028	0.029	0.048	0.049	.
Improved recycled	0.06	95	0.07	81	0.00	219	0.00	84	0.03	479	0.935	0.013	0.014	0.028	0.029	.

Table 7: First-Stage results using correct identification of improved maize varieties - CGIAR-derived germplasm

	Correct identification of improved maize variety (CI)
	(1)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.268***
	(0.069)
No. years of DSM	1.215***
	(0.279)
Log (Proximity to woreda town (KM))	-0.442
	(2.467)
Region FE	Yes
Obs.	314
Adjusted R2	0.267
F-statistic	14.970
P-value	0.000

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Control variables included are: VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

Table 8: Plot level OLS and 2SLS results on inputs use and plot selection

	Plot area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS: Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)	Plot Potential Wetness Index †	Plot elevation (m) †	Plot slope (percent) †
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: OLS													
CI	0.039*	125.407***	36.410*	60.220***	-0.122	-0.259	-0.334	-0.166	-0.174	-37.988**	-0.099	-0.018	-0.080
	(0.020)	(25.857)	(19.684)	(17.961)	(0.253)	(0.554)	(0.275)	(0.335)	(0.202)	(18.268)	(0.208)	(0.029)	(0.098)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	314	314	314	314	314	314	299	299	314	314	301	301	301
Adj- R2	0.119	0.205	0.135	0.125	0.054	0.089	0.139	0.131	0.109	0.032	0.028	-0.010	-0.016
Panel B: 2SLS													
CI	0.022	164.226*	-40.425	129.675*	1.015	2.068	-0.200	0.851*	0.883	-102.240*	-0.125	-0.177*	-0.071
	(0.036)	(95.205)	(60.068)	(68.165)	(1.108)	(2.369)	(0.859)	(0.496)	(0.775)	(54.949)	(0.674)	(0.106)	(0.221)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	314	314	314	314	314	314	299	299	314	314	301	301	301
Control mean	0.091	19.494	12.874	12.908	183.816	23.724	87.362	5.515	293.146	66.511	-0.076	0.023	-0.078

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. IHS= inverse hyperbolic sine transformation.†Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table 7.

Table 9: Plot level OLS and 2SLS results on agricultural practices

	Irrigation	Soil erosion prevention	River dispersion	Motor pump	Rotation with Legume	Crop residue cover	Minimum tillage	Zero tillage	SWC	Terracing	Water Catchment	Afforestation	Plough along plot contour	CA	Factor analysis: all practices
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: OLS															
CI	0.006	0.200**	0.007	-0.001	0.178	-0.092	-0.198**	-0.017	0.198*	0.223***	0.010	-0.006	-0.030	-0.003	-0.364**
	(0.011)	(0.092)	(0.006)	(0.009)	(0.111)	(0.111)	(0.100)	(0.068)	(0.110)	(0.084)	(0.040)	(0.026)	(0.104)	(0.003)	(0.177)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes Yes	Yes	Yes	
Obs.	314	314	314	314	314	312	314	314	314	314	314	314	314	314	312
Adjusted R2	0.001	0.180	-0.013	-0.001	0.085	0.094	0.068	0.023	0.112	0.087	0.104	-0.028	0.019	-0.009	0.084
Panel B: 2SLS															
CI	0.006	0.071	0.003	0.003	0.341	-0.458	-0.177	0.210	-0.219	0.058	0.090	-0.125	-0.243	-0.017	0.014
	(0.011)	(0.310)	(0.003)	(0.011)	(0.347)	(0.323)	(0.181)	(0.180)	(0.489)	(0.246)	(0.225)	(0.080)	(0.248)	(0.018)	(0.467)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	314	314	314	314	314	312	314	314	314	314	314	314	314	314	312
Control mean	0.011	0.718	0.000	0.011	0.351	0.371	0.379	0.215	0.648	0.204	0.095	0.024	0.324	0.007	0.071

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table 7.

Table 10: First-Stage and 2SLS results of crowding-in vs. resource reallocation at household level

	C.I.	Total Area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS: Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	Seedling rate (kg/ha)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: First-stage household level										
No. years of DSM * Log(Proximity to woreda town -KM)	-0.298***									
	(0.065)									
F-statistic	21.030									
P-value	0.000									
Panel B: 2SLS plots under maize cultivation										
CI		0.006	146.174	-18.342	110.549**	0.535	2.159	0.846	0.738**	-93.951*
		(0.038)	(93.113)	(42.693)	(54.898)	(0.894)	(2.263)	(1.190)	(0.362)	(49.717)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	289	289	289	289	289	289	289	289	289	289
Control mean	-	0.092	21.138	13.444	10.421	286.649	36.235	95.598	9.407	79.795
Panel C: 2SLS plots under all other crops except maize										
CI		-0.450	78.794*	-19.570	87.982**	0.201	1.572	0.398	0.377	41.392
		(0.385)	(42.617)	(17.205)	(41.003)	(0.612)	(2.020)	(0.571)	(0.369)	(26.658)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	289	289	289	289	289	289	289	289	289	289
Control mean	-	0.886	21.744	15.131	14.561	126.402	18.916	74.239	6.720	6.908

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Column(1): first-stage results. C.I.= Correct identification of improved varieties. Column(2)-(10) report 2SLS results. IHS= inverse hyperbolic sine transformation. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM).

Table 11: OLS and 2SLS results of correct identification on productivity and profitability of maize plots

	Log [Yield (Kg/Ha)]	Profit (USD)	Profit (USD/Ha)
	(1)	(2)	(3)
	Panel A: OLS		
CI	0.307**	10.780	12.827
	(0.140)	(8.544)	(37.317)
	Panel B: 2SLS		
CI	1.237***	21.600	252.643***
	(0.366)	(16.725)	(96.531)
Region F.E.	Yes	Yes	Yes
Obs.	312	289	289
Control mean	1,614.008	18.840	151.392

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table 7 for column(1) and Table 10 for column(2).

Table 12: Reduced-form regression at household level using ESS1, ESS2, ESS3 and ESS4 of household unbalanced panel sample of maize growing areas

		Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS: Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESS4	No. years of DSM * Log(Proximity to woreda town -KM)	-0.076*	-94.142**	-44.180	-63.008	0.163	-0.574	-0.194	-0.244**	0.117	-232.443
		(0.045)	(44.038)	(47.186)	(45.332)	(0.197)	(0.405)	(0.237)	(0.099)	(0.186)	(261.864)
	Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Obs.	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152
	Adjusted R2	0.077	0.001	-0.002	-0.001	0.022	0.078	0.082	0.054	0.033	-0.005
	Mean	0.227	108.452	60.755	75.107	298.907	39.286	191.222	5.137	38.745	708.972
ESS3	No. years of DSM * Log(Proximity to woreda town -KM)	-0.018	11.390	-5.854	10.746**	40.526	0.056	-0.265*	0.113	0.100	-146.081
		(0.031)	(12.645)	(9.370)	(4.703)	(27.326)	(0.071)	(0.144)	(0.097)	(0.125)	(99.932)
	Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Obs.	1,555	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552
	Adjusted R2	0.215	0.091	0.146	0.105	0.063	0.024	0.061	0.103	0.074	-0.005
	Mean	0.239	71.981	41.860	12.671	202.918	4.563	61.248	3.191	384.389	232.656
ESS2	No. years of DSM * Log(Proximity to woreda town -KM)	0.020	-5.656	-4.195	n.a.	-25.932	0.087	-0.118	0.040	-0.234**	n.a.
		(0.019)	(8.043)	(9.233)	n.a.	(16.236)	(0.066)	(0.129)	(0.078)	(0.103)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	1,506	1,502	1,502	n.a.	1,502	1,502	1,503	1,502	1,503	n.a.
	Adjusted R2	0.226	0.135	0.197	n.a.	0.036	0.011	0.039	0.067	0.075	n.a.
	Mean	0.229	47.491	49.599	n.a.	191.294	6.040	74.214	1.577	672.490	n.a.
ESS1	No. years of DSM * Log(Proximity to woreda town -KM)	-0.068	5.933	17.363	n.a.	0.362*	-0.138	-0.241	-0.008	0.245	n.a.
		(0.071)	(10.962)	(19.914)	n.a.	(0.196)	(0.135)	(0.204)	(0.054)	(0.227)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	1,383	1,383	1,383	n.a.	1,383	1,383	1,383	1,383	1,383	n.a.
	Adjusted R2	0.027	0.042	-0.006	n.a.	0.036	0.085	0.063	0.043	0.045	n.a.
	Mean	0.317	41.616	85.028	n.a.	290.437	6.568	51.007	0.857	947.315	n.a.

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Unbalanced household level panel. IHS: inverse hyperbolic sine transformation. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). Sample restricted to maize growing areas: Amhara, Dire Dawa, Harar, Oromia, SNNP, Tigray regions.

Table 13: Testing parallel trends development at household level based on IV

	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Family labor - planting - (days/ha)	Hired labor - planting - (days/ha)	Family labor - harvest - (days/ha)	Hired labor - harvest - (days/ha)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: ESS2-ESS1							
No. years of DSM * Log(Proximity to woreda town -KM)	-0.003	0.442	-1.952	1.489	-0.226	1.372	0.104
	(0.008)	(3.007)	(2.215)	(6.755)	(1.208)	(1.627)	(0.121)
Obs.	703	703	703	703	703	703	703
Panel B: ESS3-ESS2							
No. years of DSM * Log(Proximity to woreda town -KM)	0.002	-0.231	-0.345	-11.540**	1.909	-2.011	0.139
	(0.003)	(1.481)	(0.967)	(5.097)	(1.799)	(1.490)	(0.205)
Obs.	703	703	703	703	703	703	703
Panel C: ESS3-ESS1							
No. years of DSM * Log(Proximity to woreda town -KM)	-0.001	0.210	-2.297	-10.052	1.683	-0.639	0.243
	(0.008)	(2.668)	(2.342)	(7.180)	(1.265)	(1.257)	(0.181)
Obs.	703	703	703	703	703	703	703

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Household panel sample restricted to maize growing areas: Amhara, Dire Dawa, Harar, Oromia, SNNP, Tigray regions.

Table 14: Test for misreporting hypothesis.

	Log(Plot area (HA))	Quantity of DAP & NPS (kg/ha)	Urea (dummy)	Quantity of urea (kg/ha)	DAP (dummy)	Quantity of DAP (kg/ha)	NPS (dummy)	Quantity of NPS (kg/Ha) - wins.	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Self-reported	0.052***	102.330***	0.585***	117.156***	0.099	23.572**	0.498***	78.758***	-24.102	-2.747	-17.554	-2.324	-45.520	-120.981**
	(0.015)	(14.104)	(0.062)	(21.817)	(0.067)	(11.929)	(0.074)	(14.004)	(29.067)	(8.304)	(15.655)	(2.955)	(40.456)	(53.223)
DNA-fingerprinting	-0.032*	12.565	-0.073	-9.825	0.143**	27.982***	-0.050	-15.418	-23.603	7.572	-11.857	3.286	-22.004	-94.867*
	(0.017)	(11.604)	(0.067)	(18.942)	(0.057)	(10.800)	(0.071)	(11.739)	(29.765)	(7.351)	(14.763)	(2.533)	(41.727)	(55.855)
Naïve Wald	0.014	0.413	0.031	0.207	0.001	0.003	0.029	0.108	0.119	0.009	0.229	0.005	0.407	0.038
LM-Wald	0.031	0.502	0.073	0.306	0.006	0.023	0.068	0.176	0.190	0.030	0.285	0.013	0.482	0.079
Number of observations	479	479	479	479	479	479	479	479	479	479	452	452	479	479
Adjusted R2	0.125	0.203	0.334	0.166	0.111	0.088	0.300	0.172	0.018	0.029	0.034	0.094	0.032	0.050

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. IHS= inverse hyperbolic sine transformation.

Table 15: Machine learning algorithms performance by Mean Square Error

	MSE			
Algorithm	Average	s.e.	Min	Max
Super Learner	0.1327	0.0097	0.1078	0.1463
Discrete SL	0.1336	0.0094	0.1044	0.1537
SL.ipredbagg_All	0.1407	0.0087	0.1288	0.1532
SL.xgboost_All	0.1465	0.0120	0.1284	0.1608
SL.randomForest_1_screen.glmnet	0.1336	0.0094	0.1044	0.1537

Risk is based on: Mean Squared Error. All risk estimates are based on 5-fold cross-validation.

Table 16: First-stage results using predictions from Machine Learning application

	Predicted values of:			
	Original sample	Same Eas	Same Regions	Full sample
	(1)	(2)	(3)	(4)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.271***	-0.262***	-0.175**	-0.175**
	(0.082)	(0.080)	(0.086)	(0.086)
Region FE	Yes	Yes	Yes	Yes
Obs.	313	820	956	1,121
Adjusted R2	0.295	0.214	0.165	0.168
F-statistic for IV	11.05	10.64	4.13	4.13
P-value	0.0013	0.0016	0.0444	0.044

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

Table 17: 2SLS results using predictions from Machine Learning application

	Plot area (HA)	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS: Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)	Plot Potential Wetness Index †	Plot elevation (m) †	Plot slope (percent) †	Log(Yield (kg/ha))
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(14)
Panel A: Original Sample															
CI	0.042	0.208	159.803*	-37.999	131.267*	0.942	1.952	1.037	0.485*	0.694	-65.001*	-0.646	-0.022	-0.423	0.194
	(0.045)	(0.198)	(87.281)	(63.174)	(67.794)	(1.082)	(2.271)	(1.286)	(0.279)	(0.670)	(36.371)	(0.776)	(0.050)	(0.270)	(0.571)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	313	313	313	313	313	313	313	313	313	313	313	313	313	313	294
Control mean	0.091	0.154	19.126	13.389	10.607	187.493	21.266	71.235	1.116	281.110	69.233	-0.017	-0.004	-0.008	4,894.340
Panel B: Same EAs															
CI	0.113	0.261	207.553***	-6.678	154.086**	-0.180	4.232	0.834	0.515**	-0.039	-76.458*	-0.814	0.006	0.210	0.458
	(0.069)	(0.196)	(62.909)	(50.902)	(62.140)	(0.919)	(3.191)	(0.965)	(0.236)	(0.857)	(40.396)	(0.895)	(0.021)	(0.158)	(0.604)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	820	820	820	820	820	820	820	820	820	820	820	820	820	820	716
Control mean	0.087	0.174	20.081	9.604	14.299	211.816	22.970	64.204	1.154	300.144	109.301	-0.070	0.004	-0.003	4,281.734
Panel C: Same Regions															
CI	0.179	0.461	209.000**	-30.837	195.449*	-0.508	5.199	0.287	0.606*	-0.531	30.457	-0.706	0.113	0.247	0.962
	(0.128)	(0.368)	(83.243)	(79.881)	(100.478)	(1.153)	(3.558)	(1.495)	(0.340)	(1.260)	(148.029)	(0.894)	(0.117)	(0.200)	(0.897)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	956	956	956	956	956	956	956	956	956	956	956	956	956	956	792
Control mean	0.094	0.197	23.598	10.551	19.240	214.969	25.831	62.067	1.031	303.899	117.438	-0.058	0.005	0.002	4,667.119
Panel D: Full Sample															
CI	0.178	0.461	209.666**	-30.846	195.115*	-0.485	5.201	0.331	0.608*	-0.511	29.963	-0.707	0.114	0.247	0.963
	(0.128)	(0.368)	(83.323)	(79.875)	(100.494)	(1.150)	(3.561)	(1.491)	(0.341)	(1.258)	(148.184)	(0.893)	(0.117)	(0.200)	(0.898)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	951
Control mean	0.096	0.200	23.568	10.515	19.206	215.423	25.463	62.231	1.022	304.140	116.905	-0.057	0.005	0.002	4,816.879

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. IHS: inverse hyperbolic sine transformation. †Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Yield are calculated from self-reported data (only here). Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table 16.

13 Figures

Figure 1: Geographical roll-out of Direct Seed Marketing program for maize by year

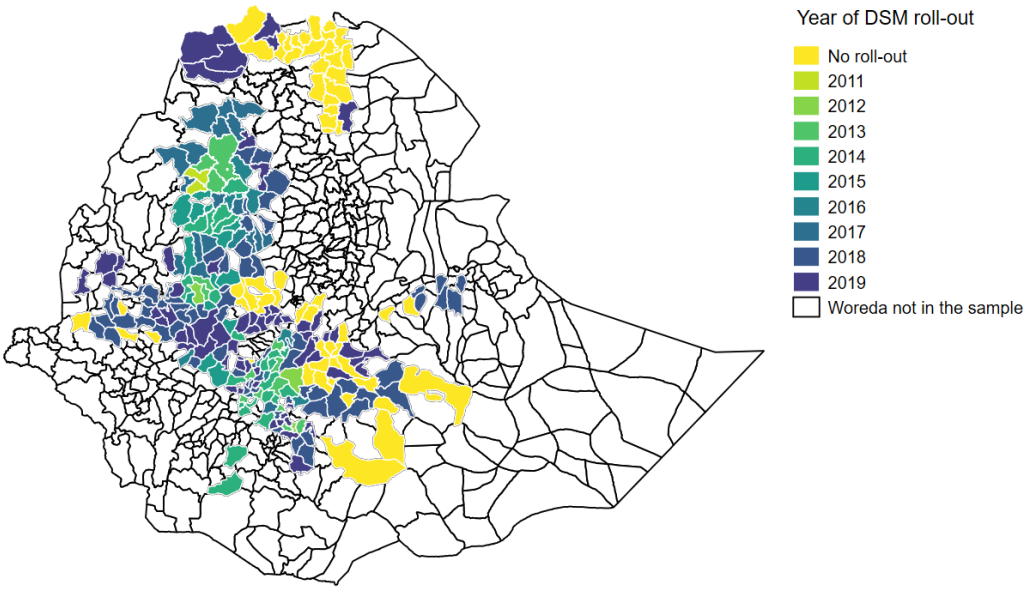
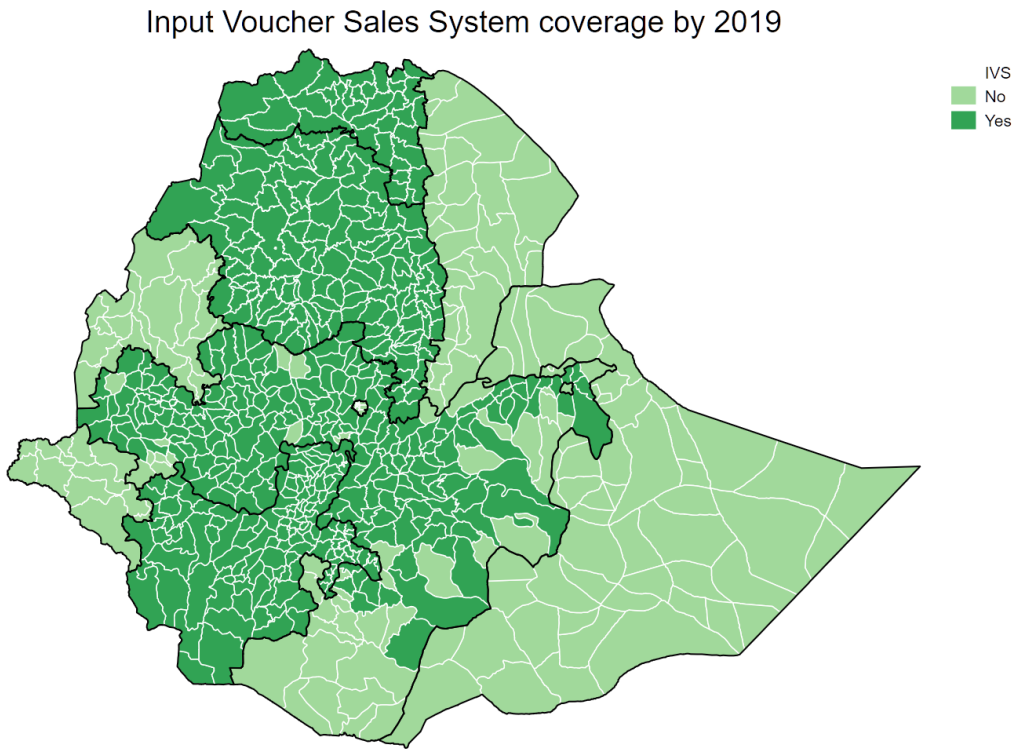


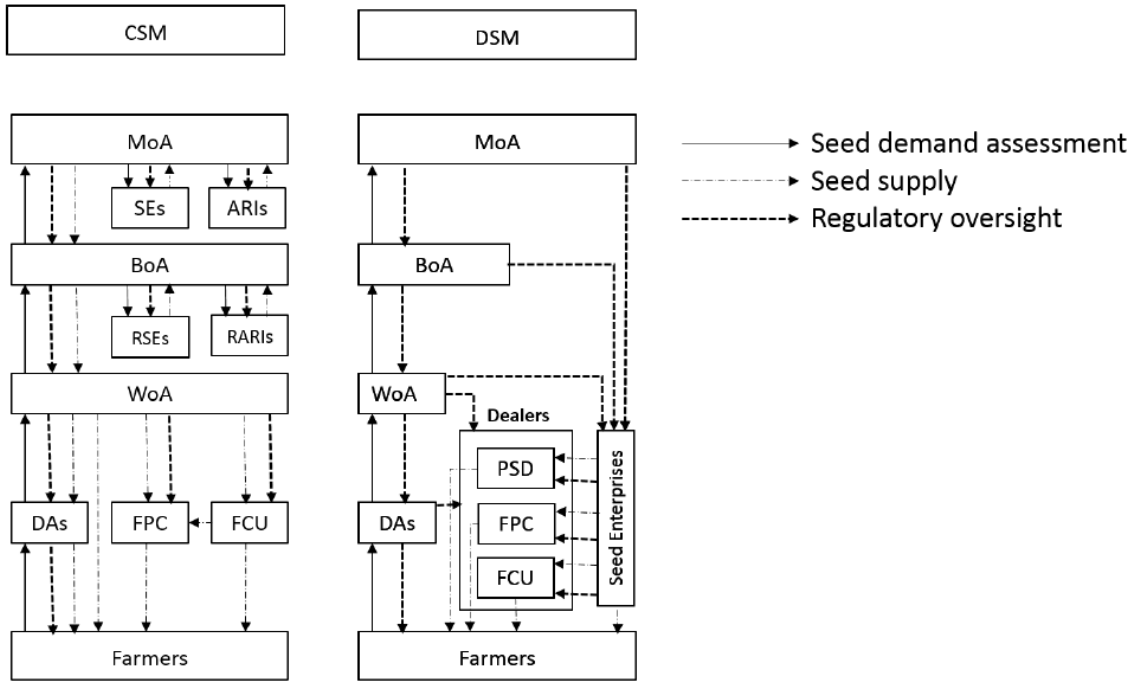
Figure 2: Geographical coverage of Input Voucher Sales System by 2019



14 Appendix

14.1 Figures

Figure A1: Seed demand assessment and supply flow through the Conventional Seed Marketing and Direct Seed Marketing systems in Ethiopia



Source: IFPRI, 2015. CSM = conventional seed marketing; DSM = direct seed marketing; MoA = Ministry of Agriculture (national level); SE = seed enterprises; ARI = agricultural research institutes; BoA = Bureau of Agriculture (region level); RSE = regional seed enterprises (regional government); RARI = regional agricultural research institutes; WoA = Woreda Office of Agriculture; DA = Development Agent (farm-level agricultural extension agent); FPC = farmers' primary cooperative; FCU = farmers' cooperative union; PSD = private seed dealer.

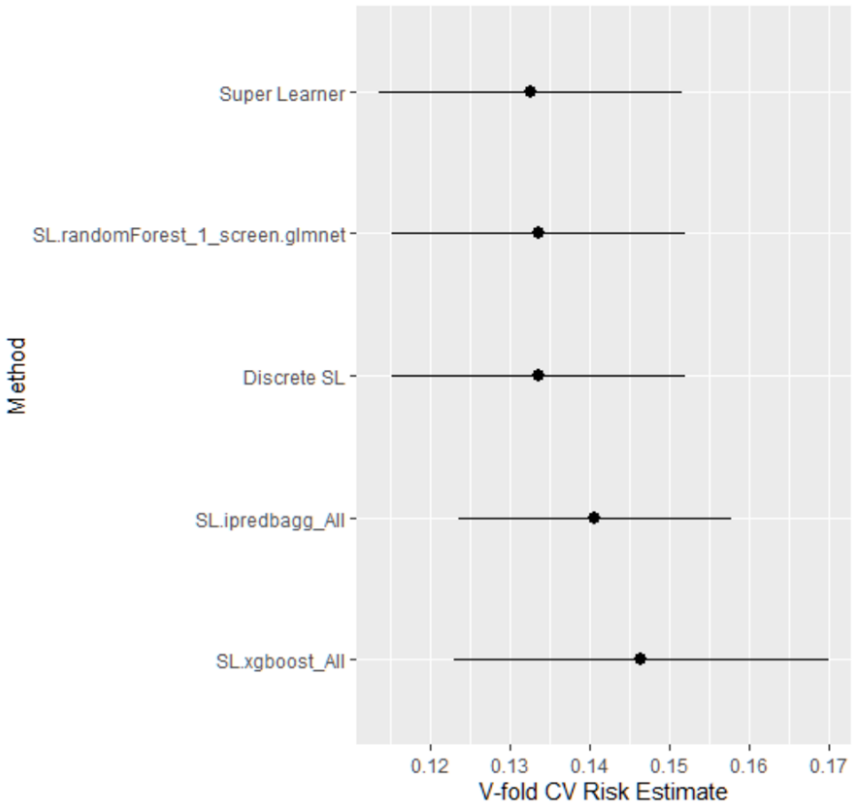


Figure A2: 5-fold CV Risk Estimate

14.2 Tables

Table A1: Testing the predictive power of the instrumental variable

	Improved variety (Self-reports)	Improved variety (DNA)	CI of improved or traditional
	(1)	(2)	(3)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.110	0.068	-0.286***
	(0.093)	(0.098)	(0.067)
Region FE	Yes	Yes	Yes
Number of observations	479	479	479
Adjusted R2	0.175	0.081	0.132
F-statistic	1.41	0.48	18.16
P-value	0.2377	0.4887	0.0000

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

Table A2: Bivariate probit model estimates of the extensive margin of input use

	Urea	NPS
	(1)	(2)
C.I.	0.182***	0.171***
	(0.032)	(0.048)
Obs.	314	314
Control mean	0.149	0.085

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Coefficients reported are marginal effects. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)). Suitability Index Suitability Index*Log(Proximity to woreda town KM) are not included to facilitate model convergence.

Table A3: IV-OLS gap decomposition

	Coefficients			Decomposition		
Variables	OLS	IV	IV-OLS	Δ_{CW}	Δ_{TW}	Δ_{ME}
Plot area (HA)	0.039	0.022	-0.017	0.008	0.000	-0.026
(s.e.)	0.020	0.036	0.043	0.041	0.000	0.040
Total maize area (HA)	0.051	0.020	-0.030	0.000	0.000	-0.030
(s.e.)	0.023	0.049	0.056	0.048	0.000	0.038
Quantity of urea (kg/ha)	125.407	164.226	38.819	6.644	0.000	32.174
(s.e.)	25.857	95.205	94.568	62.211	0.000	51.559
Quantity of DAP (kg/ha)	36.410	-40.425	-76.835	-25.782	0.000	-51.053
(s.e.)	19.684	60.068	56.882	40.524	0.000	47.846
Quantity of NPS (kg/Ha) - wins.	60.220	129.675	69.455	33.200	0.000	36.255
(s.e.)	17.961	68.165	68.450	40.251	0.000	46.229
IHS: Family labor - planting - (days/ha)	-0.122	1.015	1.137	-0.463	0.000	1.600
(s.e.)	0.253	1.108	1.123	0.638	0.000	0.874
IHS: Hired labor - planting - (days/ha)	-0.259	2.068	2.327	0.857	0.000	1.470
(s.e.)	0.554	2.369	2.195	1.079	0.000	
IHS:Family labor - harvest - (days/ha)	-0.334	-0.200	0.133	-0.070	0.000	0.203
(s.e.)	0.275	0.859	0.801	0.512	0.000	0.515
IHS: Hired labor - harvest - (days/ha)	-0.166	0.851	1.017	0.573	0.000	0.444
(s.e.)	0.335	0.496	0.572	0.449	0.000	0.436
IHS: Total labor (family and hired) (days/ha)	-0.174	0.883	1.057	-0.010	0.000	1.067
(s.e.)	0.202	0.775	0.789	0.345	0.000	0.649
Seedling rate (kg/ha)	-37.988	-102.240	-64.252	-21.241	0.000	-43.011
(s.e.)	18.268	54.949	46.092	26.069	0.000	42.074
Plot Potential Wetness Index	-0.099	-0.125	-0.026	0.573	0.000	-0.598
	0.208	0.674	0.739	0.321	0.000	0.732
Plot elevation (m)	-0.018	-0.177	-0.159	-0.098	0.000	-0.061
	0.029	0.106	0.097	0.076	0.000	0.079
Plot slope (percent)	-0.080	-0.071	0.009	-0.204	0.000	0.213
	0.098	0.221	0.227	0.188	0.000	0.241
Log(Yield (KG/Ha)	0.310	1.225	0.914	0.484	0.000	0.430
	0.141	0.367	0.371	0.248	0.000	0.279
Maize Profits (ETB)	1186.322	1801.139	614.817	-5.858	0.000	620.675
	924.930	1195.526	1505.160	1338.036	0.000	1190.140

Notes:Δ_{CW}:Covariate weight difference; Δ_{TW}: Treatment-level weight difference; Δ_{ME}: Endogeneity bias. Standard errors clustered at EA level. IHS= inverse hyperbolic sine transformation. Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM)

Table A4: Reduced-form regression results using ESS4 at household-level - DNA fingerprinting sample restricted to true positive and false negative.

	Plot area (HA)	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS: Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)	Plot Potential Wetness Index †	Plot elevation (m) †	Plot slope (percent) †	Log(Yield (kg/ha))	Profits (BIRR)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(14)	(15)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.006	-0.005	-43.931*	10.814	-34.688**	-0.272	-0.553	0.055	-0.233*	-0.236	27.349	0.034	0.047	0.019	-0.327***	-641.682*
	(0.010)	(0.014)	(22.870)	(14.388)	(13.711)	(0.248)	(0.630)	(0.245)	(0.140)	(0.165)	(18.555)	(0.181)	(0.030)	(0.061)	(0.104)	(377.445)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	314	314	314	314	314	314	314	299	299	314	314	301	301	301	312	314
R2	0.136	0.146	0.111	0.137	0.103	0.098	0.137	0.160	0.169	0.150	0.065	0.066	0.035	0.023	0.203	0.365
Adjusted R2	0.101	0.112	0.076	0.103	0.067	0.062	0.103	0.125	0.134	0.116	0.028	0.027	-0.006	-0.018	0.171	0.339

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Unbalanced household level panel. IHS: inverse hyperbolic sine transformation. †Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM).

Table A5: Reduced-form regression at household level using ESS1, ESS2, ESS3 of household balanced panel sample of maize growing areas

		Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESS3	No. years of DSM * Log(Proximity to woreda town -KM)	-0.007	9.269	-8.834	13.688**	0.192	0.017	-0.227	0.078	0.129	-254.581
		(0.035)	(14.368)	(10.940)	(5.629)	(0.167)	(0.091)	(0.182)	(0.127)	(0.158)	(188.702)
	Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Obs.	946	946	946	946	946	946	946	946	946	946
	Adjusted R2	0.209	0.072	0.106	0.128	0.047	0.023	0.034	0.099	0.063	-0.010
	Mean	0.239	71.981	41.860	12.671	202.918	4.563	61.248	3.191	384.389	232.656
ESS2	No. years of DSM * Log(Proximity to woreda town -KM)	0.003	-5.095	-5.599	n.a.	-0.178	0.080	-0.054	0.001	-0.148	n.a.
		(0.023)	(8.542)	(9.704)	n.a.	(0.125)	(0.089)	(0.152)	(0.099)	(0.109)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	937	937	937	n.a.	937	937	937	937	937	n.a.
	Adjusted R2	0.232	0.145	0.204	n.a.	0.075	0.013	0.035	0.071	0.075	n.a.
	Mean	0.229	47.491	49.599	n.a.	191.294	6.040	74.214	1.577	672.490	n.a.
ESS1	No. years of DSM * Log(Proximity to woreda town -KM)	-0.040	10.682	25.993	n.a.	0.209	-0.129	-0.124	0.029	0.114	n.a.
		(0.055)	(12.856)	(19.068)	n.a.	(0.177)	(0.132)	(0.176)	(0.064)	(0.188)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	946	946	946	n.a.	946	946	946	946	946	n.a.
	Adjusted R2	0.030	0.040	0.020	n.a.	0.019	0.050	0.109	0.053	0.019	n.a.
	Mean	0.317	41.616	85.028	n.a.	290.437	6.568	51.007	0.857	947.315	n.a.

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Unbalanced household level panel. IHS: inverse hyperbolic sine transformation. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), VSS, VSS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). Sample restricted to maize growing areas: Amhara, Dire Dawa, Harar, Oromia, SNNP, Tigray regions.

Table A6: Datasets and observations used for machine learning

Datasets		Obs
Train	short	391
	validation	100
Test		1460

14.3 Crop germplasm improvement

In the context of crop germplasm improvement, CGIAR centers play a central role in Ethiopia. Relying on its network of genebanks, centers have developed and made available to the National Agricultural Research System (NARSs) a multitude of sources of crop germplasm. Improved crop varieties developed through CGIAR research were released through a collaborative research process with the Ethiopian Institute of Agricultural Research (EIAR). Overall, 54 maize varieties have been released in Ethiopia since 1990, and 34 of them are thought to contain CIMMYT-related germplasm. Two varieties, released in 1986 and 2001, were created from parent lines from the International Institute of Tropical Agriculture (IITA). Varieties that contain CGIAR-related germplasm span the diversity of Ethiopian agroecologies. Ten are open-pollinated varieties (OPVs), and 25 are hybrids. In the past 20 years ten drought-tolerant varieties and eight quality protein maize (QPM) varieties have been released. Thus, given the importance of CGIAR's efforts both in the development of varieties and their dissemination, we use as a definition of improved variety if it contains CGIAR-derived germplasm. Based on this definition, our main explanatory variable of interest is the correct classification of improved maize varieties, that is when the crop grown on one plot is an improved variety by both farmers' elicitation and DNA fingerprinting analysis (Kosmowski et al. 2020).

14.4 Agricultural Commercialization Cluster

The Agricultural Commercialization Cluster is an initiative introduced in ATA's agenda by the Ministry of Agriculture with the overall aim to address the systemic bottlenecks in the agricultural sector. It entails forming clusters of woredas where trainings and large-scale demonstrations are provided to farmers on new farming technologies, crop management techniques, crop protection and soil fertility and health. Moreover, the initiative aimed to ensure easier access to input financing through scaling the Input Voucher Sales system and better and timely distribution of fertilizer and agrochemicals.

Table A7: First-Stage results using correct identification of improved maize varieties - CGIAR-derived germplasm

	Correct identification of improved maize variety
	(1)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.277***
	(0.065)
Region FE	Yes
Number of observations	314
Adjusted R2	0.269
F-statistic	18.240
P-value	0.000

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

Table A8: Plot level 2SLS results on inputs use and plot selection

	Plot area (HA)	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)	Plot Potential Wetness Index †	Plot elevation (m)†	Plot slope (percent)†
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: 2SLS														
CI	0.016	0.011	149.253*	-37.662	127.636**	0.942	2.253	-0.149	0.864*	0.857	-99.627*	-0.024	-0.177*	-0.125
	(0.035)	(0.045)	(77.231)	(50.371)	(53.725)	(0.903)	(2.258)	(0.723)	(0.465)	(0.570)	(53.653)	(0.532)	(0.100)	(0.212)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	314	314	314	314	314	314	314	299	299	314	314	301	301	301
Control mean	0.091	0.103	19.494	12.874	12.908	183.816	23.724	87.362	5.515	293.146	66.511	-0.076	0.023	-0.078

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. IHS= inverse hyperbolic sine transformation.†Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table A7.

Table A9: Plot level 2SLS results on agricultural practices

	Irrigation	Soil erosion prevention	River dispersion	Motor pump	Rotation with Legume	Crop residue cover	Minimum tillage	Zero tillage	SWC	Terracing	Water Catchment	Afforestation	Plough along plot contour	CA	Factor analysis: all practices
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
CI	0.004	0.033	0.003	0.001	0.310	-0.407	-0.180	0.195	-0.244	0.028	0.089	-0.121*	-0.240	-0.016	0.016
	(0.010)	(0.264)	(0.003)	(0.010)	(0.344)	(0.329)	(0.172)	(0.159)	(0.450)	(0.247)	(0.224)	(0.068)	(0.211)	(0.017)	(0.455)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	314	314	314	314	314	312	314	314	314	314	314	314	314	314	312
Control mean	0.011	0.718	0.000	0.011	0.351	0.371	0.379	0.215	0.648	0.204	0.095	0.024	0.324	0.007	0.071

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table A7.

Table A10: First-Stage and 2SLS results of crowding-in vs. resource reallocation at household level

	CI	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	Seedling rate (kg/ha)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: First-stage household level										
No. years of DSM * Log(Proximity to woreda town -KM)	-0.308***									
	(0.064)									
F-statistic	23.450									
P-value	0.000									
Panel B: 2SLS plots under maize cultivation										
CI		0.007	135.415*	-19.258	111.369**	0.544	2.384	0.740**	0.880	-93.385*
		(0.036)	(77.103)	(38.392)	(43.548)	(0.780)	(2.158)	(0.337)	(0.940)	(48.866)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	289	289	289	289	289	289	289	289	289	289
Control mean	-	0.092	21.138	13.444	10.421	286.649	36.235	95.598	9.407	79.795
Panel C: 2SLS plots under all other crops except maize										
CI		-0.478	72.383*	-19.419	82.489**	0.294	1.672	0.345	0.457	39.985
		(0.381)	(37.575)	(14.737)	(34.176)	(0.625)	(1.944)	(0.318)	(0.539)	(25.685)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	289	289	289	289	289	289	289	289	289	289
Control mean	-	0.886	21.744	15.131	14.561	126.402	18.916	74.239	6.720	6.908

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level.C.I.=Correct identification of improved varieties. IHS= inverse hyperbolic sine transformation. Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Panel A of this table.

Table A11: 2SLS results of correct identification on productivity and profitability

	Log(Yield (kg/ha))	Profits (BIRR)
	(1)	(2)
C.I.	1.167***	1,696.598*
	(0.360)	(984.461)
Region FE	Yes	Yes
Obs.	312	289
Control mean	2778.929	2140.99

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table A7 for column(1) and Table A10 for column(2).

Table A12: Reduced-form regression results using ESS4- DNA fingerprinting sample restricted to true positive and false negative.

	Plot area (HA)	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)	Plot Potential Wetness Index†	Plot elevation (m)†	Plot slope (percent)†	IHS: Yield (kg/ha)	Profits (BIRR)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(14)	(15)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.004	-0.003	-41.307**	10.423	-35.324***	-0.261	-0.624	0.042	-0.243*	-0.237*	27.573	0.007	0.049*	0.035	-0.322***	-571.955*
	(0.010)	(0.013)	(20.697)	(12.591)	(11.472)	(0.211)	(0.628)	(0.211)	(0.139)	(0.121)	(18.867)	(0.151)	(0.028)	(0.059)	(0.117)	(324.864)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	314	314	314	314	314	314	314	299	299	314	314	301	301	301	312	314
R2	0.155	0.175	0.123	0.184	0.158	0.150	0.202	0.193	0.169	0.189	0.061	0.075	0.036	0.034	0.236	0.301
Adjusted R2	0.121	0.142	0.088	0.152	0.125	0.116	0.170	0.159	0.134	0.157	0.024	0.037	-0.004	-0.006	0.206	0.274

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Unbalanced household level panel. IHS: inverse hyperbolic sine transformation. †Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM).

Table A13: Reduced-form regression at household level using ESS1, ESS2, ESS3 and ESS4 of household unbalanced panel sample of maize growing areas

		Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESS4	No. years of DSM * Log(Proximity to woreda town -KM)	-0.054	-67.024*	-31.394	-72.897	0.131	-0.477	-0.097	-0.253**	0.157	-252.364
		(0.042)	(34.308)	(34.140)	(58.271)	(0.211)	(0.569)	(0.250)	(0.100)	(0.215)	(269.043)
	Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Obs.	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152
	Adjusted R2	0.097	0.001	-0.002	-0.003	0.032	0.081	0.146	0.049	0.058	-0.006
	Mean	0.227	108.452	60.755	75.107	298.907	39.286	191.222	5.137	38.745	708.972
ESS3	No. years of DSM * Log(Proximity to woreda town -KM)	-0.049	-3.609	-15.221	12.730**	26.012	0.041	-0.157	0.077	0.160	-148.689
		(0.047)	(18.754)	(13.547)	(5.358)	(28.274)	(0.117)	(0.236)	(0.139)	(0.190)	(121.699)
	Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Obs.	1,555	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552
	Adjusted R2	0.238	0.135	0.149	0.141	0.063	0.024	0.053	0.103	0.064	-0.005
	Mean	0.239	71.981	41.860	12.671	202.918	4.563	61.248	3.191	384.389	232.656
ESS2	No. years of DSM * Log(Proximity to woreda town -KM)	0.023	-10.831	-8.384	n.a.	-31.819	-0.016	-0.243	0.015	-0.320**	n.a.
		(0.021)	(10.658)	(12.164)	n.a.	(22.432)	(0.085)	(0.171)	(0.077)	(0.157)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	1,506	1,502	1,502	n.a.	1,502	1,502	1,503	1,502	1,503	n.a.
	Adjusted R2	0.234	0.159	0.201	n.a.	0.047	0.018	0.036	0.071	0.083	n.a.
	Mean	0.229	47.491	49.599	n.a.	191.294	6.040	74.214	1.577	672.490	n.a.
ESS1	No. years of DSM * Log(Proximity to woreda town -KM)	-0.240**	-5.222	8.199	n.a.	0.565	-0.306	0.132	-0.133*	0.517	n.a.
		(0.115)	(20.734)	(42.908)	n.a.	(0.402)	(0.244)	(0.409)	(0.081)	(0.390)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	1,383	1,383	1,383	n.a.	1,383	1,383	1,383	1,383	1,383	n.a.
	Adjusted R2	0.033	0.039	-0.005	n.a.	0.043	0.088	0.057	0.054	0.059	n.a.
	Mean	0.317	41.616	85.028	n.a.	290.437	6.568	51.007	0.857	947.315	n.a.

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Unbalanced household level panel. IHS: inverse hyperbolic sine transformation. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). Sample restricted to maize growing areas: Amhara, Dire Dawa, Harar, Oromia, SNNP, Tigray regions.

Table A14: Reduced-form regression at household level using ESS1, ESS2, ESS3 of household balanced panel sample of maize growing areas

		Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS:Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESS3	No. years of DSM * Log(Proximity to woreda town -KM)	-0.031	-20.773	-26.634	16.950**	0.137	-0.065	-0.049	-0.073	0.158	-262.707
		(0.063)	(20.217)	(17.561)	(7.865)	(0.274)	(0.178)	(0.263)	(0.196)	(0.241)	(193.024)
	Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Obs.	946	946	946	946	946	946	946	946	946	946
	Adjusted R2	0.218	0.130	0.117	0.143	0.038	0.023	0.041	0.104	0.051	-0.010
	Mean	0.239	71.981	41.860	12.671	202.918	4.563	61.248	3.191	384.389	232.656
ESS2	No. years of DSM * Log(Proximity to woreda town -KM)	0.006	-11.363	-11.285	n.a.	-0.264	-0.036	-0.263	-0.038	-0.307*	n.a.
		(0.027)	(10.859)	(12.202)	n.a.	(0.170)	(0.128)	(0.198)	(0.110)	(0.162)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	937	937	937	n.a.	937	937	937	937	937	n.a.
	Adjusted R2	0.243	0.162	0.207	n.a.	0.087	0.020	0.049	0.080	0.097	n.a.
	Mean	0.229	47.491	49.599	n.a.	191.294	6.040	74.214	1.577	672.490	n.a.
ESS1	No. years of DSM * Log(Proximity to woreda town -KM)	-0.158*	3.117	23.243	n.a.	0.354	-0.406*	0.222	-0.097	0.363	n.a.
		(0.094)	(27.778)	(47.654)	n.a.	(0.415)	(0.242)	(0.384)	(0.101)	(0.401)	n.a.
	Region FE	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Obs.	946	946	946	n.a.	946	946	946	946	946	n.a.
	Adjusted R2	0.040	0.036	0.021	n.a.	0.040	0.061	0.104	0.067	0.051	n.a.
	Mean	0.317	41.616	85.028	n.a.	290.437	6.568	51.007	0.857	947.315	n.a.

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Unbalanced household level panel. IHS: inverse hyperbolic sine transformation. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). Sample restricted to maize growing areas: Amhara, Dire Dawa, Harar, Oromia, SNNP, Tigray regions.

Table A15: First-stage results using predictions from Machine Learning application

	Predicted values of:			
	Original sample	Same Eas	Same Regions	Full sample
	(1)	(2)	(3)	(4)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.283***	-0.266***	-0.192**	-0.192**
	(0.081)	(0.083)	(0.087)	(0.087)
Region FE	Yes	Yes	Yes	Yes
Obs.	313	820	956	1,121
Adjusted R2	0.269	0.203	0.164	0.167
F-statistic for IV	12.24	10.23	4.83	4.83
P-value	0.0007	0.0019	0.0299	0.0299

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

Table A16: 2SLS results using predictions from Machine Learning application

	Plot area (HA)	Total maize area (HA)	Quantity of urea (kg/ha)	Quantity of DAP (kg/ha)	Quantity of NPS (kg/Ha)	IHS: Family labor - planting - (days/ha)	IHS: Hired labor - planting - (days/ha)	IHS: Family labor - harvest - (days/ha)	IHS: Hired labor - harvest - (days/ha)	IHS: Total labor (family and hired) (days/ha)	Seedling rate (kg/ha)	Plot Potential Wetness Index †	Plot elevation (m)†	Plot slope (percent)†	Log(Yield (kg/ha))
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Original Sample															
CI	0.036	0.179	141.002*	-38.150	129.442**	0.912	2.160	1.043	0.500*	0.701	-60.754*	-0.537	-0.021	-0.447*	0.039
	(0.036)	(0.140)	(72.546)	(52.868)	(57.174)	(0.885)	(2.148)	(1.006)	(0.260)	(0.499)	(34.475)	(0.625)	(0.043)	(0.272)	(0.482)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	313	313	313	313	313	313	313	313	313	313	313	313	313	313	294
Control mean	0.091	0.154	19.126	13.389	10.607	187.493	21.266	71.235	1.116	281.110	69.233	-0.017	-0.004	-0.008	4,894.340
Panel B: Same EAs															
CI	0.118*	0.276*	199.156***	-16.683	159.621***	0.093	4.559	1.298	0.583***	0.222	-66.974*	-0.739	0.014	0.199	0.321
	(0.063)	(0.165)	(57.316)	(45.802)	(59.285)	(0.948)	(2.998)	(0.823)	(0.223)	(0.899)	(37.221)	(0.743)	(0.017)	(0.141)	(0.601)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	820	820	820	820	820	820	820	820	820	820	820	820	820	820	716
Control mean	0.087	0.174	20.081	9.604	14.299	211.816	22.970	64.204	1.154	300.144	109.301	-0.070	0.004	-0.003	4,281.734
Panel C: Same Regions															
CI	0.173	0.423	164.798**	-36.372	154.267**	-0.325	4.795	0.660	0.708**	-0.381	11.570	-0.637	0.104	0.204	0.218
	(0.110)	(0.281)	(71.766)	(65.528)	(71.376)	(1.141)	(3.426)	(1.313)	(0.326)	(1.291)	(111.988)	(0.705)	(0.089)	(0.159)	(0.669)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	956	956	956	956	956	956	956	956	956	956	956	956	956	956	792
Control mean	0.094	0.197	23.598	10.551	19.240	214.969	25.831	62.067	1.031	303.899	117.438	-0.058	0.005	0.002	4,667.119
Panel D: Full Sample															
CI	0.172	0.418	163.517**	-36.767	154.093**	-0.319	4.776	0.680	0.702**	-0.378	12.178	-0.637	0.104	0.202	0.194
	(0.109)	(0.278)	(71.249)	(65.517)	(71.244)	(1.140)	(3.420)	(1.313)	(0.322)	(1.293)	(111.723)	(0.701)	(0.089)	(0.158)	(0.674)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	1,121	951
Control mean	0.096	0.200	23.568	10.515	19.206	215.423	25.463	62.231	1.022	304.140	116.905	-0.057	0.005	0.002	4,816.879

Notes: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Standard errors (in parenthesis) clustered at EA level. C.I.=Correct identification of improved varieties. IHS: inverse hyperbolic sine transformation. †Plot potential wetness index, plot elevation and plot slope are expressed as deviation from the EA mean. Yield are calculated from self-reported data (only here). Control variables included are: No. years of DSM, Log (Proximity to woreda town (KM)), ACC, ACC*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table A15.