The Seeds of Misallocation: Fertilizer Use and Maize Varietal Misidentification in Ethiopia

Nils Bohr, Tim Deisemann, Douglas Gollin, Frederic Kosmowski, and Travis J. Lybbert

Abstract
Optimal input allocation in agriculture leverages production complementarities. For example, improved seeds are generally more responsive to fertilizer than traditional seeds. Thus, inaccurate beliefs about whether seeds sown are improved may result in sub-optimal fertilizer application. We document precisely this pattern using data from Ethiopia that includes farmer beliefs about their maize seeds and genotyping tests that identify the true genetics of these seeds. We find that 15 percent of farmers believe incorrectly that they are using improved varieties and use far more fertilizer than farmers who correctly believe that they sowed traditional varieties. Conversely, we find that about 15 percent of farmers believe incorrectly that they are growing traditional material and use far less fertilizer than those farmers who correctly believe that they are growing improved material. We extrapolate from our nationally representative sample to estimate the national-level magnitude of fertilizer misallocation due to incorrect seed beliefs.

Keywords: Crop varieties, agricultural technology, fertilizer, misallocation, African agriculture

JEL Classification Number: O1; O33; O38; Q1; Q16

Contact: Tim Deisemann, 5 Bank Street, London E14 4BG, UK. Email: tim.deisemann@ebrd.com.

Nils Bohr is an economics graduate from Oxford University; Tim Deisemann is an Associate Economist at the EBRD; Douglas Gollin is a Professor of Economics at Tufts University; Frederic Kosmowski is a Researcher at the CGIAR Standing Panel on Impact Assessment (SPIA); Travis J. Lybbert is a Professor of Agricultural and Resource Economics at University of California, Davis.

*We thank anonymous reviewers, Andrew Foster, and seminar audiences at the Center for the Study of African Economies at Oxford University, University of California-Davis and ICABR-Buenos Aires for helpful comments and feedback.

Working Paper No. 287 **Prepared in January 2024**

1 Introduction

The Green Revolution was sparked by innovations in plant breeding and fueled by complementarities between improved seeds and other inputs. Improved crop genetics were strongly complementary with inputs of chemical fertilizers and irrigation; in many contexts, they also complemented agricultural labor. Realizing the potential yield gains of better crop germplasm requires careful adjustment of other applied inputs (Foster and Rosenzweig, 1995). Hybrid maize, for example, can produce dramatically higher yields than traditional varieties when optimally fertilized, but deviations from the optimum can rapidly reduce the economic returns (Duflo, Kremer, and Robinson, 2008).

In the presence of such complementarities, farmers must know what seeds they have sown in order to know how best to manage their crops. Inaccurate information or beliefs about the seed type can impose direct or indirect costs on farmers via suboptimal input applications and lower returns. Most obviously, farmers who erroneously believe that they are growing improved seeds may purchase costly inputs in anticipation of high returns that are eventually unrealized. Conversely, farmers who incorrectly believe that they are growing unimproved (traditional) crop varieties may not purchase and apply inputs that could increase on-farm profits. In both cases, farmers' incorrect beliefs about the genetics of their seeds will lead to static inefficiencies in the application of inputs. These losses relative to optimal input use, which are likely asymmetric, are depicted in a conceptual model in the Appendix. Such input distortions are potentially important for fertilizer applied to maize, the focus of this study: while increased nitrogen use will, to a point, increase yield of both traditional and improved maize varieties, on the margin improved varieties tend to be more nitrogen-responsive. Beyond static losses, incorrect beliefs may also impose dynamic costs by undermining farmers' ability to learn about and adopt profitable technologies. Farmers may conclude that improved seeds and fertilizer represent poor investments, for instance, if they incorrectly believe that they have been using both but have experienced low or negative returns (Bold et al., 2017).

We draw on new data from Ethiopia in which we are able to compare farmer-reported seed types to the true genetic identity of their seed.[1] Ethiopia provides an interesting context for this study because improved maize seeds and chemical fertilizers have diffused widely across the country since the 1990s, in response to government programs, extension encouragement and specific fertilizer recommendations (Spielman, Mekonnen, and Alemu, 2013; Abate et al., 2015; Kosmowski et al., 2020). Official recommendations specifically encourage maize farmers to use nitrogen fertilizers and implicitly reflect input complementarities by recommending more nitrogen for hybrid (improved) maize than for non-hybrid maize (Abate et al., 2015). However, the Ethiopian context is also characterized by limited seed system regulation as well as farmer seed exchange and seed saving, which can introduce substantial on-farm uncertainty about the

[1] For a detailed discussion of this dataset and the novel insights it enables related to technology adoption see Kosmowski et al. (2020).

genetic makeup of the seeds farmers sow.

We find that (i) nearly one-third of Ethiopian maize farmers hold false beliefs; (ii) those who falsely believe they sowed *improved* seed double or triple nitrogen fertilizer application, adding an average of $84 (20 days wages) or more to their per hectare input costs; and (iii) those who falsely believe that they are growing *traditional* seeds use half or less the level of fertilizer chosen by farmers who correctly understand that their seeds are improved. The overall implication is that significant quantities of fertilizer are likely to be inefficiently allocated, with policy implications for seed systems and management of agricultural input supply chains. We conduct a scaling exercise to project fertilizer overuse and underuse at the national level; the results suggest that around 20,000-30,000 mt of nitrogen would be allocated differently if maize farmers had correct beliefs about their seeds, an amount corresponding to roughly 4-6% of Ethiopia's total nitrogen use.

We join previous researchers in highlighting that mistaken seed beliefs also pose a problem for researchers as a troubling source of non-classical measurement error (Abay, 2020; Abay, Bevis, and Barrett, 2021; Abay et al., 2022, 2019). Out of necessity, the agricultural technology adoption literature has historically relied exclusively on farmer self-reports about improved seed use and on the (often problematic) assumption that new varieties are readily distinguishable from older ones (Macours, 2019). Genotyping advances now offer a compelling alternative. Samples taken from farmers' fields can be DNA fingerprinted as a means of objective seed varietal identification (Stevenson, Macours, and Gollin, 2018; Beegle, Carletto, and Himelein, 2012). This measurement innovation has enabled a number of empirical studies that document how varietal types (e.g., improved or not) and names are commonly misperceived by farmers, introducing measurement error in data based on self-reports. We build on this emerging literature, which finds – across a variety of contexts and crops – that substantial fractions of farmers hold false seed beliefs (Wineman et al., 2020; Maredia et al., 2016; Floro IV et al., 2018; Yirga et al., 2016). Wossen, Abay, and Abdoulaye (2022) provide evidence that seed misclassification is due to misperception (i.e., false beliefs) rather than intentionally misleading misreporting, which we also assume throughout our analysis.

The link we test between seed beliefs and fertilizer use hinges on the fact that improved seeds are generally more responsive to chemical fertilizer than seeds of lower genetic quality (Ellis, 1992; Tolessa et al., 2001; Nyangena, Juma et al., 2014; Kassie et al., 2015).[2] Farmers are well aware of the complementarity between improved seed and chemical fertilizer, and official fertilizer recommendations are typically higher for improved varieties.[3] In our context, Abay et al. (2018) show that improved maize seeds and fertilizer are strong production comple-

[2]This differential responsiveness for improved germplasm is often intentionally part of the breeding process, but it may also arise as an artefact of the plant breeding process, since breeders often select improved varieties based on their performance under growing conditions characterized by high levels of input use.

[3]In Ethiopia, fertilizer recommendations for hybrid maize are 20% higher than for other varieties (see Abate et al., 2015).

ments and that Ethiopian farmers understand this interaction and manage their maize production accordingly.[4]

Our paper builds on two related sets of published studies. The first set consists of a pair of "double blind" studies in Tanzania (Bulte et al., 2014, 2023), in which researchers distributed improved and traditional seeds to farmers. Only a subset of farmers were told which seeds they had received. All other farmers only knew that there was a 50/50 chance the seeds were improved. Bulte et al. (2014) conduct this study with cowpea seeds and find that farmers who know for certain or who know there is a 50/50 chance they are sowing improved seeds are more attentive, exert greater effort and, consequently, produce high yields – even if they were in fact growing a traditional cowpea variety. Farmers who knew they were sowing traditional seeds produced significantly lower yields. In the follow-on study, Bulte et al. (2023) conducted a similar study with maize seeds, improved varieties of which tend to exhibit much greater productivity gains and stronger complementarities with other applied inputs (e.g., fertilizer) than cowpea. In the face of uncertainty about the type of seeds they are sowing, farmers reduce labor investments and thereby produce lower yields. Improved maize seeds are, however, superior enough relative to traditional seeds that yields were higher with improved seeds despite these reduced labor investments, albeit lower than when farmers knew they were sowing improved seeds and could optimize their labor allocation accordingly.

In contrast to these two studies, we leverage the natural prevalence of false seed beliefs among farmers, rather than artificially manipulating these beliefs. Because we can measure with precision the extent of false beliefs in a nationally-representative sample of farmers, we can show that this is a quantitatively important problem. Moreover, while the lack of pure experimental variation in information complicates causal identification, we are able to observe the choices that farmers make in real-world settings, using actual beliefs and studying the on-farm decisions of farmers, strengthening the external validity of our study.

The second set of studies includes two papers that are closely related to ours but that address different contexts where the issues raised are of lower salience or where data constraints limit the external validity of the analysis. Wossen, Abay, and Abdoulaye (2022) use data from Nigerian cassava farmers to demonstrate how farmer misperceptions about cassava varietal quality distort on-farm applications of fertilizer and herbicide and thereby likely introduce inefficiencies. We use a similar approach to test for distorted input applications in the context of maize in Ethiopia. The change of case and context is important. Maize is a more input-intensive crop than cassava; in many contexts in sub-Saharan Africa, fertilizer use on cassava tends to be much lower than for maize.[5] In our context, farmers are likely to be aware of the differential returns

[4] The inherent complexities of agricultural production introduce a host of other factors that affect fertilizer responsiveness, including soils and rainfall (see Burke, Jayne, and Snapp, 2022; Roobroeck et al., 2021), that are beyond the scope of this paper.

[5] As one recent review article put it, "In general, fertiliser use on roots and tuber crops in Sub-Saharan Africa is negligible" (Ezui et al., 2016). However, the Cassava Monitoring Survey on which Wossen, Abay, and Abdoulaye

that are expected to fertilizer applications on improved and traditional maize seed.

Our paper is also related to work by Wineman et al. (2020), who use a simple comparison of mean input usage between farmers with correct and incorrect beliefs about the maize seeds they sowed, in the context of Tanzania. Farmers who correctly report growing an improved variety use more fertilizer and enjoy higher yields compared to those who do not. Because this survey was constructed specifically to focus on narrow questions of varietal identification and input use, the structure of the data is limited, particularly with respect to questions about demographics and community variables. It is also unclear how well the sample represents the national population.

In contrast, our data allow us to consider the same issues in the context of a nationally-representative sample and a broad household survey that includes a richer set of household and community characteristics than those used in previous studies. This is important, because false belief about seed types is not randomly assigned, outside the experimental settings of Bulte et al. (2014) and Bulte et al. (2023). However, our rich data set and methods allow us to control for a broader set of observables than either (Wossen, Abay, and Abdoulaye, 2022) or (Wineman et al., 2020). This allows us to address selection on observables, including the use of Post-Double Selection (PDS) LASSO to optimize our choice of controls. The structure of our data also allows us then to extrapolate the results of our analysis to project national level estimates of varietal misidentification and input use.

Our analysis does not allow us to identify *why* farmers may hold false seed beliefs. Some of the related literature has suggested widespread counterfeiting or adulteration of inputs at some point in the supply chain (e.g., Bold et al., 2017). This is certainly possible. However, other explanations are also plausible. Farmers may simply not be aware of the characteristics of improved varieties (Kosmowski et al., 2019; Maredia et al., 2016). In our context, farmer-saved maize seeds tend to lose their genetic purity over time (Ilukor et al., 2017). This is particularly true of hybrid seeds, but to some degree also for non-hybrid improved maize known as open-pollinated varieties (OPVs). Farmers who purchase seed in one season and save the seed for replanting in following years will very quickly end up with seeds that are genetically unlike the original improved variety, but they may continue to view the seeds as improved. Farmers may also have purchased or been given seed that they understood, incorrectly, to be improved. Somewhat more surprising are those farmers – reasonably numerous in our data – who are growing improved seed without realizing it. These farmers may be growing OPVs, which maintain their genetic purity fairly well over time. They may assume that since they have not recently purchased seed, the genetic quality is unimproved. In this paper, we cannot explain the reasons for misidentification, although we can identify a number of correlates and covariates, as discussed below.

(2022) draw focuses on a purposively constructed sampling frame in Nigeria. In this sample, fertilizer use on cassava is both widespread and sizeable.

2 Data

We use data from the fourth wave of the Ethiopian Socioeconomic Survey (ESS4) to investigate the effects of seed misclassification on input allocation. This survey uses a two-stage sample that is nationally representative.[6] The ESS collects household data related to agricultural production and includes detailed questions at the plot, household, and community levels.In addition to eliciting detailed reports of fertilizer usage, the 2018/19 round selected a sub-sample of maize-producing households in Ethiopia's major maize-growing regions. For each of these households, usually one maize plot was selected for further analysis. The survey team visited the plot at harvest time and conducted a crop cut to measure objectively the maize yield. This produced a total sample of 506 fields, randomly selected at the level of the enumeration area (EA).[7] Maize samples from these crop-cuts were then genotyped to reveal the genetic identity of the maize varieties, making the ESS4 the first nationally-representative household survey in the public domain that incorporates DNA fingerprinting for varietal identification (Kosmowski et al., 2020).

Although ESS4 includes DNA fingerprinting data for other crops, we restrict our focus to maize for several reasons. Maize is now the most commonly grown crop by smallholder farmers in Ethiopia, and maize area has seen a sharp increase over the past 20 years (Stevenson, Macours, and Gollin, 2018). Furthermore, the crop also has the highest reported adoption rate of improved seeds in Ethiopia (Mekonen et al., 2019). Evidence from observational studies (Abay et al., 2018) is consistent with on-station experiments conducted by the national research system in finding that improved maize varieties (especially, maize hybrids) are more input-responsive than improved varieties of other crops.

Defining improved varieties in a genetic sense is not straightforward. We define a collected sample to be "improved" if it matched 95% of selected genetic markers associated with the originally-released breeder seeds in the reference library.[8] This 95% purity threshold guarantees seed uniformity and genetic proximity to the originally-released cultivar. However, we provide robustness tests to show that our main results hold qualitatively for other thresholds in the 70% to 97.5% range.[9]

We distinguish four different seed belief types, according to two binary variables. The first indicates whether the farmer believes he sowed improved or traditional seeds based on his self-report in the post-planting round of the survey. The second indicates whether the ge-

[6]For details about the sampling frame, see https://microdata.worldbank.org/index.php/catalog/3823/download/49208 (Accessed 9 August 2023).

[7]A lack of technical consistency between the different survey modules (with identifiers either missing or not matching) and missing or unrecoverable values for individual observations reduces the final sample to 432 observations when merging across these dimensions and preprocessing the data.

[8]See appendix for a detailed discussion.

[9]The 70% threshold is a low match and tends to treat as "improved" some material that may not perform very differently from unimproved material.

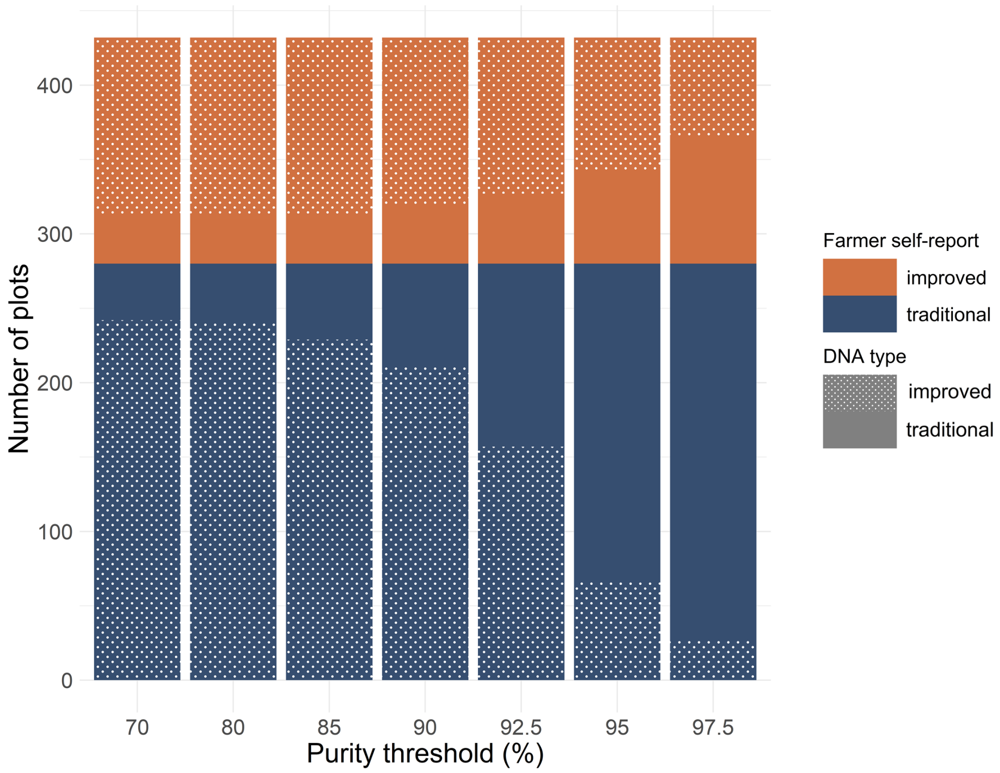

Figure 1: Distribution of seed beliefs for improved maize varieties at the plot-level based on farmer self-reports in survey and DNA fingerprinting results for different genetic purity thresholds for distinguishing improved from traditional seeds. Color indicates the farmer belief (orange = improved, blue = traditional), and pattern denotes the DNA type (dots = improved, none = traditional). Numerical results are reported in Table A1

netic fingerprinting test for the post-harvest sample revealed that it is improved ("DNA Type"), based on our chosen purity threshold. When the two variables align, the farmer has correct seed beliefs. Farmers who correctly report sowing improved seeds have "True Positive" (TP) beliefs, while farmers who correctly report sowing traditional seeds have "True Negative" (TN) beliefs. When the two variables are not aligned, the farmer has false seed beliefs. Farmers who incorrectly report sowing improved seeds have "False Positive" (FP) beliefs, while those who incorrectly report sowing traditional seeds have "False Negative" (FN) beliefs. While farmers' self-report does not change with the purity threshold discussed above, the DNA Type obviously does, as do belief types do as well. In a mechanical sense, higher purity thresholds mean that fewer samples are identified as "Positive" for genetic improvement, meaning that observations tend to shift from TP to FP and from FN to TN. Figure 1 shows how the distribution of TP, FP, TN, and FN beliefs change as the purity threshold increases from 70% to 97.5%.

We evaluate the effect of these seed beliefs on agricultural input allocation, with a specific focus on the application of purchased nitrogen fertilizer. To account for the fact that different types of fertilizer are substitutable and often used interchangeably depending on local availability, we convert all chemical fertilizer applications into total nutrient equivalents for nitrogen and

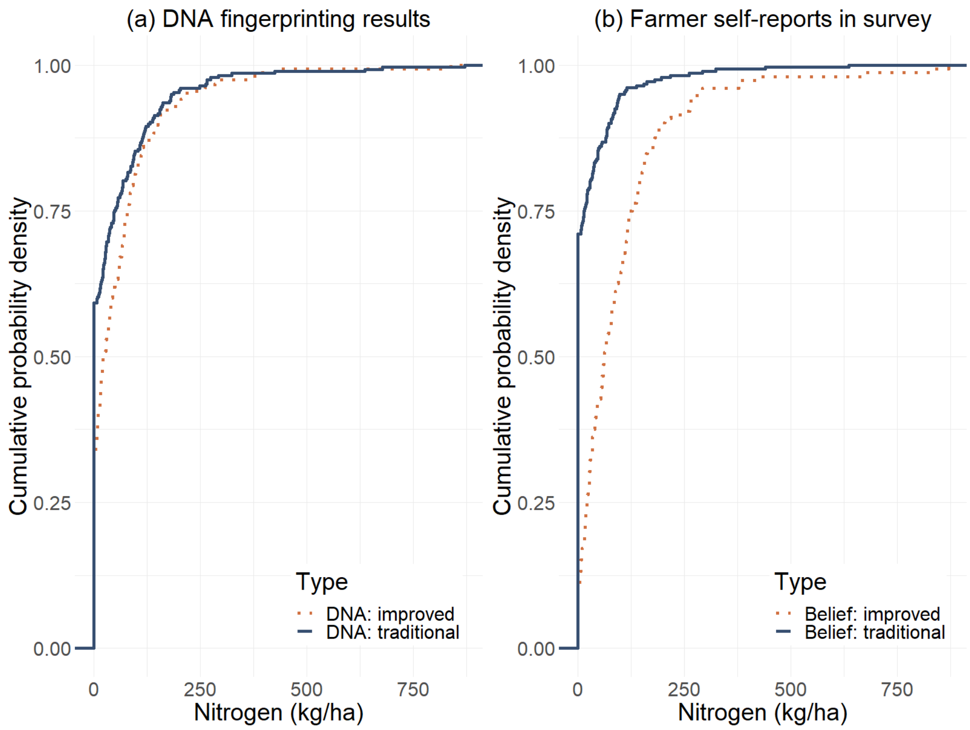

Figure 2: Cumulative distribution of purchased fertilizer measured in nitrogen equivalents by (a) DNA type at a 95% purity threshold and (b) self-reported seed belief. Official nitrogen recommendations range from 110-130kg/ha (higher for hybrid (improved) maize) (Abate et al., 2015)

phosphorus.[10] Ideally, our measure would also include nutrients applied in the form of animal manure – commonly used in our context. As is generally the case, however, we lack information on quantities of manure that are applied and on its nutrient content. As a result, we simply include a dummy variable to indicate whether farmers have applied any manure.

Figure 2 shows cumulative distributions of the nitrogen equivalents in kg per hectare. Panel (b) of this figure shows clearly that farmer seed beliefs – combined with their apparent understanding of production complementarities between improved seeds and nitrogen, which are embedded in official recommendations – drive nitrogen applications much more strongly than the actual DNA type of the seed they sowed, shown in Panel (a). As a further disaggregation of these nitrogen use distributions, Figure A2 shows the same figures by belief type (i.e., TP, TN, FP, FN). As a benchmark, 75% of farmers who report sowing improved maize apply less than the government recommendation of 130kg/ha for hybrid (improved) maize; over 90% of those who believe they are growing traditional maize apply less than their slightly lower recommendation. While it is unclear how seriously farmers (should) take this common recommendation

[10]Data from http://www.soilcropandmore.info/soil/fertiliz.htm (last accessed 9 August 2023) show the nitrogen and phosphorus contents for each respective fertilizer (for Nitrogen: Urea = 46% DAP = 18% NPS = 10%; for Phosphorus: Urea = 0% DAP = 46% NPS = 42%) Although one high-profile study found that actual nitrogen content may deviate from these expected levels (Bold et al., 2017), other studies suggest that similar results may be generated by errors in testing rather than true nutrient deficiencies (see https://blogs.worldbank.org/impactevaluations/devil-details-measuring-agricultural-input-quality (Accessed 19 April 2023).

given the pronounced heterogeneity in growing conditions, taking it at face value, costly under-use is a more prevalent problem than costly over-use of nitrogen. For median farmer of the two types in panel (b) of Figure 2, nitrogen use is less than half this recommendation (improved) and zero (traditional). If the government recommendation were close to optimal in the context of the simple model depicted in Figure A1, this means that false seed beliefs may actually 'distort' nitrogen use in a way that *increases* profits for many of the farmers in our sample.

We provide descriptive statistics for general respondent-level and agricultural production variables by belief type in Table 1. In columns (5)-(10), we report differences in these variables between specified pairs of these belief types. None of the demographic variables is systematically (statistically) different by belief type, but we do see a number of clear differences in production-related characteristics. Farmers who reported sowing improved seeds, whether this belief is true (TP) or false (FP), are more likely to have participated in extension programs, purchased the seed they sowed, and to have larger total land holdings. TP and FP farmers also apply nitrogen at much higher rates on average than those with TN and FN beliefs.

Farmers with TN beliefs rely more on manure as a source of fertilizer. Similar descriptive statistics for a broader set of variables (see Table A2), indicate virtually no other systematic differences in means between these belief types with one exception: more remote locations seem to be less likely to sow true improved seeds.

Table 1: Descriptive statistics by seed belief categories for demographic and production variables with pairwise differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	True positive (TP)	False positive (FP)	True negative (TN)	False negative (FN)	Belief = improved	DNA = Belief	DNA = improved	DNA = traditional	DNA $\neq$ Belief	Belief = traditional
	Means				FP vs TP	TN vs TP	FN vs TP	TN vs FP	FN vs FP	FN vs TN
General Respondent Variables										
Gender (female = 1)	0.18	0.17	0.17	0.12	-0.005	-0.01	-0.06	-0.01	-0.05	-0.05
	(0.39)	(0.38)	(0.38)	(0.33)	(0.06)	(0.05)	(0.06)	(0.05)	(0.07)	(0.05)
Age (years)	43.50	47.59	47.62	46.59	4.13	4.16	3.13	0.03	-1.00	-1.03
	(14.30)	(15.40)	(14.85)	(15.53)	(2.46)	(1.88)	(2.42)	(2.14)	(2.63)	(2.10)
Education	0.35	0.41	0.37	0.44	0.06	0.03	0.09	-0.04	0.03	0.07
(attended any school = 1)	(0.48)	(0.50)	(0.48)	(0.50)	(0.08)	(0.06)	(0.08)	(0.07)	(0.09)	(0.07)
Agricultural Production Variables										
Extension contact (yes = 1)	0.83	0.71	0.44	0.52	-0.12	-0.39***	-0.32***	-0.27***	-0.20*	0.07
	(0.38)	(0.46)	(0.50)	(0.50)	(0.08)	(0.06)	(0.08)	(0.07)	(0.08)	(0.07)
Seeds purchased (yes = 1)	0.93	0.76	0.17	0.21	-0.17**	-0.76***	-0.72***	-0.59***	-0.55***	0.04
	(0.25)	(0.43)	(0.37)	(0.41)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)
Land area (ha)	0.16	0.14	0.09	0.11	-0.02	-0.08***	-0.06**	-0.06**	-0.04	0.02
	(0.16)	(0.16)	(0.10)	(0.17)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Nitrogen (kg/ha)	77.65	127.85	19.05	35.62	50.20***	-58.60***	-42.03**	-108.79***	-92.23***	16.57
	(113.26)	(145.57)	(59.77)	(74.07)	(15.02)	(11.50)	(14.82)	(13.07)	(16.06)	(12.84)
Manure (yes = 1)	0.29	0.37	0.48	0.33	0.07	0.18**	0.04	0.11	-0.03	-0.14
	(0.46)	(0.49)	(0.50)	(0.48)	(0.08)	(0.06)	(0.08)	(0.07)	(0.09)	(0.07)

Notes: Number of observations: 432. Plots have been classified into farmer belief types based on farmer self-report and DNA type evaluated at a 95% threshold for genetic purity. Columns (1)–(4) present the mean values of the respective variables and columns (5)–(10) show the differences between these groups. Tukey tests for equality of mean values. * p <0.10; ** p <0.05; *** p <0.01. Amounts of nitrogen are simple aggregates of the respective nitrogen contents of three fertilizers: Urea = 46%, DAP = 18%, NPS = 10%.

3 Empirical strategy

In this section, we first present the econometric specifications that we use to estimate the impact of seed beliefs on input allocations. We discuss issues related to causal identification and related concerns. We then present a prediction exercise that allows us to project the observed on-farm effects onto data covering the entirety of Ethiopia to arrive at national-level estimates of seed misclassification and fertilizer use.

3.1 Seed beliefs and on-farm input allocation

We are interested in estimating the effects of farmer seed beliefs on their on-farm investment of other inputs. We focus our attention on nitrogen fertilizer application as the most important and most common purchased input in this context. We estimate the same specification for other purchased inputs and present results in the appendix.

In contrast to experimental approaches that artificially manipulate farmers' seed beliefs (e.g., Bulte et al., 2014, 2023), we rely on observational data. It is difficult to imagine any defensible instruments that could be used to isolate plausibly exogenous variation in seed beliefs (i.e., instruments that would satisfy the exclusion restriction). We therefore adopt a second-best empirical strategy that relies on a progressively richer set of controls and post-double selection (PDS) LASSO to account for potentially omitted variables (Belloni et al., 2014). This approach is possible given the detailed set of variables collected by the ESS4. We discuss below the plausibility of a causal interpretation of the results.

Our standard specification is as follows:

$$y_i = \alpha + \beta Belief_i + \gamma DNA_i + \delta Belief_i \times DNA_i + \mathbf{x_i}'\zeta + \varepsilon_i, \quad (1)$$

where y_i is effective nitrogen use in kg per hectare for plot i, $Belief_i$ and DNA_i are indicator variables corresponding respectively to the farmer's belief that the seed is improved, and the DNA test results for each plot. This specification nests estimates of all four belief types – TP, FP, TN, and FN – in one model. Relative to the omitted category (TN), $(\beta + \delta)$ gives the TP effect, β alone gives the FP effect, and γ alone gives the FN effect. We include a vector of control variables, $\mathbf{x_i}$, that we expand to include progressively broader sets of controls. We estimate this specification by OLS and cluster standard errors at the EA level.

Identification of the primary effects of interest in this case requires that $Belief_i$ and DNA_i are uncorrelated with ε, conditional on $\mathbf{x_i}$. Given the richness of the ESS4 data, we estimate a version of this specification in which additional controls in $\mathbf{x_i}$ are chosen using PDS LASSO. We leverage the high-dimensional nature of our dataset by offering a large number and variety of candidate variables for the algorithm to chose from. Specifically, we include $Belief_i$, DNA_i

and $Belief_i \times DNA_i$ as either 'treatment' variable or as part of the amelioration set.[11]

The full set of potential controls in $\mathbf{x_i}$ includes 288 variables from the post-planting, post-harvest, household and community questionnaires, as well as a set of spatial variables at the household and plot level; we also include all squared terms and pairwise interactions. We estimate a regression that includes only the controls selected by PDS LASSO. Information on the composition of the data set and selected variables is reported in Table A3.

3.2 Prediction and projection of national nitrogen application

We aim to extend our estimates to the national level, in order to estimate the magnitude of overuse and underuse by farmers who misidentify their seeds. This exercise consists of three stages: (1) we extrapolate the results from the subsample of maize-growing households whose plots were sampled and tested to the wider ESS4 sample of maize-growing households; (2) we scale these results to estimates of observed national-level nitrogen use by belief type; and (3) we use our empirical model to estimate nitrogen use by FP and FN farmers under a counterfactual of *corrected* beliefs.

3.2.1 Predicting DNA type beyond the sampled plots

In each ESS4 EA in the major maize-growing regions (Tigray, Amhara, Oromia, Harar, and the so-called Southern Nations, Nationalities, and People's Region), up to 10 maize plots were randomly selected for DNA fingerprinting of crop cuts. We compare key farmer and field characteristics of this DNA sample to the maize-growing plots in these regions for which DNA tests are not available. While the regional composition differs significantly in its weights, this comparison (Table A10) shows balance along most key characteristics between the DNA sample and the other maize plots. To infer the DNA type of seeds sown on plots not covered by the DNA fingerprinting, we use a so-called *SuperLearner* machine learning approach to predict the DNA type for fields in the broader ESS sample.

This approach allows us to train, validate, and optimally combine a group of candidate algorithms to generate the best-performing weighted ensemble model for classifying whether a plot is sown with improved or traditional seed.[12] We use 20-fold cross-validation to select the best-performing algorithm according to Area Under the Curve (AUC), Accuracy, Precision and Recall criteria. We then apply the best-predicting model to all remaining plots in the major

[11] To be precise, we designate either $Belief_i$ or DNA_i to be the 'treatment' variable in the PDS LASSO. We then impose the requirement that the amelioration set must include the other variable and the interaction term $Belief_i \times DNA_i$ as well as the set of extended 'controls'. All other potential control variables are then included or excluded based on the PDS LASSO estimation (Belloni, Chernozhukov, and Hansen, 2014).

[12] Our candidate models include RandomForest, glmnet, xgboost, and bagged trees.

maize-growing regions.[13] The results are plot-specific predicted probabilities that the chosen variety is genetically improved.

3.2.2 Extrapolating nitrogen use to national-level

Next, we combine self-reported beliefs for farmers outside the DNA subsample with the predicted probabilities for the DNA type of seeds that they are growing.[14] We sum these probabilities, weighted by the ESS4 sampling weights, for each belief type and use this weighted sum to obtain nationally-representative shares for each belief type. To account for differences in plot size, we weight each probabilistic observation by their respective area and calculate population-weighted shares of total maize area for each group. We then apply these shares to the estimated total area in Ethiopia devoted to maize production in 2018/2019.[15]

3.2.3 Estimating counterfactual nitrogen use under 'corrected' beliefs

In the previous step, we estimate the average *observed* nitrogen use at the national level for each of the four belief types. As an approximation exercise, we quantify the extent of over- and under-application of nitrogen for FP and FN farmers by constructing a counterfactual of 'corrected' beliefs. We first estimate a model of nitrogen application based on the full ESS4 sample and predicted DNA probabilities. We then employ this model to predict nitrogen application under an alternative of correct beliefs, as given by our estimated DNA probabilities. We then follow the probabilistic approach laid out above to predict the "corrected" nitrogen use intensity per group.

Combining area and nitrogen use intensity estimates for FP and FN farmers, we arrive at group-wise nitrogen use under observed and corrected beliefs, which allows us to quantify the extent of nitrogen overuse and underuse. We refer to these as inefficiencies, but we use this term loosely as we do not know what the optimal level of nitrogen usage is for the plots in our data.

[13] Relaxing this restriction and predicting for all maize farmers in ESS4 regardless of their region produces nearly identical results at the national level.

[14] If a farmer reported the seeds sown on a given plot as 'traditional', then these probabilities are constructed as $Pr(TP) = Pr(FP) = 0$, $Pr(FN) = P(DNA = improved)$, and $Pr(TN) = 1 - Pr(DNA = improved)$. If a farmer reported sowing 'improved' seeds, then these probabilities are: $Pr(TN) = Pr(FN) = 0$, $Pr(FP) = 1 - Pr(DNA = improved)$, and $Pr(TP) = Pr(DNA = improved)$.

[15] We take this estimate from the USDA estimate for the 2018/2019 season of 2,415,000 hectares.

4 Results

4.1 Seed beliefs and on-farm input allocation

We report our main results for effective nitrogen application (kg/ha) in Table 2. We report very similar results from the same sequence of estimations for effective phosphorus use in Table A7. In columns (1)-(3), we progressively and manually expand the set of control variables included in the specification, which we estimate using OLS. In columns (4) and (5), we use PDS LASSO to select controls with "Belief" and "DNA" as treatment variables, respectively. Across these five estimations, we see consistently large positive effects of farmer beliefs on nitrogen use, including with PDS LASSO estimation (see Table 1 for average nitrogen use).[16]

The estimated coefficient on the "Belief $\times$ DNA" interaction is consistently large and negative, which somewhat surprisingly suggests that farmers with TP beliefs systematically apply significantly less nitrogen than those with FP beliefs.[17] Summing the first three estimates in column (4), on average TP farmers apply about 54 kg/ha more than TN farmers. FP farmers, by contrast, apply 49 kg/ha *more* than these TP farmers. This puzzling pattern could perhaps reflect the fact that FP farmers are less likely to have purchased their seed than TP farmers (Table 1). If farmers have a fixed budget for agricultural inputs – either literally or as a behavioral regularity – then a reduction in seed expenditures would leave these farmers with more of their agricultural input budget available for fertilizers. This explanation is consistent with supplemental results for total expenditure on purchased inputs reported in Table A8, which show insignificant coefficients on the "Belief $\times$ DNA" interaction – suggesting that while FP farmers spend more on and apply more fertilizer than their TP peers, their overall expenditure on inputs is comparable. An alternative explanation for this pattern is suggested by differences in non-purchased seed: 17.5% of FP farmers use saved or recycled seed whereas only 3.4% of TP do. Such farmers likely face more uncertainty about seed quality or genetic purity and may try to offset a perceived deterioration with more nitrogen.

Since we adopt a second-best identification strategy out of necessity, some caution is merited when interpreting these estimates as causal. However, the robustness of the estimates to the inclusion of a full set of controls – including a set chosen agnostically with PDS LASSO – is at least suggestive of a causal relationship between beliefs and fertilizer use. This imperfect identification raises a third potential explanation for FP farmers applying more nitrogen than TP farmers: it could be that false seed beliefs are endogenous with respect to fertilizer use. Although many of the variables included as candidate and selected PDS LASSO controls (see Table A3) are likely to be correlated with fertilizer use, we cannot rule out the presence of endogeneity bias conclusively.

[16] Supplemental results indicate that these effects of beliefs on nitrogen use are driven by both extensive and intensive margin adjustments.

[17] Note, however, that this estimated coefficient is no longer significant when estimating the extensive and intensive margin adjustments separately.

Table 2: Effective nitrogen use, seed beliefs and DNA type

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Nitrogen (kg/ha)		OLS		PDS LASSO	
Belief	108.79***	135.04***	123.59***	**92.77*** **	122.02***
(improved = 1)	(29.18)	(39.61)	(36.89)	**(34.16)**	(35.79)
DNA	16.57	19.42*	18.87*	10.28	**16.33**
(improved = 1, 95% purity threshold)	(10.97)	(11.62)	(10.25)	(8.06)	(10.23)
Belief × DNA (TP = 1)	-66.77**	-62.10**	-50.59*	-48.99*	-51.30**
	(32.42)	(29.72)	(26.43)	(24.79)	(25.37)
Extension contact (yes = 1)		2.60	1.50	-33.32**	3.16
		(13.07)	(12.96)	(13.99)	(12.15)
Seeds purchased (yes = 1)		-35.82*	-29.07	-51.21	-34.05*
		(20.03)	(18.52)	(31.92)	(18.05)
Field size (ha)		-105.26***	-81.82***	-97.01***	-86.93***
		(32.93)	(29.36)	(30.95)	(29.99)
Manure use (yes = 1)		-4.22	-9.89	3.89	3.42
		(10.97)	(10.84)	(12.56)	(14.54)
OLS: Main controls (3)	no	yes	yes	yes	yes
OLS: Extended controls (12)	no	no	yes	yes	yes
PDS LASSO: No. of candidate controls				288	288
PDS LASSO: No. of selected controls				6	5
Observations	432	432	432	432	432
Adjusted R^2	0.16	0.20	0.24	0.28	0.24

Notes: The DNA indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%. Unreported 'extended controls' in model (3-5) include age, 'has attended any school', farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Amounts of nitrogen are simple aggregates of the respective nitrogen contents of three fertilizers: Urea = 46%, DAP = 18%, NPS = 10%. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * p <0.10; ** p <0.05; *** p <0.01. The set of variables selected in model 4 and 5 selected by the LASSO and used in the OLS post-LASSO estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. Designated 'treatment' variables in the PDS LASSO printed in bold. See Chernozhukov et al. (2014) for details.

A natural question that emerges from these results pertains to the production implications of false belief-based input distortions. Do the patterns that we observe lead to a loss of output or profitability? Given the substantial heterogeneity in maize production conditions in Ethiopia (e.g., soils, agroecologies, etc.), it would be heroic to estimate a production function for this context; and there is no single, uniform yield-maximizing (much less, profit-maximizing) amount of fertilizer for farmers to apply. But this has not historically prevented the Ethiopian government from providing and promoting uniform national fertilizer recommendations (Abate et al., 2015). As mentioned earlier, if interpret these recommendations as profit maximizing levels of N use and take them at face value, the vast majority of farmers in our sample under-use nitrogen. False seed beliefs may therefore 'distort' N use in a way that increases rather than decreases profits.[18] Consequently, it is impossible to impose a general productivity interpretation on these results. What we can state with greater confidence is that these results suggest that false seed beliefs distort input allocations *relative to farmer intent*. Although we do not elicit fertilizer intent directly, the substantial effect of farmer beliefs on fertilizer use suggests that farmers who believed they sowed improved (traditional) seed intended to apply much more (less) fertilizer.

We can provide descriptive evidence of yield implications of these input distortions using the crop-cut based production measures and GPS-based plot size measurements for the plots in our subsample. Together, these gold-standard measures of area and output provide reliable yield data, which we present as cumulative distributions in Figure A3. Yield tends to be higher for true improved seeds – whether aligned with farmer beliefs or not – and for plots managed by farmers who believe they sowed improved seeds – whether aligned with the genetic truth or not.

4.2 Prediction and projection of nitrogen use at national-level

We employ a *SuperLearner* to predict DNA type for those maize plots in the major maize-growing regions that are not covered in the DNA fingerprinting and use these predictions to construct probabilistic estimates of our four belief types as described above. Panel A of Table 3 shows the results of this exercise. In this larger sample, 21% of farmers are predicted to be TP, 15% FP, 48% TN, and 16% FN. Calculations in Panel A also show that farmers with false beliefs occupy around 32% of the land used for maize cultivation in Ethiopia. We consider multiple approaches to predicting DNA type for farmers outside the DNA subsample (see Table A11). The *SuperLearner* outperforms the individual candidate models and is reported in Panel C of Table 3. We use this ensemble model to predict other variables in ESS4 to test this application.

For those sowing truly improved seeds, we see in Panel A that TP farmers use about 79 kg/ha of nitrogen on average compared to 29 kg/ha for FN farmers. The disparity for truly traditional seeds is much more stark: 111 kg/ha for FP farmers compared to 25 kg/ha for TN

[18] To elaborate this point, Figure A2 suggests that while the median FP farmer applies the recommended amount of N, the median TP farmer applies about half the recommended amount.

Table 3: Predicting national-level rates of seed classification and nitrogen application.

Panel A: Population-weighted results by belief type	**True positive**	**False positive**	**True negative**	**False negative**
Shares of belief types	21%	15%	48%	16%
Shares of belief types, population-weighted	27%	15%	44%	14%
Shares of total maize area, population-weighted	41%	19%	27%	13%
Observed nitrogen use (kg/ha)	79.25	111.17	25.18	29.31
Predicted nitrogen use under correct beliefs (kg/ha)	–	65.33	–	55.14
Inefficient overuse (kg/ha)	–	45.84	–	–
Inefficient underuse (kg/ha)	–	–	–	25.83
Panel B: National-level estimates by belief type	**True positive**	**False positive**	**True negative**	**False negative**
National-level maize area (ha)	991,645	465,124	652,446	305,785
National-level nitrogen use under observed beliefs (t)	78,587	51,710	16,432	8,961
National-level nitrogen use under correct beliefs (t)	–	30,387	–	16,861
Inefficient overuse (t)	–	21,323	–	–
Inefficient underuse (t)	–	–	–	7,900
Panel C: Cross-validated model performance	**Accuracy**	**AUC**	**Precision**	**Recall**
Ensemble model (Super Learner)	0.82	0.84	0.73	0.71

Notes: 223 predictors are available to the model-generating algorithm, and we employ a 20-fold cross validation to reduce overfitting. The total national-level maize area is calculated using the percentages from Panel A and an external estimate of 2,42 million hectare used for maize cultivation in Ethiopia (USDA, 2019). Further evaluation of the performance of the algorithms underlying the prediction can be found in the appendix.

farmers. At the aggregate level, these gaps have the potential to produce quantitatively important misallocation of agricultural inputs. An optimal allocation of fertilizer should see farmers with the same seeds utilizing similar intensities of complementary agricultural inputs. Based on this assumption, we can quantify national-level over- and under-allocation of fertilizer by constructing a counterfactual scenario of 'corrected' believes for the FP and FN groups (see Appendix D) for more detail). We cannot definitively say that this represents a misallocation, since we do not know with any confidence what is the optimal level of fertilizer use. However, it is difficult to construct reasonable scenarios in which the discrepancy is efficient.

In Panel B, we report scaled estimates of nitrogen use at the national level by belief type. Using the counterfactual of 'corrected' beliefs described above, we also report what we predict total nitrogen use would have been with correct seed beliefs. FP farmers as a group overuse nitrogen by this standard by over 21,323 mt. FN farmers as a group underuse nitrogen by much less, 7,900 mt.

5 Conclusion

Our results show that significant differences in the allocation of nutrients arise depending on whether farmers have correct or incorrect beliefs about the true genetic type of the seeds they are growing. Because farmers appear to apply complementary inputs at levels reflecting their *beliefs* about seed varieties, their misidentification of the genetic type of the seeds they are growing has implications for input use. To the extent that different genetic types do, in fact, give rise to different optimal levels of input use, our results provide suggestive evidence for misallocation of agricultural inputs.

Our findings confirm earlier studies by Bulte et al. (2014) and Wineman et al. (2020) which highlight the importance of beliefs regarding the quality of agricultural technology on the allocation of complementary inputs – and subsequently, on agricultural productivity. More generally, the paper reinforces the concerns raised by Bold et al. (2017), showing the potentially negative effects of seed quality uncertainty. The potential harm that we identify is not solely from the direct cost – that farmers may have paid a price premium for seeds that are genetically low-quality – but more significantly from the indirect costs associated with inefficient input use. We also note the dynamic concerns mentioned above: farmer seed misidentification will presumably affect profitability. This may, in turn, undermine farmers' willingness to use new technologies more broadly. Farmers may be less willing in the future to purchase seeds and fertilizer. Furthermore, they may harbor unwarranted doubts about the technology packages recommended by agricultural extension services and agri-dealers.

The paper highlights the need for policy makers to closely monitor the seed quality present in the market. This is especially relevant in the context of Ethiopia's nation-wide introduction of direct seed marketing and the broader liberalization of its seed system. Our findings emphasize the need to ensure seed quality even if the government no longer controls the complete seed supply. Furthermore, interventions helping farmers to better identify the seed they are using and to ascertain the purity and quality of seeds could also be beneficial.

Our findings emphasize the importance of a deeper debate about the diffusion of newly developed seed varieties and the driving forces of farmer misinformation in this context. This paper serves as a starting point for further research efforts in this direction and demonstrates the need for more extensive data collection. Misinformation on planted seed types is widespread in rural Ethiopia. Combined with other emerging evidence, these findings raise the concern that such misinformation may be a prevalent problem elsewhere in sub-Saharan Africa. As this paper shows that farmers' beliefs play a crucial role towards the allocation of key agricultural inputs, understanding success and failure of seed identification presents a central issue to for improving agricultural productivity.

Finally, our study adds to the growing body of evidence suggesting the need for caution in using farmer self-reports of "improved" and "traditional" varieties in micro analyses. DNA evidence is increasingly pointing to the problems of relying on self-reporting, and estimates of returns to research (for example) may be biased if they are based on the self-reports. Nor can we conclude that the self-reported data would necessarily give rise to classical measurement error: there are systematic patterns of misidentification, suggesting the need for more complex adjustments. We note also that the costs and other barriers to DNA fingerprinting analysis have been greatly reduced in recent years, so that it is no longer implausible for studies to make some use of DNA-based checks on farmer seed identification (Kosmowski et al., 2019). Future work on varietal adoption and use should benefit greatly from these new measurement techniques.

References

Abate, T., B. Shiferaw, A. Menkir, D. Wegary, Y. Kebede, K. Tesfaye, M. Kassie, G. Bogale, B. Tadesse, and T. Keno. 2015. "Factors that transformed maize productivity in Ethiopia." *Food security* 7:965–981.

Abay, K.A. 2020. "Measurement errors in agricultural data and their implications on marginal returns to modern agricultural inputs." *Agricultural Economics* 51:323–341.

Abay, K.A., G.T. Abate, C.B. Barrett, and T. Bernard. 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., G. Berhane, A.S. Taffesse, K. Abay, and B. Koru. 2018. "Estimating input complementarities with unobserved heterogeneity: Evidence from Ethiopia." *Journal of Agricultural Economics* 69:495–517.

Abay, K.A., L.E. Bevis, and C.B. Barrett. 2021. "Measurement Error Mechanisms Matter: Agricultural intensification with farmer misperceptions and misreporting." *American Journal of Agricultural Economics* 103:498–522.

Abay, K.A., T. Wossen, G.A. Abate, J. Stevenson, H. Michelson, and C.B. Barrett. 2022. "Inferential and behavioral implications of measurement error in agricultural data." *Annual Review of Resource Economics*, pp. .

Beegle, K., C. Carletto, and K. Himelein. 2012. "Reliability of recall in agricultural data." *Journal of development economics* 98:34–41.

Belloni, A., V. Chernozhukov, and C. Hansen. 2014. "High-dimensional methods and inference on structural and treatment effects." *Journal of Economic Perspectives* 28(2):29–50.

Belloni, A., V. Chernozhukov, L. Wang, et al. 2014. "Pivotal estimation via square-root lasso in nonparametric regression." *Annals of Statistics* 42:757–788.

Bold, T., K.C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott. 2017. "Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda." *The Quarterly Journal of Economics* 132:1055–1100.

Breiman, L. 2001. "Random forests." *Machine learning* 45:5–32.

Bulte, E., G. Beekman, S. Di Falco, J. Hella, and P. Lei. 2014. "Behavioral responses and the impact of new agricultural technologies: evidence from a double-blind field experiment in Tanzania." *American Journal of Agricultural Economics* 96:813–830.

Bulte, E., S. Di Falco, M. Kassie, and X. Vollenweider. 2023. "Low-Quality Seeds, Labor Supply and Economic Returns: Experimental Evidence from Tanzania." *The Review of Economics and Statistics*, pp. 1–33.

Burke, W.J., T.S. Jayne, and S.S. Snapp. 2022. "Nitrogen efficiency by soil quality and management regimes on Malawi farms: Can fertilizer use remain profitable?" *World Development* 152:105792.

Chen, T., and C. Guestrin. 2016. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. pp. 785–794.

Duflo, E., M. Kremer, and J. Robinson. 2008. "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya." *American Economic Review* 98:482–88.

Ellis, F. 1992. *Agricultural policies in developing countries*. Cambridge university press.

Ezui, K., A. Franke, A. Mando, B. Ahiabor, F. Tetteh, J. Sogbedji, B. Janssen, and K. Giller. 2016. "Fertiliser requirements for balanced nutrition of cassava across eight locations in West Africa." *Field Crops Research* 185:69–78.

Floro IV, V.O., R.A. Labarta, L.A. Becerra López-Lavalle, J.M. Martinez, and T.M. Ovalle. 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* 69:518–536.

Foster, A.D., and M.R. Rosenzweig. 1995. "Learning by doing and learning from others: Human capital and technical change in agriculture." *Journal of political Economy* 103:1176–1209.

Friedman, J., T. Hastie, and R. Tibshirani. 2010. "Regularization paths for generalized linear models via coordinate descent." *Journal of statistical software* 33:1.

Ilukor, J., T. Kilic, J. Stevenson, S. Gourlay, F. Kosmowski, A. Kilian, J. Sserumaga, and G. Asea. 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*.

Kassie, M., H. Teklewold, M. Jaleta, P. Marenya, and O. Erenstein. 2015. "Understanding the adoption of a portfolio of sustainable intensification practices in eastern and southern Africa." *Land use policy* 42:400–411.

Kosmowski, F., S. Alemu, P. Mallia, J. Stevenson, and K. Macours. 2020. "Shining a Brighter Light: Comprehensive Evidence on Adoption and Diffusion of CGIAR-related Innovations in Ethiopia.", pp. .

Kosmowski, F., A. Aragaw, A. Kilian, A. Ambel, J. Ilukor, B. Yigezu, and J. Stevenson. 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.

Macours, K. 2019. "Farmers' demand and the traits and diffusion of agricultural innovations in developing countries." *Annual Review of Resource Economics* 11:483–499.

Maredia, M.K., B.A. Reyes, J. Manu-Aduening, A. Dankyi, P. Hamazakaza, K. Muimui, I. Rabbi, P. Kulakow, E. Parkes, T. Abdoulaye, et al. 2016. "Testing alternative methods of varietal identification using DNA fingerprinting: Results of pilot studies in Ghana and Zambia." Working paper.

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

Nyangena, W., M. Juma, et al. 2014. "Impact of improved farm technologies on yields: the case of improved maize varieties and inorganic fertilizer in Kenya." *Environment for Development Discussion Paper-Resources for the Future (RFF)*, pp. .

Phillips, R.V., M.J. van der Laan, H. Lee, and S. Gruber. 2022. "Practical considerations for specifying a super learner." *arXiv preprint arXiv:2204.06139*, pp. .

Roobroeck, D., C.A. Palm, G. Nziguheba, R. Weil, and B. Vanlauwe. 2021. "Assessing and understanding non-responsiveness of maize and soybean to fertilizer applications in African smallholder farms." *Agriculture, Ecosystems Environment* 305:107165.

Spielman, D.J., D.K. Mekonnen, and D. Alemu. 2013. "4 Seed, Fertilizer, and Agricultural Extension in Ethiopia." In *Food and Agriculture in Ethiopia*. University of Pennsylvania Press, pp. 84–122.

Stevenson, J., K. Macours, and D. Gollin. 2018. "The Rigor Revolution in Impact Assessment:Implications for CGIAR." Working paper.

Tolessa, D., B. Tesfa, N. Wakene, W. Tenaw, L. Minale, M. Tewodros, M. Burtukan, and M. Waga. 2001. "A Review of Fertilizer Management Research on Maize in Ethiopia." In *Proceedings of the Second National Maize Workshop of Ethiopia. 12-16, November, Addis Ababa*. pp. 46–55.

Van der Laan, M.J., E.C. Polley, and A.E. Hubbard. 2007. "Super learner." *Statistical applications in genetics and molecular biology* 6.

Wineman, A., T. Njagi, C.L. Anderson, T.W. Reynolds, D.Y. Alia, P. Wainaina, E. Njue, P. Biscaye, and M.W. Ayieko. 2020. "A case of mistaken identity? Measuring rates of improved seed adoption in Tanzania using DNA fingerprinting." *Journal of Agricultural Economics* 71:719–741.

Wossen, T., K.A. Abay, and T. Abdoulaye. 2022. "Misperceiving and misreporting input quality: Implications for input use and productivity." *Journal of Development Economics* 157:102869.

Yirga, C., K. Negisho, L. Oruko, D. Alemu, and G. Taxler. 2016. "Tracking the diffusion of crop varieties using DNA fingerprinting." Working paper, Ethiopian Institute of Agricultural Research (EIAR).

6 Appendix: Conceptual Framework

To illustrate the indirect costs associated with seed misclassification that emerge from input complementarities, consider the following stylized example. For simplicity, assume there are two varieties of maize: improved and traditional. The improved variety produces higher yield on average. While yield for both varieties increases when fertilizer is applied (at least up to a point), the improved variety is generally more responsive to fertilizer, such that the profit-maximizing level of input application is higher for the improved than the traditional variety.

For simplicity, assume farmers are risk-neutral, maximize profit, and are not liquidity constrained. Each farmer has a single maize plot and first chooses between sowing the improved or traditional variety on this plot. Farmers subsequently optimize applied inputs, which we partition into fertilizer x and a vector of all other inputs $\mathbf{z}$, to maximize expected profits based on known input costs, output price p and a production function $f(x,\mathbf{z})$ that maps inputs into expected maize production and is known by the farmer. We assume that inputs x and $\mathbf{z}$ are available and of known quality (i.e., only seeds are subject to misclassification or incomplete information).

We denote seed beliefs according to whether farmers believe the seed they sowed is improved or traditional and therefore use the belief terms $\{$Positive, Negative$\}$ to refer to the believed presence or absence of improved germplasm. Comparing these beliefs to the DNA test results enables us to distinguish whether these beliefs are $\{$True, False$\}$, thus providing four seed belief categories: True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN).

If farmers only had true seed beliefs, there would be no misclassification, and the profit maximizing level of fertilizer x^* (assuming optimal application of other inputs, $\mathbf{z}^*$) for the improved (TP) and traditional (TN) variety satisfies the following marginal value product (MVP) conditions:

$$MVP_{TP} = \frac{\partial f_P(x^*_{TP} \mid \mathbf{z}^*_P)}{\partial x} p = p_x$$

$$MVP_{TN} = \frac{\partial f_N(x^*_{TN} \mid \mathbf{z}^*_N)}{\partial x} p = p_x$$

where p is the output price of maize, p_x is the fertilizer price, f_P is the production function for the improved variety (i.e., DNA test is 'Positive'), and f_N is the production function for the traditional variety (i.e., DNA test is 'Negative').

The possibility of false seed beliefs complicates this optimization problem because the improved and traditional maize variety respond differently to applied inputs, including fertilizer, such that over a relevant fertilizer range $\frac{\partial f_P}{\partial x} > \frac{\partial f_N}{\partial x} > 0$ and $x^*_P > x^*_N$. Assume further that applied inputs have stronger complementarities for the improved variety such that $\frac{\partial^2 f_P}{\partial x \partial z_i} > \frac{\partial^2 f_N}{\partial x \partial z_i} > 0$ for at least some inputs i in $\mathbf{z}$. Thus, the discrepancy between the actual and the perceived MVP of x

emerges from both a mistaken perception of the fertilizer responsiveness of the seed germplasm and the sub-optimal application of other inputs $\mathbf{z}$. Specifically, after the farmer locks in his allocation of other inputs based on his seed beliefs, the MVP in the case of false beliefs is given by the FP and FN cases:

$$MVP_{FP} = \frac{\partial f_N(x \mid \mathbf{z}_P^*)}{\partial x} p$$

$$MVP_{FN} = \frac{\partial f_P(x \mid \mathbf{z}_N^*)}{\partial x} p$$

We graphically depict the input distortions associated with seed misclassification in this conceptual model in Figure A1. Optimal input allocations (x^* and $\mathbf{z}^*$) defined above for true beliefs TP and TN provide benchmarks in this depiction. To unpack the input distortions introduced by false beliefs, consider two potential degrees of misclassification. First, *resolved misclassification* occurs when a farmer optimizes $\mathbf{z}^*$ based on false beliefs, but subsequently chooses x^* based on correct germplasm beliefs (i.e., after learning the true identity of the seed he sowed) and taking $\mathbf{z}^*$ as given. While this scenario may not be realistic in practice, it provides a useful mid-point between no and full misclassification and helps to elucidate potential sources of input distortions. Resolved misclassification generates optimal fertilizer levels of x_{FP}^* and x_{FN}^* conditional on prior sub-optimal application of other inputs:

$$MVP_{FP} = \frac{\partial f_N(x_{FP}^* \mid \mathbf{z}_P^*)}{\partial x} p = p_x$$

$$MVP_{FN} = \frac{\partial f_P(x_{TN}^* \mid \mathbf{z}_N^*)}{\partial x} p = p_x$$

As shown graphically, the fact that $x_{TP}^* > x_{FN}^*$ even though the seeds in these cases are all improved emerges from the incorrect beliefs of the farmer who consequently under-applies other inputs $\mathbf{z}^*$, which lowers the MVP of x as a result of input complementarities. The reverse is true for the traditional seeds: $x_{TN}^* < x_{FP}^*$ because the farmer over-invests in $\mathbf{z}^*$ and inadvertently raises the MVP of x.

Second, *full (unresolved) misclassification* occurs when the farmer's false seed beliefs persist throughout the entire production season, which distorts all input applications. Without the information required to choose optimal fertilizer levels, he instead chooses sub-optimal applications $\tilde{x}_{FP} = x_{TP}^*$ and $\tilde{x}_{FN} = x_{TN}^*$.

Two distinct indirect misclassification costs are apparent in Figure A1. In the FN case, the farmer mistakenly believes he is growing the traditional variety and under-applies fertilizer. As shown by area A, at this sub-optimal level of fertilizer investment the farmer misses an opportunity to earn higher profits as even a small increase in fertilizer would increase the value of maize production net of the fertilizer cost. In the FP case, the farmer applies fertilizer beyond its optimal level and thereby reduces his profit by area B. The losses due to false seed beliefs as represented stylistically by areas A and B are generally not symmetric since they depend on the differential marginal productivity of fertilizer below and above the optimum for traditional and improved seeds. Moreover, these losses also reflect interactions with other inputs, which may be further characterized by important non-linearities, including some that differ by variety.

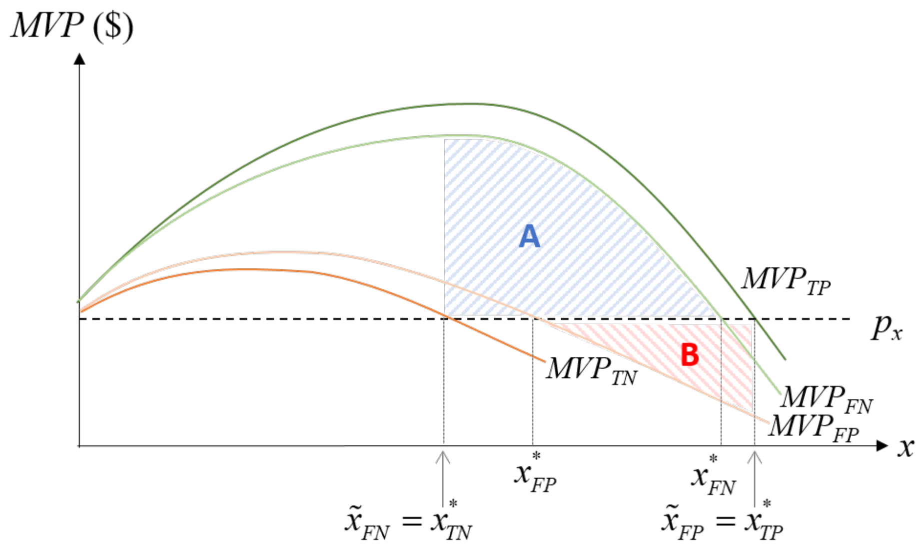

Figure A1: Stylized depiction of optimal fertilizer (x) given price p_x and marginal value product of fertilizer (MVP) for improved seeds (TP, FN) and traditional seeds (FP, TN) for true and false beliefs. Area A represents unrealized profit due to lower than optimal fertilizer application based FN seed beliefs. Area B represents lost profit due to applying fertilizer beyond its optimal level based on FP seed beliefs.

7 Appendix: Survey and Data Details

Our analysis requires a genetic definition of "improved" varieties that we can compare with farmer reports. In this paper, we consider two kinds of improved varieties: hybrids and improved OPVs. We contrast these two distinct types of improved varieties with varieties that we consider to be traditional. Varieties that are "traditional" from the perspective of modern breeding programs may nevertheless be the result of significant on-farm selection over many generations of farmers. As with many breeders and in recognition of these informal but intentional genetic selection processes, we refer to this alternative to improved varieties as "traditional" rather than "unimproved." (In the literature, the term "local" is used interchangeably with "traditional.") Since maize is not indigenous to Ethiopia, the traditional varieties are themselves very likely to bear some traces of germplasm that was viewed as "improved" at some point in the past. In more commercial settings, farmers typically purchase fresh seed every year; modern seed producers follow a set of well-understood practices to maintain the genetic purity of their varieties. However, these practices are not generally feasible for smallholders in Africa.

Given the potential for genetic drift in improved maize seeds, we need to define a purity threshold at which we would judge a sample from a farmer's plot to be improved. The genotyping analysis matched the collected material against a reference library of improved varieties, including both OPVs and hybrids. For each of the materials in the reference library, and for each collected sample of maize seed, the DNA fingerprinting process examines the genetic sequences present at around 50,000 marker locations on the genome. The metric that we rely on is how closely the collected sample matches the most closely related material from the reference library. The precise selection of these markers works as follows: When the sample is close enough to material from the reference library, the data identifies the specific hybrid or OPV that the sample resembles, along with the proportion of the 50,000 markers that are matched. Some samples matched closely to the reference varieties, while others showed more variation, reflecting either the dilution of relatively pure seed during the seed multiplication process or the genetic drift associated with saved (i.e., recycled) seed.

8 Appendix: Technical Details for Prediction Exercise

8.1 Features included in the prediction exercise

In order to predict the DNA status in the broader ESS4 sample, we select a number of candidate variables that we feed into machine learning algorithms. These include 223 variables from the post-planting, community and household questionnaire. The selected variables broadly cover the following set of features:

- Agriculture and farming management practices
- Allocation of agricultural inputs (seed quantity, fertilizer quantity, pesticides etc.)
- Demographic household information
- Distance/remoteness variables (to population center, markets etc.)
- Community-specific variables (accessibility, infrastructure etc.)
- Plot characteristics: location, distance from household etc.
- Seed characteristics: seed source, recycle status etc.
- Soil and climate specific data based on GPS an GIS measurements

8.2 Algorithms used in the prediction exercise

We use the *SuperLearner* algorithm to identify the subset of the above described features that are the best joint predictors of the DNA variable. This prediction routine is an ensemble method that employs cross-validation to determine the optimally weighted combination of predictions from a group of candidate models. The approach supports the use of machine learning algorithms in addition to standard parametric algorithms. Model outputs are the combined predictions of these candidate models (Van der Laan, Polley, and Hubbard, 2007). For a detailed exposition of the *SuperLearner* approach, see https://cran.r-project.org/web/packages/SuperLearner/vignettes/Guide-to-SuperLearner.html.

We include candidate algorithms in our *SuperLearner* that perform particularly well on classification tasks such as Random Forest, XGBoost, glmnet and a bagged tree algorithm. The Random Forest and bagged classification tree models seek to optimize prediction performance by averaging a collection of de-correlated decision trees (also known as bagging) (Breiman, 2001). The XGBoost (Extreme Gradient Boosting) algorithm employs gradient tree boosting which sequentially connects a number of decision trees whereby each tree aims at minimizing the residual error of the previous tree (Chen and Guestrin, 2016). Glmnet combines a generalised linear model with a configurable regularization parameter (elastic net or LASSO) to maximize predictive performance (Friedman, Hastie, and Tibshirani, 2010).

When specifying our model, we follow the procedure outlined in Phillips et al. (2022). For each of the models, we use stratified cross-validation to ensure that the model will generalize beyond the sample it is trained on. Specifically, we use 20-fold cross validation to select model parameters that minimize the area under the curve (AUC) on the held-out validation data across the folds [19]. Across all of our candidate models, we use Random Forest as the screener algorithm to perform feature selection before model training.

We use a number of different metrics to evaluate the performance of both the *SuperLearner* model and the algorithms included in the ensemble. Accuracy corresponds to the percentage of correct predictions in the validation data (correct predictions/all predictions). The Area Under the Curve (AUC) represents the overall performance of the classifier in terms of its ability to distinguish between positive and negative classes. Precision evaluates the proportion of true positive predictions among all positive predictions (True Positives/(True Positives + False Positives)). Recall measures the proportion of true positive predictions among all true positive instances in the data (True Positives/(True Positives + False Negatives).

8.3 Performance of prediction algorithms

Panel A of Table A11 shows the performance of the *SuperLearner* and a number of candidate models that are aggregated in the ensemble. The ensemble model outperforms the candidate models when considering all four criteria (AUC: 0.84 , Accuracy: 0.82 , Precision: 0.73, Recall: 0.71). Among the candidate models XGBoost has the highest accuracy (0.82), precision (0.71) and recall (0.73), while Random Forest has the highest AUC (0.87).

[19] Cross-validation is a commonly-used method for validating machine learning models. By dividing the data into folds and training and testing the model on these different subsets of the data, cross-validation helps to prevent overfitting and provides a more accurate estimate of model performance. It is particularly useful when working with small datasets or when the performance of the model needs to be estimated accurately.

Table A1: Distribution of seed beliefs at the plot-level

Farmer self-report	**Traditional**				**Improved**		
	64.81				35.19		
DNA purity	**70%**	**80%**	**85%**	**90%**	**92.5%**	**95%**	**97.5%**
Traditional	16.20	16.67	19.44	25.00	39.35	64.12	78.01
Improved	83.80	83.33	80.56	75.00	60.65	35.88	21.99
True Positive (TP)	27.78	27.78	27.55	26.16	24.31	20.60	15.51
False Positive (FP)	7.41	7.41	7.64	9.03	10.88	14.58	19.68
True Negative (TN)	8.80	9.26	11.81	15.97	28.47	49.54	58.33
False Negative (FN)	56.02	55.56	53.01	48.84	36.34	15.28	6.48
Correctly identified	36.57	37.04	39.35	42.13	52.78	70.14	73.84
Misidentified	63.43	62.96	60.65	57.87	47.22	29.84	26.16

Notes: Values reflect column-wise shares of the respective groups of the full sample of 432 observations

Panel B of Table A11 demonstrates the ability of the *SuperLearner* to predict variables that are measured for all observations in the broader ESS4 sample (extension program participation and seed source). Here, we train and validate the *SuperLearner* on the DNA sub-sample, then apply it to the wider ESS survey. This allows us to evaluate the general suitability of the *SuperLearner* approach to predict binary outcomes in the ESS survey. We can see that the in-sample and out-of-sample performance of the ensemble method is high across all assessed variables (Accuracy and AUC being consistently above 80%). The results show that the *SuperLearner* is able to reliably predict binary variables in the ESS4, making a machine learning approach a viable option for overcoming the scaling constraints of the DNA subsample.

8.4 Predicting fertilizer for the scaling exercise

To predict fertilizer application for FP and FN farmers under correct beliefs, we re-estimate our PDS LASSO specification based on the wider ESS4 sample, following the model structure as presented in Table 2, Column 4. This sample includes the previously predicted DNA types for the extended set of plots. We employ the estimated model to predict fertilizer application for FP and FN farmers given correct beliefs, i.e., beliefs fully aligned with the measured and predicted DNA type, respectively.

9 Appendix: Supplemental Figures and Tables

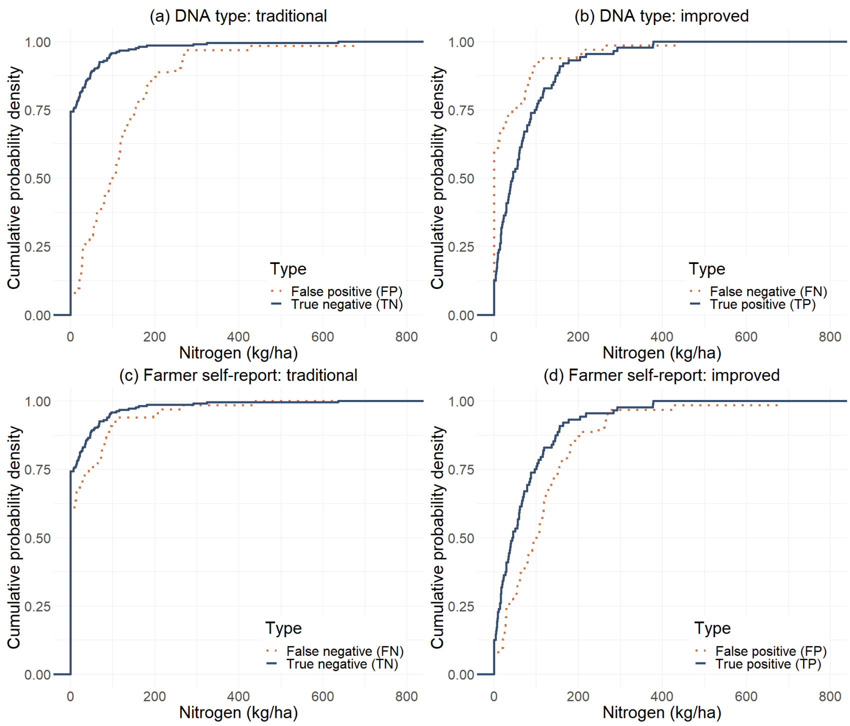

Figure A2: Cumulative distribution of applied nitrogen equivalents for pairwise comparisons of maize seed belief types at a 95% purity threshold. In panels (a) and (b), farmer beliefs differ but their seeds are the same (i.e., traditional and improved, respectively). In panels (c) and (d), farmer beliefs are the same (i.e., traditional and improved, respectively) but the genetic identity of their seeds differ.

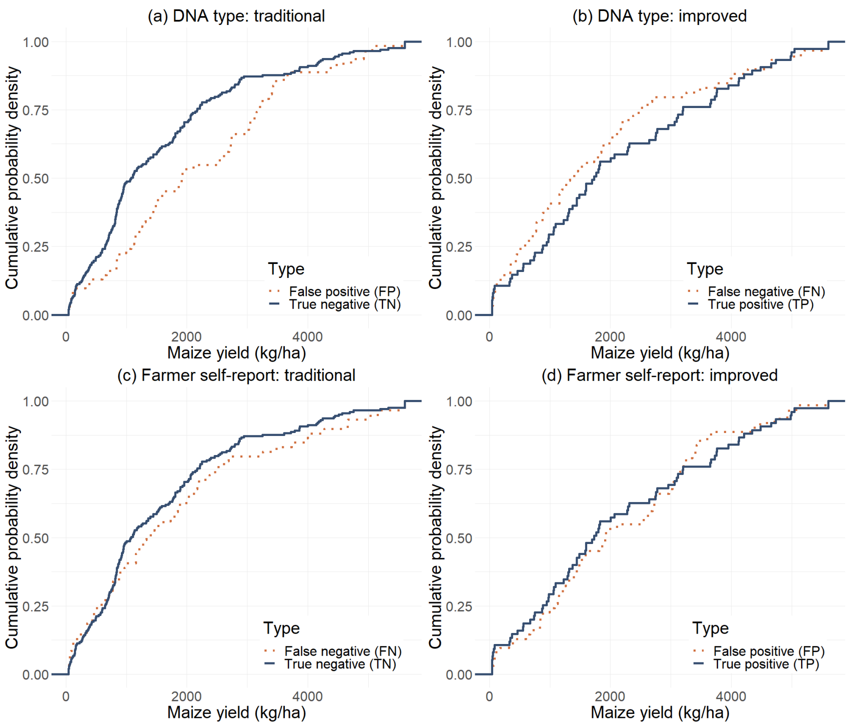

Figure A3: Cumulative distribution of maize yield for pairwise comparisons of maize seed belief types at a 95% purity threshold. In panels (a) and (b), farmer beliefs differ but their seeds are the same (i.e., traditional and improved, respectively). In panels (c) and (d), farmer beliefs are the same (i.e., traditional and improved, respectively) but the genetic identity of their seeds differ.

Table A2: Descriptive statistics by seed belief categories for additional outcome variables with pairwise differences and tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	True positive (TP)	False positive (FP)	True negative (TN)	False negative (FN)	Belief = improved	DNA = Belief	DNA = improved	DNA = traditional	DNA ≠ Belief	Belief = traditional
	Means				FP vs TP	TN vs TP	FN vs TP	TN vs FP	FN vs FP	FN vs TN
Additional Variables										
Phosphorus (kg/ha)	52.77	77.24	12.65	24.25	24.47*	-40.11***	-28.52**	-64.58***	-52.98***	11.60
	(61.66)	(96.79)	(42.97)	(51.90)	(9.69)	(7.43)	(9.56)	(8.44)	(10.37)	(8.29)
Total cost of purchased inputs (ETB/ha)	5519.53	5428.24	924.38	1639.57	-91.29	-4595.15***	-3879.96***	-4503.86***	-3788.67***	715.19
	(6267.13)	(5242.57)	(2341.67)	(3336.19)	(668.58)	(512.17)	(662.53)	(582.05)	(717.91)	(575.09)
NPS (kg/ha)	84.36	150.45	7.78	19.85	66.09***	-76.59***	-64.52***	-142.68***	-130.61***	12.07
	(134.92)	(231.83)	(34.52)	(50.09)	(18.41)	(14.10)	(18.16)	(16.03)	(19.69)	(15.74)
DAP (kg/ha)	37.68	30.53	20.41	34.60	-7.15	-17.28	-3.09	-10.12	-4.07	14.19
	(93.09)	(89.39)	(89.58)	(109.13)	(15.39)	(11.79)	(15.19)	(13.40)	(16.47)	(13.16)
UREA (kg/ha)	135.72	221.73	31.74	59.58	97.56***	-103.97***	-76.13**	-201.53***	-173.69***	27.84
	(221.73)	(273.68)	(102.81)	(124.38)	(27.83)	(21.32)	(27.46)	(24.23)	(29.77)	(23.80)
Cost of purchased NPS (ETB/ha)	1045.61	1371.98	117.17	128.92	326.37	-928.45***	-916.69***	-1254.81***	-1243.05***	11.76
	(1775.63)	(1936.83)	(592.63)	(471.34)	(194.83)	(149.25)	(192.22)	(169.61)	(208.42)	(166.61)
Cost of purchased DAP (ETB/ha)	468.63	384.16	213.00	566.76	-84.47	-255.63	98.14	-171.16	182.60	353.77
	(1140.58)	(1142.65)	(1046.96)	(1752.31)	(199.52)	(152.84)	(196.85)	(173.70)	(213.44)	(170.62)
Cost of purchased UREA (ETB/ha)	1587.43	1734.65	351.46	742.57	147.22	-1235.97	-844.86	-1383.19	-992.081	391.11
	(2973.96)	(1655.94)	(1205.26)	(1724.82)	(303.17)	(232.25)	(299.11)	(263.93)	(324.33)	(259.26)
Cost of maize seeds purchased (ETB/ha)	2336.47	1937.45	227.95	179.19	-399.02	-2108.51***	-2157.28***	-1709.50***	-1758.26***	-48.77
	(3551.00)	(3592.83)	(1014.73)	(493.07)	(369.01)	(282.69)	(365.67)	(321.25)	(396.24)	(317.41)
Total household labor (hours/ha)	1350.24	1546.81	2116.35	1462.33	196.57	766.11	112.09	569.54	-84.47	-654.02
	(1490.07)	(1601.14)	(5491.14)	(1910.61)	(667.14)	(511.07)	(661.10)	(580.79)	(716.37)	(573.85)
Total hired labor (hours/ha)	36.67	54.74	48.07	52.82	18.07	11.40	16.15	-6.67	1.92	4.75
	(88.34)	(150.95)	(151.35)	(155.13)	(23.26)	(17.82)	(22.95)	(20.25)	(24.88)	(19.89)
Access to credit services	0.15	0.29	0.11	0.11	0.14	-0.04	-0.04	-0.18	-0.18	-0.0014
	(0.36)	(0.46)	(0.31)	(0.31)	(0.06)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)
Mobile phone ownership	0.30	0.35	0.29	0.41	0.05	-0.02	0.11	-0.06	0.06	0.12
	(0.46)	(0.48)	(0.45)	(0.50)	(0.08)	(0.06)	(0.08)	(0.07)	(0.08)	(0.07)
HH distance to nearest market (in KMs)	56.95	43.09	64.55	60.26	-13.87	7.60	3.31	21.46***	17.17*	-4.29
	(29.04)	(36.44)	(44.45)	(50.70)	(6.87)	(5.26)	(6.77)	(5.98)	(7.35)	(5.87)
HH distance to nearest road (in KMs)	17.22	12.37	17.01	17.59	-4.85	-0.21	0.38	4.64	5.23	0.58
	(16.80)	(13.44)	(17.07)	(20.66)	(2.82)	(2.16)	(2.79)	(2.46)	(3.02)	(2.41)
HH distance to nearest population center	25.54	21.07	29.77	23.63	-4.47	4.23	-1.90	8.70**	2.56	-6.13
(in KMs)	(15.47)	(13.72)	(23.40)	(20.43)	(3.34)	(2.56)	(3.30)	(2.91)	(3.58)	(2.86)

Notes: Number of observations: 432. Plots have been classified into farmer belief types based on farmer self-report and DNA type evaluated at a 95% threshold for genetic purity. Columns 1 – 4 present the mean values of the respective variables and columns 5 – 10 show the differences between these groups. Tukey tests for equality of mean values. Amounts of phosphorus are simple aggregates of the respective phosphorus contents of two fertilizers: DAP = 46%, NPS = 42%. Cost outcomes are calculated as the sum of spending for maize seeds, Urea, DAP, NPS and all other inorganic fertilizer purchased.

Table A3: High-dimensional controls as candidate variables for main PDS-LASSO specification

Variable	Description
LASSO-selected covariates in model (4)	
s4q13b	When did you plant the seeds for the [Crop] on [FIELD]? (Year)
s4q14	What type of crop sowing techniques was used for [CROP] on [FIELD]?
s3q16	Is [FIELD] under Extension Program during the current agricultural season?
s5q16	Was any of the [SEED] that you used left over from a previous season?
s7q22	Total NPS fertilizer ready for use for main season in 2011 E.C
cs4q26	How many private secondary schools are there in this community?
Amelioration set in model (4)	
DNA	type indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%
Belief x DNA	Farmer self-report interacted with DNA type
s7q04	Do you participate in the extension program?
s7q06	Do you get credit services?
s5q02	Did you purchase any of the [SEED] used?
s2q04	Has [NAME] ever attended any school?
s11b_ind_01	Do you own any mobile phones, exclusively or jointly with someone else?
s3q08	Record area of the field using GPS
s1q03a	How old is [NAME]? (COMPLETED YEARS)
dist_road	HH Distance in (KMs) to Nearest Major Road
dist_market	HH Distance in (KMs) to Nearest Market
dist_popcenter	HH Distance in (KMs) to Nearest Population Center with +20,000
s1q02	What is the sex of [NAME]?
s3q25	Do you use any manure on [FIELD] in this agricultural season?
Candidate variables for PDS-LASSO models	
saq01	Region code
saq15	What is the holder's farm type?
s7q01	Do you exercise crop rotation on your land holding?
s7q02	Did you use chemical fertilizers on any one of your crop field?
s7q09	Do you not get advisory services?
s7q11_1	Who are your major suppliers of fertilizer? Supplier 1
s7q15	What type of Plough equipment do you mostly use?
s7q16	What type of machine do you mostly use to thresh the crop products?
s7q17	Do you plough any additional field other than the fields you had in the last
s7q29	Have you participated in watershed activities in your community?
s4q02	Was the area planted with [CROP] on [FIELD] pure stand or mixed?
s4q04	Was prevention measure taken to prevent damage of [CROP] on [FIELD]?
s4q08	Was [CROP] damaged on [FIELD]?
s4q13a	When did you plant the seeds for the [Crop] on [FIELD]? (Month)
s4q13b	When did you plant the seeds for the [Crop] on [FIELD]? (Year)
s4q14	What type of crop sowing techniques was used for [CROP] on [FIELD]?
s4q22	Do you intend to sell any of [CROP] to be harvested from [FIELD]?
s3q03b	During this season, what is the status of this [FIELD]?
s3q04	What was the method of cropping in this field?

Table A3 continued from previous page

Variable	Description
s3q12	ENUMERATOR: What is the appearance of this field?
individual_id	Who in HH makes primary decisions on [FIELD]?
s3q14	Are there other household members that the primary decision maker consults
s3q16	Is [FIELD] under Extension Program during the current agricultural season?
s3q17	Is [FIELD] irrigated during the current agricultural season?
s3q24	Do you use any other chemical fertilizers (other than UREA,DAP and NPS)?
s3q26	Do you use any compost on [FIELD] in this agricultural season?
s3q27	Do you use any other organic fertilizer on [FIELD]?
s3q34	During the last three years, have you planted a legume on this [FIELD]?
s3q37	What was the previous state of [FIELD] in the previous agricultural season?
s3q38	Is [Field] prevented from Erosion?
s3q40	Do you use any method for soil fertility on [FIELD]?
s2q03	Does your household have a document for this [PARCEL]?
s2q05	How did your household acquire [PARCEL]?
s2q16	What is the predominant soil type of [PARCEL]?
s5q12	Did you receive any of the [SEED] for free?
s5q16	Was any of the [SEED] that you used left over from a previous season?
s1q01	What is the relationship of [NAME] to the head of household?
s1q08	What is [NAME]'S main religion?
s1q09	What is [NAME]'s marital status?
s1q12	In what region were you born?
s1q13	Does [NAME]'s biological father live in this household?
s1q17	Does [NAME]'s biological mother live in this household?
s1q16	Highest educational level completed by [NAME]'s biological father
s1q20	Highest educational level completed by [NAME]'s biological mother
s1q21	What was the industry of occupation of [NAME]'s biological father?
s1q22	What was the industry of occupation of [NAME]'s biological mother?
s2q01	ENUMERATOR: Is this person answering for himself / herself?
s2q04	Has [NAME] ever attended any school?
s2q19	Does [NAME] plan to attend school next year?
s4q01	Is [NAME] asnwering for himself / herself?
s4q33b	Has [NAME] worked for payment in the last 12 months?
s4q45	In the past 12 months has [NAME] been employed as temporary labour by PSNP?
s4q48	Does [NAME] do any other casual/temporary labour work in past 12 months?
s4q51	Did [NAME] work for other households for free in the last 12 months?
s4q53	Did [NAME] participate in free labour contribution in the last 12 months?
cs2aq01	Do the children in this community typically wear neat clothing?
cs2aq02	Do the children under 10 in this community typically wear shoes?
cs2aq03	Do the adults in this community typically wear neat clothing?
cs2aq05	Are the house surroundings in this community swept clean?
cs2aq06	What material is most commonly used for the outside walls of the houses?
cs2aq07	What material is most commonly used for the roofs of the houses?
cs2aq09	Is there a publicly accessible notice board in this community?
cs2aq11	Is there a suggestion box in this community?

Table A3 continued from previous page

Variable	Description
cs3q01	More people moved into this community or more people moved away?
cs3q04a	What are the religions practiced by residents of this community? (1ST)
cs3q07	What is the most common use of land in this community?
cs3q08	Is the land of the community .. .?
cs3q11a	Is there gullies on agricultural land in this community?
cs3q12a_1	What are the common types of marriages witnessed in this community? (1ST)
cs4q01	What is the type of main access road surface in this community?
cs4q03	Do vehicles pass on the main road in this community throughout the year?
cs4q04__0	Main road passable by public transport: Not passable the whole year
cs4q04__1	Main road passable by public transport: SEPTEMBER
cs4q04__2	Main road passable by public transport: OCTOBER
cs4q04__3	Main road passable by public transport: NOVEMBER
cs4q04__4	Main road passable by public transport: DECEMBER
cs4q04__5	Main road passable by public transport: JANUARY
cs4q04__6	Main road passable by public transport: FEBRUARY
cs4q04__7	Main road passable by public transport: MARCH
cs4q04__8	Main road passable by public transport: APRIL
cs4q04__9	Main road passable by public transport: MAY
cs4q04__10	Main road passable by public transport: JUNE
cs4q04__11	Main road passable by public transport: JULY
cs4q04__12	Main road passable by public transport: AUGUST
cs4q04__13	Main road passable by public transport: Passable the whole year
cs4q05__0	Main road passable by a lorry: Not passable the whole year
cs4q05__1	Main road passable by a lorry: SEPTEMBER
cs4q05__2	Main road passable by a lorry: OCTOBER
cs4q05__3	Main road passable by a lorry: NOVEMBER
cs4q05__4	Main road passable by a lorry: DECEMBER
cs4q05__5	Main road passable by a lorry: JANUARY
cs4q05__6	Main road passable by a lorry: FEBRUARY
cs4q05__7	Main road passable by a lorry: MARCH
cs4q05__8	Main road passable by a lorry: APRIL
cs4q05__9	Main road passable by a lorry: MAY
cs4q05__10	Main road passable by a lorry: JUNE
cs4q05__11	Main road passable by a lorry: JULY
cs4q05__12	Main road passable by a lorry: AUGUST
cs4q05__13	Main road passable by a lorry: Passable the whole year
cs4q11	Is the community in a major urban centre (regional or zonal capital)?
cs4q14	Is there a large weekly market in this community?
cs4q19	Are all of the classrooms built of brick/stone with iron sheet roofs?
cs4q20	Is the nearest government primary school electrified?
cs4q22	Is the nearest government secondary school electrified?
cs4q27	Is there a place in this community to purchase common medicines?
cs4q29	Is there a health post in this community?
cs4q31	Nurse, midwife or trained health extension agents permanently working

Table A3 continued from previous page

Variable	Description
cs4q32	Who runs this health post?
cs4q33	Is this health post electrified?
cs4q34	Is there a hospital/health center/clinic in this community?
cs4q38	Who runs the facility where the nearest medical doctor is located?
cs4q39	Are there any groups or programs in this community providing insecticide?
cs4q41	Any groups providing support and care to people who are chronically ill
cs4q43	Is there a commercial bank in this community?
cs4q47	Is there water service in the community?
cs4q50	Is there an ATM in this community?
cs4q52	Is there a SACCO in this community?
cs4q54	Is there a Bank Agent in this community?
cs4q56	Is there an Insurance Branch in this community?
cs4q58	Has the PSNP program been operational during the past 12 months?
cs5q01_1	Most important sources of employment for individuals in the community (MOST)
cs5q02	Do people leave temporarily to look for work elsewhere?
cs5q06	Do people come to this community to look for work?
cs5q09	Is there a cooperative to create opportunities for work?
cs6q01	Do any households farm crops or keep livestock in this community?
ssa_aez09	Agro-ecological Zones
sq1	Nutrient availability
sq2	Nutrient retention capacity
sq3	Rooting conditions
sq4	Oxygen availability to roots
sq5	Excess salts
sq6	Toxicity
sq7	Workability (constraining field management)
s7q12	How many of your own oxen were used in this Meher season?
s7q14	How many oxen do you have?
s7q19	Total Chemical Fertilizers purchased for main season in 2011 E.C
s7q20	Total Dap fertilizer ready for use for main season in 2011 E.C
s7q21	Total Urea fertilizer ready for use for main season in 2011 E.C
s7q22	Total NPS fertilizer ready for use for main season in 2011 E.C
s7q23	Total other chemical fertilizer ready for use for main season in 2011 E.C
s7q32a_1	During this agriculture season, how much quantity of pesticide was used?
s7q32b_1	During this agriculture season, how much quantity of herbicide was used?
s7q32c_1	During this agriculture season, how much quantity of fungicide was used?
parcel_id	Unique Parcel Identifier
field_id	Unique Field Identifier
crop_id	Unique Crop Identifier
s4q11a	What was the quantity of Seed / Seedling used for [CROP] on [FIELD]?
s4q21a	How much of [CROP] do you expect to harvest from [FIELD]?: Quantity
s4q21b	How much of [CROP] do you expect to harvest from [FIELD]?: Unit
s3q28	For the current season, how many HH members worked on [FIELD]?
s3q30a	Hired Men (Number of Men)

Table A3 continued from previous page

Variable	Description
s3q30d	Hired Women (Number of Women)
s3q30g	Hired Children (Number of Children)
s3q31a	Other HH Labour (Number of Men)
s3q31c	Other HH Labour (Number of Women)
saq12	Household Size
saq16	What is the holder's education Level
s1q03a	How old is [NAME]? (COMPLETED YEARS)
s1q06	For how many weeks during the last 12 months was [NAME] away?
s4q03a	Total time [NAME] spent fetching water for use by HH yesterday (HOURS)
s4q03b	Total time [NAME] spent fetching water for use by HH yesterday (MINUTES)
s4q04a	Total time [NAME] spent collecting firewood for use by HH yesterday (HOURS)
s4q04b	Total time [NAME] spent collecting firewood for use by HH yesterday (MINUTES)
cs3q02	What is the population of this community?
cs3q03	How many households are found in this community?
cs3q04b	Approximately how many households practice?
cs3q06	What percentage of households within this community are polygamous?
cs3q09	Percentage of the land in this community in bush
cs3q10	Percentage of the agricultural land in this community in large scale farms
cs3q11	Percentage of the land in this community in forest, and not used for agri.
cs3q12b_1	What percentage of HH are united through this type of marriage?
cs4q02	How far is it to the nearest tar/asphalt road?
cs4q06	How far is it to the nearest bus station?
cs4q07a	Expected frequency for a bus to stop at the nearest bus station (NUMBER)
cs4q16	How many churches (congregations) are there in this community?
cs4q17	How many mosques are there in this community?
cs4q18	Distance to the nearest government primary school serving this community
cs4q21	Distance to the nearest government secondary school serving this community
cs4q23	Number of primary schools run by religious organizations in this community
cs4q24	Number of secondary schools run by religious organizations in this community
cs4q25	How many private primary schools are there in this community?
dist_border	HH Distance in (KMs) to Nearest Border Crossing
dist_admhq	HH Distance in (KMs) to Capital of Region of Residence
twi	Potential Wetness Index
af_bio_1	Annual Mean Temperature (degC * 10)
af_bio_8	Mean Temperature of Wettest Quarter (degC * 10)
af_bio_12	Annual Precipitation (mm)
af_bio_13	Precipitation of Wettest Month (mm)
af_bio_16	Precipitation of Wettest Quarter
slopepct	Slope (percent)
srtm1k	Elevation (m)
popdensity	2018 Population density per km2
cropshare	2018 Percent cropland in local area
h2018_tot	12-month total rainfall (mm) in 2018
h2018_wetQstart	Start of wettest quarter in dekads 1-36, where first dekad of 2018 =1

Table A3 continued from previous page

Variable	Description
h2018_wetQ	Total rainfall in wettest quarter in 2018
h2019_tot	12-month total rainfall (mm) in 2019
h2019_wetQstart	Start of wettest quarter in dekads 1-36, where first dekad of 2019 =1
h2019_wetQ	Total rainfall in wettest quarter in 2019
anntot_avg	Avg annual total rainfall (mm)
wetQ_avgstart	Avg start of wettest quarter in dekads 1-36, where first dekad of year =1
wetQ_avg	Avg total rainfall in wettest quarter (mm)
h2018_ndvi_avg	Average NDVI value in primary growing season (highest quarter) in 2018
h2018_ndvi_max	Maximum dekadal NDVI value in primary growing season (highest quarter) in 2018
h2019_ndvi_avg	Average NDVI value in primary growing season (highest quarter) in 2019
h2019_ndvi_max	Maximum dekadal NDVI value in primary growing season (highest quarter) in 2019
ndvi_avg	Long-term average NDVI value in primary growing season (highest quarter)
ndvi_max	Long-term maximum dekadal NDVI value in primary growing season (highest quarter)
plot_twi	Plot Potential Wetness Index
plot_srtm	Plot Elevation (m)
plot_srtmslp	Plot Slope (percent)

Notes: Selected variables and further candidate variables for PDS LASSO model, Table 2 (4).

Table A4: Regression results for nitrogen (kg/ha) across different purity thresholds

Purity threshold	**OLS**			**PDS-LASSO**		
in %	**Belief**	**DNA**	**Belief x DNA**	**Belief**	**DNA**	**Belief x DNA**
0.70	55.62*	-46.59**	53.27	53.58	-31.65	36.38
	(31.94)	(20.42)	(36.06)	(35.93)	(19.73)	(33.30)
0.80	57.22*	-44.67**	51.48	55.57	-29.34	34.15
	(31.40)	(19.45)	(35.68)	(35.23)	(18.78)	(32.93)
0.85	63.93**	-35.41**	47.40	62.64**	-19.42	28.42
	(27.45)	(15.43)	(34.20)	(31.34)	(14.81)	(31.37)
0.90	68.20***	-28.64**	45.13	67.94**	-12.65	21.08
	(23.30)	(13.16)	(33.20)	(27.52)	(12.21)	(29.34)
0.925	128.12***	-3.49	-39.99	111.85***	-2.67	-43.22
	(44.65)	(7.37)	(33.67)	(45.97)	(7.05)	(30.16)
0.95	123.59***	18.87*	-50.59*	92.77***	10.28	-48.99**
	(36.89)	(10.25)	(26.43)	(34.16)	(8.06)	(24.79)
0.975	113.47***	24.90**	-48.07**	85.53***	11.82	–35.28
	(32.49)	(12.36)	(23.75)	(30.12)	(10.36)	(22.21)

Notes: The DNA indicator is 1 if the DNA test indicated improved genetic material at the respective purity threshold. Unreported ‘extended controls’ include ‘extension contact’, ‘seeds purchased’, field size, manure use, age, ‘has attended any school’, farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Amounts of nitrogen are simple aggregates of the respective nitrogen contents of three fertilizers: Urea = 46%, DAP = 18%, NPS = 10%. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * p <0.10; ** p <0.05; *** p <0.01. The set of variables selected by the LASSO and used in the OLS post-LASSO estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. See Chernozhukov et al. (2014) for details.

Table A5: Effective nitrogen use, seed beliefs and DNA type (intensive margin)

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Nitrogen use (kg per ha)		OLS		PDS LASSO	
Belief	64.74*	117.88***	106.81***	**93.58*** **	111.81***
(improved = 1)	(33.31)	(41.94)	(34.43)	**(31.53)**	(39.06)
DNA	16.29	20.00	18.14	0.46	**8.74**
(improved = 1, threshold at 95%)	(20.02)	(18.86)	(16.98)	(17.26)	(14.35)
Belief × DNA (TP = 1)	-66.56*	-50.53*	-32.68	-32.35	-43.48*
	(37.52)	(30.03)	(24.79)	(24.08)	(25.48)
Extension contact (yes = 1)		-36.12*	-14.92	-19.65	-28.90
		(20.71)	(21.05)	(19.20)	(18.82)
Seeds purchased (yes = 1)		-75.75**	-65.98**	-82.44*	-70.46**
		(32.83)	(25.42)	(44.38)	(26.47)
Field size (ha)		-196.90***	-155.09**	-160.41**	-169.99***
		(61.49)	(61.19)	(66.77)	(62.97)
Manure use (yes = 1)		8.61	-3.54	11.09	16.12
		(20.01)	(18.26)	(20.58)	(23.99)
OLS: Main controls (3)	no	yes	yes	yes	yes
OLS: Extended controls (12)	no	no	yes	yes	yes
PDS LASSO: No. of candidate controls				278	278
PDS LASSO: No. of selected controls				20	15
Observations	217	217	217	217	217
Adjusted R^2	0.04	0.18	0.23	0.22	0.21

Notes: The DNA indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%. Unreported 'extended controls' in model (3-5) include age, 'has attended any school', farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Amounts of nitrogen are simple aggregates of the respective nitrogen contents of three fertilizers: Urea = 46%, DAP = 18%, NPS = 10%. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * p <0.10; ** p <0.05; *** p <0.01. The set of variables selected in model 4 and 5 selected by the lasso and used in the OLS post-lasso estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. See Chernozhukov et al. (2014) for details.

Table A6: Effective nitrogen use, seed beliefs and DNA type (extensive margin)

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Nitrogen use (>0 kg)		OLS		PDS LASSO	
Belief	0.66***	0.57***	0.53***	**0.24*** **	0.29***
(improved = 1)	(0.06)	(0.07)	(0.08)	**(0.08)**	(0.08)
DNA	0.14	0.10	0.09	0.07	**0.05**
(improved = 1, threshold at 95%)	0.07	(0.06)	(0.06)	(0.04)	(0.04)
Belief x DNA (TP = 1)	-0.18	-0.18**	-0.14*	-0.12	-0.12*
	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)
Extension contact (yes = 1)		0.27***	0.27***	-0.07	-0.08
		(0.06)	(0.06)	(0.06)	(0.06)
Seeds purchased (yes = 1)		0.01	0.04	-0.05	0.03
		(0.06)	(0.07)	(0.06)	(0.05)
Field size (ha)		-0.10	-0.01	-0.20	-0.16
		(0.15)	(0.16)	(0.13)	(0.13)
Manure use (yes = 1)		-0.15***	-0.15***	-0.13**	-0.11**
		(0.05)	(0.05)	(0.05)	(0.05)
OLS: Main controls (3)	no	yes	yes	yes	yes
OLS: Extended controls (12)	no	no	yes	yes	yes
PDS LASSO: No. of candidate controls				288	288
PDS LASSO: No. of selected controls				22	21
Observations	432	432	432	432	432
Adjusted R^2	0.34	0.43	0.46	0.60	0.58

Notes: The DNA indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%. Unreported 'extended controls' in model (3-5) include age, 'has attended any school', farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Amounts of nitrogen are simple aggregates of the respective nitrogen contents of three fertilizers: Urea = 46%, DAP = 18%, NPS = 10%. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. The set of variables selected in model 4 and 5 selected by the lasso and used in the OLS post-lasso estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. See Chernozhukov et al. (2014) for details.

Table A7: Effective phosphorus use, seed beliefs and DNA type

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Phosphorus (kg/ha)		OLS		PDS LASSO	
Belief	64.58***	69.82**	64.48**	**50.65****	52.26**
(improved = 1)	(19.12)	(28.13)	(25.92)	**(23.89)**	(22.21)
DNA	11.60	11.20	8.35	10.04**	**3.96**
(improved = 1, threshold at 95%)	(7.41)	(7.40)	(6.60)	(5.03)	(5.94)
Belief × DNA (TP = 1)	-36.07*	-34.90*	-29.28*	-27.25*	-25.60
	(20.74)	(19.33)	(17.08)	(15.95)	(15.72)
Extension contact (yes = 1)		15.27**	11.70	-5.19	-10.33
		(7.42)	(7.88)	(7.59)	(9.72)
Seeds purchased (yes = 1)		-11.04	-11.61	-18.46	-8.83
		(14.48)	(13.48)	(22.66)	(12.55)
Field size (ha)		-61.59***	-61.59***	-57.96***	-59.31***
		(18.82)	(18.72)	(19.32)	(19.68)
Manure use (yes = 1)		-7.02	-6.96	-1.51	3.46
		(7.31)	(6.83)	(8.51)	(10.99)
OLS: Main controls	no	yes	yes	yes	yes
OLS: Extended controls	no	no	yes	yes	yes
PDS LASSO: No. of candidate controls				288	288
PDS LASSO: No. of selected controls				11	12
Observations	432	432	432	432	432
Adjusted R^2	0.14	0.18	0.22	0.26	0.26

Notes: The DNA indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%. Unreported 'extended controls' in model (3)-(5) include age, 'has attended any school', farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Amounts of phosphorus are simple aggregates of the respective phosphorus contents of two fertilizers: DAP = 46%, NPS = 42%. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * p <0.10; ** p <0.05; *** p <0.01. The set of variables selected in model (4) and (5) selected by the LASSO and used in the OLS post-LASSO estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. See Chernozhukov et al. (2014) for details.

Table A8: Total cost of purchased inputs, seed beliefs and DNA type

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Total cost of purchased inputs (ETB/ha)		OLS		PDS LASSO	
Belief	4503.86***	4626.12***	4117.12***	**2288.48****	2431.98**
(improved = 1)	(998.75)	(1161.10)	(969.10)	**(1010.63)**	(1023.01)
DNA	715.19	727.60	706.49	302.699	**367.46**
(improved = 1, threshold at 95%)	(501.51)	(520.56)	(460.38)	(399.50)	(413.96)
Belief $\times$ DNA (TP = 1)	-623.90	-628.52	-395.66	-236.52	-192.31
	(1219.29)	(1256.22)	(1096.56)	(1014.72)	(1044.30)
Extension contact (yes = 1)		569.67	724.41	447.83	530.77
		(593.83)	(508.39)	(494.68)	(486.55)
Field size (ha)		-5571.96***	-5358.58***	-6797.20***	-6097.14***
		(1609.93)	(1417.35)	(1637.49)	(1661.03)
Manure use (yes = 1)		-442.54	-532.06	-255.01	-413.34
		(524.12)	(513.69)	(518.58)	(567.54)
OLS: Main controls	no	yes	yes	yes	yes
OLS: Extended controls	no	no	yes	yes	yes
PDS LASSO: No. of candidate controls				287	287
PDS LASSO: No. of selected controls				13	7
Observations	431	431	431	431	431
Adjusted R^2	0.21	0.24	0.30	0.32	0.31

Notes: The total cost of purchased inputs in ETP per hectare is defined as the aggregate value per hectare of all of the maize seed, NPS, UREA, DAP, and all other inorganic fertilizer purchased. The DNA indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%. Unreported 'extended controls' in model (3)-(5) include age, 'has attended any school', farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * p <0.10; ** p <0.05; *** p <0.01. The set of variables selected in model (4) and (5) selected by the LASSO and used in the OLS post-LASSO estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. See Chernozhukov et al. (2014) for details.

Table A9: Effective Urea use, seed beliefs and DNA type

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Urea (kg/ha)		OLS		PDS LASSO	
Belief	201.53***	257.24***	236.00***	**176.04*** **	230.81***
(improved = 1)	(54.72)	(73.42)	(68.57)	**(63.82)**	(63.83)
DNA	27.84	34.61*	36.83*	16.68*	**36.21***
(improved = 1, threshold at 95%)	(18.73)	(20.51)	(18.97)	(16.05)	(19.58)
Belief $\times$ DNA (TP = 1)	-125.40**	-115.59**	-94.36*	-92.85*	-104.09**
	(60.17)	(54.92)	(49.08)	(47.89)	(48.77)
Extension contact (yes = 1)		-7.27	-7.12	-71.76***	-1.73
		(24.56)	(23.76)	(27.46)	(22.33)
Seeds purchased (yes = 1)		-73.17*	-58.32*	-105.77*	-64.91**
		(37.81)	(35.10)	(59.57)	(31.72)
Field size (ha)		-184.19***	-132.04**	-171.27***	-140.13***
		(61.70)	(53.27)	(57.27)	(51.76)
Manure use (yes = 1)		-4.10	-17.17	10.79	7.33
		(20.57)	(20.20)	(23.37)	(27.35)
OLS: Main controls	no	yes	yes	yes	yes
OLS: Extended controls	no	no	yes	yes	yes
PDS LASSO: No. of candidate controls				288	288
PDS LASSO: No. of selected controls				7	6
Observations	432	432	432	432	432
Adjusted R^2	0.16	0.20	0.25	0.29	0.25

Notes: The DNA indicator is 1 if the DNA test indicated improved genetic material at a purity threshold of 95%. Unreported 'extended controls' in model (3)-(5) include age, 'has attended any school', farm type, mobile phone ownership, and household distance to nearest market, major road, and population center. Clustered standard errors robust to heteroskedasticity across enumeration areas in parentheses. * p <0.10; ** p <0.05; *** p <0.01. The set of variables selected in model (4) and (5) selected by the LASSO and used in the OLS post-LASSO estimation and in the PDS structural estimation is augmented by an amelioration set ensuring that belief, DNA, and their interaction, as well as the extended controls enter the identifying model. See Chernozhukov et al. (2014) for details.

Table A10: Comparing DNA fingerprinting and non-fingerprinting subsamples from the ESS4 along key characteristics for main maize-growing regions

	(1)	(2)	(3)
	DNA subsample	ESS4 non-DNA	(1) - (2)
Selected variables (Table 1)			
Gender (female = 1)	0.16	0.17	-0.01
	(0.37)	(0.38)	(0.02)
Age (years)	46.61	44.79	1.82**
	(14.98)	(14.41)	(0.86)
Education (attended any school = 1)	0.39	0.35	0.04
	(0.49)	(0.48)	(0.03)
Extension contact (yes = 1)	0.58	0.53	0.05*
	(0.49)	(0.50)	(0.03)
Seeds purchased (yes = 1)ˆ	0.42	0.45	-0.03
	(0.49)	(0.50)	(0.03)
Land area (ha)	0.11	0.12	-0.01
	(0.14)	(0.20)	(0.01)
Nitrogen (kg/ha)	49.64	60.68	-11.04
	(99.14)	(319.66)	(15.76)
Manure use (%, yes = 1)	0.40	0.42	-0.01
	(0.49)	(0.49)	(0.03)
PDS LASSO selected variables			
Type of crop sowing technique	0.56	0.50	0.06**
(raw (broadcast) planting = 0 (1))	(0.50)	(0.50)	(0.03)
Seed used left over from a previous season	0.51	0.47	0.03
(yes = 1)	(0.50)	(0.50)	(0.03)
Total NPS fertilizer ready for use	28.95	44.30	-15.34***
for main season in 2011	(51.43)	(89.56)	(4.65)
Number of private schools in community	0.01	0.001	0.01***
	(0.12)	(0.03)	(0.004)
Plot under extension program	0.42	0.35	0.07**
in current season (yes = 1)	(0.49)	(0.48)	(0.03)
Regional percentage shares			
Tigray	0.19	0.11	0.07***
Amhara	0.29	0.35	-0.07**
Oromia	0.19	0.25	-0.06***
SNNP	0.22	0.20	0.02
Harar	0.12	0.09	0.03*
Observations	431	877	

Notes: Interaction terms selected by PDS LASSO excluded.

Table A11: Performance of SuperLearner and candidate models

Panel A: Cross-validated in-sample performance	**Accuracy**	**AUC**	**Precision**	**Recall**
Ensemble model (SuperLearner)	0.82	0.84	0.73	0.71
XGBoost	0.82	0.84	0.71	0.73
Random Forest	0.80	0.87	0.72	0.63
GLM Net	0.75	0.76	0.64	0.56
Bagged Classification Trees	0.81	0.84	0.74	0.63
Panel B: Performance of ensemble models for alternative outcomes	**Accuracy**	**AUC**	**Precision**	**Recall**
Extension program participation (yes = 1), DNA subsample	0.83	0.92	0.77	0.82
Extension program participation (yes = 1), ESS 4	0.88	0.87	0.81	0.85
Seed source (purchased = 1), DNA subsample	0.90	0.95	0.92	0.84
Seed source (purchased = 1), ESS4	0.92	0.92	0.93	0.90

Notes: Panel A displays performance metrics for the underlying prediction model and candidate algorithms. Panel B illustrates the out-of-sample performance of ensemble models for other agricultural outcomes for which the actual outcomes are observed for the complete ESS4. In-sample performance for these models are reported next to out-of-sample results to evaluate the ability to predict agricultural outcomes accurately, and to test the extend to which these models are overfitted to the subsample on which the cross-validated training has been done.