NBER WORKING PAPER SERIES

AUTONOMY AND TECHNOLOGY ADOPTION

Benjamin Bushong
Carolina Corral
Xavier Giné
Aprajit Mahajan
Enrique Seira

Working Paper 27681
http://www.nber.org/papers/w27681

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2020, Revised July 2025

We are grateful to Qué Funciona para el Desarrollo (QFD) for managing the field experiment, to Fertilab for the soil analyses, to Agropecuaria Amozoc for providing and mixing fertilizers and to Ipampa S.C. for the provision of agricultural extension services. We would also like to thank Ned Augenblick, Pranab Bardhan, Chico Ferreira, Andy Foster, Supreet Kaur, Berk Özler and Chris Udry for very helpful comments and suggestions. We also thank seminar participants at various workshops for their comments and suggestions. Michelle Infanzón, Meritxell Martínez, Victor Manuel Pérez, Bernardo Ribeiro, Mónica Vargas and Victor Vergara provided outstanding research assistance. We also gratefully acknowledge funding from 3ie, BASIS (UC Davis), the World Bank and ISPC-SPIA (under the grant “Strengthening Impact Assessment in the CGIAR System (SIAC)"). Views expressed in this paper are those of the authors, and do not necessarily reflect the opinions of the World Bank, its executive directors, or the countries they represent. The field experiment was reviewed and approved by the ITAM IRB and UCLA IRB # 12–001756 and is registered in the AEA RCT Registry with the unique identifying number AEARCTR-0006264. The lab experiment was reviewed and approved by the MSU HRPP # 00009114. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Autonomy and Technology Adoption
Benjamin Bushong, Carolina Corral, Xavier Giné, Aprajit Mahajan, and Enrique Seira
NBER Working Paper No. 27681
August 2020, Revised July 2025
JEL No. D01, O33, Q12

ABSTRACT

We examine whether autonomy increases adherence to expert recommendations in technology adoption. In a context where farmers overuse fertilizer, we ran a field experiment that combined recommendations with either a restrictive subsidy tied to expert-recommended inputs or a flexible subsidy preserving farmer autonomy over input choice. In the short run, farmers adopted expert recommendations at similar rates regardless of subsidy autonomy and reduced fertilizer over-use by two-thirds. In the longer run, after the intervention ended, farmers with autonomy were significantly more likely to persist with the expert recommendations. We replicate these findings in a complementary laboratory experiment and find suggestive evidence that autonomy increases persistence by improving recommendation recall. Our results suggest that preserving choice can enhance the long-term effectiveness of expert advice and targeted subsidies.

Benjamin Bushong
Michigan State University
bbushong@msu.edu

Carolina Corral
Social Impact, Inc.
carocorral@gmail.com

Xavier Giné
The World Bank
xgine@worldbank.org

Aprajit Mahajan
University of California, Berkeley
and NBER
aprajit@gmail.com

Enrique Seira
University of Notre Dame
Department of Economics
enrique.seira@gmail.com

A randomized controlled trials registry entry is available at AEARCTR-0006264

1 Introduction

Technology adoption is critical for closing the productivity gap between low and high income countries.[1] Policymakers often use tightly controlled subsidies to encourage technology adoption or shape its use.[2] While such stringent subsidies can increase short-term compliance, evidence from psychology and economics suggests that such restrictions could stifle longer-term persistence by reducing the beneficial effects of free choice—that is, *autonomy*.[3]

We conducted two complementary experiments to examine this tension. First, a field experiment randomized Mexican farmers into four groups: a control group, a group receiving expert localized input recommendations, and two groups receiving subsidies in addition to the expert recommendations. In one subsidy group, farmers could only use the subsidy towards expert-recommended inputs and services (designed to correct fertilizer overuse). In the other, farmers retained autonomy to choose the inputs and services to purchase with their subsidy. This design allows us to cleanly identify whether preserving autonomy in this manner undermined or enhanced adoption both in the short term and in the longer run. We then implemented a lab experiment tied to key features of the first study to overcome some of the limitations of the field setting and explore proximate mechanisms. The lab experiment varied autonomy in the choice of an expert recommended method for a real-effort task while holding fixed preferences over production methods.

Both experiments confirm the importance of autonomy. In the field experiment short-term adoption rates were similar in both subsidy groups. Strikingly, however, farmers with autonomy were much more likely to persist with the expert recommendations up to two years after the intervention ended. Similarly, in the lab experiment, participants with autonomy over the task method were much more likely to persist with the method in a later unrestricted choice session. The lab experiment also provides suggestive evidence that autonomy increased persistence by improving recall of expert recommendations. Taken together, both experiments suggest an important role for autonomy in technology adoption. We next describe each experiment in greater detail.

Our field experiment was motivated by the fact that inappropriate technology use can reduce productivity and have significant negative consequences.[4] A striking example is the widespread overuse of fertilizer in many countries leading to unnecessary greenhouse gas emissions.[5] Accordingly, we study

[1] This productivity gap is particularly concerning in agriculture, the primary income source for approximately 500 million households (Lowder et al., 2021), representing a significant portion of the world's poor. See Foster and Rosenzweig (2010); Macours (2019); Magruder (2018) for comprehensive reviews on technology adoption, including discussions on the role of subsidies—the typical rationales for which are market failures or learning.

[2] A distinct literature—see, e.g., Dupas (2014); Fischer et al. (2019); Meriggi et al. (2021)—evaluates the role of short-term subsidies in longer run adoption. This literature typically focuses on the price elasticity effects of subsidies for new (rather than familiar) technologies, and does not examine the role of subsidy stringency or scope as considered here.

[3] The literature identifies several mechanisms through which autonomy might matter. Restrictions on choice can generate reactance and reduce engagement (Brehm, 1966; Bruns and Perino, 2023; Rosenberg and Siegel, 2018; Stolper and Walter, 2017). Self-determination theory asserts that autonomous choice increases engagement, positive affect, learning, and perseverance (Black and Deci, 2000; Deci and Ryan, 1985; Vansteenkiste et al., 2004). Finally, choice-induced-preference theory posits that the act of choosing causes agents to prefer the selected item and increases subsequent engagement (Brehm, 1956; Coppin et al., 2010; Gerard and White, 1983; Sharot et al., 2009).

[4] See e.g. Moscona and Sastry (2025) for a discussion of how agricultural technologies developed for conditions in high-income countries can be inappropriate when transferred to other contexts, contributing to global agricultural productivity differences.

[5] Globally, approximately three-fifths of applied nitrogen and half of applied phosphorus are thought to be in excess of crop

the role of autonomy in improving the use of an existing technology in an environment with widespread inappropriate use—agricultural practices among smallholder farmers. In collaboration with experts, we developed locally appropriate agricultural practices and designed temporary subsidies that varied the degree of recipient autonomy. We then examined whether the degree of autonomy in subsidy design shaped both immediate compliance with and long-term adherence to these expert recommendations.

The field experiment randomized 540 maize farmers in Tlaxcala (Mexico) to a control arm and three treatment arms. Farmers in all treatment arms received a detailed soil analysis, a set of expert (agronomist) tailored input recommendations based on the soil analyses, and a package of agricultural extension services to support their implementation. In addition, two of the three treatment arms were offered an in-kind one-season subsidy of $2,000$ MXN pesos (roughly USD 150 at the time) for use with our partner agro-dealer. Farmers with the inflexible subsidy could only apply the subsidy (or grant, we use the terms interchangeably) towards the expert recommendations and we refer to this group as the No Autonomy group. In contrast, farmers with the flexible subsidy could use it regardless of whether they followed the expert recommendations. In particular, they could use the subsidy to purchase any fertilizer from the same agro-dealer, did not need to use a precision sower or fertilize at sowing (the subsidy financed expert recommendations). We refer to this as the autonomy arm.[6] Since both arms were otherwise identical, comparisons between them allow us to infer the effect of farmer autonomy. A third treatment arm received localized soil analyses, expert recommendations, and extension services but did not receive a grant. Finally, a control group did not receive any intervention. These two additional arms allow us to test secondary domain specific hypotheses—specifically, the effect of tailored recommendations and extension services and the incremental impact of the subsidies.

Farmers implemented few of the expert recommendations at baseline—of the nine expert recommended practices only three were used by a majority, with the remainder being relatively uncommon (with implementation rates varying from 0% to 15%).[7] Most strikingly, fertilizer use rates were far from recommended dosages. Farmers used 92% more urea on average than the recommended amount and about 164% more diammonium phosphate (DAP), while using only 31% of the recommended potassium chrolide (KCl). These over-use rates are consistent with the country-level estimates in West et al. (2014) and suggest substantial greenhouse gas emissions without additional economic benefits. In particular, despite the high fertilizer use, yields were low at 2 metric tons per hectare (mt/ha).[8] The expert recommendations thus specified a substantial reduction in urea and DAP use and an increase in KCl. In addition to improved fertilizer use, the recommendations also advocated the use of a precision sower

requirements, particularly in countries like India, China and Mexico (see, e.g., Millar et al., 2018; Vitousek et al., 2009; West et al., 2014; Wu et al., 2018; Wuepper et al., 2020).

[6]The experimental variation induces autonomy in a relatively restricted sense. While this narrowly defined autonomy may seem *apriori* weak, its adequacy is fundamentally an empirical question. In addition, such constrained choice sets are common in many domains (e.g. in work settings and targeted subsidy programs).

[7]The three commonly implemented practices were: (1) ploughing, (2) using inorganic fertilizer and (3) covering the fertilizer. The six uncommon practices were: (4) deep tilling (ripping), (5) using hybrid seeds, (6) fertilizing at sowing, (7) sowing with precision machinery, (8) using pre-emergent herbicide and (9) using high-quality fertilizers. See Tables 2 and OA-5 for more details on practice rates. We classify practices (4)-(9) as "new" since they were uncommon at baseline (see Bloom et al., 2013) for a similar classification of practices for a management intervention.

[8]Almost three-quarters of Mexican farmers report using chemical fertilizers (authors' calculations from INEGI, 2007) and the figure was 98% for our study sample. However, maize yields for rain-fed farms remain relatively low at only about 2–3 mt/ha. By comparison, rain-fed maize yields in the United States are approximately 8 mt/ha (Fernández et al. 2012; Sweeney et al. 2013).

that combined sowing with optimized fertilizer placement at sowing.

We have three primary findings from the field study. First, while farmers with autonomy *did* exercise it, overall recommendation adoption rates were similar regardless of autonomy—an increase of 2.5 (1.98σ) additional practices.[9] Strikingly, both farmers with and those without autonomy reduced fertilizer use at similar rates. Farmers without autonomy reduced the (absolute) urea, DAP and KCl gap by 71%, 80% and 88% respectively. The corresponding figures for farmers with autonomy were similar and we cannot reject the null of equality. Adherence to the input recommendations is perhaps unsurprising for farmers without autonomy, given the take-it-or-leave-it nature of the subsidy. However, farmers with autonomy could ignore the expert recommendations without consequence and, therefore, it is notable that they chose to follow them in large measure. A back-of-the-envelope calculation suggests that the reduction in urea alone lowered CO_2 equivalent (CO2e) emissions by 14%.

Second, farmers with autonomy were substantially more likely to persist with the new practices two years after the end of the subsidies. Relative to those without autonomy, farmers with autonomy implemented .45 ($.55\sigma$) more practices and .87 (1.07σ) more than the control group.[10] Finally, consistent with the recommendations' intent, productivity did *not* decrease despite lower fertilizer use—yields *increased* (by 13%–16%) in both subsidy arms relative to the control arm.

Despite internal validity, field experiments may involve unobserved interactions that complicate interpretations of treatment differences. For instance, extension services or plot-level heterogeneity could interact with autonomy in ways that obscure the causal attribution to autonomy. Our lab experiment addresses these concerns by providing a clean manipulation of autonomy in a controlled environment and can also begin to shed light on mechanisms.

The two-session real-effort lab experiment examined the adoption of a potentially helpful "technology" (or method) for building abstract Lego figures. We tasked participants with building as many abstract figures as possible in twenty minutes using a very large bag of assorted Lego blocks. This task was repeated in both sessions, mimicking two "agricultural seasons" in the field. The technology—which we named "Build Better"—provided specific steps for efficient task completion (e.g. initial block sorting, using a timer). Consistent with the field experiment, the key outcomes include the adoption of Build Better in the first session, persistence in its use in the second session, and performance (the number of completed figures) in each session.

The first session began with an explanation of Build Better and an elicitation of participants' preferred technology (Build Better or "free build"). We then informed participants that there was a chance that their preferred technology would be implemented or that a computer would instead choose the technology they would use in the building session. Our primary analysis compares participants who chose Build Better and were randomized into the no computer choice group (i.e. the "Autonomy group") to those who likewise chose Build Better but who were randomized into the computer choice group and for whom the computer (randomly) chose the Build Better method (the "No Autonomy group").[11] This

[9] The figure in parentheses represents treatment effects measured in standard deviations. See Table 2 for details on the construction.

[10] Unfortunately we were not able to collect fertilizer use or yields for 2017.

[11] For those randomized into computer choice, the computer chose Build Better with a very high probability. This design is related to that of (Dal Bó et al., 2010).

comparison conditions on both builder preferences (both groups preferred Build Better) as well as the assigned implementation method (both used Build Better) and only varies by how the assigned method was chosen: computer choice or participant choice.

The results of the lab experiment mirror those from the field. First, both groups adopted Build Better at a similar rate in the first session. Second, performance in the first session was similar, with both groups completing a similar number of figures. Finally, we measured persistence by analyzing behavior in the second building session—conducted one week later—when participants had complete freedom in their choice of method. We find that those assigned to the Autonomy group in the first session were more likely to keep using Build Better in this second session compared to builders assigned to the No Autonomy group. In addition, builders in the Autonomy group assembled approximately 15% more figures than the No Autonomy group.

The lab experiment also sheds light on some potential mechanisms. At the start of the second session, builders with autonomy were better at recalling individual components of Build Better, suggesting that autonomy may operate through improved recall. Self-reported task enjoyment was not higher among builders with autonomy, suggesting a limited role for choice-induced preference based theories in explaining increased adherence. Finally, we provide some additional evidence consistent with the view that builders (both with and without autonomy) engaged in a form of motivated reasoning, which suggests interesting avenues for future research.

We contribute to several strands of literature. First, we contribute to the emerging empirical literature on autonomy. One strand of this literature is lab based and focuses the effect of autonomy on task effort—Falk and Kosfeld (2006) and Sjöström et al. (2018) demonstrate that imposing constraints (reducing autonomy) on workers in a principal-agent framework reduces effort (see also de Rochambeau, 2021). Chaudhry and Klinowski (2016) examine the effect of autonomy on motivation and effort in a lab set-up similar to ours.[12] In a group context, Dal Bó et al. (2010) use a lab study to show that cooperation is greater when the rules are chosen democratically than when they are exogenously imposed. Similarly, in an observational setting Bardhan (2000) finds that Indian farmers are less likely to violate irrigation rules when they themselves have crafted those rules.[13] Older work on effective agricultural extension emphasizes the importance of dialogue and autonomy (see e.g. Freire, 1974) but does not offer an empirical test.[14] Our work is distinct in that it combines field-experimental evidence from a high-stakes setting (without a principal-agent problem) with a lab experiment and a focus on technology adoption and persistence.

[12]A related lab experimental literature tests whether autonomy has intrinsic values, demonstrating that in a principal-agent experiment, principals reliably place more weight on decision rights even though such rights have no effect on real (i.e., financial) outcomes (Bartling et al., 2014), see also Fehr et al. (2013). This is consistent with the finding that individuals are willing to forgo income to pursue careers where their agency has fewer restrictions (see e.g. Hamilton, 2000; Stern, 2004).

[13]Relatedly, Black and Lynch (2001) use a firm fixed-effects strategy to show that implementing formal programs (like Total Quality Management) yields little benefit unless a high proportion of employees are involved in regular decision-making at the plant.

[14]The paper is also related to "motivational interviewing", a counseling approach offering information and advice while respecting autonomy by e.g. seeking permission and acknowledging the freedom to disagree (Miller, 1983). Two recent field experiments in educational settings are also related in that while they do not directly manipulate autonomy their effects may operate through it. Alan and Kubilay (2025) find that entrusting adolescents with delivering a curriculum to their junior peers improves a range of school-related outcomes. Ashraf et al. (2020) find that training teachers to think and teach using the scientific method (posing sharp questions, framing specific hypotheses, using evidence and data gathered from everyday life) led to large improvements in test scores.

A second related literature is that on the use of expert advice more broadly (Bloom et al., 2013; Easterly, 2014; Haynes et al., 1996; O'Shea and Ueda, 2021; Ronayne and Sgroi, 2018; Weizsäcker, 2010; WHO, 2003). We contribute to this literature by highlighting the role of autonomy in the adoption and persistence of the use of expert recommendations.

Finally, we contribute to the literature on the impact of tailored input recommendations in agriculture (Cole et al., 2020; Gars et al., 2022; Harou et al., 2018; Murphy et al., 2019; Tjernström, 2017). We differ in two significant respects. First, we work in a setting with significant fertilizer overuse so that explanations based on liquidity or credit constraints are less compelling. Second, we pair our recommendations with high-quality extension services.[15] Our paper also relates to the nascent literature on improving nitrogen fertilizer use, widely viewed as key to achieving several of the Sustainable Development Goals (see e.g. Ladha et al., 2020; Lin et al., 2022; West et al., 2014), by demonstrating the value of customized recommendations and autonomy in improving the adoption of more sustainable agricultural practices.

The rest of the paper is organized as follows. Sections 2 to 4 describe the field experiment—Section 2 describes the context and data used while Section 3 provides the details of the design and rationale for the various experimental arms and describes the empirical strategy. Section 4 presents the experimental results, discusses mechanisms and alternative explanations. The latter motivate the lab experiment which is discussed in Section 5 and Section 6 concludes.

2 Field Experiment Context and Data

2.1 Recruitment, Farmer Characteristics and Baseline

The project was implemented in 13 municipalities in the Mexican state of Tlaxcala (Figure OA-1). We chose Tlaxcala because, like much of Mexico, it has substantial smallholder rainfed farming, relatively low maize yields (2.7 mt/ha on average), and high fertilizer use.[16] Three parties were responsible for implementation: our partner NGO "Qué Funciona para el Desarrollo" (QFD) which was in charge of field operations and data collection, Ipampa S.C., a long-standing local commercial extension service company, and Agropecuaria Amozoc, a commercial agro-dealer.

QFD began recruitment in January 2015 using poster displays in public locations, distributing informational leaflets, and holding a number of informational meetings in each municipality. In the meetings, QFD explained the intervention, eligibility requirements, and the enrollment lottery (i.e. randomization). To be considered eligible for the program, farmers had to be aged between 18 and 70 and plan to cultivate maize in 2015 on at least one hectare of land (owned or rented) but no more than 15 hectares.

The research team visited interested farmers between February and March 2015. The team collected detailed baseline information on a range of farmer characteristics and agricultural practices during the previous (2014) growing season. After the survey, farmers were asked to register a subplot of one hectare (where they planned to grow maize) for the program. QFD marked off this subplot, recorded its GPS coordinates, and collected soil samples. Our study sample consists of 540 farmers for whom we have base-

[15] Lin et al. (2022) and Cui et al. (2018) examine tailored recommendations on fertilizer overuse in China but do not study autonomy.

[16] Table OA-3 compares our study sample to respondents from the nationally representative INEGI survey.

line and follow-up data.[17] Appendix A provides details and summary statistics for the study sample. The average annual income (self-reported) is 29,414 MXN pesos (US$ 2,200 at the time), substantially lower than the average annual income of rural households of 92,194 pesos at the time (INEGI, 2014). Farmers cultivate about two plots and an average total area of 5.8 hectares. In 2014 (the year before the intervention) fertilizer use in the registered plot was near universal (98%), with farmers carrying out an average of 1.6 fertilizations though only 6% fertilized at sowing (a key new practice). Average self-reported yields were about 2 mt/ha, and about half of the sample sold maize in the market. Only 6% had used agricultural extension services in the past, and only 15% had ever paid for a soil analysis (see Table OA-3 for a comparison with other farmers in Mexico).

2.2 Follow up Surveys

We followed farmers from 2015 to 2017, and to minimize recall bias, attempted to field surveys right after farmers completed production activities (see Appendix A for a timeline). The first mid-line survey was carried out in August 2015 and focused on labor inputs and agricultural practices during the growing season. We also collected administrative data on fertilizer purchases from our partner agro-dealer during this time. A second mid-line survey was fielded in October 2015, just before the harvest, to record agricultural activities since the first mid-line. In January 2016, we collected yield data for the 2015 growing season. Finally, we collected further information on grain sales from the 2015 harvest in June 2016. In May 2017, two years after the experiment, we carried out a final end-line survey to collect information on practices for the 2017 growing season, which allowed us to measure persistence. Due to budget constraints, we could not collect data on yields and fertilization practices for 2017.

3 Experimental Design and Hypotheses

Farmers were divided into 26 strata based on their location and agro-climatic conditions. Individual randomization was done within strata and announced after the baseline survey in March 2015.[18]

3.1 Experimental Interventions

The experimental arms combined three components: (a) soil analysis with tailored recommendations, (b) extension services, and (c) in-kind grants (with and without autonomy). These were combined into three treatment groups:

[17] Appendix A discusses program implementation. We analyze attrition in Table OA-2 and find it is not differential across experimental arms. The study also included a fifth arm receiving recommendations based on individual plot soil analyses rather than community-averaged analyses. Since the autonomy arm received recommendations based on community-averaged analyses, we focus our analysis on treatment arms receiving equivalently formulated recommendations. See Corral et al. (2020) for results with the arm receiving recommendations based on individual plot-level analyses.

[18] Specifically, study farmers came from 54 localities (*localidades*) of which 29 were quite small and thus merged with the closest *localidad* using the distance between centroids from INEGI's geographic databases. If there were ties, we chose to merge *localidades* with the closest altitude. Two of the small *localidades* were merged into one to give us a total of 26 strata. The median stratum had 21 farmers, the 25th percentile had 15, and the 75th had 32. While individual-level randomization could generate spillovers across farmers with different treatments within the same stratum, we do not find evidence for this (see details in Appendix E.4).

Autonomy group (A). Farmers received a soil analysis, tailored input recommendations, extension services and an in-kind grant that allowed farmers to choose the inputs.

No Autonomy group (NA). Farmers received a soil analysis, tailored input recommendations, extension services and a restricted in-kind grant that could only be used to purchase recommended inputs.

Recommendations group (R). Farmers received a soil analysis, tailored input recommendations and extension services.

Control group (C). Farmers did not receive any intervention during the experiment.

We included extension services in all treatment arms because an earlier pilot in the same area with a comparable sample suggested important complementarities between the soil analyses based recommendations and extension services. Farmers appeared to greatly value the ability to question and discuss the recommendations with extension agents. Grants were provided in-kind rather than in cash because organizations such as QFD are not allowed by law to disburse cash grants. We next discuss the components of the interventions in detail.

Soil analysis: The research team collected soil samples from the registered sub-plot for every study farmer. The samples were analyzed by Fertilab, a well-known soil testing laboratory in Mexico. Appendix B provides details of the soil analysis including the various micro- and macro-nutrients measured.

Tailored Input Recommendations: Farmers in all three treatment arms received a report containing soil analyses and fertilizer recommendations based on the average of soil analyses in their cluster (*localidad*). We provided averaged recommendations (instead of recommendations based on each farmer's individual program sub-plot) because these are cheaper, can be scaled up, and are commonly used in extension programs. The averaging was expressly conveyed to farmers. Control farmers received their (cluster) averaged recommendations in February 2016, after the intervention had ended.

The recommendations included the nutrient levels and the corresponding fertilizer dosages required to produce maize yields of 4.5 mt/ha under normal rain and temperature conditions. Recommendations were based on Fertilab's proprietary maize model that assumed that a certain quantity of N, P, K, and micronutrients were needed to reach a target yield per hectare. Figure OA-4 provides the delivery format and Appendix B.4 provides additional details.

Fertilab recommended fertilizer applications at sowing and 30 to 35 days after sowing. The recommended fertilizer for each application was provided by our third partner Agropecuaria Azomoc and produced by YARA, considered to be of higher quality than competing brands.[19] Based on focus group discussions, the research team designed an easy-to-read report containing a "shopping list" of the fertilizers needed in each application. The team and the extension agents were careful to explain the report and its underlying assumptions during in-person sessions with farmers (see Appendix B.4 for more details).

[19] This was confirmed in laboratory tests. YARA fertilizer samples were of higher quality, and the quality-adjusted price was actually lower compared to the business-as-usual fertilizer. See Appendix B.5 for details.

As noted in the introduction, baseline use of urea and DAP is quite high relative to recommendations. Table 1 shows fertilizer use before the year before the intervention (column 1) as well as our recommendations for 2015 (column 4) for the three main fertilizers, urea, DAP and KCl. Columns 2, 3, 5 and 6 show p-values for tests of randomization balance. On average, farmers used 92% more urea and 164% more DAP in 2014 than required by the recommendations. In contrast, they only used 31% of the recommended dosage of KCl (all differences are statistically significant). The recommended reductions in urea imply significant decreases in CO_2e emissions, as urea releases N_2O—a greenhouse gas 300 times more potent than CO_2. Finally, the total cost (per hectare) of the fertilizer recommendations (row 4) is similar to the investment actually made by farmers in 2014 (p-value 0.55).

Table 1: Fertilizer application in 2014 vs. average recommendations

	(1)	(2) Fertilizer application	(3)	(4)	(5) Average recomm. in 2015	(6)
	Control mean in 2014	p-value All T vs. C	p-value A vs NA	Mean All T	p-value All T vs. C	p-value A vs NA
Urea(kg/ha)	267.81	0.59	0.31	138.90	0.94	0.83
	(135.25)			(12.38)		
DAP(kg/ha)	49.41	0.79	0.62	18.67	0.91	0.67
	(65.27)			(16.73)		
KCL(kg/ha)	9.45	0.81	0.95	30.34	0.98	0.79
	(26.15)			(17.45)		
Cost(pesos/ha)	1,895.49	0.91	0.55	1,672.13	0.94	0.80
	(1,181.91)			(214.53)		
Observations	130			407		

This table reports amounts of fertilizers applied in 2014 and recommended in 2015, along with their costs. It also presents tests of randomization balance. The simple is comprised of the 540 farmers in the study sample (3 of which were unable to estimate this answer). All T refers to groups A, NA and R. Column 1 reports self-reported average amounts applied and the total cost of fertilizers in 2014 by the control group (standard errors in parentheses). In column 2, we regress each variable on a set of treatment dummies and report the p-value of an F-test that the three treatment indicator coefficients are jointly equal to zero and in column 3 the p-value of a t-test that the Autonomy group mean is equal to that of the No Autonomy group. Column 4 reports the average input recommendations and the corresponding input costs for all three treatment groups A, NA and R. In column 5 and 6 are analogous to columns 2 and 3, but focused on tests of equality regarding recommendations. We use administrative data on recommendations given to farmers and Baseline survey data collected in February 2014.

Recommended Practices: In addition to the fertilizer recommendations, the program recommended that farmers follow nine practices. Three of these were already common (referred to as existing practices) while six were not (referred to collectively as new practices). The three commonly implemented practices were: (1) ploughing, (2) using inorganic fertilizer and (3) covering the fertilizer. The six new practices were: (4) deep tillage (5% prevalence at baseline), (5) using hybrid seeds (5%), (6) using pre-emergent herbicide (2%), (7) fertilizing at sowing (9%), (8) sowing with precision machinery (10%), (9) using high-quality fertilizer (0%). See Tables 2 and OA-5 for more details on practice rates. Practices (7)-(9) are all linked to fertilizer use since a precision sower was recommended for carrying out sowing fertilization.

Extension Services: The extension package comprised three group training sessions and three plot visits by agricultural extension workers (AEWs). The first meeting introduced farmers to the precision sowing drill and covered the sowing recommendations. The second covered post-sowing fertilizer use and strategies to rectify nutrient deficiencies. The final meeting was held just before the harvest and emphasized field preparation. In addition, AEWs visited the program subplots thrice (before and after sowing and before harvest) and monitored nutrient deficiencies and other potential issues. AEWs were allocated by geography and balanced across treatment arms (results are unchanged when AEW fixed-effects are included). Farmers in all four groups were also informed about the high quality fertilizer agro-dealer and given its address and a map of its location.

Grants for the No Autonomy group: Since the expert recommendations differed dramatically from business-as-usual, focus group discussions suggested that a subsidy might be useful to encourage adoption. Farmers were provided in-kind grants worth 2,000 pesos (approximately US$150 at the time) in the form of vouchers (Figure OA-3)—this is roughly equivalent to the average fertilizer cost (per hectare) in 2014. Any costs above this amount had to be paid by farmers out-of-pocket.[20] The grant amount for both groups was directly transferred to our partner agro-dealer, who deducted it from the costs of each farmer's shopping list. The rental cost for the precision sower was paid directly to the machine owners.

Farmers in the No Autonomy group received a grant restricted exclusively to purchasing the expert recommended inputs from our designated provider and renting the precision sower. Each farmer received a pre-populated shopping list with the recommended amounts for each fertilizer type (see Figure 1). When farmers visited the agro-dealer they received two pre-mixed fertilizer packages prepared in proportions matching their specific recommendations—one package for sowing and another for post-sowing fertilization. This design ensured adherence to recommended fertilizer inputs.

The grant was applied sequentially to first cover the sowing costs (i.e. the rental cost of 800 MXN pesos for the precision sowing drill and the cost of the first fertilizer package). The remainder of the grant went towards the second fertilizer package. A farmer choosing not to rent the precision drill would forfeit the subsidy for that input. Analogously, a farmer could choose to forfeit either or both fertilization packages.[21]

Grants for the Autonomy group: Farmers in the Autonomy group were offered the same grant amount at the same agro-dealer, with the same unit costs as those in the No Autonomy group. The only difference was that farmers in the Autonomy group were not required to follow the expert recommendations. Specifically, farmers were not required to rent a precision sower (required for fertilization at sowing), and neither were they required to purchase the recommended inputs in pre-specified quantities. That is, the "Total Required" columns in Figure 1 were blank and farmers could fill them in with their chosen quantities with the understanding that QFD would contribute up to 2,000 pesos of their total costs. The implementation team made substantial efforts to ensure that farmers felt no pressure to fill in any specific quantities. Forms were filled without the presence of any research team members and handed directly from the farmer to the agro-dealer.[22]

We emphasize two key aspects of the grant design. First, farmers in both grant arms received the

[20] Farmers could not save the in-kind grant because the agro-dealer only provided the fertilizer packages for the 2015 season. In addition, farmers do not report saving fertilizer packages for subsequent seasons.

[21] This restrictive grant is most similar to the targeted input subsidy of many large-scale African input subsidy programs. See e.g. Carter et al. (2021); Giné et al. (2022) for randomized evaluations in Mozambique and Tanzania respectively and Jayne and Rashid (2013) for an earlier critical review of such programs.

[22] The agro-dealer was blinded to treatment assignment and had no financial (or other) incentive to alter farmer choices.

Figure 1: Shopping List

YOUR CHOICES

Rent sowing machine? ☐ Yes ☐ No

SOWING PACKAGE

	Fertilizer Dose (kg)				
Fertilizer	**Brand**	**Total Requested**	**Cost (per kg)**	**Total Cost**	**QFD share**
Precision sowing machin	YARA				
Urea (white)	YARA				
DAP (Black)	YARA				
Potassium Chloride (Red	YARA				
Micro-elements	AGROQUIMICA				
Total fertilizer expenses (per hectare)			**X pesos**		

FERTILIZATION AFTER SOWING (30-35 days after sowing)

	Fertilizer Dose (kg)				
Fertilizer	**Brand**	**Total Requested**	**Cost (per kg)**	**Total Cost**	**QFD share**
Urea (white)	YARA				
Potassium Chloride (Red	YARA				
Total fertilizer expense (per hectare)			**X pesos**		

This form was distributed to all grant farmers and reflects the key distinction between the Autonomy and No Autonomy arms. Panel A covers fertilizer use at sowing and Panel B covers fertilization after sowing. At the very top farmers checked a box if they wanted to rent a precision sowing machine (this was pre-filled with a check for farmers without autonomy). Unit costs (column 4) were pre-filled and identical for all farmers. The green column "Total Requested" in both panels was pre-filled for the No Autonomy arm, but blank for farmers in the Autonomy arm. The "Total cost" column was computed by multiplying total requested by unit costs. The last column, "QFD Share" is the peso amount covered by QFD, which covered costs at a 100% rate for each input until the 2000 pesos subsidy was exhausted (the sowing machine cost 800 pesos).

same amount to ensure income effects remained identical across the two groups. Second, the autonomy granted to farmers was circumscribed but substantive. It is circumscribed because it is limited only to the choice of inputs and sower in Figure 1. To the extent that this is "too little" autonomy we may be biased against finding effects. It is substantive in the sense that it is exercised over relatively high real-world stakes (as opposed to e.g. lab experiment manipulations).

3.2 Main Hypotheses

We focus on testing the effects of autonomy by comparing the Autonomy (A) to the No Autonomy (NA) arms. The pure control group (C) and the recommendation group (R) allow us to examine complementary questions of interest for environmental and agricultural economics. Our three primary hypotheses are:

Hypothesis 1. Autonomy reduces overall short-term compliance (A vs NA in 2015) Farmers may choose not to adopt the recommendations for a variety of reasons—e.g. choice inertia, risk aversion, or low trust in the recommendations. For farmers without autonomy, doing so would mean foregoing the subsidy. However, farmers with autonomy will not forfeit their grant if they choose to ignore the recommendations. Autonomy may thus result in lower take-up of the expert recommendations. Concerns that autonomy may lower adoption are often used to justify rigid subsidy program conditions.[23] We test if this concern has bite in our context.

[23]In many domains, including vaccination and education, for example, subsidies are conditional on following expert recommendations—sometimes to address externalities or for paternalistic reasons.

Hypothesis 2. Autonomy increases persistence (A vs NA in 2017) Autonomy has the potential to increase persistence in the use of recommendations. This is consistent with a number of theories from psychology (e.g. reactance, self-determination, choice-induced preferences) and could occur through several channels—increased attention or effort, deeper learning, increased positive affect or by limiting reactance.

Hypothesis 3. Autonomy increases performance (A vs NA in 2015) Autonomy can increase performance through at least two channels. First, by directly affecting input choices and practices. Second, autonomy could affect performance even after conditioning on inputs and practices. We cannot disentangle these two channels but can test whether their combined effect increases performance.

3.3 Secondary Hypotheses

We also test the following hypotheses:

Secondary Hypothesis 1. Reducing fertilizer use decreases yields. Study farmers expressed concerns that the recommended reductions in fertilizer use would decrease yields. Rejecting this hypothesis would be evidence of fertilizer overuse.

Secondary Hypothesis 2. The in-kind grant increases compliance (A or NA vs C) The effect of the in-kind grant could work through multiple channels, for example: (a) a direct income effect, (b) an endorsement effect (i.e. that the specific recommendations were not just cheap talk), or others. While we cannot separately identify the relative role of these (and other) mechanisms, we can test whether their combined effect increased adoption.

Secondary Hypothesis 3. Persistence with expert recommendations increases when paired with extension services (R vs C in 2017) Control farmers received their recommendations in 2016, so that by 2017 they differ from R farmers only in that the latter (a) received extension services (in 2015) and (b) received their expert recommendations a year earlier (2015 vs 2016). We hypothesize that the provision of extension services could increase recommended practice adoption by farmers in the R arm in 2017 relative to those in the control arm C. While the literature on the provision of tailored recommendations generally finds limited effects, the addition of high quality extension services may improve adoption rates.

3.4 Empirical Specification

We study the effects of our experimental interventions on the take-up of recommended inputs, practices and yields in 2015 and on the persistence of practices in 2017 (two years after the program ended). We use an ITT specification that uses the individual components of each arm as the independent variables.[24] These are (i) a *"Recommendation"* indicator equal to one if the farmer received tailored input recommendations and extension services (i.e. the farmer is in the NA, A, or R arms); (ii) a *"Grant"* indicator equal

[24] We could also have presented the results using indicator variables for each experimental group. The two are equivalent (we do not impose additional restrictions) and we are not subject to the critique in e.g. Muralidharan et al. (2019). We believe our formulation is easier to interpret. We can recover the overall impact of any treatment arm by combining the β-coefficients: the test that the coefficient on the NA indicator is equal to zero is equivalent to $\beta_R + \beta_G = 0$, that the coefficient on the A indicator is zero is equivalent to $\beta_R + \beta_G + \beta_A = 0$ and finally, and that the coefficient on the R indicator is zero is equivalent to $\beta_R = 0$. Appendix G presents the analysis using the treatment group indicators.

to one if farmer received the in-kind grant (i.e. in the NA or A arms) and (iii) an *"Autonomy"* indicator equal to 1 for farmers in the A arm. Our complete specification is:

$$Y_{it} = \beta_0 + \beta_R \text{Recommendation}_i + \beta_G \text{Grant}_i + \beta_A \text{Autonomy}_i + \alpha_s + \epsilon_{it} \tag{1}$$

where i denotes a farmer, Y is the outcome of interest and t is the time period of interest (2015 or 2017). We include randomization strata indicators α_s and compute robust standard errors. β_R compares outcomes for farmers in the no-grant arm relative to those in the control group and therefore measures the impact of the recommendations and agricultural extension services. β_G measures the incremental impact of the grant without autonomy compared to not receiving the grant. β_A measures the impact of autonomy by comparing farmers in the Autonomy group to those in the No Autonomy group, and will be the key parameter of interest for testing our primary hypotheses. In each table, we also provide p-values for the tests that the treatment effects of the No Autonomy and Autonomy groups are different from zero.

4 Results

4.1 2015 Practices

Table 2 presents ITT estimates for the adoption of recommended agricultural practices separately for new and existing practices. We aggregate practices for each type into two single indices to mitigate multiple testing concerns. The "Total Practices Applied" column simply counts the number of adopted practices, while the "Standardized Index" column subtracts the control mean from each observation and divides by the control standard deviation (for each element of the index as well as their sum). Table OA-9 reports the result for individual practices. While respondents self-reported practice adoption, these were also verified by the AEWs for some practices that allowed for it, mitigating concerns about experimenter demand effects.

New practices: Farmers who received the recommendations and extension services adopt an additional 0.34 practices over the 0.32 practices in the control group (col 3), equivalent to 0.32σ (col 4). Farmers who, in addition, were offered the grant without autonomy adopt 2.58 more practices (equivalent to 1.66σ), an almost seven-fold increase, and approximately the same increase as the 2.5 additional practices by farmers with autonomy (the latter two are not statistically distinguishable). Since autonomy did not reduce the adoption of practices on average, we reject Hypothesis 1. Given recommendations and a choice, farmers did not continue with their previous practices. Instead, they followed expert recommendations relatively closely. Columns 1 and 2 of Table 2 show that our interventions had no effect on existing practices, which is unsurprising as most farmers were already using all three practices.

While overall adoption rates were similar, adoption rates for individual practices varied between the Autonomy and No Autonomy groups so farmers did exercise autonomy. Appendix D shows that farmers with autonomy were 13 percentage points less likely to rent a precision sower (relative to those without autonomy) and somewhat more likely to use high-quality fertilizer (see Table OA-9).

Table 2: Practices 2015

	(1)	(2)	(3)	(4)
	Existing practices		All new practices	
	Total practices applied	Standardized Index	Total practices applied	Standardized Index
Recommendation (1=Yes)	0.04	0.06	0.34**	0.32**
	(0.08)	(0.13)	(0.11)	(0.14)
Grant (1=Yes)	0.04	0.12	2.24***	1.66***
	(0.08)	(0.11)	(0.14)	(0.16)
Autonomy (Yes=1)	-0.02	-0.01	-0.08	-0.07
	(0.07)	(0.08)	(0.14)	(0.19)
Observations	540	540	540	540
R-squared	0.07	0.07	0.60	0.38
Mean dep. var. control	2.38	0.00	0.32	0.00
SD dep. var. control	0.61	1.00	0.69	1.00
NA: $\beta_R + \beta_G = 0$	0.24	0.09	0.00	0.00
A: $\beta_R + \beta_G + \beta_A = 0$	0.32	0.09	0.00	0.00

Note: this table reports results on the agricultural practices performed by the farmers in our study in the 2015 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recomm_i + \beta_G Grant_i + \beta_A Autonomy + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata fixed effects and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Follow-up survey conducted in August 2015. The dependent variable in column 1 is a sum of recommidual dummies. Each dummy takes value of 1 if the farmer performed one of the so-called existing agricultural practices. In column 2, the dependent variable is the standardized index of the outcome in column 1, computed by standardizing each dummy recommidually, adding them all and standardizing the sum. We use the mean and standard deviation of the control group as reference for the standardized index. In columns 3 and 4, the dependent variables are analogous to the outcomes in columns 1 and 2, computed for the so-called new practices. The existing practices are: (a) ploughing, (b) using inorganic fertilizer and (c) covering the fertilizer. The new practices are: (a) deep tilling (ripping), (b) using hybrid seeds, (c) fertilizing at sowing, (d) sowing with precision machinery, (e) using pre-emergent herbicide and (f) using high-quality fertilizers. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.2 2015 Fertilizer Use

In Table 3 we examine fertilizer use and focus on recommendation compliance by examining the *absolute difference* between the amount of fertilizer applied and recommended for each of the three main fertilizers (urea, DAP, and KCl) as well as the total fertilizer used. We expect the intervention to reduce the gap between actual and recommended use.[25] Columns 1–3 report application at sowing, while columns 4–6 report total fertilizer application over the entire growing season.

Compliance with fertilizer recommendations: Focusing first on total fertilizer applied (cols 4–5), farmers in the Recommendations group (without a grant) did not change fertilizer use relative to control farmers. Farmers in the No Autonomy group show considerable change in fertilizer use. The gap for urea reduced by 77 Kg/ha (a 71% reduction relative to the mean control gap of 114.9 Kg/ha), the gap in DAP by 32 Kg/ha (a 80% reduction) and the gap in KCl by 28 Kg/ha (88%).

Strikingly, comparable reductions were also achieved by farmers with autonomy, and we cannot reject the null that the overall reductions are the same for both arms (and separately for urea and DAP).[26]

[25] We also measure this gap for the control group since we also have recommended amounts for them.

[26] The recommendations altered both the fertilizer mix and the timing of application (fertilization at sowing). Since control farmers carried out their first fertilization about 36 days from the time of sowing on average and did not fertilize at sowing (see Table OA-4 Panel B), columns 1 to 3 show a sowing fertilizer deficit among controls (38.7 Kg/ha for urea, 19.4 Kg/ha for DAP and 16.4 Kg/ha for KCl) at the time of sowing. Most farmers offered the grant used the precision drill and fertilized at sowing, thus reducing the recommendation gap at sowing relative to control. Farmers with autonomy reductions the gap by 28.7 Kg/ha of urea (a 74% reduction), 13.6 Kg/ha of DAP (a 70% reduction) and by 11.5 Kg/ha of KCl (a 70% reduction). These reductions are similar for farmers without autonomy. Farmers with autonomy apply 3.73 Kg/ha less of urea compared to farmers without autonomy thus increasing the gap between the amount of recommended and applied fertilizer (urea was

While columns 1–3 show that farmers with autonomy deviated more from the recommendations at sowing because fewer used the precision sower, they make up for this difference by applying more fertilizer later in the season. We thus reject Hypothesis 1 for overall fertilizer use as well.

CO_2e emissions: The closing of the gap for urea and DAP led to substantial *reductions* in fertilizer use in both grant arms. Control farmers used an average of 188 kgs/ha of Urea, similar to usage rates in the recommendations only group. In contrast, farmers with a grant reduced their average urea application by 36 kg/ha, a decrease of 19% (see Table OA-12). A back-of-the-envelope calculation shows that the reduction in fertilizer translates to a 119 kg/ha or a 14% reduction in CO2e emissions using emission factors from the IPCC (Penman et al., 2000) (see Appendix H for details). As we show in Section 4.3, this reduction in fertilizer use (and GHG emissions) did not reduce yields.

To summarize, autonomy did *not* reduce the adoption of expert recommendations as commonly feared (so we can reject Hypothesis 1). Combining expert recommendations and high quality extension services has mixed effects—we see no changes in fertilizer use relative to control farmers but a modest increase in new practice adoption ($.33\sigma$). Finally, the grant increased new practice adoption rates almost eight-fold, supporting Secondary Hypothesis 2.

4.3 2015 Yields

Column 1 in Table 4 presents results for yields on the registered subplot.[27] Yields in control plots were $2,360$ Kg/ha on average, and the provision of tailored recommendations and extension services did not increase yields; the β_R coefficient is 220 Kg/ha (a 9% increase relative to the control mean) but the estimate is not statistically significant at conventional levels. Receiving a grant, however, increased yields significantly, both in an economic and statistical sense. For farmers in the NA arm, yields rose by 300 Kg/ha (a 13% increase relative to the control mean, p=0.06). The corresponding figure for farmers in the A arm was 370 Kg/ha (a 16% increase, p=0.02). We cannot reject equality between the A and the NA arms.

We draw two conclusions from these results. First, yields did *not* decrease in either grant group despite substantial reductions in fertilizer use, thus rejecting Secondary Hypothesis 1—fertilizer use (and emissions) can be substantially reduced without decreasing yields. Second, yields were indistinguishable for farmers with and without autonomy, suggesting that autonomy does not increase yields, rejecting Hypothesis 3.

Columns 2-6 examine other outcomes, such as revenues and costs. Revenue (column 2) increases commensurately with yields, and the magnitude of the treatment effects is similar to that of column 1 since prices do not differ across groups.[28] Farming expenses (including agro-dealer purchases) increase

underused at sowing).

[27] Measuring yields is notoriously difficult (see e.g. Desiere and Jolliffe, 2017). To reduce error, we demarcated and geocoded 1 ha subplot and measured yields for that subplot. We verified self-reported yields by transporting the harvested grain to a nearby weighing station for a subset of plots. We obtain similar results from a range of different measures. See Appendix E, Table OA-13, Figure OA-5 and Corral et al. (2020) for details. In addition, rainfall was below normal after July 2015 in the study municipalities. This is a critical period for plant development, likely affecting yields, (see e.g. Sinclair and Rawlins, 1993) on the effects of drought timing on maize growth and Table OA-14 for details on rainfall during the study.

[28] We use the median localidad (farmer-level) price imputed from the quantity and value of the harvest asked during the

Table 3: Fertilizer usage in 2015: applied vs. recommended

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Absolute difference, applied vs. recommended							
	At sowing				Total			
	Urea (kg/ha)	DAP (kg/ha)	KCl (kg/ha)	All three (kg/ha)	Urea (kg/ha)	DAP (kg/ha)	KCl (kg/ha)	All three (kg/ha)
Recommendation of any type (1=Yes)	-1.83	0.52	-1.47	-2.55	-4.83	1.77	-0.93	-4.20
	(1.97)	(2.37)	(1.28)	(4.05)	(10.87)	(5.73)	(3.79)	(14.59)
Grant (1=Yes)	-30.59***	-16.81***	-13.62***	-60.32***	-77.02***	-31.98***	-27.85***	-135.33***
	(2.00)	(2.25)	(0.86)	(3.69)	(9.40)	(4.97)	(2.69)	(12.58)
Autonomy (Yes=1)	3.73**	2.70*	3.58***	10.18**	2.74	-1.55	5.15**	7.32
	(1.89)	(1.62)	(1.07)	(3.67)	(7.59)	(3.40)	(2.32)	(10.27)
Observations	532	532	532	540	532	532	532	540
R-squared	0.52	0.39	0.44	0.50	0.25	0.21	0.28	0.35
Mean dep. var. control	38.69	19.40	16.44	73.38	114.90	37.83	32.66	182.53
SD dep. var. control	14.64	21.36	14.42	35.27	82.62	47.92	33.98	118.60
NA: $\beta_R + \beta_G = 0$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A: $\beta_R + \beta_G + \beta_A = 0$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: This table reports results on the usage of fertilizers by the farmers in our study in the 2015 season, compared to the recommended dosages. Using the sample for the full sample of 540 farmers, except for 13 farmers for which we do not have data on usage of fertilizers, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recomm_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata fixed effects and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Commercialization survey conducted in June 2016. In columns 1-3, the dependent variables are the absolute differences between the amount applied by farmers at sowing and the recommended dosages of Urea, DAP and KCl, respectively. In column 4, we report the absolute difference of the addition of the amount of fertilizers applied at sowing. In columns 5-7, we report analogous outcomes for the total amount of fertilizers applied in the full season. In column 8, we report the absolute difference of the addition of the total amount of fertilizers applied in the full season. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

with the grant, although not one-for-one. We find no effects on our profit measure (see Appendix E.3 for details on variable construction) after subtracting the grant amount (column 4).[29]

4.4 2017 Practices

Using data collected in May 2017, two years after the intervention, we test whether autonomy increases persistence (Hypothesis 2). The survey took place after sowing and asked about a range of outcomes, including practices in the 2017 season. Unfortunately, we were unable to obtain fertilizer quantities (fertilizer application had not been completed at the time of the survey), so we focus on adherence to the set of new practices introduced by the intervention.

Table 5 examines practices in 2017. Farmers in the control group reported using 0.42 new practices (column 3), statistically indistinguishable from the 0.32 new practices used by the same group in the year of the intervention (see Table 2). This suggests that the provision a year later of only the soil analyses and recommendations (albeit without extension services) did little to change practices, and that spillovers from treatment farmers were likely minimal. Farmers in the R group (who had received extension services during the intervention) adopted 0.24 additional practices, an increase of 0.4 σ (column 4) providing evidence in favor of Secondary Hypothesis 3. Thus, relative to farmers in the control group, a one-time provision of extension services had persistent effects into the second growing season.

The grant for the NA group did not increase persistence beyond that of the R group (the coefficient is an insignificant increase of 0.08 practices). However, farmers with autonomy demonstrate substantially higher persistence with recommended new practices, implementing 0.77 more practices (a 1.07σ increase) than the control group. This effect is substantively larger than that of the NA arm and, to our knowledge, is the first experimental evidence of autonomy on persistence in a field setting.

We explore this result in greater detail by examining each individual practice in Table OA-10. In 2017, farmers with autonomy were more likely than those without it to use hybrid seeds, sow with a precision drill, and use YARA fertilizers. While the rental of the precision drill and YARA fertilizers were subsidized by the grant, hybrid seeds, while recommended, were not.

Table 6 summarizes the results for our primary and secondary hypotheses:

4.5 Discussion

Why did autonomy increase adherence? We discuss three specific mechanisms from the literature.

Memory: Farmers with autonomy were $5pp$ to state that they remembered the sowing recommendations relative to farmers with the inflexible grant (column 1 of Table 7).[30] This does not seem to be driven

survey.

[29]Power calculations suggest that we have 83% power to detect a 20 percent increase in mean yields, but only 9% power to detect a 20 percent increase in profits net of the subsidy.

[30]This is consistent with some work in psychology and neuroscience e.g. Ding et al. (2021); Mirty et al. (2019); Schneider et al. (2018)

Table 4: 2015 Yields and profits

	(1) Self-reported yields (t/ha)	(2) Revenue (Mex$/ha)	(3) Costs (Mex$/ha)	(4) Profits (Mex$/ha)	(5) Profits (no subsidy) (Mex$/ha)	(6) Profits (no subsidy) - Alt. (Mex$/ha)
Recommendation (1=Yes)	0.22	744.96	48.19	696.77	705.94	705.09
	(0.16)	(528.22)	(276.10)	(517.75)	(518.38)	(517.62)
Grant (1=Yes)	0.08	256.62	638.51**	-381.89	1683.35**	1301.12**
	(0.17)	(556.80)	(265.81)	(556.74)	(556.63)	(555.56)
Autonomy (1=Yes)	0.07	276.95	318.70	-41.76	-31.44	-104.40
	(0.16)	(546.14)	(250.28)	(536.65)	(542.16)	(538.90)
Observations	540	540	540	540	540	540
R-squared	0.27	0.30	0.20	0.22	0.25	0.24
Mean dep. var. control	2.36	7919.22	5280.02	2639.20	2639.20	2639.20
SD dep. var. control	1.33	4397.72	2351.52	4024.33	4024.33	4024.33
NA: $\beta_R + \beta_G = 0$	0.06	0.05	0.00	0.55	0.00	0.00
A: $\beta_R + \beta_G + \beta_A = 0$	0.02	0.01	0.00	0.59	0.00	0.00

Note: this table reports results on yields and profits earned by farmers in the 2015 season. Using the full sample of 540 farmers (in some columns for some fertilizers, the sample size is slightly smaller to reflect unknown answers by farmers), we run the following regression: $Y_{it} = \beta_0 + \beta_R Recomm_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata fixed effects and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Commercialization survey conducted in June 2016. In column 1, we use as the dependent variable the maize yields (tons/ha) self-reported by farmers in the 2015 season. In column 2, the dependent variable contains the value of farmers' maize production (per hectare) in the 2015 season. The value of the production (per hectare) is computed by multiplying the total amount of maize harvested by the farmer in the 2015 season by the price the maize could be sold in the market. We take the median price faced by farmers who sold at least a fraction of their production in the market as the price for all farmers when computing the value of the maize production. In column 3, the dependent variable is the total cost of production cost self-reported by each farmer. Total costs include the total investment in soil preparation activities, fertilizers (chemical and organic), herbicides, pesticides, and labor. We also include the cost of renting the sowing and harvest machines paid by QFD (when that was the case), as well as the subsidy for fertilizer packages, also paid by QFD. In column 4, the dependent variable is the difference between the dependent variable in columns 2 and 3. In column 5, the dependent variable is the cost of production, not including the subsidies paid by QFD. We use median market prices at the locality level to calculate revenues and profits. In localities where no farmer sold maize, we use median prices at the municipality level. To account for this imputation of prices (for 20 farmers in our sample), we include an dummy that takes value of one if prices were measured at the municipality level on the RHS of columns 2-5. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: 2017 Practices

	(1)	(2)	(3)	(4)
	Existing practices		All new practices	
	Total practices applied	Standardized Index	Total practices applied	Standardized Index
Recommendation (1=Yes)	0.11	0.12	0.24**	0.40**
	(0.11)	(0.12)	(0.11)	(0.16)
Grant (1=Yes)	-0.21**	-0.19*	0.08	0.12
	(0.10)	(0.12)	(0.13)	(0.19)
Autonomy (1=Yes)	0.05	0.05	0.45**	0.55**
	(0.10)	(0.12)	(0.14)	(0.20)
Observations	540	540	540	540
R-squared	0.09	0.08	0.22	0.18
Mean dep. var. control	2.31	0.00	0.42	0.00
SD dep. var. control	0.89	1.00	0.79	1.00
NA: $\beta_R + \beta_G = 0$	0.38	0.54	0.00	0.00
A: $\beta_R + \beta_G + \beta_A = 0$	0.65	0.83	0.00	0.00

This table reports results on the agricultural practices performed by the farmers in our study in the 2017 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recomm_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata fixed effects and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the final survey conducted in May 2017. The dependent variable in column 1 is a sum of individual dummies. Each dummy takes value of 1 if the farmer performed one of the so-called existing agricultural practices. In column 2, the dependent variable is the standardized index of the outcome in column 1, computed by standardizing each dummy individually, adding them all and standardizing the sum. We use the mean and standard deviation of the control group as reference for the standardized index. In columns 3 and 4, the dependent variables are analogous to the outcomes in columns 1 and 2, computed for the so-called new practices. The existing practices are: (a) ploughing, (b) using inorganic fertilizer and (c) covering the fertilizer. The new practices are: (a) deep tilling (ripping), (b) using hybrid seeds, (c) fertilizing at sowing, (d) sowing with precision machinery, (e) using pre-emergent herbicide and (f) using high-quality fertilizers. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: Hypotheses Results Summary

Main hypothesis	**Ho**
1. Autonomy reduces overall short-term compliance	Rejected
2. Autonomy increases persistence	Not rejected
3. Autonomy increases performance	Rejected
Secondary hypothesis	**Ho**
1. Reducing fertilizer use decreases yields	Rejected
2. In-kind grant increases compliance	Not rejected
3. Pairing recommendations with AEWs increases persistence	Not rejected

by farmers having kept the soil analysis and recommendation papers at higher rates (column 2). Farmers with autonomy were also $13pp$ more likely to report planning to use expert recommendations in the following growing season (column 3).

Trust: Reactance theory in our context posits that the deprivation of autonomy triggered negative responses including a loss of trust (see e.g. Brinson et al., 2024) that reduced adherence. We find some evidence for this in column 4 Table 7 as farmers with autonomy were 0.35σ more likely than farmers without autonomy to report trusting recommendations (measured on a five-point Likert scale) from the implementing partners—the extension service company and the agrodealer.

Choice Induced Preferences: In our setting, theories of choice-induced preferences would suggest that actively choosing to implement recommendations would increase their subsequent use by inducing greater liking of the recommendations among farmers with autonomy. We test this hypothesis by examining a hypothetical willingness to pay exercise in 2017 (2 years after the recommendations) for the three main YARA fertilizers used in the experiment. We note that this two-year difference may decrease concerns about experimenter demand effects. Farmers with autonomy expressed a 17% higher willingness to pay for KCl (307 vs 261 pesos) —the fertilizer which experts recommended to increase— relative to farmers without autonomy, consistent with choice induced preferences, and explanations such as learning. In contrast, autonomy did not increase the willingness to pay for fertilizers that experts recommended decreasing.

Heterogeneity in the Suitability of Expert Recommendations: Farmers were aware that the expert recommendations used cluster averaged soil analyses (of which their own sub-plot was a single data point). This averaging could have made farmers skeptical of the suitability of the recommendations for their individual plot. While randomization ensured that soil heterogeneity is similar across arms, some farmers without autonomy may have felt compelled to follow the recommendations (e.g. fertilize at sowing). This may have generated distrust (or reactance) and reduced their subsequent adherence. In the lab experiment we address this issue by examining adherence after conditioning on participant preferences for the expert recommendations.

Table 7: Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Knowledge of recommendations			Trust in the recomm. from input supplying institutions: standardized index	WTP for a bag of YARA fertilizer in 2017 (Mex$)		
	Remembers sowing recomm. (1=Yes)	Famers kept soil analysis (1=Yes)	Would follow the recomm. in the following year (1=Yes)		Urea	DAP	KCl
Year of data collection	*2015*	*2016*	*2016*	*2015*		*2017*	
Grant (1=Yes)	-0.01	0.07*	-0.06	0.37**	111.07***	138.54***	110.00***
	(0.02)	(0.04)	(0.06)	(0.13)	(17.78)	(22.65)	(20.35)
Autonomy (Yes=1)	0.05*	-0.02	0.13**	0.35**	14.17	25.02	46.15**
	(0.03)	(0.03)	(0.06)	(0.12)	(13.92)	(18.54)	(17.64)
Recommendation of any type (1=Yes)		0.29***	-0.04	0.61***	70.43***	64.68**	53.28**
		(0.06)	(0.06)	(0.14)	(19.57)	(24.69)	(20.63)
Observations	395	510	540	508	540	540	540
R-squared	0.09	0.19	0.06	0.29	0.31	0.27	0.29
Mean dep. var. control (or R)	0.03	0.58	0.39	-0.00	100.38	121.73	98.46
SD dep. var. control (or R)	0.18	0.50	0.49	1.00	151.12	185.92	157.18
NA: $\beta_R + \beta_G = 0$	0.79	0.00	0.10	0.00	0.00	0.00	0.00
A: $\beta_R + \beta_G + \beta_A = 0$	0.14	0.00	0.53	0.00	0.00	0.00	0.00

Note: this table reports results on the effect of the grant flexibility on a variety of outcomes. For column 1 we report the estimates from the following regression: $Y_{it} = \beta_R Recomm_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period (we do not use the control group). In columns 2-4, we report the point estimates of the following specification: $Y_{it} = \beta_0 + \beta_R Recomm_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_c + \epsilon_{it}$. Both specifications include randomization strata indicators and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the omitted group of farmers, which columns 1 is group R and in columns 2-7 is the control group. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. In column 1, the dependent variable is an indicator equal to 1 if the farmer reported remembering the sowing recommendations. In column 2, the dependent variable is an indicator equal to 1 if the farmer kept a record of the soil analysis given to them (30 farmers did not answer this question). In column 3, the dependent variable is an indicator equal to 1 if the farmer reported that they would probably follow the recommendations in the following season. Data for columns 2 and 3 were collected as part of the survey conducted in June 2016. In column 4, the dependent variable is a standardized index of two individual dummies that take value 1 if the farmers reported trusting the recommendations given by their input suppliers and IPAMPA (the AEW service). To compute the index, we standardize each dummy individually, add them all and standardize the sum. We use the mean and standard deviation of the control group as reference for the standardized index. The questions in the index in column 4 were asked as part of our follow-up survey (32 farmers did not answer both questions). In Columns 5-7, the dependent variables are the self-reported willingness to pay for a bag of each of the 3 YARA fertilizers (we impute zero for those who did not report WTPs). See Table OA-6 for the definition of variables. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

To summarize, in 2015 farmers with autonomy were more likely to self-report remembering the recommendations and had greater trust in the implementation agencies. They reported higher willingness to pay for fertilizers experts recommended vis-a-vis farmers with no autonomy. These changes could help explain why we see a greater persistence in program practices among this group two years after the program ended.

4.6 Alternative explanations

While the psychology literature suggests plausible mechanisms for the impacts of autonomy on persistence, we next explore some alternative explanations.

Reciprocity and Experimenter Demand Effects: Farmers with autonomy may have felt a stronger obligation to reciprocate or to adjust responses to perceived enumerator preferences relative to those without autonomy. We attempted to address these concerns to the extent possible by confirming implementation (e.g fertilizer use or field preparation) via surveyor observations or administrative data from the agro-dealer. While we cannot rule out such effects entirely, they are likely a reaction to subsidy stringency since they cannot be driven by subsidy amounts (farmers in both groups received the same amount) and in that sense linked to autonomy.

Interactions between Autonomy and Extension Services Since all treatment groups received high quality extension services, we cannot rule out the possibility that autonomy was effective only in the presence of these services. Anecdotally, farmers greatly valued the ability to use AEWs as a resource for understanding the recommendations. In this case, autonomy was only valuable because of access to a complemetrary resource. Such interactions may limit the generalizability of our results.

We next describe a lab experiment designed to identify the effect of autonomy in a different domain while addressing the potential confounds discussed above. In addition, the lab study can help test a few underlying mechanisms.

5 Lab Experiment

In this section, we first provide an overview of the experimental design, which closely mirrors that of the field experiment. We then describe the underlying theoretical framework and the specific ex-ante hypotheses that motivated our design choices. We conclude by presenting our results and discussing potential mechanisms that may help understand the findings from the lab and the field.

5.1 Experimental Design

Participants—who we will refer to as "builders"—were tasked with building abstract Lego figures as fast as possible and were paid \$0.50 for each successfully completed figure. The building task was designed to reflect real-world effort while providing measurable output (and some degree of entertainment for participants). Mirroring our field experiment, we introduced builders to a specific approach to building—or "technology"—that was framed as a potential way to improve construction speed. We called this method "Build Better" and the adoption of this technology and the persistence of its use was our primary interest.

After describing Build Better, we elicited participants preferred technology choice—either Build Better or a "free build" option that did not require the use of Build Better tools and techniques. We then informed approximately half of the builders that the technology they would use—i.e., whether they had to use Build Better—would in fact be selected by a computer. The remaining half of the builders were allowed to use the technology they selected. In order to ensure that this manipulation had time to sink in and to create temporal separation between revealing whether the computer would choose and revealing its choice, builders in all experimental arms completed a short block-sorting task unrelated to the technologies. To ensure that our manipulation remained salient, the experimenter asked builders which condition they were in—own choice versus computer selection—moments before they began the building phase. All but three participants were able to accurately state their treatment condition.[31]

After the block-sorting task, the builder was informed of the computer choice or reminded of their own. Figure 2 provides a visual schematic of the design. Those who utilized Build Better—either by choice or by computer selection—were provided a large checklist directly in front of them to ensure that they were saliently aware of the elements of the method.

The bulk of experiment time consisted of a 20-minute building phase. During the building phase, builders made as many replications of three different 10-by-10 abstract figures as possible and were provided a very large bag of Lego from which they could draw blocks. During this phase, we recorded overhead videos of participants, which we subsequently utilized to encode participants' adherence to the prescribed building method. This adherence was encoded by an enumerator on a five-point scale with 1 corresponding to "Did not follow Build Better" and 5 corresponding to "Almost perfectly or perfectly followed Build Better".[32] At the end of the building phase, the experimenter counted the correct, incorrect, and partial figures that the builder constructed.

Finally, builders were asked demographic questions and three questions that were not incentized about their subjective experience. Responding with a five-point Likert scale (from Strongly Disagree to Strongly Agree), participants indicated the degree to which they endorsed each of the following statements: (i) "I enjoyed the building method that I just used"; (ii) "I would have been faster using another building method"; and (iii) "I was among the top 25% of builders (in terms of total figures built)." In the analysis, we reverse-coded the second question so higher scores are more positive.

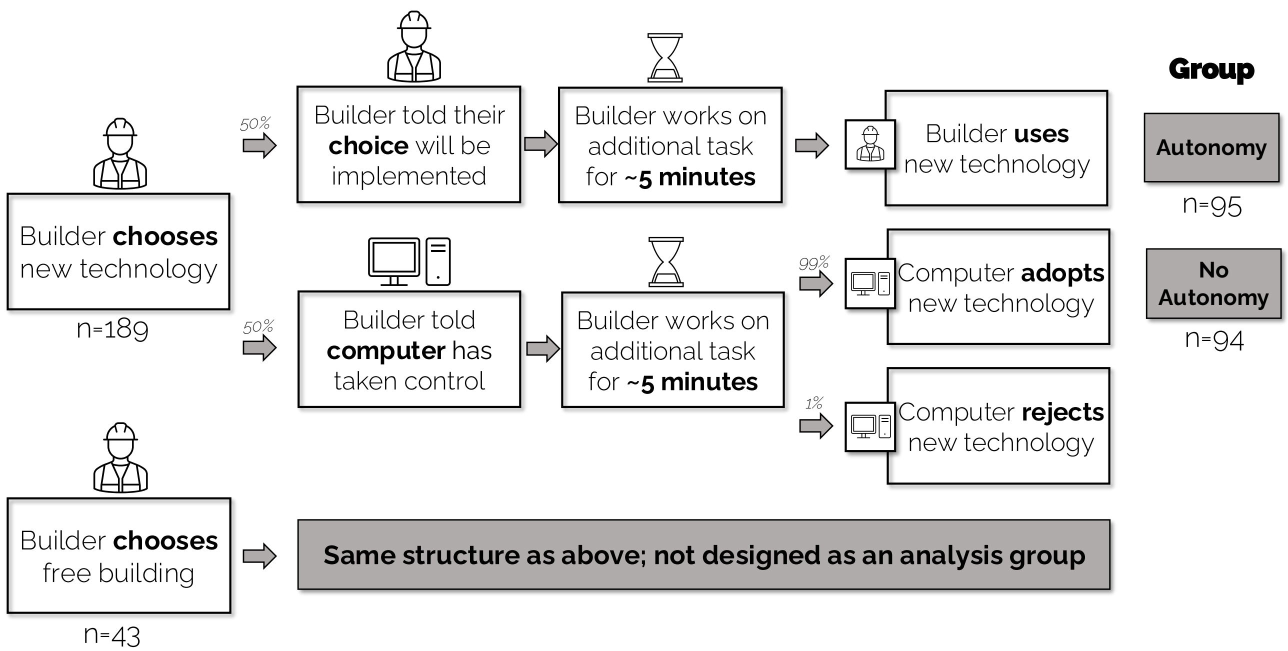
The complete experiment consisted of two sessions (or periods). In each session, participants were (re)introduced to the task, provided with the necessary construction materials and recommendations, and completed the building task. There were three key differences between the two sessions. First, and most critically, builders in Session 2 were *not* constrained in their building method. We specifically stated: "We will provide the tools for Build Better. But you are free to build however you like." Fur-

[31] During this delay, we gathered additional information about the builders by measuring their speed at sorting a specific bag of Lego and counting specific shapes and colors from that bag.

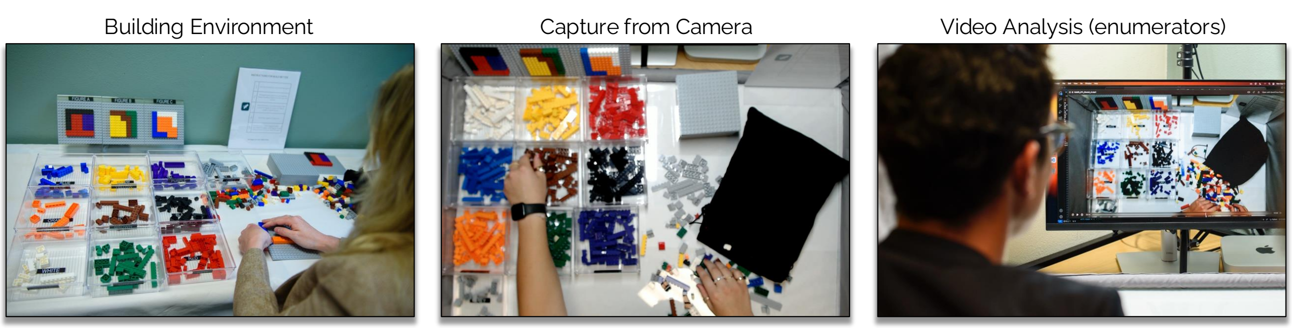
[32] This enumerator was blinded to the aims of the study as well as the builders' experimental arm. The enumerator was hired from an outside university to avoid unintentional information spillage. The enumerator was given the full instructions of the experiment and a print-out of the page that described Build Better to use during scoring.

Figure 2: Experimental Design and Captures

(a) Conceptual flow of the experiment



(b) Real building environment and analysis setup



Panel A depicts the overall experimental design, the autonomy manipulation, and the designation of our two analysis groups. Panel B shows real-world photographs of our physical setup and video analysis of the building phase.

thermore, we gave participants the option to request the checklist for Build Better (and tallied those who asked). Second, builders in Session 2 were asked two questions about their recall of the first session. In a free-text box, builders were asked to describe Build Better in their own words. Second, builders were presented with nine figures, three of which corresponded to the actual figures they built in Session 1 and asked to select the correct three. Both questions were incentivized: participants were given tokens for a $20 lottery if their answers were correct. Finally, in the first session only, we concluded by asking participants about their willingness to complete additional building tasks, for $5 additional pay, *utilizing the method they just used*. We incentivized this question using a Becker-Degroot-Marschak (BDM) mechanism. In order to avoid contaminating production in Session 2 or otherwise distort learning as a result of this mechanism, the results of this BDM were given at the very end of Session 2 (and any resulting building occurred then).

The experiment was conducted at Michigan State University. Participants were recruited from the student population, were at least 18 years old, and were required to have standard color vision. In total, 232 participants completed the first session; 214 completed both sessions, with no significant difference across treatment groups. Participants did the experiment one-at-a-time with the guidance of an experimenter in a purpose-built lab. On average, the first session took approximately 50 minutes and the second session took about 45 minutes, during which participants were not allowed to use their cell phones or other devices. Participants were paid a fixed fee of $20 in addition to their earnings, paid at the end of Session 2. In total, participants earned an average of $38.50. We relegate details on the participants and balance checks to Appendix I.

Discussion of Design. In order to isolate the impact of autonomy, we compare outcomes among participants that both (a) preferred a given technology and (b) actually used that technology—thus keeping preferences and technology used constant. That is, we will compare subjects that chose and ended up using Build Better, but the process of getting to Build Better differed: in the Autonomy arm subjects chose the recommended technology, whereas in the No Autonomy arm the computer chose for them. Our "computer choice" manipulation enables this comparison.[33]

Critically, we implemented the computer randomization after eliciting builders' preferred methods so we could compare the behavior and attitudes of builders all of whom preferred the Build Better method. This was the vast majority of participants (189/232 total participants chose Build Better). Among these builders, we compare those that were randomized into the autonomy condition and thus used their preferred method under autonomy—that is, without a computer insisting on the Build Better method (the "Autonomy" group in Figure 2)—and those who were randomized into the computer selection condition and the computer chose the Build Better method (the "No Autonomy" group in Figure 2).[34]

[33]This follows e.g.Chaudhry and Klinowski (2016); Dal Bó et al. (2010). Subjects may not have felt sufficiently constrained by the No Autonomy condition, suggesting our autonomy manipulation was weaker than intended—this would bias us toward null effects.

[34]To avoid deception but retain high statistical power, the computer implemented Build Better with probability 0.99.

5.2 Hypotheses

We first test the primary hypotheses from the field experiment in our lab setting as a form of replication and validation of our findings. In contrast with the experiment, since we elicited participant preference before the autonomy manipulation, we can condition on this preference. In particular, we focus on those subjects who expressed a preference for Build Better. We define two key groups that are the focus of our analysis:

> **Autonomy group (A)**. Those participants who chose Build Better and, during the building stage, were free to build however they determined was best.
>
> **No Autonomy group (NA)**. Those participants who chose Build Better and for whom the computer determined the method they were required to build with during the building stage (which was Build Better).

These groups comprise the majority of participants. As noted above, 82% of participants preferred to use Build Better in Session 1. 95 participants were randomized into the autonomy subgroup and the remaining 94 were in the no autonomy group. We now explicitly restate the hypotheses —which coincide with those in the field experiment— with a brief note on how we operationalize each in the new setting:

Hypothesis 1. Autonomy reduces compliance with recommendations (A vs NA in Session 1) To explore this hypothesis, we compare the adherence to Build Better across these two groups.

Hypothesis 2. Autonomy increases persistence (A vs NA in Session 2) Following Hypothesis 1 above, we compare the adherence to Build Better in Session 2 when all participants were free to build with whatever method they chose.

Hypothesis 3. Autonomy increases performance (A vs NA) We compare both the total number of figures successfully completed and the total number including incorrect or incomplete figures across the two groups.

In addition to replicating our field findings, we extend our attention to new hypotheses that are enabled by our greater control in the lab setting. We next present a few additional hypotheses designed to explore potential mechanisms motivated by the psychology literature.

Mechanism Hypothesis MH1. Autonomy improves contemporaneous liking (A vs NA in Session 1) As noted when discussing our field experiment, choice-induced preference theory would imply that the act of actively choosing the Build Better method would increase subsequent willingness to use those tools. We test this in two ways: we elicit satisfaction with the method used (in Session 1) and participants' willingness to complete additional building tasks for additional pay (incentivized and elicited at the end of Session 1).

Mechanism Hypothesis MH2. Autonomy increases contemporaneous confidence (A vs NA in Session 1) We asked participants two variants of questions about confidence after the building phase of each session. First, we asked whether they believed that they would have been faster with a different building method than the method they just used. Second, we elicited participants' confidence by asking whether they believed they were in the top 25% of builders. These questions were designed to help us explore whether autonomy enhances a perception of control and therefore increases confidence.

Mechanism Hypothesis MH3. Autonomy improves memory (A vs NA in Session 2) Our field evidence suggests that memory (or attention to recommendations) may play a role in subsequent adherence. We attempt to distinguish between overall memory effects versus memory of the recommendations with two incentivized questions asked at the start of the second session. To measure the former, we ask builders to identify the three figure shapes that they built in the first session (from a grid of nine total images). To measure the latter, we ask builders to describe (in a free text box) the steps involved in the Build Better method.

5.3 Results

We present results using a regression framework (mirroring our field results presentation). We regress outcomes on an indicator for the Autonomy group (thus comparing against those in the No Autonomy treatment) and a set of cohort indicators (γ_s) that correspond to each of the three semesters in which participants completed the experiment.[35] This yields the following estimation equation:

$$Y_{it} = \beta_0 + \beta_A \text{Autonomy}_i + \gamma_s + \epsilon_{it} \tag{2}$$

where i denotes a builder and t is the session of interest (either 1 or 2). Our key variable of interest is the coefficient on the indicator for autonomy β_A. We first present results from our primary hypotheses.

Table 8: Adherence to Build Better and Productivity in the Lab

	(1)	(2)	(3)	(4)
Time of data collection	*Session 1*		*Session 2*	
	Adherence	Completed Figures	Persistence	Completed Figures
Autonomy (Yes = 1)	-0.02	0.34	1.28***	1.08**
	(0.16)	(0.37)	(0.17)	(0.51)
Observations	189	189	177	177
R-squared	0.01	0.03	0.25	0.05
Mean dep. variable, NA	4.05	5.18	2.36	7.12
SD dep. variable, NA	1.02	2.76	1.25	3.44

This table reports results on the degree to which participants followed the Build Better approach and the number of figures they completed in the two sessions of the experiment. We report the mean of each outcome for the No Autonomy group (the omitted category). Adherence was scored on a five-point scale; completed figures corresponds to the number of ten-by-ten squares that the participant successfully replicated. The difference in observations stems from twelve participants who did not complete the second session; our results are robust to focusing on only those who complete both sessions. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

[35] Our results do not qualitatively or quantitatively change much without these indicators; we present alternative analyses in the Online Appendix.

Results on Compliance (H1): We explore the degree to which autonomy shaped participants' short-term take-up in column 1 of Table 8. As noted above, we measured this with a score on a 1-5 scale which corresponded to the degree to which the participant followed the Build Better instructions. We find that adherence to the new technology is nearly identical across the Autonomy and No Autonomy groups in the first session of the experiment. This mirrors our field result: autonomy did not decrease adherence with recommendations in Session 1 and we thus reject H1.

Results on Persistence (H2): We explore the degree to which autonomy shaped participants' longer-term take-up in column 3 of Table 8 and find that adherence to the Build Better approach in Session 2 was significantly higher for those with autonomy. The magnitudes are large 54% (1.02σ) higher persistence ($p < 0.001$). This is also consistent with our field result: autonomy increased adherence with recommendations in Session 2 and thus we cannot reject H2.

Results on Performance (H3): Finally, we look at the degree to which autonomy affected productivity (columns 2 and 4 of Table 8). Autonomy had no effect on Session 1 production. However, in Session 2 those with autonomy generated 15% more completed figures than those without autonomy ($p = 0.037$). If incorrect or partial figures are included, this session-two difference shrinks and is no longer statistically significant (0.71 difference between A and NA groups, $p = 0.135$). This result has no field counterpart since we did not measure long-term yields in the field experiment.

We now turn to our hypotheses related to the underlying mechanisms, presented in Table 9.

Results on Liking (MH1): Participants generally endorsed the statement "I enjoyed the building method I just used", which was asked near the end of Session 1 after the building phase. The level of enjoyment did not depend on whether the participant had autonomy (see column 1 of Table 9). Likewise, participants were willing to complete approximately 13 additional figures for extra pay (which would take a significant amount of time, approximately 25 minutes at the end of Session 2). This did not depend on autonomy either (column 2 of Table 9). Thus, while participants generally enjoyed the building task (in contrast to some real-effort experiments), we find no evidence for choice-induced-preference theories.

Results on Confidence (MH2): Recall, both groups used a similar building method in Session 1. As compared to those without autonomy, participants with autonomy were no more likely to believe they would have been faster with a different building method at the conclusion of Session 1 (column 2 Table 9). Likewise, builders with autonomy were no more likely to believe that they were in the top quartile of performers (column 3 Table 9) in the first session. Although this latter judgment involves an interpersonal comparison and could thus be independent from builders' confidence in the building technology, the pair of questions suggests that autonomy did not alter contemporaneous confidence.

Table 9: Enjoyment, Confidence, and Recall Questions in the Lab

Panel A: Session 1	(1) Enjoyment	(2) Faster Alternative	(3) Top 25%	(4) WTP	
Autonomy (Yes = 1)	-0.12	-0.03	0.02	-0.14	
	(0.17)	(0.16)	(0.17)	(0.71)	
Observations	189	189	189	189	
R-squared	0.01	0.00	0.08	0.03	
Mean dep. variable, NA	3.78	3.03	2.46	12.97	
SD dep. variable, NA	1.23	1.19	1.19	4.84	
Panel B: Session 2	(5) Enjoyment	(6) Faster Alternative	(7) Top 25%	(8) Recall: Method	(9) Recall: Figures
Autonomy (Yes = 1)	-0.43***	-0.59***	-0.35**	0.40**	-0.12
	(0.12)	(0.17)	(0.17)	(0.19)	(0.15)
Observations	177	177	177	177	177
R-squared	0.09	0.09	0.04	0.10	0.03
Mean dep. variable, NA	4.54	3.90	3.13	2.81	2.45
SD dep. variable, NA	0.78	1.15	1.20	1.34	0.98

Note: this table reports results on the effect of autonomy on a variety of outcomes. The first three outcomes correspond to responses to 5-point Likert scale statements: (i) "I enjoyed the building method that I just used"; (ii) "I would have been faster using another building method" [reverse coded]; and (iii) "I was among the top 25% of builders (in terms of total figures built)." Recall of method was scored on a five-point scale; recall of figures corresponds to the number of figures correctly identified (out of three). We report the mean and the standard deviation of the dependent variable for the No Autonomy group. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Results on Memory and Recall (MH3): Finally, we analyze responses to two questions related to memory, both presented at the beginning of the second session. First, we elicited free-text responses for participants, asking them to describe *Build Better*. These were scored out of five points, with higher scores indicating better recall. We find that autonomy in Session 1 improved memory along this dimension by 14% ($p = 0.038$; see column 8 of Table 9). Second, we presented participants with a grid of nine figures and tasked them with identifying the figures that they assembled in Session 1. Three of the were indeed the figures that they assembled in the first session, while the remaining six broadly resembling the figures but were different. We encode participants' responses as a score from 0 to 3, indicating the number of correctly remembered figures. Column 9 of Table 9 shows this result. On average, builders recalled 82% of the figures they had built in the first session, but we find no effect of autonomy ($p = 0.415$).

5.4 Discussion

Our results are consistent across the two experiments. This agreement across different domains suggests that the details of the field experiment (e.g., interactions between autonomy and extension services) are not the main driver for the observed effect of autonomy. In both studies, autonomy led to a long-term increase in the use of a new technology, and this did not come at any short-term cost related to lower performance or lower compliance with the recommendations. The results of the lab survey suggest that this long-term persistence likely did not stem from choice-induced preferences. Instead, both sets of results suggest that increased attention or improved memory may have played a role.

We next adduce additional results from the second session to suggest a unifying framework. Participants *without* autonomy (i) were more likely to change their building method (as evidenced by the differential persistence discussed in H2 above); (ii) liked their new method more (column 5 of Table 9) and believed this new method was effective (column 7 of Table 9), although (iii) in reality their new approach was less effective (column 4 of Table 8).

One (admittedly ex-post) narrative consistent with these results is that builders (in both groups) engaged in a form of motivated reasoning. Those with autonomy stuck with their previous building method in order to justify their previous choice. In doing so, they found their second session similarly enjoyable to their first session and, perhaps due to learning-from-doing effects, completed more figures than in the previous session. In contrast, those without autonomy did not feel as bound to their previous method. Although we do not have direct evidence for this, the fact that builders without autonomy were more likely to use a different method in Session 2 suggests that (at some point in the intervening weeks) their beliefs about the efficacy of the Build Better method diverged from those in the Autonomy group. Then, having chosen a new approach, those in the No Autonomy group felt compelled to justify this new method with higher self-reported enjoyment and confidence.

6 Conclusion

We provide novel evidence for the effect of autonomy on the adoption and persistence of expert recommendations in technology adoption. First, in a field experiment with Mexican smallholder farmers, we demonstrate that autonomy in subsidy design can significantly enhance long-term adherence to improved agricultural practices without compromising short-term adoption. Second, in a lab experiment that replicated key features of the field study and provided tighter control over the autonomy manipulation, we confirm the persistence effects and find suggestive evidence that autonomy works by increasing recall of expert recommendations.[36]

The studies yield two key insights. First, contrary to common concerns justifying limited autonomy in the use of subsidies or grants, autonomy did not reduce short-term compliance with expert recommendations. Farmers with a choice over which inputs to purchase with a grant adopted the recommended practices and reduced fertilizer overuse at similar rates as those without the choice, achieving a 14% reduction in CO_2 emissions without reducing yields. Second, in both studies, autonomy substantially improved persistence. In the field study, farmers with autonomy were significantly more likely to implement recommended practices up to two years after the intervention ended than those without autonomy. In the lab experiment, subjects were much more likely to use the expert recommendations in a later session when they had complete freedom over the choice of method. Third, the lab experiment suggests that autonomy can improve attention to and engagement with expert recommendations. The total effect of autonomy seems to stem from a combination of this and other factors (such as motivated reasoning). We leave disentangling these mechanisms for future research.

[36]The finding that autonomy improves memory also speaks to recent experimental work examining the role of other mechanisms (e.g. stories) in improving recall (Graeber et al., 2024).

These results have potentially important implications for policy design. While strict conditions on subsidies are often imposed to ensure compliance, our findings suggest that preserving autonomy, even in a limited sense, can enhance the long-term effectiveness of interventions aimed at changing behavior. Future work could explore these channels in more detail and inform the design of programs seeking to promote beneficial technologies and practices—from agricultural extension to public health interventions. More broadly, our results suggest that carefully incorporating behavioral factors like autonomy into policy design can meaningfully improve program effectiveness. As a concrete example, we believe that the central tension we highlight in this paper—increased compliance via mandate versus long-term persistence—is playing out in the tension around vaccination in the United States. Understanding how to optimize the balance between guidance and autonomy remains a crucial challenge for effective policy design.

References

Alan, S. and Kubilay, E. (2025). Empowering adolescents to transform schools: Lessons from a behavioral targeting. *American Economic Review*, 115(2):365–407. [4]

Ashraf, N., Banerjee, A., and Nourani, V. (2020). Learning to Teach by Learning to Learn. Working Paper. [4]

Bardhan, P. (2000). Irrigation and Cooperation: An Empirical Analysis Of 48 Irrigation Communities In South India. *Economic Development and Cultural Change*, 48(4):847–865. [4]

Bartling, B., Fehr, E., and Herz, H. (2014). The intrinsic value of decision rights. *Econometrica*, 82(6):2005–2039. [4]

Berck, P., Geoghegan, J., and Stohs, S. (2000). A Strong Test of the von Liebig Hypothesis. *American Journal of Agricultural Economics*, 82(4):948–955. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/0002-9092.00093. [OA - 11]

Black, A. E. and Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, 84(6):740–756. [1]

Black, S. E. and Lynch, L. M. (2001). How to Compete: The Impact of Workplace Practices and Information Technology on Productivity. *The Review of Economics and Statistics*, 83(3):434–445. [4]

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics*. [2, 5]

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3):384. [1]

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press. [1]

Brinson, N., Amini, B., Lemon, L., and Bawole, C. (2024). Electronic performance monitoring: The role of reactance, trust, and privacy concerns in predicting job satisfaction in the post-pandemic workplace. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 18. [19]

Bruns, H. and Perino, G. (2023). The Role of Autonomy and Reactance for Nudging — Experimentally Comparing Defaults to Recommendations and Mandates. *Journal of Behavioral and Experimental Economics*, 106:102047. [1]

Carter, M., Laajaj, R., and Yang, D. (2021). Subsidies and the African Green Revolution: Direct Effects and Social Network Spillovers of Randomized Input Subsidies in Mozambique. *American Economic Journal: Applied Economics*, 13(2):206–229. [9]

Chaudhry, S. J. and Klinowski, D. (2016). Enhancing Autonomy to Motivate Effort: An Experiment on the Delegation of Contract Choice. In *Experiments in Organizational Economics*, volume 19 of *Research in Experimental Economics*, pages 141–157. Emerald Group Publishing Limited. [4, 24]

Cole, S., Killeen, G., Harigaya, T., and Krishna, A. (2020). Using Satellites and Phones to Evaluate and Promote Agricultural Technology Adoption: Evidence from Smallholder Farms in India. page 66. [5]

Coppin, G., Delplanque, S., Cayeux, I., Porcherot, C., and Sander, D. (2010). I'm no longer torn after choice: How explicit choices implicitly shape preferences of odors. *Psychological science*, 21(4):489–493. [1]

Corral, C., Giné, X., Mahajan, A., and Seira, E. (2020). Autonomy and specificity in agricultural technology adoption: Evidence from mexico. Working Paper 27681, National Bureau of Economic Research. [6, 14, OA - 2]

Cui, Z., Zhang, H., Chen, X., Zhang, C., Ma, W., Huang, C., Zhang, W., Mi, G., Miao, Y., Li, X., Gao, Q., Yang, J., Wang, Z., Ye, Y., Guo, S., Lu, J., Huang, J., Lv, S., Sun, Y., Liu, Y., Peng, X., Ren, J., Li, S., Deng, X., Shi, X., Zhang, Q., Yang, Z., Tang, L., Wei, C., Jia, L., Zhang, J., He, M., Tong, Y., Tang, Q., Zhong, X., Liu, Z., Cao, N., Kou, C., Ying, H., Yin, Y., Jiao, X., Zhang, Q., Fan, M., Jiang, R., Zhang, F., and Dou, Z. (2018). Pursuing Sustainable Productivity With Millions Of Smallholder Farmers. *Nature*, 555(7696):363–366. [5]

Dal Bó, P., Foster, A., and Putterman, L. (2010). Institutions And Behavior: Experimental Evidence On The Effects Of Democracy. *American Economic Review*, 100(5):2205–29. [3, 4, 24]

de Rochambeau, G. (2021). Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms. Working Paper. [4]

Deci, E. L. and Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer Science and Business Media, Berlin. [1]

Desiere, S. and Jolliffe, D. (2017). Land Productivity And Plot Size: Is Measurement Error Driving The Inverse Relationship? Technical report, The World Bank. [14, OA - 22]

Ding, Z., Jiang, T., Chen, C., Murty, V. P., Xue, J., and Zhang, M. (2021). The effect of choice on intentional and incidental memory. *Learning & Memory*, 28(12):440–444. [16]

Dupas, P. (2014). Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment. *Econometrica*, 82(1):197–228. [1]

Easterly, W. (2014). *The Tyranny of Experts: Economists, Dictators, and the Forgotten Rights of the Poor*. Basic Books, New York. [5]

Falk, A. and Kosfeld, M. (2006). The Hidden Costs of Control. *American Economic Review*, 96(5):1611–1630. [4]

Fehr, E., Herz, H., and Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review*, 103(4):1325–59. [4]

Fernández, A. T., Wise, T. A., and Garvey, E. (2012). Achieving Mexico's Maize Potential. Global Development and Environment Institute Working Paper 12-03. [2]

Fischer, G., Karlan, D., McConnell, M., and Raffler, P. (2019). Short-term subsidies and seller type: A health products experiment in Uganda. *Journal of Development Economics*, 137:110–124. [1]

Foster, A. and Rosenzweig, M. (2010). Microeconomics of technology adoption. *Annual Review of Economics*, 2(1):395–424. [1, OA - 25]

Freire, P. (1974). *Education for Critical Consciousness*. Bloomsbury Academic, London. [4]

Gars, J., Fishman, R., Kishore, A., Rothler, Y., and Ward, P. S. (2022). Does Soil Health Testing Affect Fertilizer Usage? – Evidence from a Field Experiment in Bihar, India. Working paper. [5]

Gerard, H. B. and White, G. L. (1983). Post-decisional reevaluation of choice alternatives. *Personality and Social Psychology Bulletin*, 9(3):365–369. [1]

Giné, X., Patel, S., Ribeiro, B., and Valley, I. (2022). Efficiency and equity of input subsidies: Experimental evidence from Tanzania†. *American Journal of Agricultural Economics*, 104(5):1625–1655. [9]

Goyal, A. and Nash, J. (2017). *Reaping Richer Returns: Public Spending Priorities for African Agriculture Productivity Growth*. World Bank, Washington, DC. []

Graeber, T., Roth, C., and Zimmermann, F. (2024). Stories, statistics, and memory. *The Quarterly Journal of Economics*, 139(4):2181–2225. [29]

Hamilton, B. H. (2000). Does entrepreneurship pay? an empirical analysis of the returns to self-employment. *Journal of Political Economy*, 108(3):604–631. [4]

Harou, A. P., Madajewicz, M., Magomba, C., Michelson, H., Tschirhart, K., Amuri, N., and Palm, C. (2018). Can Information Improve Investment? Effects Of Site-Specific Soil Recommendations On Fertilizer Demand. Working paper. [5]

Haynes, R. B., McKibbon, A., and Kanani, R. (1996). Systematic review of randomised trials of interventions to assist patients to follow prescriptions for medications. *The Lancet*, 348(9024):383–386. [5]

INEGI (2014). Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2014. Accessed: 2025-03-14. [6]

IPCC (2004). N2o emissions from managed soils, and co2 emissions from lime and urea application. Technical report, Intergovernmental Panel on Climate Change. Background Paper for the IPCC Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories. [OA - 31]

Jayne, T. S. and Rashid, S. (2013). Input Subsidy Programs In Sub-Saharan Africa: A Synthesis Of Recent Evidence. *Agricultural Economics*, 44(6):547–562. [9]

Ladha, J. K., Jat, M. L., Stirling, C. M., Chakraborty, D., Pradhan, P., Krupnik, T. J., Sapkota, T. B., Pathak, H., Rana, D. S., Tesfaye, K., and Gerard, B. (2020). Chapter Two - Achieving the sustainable development goals in agriculture: The crucial role of nitrogen in cereal-based systems. In Sparks, D. L., editor, *Advances in Agronomy*, volume 163, pages 39–116. Academic Press. [5]

Lin, W., Chen, B., and Wang, A. (2022). Misspecified Learning In Technology Adoption: Experimental Evidence From Fertilizer Use In China. Working Paper. [5]

Lowder, S. K., Sánchez, M. V., and Bertini, R. (2021). Which farms feed the world and has farmland become more concentrated? *World Development*, 142:105455. [1]

Macours, K. (2019). Farmers' Demand and the Traits and Diffusion of Agricultural Innovations in Developing Countries. *Annual Review of Resource Economics*, 11(1):483–499. [1]

Magruder, J. R. (2018). An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries. *Annual Review of Resource Economics*, 10(1):299–316. [1]

Meriggi, N. F., Bulte, E., and Mobarak, A. M. (2021). Subsidies for technology adoption: Experimental evidence from rural Cameroon. *Journal of Development Economics*, 153:102710. [1]

Millar, N., Robertson, G. P., Grace, P. R., Gehl, R. J., and Hoben, J. P. (2010). Nitrogen fertilizer management for nitrous oxide (N2O) mitigation in intensive corn (Maize) production: an emissions reduction protocol for US Midwest agriculture. *Mitigation and Adaptation Strategies for Global Change*, 15(2):185–204. [OA - 31]

Millar, N., Urrea, A., Kahmark, K., Shcherbak, I., Robertson, G. P., and Ortiz-Monasterio, I. (2018). Nitrous oxide (N2O) flux responds exponentially to nitrogen fertilizer in irrigated wheat in the Yaqui Valley, Mexico. *Agriculture, Ecosystems & Environment*, 261:125–132. [2]

Miller, W. R. (1983). Motivational interviewing with problem drinkers. *Behavioural Psychotherapy*, 11(2):147–172. [4]

Mirty, V., DuBrow, S., and Davachi, L. (2019). Decision-making increases episodic memory via postencoding consolidation. *Journal of Cognitive Neuroscience*, 31(9):1308–1317. [16]

Moscona, J. and Sastry, K. A. (2025). Inappropriate technology: Evidence from global agriculture. Working Paper 33500, National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138. [1]

Muralidharan, K., Romero, M., and Wüthrich, K. (2019). Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments. Working Paper 26562, National Bureau of Economic Research. Series: Working Paper Series. [11]

Murphy, D. M. A., Roobroeck, D., Lee, D. R., and Thies, J. (2019). Underground Knowledge: Estimating the Impacts of Soil Information Transfers through Experimental Auctions. Working paper. [5]

O'Shea, B. A. and Ueda, M. (2021). Who is more likely to ignore experts' advice related to covid-19? *Preventive Medicine Reports*, 23:101470. [5]

Penman, J., Kruger, D., Galbally, I., Hiraishi, T., Nyenzi, B., Emmanul, S., Buendia, L., Hoppaus, R., Martinsen, T., Meijer, J., Miwa, K., and Tanabe, K. (2000). *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. Published by the Institute for Global Environmental Strategies (IGES) for the IPCC, Hayama, Kanagawa, Japan. OCLC: 891561395. [14, OA - 31]

Ronayne, D. and Sgroi, D. (2018). When good advice is ignored: The role of envy and stubbornness. The Warwick Economics Research Paper Series (TWERPS) 1150, University of Warwick, Department of Economics. [5]

Rosenberg, B. D. and Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, 4(4):281—-300. [1]

Schneider, S., Nebel, S., Beege, M., and Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction*, 58:161–172. [16]

Sharot, T., De Martino, B., and Dolan, R. J. (2009). How choice reveals and shapes expected hedonic outcome. *Journal of Neuroscience*, 29(12):3760–3765. [1]

Sinclair, T. R. and Rawlins, S. L. (1993). Inter-Seasonal Variation In Soybean And Maize Yields Under Global Environmental Change. *Agronomy Journal*, 85(2):406–409. [14]

Sjöström, J. T., Ulku, L., and Vadovic, R. (2018). Free to Choose: Testing the Pure Motivation Effect of Autonomous Choice. *SSRN Electronic Journal*. [4]

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, 50(6):835–853. [4]

Stolper, O. A. and Walter, A. (2017). Financial literacy, financial advice, and financial behavior. *Journal of Business Economics*, 87(5):581–643. Publisher: Springer Berlin Heidelberg. [1]

Sweeney, S., Steigerwald, D. G., Davenport, F., and Eakin, H. (2013). Mexican Maize Production: Evolving Organizational And Spatial Structures Since 1980. *Applied Geography*, 39:78–92. [2]

Tjernström, E. (2017). Learning from Others in Heterogeneous Environments. Technical Report April. [5]

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K., and Deci, E. (2004). Motivating learning, performance, and persistence: the synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology*, 87(2):246–260. [1]

Vitousek, P. M., Naylor, R., Crews, T., David, M. B., Drinkwater, L. E., Holland, E., Johnes, P. J., Katzenberger, J., Martinelli, L. A., Matson, P. A., Nziguheba, G., Ojima, D., Palm, C. A., Robertson, G. P., Sanchez, P. A., Townsend, A. R., and Zhang, F. S. (2009). Agriculture. Nutrient imbalances in agricultural development. *Science (New York, N.Y.)*, 324(5934):1519–1520. [2]

Weizsäcker, G. (2010). Do we follow others when we should? a simple test of rational expectations. *American Economic Review*, 100(5):2340–60. [5]

West, P. C., Gerber, J. S., Engstrom, P. M., Mueller, N. D., Brauman, K. A., Carlson, K. M., Cassidy, E. S., Johnston, M., MacDonald, G. K., Ray, D. K., and Siebert, S. (2014). Leverage points for improving global food security and the environment. *Science*, 345(6194):325–328. [2, 5]

WHO (2003). *Adherence to Long-Term Therapies: Evidence for Action*. World Health Organization, Geneva. [5]

Wu, Y., Xi, X., Tang, X., Luo, D., Gu, B., Lam, S. K., Vitousek, P. M., and Chen, D. (2018). Policy distortions, farm size, and the overuse of agricultural chemicals in China. *Proceedings of the National Academy of Sciences*, 115(27):7010–7015. [2]

Wuepper, D., Le Clech, S., Zilberman, D., Mueller, N., and Finger, R. (2020). Countries influence the trade-off between crop yields and nitrogen pollution. *Nature Food*, 1(11):713–719. [2]

Appendix of "Autonomy and Technology Adoption"

by Ben Bushong, Carolina Corral, Xavier Giné, Aprajit Mahajan, and Enrique Seira

FOR ONLINE PUBLICATION

Autonomy and Technology Adoption

Benjamin Bushong, Carolina Corral, Xavier Giné, Aprajit Mahajan, Enrique Seira

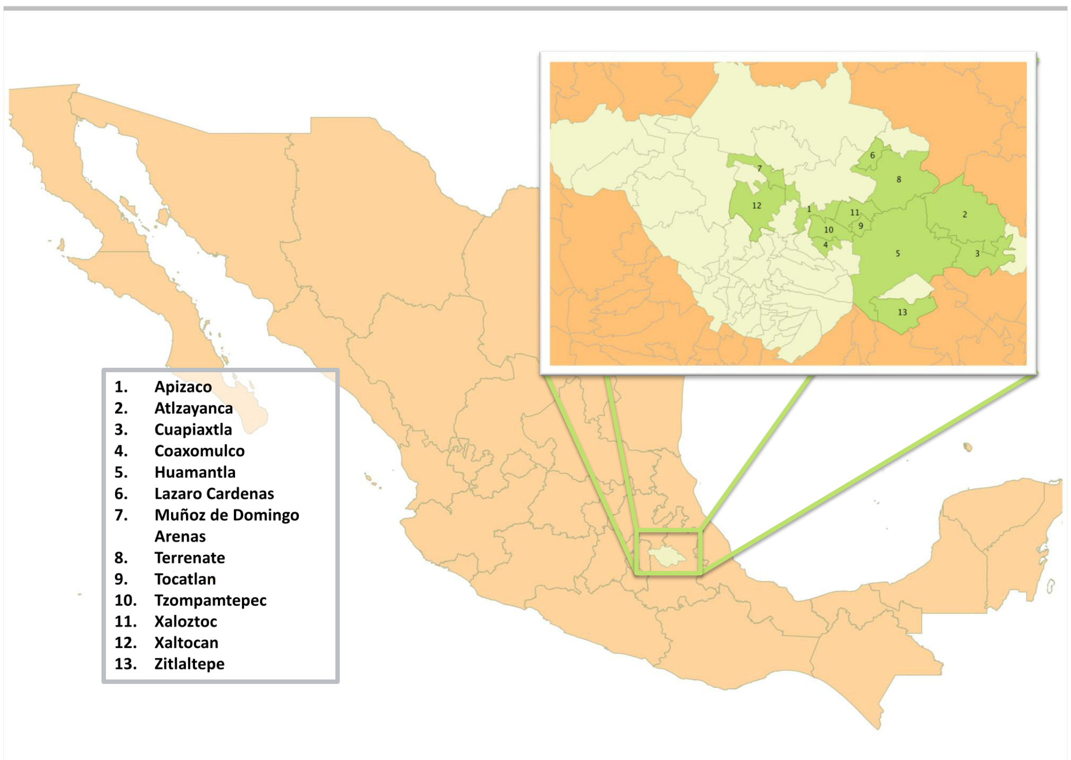
Appendix − For Online Publication

Contents

A More on Recruitment, Implementation and Attrition

In January 2015, QFD advertised the program widely by displaying posters prominently in public locations and handing out informational leaflets. QFD also organized a total of 34 promotional meetings in the principal towns in each municipality. The promotional meetings lasted between 60–90 minutes and typically took place in a large public space (e.g. a municipal auditorium). During the meetings, the research team introduced and explained the intervention, described the eligibility requirements and the grant lottery. Potentially interested farmers were asked to complete a short form.

Figure OA-1: Map of Tlaxcala



Between February and March 2015, interested farmers were visited by the research team. During the visit, the team collected a detailed baseline survey on a range of farmer characteristics and agricultural practices during the previous growing season of 2014. After the survey, farmers were asked to register a subplot of one hectare for the program where they planned to grow maize. QFD cordoned off this subplot, recorded its GPS coordinates, and collected soil samples.

Table OA-1: Timeline of Activities

Season/Date	Activity
Pre-planting 2015	
January 2015	Farmer Registration
February 2015	Soil sampling
	Baseline survey (farmer characteristics and 2014 practices)
Planting 2015	
March 2015	Delivery of soil analysis
	Orders of fertilizers
April-July 2015	Intervention
August 2015	Follow-up survey (2015 practices)
Harvest 2015	
October-December 2015	Yield estimation
February 2016	2015 Self-reported yields survey
Commercialization 2015	
June 2016	2015 Commercialization survey (prices, sales and costs)
Planting 2017	
May 2017	Follow-up survey (2017 practices)

Table OA-1 presents the timeline of the intervention. In March 2015, the team collected information on field activities to date. Farmers were divided into 26 strata based on their location and agro-climatic conditions.[37] Individual randomization was done at the stratum level and announced at the end of March.[38]

Table Table OA-3 compares our study sample to respondents from the nationally representative INEGI survey, both in Mexico and in Tlaxcala. In Panel A we include all farmers, while Panel B restricts the sample to rainfed farmers. Study farmers have lower yields than both the national and the Tlaxcala sample. In terms of agricultural practices, study farmers are less likely to use hybrid seeds relative to the national average (but comparable to the Tlaxcala sample) and are more likely to have used fertilizer and herbicide than either of the INEGI samples. They are also more likely to have used extension services in the past. Panel B shows similar patterns and both panels suggests that farmers with greater experience and perhaps those interested in improved inputs were more likely to select into the study.

We have consistent panel data on agricultural practices in 2014, 2015 and 2017 and yield information for 2014 and 2015 for 540 farmers and they comprise the core sample for the study. The panel data includes a total of 678 farmers but 138 of them were offered individualized recommendations based on the soil analysis of their own plots. In this paper we focus on the sample of farmers that received recommendations based on averages of the soil analyses in their cluster. Corral et al. (2020) explores the

[37] Specifically, study farmers came from 54 localities (*localidades*) of which 29 were quite small and thus merged with the closest *localidad* using distance between centroids from INEGI's geographic databases. If there were ties, we chose to merge *localidades* with the closest altitude. Two of the small *localidades* were merged into one to give us a total of 26 clusters. The median cluster had 21 farmers, the 25th percentile had 15, and the 75th had 32.

[38] We note that with individual level randomization, there could be spillovers across farmers in different treatments. We study the potential spillovers in Appendix E.4.

Table OA-2: Sample attrition

	(1) Main sample
Recommendation of any type (1=Yes)	0.06
	(0.04)
Grant (1=Yes)	-0.05
	(0.04)
Flexible (1=Yes)	0.01
	(0.04)
Observations	683
R-squared	0.13
Mean dep. var. control	0.78
SD dep. var. control	0.42
NA: $\beta_R + \beta_G = 0$	0.94
A: $\beta_R + \beta_G + \beta_A = 0$	0.81

This table reports results on attrition. We run the following regression: $Y_i = \beta_0 + \beta_R \text{Recomm}_i + \beta_G \text{Grant}_i + \beta_F \text{Flexible}_i + \alpha_s + \epsilon_i$, where i corresponds to a farmer, Y_i is an indicator equal to 1 if farmer i is in the panel sample. We include randomization strata indicators and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. 683 farmers registered for the four study arms and grew maize in 2015. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table OA-3: Comparing Study Sample to Mexican Farmers

	Mexico		Tlaxcala		Study Sample	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Panel A: All plots						
Plot is rain fed (1=Yes)	0.88	0.21	0.95	0.08	0.96	0.19
Plot owner uses inorganic fertilizers (1=Yes)	0.74	0.28	0.88	0.11	0.97	0.16
Plot owner uses organic fertilizers (1=Yes)	0.20	0.21	0.31	0.20	0.42	0.49
Plot owner uses hybrid seeds (1=Yes)	0.24	0.26	0.08	0.09	0.07	0.25
Plot owner uses herbicides (1=Yes)	0.35	0.30	0.38	0.31	0.72	0.45
Plot owner uses insecticides (1=Yes)	0.21	0.24	0.12	0.12	0.11	0.32
Plot owner has access to extension services (1=Yes)	0.03	0.06	0.01	0.01	0.10	0.31
Maize yields (ton/ha)	2.73	2.50	2.67	1.66	2.01	1.11
Panel B: Rain-fed plots						
Plot owner uses inorganic fertilizers (1=Yes)	0.75	0.30	0.96	0.08	0.97	0.16
Plot owner uses organic fertilizers (1=Yes)	0.18	0.23	0.13	0.13	0.42	0.49
Plot owner uses hybrid seeds (1=Yes)	0.11	0.19	0.06	0.06	0.06	0.24
Plot owner uses herbicides (1=Yes)	0.26	0.32	0.68	0.38	0.73	0.44
Plot owner uses insecticides (1=Yes)	0.12	0.21	0.13	0.15	0.11	0.31
Plot owner has access to extension services (1=Yes)	0.03	0.06	0.05	0.06	0.10	0.30

Note: This table compares the farmers in our sample to farmers Mexican state of Tlaxcala and to Mexican farmers overall. The data on the farmers in our study sample come from the Baseline survey conducted in February 2015, while data on the representative farmers of Tlaxcala and Mexico come from the INEGI Agricultural, Livestock and Forestry Census conducted in 2007. Panel A reports summary statistics of all plots in the data, while in Panel B we report the numbers for rain-fed plots only. In Columns 1 and 2, we report the average and SDs of the variables among all farmers in Mexico, while in Columns 3 and 4 we restrict the INEGI data to farmers in the state of Tlaxcala. In Columns 5 and 6 we report figures for farmers in our study sample of 540 farmers. When calculating yields using the INEGI data, we cannot distinguish between rain-fed and irrigated plots, so we do not report yields in Panel B.

Table OA-4: Summary Statistics

	Mean (SD)	p-value		Mean (SD)	p-value
Panel A: Farmer Characteristics					
Male (1=Yes)	0.83	0.78	Annual income in 2014 (000s pesos)	29.41	0.49
	(0.38)			(30.88)	
Age (years)	56.95	0.70	Reported liquidity constraints (1=Yes)	0.52	0.74
	(13.52)			(0.50)	
Years as farmer	34.90	0.31	Ever applied for a loan (1=Yes)	0.20	0.75
	(16.15)			(0.40)	
Completed primary school (1=Yes)	0.60	0.90	Never takes risks (1=Yes)	0.13	0.21
	(0.49)			(0.34)	
Panel B: 2014 Cultivation Characteristics & Yields					
Number of plots cultivated in 2014	2.09	0.33	Farmer sold maize in 2014 (1=Yes)	0.50	0.55
	(1.27)			(0.50)	
Total area cultivated (ha) in 2014	5.77	0.89	Share of maize production sold in 2014	0.28	0.97
	(4.45)			(0.34)	
Number of fertilizations in 2014	1.58	0.93	Ever had extension services (1=Yes)	0.06	0.09
	(0.54)			(0.24)	
Number of days between sowing and 1st fertilization in 2014	36.32	0.94	Ever done soil analysis (1=Yes)	0.15	0.59
	(21.48)			(0.35)	
Self-reported yield (ton/ha) in 2014	2.06	0.92	Supported by a government program in 2014 (1=Yes)	0.87	0.90
	(1.16)			(0.33)	
Panel C: Plot Characteristics					
Sand (%)	70.93	0.96	Nitrogen (N) (ppm)	14.14	0.04
	(9.96)			(7.00)	
Clay (%)	14.71	0.95	Phosphorus (P) (ppm)	28.18	0.40
	(5.43)			(25.54)	
Slope (degrees)	3.24	0.06	Potassium (K) (ppm)	198.98	0.11
	(3.43)			(137.43)	
pH (1:2 water)	6.07	0.02	CEC (cmolc/kg)	9.60	0.33
	(0.71)			(8.54)	
Coal (%)	0.15	0.16			
	(0.45)				

Note: This table reports summary statistics for the study sample of 540 farmers in 2014. For each variable, we report the mean and standard deviations. We also regress the baseline values of each variable on a set of treatment dummies and report the p-value of a F-test that all treatment coefficients are jointly equal to zero. The data were collected using our Baseline survey conducted in February 2015, before the start of the intervention. See Appendix Table
reftab:defvars for the definition of the variables.

impact of specificity of recommendations and does not find substantive differences in outcomes for an arm that received individual plot-based recommendations. In addition, we lack data for some farmers because they did not sow maize in a particular year (less than 5%) or because they could not be located. Table OA-2 shows that attrition was not differential across experimental arms. Attrition in 2017 was also uncorrelated with self-reported yields in 2014 or 2015 (results available upon request).

Table OA-5 lists recommended practices and reports their prevalence before the intervention for farmers in each treatment arm and the control arm (columns 1-4), and during the intervention for the control group (column 5).[39] Column 6 reports the p-value of a joint F-test that adoption rates are similar among all treatment and control groups in 2014, while Column 7 reports the p-value of a t-test that adoption rates among control farmers in 2014 and 2015 are similar. In Column 6, there is balance across experimental arms for all practices except for "Ripping". In Column 7 there are no substantial differences (the lowest p-value is 0.17), suggesting that spillovers between treated and control farmers during the intervention were limited.

[39]Existing practices comprise ploughing (56% in 2015), the use of inorganic fertilizers (97%) and covering the applied fertilizer (85%). New practices included deep tillage or ripping (5% in 2015), using hybrid seeds (5%), sowing with a precision drill (10%), fertilizing at sowing (9%), application of pre-emergent herbicide (2%) and using high-quality fertilizers (manufactured by YARA) (0%). We did not ask about covering the fertilizer, using high-quality fertilizers or using pre-emergent herbicide after sowing at baseline, so they are only available in 2015.

Table OA-5: Agricultural practices balance in 2014 and control 2015

	(1) NA	(2) A	(3) R	(4) Control	(5)	(6) P-values	(7)
	2014	2014	2014	2014	2015	(1)-(4)	(4) vs. (5)
Panel A: Existing practices							
Ploughing	0.56	0.59	0.56	0.49	0.56	0.35	0.23
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)		
Using inorganic fertilizer	0.98	0.96	0.98	0.98	0.97	0.59	0.45
	(0.15)	(0.20)	(0.15)	(0.12)	(0.17)		
Covering fertilizer					0.85		
					(0.36)		
Panel B: New practices							
Ripping	0.15	0.13	0.07	0.05	0.05	0.03	0.77
	(0.36)	(0.33)	(0.25)	(0.23)	(0.21)		
Using hybrid seeds	0.05	0.09	0.08	0.04	0.05	0.24	0.75
	(0.22)	(0.29)	(0.28)	(0.19)	(0.21)		
Fertilizing at sowing	0.09	0.07	0.09	0.06	0.09	0.68	0.17
	(0.29)	(0.26)	(0.28)	(0.24)	(0.29)		
Sowing with precision machinery	0.13	0.11	0.12	0.11	0.10	0.99	0.81
	(0.33)	(0.32)	(0.33)	(0.31)	(0.30)		
Using high-quality fertilizers (YARA)					0.n00		
					(0.00)		
Using pre-emergent herbicide after sowing					0.02		
					(0.15)		

Note: This table reports prevalence of existing and new agricultural practices for each treatment and control groups. Columns 1 to 4 report the percentage of farmers in each treatment and control groups that reported doing each of the existing and new practices in 2014. Column 5 reports the same percentage for farmers in the control group in 2015. Standard errors are in parentheses. In 2014, we did not collect data on 2 of the new practices and one old practice (hence the blank entries). For each of the practices, we pool the 2014 data for the farmers in each group and regress a dummy that takes value of 1 if farmers reported doing the corresponding practice against a year dummy and a set of strata fixed effects. Column 6 reports the p-value of the joint F-test that the means in columns 1-4 are equal. Column 7 reports the p-value of the t-test that the means in columns 4 and 5 are equal. Data for 2014 practices come from the Baseline survey conducted in February 2015, while data on 2015 practices were collected in August 2015 in our Follow-up survey.

A.1 Variable Definitions

Table OA-6: Definition of variables

Variable	Definition
Panel A: Farmers characteristics	
Annual income in 2014 (000s pesos)	Total income earned by farmer in 2014, including, but restricted to, sales from agricultural activities, labor earning in agricultural and non-agricultural activities, sales of animals, remmitances, pensions and cash transfers. Collected using our Baseline survey conducted in February 2015.
Reported liquidity constraints (1=Yes)	Dummy that takes value 1 if farmer reported above-average amount when asked the following question: "How much money per hectare were you missing in order to sow the way you would have wanted?". Collected using our Baseline survey conducted in February 2015.
Ever applied for a loan (1=Yes)	Dummy that takes value 1 if farmer answered "Yes" to the following question: "Have you ever, in your entire life, applied for credit or a loan for matters related to agriculture?". Collected using our Baseline survey conducted in February 2015.
Never takes risks (1=Yes)	Dummy that takes value of 1 if farmer selected the first option when asked the question "Do you consider yourself a risk taker? You would say:" and given the following options: "1. Does not like taking risks", "2. Almost never take risks", "3. Sometimes yes, sometimes no", "4. Almost always takes risks", "5. Always likes to take risks". Collected using our Baseline survey conducted in February 2015.
Panel B: 2014 Practices & Yields	
Number of plots cultivated	Number of plots farmers reported working on as owner or tenant in 2014. Collected using our Baseline survey conducted in February 2015.
Total area cultivated (ha)	Number of hectares farmers reported working on as owner or tenant in 2014. Collected using our Baseline survey conducted in February 2015.
Supported by a government program in 2014 (1=Yes)	Dummy that takes value of 1 if farmers reported being supported by any of the following input subsidy programs in 2014: PROCAMPO, PIMAF, MASAGRO or Agroincentivos. Collected using our Baseline survey conducted in February 2015.
Panel C: Grant flexibility outcomes	
Trust in the recommendation from input supplying institutions (standardized index)	Standardized index of two individual dummies that take value 1 if the farmers reported trusting the recommendations given by their input suppliers and IPAMPA, a local AES company. Computed by standardizing each dummy individually, adding them all and standardizing the sum. We use the mean and standard deviation of the control group as reference for the standardized index. Collected in August 2015 using our Follow-up survey. 40 among our sample of 678 farmers refused to answer these questions.

B Soil Analyses and Recommendations details

Soil samples were collected from the designated sub-plot during February and March 2015. Surveyors divided up the sub-plot into (up to) 6 relatively homogeneous regions and took 15 soil samples (from a depth of 30 cm). These 15 samples were then mixed and collected in bags following standard soil analysis protocols. These bags were then sent to Fertilab for analysis.[40] Based on focus group discussions conducted in December 2014 we developed a template for reporting the soil analysis and recommendations divided into three parts:

B.1 Soil Analysis

The soil analysis provided the main soil characteristics in a relatively easy to read format for farmers. The soil analysis measured a range of factors that measured the soil texture (percentage of sand, silt and clay) its ability of retain and transfer nutrients (pH levels, sand and lime concentrations, saturation points and cationic exchange capacity) as well as the levels of 13 key nutrients – the primary macro-nutrients (N, P, K), the secondary macro-nutrients (Ca, Mg, S) and selected micronutrients (Na, Fe, Zn, Mn, Cu, B, Al) – and the level of organic matter in the soil.

Nitrogen (N) affects plant growth. Many spoil microorganisms found in the soil are able to convert organic N found in plant residue, soil organic matter, or bacteria into inorganic N forms that can be taken up by plans. plant available inorganic Ammonium (NH_4^+) and nitrate (NO_3^-) are such forms of mineral or inorganic N. Nitrate NO_3^- is water soluble and does not remain in the soil.

Phosphorous (P) is critical in root development, crop maturity and seed production. P deficiency is a common problem causing crop stunting or discoloration in the field. One of the major contributing sources of P for crops comes from soil organic matter.

Potassium (K) is required for the activation of enzymes throughout. It is critical for the crop's ability to withstand extreme cold and hot temperatures, drought and pests. Potassium increases water use efficiency and transforms sugars to starch in the grain-filling process.

Calcium (Ca), magnesium (Mg), and sulfur (S), are considered secondary macronutrients, because they are required in amounts smaller than those typically needed for N, P, or K. These elements, however, are equally important for plant growth and nutrition.

Micronutrients are essential nutrients for plant growth that are used in relatively small amounts by crops. Boron (B), zinc (Zn), manganese (Mn), iron (Fe) and copper (Cu) will only make up a small proportion of a plant; however, a deficiency in any of these elements has the potential to cause a decrease in crop quality or yield.

In a following subsection we discuss the stability over time of the soil analyses.

[40] Fertilab is one of the best known laboratories in Mexico and is accredited by the North American Proficiency Testing Program (run by the Soil Sciences Society of America) that certifies laboratory operations in the United States and elsewhere.

B.2 Fertilab Original Soil analysis

The following page contains the information provided by Fertilab. Because it is technical, we provided farmers with simple interpretations and personalized fertilization recommendations based on it.

Figure OA-2: Fertilab Original Soil Analysis

Fertilab®
Fertilidad de Suelos S. de R.L.

Análisis que
Rinden Frutos

Laboratorio Acreditado en E.U. de Norteamérica — PAP-Accredited NAPT 2014
International Organization for Standardization — EMPRESA CERTIFICADA EN ISO 9001:2008

DIAGNOSTICO DE LA FERTILIDAD DEL SUELO

INFORMACIÓN GENERAL

Cliente	Ismael Zacamolpa Cerbani		
No. de Registro	SU-35440	Cultivo Anterior	Ninguno
Fecha de Recepción	09/03/2015	Cultivo a Establecer	Maiz
Fecha de Entrega	11/03/2015	Tipo de Abono Organico	N/A
Rancho o Empresa	Cuaxomulco	Tipo de Agricultura	Temporal
Municipio	Cuaxomulco	Manejo de Residuos	Retirados
Estado	Tlaxcala	Meta de Rendimiento	5 Ton/Ha Ton/Ha
Identificación	.23.01.10.01	Prof. Muestra	0-30 cm

Propiedades Físicas del Suelo

Clase Textural	Franco Arcillo Arenoso		
Punto de Saturación	31.6	%	Mediano
Capacidad de Campo	16.7	%	Mediano
Punto March. Perm.	9.94	%	Mediano
Cond. Hidráulica	6.00	cm/hr	Mod. Alto
Dens. Aparente	1.35	g/cm3	

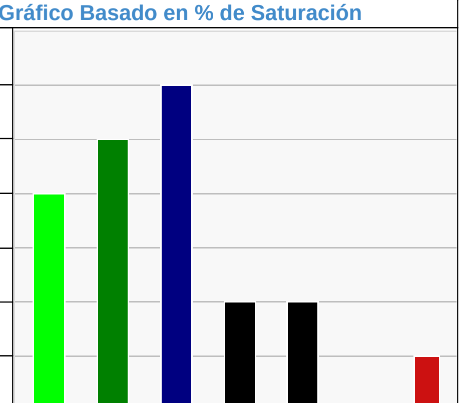
Reacción del Suelo Necesidades de Yeso y Cal Agrícola

pH (1.2 agua)	5.12		Acido
pH Buffer	6.90		
Carbonatos Totales (%)	0.01	%	Libre
Salinidad (CE Extracto)	0.30	ds/m	Muy Bajo
Requerimientos de Yeso	No Requiere		
Requerimientos de Cal	0.00		

Fertilidad del Suelo

Det	Result	Unid	Muy Bajo	Bajo	Mod. Bajo	Med.	Mod. Alto	Alto	Muy Alto
MO	1.11	%							
P-Bray	61.2	ppm							
K	121	ppm							
Ca	633	ppm							
Mg	90.0	ppm							
Na *	19.5	ppm							
Fe	34.3	ppm							
Zn	0.42	ppm							
Mn	7.70	ppm							
Cu	0.45	ppm							
B	0.13	ppm							
Al *	12.2	ppm							
S	13.8	ppm							
N-NO3	22.7	ppm							

Cationes Intercambiables
Gráfico Basado en % de Saturación



% Sat	71.3	16.7	7.00	1.81	3.16	ND	—
meq/100g	3.16	0.74	0.31	0.08	0.14	0.00	4.43
Catión	Ca	Mg	K	Na*	Al*	H*	CIC

Relacion Entre Cationes (Basadas en me/100g)

Relación	Ca/K	Mg/K	Ca+Mg/K	Ca/Mg
Resultados	10.2	2.39	12.6	4.27
Interpretación	Mediano	Mediano	Bajo	Mediano

* Es deseable que estos elementos tengan un bajo contenido

Interpretación Resumida del Diagnostico de la Fertilidad del Suelo

Suelo con pH acido. Suelo de textura media. Libre de carbonatos. Libre de sales. Bajo nivel de materia organica, es recomendable su aportacion. Bajo nivel de calcio. Muy alto suministro de fosforo disponible.Contenido bajo de potasio. Bajo nivel de magnesio. Suministro moderado en nitratos.

En cuanto a la disponibilidad de micronutrientes: Pobre en zinc. Bajo contenido de cobre. Muy pobre en boro.

Poniente 6. No. 200 Ciudad industrial
Celaya, Gto. C.P. 38010
Tel. (461) 614 5238, 614 7951
www.fertilab.com.mx

Supervisor de Análisis de Suelos
Ing. José Trinidad Guzmán M.

Programa de Calidad e Intercomparación de Análisis de Suelos y Plantas — ISP
Wageningen Evaluating Programmes for Analytical Laboratories — WEPAL

B.3 Stability of soil characteristics

In February, 2017 we visited a randomly chosen set of 99 control plots and re-did the soil analysis to measure the stability of the nutrient content in the soil. Table OA-7 in this online appendix shows that there are large and precise correlations across years, particularly for macronutrient, so that the information from the 2015 soil analysis remained relevant in 2017.[41]

Table OA-7: Soil Analysis comparison 2017 vs. 2015

$Y_{2017} = \alpha + \beta Y_{2015} + \epsilon$			
Soil characteristic	α	β	R^2
pH	2.71***	0.63***	0.51
	(0.37)	(0.06)	
Organic Matter (OM)	0.10	0.89***	0.60
	(0.06)	(0.07)	
Nitrogen (N)	3.49**	0.31**	0.22
	(0.99)	(0.06)	
Phosphorus (P)	6.84***	0.70***	0.82
	(1.74)	(0.03)	
Potassium (K)	64.99**	0.76***	0.52
	(17.27)	(0.07)	
Calcium (Ca)	1,447.66***	0.10	0.01
	(161.34)	(0.08)	
Magnesium (Mg)	37.84**	0.97***	0.36
	(28.02)	(0.13)	
Sodium (Na)	8.90***	0.44***	0.27
	(1.75)	(0.07)	
Iron (Fe)	7.27***	0.52***	0.64
	(1.48)	(0.04)	
Zinc (Zn)	0.10	0.64***	0.90
	(0.03)	(0.02)	
Manganese (Mn)	3.09**	0.26***	0.18
	(0.82)	(0.06)	
Copper (Cu)	0.17***	0.62***	0.93
	(0.02)	(0.02)	
Boron (B)	1.06***	0.19**	0.09
	(0.01)	(0.06)	

Note: n = 99 ; do-file: APPENDIX_SA_2015vs2017.do. Datasets: Soil analysis (2015 and 2017).

[41] Due to this persistence in the characteristics of the soil content, the USDA recommends that soil tests be carried out every 3-5 years (see e.g. `https://perma.cc/E8GN-GWGM`).

B.4 Recommendation Details

An sample report is available online.[42] The first page explained the program and required a signature from the farmer for consent. The second page provided basic information about the plot's physical characteristics (e.g. texture, saturation, organic matter, pH level and bulk density). It also provided the nutrient levels in the plot (e.g. N, P, K and secondary macronutrients and micronutrients) as well as the required levels of nutrients for a maize yield of 4.5 mt/ha under normal weather conditions.[43] Recommendations were based on a proprietary model that assumed that a certain quantity of N, P, K and micronutrients were needed to reach a target yield per hectare. The model is grounded in the Law of the Minimum formulated by J.V. Liebig in the 1850s which suggests that to reach a target yield, a certain quantity of each nutrient is needed (similar to a Leontief production function). The Fertilab model used this as well as a cost minimization approach given the price of available fertilizers. For example, for N one can use urea, DAP or ammonium sulfate. Taking into account the cost of the different fertilizers and the soil absorption capacity, the model selected the cheapest fertilizer package that met the nutrient requirements. For instance, if the soil were pH negative (alkaline), then ammonium sulfate was preferred to urea, but if the soil were pH positive (acidic) then urea was preferred. There are other maize yield models such as CERES and NLEAP but many of the variables and parameters required by these models are unknown for Tlaxcala (and Mexico in general). Empirical tests of the appropriateness of the Von-Leibig type production function in agriculture typically reject it (see e.g. Berck et al., 2000).

The third page provided a "shopping list", that is, the list of recommended fertilizer amounts (DAP, urea, KCl and micronutrients) and its cost at our partner agro-dealer. Costs were divided into the portion paid by the research team and the remainder which the farmer was expected to provide (if they went above the subsidy amount). The fourth and fifth pages compared the farmer's own 2014 input use and costs (from the baseline survey) to the recommended input mix and costs. They also provided sub-plot yields and prices from 2014. These 2014 costs and revenues were compared with the expected yields, revenues (using 2014 prices) and costs of inputs if the recommendations were followed and Fertilab's assumptions (about weather and temperature) proved accurate – the research teams were careful to explain the assumptions underlying the yield predictions.

[42] https://are.berkeley.edu/~aprajit/autonomyreport.pdf

[43] In theory, the target yield should be chosen to maximize farmer profits. The model, however, assumed that yields were roughly linear in inputs and we chose the target of 4.5 mt/ha because it equated the average cost of fertilizers (according to the model) to the average baseline farmer expenditure in fertilizer .

Figure OA-3: Cupon

Qfd

jose luis castillo montes

ignacio allende

cuapiaxtla

Vale por 2000 pesos para fertilizante y/o maquinaria para la siembra*

Lugar para pasar a recoger los fertilizantes: Fertilizantes YARA Carretera Mexico-Veracruz Km 145.5, Huamantla, Tlaxcala en la fecha que se lo indique su ingeniero agrónomo

Número de FOLIO: 842 ID: .22.03.02.

Si tiene dudas contacte a los teléfonos: (246) 4626577 o (247) 4720603

*El costo de la maquinaria es de $800 pesos si decide rentarla con nosotros; sólo la cantidad de dinero restante podrá ser usada para la compra de fertilizante

Figure OA-4: Fertilization recommendations

Fertilizer Dosis in kg/ha	TIME OF APPLICATION		
	Sowing (kg/ha)	**1st fertilization (kg/ha)**	**Total Kg**
Urea (white)			
DAP (Black)			
Potassium Chloride (Red)			
Microelemets			
Cost	$	$	$

Farmers were provided the fertilization recommendation form above (filled out with their recommendation). Fertilizer requirements were specified for each of the four different types of fertilizer and two different application timings.

B.5 Fertilizer Quality

Fertilizer quality: In addition to differences in the fertilizer mix and timing, farmers were unfamiliar with the recommended fertilizer brand YARA, a reputed, high-quality manufacturer, stocked by the agro-dealer. In order to assess fertilizer quality, the research team tested samples in a laboratory for each of the three main fertilizers (urea, DAP and KCl) manufactured by YARA (from five different locations) and by the most popular manufacturer of government subsidized fertilizer. We found that the urea and KCl content was comparable across the two types of manufacturer and generally matched the labelled concentrations. However, DAP concentrations were lower than advertised in the commonly subsidized brand relative to YARA. In fact after accounting for differences in concentrations, the cost per kilogram of nutrient was actually lower for YARA.

Table OA-8: Lab analysis of nutrient content of YARA and government fertilizers

	(1)	(2)	(3)	(4)	(5)
		Government		YARA	
	Label (%)	Lab test (%)	Cost per kg of nutrient (Mex$)	Lab test (%)	Cost per kg of nutrient (Mex$)
Panel A: Urea					
Nitrogen (N)	46	46.72	13.02	47.00	13.62
Phosphorus (P)	0	0.00	.	0.00	.
Potassium (K)	0	0.00	.	0.00	.
Cost of 50kg bag (Mex$)		304		320	
Panel B: DAP					
Nitrogen (N)	18	10.40	82.69	16.70	55.09
Phosphorus (P)	46	14.00	61.43	36.20	25.41
Potassium (K)	0	0.00	.	0.00	.
Cost of 50kg bag (Mex$)		430		460	
Panel C: KCl					
Nitrogen (N)	0	0.00	.	0.00	.
Phosphorus (P)	0	0.00	.	0.00	.
Potassium (K)	60	51.23	12.26	53.10	15.82
Cost of 50kg bag (Mex$)		314		420	

Notes: This table reports the nutrient content advertised by our partner YARA and the fertilizer brand subsidized by the Mexican government to actual nutrient content measured in a laboratory test. Using the price of a 50kg bag of each fertilizer, we also compare the average cost per kg of nutrient between YARA and the government-subsidized brand. Panel A reports the figures for Urea bags, while Panels B and report numbers for DAP and KCl bags. In column 1 we show the percentages of each nutrient (Nitrogen, Phosphorus and Potassium) reported on commercial labels of each bag. In columns 2 and 4 we report the percentages measured in the lab. In the last column of each panel we report the (average) price of each bag of fertilizer. We then divide the price of the bags by the nutrient percentages and report the cost per nutrient percentage in columns 3 and 5.

The research team tested samples for each of the three main fertilizers —urea, DAP and KCl— in the laboratory Laboratorios A-L de México, in Guadalajara, México. Samples came from five different locations – Altlzayanca, Apizaco, Calpulalpan, Cuapiaxtla and Muñoz – from YARA and the most popular manufacturer who provides government subsidized fertilizer.

Table OA-8 presents the results from our fertilizer testing exercise. The label of a bag of urea (Panel A) shows an NPK content of 46-0-0, so that 46% of the contents should be N. According to the laboratory tests, the commonly used bag had a content of 46.7% while the YARA bag had a content of 47%. Panel A also reports the total cost per bag which allows us to compute the cost per kilogram of nutrient at 13 pesos for the government subsidized bag compared to 13.6 pesos per kg of N in the YARA bag. We conclude that both urea bags have similar content and price per unit of nutrient. The results are similar for KCl (Panel C) although both bags have lower content of K than advertised. The YARA bag is a bit more expensive and thus its cost per kg of K is slightly higher. In Panel B however, we see that the subsidized DAP bag has much lower content of N and P than advertised. The label for DAP is 18-46-0, indicating that there should be 18% N and 46% P. According to the laboratory test, however, the government bag only had 10.4% of N and 14% of P. In contrast, the YARA bag had 16.7% of N and 36.2% of P. Therefore, even though the YARA bag was more expensive, its cost per kg of nutrient was in fact lower. We conclude that the YARA bag of DAP was of higher quality than the government subsidized one (and was in fact cheaper after adjusting for quality)

C Practices

Table OA-9: Individual practices 2015

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Existing practices			New practices					
	Ploughing	Using inorganic fertilizer	Covering fertilizer	Ripping	Using hybrid seeds	Fertilizing at sowing	Sowing with precision machinery	Using high-quality fertilizer (YARA)	Using pre-emergent herbicide after sowing
Recommendation (1=Yes)	0.03	0.00	0.01	0.02	0.05*	0.08**	0.08*	0.10***	0.01
	(0.06)	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)
Grant (1=Yes)	-0.03	0.02*	0.05	0.05	-0.02	0.69***	0.66***	0.81***	0.04
	(0.06)	(0.01)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)
Autonomy (1=Yes)	-0.04	0.00	0.03	-0.00	0.02	-0.10**	-0.11**	0.06*	0.04
	(0.06)	(0.00)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)
Observations	540	540	540	540	540	540	540	540	540
R-squared	0.09	0.06	0.07	0.21	0.11	0.53	0.47	0.81	0.09
Mean dep. var. control	0.56	0.97	0.85	0.05	0.05	0.09	0.10	0.00	0.02
SD dep. var. control	0.50	0.17	0.36	0.21	0.21	0.29	0.30	0.00	0.15
NA: $\beta_R + \beta_G = 0$	0.92	0.08	0.12	0.04	0.24	0.00	0.00	0.00	0.04
A: $\beta_R + \beta_G + \beta_A = 0$	0.42	0.07	0.02	0.04	0.08	0.00	0.00	0.00	0.00

Note: this table reports results on the agricultural practices performed by the farmers in our study in the 2015 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recommendation_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Follow-up survey conducted in August 2015. The dependent variable in each column is a dummy that takes value of 1 if farmers reported performing each practice. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table OA-10: 2017 Individual practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Existing practices			New practices					
	Ploughing	Using inorganic fertilizer	Covering fertilizer	Ripping	Using hybrid seeds	Fertilizing at sowing	Sowing with precision machinery	Using high-quality fertilizer (YARA)	Using pre-emergent herbicide after sowing
Recommendation (1=Yes)	0.05	0.03	0.03	0.02	0.02	0.07**	0.02	0.05	0.03
	(0.06)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)
Grant (1=Yes)	-0.16**	-0.03	-0.02	0.04	0.00	-0.00	-0.05	0.05	0.04
	(0.06)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.03)
Autonomy (1=Yes)	0.03	0.01	0.01	0.02	0.07*	0.06	0.10**	0.17**	0.04
	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
Observations	540	540	540	540	540	540	540	540	540
R-squared	0.12	0.06	0.05	0.08	0.13	0.20	0.22	0.12	0.09
Mean dep. var. control	0.52	0.91	0.88	0.07	0.06	0.06	0.12	0.07	0.04
SD dep. var. control	0.50	0.29	0.32	0.25	0.24	0.24	0.32	0.25	0.19
NA: $\beta_R + \beta_G = 0$	0.09	0.91	0.79	0.08	0.53	0.05	0.37	0.02	0.02
A: $\beta_R + \beta_G + \beta_A = 0$	0.22	0.95	0.51	0.03	0.01	0.00	0.08	0.00	0.00

Note: this table reports results on the agricultural practices performed by the farmers in our study in the 2017 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recommendation_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Final survey conducted in May 2017. The dependent variable in each column is a dummy that takes value of 1 if farmers reported performing each practice. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

D More on Take up

Table OA-11 uses the sample of farmers that received soil analyses and recommendations during the intervention and examines the take up of the precision drill during sowing (column 1), the two fertilizer packages (columns 2 and 3), attendance at AEW group meetings and the total number of AEW plot visits (columns 5 and 6). Column 7 uses as the dependent variable the sum of the dependent variables in columns 1-3, 5 and 6 while column 8 uses a standardized index of the outcome in column 7. The take up of these items was verified both from farmer reports as well as administrative data and as a result, we do not believe mis-classification is a serious concern. The penultimate row of Table OA-11 reports the mean of the dependent variable among farmers in the no-grant arm with the corresponding standard errors.[44]

Precision drill. The take up of the precision drill among farmers in recommendation only arm was 8%. Receiving the grant increased the probability of take up among farmers in NA by 78 percentage points (pp), a near ten-fold increase. Recall that farmers with the inflexible grant forfeited the rental amount for the precision drill if they did not use it, unlike farmers in A. Perhaps unsurprisingly, farmers with the flexible grant were 13pp less likely to rent the precision drill than farmers with the inflexible grant (although 57pp more likely to rent the drill relative to no-grant farmers).

Fertilizer packages. The take up of the first fertilizer package (column 2) among farmers in the recommendation only arm (Arm R) is 7%, but increases by 83pp for farmers with the inflexible grant, a more than a ten-fold increase. The increase was comparably large, at 89pp, for farmers with the flexible grant.[45] Take-up rates for the second package (column 3) are somewhat lower than those for the first package: 4% for farmers who received just the recommendations and extension services and an increase of 75pp for farmers who received the inflexible grant. Farmers with the flexible grant took up the second package at very similar rates.

Costs of Second fertilizer package. Take-up rates for the second package among grant recipients are lower than those for the precision drill and the first package likely because the grant typically did not cover the full cost of the second fertilizer package (while it typically fully covered the costs of the sowing machinery rental and the first package). In fact, farmers with the inflexible grant needed to pay 319 pesos (on average) out-of-pocket to purchase the second package. Column 4 shows these out-of-pocket expenses that farmers with the grant made to cover the cost of the second package. As expected, farmers with autonomy spend less out of pocket on the second package compared to farmers without autonomy because they were less likely to rent the precision sowing drill and thus used the rental amount towards the second package.

Extension services. Turning next to extension services, columns 5 and 6 record the number of AEW led group sessions attended by the farmer and the number of visits by AEWs at farmer plots where the farmer was present, respectively. These sessions and plot visits were described in Section 3 and functioned as tutorials and Q&A sessions on best practices for maize cultivation. Farmers in the no-

[44]The control group is not included in these regressions as no intervention was offered to them.

[45]While the first package was to be applied at sowing with the precision drill to guarantee an optimal spread of fertilizer, farmers who did not use the precision drill were instead advised to use the first package 30-60 days after sowing depending upon plant growth.

Table OA-11: Take up

	(1) Precision drill (1=Yes)	(2) 1^{st} Package (1=Yes)	(3) 2^{nd} Package (1=Yes)	(4) Out-of-pocket (Mex$/ha)	(5) # training sessions (0-3)	(6) # AEW visits (0-3)	(7) Total (0-9)	(8) Total (Std. Index)
Grant (1=Yes)	0.78***	0.83***	0.75***	319.59***	1.28***	1.25***	4.89***	3.73***
	(0.04)	(0.03)	(0.04)	(22.13)	(0.11)	(0.13)	(0.26)	(0.16)
Autonomy (1=Yes)	-0.13**	0.06**	-0.02	-107.21***	0.19**	-0.01	0.08	-0.06
	(0.05)	(0.03)	(0.05)	(28.00)	(0.09)	(0.09)	(0.24)	(0.16)
Observations	410	410	410	410	410	410	410	410
R-squared	0.50	0.75	0.54	0.41	0.44	0.40	0.62	0.69
Mean dep. var. R	0.08	0.07	0.04	0.00	0.76	1.40	2.34	0.00
SD dep. var. R	0.27	0.25	0.19	0.00	0.95	1.28	2.11	1.00

Note: this table reports results on the take-up of our proposed treatment by farmers in our sample. Using only the set of 410 treated farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome variable for the group of farmers who did not receive the grant (R), the omitted category in our regression. In column 1, the dependent variable is a dummy that takes value 1 if the farmer used the precision machinery provided by our team to fertilize at sowing. In column 2, the dependent variable is a dummy that takes value 1 if the farmer took up the first package of YARA fertilizers, that should be applied at sowing using the precision machinery. Farmers who did not use the sowing machinery were advised to use this package 30-60 days after sowing depending on how their crop grew. In column 3, the dependent variable is a dummy that takes value 1 if the farmer took up the second package of YARA fertilizers, that should be applied 45 days after sowing. In column 4, the dependent variable is the amount that each farmer had to pay for the packages of fertilizers, on top of the QFD subsidies, if they received them. In column 5, the dependent variable counts the number of training sessions each farmer attended, out of 3 our team organized. The first training session introduced farmers on how to fertilize at sowing. The second one aimed at on harvesting and preparations for yield measurement, as well as soil preparation for the following season. The third training session covered the importance of using quality fertilizers and herbicides, as well as on the right timing to fertilize during plant development. In column 6, the dependent variable counts how many times the farmer was visited by the our team to be provided with technical assistance (out of 3 scheduled visits). In column 7, the dependent variable is the sum of columns 1, 2, 3, 5 and 6, varying from 0 to 9. In column 8, the dependent variable is a standardized index of the outcome in column 8, computed by standardizing each variable individually, adding them all and standardizing the sum. We use the mean and standard deviation of R as reference for the standardized index. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

grant arm attended an average of 0.76 sessions (column 5) and had 1.4 plot visits (column 6). Farmers with the inflexible grant attended 1.28 additional sessions and had 1.25 additional plot visits. Strikingly, farmers with the flexible grant attended 1.47 more sessions (relative to no-grant farmers), about a fifth of a session more than farmers with the inflexible grant (the number of AEW plot visits were the same for both groups). Autonomy, therefore, seems to generate more engagement.[46]

Indices of practices. Column 7 summarizes the previous columns by simply recording the total number of adopted program components (nine in total).[47] Farmers with just the recommendations and extension services adopted 2.34 components while farmers who also received the inflexible grant adopted an additional 4.89 components, confirming the importance of the in-kind grant. Interestingly, farmers with the flexible grant (who had no obligation to choose any of the nine components) adopted the same number on average as those who did not have autonomy – the point-estimate is in fact slightly higher (4.97 components). It thus appears that there was no trade-off between autonomy and short-term compliance. Finally, column 8 reports results using a standardized index (measured in standard deviations) with similar results .[48]

[46]The overall take up of the group training sessions declines over time with the most attendance around sowing and the least attendance before the harvest (results available upon request).

[47]These were the use of precision drill, 1st and 2nd fertilizer package, 3 group sessions and being present in the 3 plot visits by AEWs.

[48]We follow the convention of standardizing each variable, summing the standardized variables, and re-standardizing again so that the index has mean zero and variance 1 in *E*3.

Table OA-12: Fertilizer usage in 2015: applied

	(1)	(2)	(3)	(4)	(5)	(6)
	Applied fertilizer, 2015					
	At sowing			Total		
	Urea (kg/ha)	DAP (kg/ha)	KCl (kg/ha)	Urea (kg/ha)	DAP (kg/ha)	KCl (kg/ha)
Recommendation	3.74	1.64	-1.25	10.38	3.71	3.93
	(3.45)	(2.83)	(1.74)	(15.70)	(6.81)	(5.24)
Grant (1=Yes)	23.54***	10.06***	11.81***	-36.07**	-14.37**	9.56**
	(3.00)	(2.48)	(1.03)	(12.46)	(5.58)	(3.78)
Autonomy (1=Yes)	-5.15**	-1.71	-1.01	0.55	-0.64	3.97
	(2.12)	(1.70)	(1.10)	(8.59)	(3.53)	(2.55)
Observations	540	540	540	540	540	540
R-squared	0.28	0.26	0.30	0.10	0.13	0.13
Mean dep. var. control	5.77	4.50	2.88	188.35	34.65	14.96
SD dep. var. control	26.88	21.85	16.95	131.38	57.21	42.06
NA: $\beta_R + \beta_G = 0$	0.00	0.00	0.00	0.04	0.06	0.00
A: $\beta_R + \beta_G + \beta_A = 0$	0.00	0.00	0.00	0.05	0.04	0.00

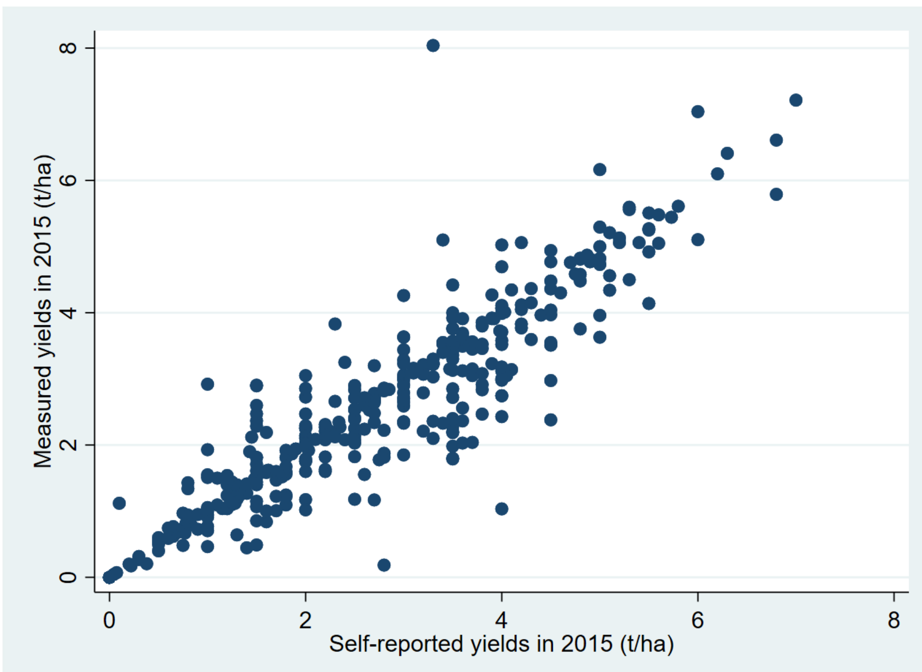
Note: This table reports results on the usage of fertilizers by the farmers in our study in the 2015 season, in absolute value dosages. Using the sample for the full sample of 540 farmers, except for 13 farmers for which we do not have data on usage of fertilizers, we run the following regression: $Y_{it} = \beta_0 + \beta_R \text{Recommendation}_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata fixed effects and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Commercialization survey conducted in June 2016. In columns 1-3, the dependent variables are the amounts applied by farmers at sowing of Urea, DAP and KCl, respectively. In columns 4-6, we report analogous outcomes for the total amount of fertilizers applied in the full season. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

E Yield measurement protocol

As discussed in Desiere and Jolliffe (2017) among others, self-reported yields are plagued by measurement error both in the numerator (the quantity harvested) as well as in the denominator (area sown). We took two steps to minimize this problem. First, the research team demarcated the registered subplot (which was one hectare in most cases) using GPS devices. For farmers with a plot area of less than 1 ha, the research team measured how much they had and adjusted the denominator appropriately. Results are robust to excluding these plots. Second, we attempted to verify self-reported yields by transporting the harvested grain to a nearby weighing station. We were able to do so for a subset of plots.

The harvesting and weighing of yields followed different protocols depending on whether farmers had already harvested the crop by the time of the team visit or whether the harvester/thresher could reach the program plot.

Figure OA-5: Measured yield against self reported yield 2015



The correlation between the actual yield and the self reported yield was also measured, resulting in a 0.912 coefficient. This scatter plot is constructed using the sub sample of farmers for which we have both the self reported and the measured yields, 392 out of the total sample of 540 (72.6%)

E.1 Harvest by QFD

For the 376 farmers that had not yet harvested the crop and with a program plot that could be reached by the mechanized harvester/thresher, the size of the plot was verified with the pre-registered GPS coordinates and the maize production on the registered plot was harvested and threshed. The grain was then collected and loaded onto a truck and weighed in the nearest weighing station.

E.2 Harvest by farmer

For the remaining farmers that had harvested by the time the team visited the registered plot or for those farmers that had not yet harvested but whose plot could not be reached by the harvester/thresher, the following procedure was used during the QFD team visit (the QFD comprised of an agronomist, a supervisor and 2 field assistants).

If the harvested cobs were in the field, all the cobs were packed in burlap sacks provided by QFD. Each sack was sealed and stitched with raffia ribbon provided by QFD and properly identified with a label including the producer's ID, the plot's name, locality and number of harvested sack. Once all the cobs were collected, the producer moved the bags to their q home, where they were placed in a ventilated and moisture-free room for drying.

If the harvested cobs were already at the farmer's home, the QFD supervisor had to verify that the cobs from the registered plot could be identified. This was the case when the cobs were stored in a separate location from other maize production or the program plot had produced maize that could be distinguished due to color or maize variety (hybrid or creole). If identification was not possible, then the team was instructed not to proceed with the yield measurement protocol (and for these farmers we only have self-reported yields).

A day before the shelling of maize, a QFD team visited the farmer to verify that moisture content (ideally less than 16%) for the shelling.[49] The team also verified that all the bags were still sealed and unaltered. For the shelling visit, the team arrived with a freight truck to transport the grain to the weighing station after shelling.

The shelling was done with a mechanical sheller in an open space, placing a a blanket below the machine to avoid loss of grain, and placing a container to collect the grain and a sack to collect maize stalks. Cobs were fed slowly to the sheller and impurities of the threshed grain (such as maize stalk, leaves, etc) were removed.

[49] To test moisture, five cobs from different parts of a burlap were collected and a few grains from each cob were collected at random. Grain moisture was then measured with a portable grain moisture tester MT- 16.

Table OA-13: Measured yields 2015

	(1) Measured yields (t/ha)	(2) Self-reported yields (t/ha) (Measurement sample)
Recommendation (1=Yes)	0.37*	0.35
	(0.20)	(0.22)
Grant (1=Yes)	-0.12	-0.12
	(0.21)	(0.21)
Autonomy (1=Yes)	0.17	0.09
	(0.18)	(0.18)
Observations	392	392
R-squared	0.29	0.26
Mean dep. var. control	2.41	2.30
SD dep. var. control	1.30	1.41
NA: $\beta_R + \beta_G = 0$	0.18	0.22
A: $\beta_R + \beta_G + \beta_A = 0$	0.02	0.09

Note: this table reports results on the maize yields in the 2015 season, measured by our team. Using only the sample of 392 for which we measured yields using our own machinery, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recommendation_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. In column 1, the outcomes are 2015 yields measured by our team. The yield measurement was only done for a subsample of the farmers in our study. In column 2, the outcome is 2015 self-reported yields, restricting the sample to the set of farmers who had their yields measured by our team. For self-reported yields, we use data from the Commercialization survey conducted in June 2016, while the data on measured yields by our team was collected in February 2016. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table OA-14: Drought 2014-2016

	(1)	(2)	(3)	(4)	(5)	(6)
	2014		2015		2016	
	Mean	p-value	Mean	p-value	Mean	p-value
Precipitation at initial stage (mm)	77.51	0.65	78.69	0.92	65.19	0.13
	(38.55)		(36.04)		(40.27)	
Total precipitation (mm)	698.64	0.65	600.29	0.45	659.15	0.67
	(118.99)		(75.31)		(55.49)	
Suffered drought (1=Yes)	.	0.67	0.72	0.29	0.31	0.64
	(.)		(0.45)		(0.46)	

Note: this table shows precipitation measures and drought reports by farmers during the time of our study. For the 2014, 2015 and 2016 seasons, we report the average precipitation (in mm) during the 30 days following sowing and average precipitation (in mm) faced by each farmer during the whole season. For each farmer, the precipitation figures are measured by the closest station to the registered plot. Data is provided by CONAGUA. For the 2015 and 2016, we also report the share of farmers who reported facing a drought at some point in the season. Data for these reports come from the Baseline and Commercialization surveys. Columns 1, 3 and 5 show means and standard deviations of each variable for the 2014, 2015 and 2016 seasons, respectively, for our full sample. For each season, we take each variable and regress it against the set of treatment dummies. In columns 2, 4 and 6, we report the p-values of the F-tests that the dummy coefficients are all equal.

E.3 Profits 2015

Measuring profits is notably challenging for smallholder farmers (see e.g. Foster and Rosenzweig, 2010). We measured revenues and expenditures on a comprehensive set of agricultural inputs using frequent, detailed surveys throughout the growing season.[50] To calculate revenues we multiplied the price received in the sale of maize by the self-reported quantity harvested.[51] Revenues are reported in column 2 of Table 4 and not surprisingly, show a similar pattern to that for yields in column 1. Expenditures are reported in column 3. For each stage of the growing season (soil preparation, sowing, plant maintenance, and harvesting) we measured labor days in the one hectare subplot, whether it was provided by a family member or hired-in labor, and the wage paid for hired labor. We also measured other inputs, including seeds, fertilizers, sowing machinery, pesticides, herbicides, and harvest machinery and whether the cost was paid by the farmer or by the research team (i.e. we include the $2,000$ in-kind grant in the costs and impute harvesting costs if they were paid by the research team).[52] Column 3 shows that grant recipients without autonomy invested 639 more pesos/ha than farmers either the control arm or the recommendation only arm, who spent on average $5,280$ pesos/ha. Column 4 reports profits as the difference between revenues and costs. Although all the point estimates are positive and suggest profit increases in the range of 10% (A), 12% (NA), and 23% (R), they are imprecisely estimated and none of the estimates are significant at conventional levels.

In column 5, we remove the amount of subsidy and harvesting costs paid by the program from the costs to consider only the out-of-pocket investment made by each farmer. We find, unsurprisingly, an increase in profits in the range of $1,300$ pesos/ha among farmers with the grant relative to control farmers.

E.4 Spillovers

Using GPS coordinates for each subplot we also assessed whether control farmers with plots located close to those of treated farmers (controlling for the total number of nearby study plots defined with reference to a 500m or 1000m radius) are more likely to adopt the new practices. The intuition is that while the density of study farmers nearby is endogenous, the share of those farmers that is treated is exogenous by virtue of randomization, and so if spillovers were significant, one should detect larger changes in the adoption of recommended practices among control farmers near treated farmers. We find no evidence of any such spillovers (results available upon request).

[50] Plots were visited several times by the team during the year. We note, however, that unpaid labor is however not taken into account in our calculations because of the difficulty in imputing a shadow wage.

[51] Only about 70% of farmers sold maize, and we imputed the price of maize for the remaining farmers using the median price in their cluster.

[52] Table OA-15 disaggregates expenditures into different categories.

Table OA-15: Cost disaggregation 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Labor costs		Capital costs		Input costs	
	Mex$	% total	Mex$	% total	Mex$	% total
Recommendation of any type (1=Yes)	-49.95	-0.01	35.11	0.00	39.43	0.02
	(215.59)	(0.02)	(82.23)	(0.02)	(128.93)	(0.02)
Grant (1=Yes)	-111.07	-0.08***	546.73***	0.08***	233.31**	-0.00
	(224.56)	(0.02)	(79.35)	(0.02)	(115.39)	(0.02)
Autonomy (1=Yes)	317.13	0.03	-141.70*	-0.04**	138.40	0.01
	(224.25)	(0.02)	(72.81)	(0.01)	(94.86)	(0.01)
Observations	540	540	540	540	540	540
R-squared	0.11	0.11	0.27	0.14	0.15	0.08
Mean dep. var. control	2458.88	0.44	903.86	0.18	1917.28	0.37
SD dep. var. control	1788.69	0.19	713.63	0.15	1070.96	0.16
NA: $\beta_R + \beta_G = 0$	0.45	0.00	0.00	0.00	0.01	0.46
A: $\beta_R + \beta_G + \beta_A = 0$	0.48	0.00	0.00	0.00	0.00	0.19

Note: This table reported results on profits earned by farmers in the 2015 season, breaking them down by cost and revenue components. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_R Recommendation_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Commercialization survey conducted in June 2016. In Column 1, the dependent variable is the sum of all labor expenses incurred by farmers in the 2015 season. In Column 2, the dependent variable reports the labor costs as a share of the total cost paid by farmers in the 2015 season. In Columns 3 and 4, we report analogous dependent variables for capital costs. In Columns 5 and 6, the dependent variables are the input costs, such as, but not restricted to, expenses on fertilizers, herbicides, and seeds. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

F Mechanisms Autonomy

Table OA-16: WTP for fertilizers

	(1)	(2)	(3)	(4)
	Reported WTP for YARA fertilizers	WTP for a bag of YARA fertilizer in 2017 (Mex$)		
	(1=Yes)	Urea	DAP	KCl
Recommendation (1=Yes)	0.20***	70.43***	64.68**	53.28**
	(0.06)	(19.57)	(24.69)	(20.63)
Grant (1=Yes)	0.37***	111.07***	138.54***	110.00***
	(0.05)	(17.78)	(22.65)	(20.35)
Autonomy (1=Yes)	0.04	14.17	25.02	46.15**
	(0.03)	(13.92)	(18.54)	(17.64)
Observations	540	540	540	540
R-squared	0.36	0.31	0.27	0.29
Mean dep. var. control	0.33	100.38	121.73	98.46
SD dep. var. control	0.47	151.12	185.92	157.18
NA: $\beta_R + \beta_G = 0$	0.00	0.00	0.00	0.00
A: $\beta_R + \beta_G + \beta_F = 0$	0.00	0.00	0.00	0.00

Note: this table reports results willingness to pay for YARA fertilizers. Using the full sample of 540 farmers in our study, we report the point estimates of the following specification: $Y_{it} = \beta_0 + \beta_R Recommendation_i + \beta_G Grant_i + \beta_A Autonomy_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and standard deviation of the outcome for the control group, the omitted category in our regression. We also report p-values of linear combinations of the estimates coefficients that map into the original study design. We use data from the Final survey conducted in May 2017. In Column 1, the dependent variable is a dummy that takes value 1 if the farmer reported his or her willingnes to pay for any of the 3 YARA fertilizers (Urea, DAP and KCl). In Columns 2-4, the dependent variables are the self-reported willingness to pay for a bag of each of the 3 YARA fertilizers. We imput WTP equal to zero for those who did not report their WTPs. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

G Specification using group indicators

Table OA-17: 2015 Practices using the specification with group indicators

	(1)	(2)	(3)	(4)
	Existing practices		All new practices	
	Total practices applied	Standardized Index	Total practices applied	Standardized Index
A	0.07	0.17*	2.49***	1.90***
	(0.07)	(0.10)	(0.11)	(0.16)
NA	0.09	0.18*	2.57***	1.98***
	(0.07)	(0.11)	(0.11)	(0.14)
R	0.04	0.06	0.34**	0.32**
	(0.08)	(0.13)	(0.11)	(0.14)
Observations	540	540	540	540
R-squared	0.07	0.07	0.60	0.38
Mean dep. var. control	2.38	0.00	0.32	0.00
SD dep. var. control	0.61	1.00	0.69	1.00

Note: this table reports results on the agricultural practices performed by the farmers in our study in the 2015 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_1 A_i + \beta_2 NA_i + \beta_3 R_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We use data from the Follow-up survey conducted in August 2015. The dependent variable in column 1 is a sum of individual dummies. Each dummy takes value of 1 if the farmer performed one of the so-called existing agricultural practices. In column 2, the dependent variable is the standardized index of the outcome in column 1, computed by standardizing each dummy individually, adding them all and standardizing the sum. We use the mean and standard deviation of the control group as reference for the standardized index. In columns 3 and 4, the dependent variables are analogous to the outcomes in columns 1 and 2, computed for the so-called new practices. The existing practices are: (a) ploughing, (b) using inorganic fertilizer and (c) covering the fertilizer. The new practices are: (a) deep tilling (ripping), (b) using hybrid seeds, (c) fertilizing at sowing, (d) sowing with precision machinery, (e) using pre-emergent herbicide and (f) using high-quality fertilizers. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table OA-18: 2015 Fertilizer use using the specification with group indicators

	(1)	(2)	(3)	(4)	(5)	(6)
	Absolute difference, applied vs. recommended					
	At sowing			Total		
	Urea (kg/ha)	DAP (kg/ha)	KCL (kg/ha)	Urea (kg/ha)	DAP (kg/ha)	KCL (kg/ha)
A	-28.69***	-13.59***	-11.51***	-79.10***	-31.76***	-23.63***
	(1.78)	(1.81)	(1.42)	(9.21)	(4.53)	(3.49)
NA	-32.42***	-16.29***	-15.09***	-81.85***	-30.21***	-28.78***
	(1.76)	(1.76)	(1.22)	(8.97)	(4.98)	(3.08)
R	-1.83	0.52	-1.47	-4.83	1.77	-0.93
	(1.97)	(2.37)	(1.28)	(10.87)	(5.73)	(3.79)
Observations	532	532	532	532	532	532
R-squared	0.52	0.39	0.44	0.25	0.21	0.28
Mean dep. var. control	38.69	19.40	16.44	114.90	37.83	32.66
SD dep. var. control	14.64	21.36	14.42	82.62	47.92	33.98

Note: This table reports results on the usage of fertilizers by the farmers in our study in the 2015 season, compared to the recommended dosages. Using the sample for the full sample of 540 farmers, except for 13 farmers for which we do not have data on usage of fertilizers, we run the following regression: $Y_{it} = \beta_0 + \beta_1 A_i + \beta_2 NA_i + \beta_3 R_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We use data from the Commercialization survey conducted in June 2016. In columns 1-3, the dependent variables are the absolute differences between the amount applied by farmers at sowing and the recommended dosages of Urea, DAP and KCl, respectively. In columns 4-6, we report analogous outcomes for the total amount of fertilizers applied in the full season. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table OA-19: 2015 Yields and profits 2015 using the specification with group indicators

	(1) Self-reported yields (t/ha)	(2) Revenue (Mex$/ha)	(3) Costs (Mex$/ha)	(4) Profits (Mex$/ha)	(5) Profits (no subsidy) (Mex$/ha)
A	0.36**	1278.53**	1005.40***	273.13	2357.86***
	(0.15)	(523.15)	(260.94)	(503.28)	(509.91)
NA	0.29*	1001.58*	686.70**	314.88	2389.30***
	(0.15)	(517.66)	(239.87)	(522.47)	(522.66)
R	0.22	744.96	48.19	696.77	705.94
	(0.16)	(528.22)	(276.10)	(517.75)	(518.38)
Observations	540	540	540	540	540
R-squared	0.27	0.30	0.20	0.22	0.25
Mean dep. var. control	2.36	7919.22	5280.02	2639.20	2639.20
SD dep. var. control	1.33	4397.72	2351.52	4024.33	4024.33

Note: this table reports results on yields and profits earned by farmers in the 2015 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_1 E1_i + \beta_2 E2_i + \beta_3 E3 + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We use data from the Commercialization survey conducted in June 2016. In column 1, we use as the dependent variable the maize yields (tons/ha) self-reported by farmers in the 2015 season. In column 2, the dependent variable contains the value of farmers' maize production (per hectare) in the 2015 season. The value of the production (per hectare) is computed by multiplying the total amount of maize harvested by the farmer in the 2015 season by the price the maize could be sold in the market. We take the median price faced by farmers who sold at least a fraction of their production in the market as the price for all farmers when computing the value of the maize production. In column 3, the dependent variable is the total cost of production cost self-reported by each farmer. Total costs include the total investment in soil preparation activities, fertilizers (chemical and organic), herbicides, pesticides, and labor. We also include the cost of renting the sowing and harvest machines paid by QFD (when that was the case), as well as the subsidy for fertilizer packages, also paid by QFD. In column 4, the dependent variable is the difference between the dependent variable in columns 2 and 3. In column 5, the dependent variable is the cost of production, not including the subsidies paid by QFD. We use median market prices at the locality level to calculate revenues and profits. In localities where no farmer sold maize, we use median prices at the municipality level. To account for this imputation of prices (for 20 farmers in our sample), we include an dummy that takes value of one if prices were measured at the municipality level on the RHS of columns 2-5. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table OA-20: 2017 Practices using the specification with group indicators

	(1)	(2)	(3)	(4)
	Existing practices 2017		All new practices 2017	
	Total practices applied	Standardized Index	Total practices applied	Standardized Index
A	-0.05	-0.03	0.77***	1.06***
	(0.10)	(0.12)	(0.13)	(0.17)
NA	-0.10	-0.08	0.32**	0.51**
	(0.11)	(0.12)	(0.11)	(0.16)
R	0.11	0.12	0.24**	0.40**
	(0.11)	(0.12)	(0.11)	(0.16)
Observations	540	540	540	540
R-squared	0.09	0.08	0.22	0.18
Mean dep. var. control	2.31	0.00	0.42	0.00
SD dep. var. control	0.89	1.00	0.79	1.00

Note: This table reports results on the agricultural practices performed by the farmers in our study in the 2017 season. Using the full sample of 540 farmers, we run the following regression: $Y_{it} = \beta_0 + \beta_1 A_i + \beta_2 NA_i + \beta_3 R_i + \alpha_s + \epsilon_{it}$, where i corresponds to a farmer, Y is the outcome of interest and t is the time period. We include randomization strata indicators α_s and compute robust standard errors. At the bottom of the table, we report the mean and the standard deviation of the outcome for the control group, the omitted category in our regression. We use data from the Final survey conducted in May 2017. The dependent variable in column 1 is a sum of individual dummies. Each dummy takes value of 1 if the farmer performed one of the so-called existing agricultural practices. In column 2, the dependent variable is the standardized index of the outcome in column 1, computed by standardizing each dummy individually, adding them all and standardizing the sum. We use the mean and standard deviation of the control group as reference for the standardized index. In columns 3 and 4, the dependent variables are analogous to the outcomes in columns 1 and 2, computed for the so-called new practices. The existing practices are: (a) ploughing, (b) using inorganic fertilizer and (c) covering the fertilizer. The new practices are: (a) deep tilling (ripping), (b) using hybrid seeds, (c) fertilizing at sowing, (d) sowing with precision machinery, (e) using pre-emergent herbicide and (f) using high-quality fertilizers. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

H CO_2e Calculations

We convert N fertilizer used to N_2O emissions using the formula $E = 1 + .0125 * N$ from IPCC (2004); Penman et al. (2000) (see Millar et al., 2010, for a higher conversion number), convert Urea (measured in kilograms per hectare) to N using a .46 conversion factor and DAP to N using a .18 conversion factor. We then convert emissions (measured in kg/ha) into CO_2 equivalent (CO_2e) emissions using a conversion rate of 310 from UNFCC.

I Additional Results, Lab Experiment

Table OA-21: Adherence to Build Better and Productivity in the Lab, Full Sample

	(1)	(2)	(3)	(4)
Time of data collection	*Session 1*		*Session 2*	
	Adherence	Completed Figures	Adherence	Completed Figures
No Computer (Yes = 1)	-0.02	0.70*	1.27***	1.22**
	(0.16)	(0.38)	(0.17)	(0.49)
Chose Build Better (Yes = 1)	0.58*	-2.95***	0.59	-1.02
	(0.35)	(0.49)	(0.38)	(0.64)
Dropout (Yes = 1)	-0.29	0.57		
	(0.31)	(0.72)		
Constant	3.43***	8.29***	1.86***	8.55***
	(0.34)	(0.56)	(0.37)	(0.72)
Observations	200	232	187	214
R-squared	0.03	0.19	0.28	0.07
Mean dep. variable, NA	4.05	5.18	2.36	7.12
SD dep. variable, NA	1.02	2.76	1.25	3.44

This table reports results on the degree to which participants followed the Build Better approach and the number of figures they completed in the two sessions of the experiment. We report the mean of each outcome for the No Autonomy group (the omitted category). Adherence was scored on a five-point scale; completed figures corresponds to the number of ten-by-ten squares that the participant successfully replicated. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table OA-22: Balance Table Across Autonomy Groups

	Autonomy = 0	Autonomy = 1	*p-value*
Birth Year (Mean (SD))	2001.84 (5.40) [N=94]	2003.04 (3.76) [N=95]	0.077
Sex (Mean % Female)	50.0% [N=47]	49.5% [N=47]	0.997
Dropout (Mean % Yes)	5.32% [N=94]	7.37% [N=95]	0.563
School Year (counts):			0.190
First year (freshman)	24	21	
Fourth or fifth year (senior)	13	16	
Graduate student or non-student	14	5	
Second year (sophomore)	26	29	
Third year (junior)	17	24	

J Instructions, Lab Experiment

Day 1 Instructions

Participant Registration

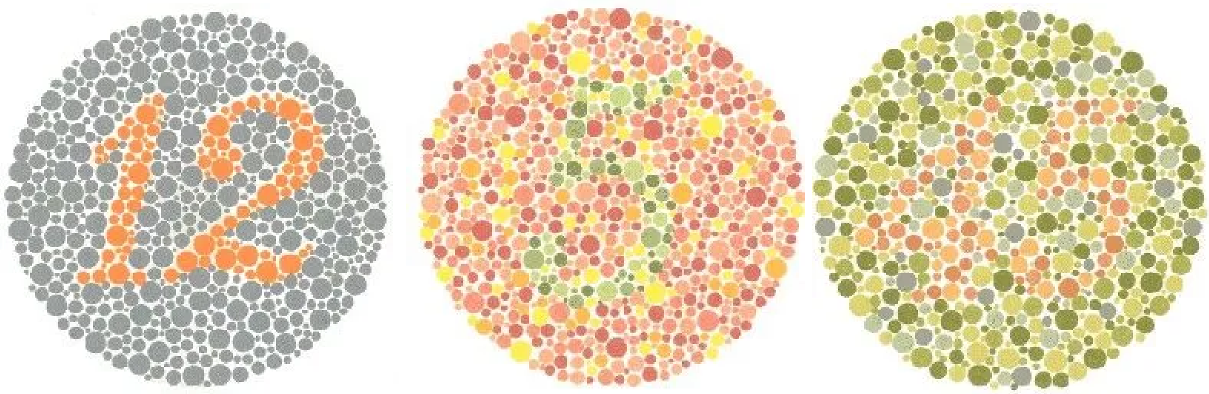
Experimenter: You must initialize the study by entering the Participant ID below. This is critical to link the answers across the two sessions.

Welcome and Preliminary Instructions

Welcome to this two-session economics experiment. We appreciate your participation. First, we want to remind you that you must complete both sessions to earn any pay for this study. You will receive payment after the second session. There will be no exceptions to this rule. All payments will be made in cash immediately after completing the second session of the study.

Before we begin, we want to ensure you can complete the tasks in this experiment. Although we use color pairings that are intended to be easy to distinguish, this may be difficult for people who suffer severe forms of color blindness. Below are three images; for each, look at the picture and enter the numbers that you see in the corresponding boxes. If you find that you cannot complete this task, please alert the experimenter.

[Participants were required to answer correctly to proceed; all participants succeeded.]



Overview of the Experiment

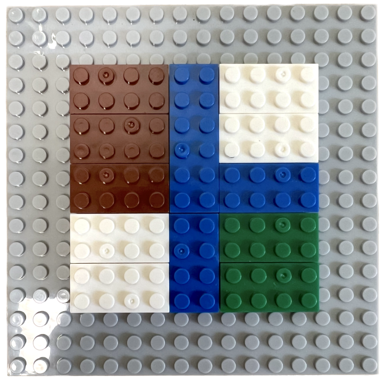
We'll now take some time to carefully explain the experiment and ensure that you fully understand the process. In this experiment, you will be working with Lego-like blocks to perform a construction task. Your main objective will be to replicate figures that we will provide. Your goal overall is to build as many copies of the figures as you can in twenty minutes. We will supply you with more than enough blocks to complete the task.

Below is an example of the type of figure that you will build:

About the Building Task

As mentioned above, you will try to build as many copies of the figure as you can in twenty minutes. There are a few simple rules:

1. You must keep the bag of blocks and the blocks themselves on the table at all times.
2. For each figure, you must replicate the arrangement of the blocks. However, you do not need to use the same shape blocks as in the provided sample. You may use any blocks (including flatter,



half-height blocks and "ramps") so long as the final figure matches the arrangement of the colors.

3. Each square figure must be separated from the next figure by at least one peg – they cannot be directly adjacent.

You may keep building until time is up or until you run out of blocks.

You can substantially increase your payment for participating in this experiment by producing lots of figures. Put simply: the more figures you successfully assemble, the more you'll be paid. So, it's in your best interest to try your best and build as many figures as you can during the experiment. We will go into details on each step of the experiment as we go through, but the overall flow of today's experiment will go as follows:

1. We will tell you about **Build Better**, a specific set of tools and techniques designed to help you build as quickly as possible.
2. You will say whether you would like to use Build Better or to freely build without the aid of the tools from Build Better.
3. A computer may take control and determine what method you will actually use. This will happen 50% of the time. If the computer takes control, you will learn which method it chose a few moments before you start the main building session.
4. You will complete a counting task.
5. (If applicable) You will learn the method the computer selected for you.
6. You will then build figures as quickly as possible.

Build Better

Let's discuss the specific building approach we have designed to assist you during the actual experiment. This has been designed to make the building process more efficient, helping you to complete more figures within the given time. We will explain each bit in detail to ensure you have a clear understanding of how they can benefit you.

1. **Bins for sorting Legos:** To help you stay organized and quickly locate the pieces you need, we provide sorting bins. If you choose Build Better, you will use these bins to separate the Lego blocks based on their color. You will sort all colors into labeled bins. By having a well-organized

workspace, you can minimize the time spent searching for specific pieces and enhance your overall productivity.

2. **Timer to manage time:** To help you keep track of the time during the task, we provide a simple kitchen timer. If you choose Build Better, you will use the timer to divide the amount of time you spend sorting versus building. Specifically, you will set the timer to exactly ten minutes, which you will spend sorting. The remaining time you will spend building. When building, you will build one figure until you cannot reasonably continue, and then move to the next. The timer can also serve as a motivational tool, encouraging you to work efficiently to beat the clock during the sorting phase.

3. **Base plate utilization techniques:** To improve your building speed, the base plates provided in the experiment can be used in specific ways. If you choose Build Better, you will join four base plates (making a 2x2 square) and you will build nine figures on this joined plate. You will also build the figures in order: you will first build copies of Figure A, then copies of Figure B, then copies of Figure C.

By using Build Better, you may be able to increase your productivity, allowing you to assemble more figures within the given time and potentially earn a higher payment. If you choose to use Build Better (or the computer forces you to use Build Better), we will give you a checklist to remind you of the process and we will monitor you to make sure you follow the correct steps. If you use your own method (or the computer forces you to use your own method), you are free to build in whatever manner you feel will be fastest. However, you will not have access to the timer or bins.

You can either choose to use Build Better or to freely build. Please consider carefully and select the alternative you believe would be most beneficial for your performance below. However—as mentioned above—there is a chance that the computer will take control and pick for you. Nevertheless, you should choose carefully.

Choice

(1) I would prefer to use Build Better.

(2) I would prefer to build on my own.

On the next screen, you will learn whether your choice counts or whether the computer has taken control.

Conditional Instructions Based on Computer Control

If computer-control = 1:

THE COMPUTER HAS TAKEN CONTROL AND WILL CHOOSE FOR YOU. You will learn which method the computer chose (Build Better or free build) in a few minutes. In the meantime, you will now do a short task designed to measure your speed at working with blocks. You will be paid an additional $5 if your performance puts you in the top 25% of participants, so try your best.

Your task is a simple sorting and counting task. We will hand you a list of blocks (where each block's shape and color matters). You must determine how many blocks in the bag match the listed blocks and write it on the list. Your time will only count if you are correct in your count, so be careful.

When you are ready, please raise your hand and the researcher will set up the computer for you. Your time will begin when you are handed the list. When you are finished, raise your hand again.

If computer-control = 0:
THE COMPUTER HAS NOT TAKEN CONTROL. Your previous choice (between Build Better or free build) will be what you use to build in a few minutes. In the meantime, you will now do a short task designed to measure your speed at working with blocks. You will be paid an additional $5 if your performance puts you in the top 25% of participants, so try your best.

Your task is a simple sorting and counting task. We will hand you a list of blocks (where each block's shape and color matters). You must determine how many blocks in the bag match the listed blocks and write it on the list. Your time will only count if you are correct in your count, so be careful.

When you are ready, please raise your hand or alert the experimenter and they will set up the computer for you. Your time will begin when you are handed the list. When you are finished, raise your hand again.

Counting Task Hold Screen

Alert the experimenter if you have not done so already. Your time will stop after you tell the experimenter how many of each block there are. You must answer correctly to continue. When you have finished, the experimenter will type "continue" (without quotation marks) below.

If computer-control = 1:
Remember, the computer took control and will choose for you. You do not get to choose how to build. In a few seconds we will reveal what the computer chose for you. You will be able to proceed to the next page in 15 seconds.

If computer-control = 0:
Remember, you got to choose your building method. In a few seconds we will remind you of the rules for building. You will be able to proceed to the next page in 15 seconds.

Building Session Instructions

For Free Build:
We are now entering the building portion of the experiment. **You chose Free Build.** This means that you can use any method you like to build, but you will not have access to the tools from Build Better. As a reminder of the rules:

1. You must keep the bag of blocks and the blocks themselves on the table at all times.
2. For each figure, you must replicate the arrangement of the blocks. However, you do not need to use the same shape blocks as in the provided sample. You may use any blocks so long as the final figure matches the arrangement of the colors.
3. Each figure must be separated from the next figure by at least one peg – they cannot be directly adjacent.

For Build Better:

If computer-control = 1:

We will now reveal which building method the computer chose. The computer chose Build Better. **If computer-control = 0:**

You chose to use Build Better.

Build Better means you will use the following techniques:

- Bins for sorting Legos: To help you stay organized and quickly locate the pieces you need, we provide labeled sorting bins. You will sort all colors – even those not utilized in the figure you will build.
- Timer to manage time: To help you keep track of the time during the task, we provide a simple kitchen timer. You will set the timer to exactly ten minutes, which you will spend sorting. Once it goes off, you will switch to assembling for the remaining time. If you do not finish sorting within ten minutes, you will push aside all unsorted pieces. If you finish early, you will immediately begin building.
- Base plate utilization techniques: To improve your building speed, you will join the base plates together. After the timer goes off, you will immediately grab four base plates. You will then build figures on four combined base plates (2x2). Separate figures from one another by one row of pegs. You will build nine figures on this joined plate. You will also build the figures in order: you will first build three copies of Figure A, then three copies of Figure B, then three copies of Figure C. If you complete this, you will build remaining figures one at a time. See the image below for an example of how to use the baseplates.

As a reminder of the rules:

1. You must keep the bag of blocks and the blocks themselves on the table at all times.
2. For each figure, you must replicate the arrangement of the blocks. However, you do not need to use the same shape blocks as in the provided sample. You may use any blocks so long as the final figure matches the arrangement of the colors.
3. Each figure must be separated from the next figure by at least one peg – they cannot be directly adjacent.
4. Finally, you must follow the Build Better approach. You will be provided a checklist to remind you.

(Presented to all groups)

As previously mentioned, there are three figures that you can build. These will be placed in a box on the table in a moment. When your time begins, you can remove the box and place the figures wherever you like. You will be paid $0.50 for each completed figure. You will be paid for completed figures that are correct. If you end with a partially completed figure, your payment will be rounded to the nearest $0.25, so keep going until time is up. The computer screen will provide a countdown for the last 30 seconds of the building time and an alarm will sound at the very end. When you are ready to begin, please alert the experimenter to give you the remaining materials and set up the computer.

Survey Questions

Thank you. To conclude this survey, we will ask some simple questions. We will begin with some demographic questions below.

1. What is your year of birth? ________________________
2. What year in school are you?
   (1) First year (freshman)
   (2) Second year (sophomore)
   (3) Third year (junior)
   (4) Fourth or fifth year (senior)
   (5) Graduate student or non-student
3. Which of the following best describes your sex assigned at birth?
   (1) Male
   (2) Female
   (3) Prefer not to say

Next, we would like to ask you a few simple questions about your experience building. Please indicate how strongly you agree or disagree with each statement by choosing the appropriate response on the scale provided. The scale ranges from Strongly Disagree to Strongly Agree. There are no right or wrong answers, just answer honestly.

1. I enjoyed the building method that I just used.
2. I would have been faster using another building method. [*Reverse coded*]
3. I was among the top 25% of builders (in terms of total figures built).

Willingness to Complete Additional Tasks for Additional Pay

Finally, we'd like about your willingness to construct additional figures for additional payment. If you are willing to do additional figures for additional payment, you will do them at the end of Session 2, so make sure you plan accordingly. The method we use to determine whether you will complete extra tasks may seem complicated. But, we'll walk through it step-by-step. The punchline will be that it's in your best interest to just answer truthfully. Here's how the system works. First, we will ask you how many additional figures you are willing to build for a fixed amount of money. For instance, we might ask: "What is the maximum number of extra figures you are willing to build for $1.00?"

On the decision screen, you will be presented a slider that goes between 0 and 20. You will move the slider to indicate the maximal number of figures you'd be willing to do for that amount of money. The amount of money is the total payment, not the amount per figure. Suppose you would be willing to do 15 additional figures and get paid $1.00 for that effort, but you would not be willing to complete 16 for $1.00. Then you should move the slider to 15.

These would be the same figures as you previously completed. In this scenario, you would have to utilize the same building technique as you did a few moments ago. You would also begin (as before) with unsorted blocks. We would give you extra blocks as needed to ensure you could build as many figures as you'd like.

We will then draw a random number between 0 and 20. If your answer is less than that random number, you will not do additional figures. However, if your answer is greater than or equal to that random number, you will do a number of additional figures equal to the random number.

Example: Suppose you indicated you were willing to do 15 additional figures if you were paid $1.00 (total). If the random number was 16 or higher, you would do no additional figures. However, if the random number was 12, you would do 12 additional figures. The next pages have a short quiz to help clarify how this works.

1. Suppose you were asked "What is the maximum number of additional figures you are willing to do for $1.00?" and you responded 18. If the random number is 17, how many figures will you complete?
   (1) 0 and I will be paid $0 in supplementary payments.
   (2) 18 and I will be paid $1.00 in supplementary payments.
   (3) 17 and I will be paid $1.00 in supplementary payments.
   (4) 17 and I will be paid $2.67 in supplementary payments.
2. Suppose you were asked "What is the maximum number of additional figures you are willing to do for $1.00?" and you responded 18. If the random number is 19, how many additional tasks will you complete?
   (1) 0 and I will be paid $0 in supplementary payments.
   (2) 18 and I will be paid $1.00 in supplementary payments.
   (3) 19 and I will be paid $1.00 in supplementary payments.
   (4) 19 and I will be paid $2.67 in supplementary payments.

This method of selecting how many additional tasks you will do might seem complicated, but as we previously highlighted, there's a great feature to it: your best strategy is to simple answer honestly.

As a reminder: If you are willing to do additional figures for additional payment, you will do them at the end of Session 2, so make sure you plan accordingly. These would be the same figures as you previously completed. In this scenario, you would have to utilize the same building technique as you did a few moments ago. You would also begin (as before) with unsorted blocks. As needed, we will give you extra blocks to ensure you could build as many figures as you'd like.

On the next screen you will see the real question.

Final Additional Question:
What is the maximal number of additional figures you'd be willing to complete for $5.00?
(Slider)

End of Day 1

Thank you for participating. This is the last screen before the end.

Please click the final button below to submit your work. Remember, you must return for the second session for payment.

Day 2 Instructions

Participant Registration

Experimenter: Please initialize the study by entering your Participant ID below to link your answers with Day 1.

Session Introduction

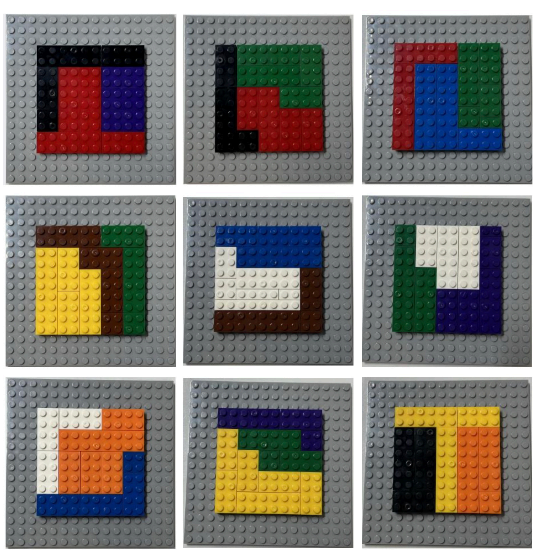
Thank you. To begin this session, we will ask you some simple questions about your experience last time. For each question, you will be asked specific details about the building process. If your answers are correct (as scored by an AI language model) your name will be credited with one token. After the experiment is completed, we will draw 50 tokens and give winners an additional $20. So please answer these questions carefully and thoughtfully.

Recall of Build Better Steps

Using the space below, describe (in your own words) the steps to **Build Better**. For each element that you correctly describe, you will be awarded one token.

Image Selection Task

Below are nine images. Please select the images that correspond to the figures that you were supposed to build in the first session. For each correct response you will receive one token. For each incorrect response you will lose one token. [Scrambled; correct answers shown in left column.]



Reminder of the Building Task

We'll now take some time to remind you of the experiment and ensure that you fully understand the process. As before, you will be working with Lego-like blocks to perform a construction task. Your main objective is to replicate figures that we will provide. Your goal overall is to build as many copies of the figures as you can in twenty minutes. We will supply you with more than enough blocks to complete the task.

You will have three figures to build. As in Session 1, you will try to build as many copies of the figure as you can in twenty minutes. There are a few simple rules:

1. You must keep the bag of blocks and the blocks themselves on the table at all times.
2. For each figure, you must replicate the arrangement of the blocks. However, you do not need to use the same shape blocks as in the provided sample. You may use any blocks (including flatter, half-height blocks and "ramps") so long as the final figure matches the arrangement of the colors.
3. Each square figure must be separated from the next figure by at least one peg – they cannot be directly adjacent.

You may keep building until time is up or until you run out of blocks.

Build Better Checklist Option

We will provide the tools for Build Better. But you are free to build however you like. If you would like the checklist for Build Better, please select the box below. If you do not want the checklist, click the arrow to continue.

Building and Payment Instructions

As in the first session, there are three figures that you can build. These will be placed on the table in a moment. You will be paid $0.50 for each completed figure. You will be paid for completed figures that are correct. If you end with a partially completed figure, your payment will be rounded to the nearest $0.25, so keep going until time is up. The computer screen will provide a countdown for the last 30 seconds of the building time and an alarm will sound at the very end.

Conditional Display for Checklist

Please alert the experimenter. Do not continue with this survey until you have finished building. Experimenter: please include the Build Better checklist on the table with the other materials. When you have finished, type "continue" (without quotation marks) below.

If you did not select the checklist option:
Alert the experimenter once you have finished building, then type “continue” to proceed.

Post-Build Survey

As in the first session, we would like to ask you a few simple questions about your experience building. Please indicate how strongly you agree or disagree with each statement by choosing the appropriate response on the scale provided. The scale ranges from Strongly Disagree to Strongly Agree. There are no right or wrong answers, just answer honestly

1. I enjoyed the building method that I just used.
2. I would have been faster using another building method.
3. I was among the top 25% of builders.

Additional Figures Decision

Based on your responses from Session 1, you will be informed whether you must complete additional figures. [BDM was resolved]

Final Submission

Thank you for participating. This is the last screen before the end.

Please click the final button below to submit your work.