Pre-Analysis Plan: Cultivating Connectivity: Measuring the Impact of a Digital Agricultural Extension Platform on Millennial Farmers and Extension Workers in Indonesia (LenteraDigiEx)

Irene Furlani and Michael Grimm

January 3, 2026

Abstract

To attract Millennial Farmers (MFs) into an aging agricultural sector, Indonesian Extension Workers (EWs) must match their digital literacy levels and offer digital solutions to their pool of farmers. Moreover, digital platforms offer a solution to reduce the cost of capacity building and activity support for EWs, as well as the provision of information and services for MFs. We assess the impact of two formats (fully online and blended offline–online) of a double-sided intervention on digital literacy, as well as knowledge and adoption of the digital platform Lentera DESA (where the self-paced course component is offered), using an experimental design. On the supply side, we also examine the impact of the project on extension performance and capabilities. On the demand side, we consider the impact on knowledge and adoption of agricultural and business practices, agricultural income, loans and investments, and employment. We hypothesize that the interventions will increase digital literacy, platform adoption, and usage. The use of the Lentera DESA platform may increase EWs' performance and capabilities and may also increase MFs' knowledge and use of sustainable agricultural and business practices, ultimately improving MFs' overall welfare.

Revision of the Pre-Analysis Plan (PAP)

This version of the pre-analysis plan (PAP) revises the initial version for three primary reasons: (i) insights obtained from the baseline data collection and the inception workshop; (ii) changes to the planned study timeline due to financial constraints; and (iii) the need to improve completeness and clarity.

Insights from activities conducted at baseline, informed refinements to the intervention design, allowing the intervention content to better reflect the needs of the study's direct beneficiaries, namely extension workers (EWs) and Millennial Farmers (MFs). While the structure of the self-paced course for both beneficiaries includes the first modules related to digital literacy, EWs have and extensive part related to social media content, while the MFs had more focus on the automation of agricultural practices, e-commerce, and digital farm registry (already mentioned before). These changes are reflected in the revised description of the intervention and entail corresponding adjustments to the research questions, their prioritization, and the inclusion of additional survey items.

In addition, due to budget constraints, the planned midline data collection was not implemented. For midline data collection we wanted to deploy a short survey to ask only about digital literacy and training received (not a full survey like at baseline and endline). To partially compensate for the absence of a midline survey, additional questions were incorporated into the remaining survey instruments.

Finally, the revised PAP provides greater detail on the theoretical framework and empirical analysis in order to clarify the study's analytical approach and enhance transparency regarding the planned analyses.

1 Introduction

1.1 Motivation

Indonesia's agricultural sector faces an aging workforce, a shortage of extension workers (EWs), and the need to adopt sustainable farming practices to achieve food security. Attracting younger cohorts to the agricultural sector and integrating digital tools into extension services could address both challenges simultaneously. Although the government offers economic incentives to young farmers, extension services struggle to engage them due to digital literacy gaps and limited outreach capacity. Many EWs lack digital skills, creating a mismatch in knowledge transfer (Kaliky et al., 2023; Sugihono et al., 2024). At the same time, relatively high education levels and adequate digital literacy among younger cohorts present an opportunity to leverage digital platforms and make agriculture more appealing to young farmers.

The integration of digitization into agricultural extension services has attracted increasing attention, although evidence on its effectiveness remains limited. Existing studies suggest that digital extension services can improve awareness, use, and integration of digital resources, as well as enhance the digital literacy of extension workers (Tata and Mcnamara, 2017; Enwelu et al., 2017). In the context of Indonesia, Suswadi and Irawan (2023) find a positive association between digital extension services and the performance of agricultural extension systems. From the farmers' perspective, digitization also holds promise for improving productivity and welfare outcomes. However, these benefits may vary across demographic groups, as lower levels of digital literacy among older and less-educated farmers may hinder equitable access and utilization (Aker, 2011; Baumüller, 2017). Addressing such disparities, Ngadi et al. (2023) suggest that integrating younger farmers into the sector could enhance productivity by leveraging their greater facility with information and communication technologies (ICTs) to bridge generational gaps in digital adoption.

According to the meta-analysis by Beach et al. (2025), experimental and quasi-experimental evidence on the impact of digital information interventions on agricultural development remains limited, comprising only 20 studies. The authors report an average yield and income increase of approximately six percent, based on a small number of studies that capture effects only in Sub-Saharan Africa, India, and Cambodia. Moreover, the intervention characteristics do not include any self-paced courses delivered via digital platforms. The most common digital mode is video, typically shown on a tablet or sent to recipients. Only four interventions are based on digital platforms (three mobile applications and one website), all implemented in Sub-Saharan Africa. Importantly, none of these platforms provides agricultural information directly to all end users, as they primarily target extension workers or other types of officers. The only platform that disseminates information directly to farmers focuses exclusively on price information.[1]

1.2 Research Questions

Table 1 summarizes the overarching questions and sub-questions of this project, with the related indicators (with information on the recall period and respondent), and whether they are part of the primary or secondary analysis. The first overarching question relates to the compliance with treatment assignment. While the answer to this question does not provide valuable information from an academic perspective, it

[1] Their paper relies solely on published experimental and quasi-experimental evidence; therefore, other platforms may exist whose effects have not been rigorously evaluated.

Table 1: Summary of research questions and related indicators.

Overarching (Research) Questions	Subquestion	Indicator(s)	Period	Focus	Importance
What is the level of compliance of LenteraDigiEx?	What is the effect of LenteraDigiEx on the respondent's training uptake?	From the survey data:Training invitation received (0/1), training attendance (0/1), training completion (0/1), any certificate obtained (0/1). Additionally, these question can exploit monitoring data related to training attendance, self-paced course completion rate, and pre- and post-test scores.	-	Both	Primary
RQ1: What is the effect of digital literacy and agricultural digitalization training on digital literacy and platform knowledge and usage of the respondent?	RQ1.1: What is the effect of digital literacy and agricultural digitalization training on digital literacy of the respondent?	Composite digital literacy indicator (0-3)	-	Both	Primary
	RQ1.1a: What is the effect of digital literacy and agricultural digitalization training on each dimension of digital literacy?	Composite digital literacy indicator (0-3) for each dimension (digital information and data literacy, digital communication and collaboration, digital content creation, digital safety digital problem solving)	-	Both	Secondary
	RQ1.2: What is the effect of digital literacy and agricultural digitalization training on LD platform knowledge?	Knowledge of LD (0/1)	-	Both	Primary
	RQ1.3: What is the effect of digital literacy and agricultural digitalization training on LD platform use?	Use of LD (0/1)	-	Both	Primary
RQ2: What is the effect of digital literacy and agricultural digitalization training on EWs' performance and capabilities?	RQ2.1: What is the effect of digital literacy and agricultural digitalization training on EWs' performance?	Integration of digital tools into activities (0/1), Integration of LD into activities (0/1), number of extension activities done in total	Last 12 months	EWs	Primary
	RQ2.1a: What is the effect of digital literacy and agricultural digitalization training on EWs' performance?	Number of extension activities done by activity type	Last 12 months	EWs	Secondary
	RQ2.2: What is the effect of digital literacy and agricultural digitalization training on EWs' capabilities?	Number of farmers reached, number of MFs reached, area covered in acres	Last 12 months	EWs	Primary
	RQ2.2a: What is the effect of digital literacy and agricultural digitalization training on EWs' capability to reach farmers of different gender?	Number of male farmers reached, number of female farmers reached, number of male MFs reached, number of female MFs reached	Last 12 months	EWs	Secondary
RQ3: What is the effect of digital literacy and agricultural digitalization training on MFs' knowledge and adoption of practices, and their welfare?	RQ3.1: What is the effect of digital literacy and agricultural digitalization training on MFs' knowledge of agricultural and business practices?	Knowledge of each main practice (0/1)	-	MFs	Primary
	RQ3.2: What is the effect of digital literacy and agricultural digitalization training on MFs' adoption of agricultural and business practices?	Use of each main practice (0/1)	-	MFs	Primary
	RQ3.2: What is the effect of digital literacy and agricultural digitalization training on MFs' agricultural income?	Agricultural income (1,000 IDR - Indonesian Rupiah), agricultural sales (1,000 IDR)	Last 12 months	MFs	Primary
	RQ3.2a: What is the effect of digital literacy and agricultural digitalization training on MFs' agricultural productivity?	Agricultural productivity (1,000 IDR per acre)	Last 12 months	MFs	Secondary
	RQ3.3 What is the effect of digital literacy and agricultural digitalization training on MFs' food security and resilience?	Reduced Coping Strategies Index (rCSI - 0-27)	Last 7 days	MFs	Secondary
	RQ3.4 What is the effect of digital literacy and agricultural digitalization training on MFs' loan and investments?	Number of Loans, total loan amount (1,000 IDR), Investment amount (1,000 IDR)	Last 12 months	MFs	Secondary
	RQ3.5 What is the effect of digital literacy and agricultural digitalization training on employment within the MF's activity?	Presence of additional workers (0/1), number of people employed for a wage, number of family members not paid for work, total number of people engaged in the agricultural activity (sum of the last two)	-	MFs	Secondary
	RQ3.6 What is the effect of digital literacy and agricultural digitalization training on MFs' formalization?	Presence of any type of formalization (0/1), The MF agricultural activity has a specific type of formalization (0/1 - for each license or certification).	-	MFs	Secondary

gives us a basis for cost-benefit analysis, which has policy relevance. The other three questions analyze the impact of digital literacy and agricultural digitalization training on: (i) digital literacy and platform knowledge and use for MFs and EWs; (ii) EWs' performance and capabilities; and (iii) MFs' knowledge and adoption of practices, as well as their welfare. These three research questions allow us to contribute to the literature agricultural digital information interventions in middle income countries reporting effects both for the demand and supply side of agricultural extension services.

2 Research Strategy

2.1 Sampling

2.1.1 Sampling Frame

The eligible population consists of registered governmental extension workers (EWs) and registered Millennial Farmers (MFs) in the Special Region of Yogyakarta, serving subdistricts with at least two MFs. Our sample of respondents is drawn from 51 subdistricts corresponding to agricultural extension service offices,[2] located across five regencies in the Special Region of Yogyakarta, Indonesia. These 51 subdistricts cover almost the entire area under analysis. Coverage is incomplete because five offices were excluded from the analysis, as they serve at most two MFs. The exclusions were made for reasons of implementation efficiency and cost–benefit considerations.

The planned proportional random sample consists of 839 MFs and 171 EWs, drawn from a sampling frame of 1,624 MFs and 303 EWs. The sampling frame is based on administrative data collected by the Agriculture and Food Security Office of the local government[3] and includes governmental EWs and registered MFs.

We designed the sample to include a similar number of respondents across treatment arms to increase statistical power. To this end, we first allocated 50% of individuals in the sampling frame across the three study arms (two treatment groups

[2]The subdistrict level coincides in most cases with the same administrative categorization, except for the regency of Sleman, where administrative subdistricts are grouped into eight offices.

[3]Specifically, the data originate from the *Unit Pelaksana Teknis Daerah (UPTD) Balai Pengembangan Sumber Daya Manusia Pertanian (BPSDMP) Dinas Pertanian dan Ketahanan Pangan DIY*, which is the Regional Technical Implementation Unit (UPTD) of the Agricultural Human Resources Development Center (BPSDMP) within the Agriculture and Food Security Office of the Special Region of Yogyakarta (DIY).

and one control group), such that

$$R_T = \frac{N \times 0.5}{3},$$

where R_T denotes the ideal number of respondents per treatment arm, and N is the total number of individuals in the sampling frame. In a second step, we determined the number of respondents per subdistrict according to

$$\lceil R_{TS} \rceil = \frac{N_{TS}}{N_T} \times R_T.$$

Specifically, the number of selected respondents R in subdistrict S within treatment arm T is given by the ceiling of the proportion of individuals in that subdistrict and treatment arm (N_{TS}) relative to the total population in the treatment arm (N_T), multiplied by the total number of respondents assigned to that treatment arm (R_T). We rounded up all fractional values to ensure that at least the target number of respondents per treatment arm was reached.

During baseline data collection (February 2025), we identified concerns regarding the reliability of the MF data, as many potential respondents had only attended one or more events organized by BPSDMP. We therefore requested additional data from the local agricultural department (DINAS). After randomly selecting potential replacements from this supplementary list, we obtained a final baseline sample of 784 MFs and 170 EWs. Due to budget constraints, the planned midline data collection was canceled. Consequently, we aim to obtain 1,908 observations from 954 respondents through the endline data collection.

2.1.2 Statistical Power

We performed power calculations using our baseline data collection for the composite digital literacy indicator (scale 0–4), assuming a 95% confidence interval and a statistical power of 80%. The baseline data show a mean digital literacy score of 1.74, with a standard deviation of 0.49 and an intra-cluster correlation (ICC) of 0.069. Given an average cluster size of 18 respondents per cluster and 17 clusters per treatment arm, the study is powered to detect a standardized Minimum Effect Sizes of 0.165. In separate analyses, we can detect minimum effects of 0.189 for MFs and 0.205 for EWs.

2.1.3 Assignment to Treatment

We implemented a stratified random assignment at the subdistrict level, with misfit corrections to account for imbalances in the distribution of extension workers (EWs) across subdistricts. Stratification was based on whether the number of EWs in a subdistrict was above or below the median number of EWs per subdistrict in the sampling frame. Within each stratum, subdistricts were randomly assigned in equal proportions to one of three study arms (two treatment arms and one control arm), with 17 subdistricts allocated to each arm. Misfit corrections were applied globally to ensure balance in the distribution of EWs across treatment arms.

2.1.4 Attrition from the Sample

We do not know what the attrition rate will be, but it could be the case that due to higher non-compliance, some MFs may refuse to take part in the endline questionnaire. We will train enumerators and try to keep attrition at its minimum, while also checking the reasons for it. If overall attrition exceeds 10% and there is evidence of differential attrition by treatment status, we will assess the robustness of the estimated treatment effects by computing pairwise Lee bounds.

2.2 Fieldwork

2.3 Intervention

The intervention includes two treatment arms. In the first treatment arm (light training), participants are offered an online training session followed by an online self-paced course. In the second treatment arm (intensive training), participants receive the same training session delivered in person (offline) followed by the same online self-paced course. The two formats are identical in content. The training sessions include an introduction to the project, an overview of the Lentera DESA platform, and guidance on account creation.

In both treatment arms, the initial training session is followed by a self-paced online course delivered via the Lentera DESA platform. This course covers topics related to digital literacy and the digitalization of agricultural practices. The training concludes with participant presentations, and certificates are awarded upon successful completion of the course.

Two versions of the self-paced course are implemented: one tailored to millennial farmers (MFs) and one tailored to extension workers (EWs). Each self-paced course consists of four modules. The first two modules focus on digital literacy and agricultural digitalization through the automation of simple prototypes (e.g., irrigation systems or livestock feeding systems). The remaining two modules are tailored to participant type.

For EWs, the course includes an introduction to digital farm record-keeping (Buku Tani)[4] followed by an in-depth module on digital content creation, with a particular focus on social media and its associated ethical considerations. For MFs, the course includes modules on e-commerce and e-banking, as well as basic financial literacy, which serve as an introduction to the use of Buku Tani for agribusiness management purposes.

2.3.1 Instruments

The primary data sources for this study are survey data and program monitoring data. Survey instruments are harmonized across waves and use consistent recall periods. For the endline survey, selected modules and questions that were no longer required were removed (or added, e.g., the reduced Coping Strategy Index -rCSI), and remaining items were adjusted where necessary to address inconsistencies arising from the absence of a midline data collection. The majority of survey questions are adapted from established instruments and prior surveys (e.g., the DigComp framework, the Agricultural Census, and the rCSI). The questionnaire was piloted during the baseline survey.

2.3.2 Data Collection

Data collection was initially planned in three waves: baseline, midline, and endline. Due to budget constraints, the midline data collection was not implemented, and the study therefore relies on baseline and endline surveys only. Baseline data were collected between late January and early March 2025. Endline data collection is scheduled for January–February 2026 and has been slightly advanced to accommodate respondents' availability and leave them free during the Ramadan period.

2.3.3 Data Processing

Data are collected using the SurveyCTO platform, which provides encrypted data transmission and storage. During data collection, high-frequency data quality checks are conducted on key variables. Collected data are processed using Stata 18. All

[4] The digital farm record (*Buku Tani*) allows MFs to generate reports that can be used as financial documentation for loan applications.

analyses are performed on anonymized datasets stored locally. Final datasets are stored on secure servers.

3 Empirical Analysis

3.1 Intent to Treat effects

We estimate the intent-to-treat (ITT) effects of the LenteraDigiEx training intervention on a set of predefined outcomes using an ordinary least squares (OLS) analysis of covariance (ANCOVA) specification. Given the randomized assignment of treatment at the subdistrict level, the ANCOVA estimator is used to improve statistical power by controlling for baseline levels of the outcome variables. Our main estimating equation is:

$$Y_{ist} = \beta_0 + \beta_1 Light_s + \beta_2 Intensive_s + \beta_3 Y_{is0} + X'is0\gamma + \epsilon st, \quad t = 1, \tag{1}$$

where Y_{ist} denotes the post-treatment outcome for respondent i in subdistrict s at time t. $Light_s$ is an indicator equal to one if subdistrict s was assigned to the light training treatment, and $Intensive_s$ is an indicator equal to one if subdistrict s was assigned to the intensive training treatment. The coefficients β_1 and β_2 capture the ITT effects of the respective treatment arms relative to the control group.

Y_{is0} denotes the baseline value of the outcome variable. The vector X_{is0} includes pre-treatment covariates such as age, gender, education, marital status, and household size, with additional covariates included depending on respondent type. For extension workers (EWs), these covariates include work experience and government contract type. For millennial farmers (MFs), they include asset-based wealth proxies, additional household income, entrepreneurial ability, and membership in a farmer organization.

In alternative specifications, we include regency fixed effects and subsector fixed effects (crop cultivation, livestock, fishery, forestry, and agro-processing). For outcomes without baseline measurements, we estimate cross-sectional models comparing treatment and control groups at $t = 1$, as well as pooled specifications where appropriate.

We also plan to include Local Average Treatment Effect (LATE) estimates to look into the actual participation effects. In this estimation we will use the randomized treatment assignment indicator as an instrument for actual participation.

3.2 Heterogeneous Effects

We also examine treatment effect heterogeneity along pre-specified dimensions, including gender, agricultural subsector (for MFs), age, education, and baseline digital literacy. Heterogeneous effects are estimated by interacting each treatment indicator with the corresponding characteristic of interest. For example, heterogeneity by gender is assessed by interacting the treatment indicators with a female indicator variable. In the main analysis, heterogeneous treatment effects will be reported for at least one of these dimensions.

3.3 Standard Error Adjustments

Standard errors are clustered at the subdistrict level to account for the unit of randomization. Statistical inference is based on these clustered standard errors, with statistical significance evaluated at a two-sided significance level of $\alpha = 0.05$. For completeness, we also report results at alternative conventional significance levels ($\alpha = 0.10$ and $\alpha = 0.01$).

Given the large number of outcome variables considered in the study, there is an increased risk of false rejections of the null hypothesis due to multiple hypothesis testing. To address this concern, we adjust for multiple testing by controlling the false discovery rate following Benjamini and Hochberg (1995). Specifically, we compute sharpened q-values within pre-defined families of outcomes.

References

Aker, J. (2011, 11). Dial "a" for agriculture: A review of information and communication technologies for agricultural extension in developing countries - working paper 269. *Agricultural Economics 42*, 631 – 647.

Baumüller, H. (2017, 01). *Towards Smart Farming? Mobile Technology Trends and their Potential for Developing Country Agriculture*, pp. 191–201.

Beach, R. H., C. Milliken, K. Franzen, and D. Lapidus (2025). Meta-analysis of the impacts of digital information interventions on agricultural development. *Global Food Security 45*, 100866.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: series B 57*(1), 289–300.

Enwelu, I., S. Enwereuzor, A. Asadu, H. Nwalieji, and B. Ugwuoke (2017, 07). Access and use of information and communication technologies by extension workers in anambra state agricultural development programme, nigeria. *Journal of Agricultural Extension 21*, 152–162.

Kaliky, R., H. A. Juniarti, and S. Kusumawati (2023). Factors influencing intention and behaviour of agricultural extension staff concerting post-covid-19 digital technical guidance: A study in yogyakarta, indonesia. *Research in Social Sciences and Technology 8*(4), 147–171.

Ngadi, N., A. Zaelany, A. Latifa, D. Harfina, D. Asiati, B. Setiawan, F. Fitranita, T. Triyono, and Z. Rajagukguk (2023, 01). Challenge of agriculture development in indonesia: Rural youth mobility and aging workers in agriculture sector. *Sustainability 15*, 922.

Sugihono, C., S. S. Hariadi, and S. P. Wastutiningsih (2024). Digital literacy among agricultural extension agents in north maluku, indonesia. *Asian Journal of Agricultural Extension, Economics and Sociology 42*(6), 417–429.

Suswadi and N. C. Irawan (2023, December). Digital Transformation of Agricultural Extension in Indonesia: A Comprehensive Analysis. *AGRIS on-line Papers in Economics and Informatics 15*(4).

Tata, J. and P. Mcnamara (2017, 10). Impact of ict on agricultural extension services delivery: evidence from the catholic relief services smart skills and farmbook project in kenya. *The Journal of Agricultural Education and Extension 24*, 1–22.