



HAL
open science

You reap what (you think) you sow? Evidence on farmers' behavioral adjustments in the case of correct crop varietal identification

Paola Mallia

► **To cite this version:**

Paola Mallia. You reap what (you think) you sow? Evidence on farmers' behavioral adjustments in the case of correct crop varietal identification. 2022. hal-03597332v2

HAL Id: hal-03597332

<https://pse.hal.science/hal-03597332v2>

Preprint submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2022 – 08

You reap what (you think) you sow? Evidence on farmers' behavioral adjustments in the case of correct crop varietal identification

Paola Mallia

JEL Codes:

Keywords:

anr 
agence nationale
de la recherche
AU SERVICE DE LA SCIENCE



You reap what (you think) you sow?
Evidence on farmers' behavioral adjustments in the case
of correct crop varietal identification

Paola Mallia *

June 30, 2022

[Latest version here](#)

Abstract

Adoption of improved seed varieties has the potential to lead to substantial productivity increases in agriculture. However, only 36 percent of the farmers that grow an improved maize variety report doing so in Ethiopia. This paper provides the first causal evidence of the impact of misperception in improved maize varieties on farmers' production decisions, productivity and profitability. We employ an Instrumental Variable approach that takes advantage of the roll-out of a governmental program that increases transparency in the seed sector. We find that farmers who correctly classify the improved maize variety grown experience large increases in inputs usage (urea, NPS, labor) and yields, but no statistically significant changes in other agricultural practices or profits. Using machine learning techniques, we develop a model of interpolation to predict objectively measured varietal identification from farmers' self-reported data which provides proof-of-concept towards scalable approaches to obtain reliable measures of crop varieties and allows us to extend the analysis to the nationally representative sample.

1 Introduction

It is widely recognized that agriculture in Sub-Saharan Africa has experienced limited productivity improvements in the last 60 years compared to other parts of the world such as Asia and Latin America. Given that the share of employment in agriculture is about 50 percent, this is broadly conceived as a crucial reason why the region has failed to embark on a path of sustained economic growth and mass poverty is still widespread ([Suri and Udry, 2022](#)). Low level of improved technology adoption is often suggested as an explanation for why agricultural productivity has remained stagnant in Sub-Saharan Africa. Despite the

*Paris School of Economics, paola.mallia@psemail.eu

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 am grateful to Karen Macours for her guidance and advice, Sylvie Lambert for her useful comments and Gero Carletto for his constructive critiques. I thank ANR for the support received through the EUR grant ANR-17-EURE-0001. All remaining errors are my own.

availability of seemingly profitable agricultural technologies, the reason why some technologies are not widely adopted still remains an open question (Macours, 2019). Therefore, obtaining a better understanding of the constraints on adoption and optimal allocation of complementary inputs is deemed to be a crucial step to comprehend a major component of growth (de Janvry and Sadoulet, 2020).

The literature on agricultural technology adoption has investigated different potential constraints that farmers may face¹, such as credit and risk market inefficiencies (Karlan et al., 2014); input and output market inefficiencies (Suri, 2011; Shamdasani, 2021; Bold et al., 2017; Ashour et al., 2019; Michelson et al., 2021); behavioural constraints (Duflo et al., 2011); and lack of knowledge about profitability and proper use (Foster and Rosenzweig, 2010). More recent assessments of the state of technology adoption in SSA claim that, on the one hand the use of inorganic fertilizer and agro-chemicals remains relatively low on average, while on the other hand governments' efforts to stimulate modern input use, especially maize, seem to have had some success. Yet, these inputs are rarely used together on plots, despite widespread evidence of agronomic synergies (Sheahan and Barrett, 2017).

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 et al. 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 et al. 2020 report that farmers correctly classify the variety grown on 36 percent of maize plots in Ethiopia². Despite the importance of investments in complementary inputs to obtain benefits for improved varieties (Emerick et al., 2016), little attention has been devoted to investigate the effect of farmers' mistaken beliefs of the variety grown on decision-making processes and behaviors.

This paper provides causal evidence on the effect of farmers' misperception of the maize varieties used on production decisions, productivity and profitability. Consistently with the behavioral economics literature (Tversky and Kahneman, 1973; Kahneman and Tversky, 1984; Angner and Loewenstein, 2007), we show that farmers' act upon their beliefs of the improved traits of the varieties they grow through the allocation of complementary inputs, which in turn affects the levels of yields and profits achieved. We rely on the Ethiopia Socioeconomic Survey collected in 2018/2019 (ESS4) to recover a measure of farmers' correct classification 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). In order to isolate the behavioral adjustment of farmers' beliefs of the variety grown, we restrict the

¹See Foster and Rosenzweig 2010, de Janvry et al. 2017, Magruder 2018 for literature reviews.

²In this paper we rely on the same data as in Kosmowski et al. 2020, which can be found at <https://www.openicpsr.org/openicpsr/project/124681/version/V9/view>.

sample to farmers who grow an improved maize variety according to the objective measures, but who have either correct or wrong perceptions about it (i.e. true positive and false negative). We take advantage of administrative data of the roll-out of 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. We argue that longer exposure to the program, facilitates the creation of an environment where farmers become aware of the variety grown and are able to experiment dynamically the agricultural investments to make. We find that correct classification makes farmers adopt more fertilizer 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. We 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 are quite substantial, accounting for 120 percent increase. Given that both true positives and false negatives cultivate an improved variety, we can attribute the differences in yields to the different production decisions. We do not find statistically significant increases in profits, although the effect is positive and big in size. 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 expenses of the other crops, but instead they are crowding-in additional resources. Finally, applying machine learning techniques we predict the objective measure for varietal identification with 85 percent of accuracy using farmers' self-reported data. This allows us to extend the analysis to the nationally representative sample, thus addressing potential concerns linked to low statistical power.

We contribute to the literature in multiple ways. First, we take a step back on the basic assumption common to all the literature related to agricultural technology adoption: that is, farmers are fully aware of *which* technology they use. Indeed, not all agricultural technologies have traits that are easily distinguishable and 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 correctly identify them. Farmers may have the wrong belief when using or, alternatively, foregoing to use an improved variety. Most of improved varieties perform well relative to traditional varieties even under unfavorable production conditions, but they express their full yield potential only with favorable management (Gollin et al., 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 inputs allocation to achieve the yield gains promised by the technology. Whenever farmers misperceive which technology is being used, an observed

input misallocation might reflect either that farmers are failing to behaviorally adjust to the technology or that farmers are successfully optimizing the objective function but based on wrong beliefs on the true technology. Thus, we contribute to the literature by relaxing this assumption and providing the first causal evidence of the impact of correct classification of improved varieties on a number of complementary production decisions, productivity and profitability. Our findings suggest that farmers are able to obtain higher productivity levels when they have correct beliefs about the technology used. While we refrain from drawing conclusions on the optimality of the new inputs 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.

Second, our instrumental variable, since it is identified by the cumulative length of exposure to a policy intervention that favors transparency in the seed system, provides suggestive evidence of learning about agricultural technologies and complementary inputs allocation. This contribution relates to the literature that recognizes farmers own's learning process among the factors affecting decisions pertaining to technology choice and inputs allocation. In a learning-about-productivity model, beliefs about the new technology evolve over time with experience, thus individuals learn about the overall profitability of a new technology from experience and compare it to the profitability of the existing technology that is assumed to be well established (Besley and Case, 1993; Besley et al., 1994). However, uncertainty about profitability of a new technology is not the only challenge that needs to be overcome by farmers. Learning may also involve acquiring information about how to optimally manage the new technology (Foster and Rosenzweig, 1995). Agricultural research organization and extension agents carry out controlled experiments on a new technology and can thus determine the maximal possible yields and even, for a given set of prices, maximal profitability. What they cannot necessarily do is provide information on how best to achieve these yields given the specific characteristics of farmers' plots (Laajaj et al., 2020). As such, a farmer may have to experiment with a crop on his own land to sort out the correct inputs application. By deciphering the optimal management of a new technology, the profitability of a technology grows over time as knowledge accumulates.

Third, we show the possibility to predict 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 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 are 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.

Finally, we contribute to the wider literature on measurement error in agricultural survey that investigates the extent of nonclassical measurement error in land size, soil quality, yields and crop variety identification (Carletto et al., 2013; Gourlay et al., 2019; Ilukor et al., 2017;

Abay et al., 2019, 2021; Kosmowski et al., 2019; Wossen et al., 2019), by providing empirical evidence that the implications of measurement error can go beyond statistical inference.

The rest of the paper is organized as follows. In Section 2, we present the data used in the analysis and provide an overview of the government’s programs relevant for this analysis that target the Ethiopian seed sector. In Section 3, we specify the empirical strategy employed and in Section 4 we discuss the results. In Section 5, we provide empirical evidence in support of the validity of the exclusion restriction of the instrumental variable approach. Section 6, we develop a machine learning model to predict objectively measured data from farmers’ self-report and employ the predictions to extend the analysis to the full nationally representative sample. Finally, in Section 7 we conclude and acknowledge the limitations of this study.

2 Data and descriptive statistics

2.1 The Ethiopia Socioeconomic Survey

Our 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, we mainly rely on the wave collected in 2018/19 (ESS4) and we take advantage of the previous data collections to provide evidence in support of the validity of the identifying assumptions. The 2018/19 ESS (ESS4) is the first wave of a new panel, not a follow-up to previous ESS waves. It is a baseline survey for the waves to follow and it covers all nine states and two cities, Addis Ababa and Dire Dawa. The ESS4 data collection was conducted in 565 EAs, of which 316 are rural and 219 urban. 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 a 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, harvested quantity and crop variety for maize, barley and sorghum, obtained through GPS measurement, crop-cut and DNA fingerprinting, respectively. The collection of samples for DNA fingerprinting was restricted to the subsample of most growing regions of maize, barley and sorghum: 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 measures³. We combine the data of the DNA fingerprinting sample 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) and funded by the Bill and Melinda Gates Foundation. The total sample of maize plots with DNA fingerprinting measures in ESS4 consist of 479 plots cultivated by 423 households⁴. For the purpose of this analysis, we restrict the sample to the plots where farmers either:

- Correctly identified the variety grown as improved (true positive);
- Grow an improved variety, but wrongly believe to be growing a traditional variety (false negative).

The analysis is thus informed by a sample of 314 plots cultivated by 289 households.

2.2 Correct classification of improved maize varieties

While the DNA fingerprinting measure was collected for barley, sorghum and maize, we focus our analysis only on the latter crop. There are at least two reasons for this choice. On one hand, maize is the most common cereal crop in Ethiopia, cultivated by 9.8 million farmers (CSA, 2019), it is considered crucial for food security and it is adapted to all the agroecologies of the country. On the other hand, while improved varieties of sorghum and barley exhibit low adoption levels, less than 1 and 17.7 percent respectively, maize improved varieties are widely available, with 62.6 percent of households growing one on at least one plot, providing meaningful variation for our analysis.

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

³The 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.

⁴The 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.

variable of interest is 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. The specific focus on farmers' correct classification of varieties with an improved trait and not on landraces is motivated by the fact that they require a specific management. In particular, agronomists highlight the key complementarity between growing yield-enhancing varieties and fertilizers application. Moreover, as already mentioned, tolerance traits to adverse climatic events or pests are becoming more and more relevant and thus, correct knowledge is key to exploit their full potential, but also to correctly complement the crop choice with the relevant climate-smart practices. Thus, we hypothesize that farmers' beliefs of the crop variety grown (improved vis-à-vis traditional) drives the production decisions on complementary inputs allocation and, consequently, based on the true variety planted, the level of productivity achieved.

2.3 Seed supply and distribution in Ethiopia

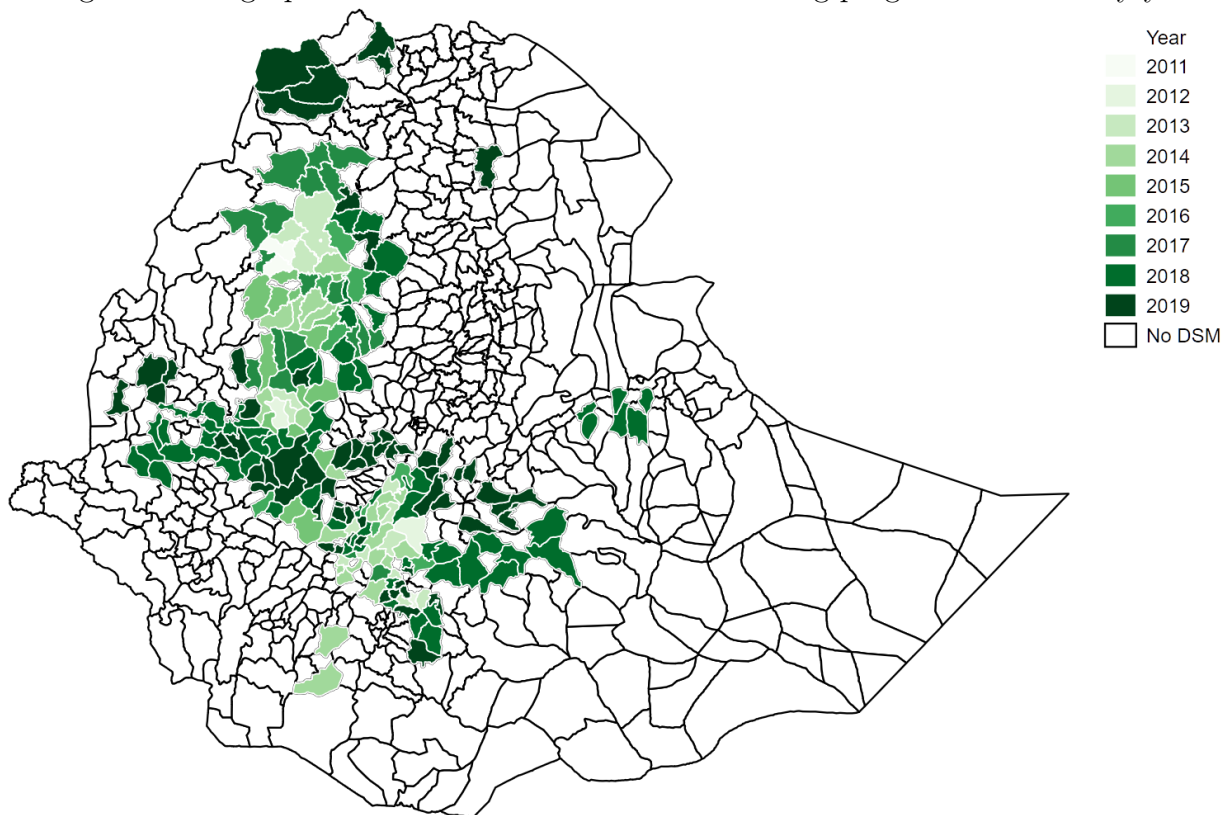
We also rely on administrative data of the roll-out of two governmental programs, namely the Direct Seed Marketing (DSM) and the Input Voucher Sales System (IVS) ⁵. DSM was designed as a mechanism through which both public and private seed producers directly supply adequate amounts of quality seeds to 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 increased access and use of improved seeds of high quality by smallholder farmers (ATA, 2015). 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's⁶ 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 involved in certified seed production, producing mainly hybrid maize seed. The largest private seed company is Pioneer Hi-Bred plc, which primarily produces hybrid maize seed (Alemu et al., 2019; Benson et al., 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 own 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, including hybrid maize. At the same time, a growing number of private seed producers and agricultural cooperatives emerged to supply hybrid maize seed primarily to the regional Bureaus of Agriculture (BoA) for distribution through their supply channels to farmers (Mekonen et al., 2019). While the design did not embed a specific targeting criteria to define the roll-out of the program across the country, it

⁵The data was kindly provided by the Ethiopian Agricultural Transformation Agency (ATA).

⁶Woreda are the third-level of the administrative division of Ethiopia - after zones and the regional states.

emerges from various official sources that regions were chosen based on their seed production capacity and of 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. Currently, DSM coverage has increased to more than 290 woredas. Crops now operational 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 for this study. Our empirical strategy relies on the fact

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



Data source: Agricultural Transformation Agency

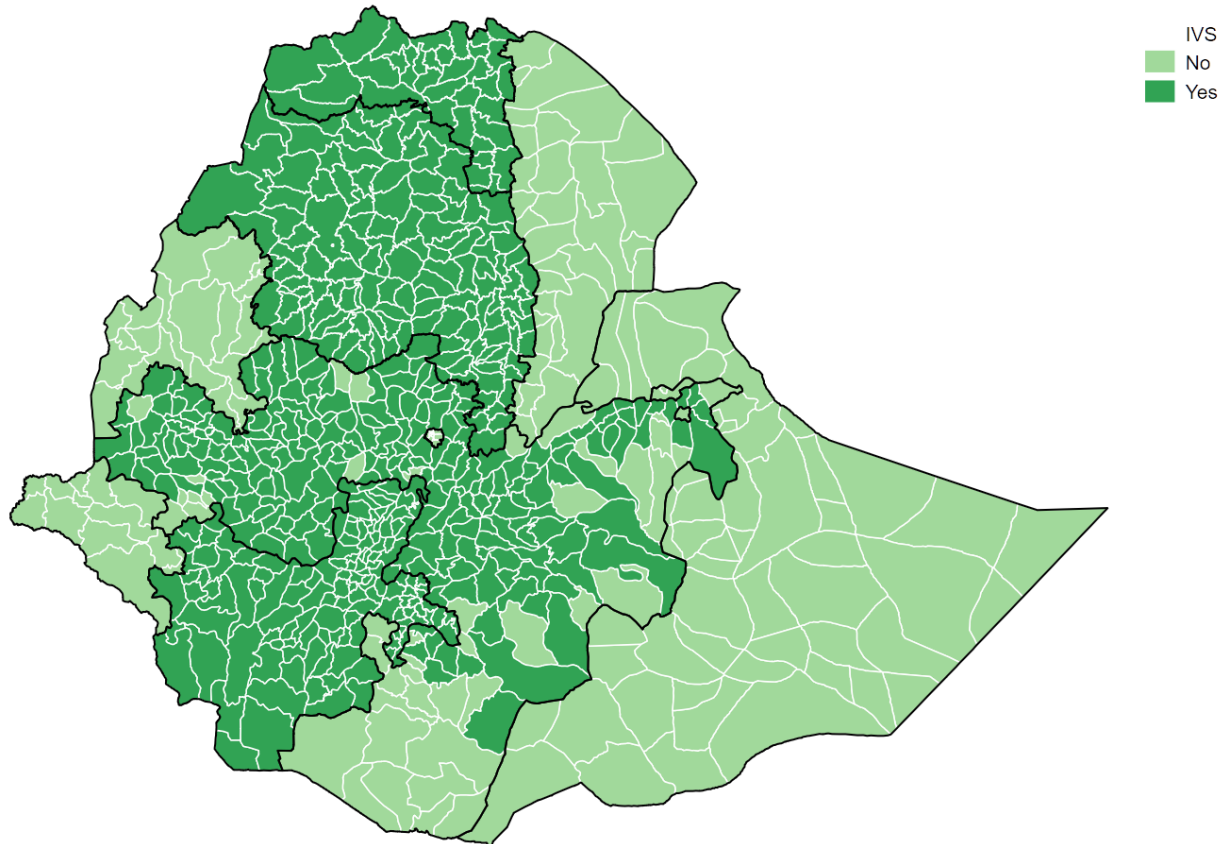
that the number of years since the DSM roll-out in one woreda, should increase both the probability that farmers can access improved maize varieties, while reducing the likelihood of counterfeiting given the fewer steps in the supply chain in comparison to the CSM system. The length of exposure to the program, paired with the proximity of households to the closest woreda town, provides the exogenous variation in farmers' correct classification of improved maize varieties.

The Input Voucher Sales System is an 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 ⁷. 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 IVS 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

⁷<http://www.ata.gov.et/programs/highlighted-deliverables/input-voucher-sales-system-ivs-2/>

different but complementary issues of the Ethiopian agricultural sector, we control for the IVS presence when exploiting the length of exposure to the DSM program⁸.

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



Data source: Agricultural Transformation Agency

2.4 Maize suitability

To relax potential endogeneity concerns regarding the targeting of the roll-out of DSM program, we control for maize suitability through the Land Suitability Index at woreda level. We rely on the data calculated in [Zabel et al. 2014](#) “Global agricultural land resources—a high-resolution suitability evaluation and its perspectives until 2100 under climate change conditions”. The methodology applied is a fuzzy logic approach to compute global agricultural suitability to grow the 16 most important food and energy crops according to the climatic, soil and topographic conditions at a spatial resolution of 30 arc seconds. We use the results of this approach, taking into account rainfed and irrigated conditions, the start of the growing cycles and the number of crop cycles, to control for the woreda’s average land suitability to maize cultivation.

⁸We also performed the analysis controlling for the presence of the Agricultural Commercialization Cluster and obtain very similar results which we report in Appendix [A.3](#).

2.5 Descriptive statistics

Table 1 presents summary statistics of the main plot and household level variables used in the analysis. The first part of the table displays four possible scenarios when comparing farmers' self-reported crop variety grown on the plot with the objectively measured data. Correct classification of improved maize varieties (true positive) only occurs on 22 percent of the cultivated plots, while in 39 percent of the cases, farmers' grow an improved variety but believe to use a traditional variety (false negative). Instead, the opposite is true for 21 percent of the full sample (false positive), and the remaining 18 percent has correct beliefs about growing a traditional variety (true negative). The rest of the table shows summary statistics using the sample of analysis: that is, true positive and false negative plots. The unconditional average length of exposure to DSM is 1.21 years, where the program was implemented on 34.24 percent of the sample of woredas. We define household's proximity to the closest woreda town as the inverse of the self-reported distance in kilometers. In this way, higher values of the variable indicate that households are closer to the urban center in comparison to the rest of the sample, where the average distance amounts to 21 kilometers⁹. The IVS initiative is implemented in almost the totality of the woredas in the sample (99 percent). The plot average suitability to maize cultivation is 20 on a scale from 0 (less suitable) to 100 (highly suitable), where the maximum in the sample is 55. The average plot area under maize, measured using GPS devices, is 0.11 hectares while the average land size per household is 0.13 hectares¹⁰. Regarding fertilizers usage, farmers apply on average 63 kg of urea, 32 kg of DAP and 32 kg of NPS per hectare of land. Labor figures, calculated with the assumption that 1 day equals to 8 hours of work, show that farmers dedicate on average 186 days of family and only 24.5 days of hired labor for each hectare of land during preparation and planting. While on the other hand, labor used at harvest is on average 82 and 6 days per hectare for family and hired labor, respectively. Farmers plant on average 65.5 kg of maize seed per hectare. Regarding plot characteristics, the sample exhibits limited variation from the EA mean in terms of potential wetness, elevation and slope. In terms of agricultural practices, farmers practice on 20 to 40 percent of their plots crop rotation with a legume, crop residue cover, minimum or zero tillage, terracing and ploughing along the plot contour, separately. Soil erosion prevention measures are taken on 77 percent of the plots and 70 percent of farmers employ Soil Water Conservation Practices. Notably, only 2 percent of the plots are irrigated. In order to measure yield, we take advantage of the harvest measure obtained through crop-cut (dry weighted crops) on 4x4 meters sub-plots. Every time that is necessary, we adjust the plot area taking into account the percentage of land dedicated to maize in case of intercropping with other crops. We calculate maize cultivation profits by assigning the monetary value of the harvested quantity using prices collected in the first closest market center. The granularity of the data allows to assign location specific prices by different units of measurement, also non-standard, which captures the geographical variations in conversion factors. While self-reported quantities of crop sales

⁹Proximity of household h to the closest woreda towns is given by: $Proximity_h = \max(distance_h) - distance_h$ where $distance_h$ is the distance of household h to the closest woreda town.

¹⁰This is consistent with what is reported by (Desiere and Jolliffe, 2018).

Table 1: Descriptive statistics of selected variables

	Mean	SE	Min	Max	N	Sample
Farmers' self-reported vs. DNA-fingerprinting						
True positive (TP)	0.22	0.02	0	1	479	Full DNA Fingerprint sample
False negative (FN)	0.39	0.03	0	1	479	Full DNA Fingerprint sample
False positive (FP)	0.21	0.02	0	1	479	Full DNA Fingerprint sample
True negative (TN)	0.18	0.02	0	1	479	Full DNA Fingerprint sample
Control variables						
Correct identification - CGIAR germplasm (True positive)	0.36	0.03	0	1	314	Sample of analysis (TP + FN)
Proximity to woreda town (KM)	68.40	1.07	0	89	314	Sample of analysis (TP + FN)
No. years of DSM	1.21	0.13	0	8	314	Sample of analysis (TP + FN)
IVS	0.99	0.01	0	1	314	Sample of analysis (TP + FN)
Suitability Index	20.89	0.94	2	55	314	Sample of analysis (TP + FN)
Input application						
Plot area in HA (GPS measure)	0.11	0.01	0	1	314	Sample of analysis (TP + FN)
Total household area under maize cultivation (HA)	0.13	0.01	0	1	314	Sample of analysis (TP + FN)
Quantity of UREA per HA	63.32	8.89	0	995	314	Sample of analysis (TP + FN)
Quantity of DAP per HA	32.10	6.27	0	435	314	Sample of analysis (TP + FN)
Quantity of NPS per HA	32.39	5.01	0	562	314	Sample of analysis (TP + FN)
Pre-harvest family labor (days/ha)	186.03	14.09	0	927	314	Sample of analysis (TP + FN)
Pre-harvest hired labor (days/ha)	24.48	4.13	0	239	314	Sample of analysis (TP + FN)
Harvest family labor (days/ha)	82.08	6.35	0	459	299	Sample of analysis (TP + FN)
Harvest hired labor (days/ha)	6.41	1.46	0	87	299	Sample of analysis (TP + FN)
Total labor per plot (Days/Ha)	294.38	18.73	0	1,409	314	Sample of analysis (TP + FN)
Seedling rate (KG/Ha)	65.51	9.99	2	2,096	314	Sample of analysis (TP + FN)
Plot characteristics (deviation from EA average)						
Potential Wetness Index	-0.04	0.08	-5	7	301	Sample of analysis (TP + FN)
Elevation (m)	0.01	0.02	-1	3	301	Sample of analysis (TP + FN)
Slope (percent)	-0.09	0.05	-2	4	301	Sample of analysis (TP + FN)
Agricultural Practices						
Plot is irrigated	0.02	0.01	0	1	314	Sample of analysis (TP + FN)
Plot prevented from soil erosion	0.77	0.03	0	1	314	Sample of analysis (TP + FN)
River dispersion	0.00	0.00	0	1	314	Sample of analysis (TP + FN)
Motor pump used for irrigation	0.01	0.01	0	1	314	Sample of analysis (TP + FN)
Crop rotation with a legume	0.39	0.03	0	1	314	Sample of analysis (TP + FN)
Crop residue cover - visual aid	0.38	0.03	0	1	312	Sample of analysis (TP + FN)
Minimum tillage	0.31	0.03	0	1	314	Sample of analysis (TP + FN)
Zero tillage	0.20	0.03	0	1	314	Sample of analysis (TP + FN)
Soil Water Conservation practices (SWC)	0.70	0.03	0	1	314	Sample of analysis (TP + FN)
Terracing	0.28	0.03	0	1	314	Sample of analysis (TP + FN)
Water catchments	0.08	0.02	0	1	314	Sample of analysis (TP + FN)
Afforestation	0.03	0.01	0	1	314	Sample of analysis (TP + FN)
Plough along the contour	0.31	0.03	0	1	314	Sample of analysis (TP + FN)
Conservation Agriculture (CA)	0.00	0.00	0	1	314	Sample of analysis (TP + FN)
Productivity and Profitability						
Maize yields (Kg/Ha)	2,904.63	303.02	0	2,175,000	314	Sample of analysis (TP + FN)
Household Maize profits (ETB)	2,577.15	328.52	-1,102	1,673,845	289	Sample of analysis (TP + FN)

Notes: Point estimates are weighted sample means.

are available in the data, we prefer to rely on the objectively collected measures of harvest, since we found inconsistencies in the break-down of quantities by different uses probably due to recalling bias or enumerators' typing error. On the side of costs, we take into consideration all the expenses related to the cultivation of maize, that is hired labor and inputs purchase, including transportation costs. As it is common practice, we did not include the costs of items that were not purchased, such as family labor or quantity of recycled seed, due to data limitations. Maize yields are on average almost 2.9 tons per hectare while household's profits from maize cultivation are around 2577 ETB per year (around 88 USD¹¹).

3 Empirical strategy

Improved seed varieties are a key input in agriculture and exhibit strong complementarities with other inputs allocation, such as fertilizer, labor and, agricultural practices. Moreover, they are characterized by latent benefits: the signal of the characteristic trait may not always manifest itself or become stronger only after germination. On one hand, yield-enhancing varieties may fail to deliver higher yields one year due to incorrect complementary inputs management. On the other, the resistance trait to negative climate related events of a variety will be evident only in case of the occurrence of this scenario. Moreover, as shown by the literature, knowledge of agricultural technologies accumulates through continuous experimentation in: i) learning how to recognize the improved variety and, ii) the optimal inputs allocation. For this reason, observing farmers correctly identifying the variety in the current period is likely correlated with past correct classification. Therefore, inputs allocation in the current period should not only be determined by today's current classification, but also and foremost by the knowledge accumulated through longer experimentation with the improved variety. For this reason, we focus on farmers who correctly classify the improved maize variety vis-à-vis farmers who grow improved varieties, but believe to cultivate a traditional variety. This allows us to test if farmers who correctly classify the improved maize variety in the year of observation, behave differently in comparison to those who wrongly believe to grow an improved variety. By identifying correct classification based on past exposure to an intervention that increases the likelihood of correct identification in previous years, we aim to provide evidence of the behavioral adjustment deriving from farmers' learning through experimentation.

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

$$Y_{phw} = \beta_0 + \beta_1 CI_{phw} + \beta_2 P_{hw} + \beta_3 DSM_w + \beta_4 IVS_w + \beta_5 IVS_w * P_{hw} + \beta_6 Suitability_w + \beta_7 Suitability_w * P_{hw} + \alpha_r + \epsilon_{phw} \quad (1)$$

Where Y_{phw} is the input allocation observed on plot p of household h in woreda w , CI_{phw} is improved varieties correct identification 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

¹¹According to the average exchange rate in 2019.

h residing in woreda w , DSM_w is the number of years of exposure to the DSM program of woreda w , IVS_w is the presence of the IVS system in woreda w , $Suitability_w$ is a measure of maize suitability in woreda w and, α_r are regional fixed effects. Both IVS 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 Enumeration Area (EA) level which corresponds to the first level of sampling.

As mentioned, the idea we are trying to model is that true positives, i.e. farmers that are correctly aware of growing improved maize varieties are also better optimizing inputs allocation in comparison to false negatives. Thus, we define the variable of interest, CI , as taking value one when the farmer's self-report of adoption of maize improved varied is confirmed by the DNA fingerprinting measure on the plot, while zero if farmers report to grow a traditional variety but the objective measures identifies it as improved. Our main hypothesis is that farmers growing improved varieties should allocate inputs differently in comparison to those who believe to grow landraces, but cultivate an improved variety. Indeed, those who are rightfully aware in the season we observe were also more likely to have had correct perceptions in the past, and hence had a better chance to learn about optimal complementary inputs allocation and thus, act on their correct beliefs. In this setting, we will focus on a broad definition of inputs which includes land area, fertilizers application, labor activities, but also field characteristics, that is elevation, slope and potential wetness of the plot. Since farmers in our sample cultivate on average 3 maize plots, it is of interest to consider plot selection based on its characteristics as one dimension among which farmers can optimize their objective function.

The correct identification of adoption status is potentially endogenous since farmers who misperceive adopting an improved variety are likely to be different from those who do not in both observed and unobserved characteristics that affect inputs allocation. Thus, the ordinary least squares (OLS) estimates of equation (1) would be biased. To solve the endogeneity issue, we implement an Instrumental Variable (IV) approach. Hence, we instrument correct identification implementing the following first-stage:

$$CI_{phw} = \theta_0 + \theta_1 P_{hw} + \theta_2 DSM_w + \theta_3 P_{hw} * DSM_w + \theta_4 IVS_w + \theta_5 IVS_w * P_{hw} + \theta_6 Suitability_w + \theta_7 Suitability_w * P_{hw} + \alpha_r + \eta_{phw} \quad (2)$$

Where $P_{hw} * DSM_w$ is the interaction between exposure to DSM and proximity to woreda town and represents our instrumental variable, while the other variables are defined as above. As already mentioned, the DSM program allows both private and public seed suppliers to market seeds directly to farmers. The competition between different suppliers calls for transparency in the characteristics of the varieties sold in order to gain market shares among farmers. Based on this, we could expect that longer exposure to the program, should facilitate the creation of an environment where farmers become aware of the variety grown and are able to experiment dynamically the agricultural investments to make. The length of exposure to 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 their characteristics, which should ultimately increase awareness in inputs allocation. At the same time, farmers that 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 too costly to travel to urban centers and thus, end up relying in informal channels that are not subject to quality controls and more likely to engage in counterfeiting than officially recognized suppliers. While both channels, proximity to woreda town and exposure to DSM are expected to enhance correct classification of improved varieties alone, the direction of the interaction between the two is not obvious and ultimately an empirical question. We hypothesize that since the exposure to the DSM program is meant to address potential market failures in the seed supply chain, the value of developing trustworthy relations should be stronger when households reside in remote rural areas which are harder to connect and therefore more vulnerable to the inefficiencies of the seed sector. On the other hand, the closer the households is to key centers of seed supply sales, the less we expect the effect of exposure to DSM to play a role.

From an identification point of view, given the non-random targeting of the DSM program, we cannot rely on the 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 (as the DSM), which affect inputs allocation. Therefore, we exploit the interaction between exposure to DSM and proximity to woreda towns to identify the exogenous variation that affects inputs allocation only through correct classification, while controlling for other confounding factors. For this reasons, in both specifications we include a measure of maize suitability and its interaction with proximity to woreda towns¹². In addition, we also control for the presence of other major agricultural programs in the woreda promoted by the Agricultural Transformation Agency (ATA) and likely to have followed similar targeting criteria than DSM. In particular, we focus on the IVS described in Section 2.3. Using administrative data on the geographical coverage of the program allows us to take into account any difference across woredas derived from IVS's activities in terms of input use through access to credit.

Controlling for some the targeting criteria of DSM, that is IVS and maize suitability, conditioning on proximity to woreda town, relaxes the concerns of the direct effect of the instrumental variable to inputs allocation, lending confidence to the validity of our identifying assumptions. We discuss in greater detail the validity of the exclusion restriction in Section 5.

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

4 Results

4.1 First-stage results

Table 2 shows the estimates of the first-stage specification in Equation 2. The coefficient of the interaction term between proximity to a woreda town and years of exposure to DSM is negative. This result indicates that, as farmers are more exposed to transparent supply channels, such as the ones provided by DSM, the less important is the effect of being closer to a urban center, and vice versa.

Table 2: 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)	-0.268*** (0.069)
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: 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).

The coefficient of the instrumental variable is large and statistically significant at 1 percent level of confidence. Our results show that our specification explains around 27 percent of the variation in farmers correct classification of maize varieties. The size of the F-statistic on the instrument is 14.97 which suggests that our approach should not suffer from the concern of weak instrumental variable.

As we show in Table A1 (Appendix), the results are robust even when considering the correct classification of a different definition of improved variety, which indicates that the transparency effect that we are modeling is not specific to the CGIAR-derived germplasm definition of improved. Moreover, Table A2 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 the objective measure (column (2)). Instead, the instrument is still relevant in predicting correct identification of both improved and traditional varieties, thus corroborating the transparency channel argument.

4.2 Inputs use and plot selection

Table 3 reports the OLS and Two-Stage-Least-Square (2SLS) estimates of Equation 1, 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 correct classification is 13 kg per hectare¹³. Both DAP and NPS are blended fertilizers based on nitrogen and phosphorus, but the latter also contains sulphur. With the introduction of NPS in Ethiopia, farmers have progressively shifted from DAP to NPS (Balemi et al., 2019), which is consistent with our results.

Despite the increase in fertilizers use, we do not find any statistically significant effect on the measures of labor at planting, although the coefficients are positive and big in size. However, we find a positive and statically 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 a 85 percent increase with respect to the (control) average. As in Beaman et al. 2013, 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, our results are in line with the complementarity between improved varieties and fertilizer use predicated by agronomists and, cultivating an yield-enhancing variety translates into a greater amount of work at harvest.

Moreover, we 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.

¹³We report the effects on the extensive margin for urea and NPS use in Table A3. The large impact reflects important changes at the extensive margin, with an increase of 18 percentage point and 17 percentage point, starting from 15 for urea and 9 for NPS.

Table 3: Plot level OLS and 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 Wetness Index †	Plot elevation (m) †	Plot slope (percent) †
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: OLS													
CI	0.039* (0.020)	125.407*** (25.857)	36.410* (19.684)	60.220*** (17.961)	-0.122 (0.253)	-0.259 (0.554)	-0.334 (0.275)	-0.166 (0.335)	-0.174 (0.202)	-37.988** (18.268)	-0.099 (0.208)	-0.018 (0.029)	-0.080 (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 (0.036)	164.226* (95.205)	-40.425 (60.068)	129.675* (68.165)	1.015 (1.108)	2.068 (2.369)	-0.200 (0.859)	0.851* (0.496)	0.883 (0.775)	-102.240* (54.949)	-0.125 (0.674)	-0.177* (0.106)	-0.071 (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. 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)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table 2.

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. We do not find any statistically significant effect neither on plot area or on total area under maize cultivation from correct classification, although the coefficients are 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 accomplish different things and thus they 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 mechanisms at play discussed for the results of the First-Stage.

By comparing the coefficients reported in Panel A and Panel B of Table 3, we find that the IV coefficients of the different outcomes of interest are larger than the OLS estimates most of the time. The difference 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 bias¹⁴. We 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 A4 show that in the vast majority of the cases, most of the gap can be attributed to the endogeneity bias from omitted variable. Indeed we can expect 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 incur in misperception, given that they only adopt traditional varieties. The interaction between these two effects is consistent with 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 et al., 2022](#)) or assume selection on observable characteristics ([Bohr et al., 2022](#)), our results suggest that ignoring the endogeneity bias due to unobservables would lead to an underestimation of the true effect of correct identification on inputs application¹⁵.

¹⁴Given that the treatment is binary, the difference attributable to weights on treatment levels is mechanically zero.

¹⁵Similar considerations apply to the results on agricultural productivity reported in Section 4.5

4.3 Cultivation practices

Departing from inputs allocation, another interesting hypothesis to investigate relates to whether 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 on increasing farmers income and reducing production costs (Teklewold et al., 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 4 shows no impact on correct identification of improved varieties on the adoption of any of the practices under analysis. 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 neither when using an aggregate index of all practices. However, the estimates of some of the practices 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.

Table 4: Plot level OLS and 2SLS results on agricultural practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	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
Panel A: OLS															
CI	0.006 (0.011)	0.200** (0.092)	0.007 (0.006)	-0.001 (0.009)	0.178 (0.111)	-0.092 (0.111)	-0.198** (0.100)	-0.017 (0.068)	0.198* (0.110)	0.223*** (0.084)	0.010 (0.040)	-0.006 (0.026)	-0.030 (0.104)	-0.003 (0.003)	-0.364** (0.177)
Region FE	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 (0.011)	0.071 (0.310)	0.003 (0.003)	0.003 (0.011)	0.341 (0.347)	-0.458 (0.323)	-0.177 (0.181)	0.210 (0.180)	-0.219 (0.489)	0.058 (0.246)	0.090 (0.225)	-0.125 (0.080)	-0.243 (0.248)	-0.017 (0.018)	0.014 (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. 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 2.

4.4 Crowd-in vs. reallocation of resources from other plots

After having shown that farmers allocate more inputs on those plots where the improved variety grown is correctly identified, a natural next question to investigate is whether the additional amount of resources comes from reallocation from other households' plots under crops other than maize. In order to do so, we estimate the 2SLS at the household level - instead of plot level as before - where we aggregate the outcome measures across all farmers' plots. The correct identification variable will take value one if farmers correctly identify the improved variety grown on at least one of their maize cultivated plots. Thus, we compare the results of these estimations of the aggregate values of inputs use on all maize plots cultivated versus inputs allocation on all the plots under other crops but 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 we are not able to establish correct classification for each of the household's plot. Thus, in order to conduct this part of the analysis at the household level, we need to make the assumption 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. In addition, it should be noted that in some cases the random procedure selected more than one plot for DNA fingerprinting for each household and for this reason when we collapse at household level the number of observations at our disposal decreases in comparison to the plot level analysis. Results are shown in Table 5: Panel A shows the first-stage results at household level. Our instrumental variable has a slightly stronger negative effect on correct classification than the plot level analysis and it 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 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 effects on urea application are also positive though no longer significant statistically significant at the household level. The fact that we find consistent results when conducting the analysis at household level in comparison to the plot level lends confidence to the validity of the assumption made above regarding the likelihood that farmers plant the same maize seed on all their plots. In Panel C, we look at what happens on other crops cultivations when the correct identification condition on maize plots is met. As we see in column (2), farmers who correctly classify use a higher quantity of urea even on other plots, although the increase is smaller than the one observed on maize plots. We also see any effect on NPS, but there is not a statistically significant increase in labor use, neither at planting or at harvest. The overall picture provided by these results seems to suggest that in households where farmers correctly identify the improved variety grown on at least one of their maize plots, are not systematically reallocating resources to maize cultivation at the expenses of the other crops. Instead, the findings point towards the

Table 5: First-Stage and 2SLS results of crowding-in vs. resource reallocation at household level

C.I.	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	-0.298***								
* Log(Proximity to woreda town -KM)									
F-statistic	(0.065)								
P-value	21.030								
	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
Obs.	289	289	289	289	289	289	289	289	289
Control mean	-	0.092	21.138	13.444	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
Obs.	289	289	289	289	289	289	289	289	289
Control mean	-	0.886	21.744	15.131	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. IHS= inverse hyperbolic sine transformation. 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).

direction of households crowding-in additional resources. Since effects are identified through the number of years of exposure to DSM, this is in line with a learning mechanism where farmers dynamically learn how to optimize inputs allocation, starting from maize cultivation where correct classification happens, and transfer the management practices learned also to other cultivations. Another potential explanation for the results observed in Panel C is linked to the exposure to DSM. Although our 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, we 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.

4.5 Yield and profits

The results discussed so far indicate that correctly classifying the variety grown on one plot as improved causes several changes in inputs 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 inputs allocation changes translate into higher productivity and profitability levels. Table 6 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 results highlights that utilizing more inputs as a result of correct beliefs, leads to substantial productivity gains.

Table 6: OLS and 2SLS results of correct identification on productivity and profitability

	Log(Yield (kg/ha))	Profits (ETB)
	(1)	(2)
Panel A: OLS		
CI	0.310** (0.141)	1,186.322 (924.930)
Panel B: 2SLS		
CI	1.225*** (0.367)	1,801.139 (1,195.526)
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. IHS= inverse hyperbolic sine transformation. 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 2 for column(1) and Table 5 for column(2).

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 fertilizers use. Although, we do not detect any statistically significant effect of correct classification on household's profits from maize cultivation, the coefficient is positive and big in magnitude, as shown in column (2). Suggestively, we could argue that there is a clear path towards higher profit gains, but we are not able to detect it due to the noisy measure.

5 Exclusion restriction

The exclusion restriction implied by our instrumental variable approach is that, conditional on the controls included in the regression, the interaction effect between the duration of exposure to the DSM program and households' proximity to woreda towns has no effect on production decisions, other than their effect through farmers' correct identification of the improved variety grown on their plot. Hence, the obvious potential threat to the identifying assumption is if DSM was implemented earlier in some woredas rather than in others for possible unobserved reasons that may affect production decisions differently based on proximity to woreda towns. For example, if DSM was rolled-out earlier in places where demand for inputs was increasing for those close to woreda towns, but not for those far away. In that case, the length of exposure to DSM would be directly associated to our outcome variables independently of the channel of correct classification of the seed variety. Given the targeting criteria of the roll-out of the program discussed in Section 2.3, we can exclude a random assignment of DSM to woredas. Nevertheless, we can still provide empirical evidence in support of the exclusion restriction. In order to do so, we employ two strategies.

First, we perform a placebo test running the reduced form regression using the waves collected prior to ESS4: namely, ESS1 in 2011/2012, ESS2 in 2013/2014 and ESS3 in 2015/2016. The reduced-form regression at household-level is the following:

$$Y_{hw} = \beta_0 + \beta_1 P_{hw} * DSM_w^{2019} + \beta_2 P_{hw} + \beta_3 DSM_w^{2019} + \beta_4 IVS_w + \beta_5 IVS_w * P_{hw} + \beta_6 Suitability_w + \beta_7 Suitability_w * P_{hw} + \alpha_r + \epsilon_{phw} \quad (3)$$

Where the variables are defined as in equation (1) and (2), with DSM_w^{2019} being the length of exposure to DSM in the year 2019 for woreda w .

All the waves share almost identical questionnaires, with the only exception of quantity applied of NPS and seeds which were not collected in ESS1 and ESS2. Moreover, although the quantity of harvested crop was meant to be collected in all the waves, in practice, most crop/fields do not have self-reported harvest information in ESS1¹⁶. In addition, the crop-cut measure, on which we heavily rely for the main results on productivity and profitability, is only available in ESS4 and for a very limited subsample in the previous waves. The measurement error that affects the self-reported harvested quantity, while not the crop-cut measure, makes the two items less comparable. For this reason we cannot use yields or profits neither for this placebo test or the next strategy performed.

¹⁶Source: ESS1 Basic Information Document.

Table 7: Reduced-form regression at household level using ESS1, ESS2, ESS3 and ESS4 of household unbalanced panel sample of maize growing areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
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: Family labor planting (days/ha)	IHS: Family labor planting (days/ha)	IHS: Family labor planting (days/ha)	IHS: Family labor planting (days/ha)	IHS: Family labor planting (days/ha)	IHS: Family labor planting (days/ha)
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
Region FE	(0.045)	(44.038)	(47.186)	(45.332)	(0.197)	(0.405)	(0.237)	(0.099)	(0.186)	(261.864)
Obs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R2	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152
Mean	0.077	0.001	-0.002	-0.001	0.022	0.078	0.082	0.054	0.033	-0.005
	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
Region FE	(0.031)	(12.645)	(9.370)	(4.703)	(27.326)	(0.071)	(0.144)	(0.097)	(0.125)	(99.932)
Obs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R2	1,555	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552
Mean	0.215	0.091	0.146	0.105	0.063	0.024	0.061	0.103	0.074	-0.005
	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.
Region FE	(0.019)	(8.043)	(9.233)	n.a.	(16.236)	(0.066)	(0.129)	(0.078)	(0.103)	n.a.
Obs.	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
Adjusted R2	1,506	1,502	1,502	n.a.	1,502	1,502	1,503	1,502	1,503	n.a.
Mean	0.226	0.135	0.197	n.a.	0.036	0.011	0.039	0.067	0.075	n.a.
	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.
Region FE	(0.071)	(10.962)	(19.914)	n.a.	(0.196)	(0.135)	(0.204)	(0.054)	(0.227)	n.a.
Obs.	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
Adjusted R2	1,383	1,383	1,383	n.a.	1,383	1,383	1,383	1,383	1,383	n.a.
Mean	0.027	0.042	-0.006	n.a.	0.036	0.085	0.063	0.043	0.045	n.a.
	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)), IVS, IVS*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.

For the purpose of comparison, we restrict the samples of the different waves to the maize cultivated plots in the regions where the DNA fingerprint sample was collected for ESS4, i.e. Amhara, Dire Dawa, Harar, Oromia, SNNP, Tigray. The estimates from reduced-form regression should be proportional to the casual effect of interest, that is the effect of correct classification on inputs allocation identified through our instrumental variable. Obtaining similar results in ESS1, ESS2, ESS3 compared to ESS4 would signal that our instrumental variable conditional on the covariates included is capturing baseline differences between close by and farther localities in woredas that ended up with greater exposure to DSM in 2019 and those with limited or zero exposure, thus invalidating the exclusion restriction. Table 7 shows the estimates of the placebo test¹⁷. Reassuringly, the instrumental variable, that is the interaction between years of exposure to DSM in 2019 and proximity to woreda towns, does not predict fertilizer use and hired labor in ESS1, ESS2 and ESS3, while it does in ESS4. Specifically, while we observe some significant results for the waves prior to ESS4, they do not concern the variables of interest (urea, NPS and hired labor at harvest) or have opposite sign. Across the different specifications, the estimates of the ESS3 sample are the closest to the ESS4. This is because there are only 2 years of difference in terms of exposure to DSM based on proximity to woreda town between the two waves and thus, it is likely that the learning mechanism starts to play a role¹⁸. Results are consistent when using the household balanced panel sample, which we report in the Appendix (Table A6)¹⁹ showing that the evidence provided is not driven by sample selection across waves. Finally, we test whether we can attribute any differential trends in inputs use based on the instrumental variable. For this purpose, we look at potential differential trends in outcome variables for the panel of households in woredas targeted after 2016 and those who have never been targeted by DSM. We take advantage of the panel component of the three waves prior to ESS4 to estimate the following specification at the household level:

$$\Delta Y_{t,t-1} = \beta_0 + \beta_1 P_{hw} * DSM_{ESS4} + \epsilon \quad (4)$$

Where $\Delta Y_{t,t-1}$ represents the difference in input use between ESS2 and ESS1, ESS3 and ESS2 and ESS3 and ESS1 and the independent variable is number of years of DSM exposure in ESS4, interacted with proximity to woreda town. The results reported in Table 8 show that pre-trends in inputs allocation are not correlated with the instrumental variable, with the only exception of family labor at harvest. This provides further evidence in support of the exclusion restriction.

¹⁷To ensure that results are comparable across waves, we use the sample of all households growing maize in the regions where DNA fingerprinting was collected. This is a larger sample for ESS4 in comparison to the one used in the main analysis, since for the purpose of this exercise we do not need to limit to the observations where DNA fingerprinting was collected. We report the estimates of the reduced form regression for the sample of DNA fingerprinting of ESS4 in Table A5 (Appendix).

¹⁸We acknowledge the fact that the results in Panel A are weaker than the results in Table A5 due to the sample composition. Indeed, the larger sample size comes at the "expenses" of including EAs in areas that are less important for maize production and thus not sampled for DNA fingerprinting.

¹⁹The panel dimension mentioned refers to ESS1, ESS2 and ESS3, but not ESS4.

Table 8: 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.

6 Application on the nationally representative sample

The analysis conducted so far suffers from the limitation 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, we apply machine learning (ML) techniques to predict the objective measure of improved maize variety. In this way, we can extend our 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 we 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 observations²⁰ where both the DNA fingerprinting measure and the self-reported measure of improved maize variety are available (henceforth the training set), we try to predict the former measure on the remaining sample of 1490 plots under maize cultivation (the test set).

As recommended by the literature, the first step consists of selecting the best performing model. In order to do so, we 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, we choose the best performing model and predict the variable of interest on the test set (Athey and Imbens, 2019). Before performing this procedure, we exclude from the original training set, 100 observations that we use to estimate

²⁰To train the algorithm we take advantage of the full sample available which also includes false positive and true negative. Moreover, we are able to recover 12 additional observations which were excluded due to missing sampling weights.

the accuracy of our preferred model. That is, we 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, we are able to assess the percentage of correctly predicted data points using the models of choice. Table A7 provides an overview of the different samples used in the procedure just described.

We take advantage of the SuperLearner package available for R, which provides an easy and efficient way to perform the training and the prediction processes using a wide library of machine learning algorithms. We feed the training procedure 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 and rainfall variables based on GPS plot location;
- Soil fertility measures based from GIS data;
- Farmer’s agricultural practices adopted on the plot.

We pay particular attention to 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, since we replicate the same specifications using the predicted sample.

For each of the machine learning algorithms, we 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, we rely on the Xg.boost, Ipredbagg and Random Forest learners ²¹. Table 9 reports the average MSE based on 5-fold cross-validation. The best performing individual model is Random Forest, with a risk coefficient (MSE) of 0.1336. However, as shown also in Figure ??, SuperLearner selects as 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 combinations 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, we are able 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 the 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 relying on

²¹Xg.boost, Random Forest and Ipredbagg are decision-tree-based ensemble machine learning algorithms that perform well in classification problems.

Table 9: Machine learning algorithms performance by Mean Square Error

Algorithm	MSE			
	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.

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 we discuss the results of the first-stage and 2SLS estimations of correct identification, constructed using the predicted DNA fingerprinting measure on inputs usage. To assess the validity of the predictions and to disentangle the potential source of error, we replicate the analysis using the *predictions* of: i) the original sample used so far ²²; 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, we replicate the analysis on the sample true positive (correct identification) and false negative cases.

Table 10 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 2. All the coefficients in the regression are of the same sign and magnitude of 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 we are comforted by the results just discussed, they do not provide an exhaustive picture of the model’s performance. This is because we are using the predictions obtained on the same sample that was used to train the algorithms. Thus, in column (2)-(4), we 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

Table 10: 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*** (0.082)	-0.262*** (0.080)	-0.175** (0.086)	-0.175** (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)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

²²In this Section, we 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.

relevance of the IV in the different contexts. In column (2), we 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 column (3) and (4), respectively. The values of the F-statistic for IV are 4.13, which suggest that the instrument is now less relevant than it was in the original sample. There are at least two potential motivations for why this is the case. The woredas and regions included in column (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.

In Table 11, following the same logic, we report the results of the 2SLS. As shown in Panel A, the results obtained from the predictions are almost identical to the original sample discussed in the previous sections (Table 3). 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 correct identified varieties on quantity of urea are 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 is no longer statistically significant and positive in sign. However, since we are by design adding woredas (and regions) where maize is much less important and where only 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 our 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.

Table 11: 2SLS results using predictions from Machine Learning application

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(14)
	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: Hiring labor planting (days/ha)	IHS: Family labor harvest (days/ha)	IHS: Hiring labor harvest (days/ha)	IHS: Total family labor hired (days/ha)	Seedling rate (kg/ha)	Plot potential Wetness Index †	Plot elevation (m) †	Plot slope (percent) †	Log(Yield (kg/ha))
Panel A: Original Sample															
CI	0.042 (0.045)	0.208 (0.198)	159.803* (87.281)	-37.999 (63.174)	131.267* (67.794)	0.942 (1.082)	1.952 (2.271)	1.037 (1.286)	0.485* (0.279)	0.694 (0.670)	-65.001* (36.371)	-0.646 (0.776)	-0.022 (0.050)	-0.423 (0.270)	0.194 (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 (0.069)	0.261 (0.196)	207.553*** (62.909)	-6.678 (50.902)	154.086** (62.140)	-0.180 (0.919)	4.232 (3.191)	0.834 (0.965)	0.515** (0.236)	-0.039 (0.857)	-76.458* (40.396)	-0.814 (0.895)	0.006 (0.021)	0.210 (0.158)	0.458 (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 (0.128)	0.461 (0.368)	209.000** (83.243)	-30.837 (79.881)	195.449* (100.478)	-0.508 (1.153)	5.199 (3.558)	0.287 (1.495)	0.606* (0.340)	-0.531 (1.260)	30.457 (148.029)	-0.706 (0.894)	0.113 (0.117)	0.247 (0.200)	0.962 (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 (0.128)	0.461 (0.368)	209.666** (83.323)	-30.846 (79.875)	195.115* (100.494)	-0.485 (1.150)	5.201 (3.561)	0.331 (1.491)	0.608* (0.341)	-0.511 (1.258)	29.963 (148.184)	-0.707 (0.893)	0.114 (0.117)	0.247 (0.200)	0.963 (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. 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)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM). First-stage results reported in Table 10.

7 Concluding remarks

The analysis conducted in this paper provides causal evidence of the impact of correct classification of improved maize varieties on inputs allocation, agricultural practices, yields and profits. We find that farmers who correctly classify the variety grown as improved use more fertilizers and hired labor at harvest, while we do not find any detectable changes in agricultural practices in comparison to farmers who have wrong beliefs of the variety planted. We observe large productivity increases when there is correct identification, however we are not able to detect the potential gains in profitability. Given that we identify the effect of correct classification through the cumulative length of exposure to a program that favors transparency in the seed sector, we 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 we can draw from this analysis. First, misperception in which agricultural technology is used has testable implications for farmers' production decisions. We argue that 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 on 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, we contribute 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 collecting and using farmers' self-reported measures in specific contexts. We argue that objective measures in agricultural surveys should not be considered the panacea to all observed inefficiencies. Instead, as we show here, self-reported data offer a window to farmers' perceptions that can provide insights to the inner causes of observed behaviors.

The quasi-experimental nature of the empirical strategy suffers from several limitations to date. 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 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. We argue that the interest in improved seed varieties and optimal inputs 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 fa-

vor 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 et al., 2019). For this reasons, 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 pest 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 take 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 pursue of climate adaptation.

References

- Abay, K. A., Abate, G. T., Barrett, C. B., and Bernard, T. (2019). Correlated non-classical measurement errors, ‘second best’ policy inference, and the inverse size-productivity relationship in agriculture. *Journal of Development Economics*, 139:171–184.
- Abay, K. A., Bevis, L. E., and Barrett, C. B. (2021). Measurement error mechanisms matter: Agricultural intensification with farmer misperceptions and misreporting. *American Journal of Agricultural Economics*, 103:498–522.
- Alemu, D., Rashid, S., Tripp, R., Alemu, D., Rashid, S., and Tripp, R. (2019). Seed system potential in ethiopia: constraints and opportunities for enhancing the seed sector. *Gates Open Research 2019 3:948*, 3:948.
- Angner, E. and Loewenstein, G. (2007). Behavioral economics. *Handbook of the philosophy of science: Philosophy of economic*, pages 641–690.
- Ashour, M., Gilligan, D. O., Hoel, J. B., and 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., and Karlan, D. (2009). 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.
- ATA (2015). Seed system development strategy: Vision, systemic challenges, and prioritized interventions.
- Athey, S. and 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., Gurm, G., and Getaneh, F. (2019). 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.
- Beaman, L., Karlan, D., Thuysbaert, B., and Udry, C. (2013). Profitability of fertilizer: Experimental evidence from female rice farmers in mali. *American Economic Review*, 103:381–86.
- Benson, T., Spielman, D., and 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. and 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. Technical report.

- Bohr, N., Deisemann, T., and Gollin, D. (2022). Seeds and false beliefs: seed misidentification, input allocation, and agricultural productivity in ethiopia. *Mimeo*.
- Bold, T., Kaizzi, K. C., Svensson, J., and Yanagizawa-Drott, D. (2017). Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in uganda. *The Quarterly Journal of Economics*, 132:1055–1100.
- Burke, M., Bergquist, L. F., and Miguel, E. (2019). Sell low and buy high: Arbitrage and local price effects in kenyan markets. *The Quarterly Journal of Economics*, 134:785–842.
- Carletto, C., Savastano, S., and Zezza, A. (2013). Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. *Journal of Development Economics*, 103:254–261.
- Cole, S., Giné, X., and Vickery, J. (2017). How does risk management influence production decisions? evidence from a field experiment. *The Review of Financial Studies*, 30:1935–1970.
- Cole, S. A. and Fernando, A. N. (2021). ‘mobile’izing agricultural advice technology adoption diffusion and sustainability. *The Economic Journal*, 131:192–219.
- CSA (2019). Sample survey area and production of major crops.
- de Janvry, A. and Sadoulet, E. (2020). Using agriculture for development: Supply- and demand-side approaches. *World Development*, 133:105003.
- de Janvry, A., Sadoulet, E., and Suri, T. (2017). Field experiments in developing country agriculture. In *Handbook of economic field experiments*, volume 2, pages 427–466. Elsevier.
- Desiere, S. and Jolliffe, D. (2018). Land productivity and plot size: Is measurement error driving the inverse relationship? *Journal of Development Economics*, 130:84–98.
- Dufló, E., Kremer, M., and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review*, 101:2350–90.
- Emerick, K., de Janvry, A., Sadoulet, E., and Dar, M. H. (2016). Technological innovations, downside risk, and the modernization of agriculture. *American Economic Review*, 106(6):1537–61.
- Evenson, R. E. and Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science*, 300:758–762.
- Fink, G., Jack, B. K., and Masiye, F. (2020). Seasonal liquidity, rural labor markets, and agricultural production. *American Economic Review*, 110:3351–92.
- Floro IV, V. O., Labarta, R. A., Becerra López-Lavalle, L. A., Martínez, J. M., and 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. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, 103:1176–1209.
- Foster, A. D. and Rosenzweig, M. R. (2010). Microeconomics of technology adoption. *Annual Reviews Economics*, 2:395–424.
- Gollin, D., Hansen, C. W., and Wingender, A. M. (2021). Two blades of grass: The impact of the green revolution. *Journal of Political Economy*, 129:2344–2384.
- Gollin, D., Morris, M., and Byerlee, D. (2005). Technology adoption in intensive post-green revolution systems. *American Journal of Agricultural Economics*, 87(5):1310–1316.
- Gourlay, S., Kilic, T., and Lobell, D. B. (2019). 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.
- Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129:1311–1353.
- Ilukor, J., Kilic, T., Stevenson, J., Gourlay, S., Kosmowski, F., Kilian, A., Sserumaga, J., and 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). Empirical Decomposition of the IV–OLS Gap with Heterogeneous and Nonlinear Effects. *The Review of Economics and Statistics*, pages 1–45.
- Kahneman, D. and Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39:341–350.
- Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129:597–652.
- Kosmowski, F., Alemu, S., Mallia, P., Stevenson, J., and Macours, K. (2020). Shining a brighter light: Comprehensive evidence on adoption and diffusion of cgiar-related innovations in ethiopia. *SPIA Synthesis Report*.
- Kosmowski, F., Aragaw, A., Kilian, A., Ambel, A., Ilukor, J., Yigezu, B., and Stevenson, J. (2019). 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.
- Laajaj, R., Macours, K., Masso, C., Thuita, M., and Vanlauwe, B. (2020). Reconciling yield gains in agronomic trials with returns under african smallholder conditions. *Scientific Reports 2020 10:1*, 10:1–15.

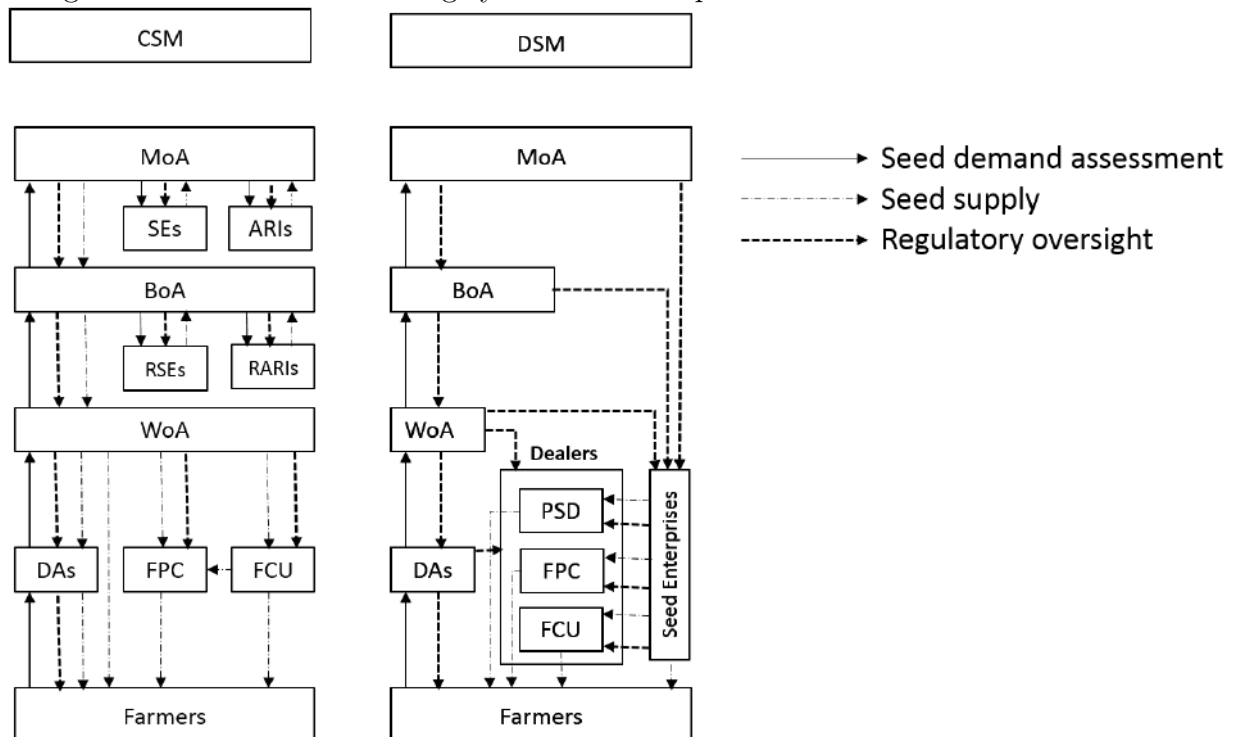
- Lemessa, S. D., Watebaji, M. D., and Yismaw, M. A. (2019). Climate change adaptation strategies in response to food insecurity: The paradox of improved potato varieties adoption in eastern ethiopia. *Cogent Food Agriculture*, 5.
- Macours, K. (2019). Farmers' demand and the traits and diffusion of agricultural innovations in developing countries. *Annual Review of Resource Economics*, 11:483–499.
- Magruder, J. R. (2018). An assessment of experimental evidence on agricultural technology adoption in developing countries. *Annual Review of Resource Economics*, 10:299–316.
- Maredia, M., Reyes, B., Manu-Aduening, J., Dankyi, A., Hamazakaza, P., Muimui, K., Rabbi, I., Kulako, P., Parkes, E., Abdoulaye, T., et al. (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., and Abate, G. T. (2019). *Performance of direct seed marketing pilot program in Ethiopia: Lessons for scaling-up*, volume 132. Intl Food Policy Res Inst.
- Michelson, H., Fairbairn, A., Ellison, B., Maertens, A., and Manyong, V. (2021). Misperceived quality: Fertilizer in tanzania. *Journal of Development Economics*, 148:102579.
- Mobarak, A. M. and 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., Ringler, C., Msangi, S., Palazzo, A., Batka, M., et al. (2009). *Climate change: Impact on agriculture and costs of adaptation*, volume 21. Intl Food Policy Res Inst.
- Shamdasani, Y. (2021). Rural road infrastructure agricultural production: Evidence from india. *Journal of Development Economics*, 152:102686.
- Sheahan, M. and Barrett, C. B. (2017). Ten striking facts about agricultural input use in sub-saharan africa. *Food Policy*, 67:12–25.
- Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica*, 79:159–209.
- Suri, T. and Udry, C. (2022). Agricultural technology in africa. *Journal of Economic Perspectives*, 36(1):33–56.
- Teklewold, H., Kassie, M., Shiferaw, B., and Köhlin, G. (2013). 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.
- Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5:207–232.

- Wossen, T., Abay, K. A., and Abdoulaye, T. (2022). Misperceiving and misreporting input quality: Implications for input use and productivity. *Journal of Development Economics*, 157:102869.
- Wossen, T., Abdoulaye, T., Alene, A., Nguimkeu, P., Feleke, S., Rabbi, I. Y., Haile, M. G., and Manyong, V. (2019). Estimating the productivity impacts of technology adoption in the presence of misclassification. *American Journal of Agricultural Economics*, 101:1–16.
- Zabel, F., Putzenlechner, B., and 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.

A Appendix

A.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.

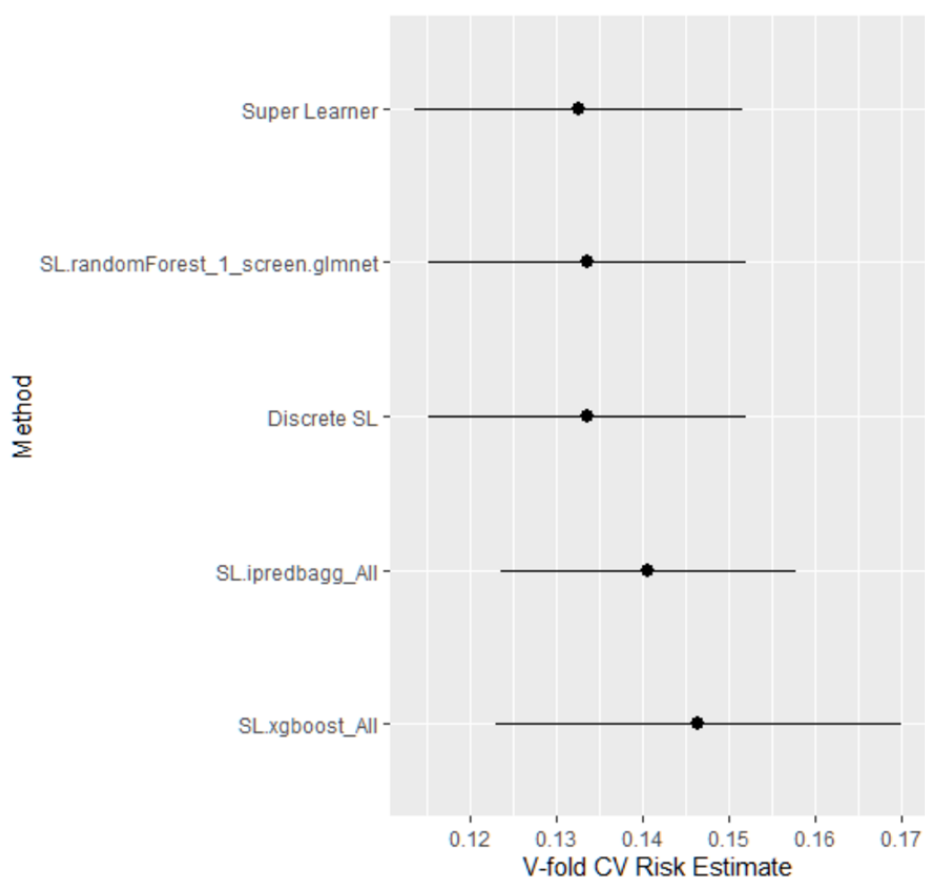


Figure A2: 5-fold CV Risk Estimate

A.2 Tables

Table A1: First-Stage results using correct identification of improved varieties - Exotic germplasm

	Correct identification of improved maize variety - Exotic
	(1)
No. years of DSM * Log(Proximity to woreda town -KM)	-0.231*** (0.056)
Region FE	Yes
Obs.	379
Adjusted R2	0.285
F-statistic	16.790
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)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town KM).

Table A2: 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 (0.093)	0.068 (0.098)	-0.286*** (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. 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).

Table A3: Bivariate probit model estimates of the extensive margin of input use

	Urea	NPS
	(1)	(2)
C.I.	0.182*** (0.032)	0.171*** (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. 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) are not included to facilitate model convergence.

Table A4: IV-OLS gap decomposition

Variables	Coefficients				Decomposition		
	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	-76.835	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	
(s.e.)	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	
(s.e.)	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	
(s.e.)	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	
(s.e.)	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	
(s.e.)	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)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM)

Table A5: Reduced-form regression results using ESS4- DNA fingerprinting sample restricted to true positive and false negative.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(14)	(15)
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)	Plot Wetness Index †	Plot elevation (m) †	Plot slope (percent) †	Log(Yield (kg/ha))	Profits (BIRR)			
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*
Region FE	(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)
Number of observations	Yes 314	Yes 314	Yes 314	Yes 314	Yes 314	Yes 314	Yes 314	Yes 299	Yes 299	Yes 314	Yes 314	Yes 301	Yes 301	Yes 301	Yes 312	Yes 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)), IVS, IVS*Log(Proximity to woreda town(KM)), Suitability Index, Suitability Index*Log(Proximity to woreda town -KM).

Table A6: Reduced-form regression at household level using ESS1, ESS2, ESS3 of household balanced panel sample of maize growing areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total maize area (HA)	Quantity of area (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)	To- tal labor - (fam- ily and hired) (days/ha)	Seedling rate (kg/ha)
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
Region FE	(0.035)	(14.368)	(10.940)	(5.629)	(0.167)	(0.091)	(0.182)	(0.127)	(0.158)	(188.702)
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.
Region FE	(0.023)	(8.542)	(9.704)	n.a.	(0.125)	(0.089)	(0.152)	(0.099)	(0.109)	n.a.
Obs.	937	937	937	937	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.
Region FE	(0.055)	(12.856)	(19.068)	n.a.	(0.177)	(0.132)	(0.176)	(0.064)	(0.188)	n.a.
Obs.	946	946	946	946	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)), IVS, IVS*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 A7: Datasets and observations used for machine learning

Datasets		Obs
Train	short	391
	validation	100
Test		1460

A.3 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 A8: 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 A9: 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)
CI	0.016 (0.035)	149.253* (77.231)	-37.662 (50.371)	127.636** (53.725)	0.942 (0.903)	2.253 (2.258)	-0.149 (0.723)	0.864* (0.465)	0.857 (0.570)	-99.627* (53.653)	-0.024 (0.532)	-0.177* (0.100)	-0.125 (0.212)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	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

Panel A: 2SLS

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. †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 A8.

Table A10: Plot level 2SLS results on agricultural practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Irrigation	Soil ero- sion pre- vention	River dis- persion	Motor pump	Rotation with Legume	Crop residue cover	Minimum tillage	Zero tillage	SWC	Terracing	Water Catch- ment	Afforestation along plot con- tour	CA	Factor analysis: all prac- tices	
CI	0.004 (0.010)	0.033 (0.264)	0.003 (0.003)	0.001 (0.010)	0.310 (0.344)	-0.407 (0.329)	-0.180 (0.172)	0.195 (0.159)	-0.244 (0.450)	0.028 (0.247)	0.089 (0.224)	-0.121* (0.068)	-0.240 (0.211)	-0.016 (0.017)	0.016 (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. 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 A8.

Table A11: 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)	Hired labor harvest - (days/ha)	Seedling rate (kg/ha)
	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: First-stage household level									
No. years of DSM	-0.308***								
* Log(Proximity to woreda town -KM)	(0.064)								
F-statistic	23.450								
P-value	0.000								
Panel B: 2SLS plots under maize cultivation									
CI	0.007 (0.036)	135.415* (77.103)	-19.258 (38.392)	111.369** (43.548)	0.544 (0.780)	2.384 (2.158)	0.740** (0.337)	0.880 (0.940)	-93.385* (48.866)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	289	289	289	289	289	289	289	289	289
Control mean	-	0.092	21.138	13.444	286.649	36.235	95.598	9.407	79.795
Panel C: 2SLS plots under all other crops except maize									
CI	-0.478 (0.381)	72.383* (37.575)	-19.419 (14.737)	82.489** (34.176)	0.294 (0.625)	1.672 (1.944)	0.345 (0.318)	0.457 (0.539)	39.985 (25.685)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	289	289	289	289	289	289	289	289	289
Control mean	-	0.886	21.744	15.131	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. 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 2.

Table A12: 2SLS results of correct identification on productivity and profitability

	Log(Yield (kg/ha))	Profits (BIRR)
	(1)	(2)
C.I.	1.167*** (0.360)	1,696.598* (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. 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 A8 for column(1) and Table A11 for column(2).

Table A13: Reduced-form regression results using ESS4- DNA fingerprinting sample restricted to true positive and false negative.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Plot area (HA)	-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***
Total maize area (HA)	(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)
Quantity of urea (kg/ha)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quantity of DAP (kg/ha)	314	314	314	314	314	314	314	299	299	314	314	301	301	301	312
Quantity of NPS (kg/ha)	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	0.301
IHS: Family labor - planting - (days/ha)															
IHS: Family labor - harvest - (days/ha)															
IHS: Family labor - harvest - (days/ha) hired															
IHS: Total labor (days/ha)															
Seeding rate (kg/ha)															
Plot elevation (m)†															
Plot elevation slope (percent)†															
Plot potential Wetness Index†															
IHS: Yield (kg/ha)															
Profits (BIRR)															
No. years of DSM Log(Proximity to town -KM)	*														
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	314	314	299	299	314	314	301	301	301	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	0.301
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	0.274

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 A14: Reduced-form regression at household level using ESS1, ESS2, ESS3 and ESS4 of household unbalanced panel sample of maize growing areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
	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)	To- tal labor rate (kg/ha)	Seedling rate (kg/ha)
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
	Region FE	(0.042)	(34.308)	(34.140)	(58.271)	(0.211)	(0.569)	(0.250)	(0.100)	(0.215)	(269.043)
	Obs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Adjusted R2	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152	1,152
	Mean	0.097	0.001	-0.002	-0.003	0.032	0.081	0.146	0.049	0.058	-0.006
		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
	Region FE	(0.047)	(18.754)	(13.547)	(5.358)	(28.274)	(0.117)	(0.236)	(0.139)	(0.190)	(121.699)
	Obs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Adjusted R2	1,555	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552	1,552
	Mean	0.238	0.135	0.149	0.141	0.063	0.024	0.053	0.103	0.064	-0.005
		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.
	Region FE	(0.021)	(10.658)	(12.164)	n.a.	(22.432)	(0.085)	(0.171)	(0.077)	(0.157)	n.a.
	Obs.	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Adjusted R2	1,506	1,502	1,502	n.a.	1,502	1,502	1,503	1,502	1,503	n.a.
	Mean	0.234	0.159	0.201	n.a.	0.047	0.018	0.036	0.071	0.083	n.a.
		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.
	Region FE	(0.115)	(20.734)	(42.908)	n.a.	(0.402)	(0.244)	(0.409)	(0.081)	(0.390)	n.a.
	Obs.	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	n.a.
	Adjusted R2	1,383	1,383	1,383	n.a.	1,383	1,383	1,383	1,383	1,383	n.a.
	Mean	0.033	0.039	-0.005	n.a.	0.043	0.088	0.057	0.054	0.059	n.a.
		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: Reduced-form regression at household level using ESS1, ESS2, ESS3 of household balanced panel sample of maize growing areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	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 (fam- ily and hired) (days/ha)	To- Seedling rate (kg/ha)
ESS3										
No. years of DSM	-0.031	-20.773	-26.634	16.950**	0.137	-0.065	-0.049	-0.073	0.158	-262.707
woreda town -KM)										
Region FE	(0.063)	(20.217)	(17.561)	(7.865)	(0.274)	(0.178)	(0.263)	(0.196)	(0.241)	(193.024)
Obs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
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	0.006	-11.363	-11.285	n.a.	-0.264	-0.036	-0.263	-0.038	-0.307*	n.a.
woreda town -KM)										
Region FE	(0.027)	(10.859)	(12.202)	n.a.	(0.170)	(0.128)	(0.198)	(0.110)	(0.162)	n.a.
Obs.	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	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	-0.158*	3.117	23.243	n.a.	0.354	-0.406*	0.222	-0.097	0.363	n.a.
woreda town -KM)										
Region FE	(0.094)	(27.778)	(47.654)	n.a.	(0.415)	(0.242)	(0.384)	(0.101)	(0.401)	n.a.
Obs.	Yes	Yes	Yes	n.a.	Yes	Yes	Yes	Yes	Yes	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 A16: 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*** (0.081)	-0.266*** (0.083)	-0.192** (0.087)	-0.192** (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 A17: 2SLS results using predictions from Machine Learning application

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	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))
Panel A: Original Sample															
CI	0.036 (0.036)	0.179 (0.140)	141.002* (72.546)	-38.150 (52.868)	129.442** (57.174)	0.912 (0.885)	2.160 (2.148)	1.043 (1.006)	0.500* (0.260)	0.701 (0.499)	-60.754* (34.475)	-0.537 (0.625)	-0.021 (0.043)	-0.447* (0.272)	0.039 (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* (0.063)	0.276* (0.165)	199.156*** (57.316)	-16.683 (45.802)	159.621*** (59.285)	0.093 (0.948)	4.559 (2.998)	1.298 (0.823)	0.583*** (0.223)	0.222 (0.899)	-66.974* (37.221)	-0.739 (0.743)	0.014 (0.017)	0.199 (0.141)	0.321 (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 (0.110)	0.423 (0.281)	164.798** (71.766)	-36.372 (65.528)	154.267** (71.376)	-0.325 (1.141)	4.795 (3.426)	0.660 (1.313)	0.708** (0.326)	-0.381 (1.291)	11.570 (111.988)	-0.637 (0.705)	0.104 (0.089)	0.204 (0.159)	0.218 (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 (0.109)	0.418 (0.278)	163.517** (71.249)	-36.767 (65.517)	154.093** (71.244)	-0.319 (1.140)	4.776 (3.420)	0.680 (1.313)	0.702** (0.322)	-0.378 (1.293)	12.178 (111.723)	-0.637 (0.701)	0.104 (0.089)	0.202 (0.158)	0.194 (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. 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 A16.