

Improving Input Measurement in Agriculture: Evidence from a mixed-mode Survey Experiment in Nigeria¹

Ivette Contreras²

Pauline Castaing³

Amparo Palacios-Lopez⁴

Steffanny Romero⁵

October 2025

Abstract

Mixed-mode data collection strategies, combining phone and in-person interviews, are increasingly used in longitudinal surveys in low- and middle-income countries to reduce costs while maintaining data quality. However, there is limited evidence of their effectiveness in improving measurement accuracy. This paper uses experimental data from Nigeria to assess whether mixed-mode surveys can reduce recall bias in reporting agricultural labor and input use. We compare estimates from a mixed-mode approach that builds on high-frequency phone interviews between two in-person visits with those from a standard end-of-season recall survey. The findings reveal substantial recall bias in the standard method: plots in the mixed-mode group were 92.6% more likely to report individual labor participation, and the number of household members reported as working increased by 101.5%. Conversely, hours and days worked per person were overreported in the recall survey, with reductions of 29–32% observed under the mixed-mode approach. Non-labor inputs such as fertilizers, herbicides, and pesticides were underreported by up to 198% in the recall survey. Differences were especially pronounced on larger plots and among male-headed households. These results underscore the potential of mixed-mode surveys to dramatically improve the accuracy of agricultural input data and inform more effective policy design.

JEL classification: C8, C9, J43, O1

Keywords: Survey methods, mode effects, recall bias, agricultural input measurement

¹The findings, interpretations, and conclusions expressed in this report are entirely the authors and do not necessarily represent the views of the World Bank, its affiliated organizations, Executive Directors, or the governments that they represent. This paper was produced with financial support from the 50x2030 Initiative to Close the Agricultural Data Gap, a multi-partner program that seeks to bridge the global agricultural data gap by transforming data systems in 50 countries in Africa, Asia, the Middle East and Latin America by 2030. For more information on the Initiative, visit 50x2030.org.

² Corresponding Author. The World Bank. Email: icontreras@worldbank.org.

³ The World Bank. Email: pcastaing@worldbank.org

⁴ The World Bank. Email: apalacioslopez@worldbank.org.

⁵ The World Bank. Email: sromeroesteban@worldbank.org.

1. Introduction

The importance of boosting agricultural productivity and incomes of small food producers to alleviating poverty and improving food security makes these issues a policy focus for many developing countries. Indeed, research suggests that growth in the agricultural sector translates into higher rates of poverty reduction than in other sectors (Dorosh and Thurlow, 2018). High-quality, input data at scale are indispensable for effective policy design and analysis. Reliable measurement of labor and non-labor inputs has been of research and policy interest (see e.g. Abay et al., 2021; Arthi et al., 2018; Desiere and Jolliffe, 2018; Gollin et al., 2014; Gollin and Udry, 2021; Kilic et al., 2021).

The most common approach to collect survey data on farm labor, and non-labor inputs is through respondent recall at the end of the agricultural season (end-of-season recall). With this approach, respondents are asked to estimate the amount of time worked, the quantity of inputs applied, or the amount harvested, in the previous agricultural season. Smallholder farming activities, such as planting, plot maintenance, or harvest, are often highly irregular and take place at different times during the agricultural year. Recalling them forces respondents into cognitively taxing calculations (Arthi et al., 2018) and methodological research has shown that end-of-season recall in farm labor, inputs, and harvest is rife with non-random measurement error (Wollburg et al., 2021). For example, studies in Tanzania and Ghana indicate that the recall method overestimates farm labor per person per plot (Arthi et al., 2018; Gaddis et al., 2021); and evidence from Ethiopia, Uganda, and Malawi suggests farmer-reported harvest is subject to non-classical measurement error (Desiere and Jolliffe, 2018; Gourlay et al., 2017; Kilic et al., 2021). This may distort policies targeted at improving productivity and alleviating poverty.

High-frequency visits to households (e.g. weekly, bi-weekly) present a more reliable alternative to end-of-season recall, as they significantly reduce recall period and respondent burden, and are often used for benchmark estimates (e.g. Arthi et al., 2018; Beegle et al., 2020; Gaddis et al., 2021; Kilic et al., 2021). In practice, this data collection method is too expensive to be implemented at scale as part of nationally representative household surveys. However, several recent studies have shown that high-frequency mobile phone surveys hold promise to produce similarly reliable data, while being considerably less expensive than high-frequency in-person visits. Arthi et al. (2018) note, in the context of farm labor data, the strong performance of phone surveys, which show little difference in labor estimates relative to the in-person weekly visit survey. Kilic et al. (2021) even consider harvest estimates from twice-weekly phone surveys superior to estimates from twice-weekly in-person visits in the context of extended harvest crops in Malawi.

The rapid accumulation of experience with phone surveys by national statistical offices and survey practitioners in low- and middle-income countries has laid a strong foundation for the future implementation of phone surveys (Gourley et al., 2021). Phone interviews are increasingly recognized as a valuable complement to face-

to-face surveys, with phone survey estimates proven remarkably robust to several commonly claimed biases (Markhof et al., 2023). However, to fully integrate phone surveys into nationally representative surveys, it is essential to conduct further research and develop comprehensive guidelines. Numerous questions remain under-researched on phone survey design and implementation as part of longitudinal household surveys.

This paper investigates the extent and nature of recall bias in the measurement of agricultural inputs by leveraging a randomized survey experiment embedded in Nigeria's General Household Survey Panel. By comparing data collected through traditional end-of-season recall interviews with data gathered via high-frequency phone calls combined with an in-person interview, the study evaluates how a mixed-mode data collection strategy affect the reporting of both labor and non-labor inputs. The analysis spans multiple levels of aggregation—individual-plot, plot, individual, and household—allowing for a nuanced understanding of how recall bias manifests across different units of analysis. The paper not only quantifies the magnitude of misreporting but also identifies which types of inputs and respondent characteristics are most affected, offering practical insights for improving the design of agricultural surveys in low- and middle-income countries.

We find that, at the extensive margin, the mixed-mode data collection strategy substantially increases the likelihood of reporting labor engagement in agricultural plots—particularly among women and children—as well as the use of non-labor inputs such as fertilizers, herbicides, pesticides, and animal traction. At the intensive margin, it reveals that the traditional approach tends to overstate labor input quantities, with mixed-mode respondents reporting significantly fewer hours and days worked per individual. In contrast, the quantities of non-labor inputs reported are markedly higher under the mixed-mode approach, suggesting that recall bias leads to underestimation of input intensity. These effects are most pronounced at the plot level and diminish with higher levels of aggregation, underscoring the importance of disaggregated data collection for capturing accurate input use patterns.

This study contributes to the literature on agricultural data collection by providing experimental evidence on the effects of mixed-mode data collection strategies on the measurement of both labor and non-labor inputs in a lower-middle income country context. While prior research has documented recall bias in labor reporting, particularly for irregular and unpaid work, this study extends the analysis to include non-labor inputs such as fertilizers, pesticides, and animal traction. By embedding a randomized survey experiment within a nationally representative panel survey in Nigeria, the study isolates the effects of integrating high-frequency phone interviews into standard face-to-face data collection. The findings demonstrate that mixed-mode data collection strategies influence not only the likelihood of reporting input use but also the reported quantities, with implications for productivity analysis and policy design. Moreover, the study provides new insights into how these effects vary by plot size, education level of the plot manager, and household head characteristics, offering a more nuanced understanding of the mechanisms underlying measurement error in agricultural surveys.

The rest of the paper is organized as follows. Section 2 discusses the different approaches to measuring agricultural labor and non-labor inputs. Section 3 describes the experimental design, the data, and the main outcomes of interest. Section 4 presents the methodology. Section 5 outlines the results, and section 6 concludes.

2. Measuring Agricultural Inputs

Understanding the agricultural sector is essential for analyzing global production, employment, poverty, hunger, and food security, particularly in developing countries (Timmer, 1995; Datt and Ravallion, 1998; Minten and Barrett, 2008). In 2023, agriculture employed approximately one-quarter of the world's labor force (International Labor Organization, ILOSTAT). The sector's importance is even more pronounced in low-income and lower-middle-income countries, where 58% and 39% of the workforce, respectively, were engaged in agriculture. Simultaneously, agriculture contributed 4.1% of the global gross domestic product (World Bank, 2025), and 24.8% in low-income countries and 15.4% in lower-middle-income countries. Notably, around 80% of the world's poor live in rural areas and depend on agriculture for their livelihoods. Moreover, growth in the agricultural sector has been shown to be 3.2 times more effective at reducing extreme poverty than equivalent growth in other sectors (Christiaensen, Demery, and Kuhl, 2011).

As a result, substantial investments and multiple interventions have targeted the agricultural sector. For instance, the World Bank (2025) alone financed \$22.3 billion for food and agriculture between 2019 and 2024, including operations promoting the adoption of technology to efficiently using inputs in Uruguay and Angola (Jeronimo, 2025; Freeman, 2021). Furthermore, the Consultative Group on International Agricultural Research (CGIAR) allocated approximately \$60 billion in investments over five decades to agricultural research and collaboration with national agricultural research systems, international research centers, and development partners in low- and middle-income countries. These investments focused on developing crop varieties, improving productivity, and enhancing food security through science-driven innovation (Anderson and Alston, 2020). CGIAR estimates that such investments yield a benefit-cost ratio of up to 10-to-1.

Consequently, both the academic literature and policy debates have devoted considerable attention to agricultural production, with a particular emphasis on the evolution of total factor productivity (TFP) and its underlying determinants. Much of this focus stems from the recognition that growth in agricultural output over the past several decades has relied not only on the expansion of cultivated area or increased use of conventional inputs but increasingly on improvements in the efficiency with which these inputs are transformed into output. Huang et al. (2002) documents how the Green Revolution ushered in a period of rapid yield gains, especially in Asia, driven by the widespread adoption of modern crop varieties and complementary investments in irrigation and fertilizers. However, they caution that many of the gains from input intensification have now been

exhausted, particularly in high-performing regions, and that future productivity improvements will require new technological breakthroughs, including those emerging from biotechnology and advanced breeding methods.

Accurate measurement of these inputs, such as labor, fertilizers, pesticides, and seeds, is critical for assessing productivity measures. For example, Desiere, Jolliffe, and Blanc (2018) show that land size mismeasurement in household surveys distorts yield estimates and can produce spurious conclusions about productivity. Abay, Bevis, and Barrett (2021) and Abay et al. (2023) emphasize that non-classical measurement errors, particularly in labor and land inputs, can significantly bias TFP estimations and policy inferences. Similarly, Michelson, Gourlay, and Wollburg (2022) review evidence on fertilizer and pesticide quality in Sub-Saharan markets and emphasize that both quantity and quality of these inputs must be measured to assess economic returns and productivity correctly. Thus, accurate quantification of agricultural inputs stands crucial for understanding productivity dynamics, evaluating the impacts of interventions, and informing input policies in developing countries. However, household and farm survey instruments continue to rely on recall-based data collection methods that are being found to introduce systematic biases.

Multi-topic household surveys that also collect agricultural survey modules often have a low frequency of data collection with long recall periods, resulting in substantial measurement errors in labor reporting Arthi et al. (2018) and Gaddis et al. (2021) both investigate recall bias in measuring household smallholder farm labor through randomized survey experiments, but in different contexts: Tanzania and Ghana, respectively. Arthi et al. (2018) finds substantial recall bias in Tanzania: end-of-season surveys overreport labor hours per person-plot by up to a factor of 3.7 compared to weekly surveys, while simultaneously underreporting the number of plots and workers. This dual bias, at both the intensive and extensive margins, leads to a significant underestimation of agricultural labor productivity and is attributed to the cognitive difficulty of recalling irregular and non-salient farm work. In contrast, Gaddis et al. (2021) finds a much smaller recall bias in Ghana—about 10% overreporting—driven almost entirely by “listing bias,” where marginal plots and workers are omitted from recall surveys, inflating average labor estimates. Once listing differences are accounted for, the remaining bias in hours nearly vanishes. They also find that better-educated respondents show no significant recall bias, and that gender disparities in listing are present but somewhat mitigated by differences in hours worked.

Building on this literature, this paper contributes to the understanding of measurement error by examining recall biases using a mixed-mode data collection strategy not only in labor inputs, which have received the most attention to date, but also in non-labor inputs, such as animal traction, fertilizers, and pesticides. By incorporating both types of inputs, this study broadens the scope of existing analyses and highlights underexplored dimensions of recall bias in agricultural data.

3. Experimental Design, and Main Outcomes

3.1 Experimental Design

We conducted a randomized survey experiment during the fieldwork of the Nigeria's General Household Survey Panel Wave 5 (GHS-Panel W5). The GHS-Panel is a nationally representative longitudinal survey implemented by Nigeria's National Bureau of Statistics (NBS) in collaboration with the World Bank and the Bill and Melinda Gates Foundation. It collects rich data on household characteristics and agricultural practices, including labor inputs, inputs use, crop types, and plot features. The GHS-Panel W5 targeted 5,067 households in 518 Enumeration Areas (EAs) and was administered in two visits: the post-planting period (July–September 2023) and the post-harvest period (January–March 2024).

Using the GHS-Panel Wave 4, we identified the Enumeration Areas (EAs) with a minimum of 60% of households engaged in agricultural activities. Among the 401 EAs identified, we randomly selected 106 as treated EAs for the experiment consisting of monthly phone calls between the two in-person visits of the GHS-Panel W5. All households within these selected enumeration areas were eligible to participate in the experiment, resulting in a total of 995 eligible households. The 2,295 households located in the remaining 295 enumeration areas were not invited to participate to phone interviews and were solely part of the GHS-Panel Survey Wave 5. These households serve as the control group⁶.

During the post-planting visit, enumerators were able to locate and interview 922 households eligible to the experiment. Of these, 911 agreed to participate in the experiment and received a phone and a SIM card. Enumerators were instructed to deliver the phone to the household member most knowledgeable about agricultural activities⁷, (See Figure A1 in Appendix A for a picture of the phone provided). The mobile phones and SIM cards were provided at the end of the interview to not affect any of the previous responses in the in-person post-planting visit. Follow-up phone interviews were scheduled approximately every three to four weeks starting after the post-planting visit. Among participating households, 86.7% (N = 790) were successfully reached in both in-person visits and all four rounds of phone interviews. For our analysis, we restrict the mixed-mode group to a balanced panel of these 790 households and the control group to a balanced panel of 2,279 households successfully interviewed in both in-person visits. Figure 1 summarizes our experimental design.

All households (regardless of their treatment status) were asked to list their members and plots at baseline and to provide key demographic and plot-specific information. In the four subsequent phone interviews, households in the experiment were asked about their agricultural activities. To reduce respondent fatigue, these

⁶ The experiment was conducted in Nigeria, a country with a high phone coverage rate and widespread recent experience with high frequency phone surveys (Gourlay et al., 2021). Using the data of the GHS-Panel Wave 5 Post-Planting Visit (data collected before our experiment), 93.3% of the households have at least one mobile phone.

⁷ If the household had not engaged in agricultural activities during the most recent season, the preference was to select the member most informed about the work activities of household members.

interviews focused on up to three plots reported at baseline.⁸ For households with more than three plots, a random subset of three plots was selected during the baseline and remained fixed throughout the experiment. 18.6% of the households in the mixed-mode group had more than three plots. This resulted in 440 plots being classified as control plots, representing 5.4% of all plots included in the analysis.

Table 1 presents household, individual, and plot-level characteristics for the final sample in both the mixed-mode and control groups, based on data from the baseline face-to-face survey. Overall, the two groups appear well balanced, though a few statistically significant differences emerge. The dependency ratio is slightly lower in the mixed-mode group (0.416) than in the control group (0.443), and mixed-mode households report a higher average number of agricultural plots (2.51 vs. 2.36). Household members in the mixed-mode group are also marginally older (25.1 vs. 24.4 years). At the plot level, while maize cultivation rates are similar, mixed-mode plots are more likely to be planted with yam (10.8% vs. 8.6%), cassava (27.9% vs. 22.2%), and rice (10.7% vs. 6.8%). Average plot size is smaller in the mixed-mode group (5,263 m² vs. 6,472 m²). To account for these baseline differences, the empirical analysis includes some control variables, as discussed in the next section. Baseline labor inputs, measured in both hours and days worked, do not differ significantly between groups.

4. Methodology

4.1 Baseline estimates

For all the outcomes identified in the previous section (both in the intensive margin analysis as well as in the extensive margin analysis), we conduct the following regression analysis using a double-lasso procedure for identifying the covariates that are relevant to a specific outcome. We conduct the analysis in four different levels of data aggregation: (i) individual-plot level, (ii) plot level, (iii) individual level, and (iv) household level. We start the analysis with the most disaggregated data at the individual-plot level, and we estimate Equation 1.

$$Y_{iphez} = \beta_0 + \beta_1 mm_plot_{phez} + \beta_2 mm_hh_{hez} + \beta_3 X_{iphez} + \delta_z + \varepsilon_{iphez} \quad (1)$$

Where Y_{iphez} is the outcome variable for individual i in a plot p which is own or used by household b that is located at enumeration area e in the geographic zone z , mm_plot is a dummy variable that takes the value of 1 if the plot p in household b was included in the mixed-mode treatment group and 0 if the plot belongs to the control group.⁹ In addition, mm_hb is a dummy variable that takes the value of 1 if the household b was included

⁸ As Table 1 shows, the difference in the distributions of households with more than 3 plots in both control and mixed-mode groups is not statistically significant.

⁹ As discussed above, to reduce respondent fatigue, the experiment focused on up to three plots reported at baseline. For households with more than three plots, a random subset of three plots was selected during the baseline and remained fixed throughout the experiment. The other plots were not part of the mixed-mode group and were treated as control plots.

in the mixed-mode experiment and 0 if the household belongs to the control group. X is a vector of individual, plot, household, and plot-owner characteristics, δ_z are zone fixed effects while ε_{iphez} is an error term. Standard errors are clustered at the enumeration-area level.

The second equation explores the data at the plot level. Where Y_{phez} is the outcome variable for plot p which is own or used by household b that is located at enumeration area e , mm_plot and mm_hb follow the same definition as in equation 1. X is a vector of plot, household and plot-owner characteristics, δ_z are zone fixed effects while ε_{phez} is an error term. Standard errors are clustered at the enumeration-area level.

$$Y_{phez} = \beta_0 + \beta_1 mm_plot_{phez} + \beta_2 mm_hh_{hez} + \beta_3 X_{phez} + \delta_z + \varepsilon_{phez} \quad (2)$$

The third specification analyzes the data at the individual level. Where Y_{ihez} is the outcome variable for individual i in household b located at enumeration area e . In addition, mm_hb follows the same definition as in equation 1. X is a vector of individual and household characteristics, δ_z are zone fixed effects, while ε_{ihez} is an error term. Standard errors are clustered at the enumeration-area level.

$$Y_{ihez} = \beta_0 + \beta_1 mm_hh_{hez} + \beta_2 X_{ihez} + \delta_z + \varepsilon_{ihez} \quad (3)$$

Finally, the fourth specification explores the data at the household level. Where Y_{hez} is the outcome variable for household b located at enumeration area e . The variable mm_hb follows the same definition as in equation 1. X is a vector of household characteristics, δ_z are zone fixed effects, while ε_{hez} is an error term. Standard errors are clustered at the enumeration-area level.

$$Y_{hez} = \beta_0 + \beta_1 mm_hh_{hez} + \beta_2 X_{hez} + \delta_z + \varepsilon_{hez} \quad (4)$$

4.2 Heterogeneity analysis

We expand our analysis by estimating heterogeneous treatment effects by plot size, plot manager's educational level, and sex of household head. For the first two variables, we extend Equation (2) at the plot level by interacting each treatment indicator with the heterogeneity variable denoted by D_{phez} , the new specification is presented in Equation (5). Separate models are estimated for plot size, and plot manager's educational level, as follows:

$$Y_{phez} = \beta_0 + \beta_1 mm_plot_{phez} + \beta_2 mm_hh_{hez} + \beta_3 mm_plot_{phez} \times D_{phez} + \beta_4 mm_hh_{hez} \times D_{phez} + \beta_5 X_{phez} + \delta_z + \varepsilon_{phez} \quad (5)$$

At the household level, we extend Equation (4) by interacting the treatment indicator with the heterogeneity variable denoted by D_{hez} , the new specification is presented in Equation (6). Separate models are estimated for sex of household head, as follows:

$$Y_{hez} = \beta_0 + \beta_1 mm_hh_{hez} + \beta_2 mm_hh_{hez} \times D_{hez} + \beta_3 X_{hez} + \delta_z + \varepsilon_{hez} \quad (6)$$

5. Results

5.1. *Impacts of the mixed-mode data collection strategy on measuring agricultural inputs*

Extensive Margin Effects. We find that the mixed-mode treatment significantly increased the likelihood of reporting individuals as having worked on plots and increased the number of household members reported as working. Table 2 (at the individual-plot level) shows that plots included in the mixed-mode treatment group increased the likelihood of reporting that an individual worked on a plot by 40.2 percentage points (92.6% increase over the control mean). As discussed above, 5.4% of the total plots from the mixed-mode households were randomly excluded from the high-frequency phone calls to minimize respondents' fatigue. When the household (but not the specific plot) was included in the mixed-mode treatment, the likelihood of reporting an individual as having worked on a plot decreased by 13.4 percentage points. In other words, non-selected plots of the mixed-mode households reported a lower probability of having an individual that had worked on the plot. The negative coefficient for households included in the mixed-mode study may suggest that the effect is more pronounced when the treatment is applied at the plot level rather than the household level. This could reflect differences in how respondents recall and report labor when prompted about specific plots versus general household activity.

At the plot-level, Table 3a shows that the mixed-mode treatment significantly increased the number of household members reported as working on a plot by 2.509, a 101.5% increase over the control mean of 2.472. As in the previous table, we observe a negative coefficient for the variable that controls if the household belongs to the mixed-mode group. Similar results are observed for the number of women, and children household members that are reported to work in the plot with 95.3% and 199.8% increase in the over the outcome mean due to the inclusion of the plot in the mixed-mode study, respectively. However, for male household members working in the plot the mixed-mode treatment at the household level is not statistically significant, this evidence may show that when reporting male household members workers, the differences of the treatment level does not make significant changes. Furthermore, the mixed-mode treatment at the plot level shows significant increases in the probability of reporting to hire workers or have unpaid workers in the plot (women, men or children). The percentage change in reported labor is more pronounced for women and children than for male

across all categories of workers (family, hired, or unpaid). No significant effects for mixed-mode treatment at the household level were reported for hiring workers or unpaid workers.

Table 3b focuses on non-labor inputs reported usage at the plot level. The mixed-mode treatment significantly increases the reporting of all non-labor inputs captured on the table. Specifically, plots included in the mixed-mode group were 34 percentage points more likely to report using inorganic fertilizer, 25.4 points more likely for organic fertilizer, 34.6 points for herbicides, 30.1 points for pesticides, and 7.6 points for animal traction. These effects translate into substantial relative increases compared to the control group means—ranging from 38.0% for animal traction to 183.8% for pesticide use. The table also shows that these effects are driven by the inclusion of the plots in the mixed-mode study rather than the inclusion of the entire household into the study.

At the individual-level, table 4 shows that the mixed-mode treatment increased the likelihood that an individual was reported to have worked on any plot by 22.3 percentage points (56.7% increase). The effect is smaller compared to the one in Table 2, likely because it captures a broader and less detailed measure of labor participation.

Table 5a focuses on labor inputs measured at the household level. We observe that the inclusion in the mixed-mode study increases significantly all the labor inputs at the extensive margin. The increases are particularly pronounced for women and children, especially among the household members and hired workers, suggesting that traditional recall-based surveys may systematically under-report these groups. These household-level estimates are still substantial and statistically significant, but consistently smaller than those at the plot level (reported in Table 3a). This may reflect aggregation effects or the dilution of detail when labor is reported at the household rather than plot level.

Table 5b shows that the mixed-mode survey treatment significantly increased the likelihood that households reported using key agricultural inputs. The largest relative increase was observed for pesticide use (175.9%), followed by organic fertilizer (110.2%), herbicide (76.9%), inorganic fertilizer (65.3%), and animal traction (52.1%). The effects are generally smaller at the household level for inorganic fertilizer, herbicide, and pesticide use compared to the effects reported at the plot level (table 3b), while organic fertilizer (110.2% increase at the household level compared to 107.6% increase at the plot level) and animal traction (52.1% vs 38%) show slightly higher or comparable effects at the household level.

Intensive Margin Effects. Table 6 shows that the mixed-mode treatment generally led to lower reported hours and days worked per individual in each plot. The person/plot hours worked in the treated plots shows a decrease of 204.9 hours in the entire agricultural season (29.3% decrease). Similarly, the person/days worked in the treated plots decreased by 46.8 days (31.9% decrease). The effect is significant only when the plot is included in the mixed-mode treatment—and not when only the household is included.

Table 7 presents how the mixed-mode treatment affects the quantities of labor and non-labor inputs reported. The effects on labor inputs—measured as hours and days worked by household members on the plot—are positive but not statistically significant. In the case of non-labor inputs, the mixed-mode treatment leads to significant increases in the reported quantities. For instance, the quantity of inorganic fertilizer used increased by 117.4%, and organic fertilizer use rose by 149.6%. Even more striking are the increases in herbicide and pesticide use, which rose by 198.4% and 116.9%, respectively. The number of days with animal traction also increased by 161.8%.

Table 8 presents the intensive margin effects at the individual level. It shows that mixed-mode treatment leads to a statistically significant reduction in both hours and days worked per individual. Specifically, individuals in the mixed-mode group reported working 300.1 fewer hours and 68.2 fewer days compared to those in the control group. These reductions are substantial, representing a 23.4% decrease in hours and a 25.3% decrease in days relative to the control group means of 1,284.5 hours and 269.4 days, respectively.

Table 9 shows the intensive margins effects at the household level; it aggregates input use across all plots and individuals within a household. The reported hours and days worked by household members show minimal changes that are not statistically significant. Even at the household level, the mixed-mode treatment leads to significant increases in the reported quantities of non-labor inputs. The reported quantity of inorganic fertilizer used increased by 127.3%, organic fertilizer by 135.0%, herbicides by 190.9%, pesticides by 118.2%, and the number of days with animal traction by 144.5%.

6. Discussion

This study provides robust evidence on the impact of mixed-mode data collection—combining in-person and high-frequency phone interviews—on the measurement of agricultural inputs in Nigeria. The findings underscore the substantial recall bias present in traditional end-of-season surveys and demonstrate the potential of mixed-mode approaches to improve the accuracy of both labor and non-labor input reporting.

One of the most striking findings is the improvement in reporting at the extensive margin. Households in the mixed-mode group were significantly more likely to report that individuals worked on plots, with particularly large increases in the reporting of women and children’s labor. Similarly, the likelihood of reporting the use of key non-labor inputs—such as fertilizers, herbicides, pesticides, and animal traction—rose significantly. These results suggest that traditional recall surveys may systematically underreport both labor participation and input use, especially for less salient or irregular activities.

At the same time, the mixed-mode approach led to lower reported hours and days worked per individual on plots, indicating that recall-based surveys may overestimate labor intensity. This pattern is consistent with cognitive biases in long recall periods, where respondents may rely on assumptions about typical behavior rather

than actual events. The reductions in reported labor—up to 30% in hours and days worked—underscore the importance of shorter recall windows for capturing accurate measures of labor input.

The effects of the mixed-mode treatment on non-labor inputs at the intensive margin were particularly striking. Across all levels of aggregation, the reported quantities of inputs such as fertilizers, herbicides, pesticides, and animal traction were significantly higher among households in the mixed-mode group. At the plot level, the use of inorganic fertilizer increased by over 117%, while organic fertilizer rose by nearly 150%. Herbicide and pesticide use saw even larger gains—196.9% and 116.9%, respectively—suggesting that traditional recall surveys may severely underestimate the intensity of input application. These effects were consistent at the household level, where increases in reported input quantities remained large and statistically significant. The magnitude of these differences underscores the extent to which recall bias can distort input use estimates, particularly for inputs that are applied intermittently or in varying quantities. These findings reinforce the value of high-frequency data collection in capturing more accurate and granular information on input intensity, which is critical for evaluating productivity, profitability, and the effectiveness of agricultural interventions.

The effects of the mixed-mode treatment varied by the level of data aggregation. At the plot level, the treatment effects were most pronounced, particularly for extensive margin indicators. This level of granularity allowed for more precise attribution of labor and input use, revealing substantial underreporting in traditional surveys. At the individual level, similar patterns emerged, though the magnitudes were slightly attenuated. At the household level, the effects were generally smaller, likely due to aggregation diluting the variation captured at finer levels. Nonetheless, significant increases in reported use of non-labor inputs and labor participation by women and children were still observed, reinforcing the value of disaggregated data collection.

These findings have important implications for survey design in agricultural research and policy. Mixed-mode approaches offer a cost-effective and scalable alternative to high-frequency in-person visits, particularly in settings with high mobile phone penetration. They can substantially reduce recall bias and improve the measurement of inputs that are critical for productivity analysis and policy targeting. However, the results also caution against relying solely on household-level aggregates, which may obscure important heterogeneity and underreporting.

7. References

- Abay, K. A., Bevis, L. E., & Barrett, C. B. (2021). Measurement Error Mechanisms Matter: Agricultural intensification with farmer misperceptions and misreporting. *American Journal of Agricultural Economics*, 103(2), 498-522.
- Abay, K. A., Barrett, C. B., Kilic, T., Moylan, H., Ilukor, J., & Vundru, W. D. (2023). Nonclassical measurement error and farmers' response to information treatment. *Journal of Development Economics*, 164, 103136.
- Aker, J. C., & Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of economic Perspectives*, 24(3), 207-232.
- Alston, J. M., & Pardey, P. G. (2014). Agriculture in the global economy. *Journal of Economic Perspectives*, 28(1), 121-146.
- Alston, J. M., Pardey, P. G., & Rao, X. (2020). The payoff to investing in CGIAR research.
- Amankwah, A., Ambel, A., Gourlay, S., Kilic, T., Markhof, Y., & Wollburg, P. (2025). Smallholder farming, fertilizer use, and the polycrisis period: Cross-country evidence from longitudinal surveys in Sub-Saharan Africa. *Food Policy*, 133, 102885.
- Arthi, V., Beegle, K., De Weerd, J., & Palacios-López, A. (2018). Not your average job: Measuring farm labor in Tanzania. *Journal of Development Economics*, 130, 160-172.
- Avila, A. F. D., & Evenson, R. E. (2010). Total factor productivity growth in agriculture: The role of technological capital. *Handbook of agricultural economics*, 4, 3769-3822.
- Beegle, K., Carletto, C., & Himelein, K. (2012). Reliability of recall in agricultural data. *Journal of Development Economics*, 98(1), 34-41.
- Beegle, K., Gibson, J., & De Weerd, J. (2020). What Can We Learn from Experimenting with Survey Methods? *Annual Review of Resource Economics*, 12.
- Carletto, C., Dillon, A., & Zezza, A. (2021). Agricultural data collection to minimize measurement error and maximize coverage. In *Handbook of agricultural economics* (Vol. 5, pp. 4407-4480). Elsevier.
- Christiaensen, L., Demery, L., & Kuhl, J. (2011). The (evolving) role of agriculture in poverty reduction—An empirical perspective. *Journal of development economics*, 96(2), 239-254.
- Datt, G., & Ravallion, M. (1998). Farm productivity and rural poverty in India. *The Journal of Development Studies*, 34(4), 62-85.
- Desiere, S., & Jolliffe, D. (2018). Land productivity and plot size: Is measurement error driving the inverse relationship?. *Journal of Development Economics*, 130, 84-98.
- Dillon, A., Lybbert, T. J., Michelson, H., & Rudder, J. (2025). Agricultural Input Markets in Sub-Saharan Africa: Theory and Evidence from the (Underappreciated) Supply Side. *Annual Review of Resource Economics*, 17.
- Djima, I. Y., & Kilic, T. (2024). Attenuating measurement errors in agricultural productivity analysis by combining objective and self-reported survey data. *Journal of Development Economics*, 168, 103249.

- Dorosh, P., & Thurlow, J. (2018). Beyond agriculture versus non-agriculture: decomposing sectoral growth–poverty linkages in five African countries. *World Development*, 109, 440-451.
- Fabregas, R., Kremer, M., Lowes, M., On, R., & Zane, G. (2025). Digital information provision and behavior change: Lessons from six experiments in East Africa. *American Economic Journal: Applied Economics*, 17(1), 527-566.
- Freeman, Katie Kennedy, 2021. *Concept Project Information Document (PID) - Uruguay Agro-Ecological and Climate Resilient Systems Project - P176232 (English)*. Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/388571635441090360>
- Gaddis, I., Oseni, G., Palacios-Lopez, A., & Pieters, J. (2021). Measuring farm labor: survey experimental evidence from Ghana. *The World Bank Economic Review*, 35(3), 604-634.
- Gollin, D., Lagakos, D., & Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2), 939-993.
- Gollin, D., & Udry, C. (2021). Heterogeneity, measurement error, and misallocation: Evidence from African agriculture. *Journal of Political Economy*, 129(1), 1-80.
- Gourlay, S., Kilic, T., & Lobell, D. (2017). Could the debate be over? Errors in farmer-reported production and their implications for the inverse scale-productivity relationship in Uganda. *Errors in Farmer-Reported Production and Their Implications for the Inverse Scale-Productivity Relationship in Uganda (September 12, 2017)*. *World Bank Policy Research Working Paper*, (8192).
- Gourlay, S., Kilic, T., Martuscelli, A., Wollburg, P., & Zezza, A. (2021). High-frequency phone surveys on COVID-19: good practices, open questions. *Food policy*, 105, 102153.
- Huang, J., Pray, C., & Rozelle, S. (2002). Enhancing the crops to feed the poor. *Nature*, 418(6898), 678-684.
- Jeronimo, Anderson Renato de Brito., 2025. *Angola - EASTERN AND SOUTHERN AFRICA- P159052- Angola Commercial Agriculture Development Project (PDAC) - Procurement Plan (English)*. Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/099060525201558766>
- Kilic, T., Moylan, H., Ilukor, J., Mtengula, C., & Pangapanga-Phiri, I. (2021). Root for the tubers: Extended-harvest crop production and productivity measurement in surveys. *Food Policy*, 102, 102033.
- Kilic, T., Palacios-López, A., & Goldstein, M. (2015). Caught in a productivity trap: A distributional perspective on gender differences in Malawian agriculture. *World development*, 70, 416-463.
- Michelson, H., Gourlay, S., & Wollburg, P. (2022). Non-Labor Input Quality and Small Farms in Sub-Saharan Africa.
- Minten, B., & Barrett, C. B. (2008). Agricultural technology, productivity, and poverty in Madagascar. *World Development*, 36(5), 797-822.
- Markhof, Y. V., Wollburg, P. R., & Zezza, A. (2023). *Are vaccination campaigns misinformed? Experimental evidence from COVID-19 in low-and middle-income countries* (No. 10443). The World Bank.
- Miyajima, K. (2022). Mobile phone ownership and welfare: Evidence from South Africa's household survey. *World Development*, 154, 105863.

- National Bureau of Statistics (NBS). (2024). General Household Survey, Panel 2023-2024 [sect5b_plantingw4]. World Bank, Development Data Group.
- Oseni, G., Corral, P., Goldstein, M., & Winters, P. (2015). Explaining gender differentials in agricultural production in Nigeria. *Agricultural Economics*, 46(3), 285-310.
- Reardon, T. and P. Glewwe. 2000. "Agriculture," in M. Grosh and P. Glewwe (eds.), *Designing household survey questionnaires for developing countries: lessons from ten years of LSMS experience* (Washington: The World Bank).
- Rezek, J. P., Campbell, R. C., & Rogers, K. E. (2011). Assessing total factor productivity growth in Sub-Saharan African agriculture. *Journal of agricultural economics*, 62(2), 357-374.
- Timmer, C. P. (1995). Getting agriculture moving: do markets provide the right signals?. *Food policy*, 20(5), 455-472.
- World Bank. (2025). World Development Indicators: Structure of Value Added. <https://wdi.worldbank.org/table/4.2>
- Wollburg, P., Tiberti, M., & Zezza, A. (2021). Recall length and measurement error in agricultural surveys. *Food Policy*, 100, 102003.

8. Tables and Figures

Figure 1. Experimental Design

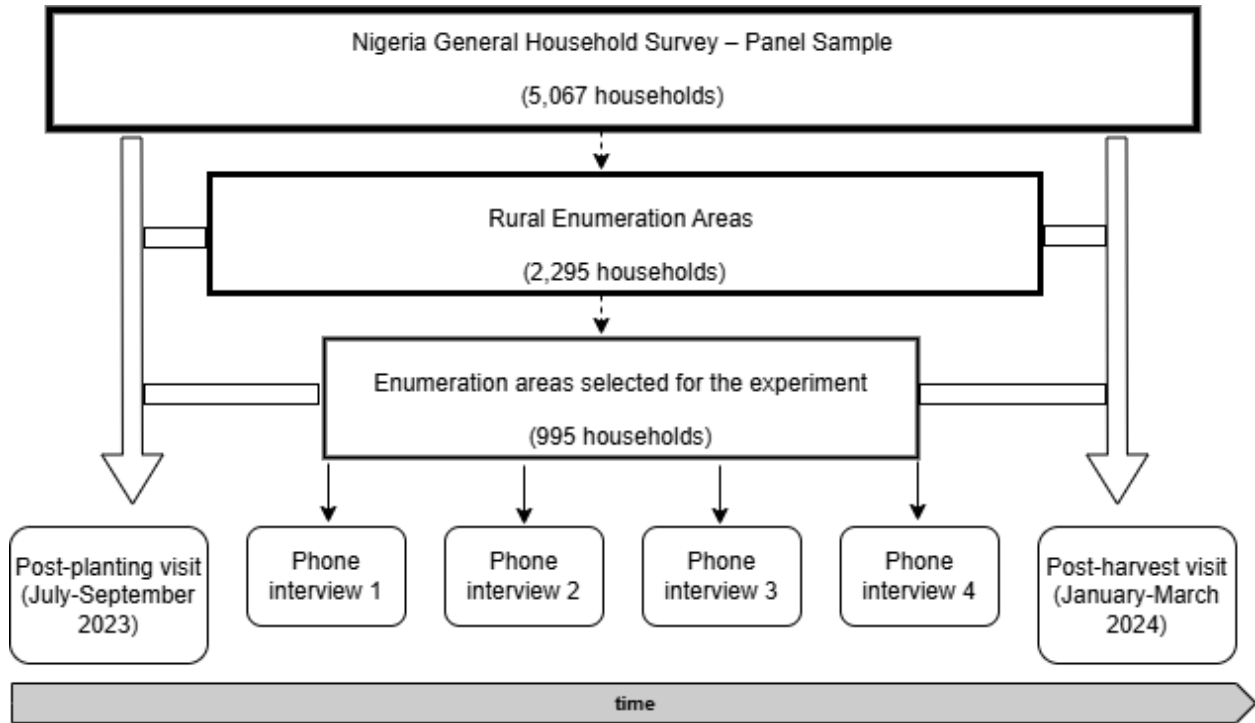


Table 1. Sample Characteristics in Baseline Survey by Treatment Arm – Post-Planting visit

	(1) Control	(2) Mixed-Mode	(3) Diff.
Household in rural area	0.864 (0.343)	0.875 (0.331)	0.011 (0.039)
Household size (# of members)	6.208 (3.779)	6.020 (3.745)	-0.188 (0.276)
Male household head	0.786 (0.410)	0.776 (0.417)	-0.010 (0.024)
Age household head	53.332 (15.210)	53.418 (14.554)	0.086 (0.842)
Head completed more than primary school	0.361 (0.480)	0.400 (0.490)	0.039 (0.029)
Dependency ratio	0.443 (0.259)	0.416 (0.253)	-0.027** (0.013)
Assets index	-0.350 (0.687)	-0.281 (0.725)	0.069 (0.057)
Plots owned by the household (# of plots)	2.356 (1.281)	2.514 (1.332)	0.158 (0.102)
Household had more than 3 plots	0.168 (0.374)	0.186 (0.389)	0.018 (0.025)
Household raised or owned any animal	0.497 (0.500)	0.487 (0.500)	-0.009 (0.036)
Household members who worked on plots (# of members)	4.894 (5.143)	4.919 (4.745)	0.025 (0.314)
<i>N (households)</i>	<i>2,295</i>	<i>790</i>	<i>3,085</i>
Age	24.422 (19.954)	25.142 (19.991)	0.721 (0.624)
Male	0.503 (0.500)	0.492 (0.500)	-0.011 (0.009)
Never Married	0.341 (0.474)	0.347 (0.476)	0.007 (0.013)
<i>N (individuals)</i>	<i>14,247</i>	<i>4,756</i>	<i>19,003</i>
Plot area (in Sq meters)	5,328.458 (10,556.748)	5,044.993 (9,906.948)	-283.465 (528.611)
Log plot area above the median	0.504 (0.500)	0.467 (0.499)	-0.037 (0.034)
Proportion of plots cultivating maize	0.220 (0.414)	0.227 (0.419)	0.007 (0.023)
Proportion of plots cultivating yam	0.086 (0.280)	0.108 (0.311)	0.022 (0.018)

Proportion of plots cultivating cassava	0.222 (0.416)	0.279 (0.449)	0.057* (0.034)
Proportion of plots cultivating rice	0.068 (0.251)	0.107 (0.310)	0.040** (0.016)
Proportion of plots intercropped	0.327 (0.469)	0.319 (0.466)	-0.008 (0.024)
Proportion of plots irrigated	0.015 (0.122)	0.017 (0.128)	0.001 (0.006)
<i>N (plots)</i>	<i>5,406</i>	<i>1,986</i>	<i>7,392</i>
Hours worked on the plot, reported in PP visit (# of hours)	330.354 (316.411)	328.764 (301.913)	-1.590 (25.320)
Days worked on the plot, reported in PP visit (# of hours)	68.331 (46.828)	67.367 (43.049)	-0.964 (3.865)
<i>N (Individuals-plots)</i>	<i>10,229</i>	<i>3,550</i>	<i>13,779</i>

Notes: This table presents baseline characteristics across treatment groups. Column (1) reports the mean for the Control group, Column (2) the mean for the Mixed-mode group, and Column (3) reports the difference between the two, with standard errors clustered at the EA level. All variables were collected during the GHS-Panel Wave 5 post-planting visit. Variables are binary indicators unless a unit of measurement is specified in parentheses. *Plot area* reflects the average size of each plot, with values winsorized at the 99th percentile to reduce the influence of outliers. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2. Extensive Margin Effects at the Individual-Plot Level

	(1)
	Individual worked on the plot
Plot included in the Mixed-Mode	0.402*** (0.071)
Household included in the Mixed-Mode	-0.134* (0.073)
Observations	28,792
Outcome Control Mean	0.436
Percent Change (Plot included in the Mixed-Mode)	92.2

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome, in this case the control group refers to all the plots excluded from the mixed-mode study, even if they belonged to a household included in the mixed-mode study. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3a. Extensive Margin Effects at the Plot Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of household members working in the plot				Hired workers			Unpaid workers		
	Total	Women	Men	Children	Women	Men	Children	Women	Men	Children
Plot included in the Mixed-Mode	2.509*** (0.391)	0.731*** (0.132)	0.844*** (0.267)	0.915*** (0.102)	0.288*** (0.074)	0.204*** (0.073)	0.083*** (0.029)	0.146*** (0.030)	0.216*** (0.045)	0.062*** (0.018)
Household included in the Mixed-Mode	-1.029*** (0.391)	-0.235* (0.139)	-0.428 (0.266)	-0.357*** (0.096)	0.012 (0.072)	0.006 (0.074)	0.016 (0.026)	-0.020 (0.027)	-0.012 (0.043)	0.003 (0.017)
Observations	5,165	5,165	5,165	5,165	5,165	5,165	5,165	5,165	5,165	5,165
Outcome Control Mean	2.472	0.767	1.247	0.458	0.195	0.508	0.031	0.056	0.084	0.020
Percent change (Plot included in the Mixed-Mode)	101.5	95.3	67.7	199.8	147.9	40.1	267.6	260.0	257.6	309.0

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable, Columns (5)–(10) present binary indicators (dummy variables) that take the value of 1 if the household satisfies the corresponding condition, and 0 otherwise. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3b. Extensive Margin Effects at the Plot Level

	(1)	(2)	(3)	(4)	(5)
	Inorganic fertilizer was used	Organic fertilizer was used	Herbicide fertilizer was used	Pesticide fertilizer was used	Animal traction was used
Plot included in the Mixed- Mode	0.340*** (0.098)	0.254*** (0.047)	0.346*** (0.114)	0.301*** (0.069)	0.076* (0.039)
Household included in the Mixed-Mode	-0.094 (0.104)	0.012 (0.045)	-0.049 (0.117)	0.026 (0.069)	0.040 (0.036)
Observations	5,165	5,165	5,165	5,165	5,165
Outcome Control Mean	0.369	0.236	0.382	0.164	0.201
Percent change	92.3	107.6	90.6	183.8	38.0

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4. Extensive Margin Effects at the Individual Level

	(1)
	Individual worked at any plot
Mixed-Mode Treatment	0.223*** (0.019)
Observations	17,798
Outcome Control Mean	0.394
Percent Change	56.7

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5a. Extensive Margin Effects at the Household Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of household members working in the plots				Hired workers			Unpaid workers		
	Total	Women	Men	Children	Women	Men	Children	Women	Men	Children
Mixed-Mode Treatment	1.293*** (0.114)	0.445*** (0.057)	0.328*** (0.049)	0.513*** (0.054)	0.284*** (0.027)	0.200*** (0.028)	0.112*** (0.017)	0.149*** (0.020)	0.252*** (0.024)	0.088*** (0.014)
Observations	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068
Outcome Control Mean	2.299	0.720	1.125	0.454	0.222	0.512	0.039	0.071	0.103	0.026
Percent change	56.2	61.8	29.1	113.0	127.8	39.0	287.8	210.5	244.6	337.8

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses, and the list of control variables selected using a double-LASSO procedure for each outcome. All estimations include zone fixed effects. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5b. Extensive Margin Effects at the Household Level

	(1)	(2)	(3)	(4)	(5)
	Inorganic fertilizer was used	Organic fertilizer was used	Herbicide was used	Pesticide was used	Animal traction was used
Mixed-Mode Treatment	0.251*** (0.031)	0.277*** (0.029)	0.289*** (0.033)	0.331*** (0.024)	0.104*** (0.027)
Observations	3,068	3,068	3,068	3,068	3,068
Outcome Control Mean	0.384	0.251	0.376	0.188	0.199
Percent change	65.3	110.2	76.9	175.9	52.1

Notes: OLS

regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses, and the list of control variables selected using a double-LASSO procedure for each outcome. All estimations include zone fixed effects. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6. Intensive Margin Effects at the Individual-Plot Level

	(1)	(2)
	Hours worked per individual-plot	Days worked per individual-plot
Plot included in the Mixed-Mode	-204.938** (101.195)	-46.816*** (16.015)
Household included in the Mixed-Mode	35.420 (107.608)	8.165 (17.267)
Observations	14,533	14,533
Outcome Control Mean	699.514	146.577
Percent change	-29.3	-31.9

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. Both outcomes include the hours/days reported in the post-planting visit. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7. Intensive Margin Effects at the Plot Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Hours worked by HH members	Days worked by HH members	Qty of inorganic fertilizer used	Qty of organic fertilizer used	Qty of herbicides used	Qty of pesticides used	Number of days with animal traction
Plot included in the Mixed-Mode	175.420 (281.824)	33.894 (51.137)	117.544*** (31.461)	2,266.675*** (636.122)	6.974*** (0.983)	3.455*** (1.182)	3.679*** (0.439)
Household included in the Mixed-Mode	-80.467 (332.960)	-21.104 (60.584)	6.800 (27.217)	-415.241 (644.500)	-0.820 (0.735)	-0.612 (1.101)	-0.755*** (0.234)
Observations	4,860	4,860	2,203	1,536	2,341	1,258	1,137
Outcome Control Mean	2159.646	456.112	100.148	1514.927	3.515	2.957	2.274
Percent change	8.1	7.4	117.4	149.6	198.4	116.9	161.8

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. The hours and days worked include the hours/days reported in the post-planting visit. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8. Intensive Margin Effects at the Individual Level

	(1)	(2)
	Hours worked per individual	Days worked per individual
Mixed-Mode Treatment	-300.148*** (49.807)	-68.241*** (9.234)
Observations	8,040	8,040
Outcome Control Mean	1284.499	269.378
Percent change	-23.4	-25.3

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. The hours and days worked include the hours/days reported in the post-planting visit. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9. Intensive Margin Effects at the Household Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Hours worked by HH members	Days worked by HH members	Quantity of inorganic fertilizer used	Quantity of organic fertilizer used	Quantity of herbicides used	Quantity of pesticides used	Number of days with animal traction
Mixed-Mode Treatment	69.340 (186.745)	-3.178 (32.827)	211.579*** (40.764)	3,021.176*** (627.105)	11.469*** (1.422)	5.063*** (1.064)	5.414*** (0.794)
Observations	2,603	2,603	1,353	979	1,360	824	666
Outcome Control Mean	4037.594	853.179	167.755	2397.804	6.007	4.284	3.826
Percent change	1.7	-0.4	126.1	126.0	190.9	118.2	141.5

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. The hours and days worked include the hours/days reported in the post-planting visit. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10. Heterogeneous Effects by Plot Size at the Plot Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of household members working in the plot				Hired workers			Unpaid workers		
	Total	Women	Men	Children	Women	Men	Children	Women	Men	Children
Plot included in the Mixed-Mode	2.440*** (0.359)	0.710*** (0.182)	0.888*** (0.237)	0.816*** (0.156)	0.247** (0.107)	0.216** (0.087)	-0.008 (0.058)	0.042 (0.054)	0.158*** (0.045)	0.020 (0.041)
Household included in the Mixed-Mode	-1.005*** (0.363)	-0.303 (0.188)	-0.417* (0.225)	-0.288* (0.147)	0.073 (0.111)	0.042 (0.085)	0.068 (0.059)	0.044 (0.054)	0.010 (0.043)	0.027 (0.042)
Plot in Mixed-Mode x Log plot area above the median	0.136 (0.384)	0.045 (0.197)	-0.065 (0.299)	0.164 (0.172)	0.069 (0.108)	-0.010 (0.108)	0.153** (0.063)	0.175*** (0.057)	0.099** (0.045)	0.068 (0.045)
Household in Mixed-Mode x Log plot area above the median	-0.064 (0.390)	0.087 (0.186)	-0.030 (0.293)	-0.106 (0.160)	-0.103 (0.107)	-0.087 (0.106)	-0.086 (0.061)	-0.108** (0.053)	-0.039 (0.040)	-0.039 (0.044)
Log plot area above the median	0.176** (0.071)	0.022 (0.041)	0.122*** (0.039)	0.035 (0.041)	0.054*** (0.019)	0.124*** (0.025)	-0.017* (0.009)	0.005 (0.012)	-0.008 (0.013)	-0.022*** (0.007)
Plot in Mixed Mode + Plot in Mixed-Mode x Log plot area above the median	2.576*** (0.463)	0.755*** (0.159)	0.823** (0.338)	0.980*** (0.114)	0.317*** (0.074)	0.206** (0.093)	0.145*** (0.025)	0.217*** (0.023)	0.257*** (0.051)	0.089*** (0.016)
Observations	5,165	5,165	5,165	5,165	5,165	5,165	5,165	5,165	5,165	5,165
Outcome Control Mean	2.472	0.767	1.247	0.458	0.195	0.508	0.031	0.056	0.084	0.020

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 11. Heterogeneous Effects by Plot Size at the Plot Level

	(1)	(2)	(3)	(4)	(5)
	Inorganic fertilizer was used	Organic fertilizer was used	Herbicide fertilizer was used	Pesticide fertilizer was used	Animal traction was used
Plot included in the Mixed-Mode	0.259*** (0.082)	0.377*** (0.053)	0.228 (0.165)	0.254*** (0.055)	0.077** (0.036)
Household included in the Mixed-Mode	-0.010 (0.093)	-0.068 (0.051)	0.014 (0.174)	0.021 (0.049)	-0.017 (0.035)
Plot in Mixed-Mode x Log plot area above the median	0.136 (0.093)	-0.206** (0.082)	0.204 (0.140)	0.088 (0.080)	0.010 (0.062)
Household in Mixed- Mode x Log plot area above the median	-0.140 (0.099)	0.131* (0.076)	-0.111 (0.148)	-0.001 (0.077)	0.082 (0.055)
Log plot area above the median	0.034 (0.022)	0.017 (0.020)	0.049** (0.024)	0.047*** (0.018)	0.053*** (0.020)
Plot in Mixed-Mode + Plot in Mixed-Mode x Log plot area above the median	0.395*** (0.116)	0.171*** (0.064)	0.432*** (0.098)	0.342*** (0.091)	0.087 (0.058)
Observations	5,165	5,165	5,165	5,165	5,165
Outcome Control Mean	0.369	0.236	0.382	0.164	0.201

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 12. Heterogeneous Effects by Plot Size at the Plot Level

	(1) Hours worked by HH members including reported in PP visit	(2) Days worked by HH members including reported in PP visit	(3) Quantity of inorganic fertilizer used	(4) Quantity of organic fertilizer used	(5) Quantity of herbicides used	(6) Quantity of pesticides used	(7) Number of days with animal traction
Plot included in the Mixed-Mode	573.715** (235.108)	98.849* (55.512)	72.360*** (20.393)	1,582.123* (830.403)	4.349*** (0.775)	1.135 (1.048)	3.747** (1.442)
Household included in the Mixed-Mode	-477.359* (282.942)	-79.461 (63.341)	10.239 (18.082)	-741.109 (840.797)	0.022 (0.576)	-0.355 (1.319)	-0.699*** (0.164)
Plot in Mixed-Mode x Log plot area above the median	-606.033* (332.134)	-97.318 (74.015)	77.767* (41.470)	1,277.727 (833.016)	3.754*** (1.321)	2.964 (1.987)	-0.069 (1.454)
Household in Mixed-Mode x Log plot area above the median	598.351* (354.754)	84.360 (79.247)	-17.477 (28.957)	210.443 (783.892)	-1.370 (0.874)	-0.354 (1.874)	-0.078 (0.307)
Log plot area above the median	436.128*** (78.334)	66.669*** (12.581)	32.773*** (9.313)	192.990 (227.787)	0.877*** (0.282)	-1.103 (0.943)	0.579*** (0.149)
Plot in Mixed Mode + Plot in Mixed-Mode x Log plot area above the median	-32.318 (353.869)	1.531 (65.319)	150.127*** (41.267)	2859.849*** (574.509)	8.103*** (1.245)	4.099*** (1.503)	3.678*** (0.422)
Observations	4,860	4,860	2,203	1,536	2,341	1,258	1,137
Outcome Control Mean	2159.646	456.112	100.148	1514.927	3.515	2.957	2.274

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 13. Heterogeneous Effects by Plot Manager Education at the Plot Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Hours worked by HH members including reported in PP visit	Days worked by HH members including reported in PP visit	Quantity of inorganic fertilizer used	Quantity of organic fertilizer used	Quantity of herbicides used	Quantity of pesticides used	Number of days with animal traction
Plot included in the Mixed-Mode	81.889 (344.334)	39.837 (64.337)	111.413*** (40.946)	2,412.424*** (803.551)	6.582*** (1.283)	2.729*** (1.006)	4.063*** (0.607)
Household included in the Mixed-Mode	81.429 (374.574)	-7.318 (71.940)	17.568 (36.248)	-217.238 (798.783)	-0.031 (0.849)	0.578 (0.779)	-0.716** (0.348)
Plot in Mixed-Mode x Manager completed more than primary school	116.650 (526.901)	-21.724 (92.029)	21.138 (50.415)	-9.550 (896.481)	0.721 (1.737)	3.813 (2.849)	-1.453* (0.848)
Household in Mixed-Mode x Manager completed more than primary school	-269.118 (564.810)	-25.688 (99.908)	-35.342 (39.528)	-1,019.192 (852.027)	-1.685 (1.099)	-5.148** (2.469)	-0.038 (0.441)
Manager completed more than primary school	-115.369 (83.114)	-22.394 (15.679)	19.228* (10.273)	-153.062 (207.679)	0.183 (0.290)	0.256 (0.503)	-0.153 (0.165)
Plot in Mixed Mode + Plot in Mixed-Mode x Manager completed more than primary school	198.539 (414.276)	18.113 (71.958)	132.551*** (31.904)	2402.874*** (364.562)	7.303*** (1.279)	6.542** (2.538)	2.609*** (0.449)
Observations	4,860	4,860	2,203	1,536	2,341	1,258	1,137
Outcome Control Mean	2159.646	456.112	100.148	1514.927	3.515	2.957	2.274

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 14. Heterogeneous Effects by Sex of Household Head at the Household Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of household members working in the plot				Hired workers			Unpaid workers		
	Total	Women	Men	Children	Women	Men	Children	Women	Men	Children
Mixed-Mode Treatment	0.631*** (0.128)	0.262*** (0.082)	0.171** (0.068)	0.194*** (0.057)	0.257*** (0.040)	0.194*** (0.051)	0.096*** (0.027)	0.120*** (0.036)	0.166*** (0.033)	0.067*** (0.023)
Male household head	-0.164** (0.084)	-0.326*** (0.046)	0.336*** (0.048)	-0.169*** (0.042)	0.008 (0.022)	0.081*** (0.025)	-0.016 (0.010)	-0.031* (0.016)	0.001 (0.014)	-0.015* (0.008)
Mixed-Mode x Male household head	0.839*** (0.172)	0.232** (0.099)	0.202** (0.078)	0.411*** (0.082)	0.034 (0.044)	0.006 (0.052)	0.020 (0.030)	0.038 (0.039)	0.110*** (0.039)	0.027 (0.026)
Mixed Mode + Mixed Mode x Male household head	1.470*** (0.131)	0.494*** (0.067)	0.373*** (0.056)	0.605*** (0.066)	0.291*** (0.030)	0.200*** (0.030)	0.117*** (0.019)	0.158*** (0.022)	0.276*** (0.028)	0.094*** (0.015)
Observations	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068
Outcome Control Mean	2.299	0.720	1.125	0.454	0.222	0.512	0.039	0.071	0.103	0.026

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 15. Heterogeneous Effects by Sex of Household Head at the Household Level

	(1)	(2)	(3)	(4)	(5)
	Inorganic fertilizer was used	Organic fertilizer was used	Herbicide was used	Pesticide was used	Animal traction was used
Mixed-Mode Treatment	0.262*** (0.049)	0.287*** (0.043)	0.190*** (0.047)	0.232*** (0.040)	0.029 (0.022)
Male household head	0.047* (0.024)	0.052*** (0.020)	0.021 (0.024)	0.003 (0.019)	0.006 (0.016)
Mixed-Mode x Male household head	-0.015 (0.049)	-0.013 (0.045)	0.126*** (0.049)	0.124*** (0.043)	0.096*** (0.032)
Mixed Mode + Mixed Mode x Male household head	0.247*** (0.033)	0.273*** (0.031)	0.317*** (0.035)	0.358*** (0.028)	0.125*** (0.031)
Observations	3,068	3,068	3,068	3,068	3,068
Outcome Control Mean	0.384	0.251	0.376	0.188	0.199

Notes: OLS regressions of various outcomes on a dummy for the survey mode of the interview. Standard errors clustered at the EA level in parentheses. All estimations include zone fixed effects, and the list of control variables selected using a double-LASSO procedure for each outcome. Each column is a separate dependent variable. Outcome Control Mean refers to the mean of the control group for each outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$