



## **Harnessing Digital Bundles: Yield and Efficiency Gains from Internet-Based Farming Information and E-Commerce among Chili Farmers in Indonesia<sup>1</sup>**

Kadir Ruslan

BPS-Statistics Indonesia, Jakarta, Indonesia – [kadirsst@bps.go.id](mailto:kadirsst@bps.go.id)

Weni Lidya Sukma

BPS-Statistics Indonesia, Jakarta, Indonesia – [wenilidya@bps.go.id](mailto:wenilidya@bps.go.id)

### **Abstract**

This study investigates the impact of internet use on chili farmers' performance and the gender yield gap in Indonesia, utilizing nationally representative data from the 2018 Horticulture Cost Structure Survey. Employing OLS, Propensity Score Matching (PSM), Multidimensional Distance Matching (MDM), and Inverse Probability Weighted Regression (IPWRA), the results show that internet use significantly increases yield (0.78–1.39 tons/ha) and improves technical efficiency. Importantly, the study finds that digital engagement is not homogeneous; farmers who "bundle" informational and transactional uses—combining cultivation data with both input and output e-commerce—achieved the highest gains. Conversely, a clear gender gap in adoption exists, explaining a meaningful share of the yield differential. These findings suggest that policy must shift from providing basic digital access to fostering integrated digital ecosystems. To maximize productivity and bridge gender inequalities, interventions should prioritize "all-in-one" platforms and targeted training that move farmers beyond simple information-seeking toward multifaceted, high-intensity digital integration.

**Keywords:** internet adoption, chilli farming, technical efficiency, gender yield gap, Indonesia

### **1. Introduction**

Chili is a strategic agricultural commodity in Indonesia, driven by widespread culinary preferences for spicy food (1, 2, 3, 4). Despite seasonal supply fluctuations, persistent high demand throughout the year makes chili production essential for food price stability (5). Indeed, chili price volatility remains a significant contributor to food inflation in the country (1).

Improving chili farmers' performance, particularly in terms of yield and technical efficiency, is therefore both an economic and a food security priority. Digital technologies, especially internet use, offer promising tools to reduce inefficiencies by enhancing farmers' access to information, markets, and innovations (6, 7, 8, 9, 10, 11). Studies from countries such as China, Vietnam, and Pakistan suggest that internet adoption can improve technical efficiency through better resource allocation, risk management, and the uptake of modern inputs (7, 12, 13).

However, evidence on the impact of internet use for agricultural purposes among Indonesian chili farmers is limited. Most existing studies focus on general agriculture or staple crops like rice (14, 15, 16) and often fail to distinguish between general internet access and its specific use in agriculture. This distinction is critical, as evidenced in China, where not all farmers with

---

<sup>1</sup> The text and materials are free from any copyright violations.

internet access utilize it for agricultural purposes (12). Moreover, emerging evidence suggests that different types and intensities of digital use may generate heterogeneous productivity effects (17, 18). Failure to capture this—as well as the specific digital skills required for different tasks—may therefore attenuate estimated effects and weaken policy relevance (19, 20). The existing studies also often lacked robust methodological approaches to address selection bias, as in Kadir and Prasetyo and Ruslan (14, 15).

Internet adoption is influenced by socio-demographic factors such as age, education, and household composition (21, 22). Gender also plays a significant role: female farmers in Indonesia are less likely to adopt digital tools and face greater barriers to digital inclusion (23, 24, 25, 26). These disparities have implications for productivity. For instance, Ruslan and Prasetyo (27) found that differential internet adoption helps explain the gender yield gap among paddy farmers, a pattern that may extend to chili farming. However, to the best of our knowledge, no study on this area for chili farming has been conducted.

This study addresses these gaps by using nationally representative data from the 2018 Horticulture Cost Structure Survey (SOUH 2018) and robust econometric methods to control for self-selection. It seeks to answer two key questions: (1) Does internet use for agriculture improve chili farmers' yield and technical efficiency in Indonesia? (2) Is there a gender gap in chili productivity, and to what extent do gendered differences in internet use drive it?

This study contributes to the literature in three ways. First, it provides the first nationally representative evidence on internet use in chili farming in Indonesia. Second, it moves beyond binary adoption by examining heterogeneous impacts across levels of digital engagement. Third, it integrates productivity analysis with gender decomposition to assess how digital divides contribute to yield inequality. The rest of the paper proceeds as follows. Section 2 describes the study's methodology. Section 3 reports and discusses the empirical results, and Section 4 offers concluding remarks, policy recommendations, and the limitations of this study.

## 2. Methodology

### 2.1 Data

This study draws on microdata from the nationally representative SOUH 2018 by Statistics Indonesia, the most recent dataset detailing input use, cost structure, farm characteristics, and farmer demographics. The survey enables analysis of farm-level performance through yield and technical efficiency metrics. The analysis includes 15,320 chili farmers, with sampling weights applied to ensure national representativeness. Among the total samples, 1,111 samples are internet users and 14,209 samples are non-users. Variable definitions are provided in Table 1.

Table 1. Description of Research Variables

No.	Variables	Description	Mean	S.D.
1.	Yield	The last harvest in the form of fresh chili, expressed in tons/ha.	4.8278	6.1694
2.	Harvested area	The area of the last harvest in hectares	0.1953	0.2131
3.	Gender	Male (ref. = 0), female (= 1).	0.9727	0.2964
4.	Age	Age of farmers in years.	47.0018	11.6777
5.	Education Level	The level of education completed: no education or only completing primary school (ref. = 0), completing high school (junior or senior) (=1), and completing a university degree (= 2).	1.4208	0.5440
6.	Position in the household	Non-household head (ref. = 0); household head (= 1).	0.9350	0.2466

No.	Variables	Description	Mean	S.D.
7.	Number of farmers in the household	At least two (ref. = 0) and more than two (= 1).	1.2639	0.8266
8.	Farm scale	Less than 2 ha/small scale (ref. category = 0), more than 2 ha (= 1).	0.0005	0.0228
9.	Land ownership	Non-owner (ref. = 0), owner (= 1).	0.9422	0.2333
10.	Fertilizer use	Fertilizer used in the last harvest (ton/ha).	1.7038	3.7428
11.	Seed use	Seed used in the last harvest (gr/ha).	247.8440	1,536.3904
12.	Pesticide use	Pesticide used in the last harvest (kg/ha).	2.7886	6.4944
13.	Labour use	Labour used in the last harvest (man-day/ha).	186.9624	278.4316
14.	Mulch plastic	Mulch plastic used in the last harvest (m/ha).	634.694	1,419.639
15.	Partnership or contract	Not having a contract (ref. = 0) and having a contract (= 1).	0.0133	0.1144
16.	Extension services	Not having access (ref. = 0) and having access (= 1).	0.1513	0.3584
17.	Access to credit	Not having access (ref.= 0) and having access (= 1).	0.0686	0.2528
18.	Certified seed	Users (ref. = 0), non-users (= 1).	0.4316	0.4953
19.	Mobile phone use	Users (ref. = 0), non-users (= 1).	0.3993	0.4898
20.	Internet use	Non-user (ref. = 0) and user (= 1).	0.0725	0.2594
21.	Chile variety	Small (ref. = 0), big (= 1).	0.4785	0.4996
22.	Regional Dummy	Outside Jawa (ref. = 0), Jawa (= 1).	0.2005	0.4004

Note: S.D. refers to standard deviation; ref. refers to the reference category.

## 2.2 Empirical Strategy

In estimating the technical efficiency (TE) of chili crop cultivation, we applied a well-known stochastic analysis frontier (SAF) (28, 29), specifically, the trans-log production function is

$$\ln(Y_i) = \beta_0 + \sum_{k=1}^5 \beta_k \ln(X_{ik}) + \frac{1}{2} \sum_{k=1}^5 \beta_{kk} \ln(X_{ik})^2 + \frac{1}{2} \sum_{j=1}^5 \sum_{k=1}^5 \beta_{jk} \ln(X_{ij}) \ln(X_{ik}) + v_i - u_i \quad (1)$$

where the subscript  $i$  denotes the  $i$ th chili farmer, the subscripts  $k$  and  $j$  both are factor indices,  $Y_i$  denotes the yield of the  $i$ th chili farmer,  $X_{ik}$  or  $X_{ij}$  denotes the input of factor  $j$  or  $k$  of farmer  $i$ ,  $v_i$  denotes the random error term, and  $u_i$  denotes the non-negative random errors related to inefficiency. We use five input factors, including labour, fertilizer, seedings, pesticide, and mulch plastic. Land input and its interaction were not included as explanatory variables since all input factors are measured at the per unit of land. The TE was calculated using

$$TE = e^{-u_i} \quad (2)$$

To analyse the impact of internet adoption on chili farmers' yield and technical efficiency, we apply multiple methods for assessment to obtain a robust result. First, following (30), we estimate the OLS benchmark model as follows

$$P_i = \gamma_0 + \gamma_1 D_i + \sum_j \gamma_j X_{ij} + \varepsilon_i \quad (3)$$

where  $P_i$  denotes the performance of the  $i$ th chili farmer measured in yield and technical efficiency,  $D_i \in \{0,1\}$  denotes a dummy variable indicating whether the chili farmer  $i$  adopt internet for agricultural purposes or not,  $X_{ij}$  denotes the  $j$  control variables (farmer socio-demographic characteristic and farm characteristics) of the  $i$ th chili farmer,  $\gamma_0$  denotes the intercept of the model,  $\gamma_1$  the regression coefficient of the internet adoption, and  $\varepsilon_i$  denotes the error term.

Equation (3) was estimated separately for yield and technical efficiency. Since yield is log-transformed, the coefficient  $\gamma_1$  reflects the average percentage difference in yield between internet adopters and non-adopters, holding other variables constant. Following Halvorsen and Palmquist (31), the impact estimation is corrected using  $(e^{\gamma_1} - 1) \times 100\%$ . Farmers are defined as adopters if they self-reported using the internet in the last 12 months for at least one agricultural purpose: accessing price or cultivation information, purchasing inputs, or selling produced agricultural products.

To capture the heterogeneous impacts of different levels of digital engagement, we extend our analysis beyond a binary adoption indicator. Following the logic that the benefits of digital technology may scale with the sophistication or frequency of use, we specify an ordinal treatment variable,  $D_i^{ord}$ . This variable categorizes internet use into discrete levels (e.g., 0 = No use, 1 = searching information or single e-commerce activity, 2 = searching information combined with purchasing inputs or selling agricultural products, and 3 = multiple use-information and e-commerce). We estimate this relationship using an OLS model:

$$P_i = \theta_0 + \sum_{k=1}^3 \theta_k \mathbb{I}(D_i^{ord} = k) + \sum_j \gamma_j X_{ij} + \varepsilon_i \quad (4)$$

where  $\mathbb{I}(D_i^{ord} = k)$  is an indicator (dummy) variable for each level  $k$  of the ordinal internet adoption scale, with "No use" ( $k=0$ ) serving as the reference category and  $\theta_k$  represents the incremental impact of reaching adoption level  $k$  relative to non-adopters.

While OLS estimates provide initial evidence of a positive association between internet use and farm performance, they may suffer from methodological limitations, most notably selection bias. Internet adopters may differ systematically from non-adopters in observable traits (e.g., education or access to extension services), which can confound results. To address these issues, we also apply multiple econometric methods aimed at mitigating selection bias due to observed confounders. Following Rosenbaum & Rubin (32), Propensity Score Matching (PSM) was employed to pair adopters with non-adopters based on similar observable characteristics, using propensity scores estimated via a probit model (Equation 5).

$$p(X_i) = pr(D_i = 1 | X_i) \quad (5)$$

where  $X_i$  denotes observable characteristics influencing the decision of a farmer to use the internet for agricultural purposes or not. To have robust pairing outcomes, we employed two matching algorithms, which are Nearest-neighbour matching ( $n = 5$ ) combined with Caliper or radius matching (threshold = 0.005) and Kernel matching. Based on the matching results, the impact of internet adoption on yield and technical efficiency was estimated by computing the Average Treatment Effect (ATE) and Average Treatment Effects on the Treated (ATT). ATE shows the average productivity and efficiency gains the chili farmers could achieve if all of them used the internet for agricultural purposes, while ATT measures the gains in the user group with or without internet adoption. The two are estimated using

$$ATE = E(Y_{1i} | p(X_i)) - E(Y_{0i} | p(X_i)) \quad (6)$$

$$ATT = E(Y_{1i} | D_i = 1, p(X_i)) - E(Y_{0i} | D_i = 1, p(X_i)) \quad (7)$$

where  $Y_{1i}$  is the yield or technical efficiency score of chili farmers for internet users and vice versa,  $Y_{0i}$  is the yield or technical efficiency score of chili farmers for non-users. To address

limitations of propensity score–based methods, we also employed Multivariate Distance Matching (MDM) using the Mahalanobis Kernel algorithm. MDM mitigates issues such as model misspecification in propensity score estimation, loss of information from collapsing high-dimensional covariates into a single score, and poor covariate balance. Unlike PSM, MDM directly matches on multiple covariates, offering greater flexibility and transparency, especially when treatment assignment is influenced by complex or nonlinear relationships (33, 34, 35).

Given the substantial imbalance in the sample, where internet adopters are notably fewer than non-adopters, we further applied Inverse Probability Weighted Regression (IPWR) to estimate ATE and ATT. Unlike PSM and MDM, which may exclude unmatched units and thus reduce statistical power, IPWR retains the full sample by reweighting observations based on treatment probabilities (36, 37, 38). It integrates the inverse probability weights into a regression framework, which is estimated in Equation (3), to account for the non-random selection into treatment while modeling the outcome, provided that either the matching model or the outcome regression model is correctly specified (double robustness property) (39, 40).

A statistically significant coefficient on gender in Equation (3) indicates yield differences between male and female farmers, controlling for other factors. To further explore the role of internet adoption in explaining this gap, we employ the Blinder-Oaxaca decomposition method (41, 42) to separate the observed yield gap into two components: an explained component, attributable to differences in observable characteristics, including internet use, and an unexplained component, reflecting differences in returns to those characteristics or unobserved factors. The procedure involves estimating Equation (3) separately for male and female farmers and decomposing the mean yield difference accordingly.

We designate the two groups of farmers based on gender, labelling female farmers as group  $F$  and male farmers as group  $M$ . The mean yield difference to be decomposed corresponds to the difference in average yields between these two groups, denoted as  $\bar{Y}_M$  for male farmers and  $\bar{Y}_F$  for female farmers. Following Jann (43), we applied Equation (8) for the decomposition.

$$\Delta\bar{Y} = (\bar{\mathbf{X}}_M - \bar{\mathbf{X}}_F)' \hat{\beta}_M + \bar{\mathbf{X}}_F' (\hat{\beta}_M - \hat{\beta}_F) \quad (8)$$

The first term on the right-hand side represents the explained component of the yield gap, and the second term captures the unexplained component, commonly interpreted as structural factors or potential discrimination (44, 43). Following Neumark (45), we applied the pool coefficient vector for each group.

### 3. Results and Discussion

#### 3.1 Internet Adoption Impact on Yield and Technical Efficiency

The empirical findings strongly support the theoretical and empirical arguments outlined in the Introduction that digital technologies enhance agricultural productivity through improved information access, market participation, and input optimization (6, 7, 9, 10, 11). The higher mean yield and technical efficiency observed among internet adopters are consistent with evidence from Vietnam and China, where internet use improved farm productivity by facilitating better production planning and market integration (12, 10).

The average yield difference of 1.79 tons/ha between users and non-users in Table 2 aligns with prior findings that digital connectivity reduces information asymmetries and enhances farmers' responsiveness to price signals and agronomic advice (9, 11). In the Indonesian context—where chili price volatility significantly contributes to food inflation (1, 5)—the ability to access real-time market information likely enables farmers to better time harvests and marketing decisions, thereby stabilizing income and improving output performance.

Table 2. Mean of Yield and Technical Efficiency Score by Internet Use

Variables	User	Non-user	Total	Difference
Yield (ton/ha)	6.5601	4.7709	4.8907	1.7892***
Technical efficiency	0.6433	0.6002	0.6031	0.0431***

Note: Number of observations is 15,320 farmers; sampling weights were applied; \*\*\*p<0.01, \*\*p<0.05, \*p<0.1

The technical efficiency gains observed in the stochastic frontier analysis are also consistent with studies showing that internet use improves allocative efficiency and resource management (7, 13). The relatively low national average efficiency score (0.603) suggests substantial scope for improvement, echoing Batiese's (28) argument that inefficiency in agriculture often stems from imperfect information and suboptimal input combinations. Internet adoption appears to partially close this information gap.

The estimation results of the SPF model presented in Table 3 indicate that several input variables significantly influence production efficiency. Fertilizer and mulch show positive and statistically significant first-order effects, underscoring their critical roles in enhancing productivity. In contrast, seed, labour, and pesticide exhibit negative and significant coefficients, suggesting potential overuse or inefficiencies in their application. Importantly, the nonlinearities and significant interaction terms in the production function reinforce the idea that digital technologies do not operate in isolation but interact with complementary inputs such as fertilizer, mulch, and labor. This complements the findings of Zheng et al. (11), who demonstrate that internet use promotes improved technology adoption, thereby enhancing productivity indirectly through input complementarities.

Table 3. SPF Model Estimation

Variables	SPF Models
Ln(Seedlings)	-0.0583*** (0.0172)
Ln(Fertilizers)	0.0417** (0.0174)
Ln(Pesticides)	-0.0191** (0.0083)
Ln(Mulch)	0.1487*** (0.0129)
Ln(Labours)	-0.2104*** (0.0719)
Input interactions	yes (sig.)
Quadratic terms	yes (sig.)
Constant	1.3765*** (0.1646)
$\sigma(u)$	-0.5868** (0.2646)
$\sigma(v)$	-0.7957*** (0.1430)
Number of Observations	15,320

Note: Half normal distribution and sampling weights were applied; \*\*\*p<0.01, \*\*p<0.05, \*p<0.1

To examine distributional differences, a kernel density analysis of log-transformed yield and technical efficiency was conducted for chili farmers who use the internet and those who do not. Figure 1 shows a clear rightward shift in the yield distribution among internet users, indicating higher average yields compared to non-users. The distribution for users is also more concentrated, suggesting more consistent performance. In contrast, non-users display a broader

distribution with a longer left tail, reflecting greater variability and more frequent low-yield outcomes—potentially linked to limited access to timely information, weaker market signals, and suboptimal input use.

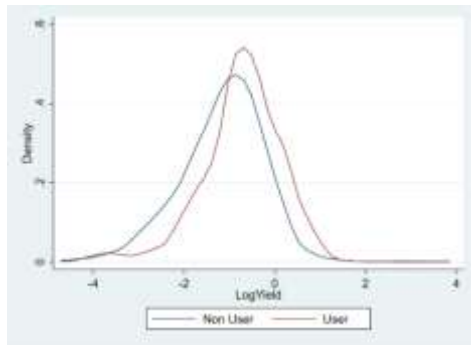


Figure 1. Kernel Density of Chili Yield Distribution by Internet Adoption

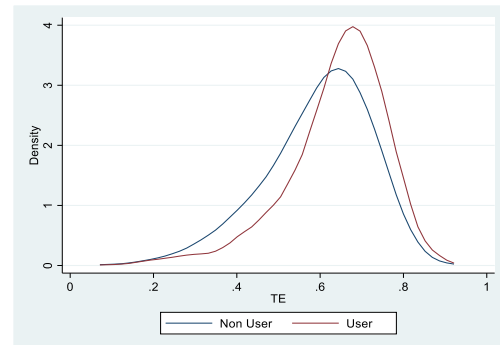


Figure 2. Kernel Density of Technical Efficiency by Internet Adoption

Figure 2 presents a similar pattern for technical efficiency. The distribution for users is shifted to the right, with a more concentrated peak, indicating higher and more consistent efficiency levels. Non-users exhibit a wider spread and a longer left tail, suggesting more frequent low-efficiency outcomes. Overall, both the mean comparisons and graphical evidence provide strong preliminary support for a positive association between internet adoption and higher, more stable yields and technical efficiency.

Table 4. OLS Estimates of the Impact of Internet Use on Farmers' Performance

Variables	Yield		Technical efficiency	
	Base model	Full model	Base model	Full model
Internet use (user)	0.4190*** (0.1116)	0.2410*** (0.0789)	0.0431*** (0.0109)	0.0299*** (0.0106)
Gender (Female)	-	-0.1329** (0.0631)	-	-0.0065 (0.0103)
Inputs	-	yes (sig.)	-	-
Other control variables	-	yes	-	yes
Constant	1.1488*** (0.0575)	-0.5963** (0.2449)	0.6002*** (0.0068)	0.5051*** (0.0224)
Number of observations	15,320	15,320	15,320	15,320
Adjusted R-squared	0.0130	0.2305	0.0079	0.0545

Note: Robust standard errors in parentheses; \*\*\*p<0.01, \*\*p<0.05, \*p<0.1; sampling weights were applied.

The regression results in Table 4 align with the descriptive evidence from the kernel density distributions of log-transformed yield and technical efficiency by internet use. In the baseline model, internet adoption is associated with a statistically significant 52.04 percent  $[(e^{0.419} - 1) \times 100]$  increase in yield. After controlling for socioeconomic and farm characteristics, the effect remains robust at 27.25 percent. Internet use also has a positive and statistically significant impact on technical efficiency, increasing scores by up to 0.03 points after adjustments. These findings support the view that internet access enhances farmers' ability to obtain agronomic information, access input markets, and improve decision-making, leading to more efficient resource allocation.

The OLS results further confirm a persistent gender yield gap. Even after controlling for observable characteristics, female farmers have approximately 12.44 percent lower yields than male farmers. Significant effects of complementary inputs—such as fertilizer, labor, certified seeds, and education—indicate that internet use operates synergistically with these factors.

Farmers with digital access are better positioned to adopt improved practices and optimize input use, consistent with the observed tighter yield distribution among internet users.

The dose–response relationship identified in Table 5 further extends the literature by demonstrating that productivity gains scale with the intensity and diversity of digital engagement. This finding supports recent evidence that heterogeneous forms of digital adoption generate differentiated productivity outcomes (17, 18). Farmers who combine informational and transactional uses (“digital bundling”) achieve the highest gains, suggesting complementarity effects between knowledge acquisition and market participation.

This result also resonates with the digital literacy argument advanced by Xu et al. (20), who emphasize that the effectiveness of digital technologies depends not merely on access but on the ability to integrate multiple digital functions into farm management. Farmers engaging only in basic information searches may face information overload or limited absorptive capacity, as discussed by Eppler and Mengis (19), potentially explaining why the largest productivity gains accrue to high-intensity users capable of translating information into commercial transactions.

Table 5. OLS Estimates of the Impact of Internet Use Intensity on Farmers’ Performance

Variables	Yield	Technical efficiency
Intensity of Internet Use		
Information or a single e-commerce activity only	0.162** (0.078)	0.022* (0.012)
Combined with purchasing or selling	0.208*** (0.077)	0.018 (0.011)
Multiple use	0.416*** (0.085)	0.079*** (0.015)
Other control variables	yes	yes
Constant	-0.792*** (0.229)	0.424*** (0.035)
Number of observations	15,320	15,320
Adjusted R-squared	0.288	0.106

Note: Robust standard errors in parentheses; \*\*\*p<0.01, \*\*p<0.05, \*p<0.1; sampling weights were applied.

The robustness checks using PSM, MDM, and IPWRA in Table 6 confirm that the impact of adoption on production outcomes is not merely driven by observable selection bias. The consistency of ATE and ATT estimates across matching algorithms supports the argument of Rosenbaum and Rubin (32) that credible causal inference in observational settings requires balancing covariates between treated and control groups. The slightly larger effects observed under MDM and IPWRA are consistent with concerns raised by King and Nielsen (34) regarding potential inefficiencies in standard PSM.

Overall, the findings reinforce cross-country evidence that internet adoption improves agricultural productivity (6, 12, 13), while contributing novel evidence from Indonesia’s chili farming, which has been underexplored relative to rice (14, 15).

Table 6. Impact of Internet Adoption on Farmers’ Performance

Method	Yield (ton/ha)		Technical Efficiency	
	ATE	ATT	ATE	ATT
PSM				
- Nearest-Neighbour (5) with Calliper (0.005)	0.8398***	1.1274***	0.0056	0.0150***
- Kernel	0.7815***	1.0519***	0.0055	0.0119**
MDM	1.3921***	1.2749***	0.0176***	0.0167***

Note: \*\*\*p<0.01, \*\*p<0.05. Sampling weights were used in estimation. Kernel matching for PSM and MDM applied bootstrapping (2,000 iterations). Robust standard errors were applied when employing IPWRA.

### 3.2 Internet adoption and gender yield gap

As shown in Table 7, the results also confirm the presence of a statistically significant gender yield gap among chili farmers, consistent with broader agricultural evidence highlighting gender-based productivity disparities (44). The Blinder–Oaxaca decomposition reveals that more than one-third of the gap is attributable to differences in observable characteristics, including internet adoption.

The positive and significant contribution of internet use in the explained component supports prior findings that differential technology adoption contributes to gender-based productivity gaps (27). Given documented disparities in digital literacy and access among women farmers (23, 24, 25), the lower adoption rates observed among female chili farmers likely constrain their ability to benefit from market information, extension services, and e-commerce opportunities.

This pattern aligns with evidence from Sub-Saharan Africa and South Asia showing that women face structural constraints—including limited asset ownership, restricted mobility, and weaker institutional linkages—that reduce both adoption rates and returns to digital technologies (24, 26). The fact that 63.26 percent of the yield gap remains unexplained suggests that structural and institutional factors continue to shape differential returns, consistent with decomposition studies in other agricultural contexts (44).

Importantly, the finding that internet adoption explains 5.52 percent of the total yield gap indicates that digital inclusion policies could function as gender-equalizing mechanisms. However, consistent with the literature on digital divides (20, 24), access alone is insufficient; interventions must address digital skills, control over productive resources, and intra-household decision-making dynamics. Thus, the results position digital agriculture not only as a productivity-enhancing innovation but also as a potential instrument for inclusive rural development—provided that policy design explicitly integrates gender-sensitive strategies.

Table 7. Blinder-Oaxaca Decomposition of Gender Yield Differential

	Pooled coefficient	
<b>1. Gender differential</b>		
Mean male log(chili yield)	1.1989*** (0.0190)	
Mean female log(chili yield)	0.9888*** (0.0551)	
Log(chili yield) difference	0.2101*** (0.0583)	
<b>2. Aggregate decomposition</b>		
Total	0.0772*** (0.0235)	0.1329** (0.0546)
Share of the total gap (%)	36.74	63.26
<b>3. Detailed decomposition</b>		
Internet	0.0116*** (0.0037)	0.0009 (0.0055)
Other characteristics	yes	yes
Constant	-	-0.6104 (0.7174)

Note: Robust standard errors in parentheses (clustered at subdistrict level); \*\*\*p<0.01, \*\*p<0.05, \*p<0.1; sampling weights were applied.

#### 4. Conclusion, policy implications, and limitations

This study makes a substantive contribution to the digital agriculture literature by providing nationally representative evidence on the productivity effects of internet adoption among chili farmers in Indonesia. Using rigorous identification strategies—including PSM, MDM, and IPWRA—the analysis shows that internet use for agricultural purposes increases yield by 0.78–1.39 tons per hectare and improves technical efficiency by up to three percentage points. These gains are statistically significant and economically meaningful, particularly for a high-value and price-sensitive commodity such as chili.

Beyond documenting an average treatment effect, the study shows that digital engagement is heterogeneous. Productivity gains are driven not only by access but by the depth and diversity of digital use. Farmers who bundle informational functions (e.g., cultivation and price information) with transactional activities (input procurement and output marketing) achieve the largest improvements in yield and efficiency. This complementarity suggests that digital technologies function as integrated production systems rather than isolated tools. In other words, the returns to internet adoption depend on how comprehensively it is embedded into farm management decisions.

Kernel density analysis further indicates that internet adoption reduces yield dispersion, implying not only higher average productivity but also more stable performance. This is particularly relevant in chili production, where volatility and risk exposure are substantial. Improved access to timely information and markets likely enhances decision-making, reduces input misallocation, and mitigates downside production risk.

The study also identifies persistent gender disparities. Female farmers are significantly less likely to adopt internet-based tools and record lower yields than male farmers. Blinder–Oaxaca decomposition shows that a meaningful share of the gender yield gap is explained by differences in digital access, education, input use, and institutional support, although a substantial unexplained component remains, pointing to deeper structural constraints. These findings underscore that digital transformation in agriculture is not gender-neutral.

Policy implications are clear. Expanding rural internet infrastructure and strengthening digital literacy—especially for women—can enhance productivity and contribute to food price stability. However, policies should move beyond a narrow focus on connectivity toward improving the quality and intensity of digital engagement. Rather than promoting fragmented applications, governments and development agencies should support interoperable “all-in-one” platforms that combine agronomic information with e-commerce for input and output markets. Extension services should also move beyond basic digital awareness toward advanced integration training, particularly for farmers at the information-only stage.

Despite its contributions, this study has limitations. Reliance on self-reported input data may introduce measurement error, and the cross-sectional nature of the 2018 SOUH dataset limits analysis of long-term dynamics. Although matching and weighting techniques mitigate observable selection bias, unobserved factors—such as entrepreneurial ability and digital literacy—may still influence results. Finally, measuring internet adoption as a binary indicator masks differences in usage intensity, suggesting the need for panel data and more granular measures in future research.

#### 5. References

1. Lukas A, Kairupan AN, Hendriadi A, Arianto A, Manalu LP, Sumarno L, et al. Fresh chili agribusiness: opportunities and problems in Indonesia. In: Özçatalbaş O, editor. *Agricultural Economics and Agri-Food Business*. London: IntechOpen; 2024.
2. Muflikh YN, Kiloes AM. Insight into the buying behaviour of consumers for chilli in Indonesia: households and food businesses in selected cities. *Appl Food Res*. 2024;4(1):100413

3. Sumaryanto, Susilowati SH, Ashari, Sudaryanto T, Ariani M, Permana D, et al. Determining drivers of chili farming participation: insights from Upper Citarum Watershed, Indonesia. *Int J Des Nat Ecodyn*. 2024;19(2):581–590.
4. Sundari MT, Darsono, Sutisno J, Antriandarti E. Analysis of chili supply in Indonesia. *IOP Conf Ser Earth Environ Sci*. 2023;1200(1):012036.
5. Wardhono A, Nasir MA, Qori'ah CG, Indrawati Y. Perfecting policies of chili agribusiness to support food security: evidence from Indonesia districts. *IOP Conf Ser Earth Environ Sci*. 2021;759(1):012048.
6. Ahmad B, Zhao Z, Jile X, Gultaj H, Khan N, Yunxian Y. Exploring the influence of internet technology adoption on the technical efficiency of food production: insight from wheat farmers. *Front Sustain Food Syst*. 2024;8.
7. Duguma AL, Bai X. How the internet of things technology improves agricultural efficiency. *Artif Intell Rev*. 2024;58(2):63.
8. Kusumaningsih N. The technical efficiency of rice farming and mobile phone usage: a stochastic frontier analysis. *Food Res*. 2023;7(1):93–103.
9. Lio M, Liu M. ICT and agricultural productivity: evidence from cross-country data. *Agric Econ*. 2006;34(3):221–228.
10. Zheng H, Ma W, Wang F, Li G. Does internet use improve technical efficiency of banana production in China? Evidence from a selectivity-corrected analysis. *Food Policy*. 2021;102:102044.
11. Zheng Y, Zhu T, Jia W. Does internet use promote the adoption of agricultural technology? Evidence from 1,449 farm households in 14 Chinese provinces. *J Integr Agric*. 2022;21(1):282–292.
12. Kaila H, Tarp F. Can the internet improve agricultural production? Evidence from Viet Nam. *Agric Econ*. 2019;50(6):675–691.
13. Li X, Xiong H, Hao J, Li G. Impacts of internet access and use on grain productivity: evidence from Central China. *Humanit Soc Sci Commun*. 2024;11(1):90.
14. Kadir K, Prasetyo OR. Determinan demografi penggunaan internet petani padi di Indonesia dan kaitannya dengan produktivitas. *Semin Nas Off Stat*. 2021;2021(1):166–175.
15. Ruslan K. Food and horticulture crop productivity in Indonesia. Jakarta: 2021.
16. Sukma WL, Puspitasari MD, Prasetyoputra P. Internet use in agriculture and farm earnings: an analysis of the Indonesia Labor Force Survey. *Asian Dev Rev*. 2025;42(1):233–257.
17. Khan N, Ray RL, Kassem HS, Zhang S. Mobile internet technology adoption for sustainable agriculture: evidence from wheat farmers. *Appl Sci*. 2022;12(10):4902.
18. Manzoor F, Wei L, Siraj M, Lu X, Qiyang G. Digital agriculture technology adoption in low and middle-income countries—a review of contemporary literature. *Front Sustain Food Syst*. 2025;9:1621851.
19. Eppler MJ, Mengis J. The concept of information overload: a review of literature from organization science, accounting, marketing, MIS, and related disciplines. *Inf Soc*. 2004;20(5):325–344.
20. Xu A, Radzi NAM, Liu Y, Sieng LW. The role of digital literacy in agricultural technology adoption and efficiency: a systematic literature review. *Sustainability*. 2026;18(2):1138
21. Mishra AK, Williams RP. Internet access and use by farm households. Presented at: AAEA Annual Meeting; 2006.
22. Sekabira H, Bonabana J, Asingwire N. Determinants for adoption of information and communications technology (ICT)-based market information services by smallholder farmers and traders in Mayuge District, Uganda. *J Dev Agric Econ*. 2014;4(14):404–415.
23. BPS. Hasil Survei Pertanian Antar Sensus (SUTAS) 2018. Jakarta: BPS; 2018.

24. Karakara AAW, Daudu AK, Koomson I. Digital literacy and financial inclusion among smallholder farmers: exploring gender dichotomy in Cote d'Ivoire and Nigeria. In: Koomson I, Villano RA, editors. *Financial Inclusion and Sustainable Rural Development*. Singapore: Palgrave Macmillan; 2024. p. 201–227.
25. WEF. *Agritech for Women Farmers: A Business Case for Inclusive Growth*. 2024.
26. Yadav H, Paliwal M, Chatradhi N. Entrepreneurship development of rural women through digital inclusion: examining the contributions of public programs. In: Rajagopal, Behl R, editors. *Inclusive Businesses in Developing Economies*. Cham: Palgrave Macmillan; 2022. p. 287–309.
27. Ruslan K, Prasetyo OR. Development of paddy yield gap between Java and outside Java: does it have a contribution to paddy yield improvement from 2018 to 2021? *Proc Int Conf Data Sci Off Stat*. 2023;2023(1):767–779.
28. Batiase GE. Frontier production functions and technical efficiency: a survey of empirical applications in agricultural economics. *Agric Econ*. 1992;7(3–4):185–208.
29. Førsund FR, Lovell CAK, Schmidt P. A survey of frontier production functions and of their relationship to efficiency measurement. *J Econom*. 1980;13(1):5–25.
30. Deng X, Peng J, Wan C. The impact of internet use on land productivity: evidence from China Land Economy Survey. *Land*. 2024;13(2):262.
31. Halvorsen R, Palmquist R. The interpretation of dummy variables in semilogarithmic equations. *Am Econ Rev*. 1980;70(3):474–475.
32. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41–55.
33. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399–424.
34. King G, Nielsen R. Why propensity scores should not be used for matching. *Polit Anal*. 2019;27(4):435–454.
35. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010;25(1).
36. Hirano K, Imbens GW, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*. 2003;71(4):1161–1189.
37. Wooldridge JM. Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Port Econ J*. 2002;1(2):117–139.
38. Zhang JL, Rubin DB. Estimation of causal effects via principal stratification when some outcomes are truncated by 'death'. *J Educ Behav Stat*. 2003;28(4):353–368.
39. Imbens GW, Wooldridge JM. Recent developments in the econometrics of program evaluation. *J Econ Lit*. 2009;47(1):5–86.
40. Wooldridge JM. Inverse probability weighted estimation for general missing data problems. *J Econom*. 2007;141(2):1281–1301.
41. Blinder AS. Wage discrimination: reduced form and structural estimates. *J Hum Resour*. 1973;8(4):436–455.
42. Oaxaca R. Male-female wage differentials in urban labor markets. *Int Econ Rev*. 1973;14(3):693.
43. Jann B. The Blinder–Oaxaca decomposition for linear regression models. *Stata J*. 2008;8(4):453–479.
44. Fukase E, Kim YS, Chiarella C. Exploring the source of the agricultural productivity gap: evidence from Sri Lanka. 2022; No. 10025.
45. Neumark D. Employers' discriminatory behavior and the estimation of wage discrimination. *J Hum Resour*. 1988;23(3):279.