CELEBES Agricultural 7(1): 13 - 37 August (2026)

Research Article

Extension Contact and Rice Productivity of Non-Swamp Farmer Groups in Mamasa Regency, West Sulawesi

Kontak Penyuluhan dan Produktivitas Padi Kelompok Tani Lahan Non-Rawa di Kabupaten Mamasa, Sulawesi Barat

Muhammad Ridwan[1,2]*, Andi Syastiawan[1], Andi Emelda[1]

[1] Faculty of Vocational Studies, Hasanuddin University, Makassar, 90245, South Sulawesi, Indonesia

[2] Applied Agriculture Research Group, Makassar, 90245, South Sulawesi, Indonesia

*email: muhammadridwanrani@unhas.ac.id

Recived: August, 2026
Accepted: October, 2026
Published: October, 2026

p-ISSN: 2723-7974
e-ISSN: 2723-7966
doi: 10.52045/jca.v7i1.1505

Website:

Abstract: Rice production in West Sulawesi fell by 16.83% in 2023, driven mainly by a 14.71% contraction in harvested area, while implied provincial productivity declined only slightly (from about 5.10 to 4.97 t ha^{-1}). Extension services are expected to raise productivity, yet their association with yields in heterogeneous highland rice systems is poorly documented. This study examined the cross-sectional association between extension contact and rice productivity among non-swamp farmer groups in Mamasa Regency, defined as lowland rice (*sawah*) groups outside tidal and inland swamps, cultivating rainfed, semi-technically irrigated, or technically irrigated fields. Field-survey data from 179 farmer groups in all 17 sub-districts were analyzed. Extension exposure was measured by three categorical items (meeting regularity, frequency of contact with extension workers, and mentoring activeness) and productivity as the chairperson-reported mean yield across cultivated seasons. Because a reflective latent specification showed inadequate reliability (ρ_C=0.629; AVE=0.411) and a redundant outcome indicator, the primary analysis used regression on the observed indicators with robust and bootstrap confidence intervals, with PLS-SEM as a secondary check. Routine contact with extension workers (88.3% of groups) was associated with 1.59 t ha^{-1} lower productivity than rare or occasional contact (95% CI −2.65 to −0.54; p=0.003), whereas meeting regularity and mentoring activeness showed no association. The association was stable across alternative outcome definitions, influential-case exclusions, sub-district fixed effects, and PLS-SEM specifications but attenuated and became non-significant after exploratory adjustment for agroecological, technology-access, and group covariates (−0.88; 95% CI −2.05 to 0.29). Out-of-sample predictive accuracy was low ($Q^2_{predict}$=0.070 against a naive mean benchmark). The negative association should not be read as a harmful effect of extension; need-based targeting, poorly matched extension content, missing complementary inputs, and unmeasured agroecological differences remain competing hypotheses. Documenting targeting criteria, collecting baseline productivity and agroecological data, and evaluating context-specific extension packages over time are recommended.

Keywords: Agricultural extension; extension contact; farmer groups; lowland rice; Mamasa Regency

Abstrak: Produksi padi di Sulawesi Barat turun 16,83% pada tahun 2023, terutama akibat penyusutan luas panen sebesar 14,71%, sedangkan produktivitas provinsi yang tersirat hanya sedikit menurun (dari sekitar 5,10 menjadi 4,97 t ha^{-1}). Penyuluhan diharapkan meningkatkan produktivitas, tetapi hubungannya dengan hasil pada sistem padi dataran tinggi yang heterogen belum banyak didokumentasikan. Penelitian ini menganalisis hubungan lintas-seksi (*cross-sectional*) antara kontak penyuluhan dan produktivitas padi pada kelompok tani lahan non-rawa di Kabupaten Mamasa, yaitu kelompok padi sawah di luar lahan rawa pasang surut dan lebak yang mengusahakan sawah tadah hujan, setengah teknis, atau irigasi teknis. Data survei lapangan dari 179 kelompok tani di seluruh 17 kecamatan dianalisis. Keterpaparan penyuluhan diukur dengan tiga butir kategorik (keteraturan pertemuan, frekuensi kontak dengan penyuluh, dan keaktifan pendampingan), sedangkan produktivitas diukur sebagai rata-rata hasil lintas musim tanam yang dilaporkan ketua kelompok. Karena spesifikasi laten reflektif menunjukkan reliabilitas yang tidak

Citation:
Ridwan M, Syastiawan A, Emelda A. 2026. Extension Contact and Rice Productivity of Non-Swamp Farmer Groups in Mamasa Regency, West Sulawesi. *CELEBES Agricultural*. 7(1): 13-37. doi: 10.52045/jca.v7i1.1505

https://ojs.untika.ac.id/index.php/faperta

memadai (Q_c=0,629; AVE=0,411) dan indikator hasil yang redundan, analisis utama menggunakan regresi pada indikator teramati dengan selang kepercayaan robust dan bootstrap, dengan PLS-SEM sebagai pemeriksaan sekunder. Kontak rutin dengan penyuluh (88,3% kelompok) berasosiasi dengan produktivitas 1,59 t ha^{-1} lebih rendah dibandingkan kontak jarang atau kadang-kadang (95% CI −2,65 hingga −0,54; p=0,003), sedangkan keteraturan pertemuan dan keaktifan pendampingan tidak berasosiasi. Asosiasi ini stabil pada definisi hasil alternatif, pengeluaran observasi berpengaruh, efek tetap kecamatan, dan spesifikasi PLS-SEM, tetapi melemah dan tidak signifikan setelah penyesuaian eksploratif terhadap kovariat agroekologi, akses teknologi, dan karakteristik kelompok (−0,88; 95% CI −2,05 hingga 0,29). Akurasi prediksi di luar sampel rendah ($Q^2_{predict}$=0,070 terhadap tolok ukur rata-rata). Asosiasi negatif ini tidak boleh dimaknai sebagai dampak merugikan penyuluhan; penargetan berbasis kebutuhan, materi penyuluhan yang kurang sesuai, keterbatasan input pelengkap, dan perbedaan agroekologi yang tidak terukur tetap merupakan hipotesis yang bersaing. Disarankan agar kriteria penargetan didokumentasikan, data dasar produktivitas dan agroekologi dikumpulkan, dan paket penyuluhan spesifik lokasi dievaluasi dari waktu ke waktu.

Kata kunci: Kabupaten Mamasa; kelompok tani; kontak penyuluhan; padi sawah; penyuluhan pertanian

INTRODUCTION

Rice is a strategic food commodity for Indonesia's food security; however, production gains have not been sustained in several producing regions. In West Sulawesi Province, the harvested rice area in 2023 was estimated at 59.12 thousand hectares, 14.71% lower than in 2022, and production fell by 16.83% to 294.03 thousand tonnes of milled dry grain (*gabah kering giling*, GKG) (BPS, 2023). Because harvested area, production, and productivity are distinct quantities, these figures do not indicate a collapse in productivity; they imply that the provincial decline in production was driven mainly by the contraction in harvested area, while mean productivity fell only slightly, from approximately 5.10 to 4.97 t GKG ha^{-1} (about −2.5%). Provincial production trends therefore provide context rather than direct evidence on the determinants of yield in individual districts. Mamasa Regency is a highland district in which rice is grown in small lowland fields under heterogeneous topographic, drainage, and water-supply conditions; in the survey analyzed here, 41.3% of farmer groups reported undulating or steep field topography and 50.3% reported insufficient or no water supply.

Agricultural extension is expected to raise productivity by transferring knowledge and technology to farmers and encouraging the adoption of improved practices (Hernalius et al., 2018; Latif et al., 2022). Extension workers act as change agents linking sources of innovation with prospective adopters, consistent with Rogers' (2003) diffusion of innovations theory, in which adoption depends on the attributes of the innovation, the communication channels, time, and the social system; more recent syntheses also integrate farmers' behavioral factors into adoption models (Hidayati et al., 2025). Empirical studies in Indonesia have generally reported favorable associations. Hernalius et al. (2018) linked agricultural extension to the productivity of irrigated rice fields in Sukabumi Regency, Latif et al. (2022) reported a relationship between the role of extension workers and the productivity of rice farmers, and structural-equation studies in

Southeast and South Sulawesi identified farming capital, farmer education, family resources, and farmer engagement as drivers of extension effectiveness (Jamil et al., 2023; Salam et al., 2024).

However, extension effects are not uniformly positive. Reviews of extension in developing countries show that measured impacts vary widely and are often smaller than expected (Birkhaeuser et al., 1991; Anderson & Feder, 2004; Takahashi et al., 2020). Whether extension contact translates into higher yields depends on the relevance of the advice to local agroecological constraints, farmers' access to complementary inputs such as improved seeds, fertilizers, machinery, and water, and market, infrastructure, and institutional conditions (Anderson & Feder, 2004; Takahashi et al., 2020). In Indonesia, an evaluation of the national farmer field school programme found no significant impact on the yields or pesticide use of participants and their neighbors (Feder et al., 2004). Extension resources are also frequently allocated in response to local conditions. When programmes are placed deliberately where particular problems exist, simple comparisons between more and less exposed units confound the programme with the reasons for its placement and can produce null or even negative associations even when the programme is beneficial (Khandker et al., 2010). Therefore, there is a theoretical basis for heterogeneous, null, or negative cross-sectional associations between extension exposure and yields, as well as positive ones.

The agroecological setting is central to this question. Indonesia's land optimization programme distinguishes between swamp land (tidal and inland swamps) and non-swamp land, and its phase assesses the physical, technical, and socio-economic conditions and problems of under-utilized non-swamp rice land, such as water shortage, waterlogging, and erosion risk (BBPP Ketindan, 2025). In highland districts such as Mamasa, non-swamp rice land comprises rainfed fields and fields served by semi-technical or technical irrigation. Rainfed and poorly drained lowland rice systems are constrained by rainfall variability, water shortage, and soil fertility limitations (Haefele et al., 2014), and the adoption of yield-increasing technology is conditioned by access to seed, fertilizer, and machinery (Takahashi et al., 2020). Under such heterogeneous conditions, blanket recommendations developed for well-irrigated fields may be poorly matched to local constraints, and the relationship between extension contact and productivity may differ between water regimes.

The Indonesian studies cited above were conducted mainly in irrigated or lowland settings and used farmer-level surveys, with outcomes ranging from perceived extension effectiveness to yields, which limits their transferability to highland non-swamp systems. The present study contributes district-wide evidence from farmer groups in all 17 sub-districts of Mamasa Regency, a relatively under-documented non-swamp rice context, and examines a counter-intuitive pattern in the relationship between extension contact and rice productivity. The survey measured extension exposure the regularity of group extension meetings, the frequency of contact with extension workers, and the activeness of mentoring rather than the content, quality, or learning outcomes of extension training; the construct is therefore referred to as *extension contact* throughout. Particular attention is given to transparent measurement assessment, sensitivity to

alternative specifications, and the agroecological information needed to interpret the association responsibly.

Two competing expectations follow from this literature.

H1: Consistent with diffusion-based reasoning, farmer groups with more regular extension contact report higher rice productivity levels.

H2: Consistent with program-placement and context-constraint reasoning, the cross-sectional association between extension contact and productivity is null or negative because contact may be directed toward more constrained groups or advice may be poorly matched to local constraints.

Because the data are cross-sectional, this study tests associations and does not estimate causal effects. The objectives were to (i) describe extension contact, rice productivity, and agroecological conditions among non-swamp farmer groups in the Mamasa Regency; (ii) estimate the association between extension contact and productivity and assess its sensitivity to measurement and model specification; and (iii) explore whether the association differs across land and water categories.

MATERIALS AND METHODS

Study Area and Research Design

This study was conducted in the Mamasa Regency, West Sulawesi Province, which comprises 17 subdistricts. A quantitative cross-sectional design was used. All estimates describe associations between extension contact and productivity observed at one point in time and are not interpreted as causal effects. Neither PLS-SEM nor bootstrapping resolves endogeneity, reverse causality, or omitted-variable bias; potential confounders are addressed only to the extent that they are explicitly included in a model, which is done here in an exploratory adjusted model.

Data Source and Survey Procedures

Data were collected through a structured field survey of rice farmer groups in Mamasa Regency, in which the authors were directly involved in field data collection. The survey was conducted in June–July 2026. In each farmer group was interviewed face to face and reported group-level information on membership, cultivated area, land and water conditions, input access, extension activities, and rice production. Only the survey fields relevant to the present research questions were extracted; their item labels, response options, and coding are documented in Table 1. Because the instrument was designed for programme mapping rather than as a psychometric scale for this study, its measurement properties were assessed explicitly.

All variables describe the collective attributes of a farmer group (for example, the regularity of group meetings or the typical yield of the group's fields) rather than the attitudes of individual members. The farmer group was therefore the unit of analysis, and its chairperson served as the key informant, because chairpersons coordinate group meetings and planting plans and are the

usual contact point for extension workers, making them the members most likely to be knowledgeable about these attributes (Kumar et al., 1993). No aggregation of multiple respondents was performed; each group-level value is the chairperson's single report. Single-informant reporting may introduce recall error and common-method bias (Podsakoff et al., 2003), and this limitation is considered in the interpretation.

Extension-contact items and seasonal productivity were recorded in the same interview. The productivity items refer to the group's first (MT1) and second (MT2) planting seasons of the most recent cropping year, whereas the extension items describe the group's usual pattern of extension activity without a separate reference period. The temporal ordering of extension contact and reported harvests, therefore, cannot be established: contact may have preceded, accompanied, or responded to the reported yields.

Operational Definition of Non-Swamp Land and Agroecological Information

Non-swamp land was operationally defined by the dominant land typology recorded for each farmer group in the survey, and groups whose dominant typology was swamp land were excluded. The retained groups cultivated lowland rice fields (*sawah*) classified as rainfed (*sawah tadah hujan*), semi-technical irrigation (*sawah setengah teknis),* or technical irrigation (*sawah irigasi teknis*), except for three groups recorded only as non-swamp without further specification. No upland (dryland) rice groups were present in the data. The survey also recorded field topography, drainage condition, adequacy and continuity of water supply, exposure to natural hazards, cropping index, dominant variety, main water source, and access to fertilizer and agricultural machinery (Table 1). These variables were used to describe the agroecological context, explore heterogeneity, and in an exploratory adjusted model. Elevation, slope gradient, soil fertility, rainfall, input application rates, and pest and disease pressure were not recorded and could not be analyzed.

Population, Sample Screening and Missing Data

The survey produced 244 farmer group records. The screening process was conducted in three steps. First, records were retained if the respondent had chairperson status and a complete name, and one record with swamp land typology was excluded, yielding 195 valid farmer groups located in all 17 sub-districts of the regency. Second, productivity values were checked for plausibility: 12 records with values outside 0.1–12 t ha^{-1} (200–2,000, probably entered in kg ha^{-1}, or 0.01–0.09) were treated as invalid, and three records had no productivity data. Third, one record lacked a contact frequency item. The final analytical sample comprised 179 farmer groups (91.8% of valid groups) from all 17 sub-districts in the province.

To assess whether complete-case analysis was plausibly unbiased, the retained and excluded groups were compared on all available characteristics using standardized mean differences (SMD), Fisher's exact tests for binary variables, and Mann–Whitney tests for ordinal and continuous variables (Austin, 2009). The earlier version of this analysis required valid values for both planting seasons and retained 161 groups from 15 sub-districts. This rule systematically excluded 18 groups that planted only one rice season per year (cropping index IP100), for which

second-season productivity was not missing but not applicable, and removed all groups in Nosu and Pana Sub-districts. The revised outcome definition described below retains these groups, and the original 161-group sample was analyzed as a sensitivity check.

Sample adequacy was evaluated using the inverse square root method (Kock & Hadaya, 2018), $n_{min} > (2.486/|\beta_{min}|)^2$, which corresponds to a one-tailed test at α=0.05 with 80% power. Assuming a minimum relevant standardized association of $|\beta_{min}|$=0.20, at least 155 groups were required (197 for a two-tailed test). With n=179, the minimum detectable standardized association at two-tailed α=0.05 and 80% power was 0.209. For the regression models, a noncentral F power calculation showed that n=179 provides 80% power at α=0.05 to detect f^2=0.044 for a single coefficient and f^2=0.069 for the joint test of the four extension-contact terms.

Measurement of Variables

Extension contact was measured by three categorical items (Table 1): the regularity of group extension meetings (PP1: incidental, seasonal, or routine), the frequency of contact with field extension workers (PP2: rare, occasional, or routine), and the activeness of mentoring by extension workers (PP3: fairly active or very active). Response options were verbal categories selected by the chairperson; they did not specify a numeric frequency or reference period (for example, meetings per month), and no inactive category was recorded for PP3. Because the categories are few and unevenly distributed, the items were analyzed as observed categorical indicators: PP1 as two dummy variables (incidental and routine, with seasonal meetings as reference), PP2 as routine versus rare or occasional contact (only one group reported rare contact), and PP3 as very active versus fairly active mentoring. These items capture exposure to extension, not the content, quality, timing, or learning outcomes of extension training.

Chairpersons reported the typical group-level yield (t ha^{-1}) for the first (MT1) and, where cultivated, the second (MT2) planting season. Yields were reported as harvested unhulled grain (*gabah kering panen*, GKP), were not converted to a standard moisture content, and were not based on crop-cut (*ubinan*) measurements or demonstration plots; therefore, they are recall-based estimates. The mean-productivity field was the arithmetic mean of MT1 and MT2 for every two-season group and was not used as a separate indicator. The outcome was defined as a single annual mean productivity across the seasons actually cultivated: the mean of MT1 and MT2 for groups planting two or more seasons and MT1 for single-season groups. Because the survey recorded one cultivated area per group rather than season-specific areas, equal seasonal weights were used, which equals an area-weighted annual mean when the same fields are planted in each season. Season-specific yields were analyzed separately in the sensitivity analyses.

Table 1. Survey items, response options and coding of the variables analysed

Variable (code)	Survey field label (Indonesian)	Response options / unit	Coding in analysis
Extension contact			
Regularity of group extension meetings (PP1)	Pertemuan rutin	Insidental (incidental); Musiman (seasonal); Rutin (routine)	Dummies: incidental, routine (ref. seasonal); score 1–3 in PLS-SEM and sensitivity
Frequency of contact with extension workers (PP2)	Penyuluhan	Jarang (rare); Kadang-kadang (occasional); Rutin (routine)	1 = routine, 0 = rare/occasional; score 1–3 in PLS-SEM and sensitivity
Activeness of extension mentoring (PP3)	Pendampingan penyuluh	Cukup aktif (fairly active); Sangat aktif (very active)	1 = very active, 0 = fairly active
Rice productivity			
First-season productivity (MT1)	Produksi MT1	t ha^{-1}, harvested unhulled grain (GKP), chairperson report	Continuous; valid range 0.1–12
Second-season productivity (MT2)	Produksi MT2	t ha^{-1} GKP; not applicable to single-season groups	Continuous; valid range 0.1–12
Annual mean productivity (outcome)	Derived	t ha^{-1} GKP	Mean of MT1 and MT2; MT1 for IP100 groups
Agroecological and contextual variables			
Land typology	Tipologi dominan	Sawah tadah hujan (rainfed); setengah teknis (semi-technical); irigasi teknis (technical)	Dummies (ref. rainfed)
Topography	Topografi	Curam (steep); bergelombang (undulating); landai (gently sloping); datar (flat)	Score 1–4
Drainage condition	Kondisi drainase	Buruk (poor); sedang (moderate); baik (good)	Score 1–3
Water adequacy	Ketersediaan air	Tidak tersedia (none); kurang (insufficient); cukup (adequate)	1 = adequate
Water continuity	Ketersediaan air (kontinuitas)	Sering kering (often dry); musiman (seasonal); sepanjang tahun (year-round)	1 = year-round
Hazard exposure	Kebencanaan	Aman (safe); rawan banjir, kekeringan, longsor or erosi (prone)	1 = hazard-prone
Cropping index	Indeks pertanaman	IP100; IP200; IP300	1 = two or more seasons
Dominant variety	Varietas dominan	Lokal; Inpari; Ciherang/Mekongga; hibrida; combinations	1 = any improved variety
Fertilizer access	Akses pupuk	Sangat sulit; sulit; cukup; mudah (very difficult to easy)	Score 1–4
Machinery access	Alsintan	Tidak ada (none); kurang (insufficient); cukup (adequate)	Score 1–3
Chairperson and group characteristics	Umur; pendidikan; jumlah anggota; luas garapan	Years; five education levels; persons; ha	Age; education score 1–5; active members; ha per member

Note: GKP = *gabah kering panen*; IP = cropping index; ref. = reference category.

Data Analysis

The analysis proceeded in six steps. **(1) Descriptive analysis.** The distributions of each extension-contact item and productivity were summarized by item category, sub-district, land typology, water availability, and productivity quartile, using Kruskal–Wallis, Mann–Whitney, chi-square, and Cochran–Armitage trend tests and Spearman correlations with bootstrap 95% confidence intervals (CIs).

(2) Measurement assessment of the original latent specification. The earlier PLS-SEM model (reflective Extension Training measured by PP1–PP3 and reflective Rice Productivity measured by mean, MT1 and MT2 productivity; n=161) was re-estimated to document its measurement properties: outer loadings with bootstrap CIs, Cronbach's α, ρ_A, composite reliability (ρ_c), average variance extracted (AVE), the Fornell–Larcker criterion, and the HTMT (Fornell & Larcker, 1981; Henseler et al., 2015; Hair et al., 2019).

(3) Primary association models. The predictor side consists of three coarse categorical items with weak inter-item correlations, and the outcome is a single observed quantity. In this situation, a latent-variable framework adds no measurement advantage, while obscuring which item carries the association and requiring a construct specification that the data do not support. The primary analysis therefore used ordinary least squares (OLS) regression of annual mean productivity on the observed categorical indicators: Model M1 included all three items, and Model M2 included contact frequency (PP2) alone. Standard errors were heteroscedasticity-consistent (HC3; Long & Ervin, 2000), and 95% CIs were also obtained by nonparametric case-resampling bootstrap with 10,000 resamples (percentile and bias-corrected and accelerated [BCa] intervals; Efron & Tibshirani, 1993). PLS-SEM was retained as a secondary specification check, with extension contact specified as a composite (Mode B) construct, as a reflective (Mode A) construct, or by PP2 as a single indicator, each predicting single-item annual productivity (Diamantopoulos & Winklhofer, 2001; Jarvis et al., 2003).

(4) Sensitivity analysis. Model M1 was re-estimated after excluding influential observations (Cook's distance), with log-transformed productivity, season-specific outcomes (MT1 and MT2 separately, and a season-stacked model with standard errors clustered by group), in the original 161-group complete-case sample, after recovering eight records with probable kg ha^{-1} entry errors (values of 200–2,000 divided by 1,000), with the ordinal items entered as linear scores, with sub-district fixed effects (HC1 standard errors), and with standard errors clustered by sub-district using a wild cluster bootstrap p-value (9,999 replications, Webb weights; Cameron et al., 2008; Webb, 2023). Leave-one-sub-district-out estimates were also calculated.

(5) Exploratory adjustment and heterogeneity. An adjusted model (M3) added land typology, topography, drainage, water adequacy and continuity, hazard exposure, cropping index, improved-variety use, fertilizer and machinery access, and chairperson and group characteristics, and was estimated for groups with complete covariate data (n=161). Covariates were also added in blocks (agroecological, technology access, and group characteristics) to identify which block changed the contact coefficient. Because these covariates are also reported

by the chairperson and some may lie on the pathway between extension and productivity, M3 is descriptive rather than causal. Heterogeneity was explored with models stratified by land typology and water adequacy and continuity, with interaction terms, and with quantile regression at the 25th, 50th, and 75th percentiles (Koenker & Bassett, 1978; 2,000 bootstrap resamples).

(6) Predictive assessment. Out-of-sample performance was evaluated with 10-fold cross-validation repeated 20 times. Root mean squared error (RMSE), mean absolute error (MAE), and Q^2_{predict} (one minus the ratio of the model's squared prediction error to that of the training-fold mean) were compared with a naive mean benchmark. For the original PLS-SEM model, a PLSpredict-type procedure compared PLS predictions of each productivity indicator with those of a linear-model (LM) benchmark (Shmueli et al., 2019). The Q^2 value reported in the earlier version was not retained because it was obtained with a leave-one-out procedure that does not correspond to standard blindfolding or PLSpredict.

All tests were two-tailed at $\alpha=0.05$, and exact p-values are reported ($p<0.001$ where applicable). Computations were performed in Python 3.11.15 with NumPy 2.4.4, SciPy 1.17.1, pandas 3.0.2, and scikit-learn 1.8.0, using purpose-written PLS and regression routines with a fixed random seed (1505). The PLS algorithm used standardized indicators and the path weighting scheme, with a stop criterion of 10^{-10} for the change in outer weights and a maximum of 500 iterations (all full-sample models converged within nine iterations); each bootstrap solution was sign-aligned with the full-sample outer weights at the construct level. The analysis code is available from the corresponding author upon reasonable request.

RESULTS AND DISCUSSION

Sample and Agroecological Context

The 179 farmer groups were distributed across all 17 sub-districts, with the largest numbers in Tanduk Kalua (16.2%), Mamasa (14.0%) and Rantebulahan Timur (13.4%) (Appendix Table A2). Chairpersons were predominantly male (96.6%), with a mean age of 48.4 years, and senior secondary education was the most common attainment (42.5%). Groups had on average 24.3 active members and 0.33 ha of cultivated land per member (Table 2). Groups cultivated lowland rice fields under semi-technical irrigation (43.6%), rainfed conditions (27.9%) or technical irrigation (26.8%), and three groups (1.7%) had an unspecified non-swamp typology. Half of the groups (50.3%) reported insufficient or no water supply, 44.7% reported seasonal or frequently drying water sources, 41.9% reported poor drainage and 41.3% reported undulating or steep topography. Most groups planted two rice seasons per year (IP200, 87.2%) and relied on local varieties (87.2%); most reported easy access to fertilizer (73.2%) but no access to agricultural machinery (68.2%), and one quarter (25.7%) were exposed to flood, drought, landslide or erosion hazards. Rivers were a main water source for 82.7% of groups.

Mean annual productivity was 2.67 t ha^{-1} (SD=1.75; median 2.00; interquartile range 1.05–4.00; range 0.15–10.00), with a right-skewed distribution. This level is well below the implied provincial mean of approximately 4.97 t GKG ha^{-1} (BPS, 2023). The gap may reflect highland conditions and the predominance of local varieties, but recall-based reporting is also likely to contribute (Gourlay et al., 2019): yields were identical in both seasons for 71.4% of two-season groups, suggesting that chairpersons often reported a typical rather than a season-specific yield. Absolute yield levels should therefore be interpreted with caution.

Table 2. Characteristics and agroecological context of the farmer groups (n=179)

Characteristic	Category	n	% or mean (SD)
Chairperson's sex	Male	173	96.6
	Female	6	3.4
Chairperson's age (years)	Mean (SD)	178	48.4 (11.1)
Chairperson's education	Incomplete primary	5	2.8
	Primary	27	15.1
	Junior secondary	34	19.0
	Senior secondary	76	42.5
	Diploma/bachelor	37	20.7
Active members (persons)	Mean (SD)	179	24.3 (5.9)
Group cultivated area (ha)	Mean (SD)	177	7.64 (6.20)
Cultivated area per member (ha)	Mean (SD)	176	0.33 (0.27)
Land typology	Rainfed lowland	50	27.9
	Semi-technical irrigation	78	43.6
	Technical irrigation	48	26.8
	Non-swamp, unspecified	3	1.7
Topography	Flat	61	34.1
	Gently sloping	44	24.6
	Undulating	67	37.4
	Steep	7	3.9
Drainage condition	Good	48	26.8
	Moderate	50	27.9
	Poor	75	41.9
	Not recorded	6	3.4
Water adequacy	Adequate	88	49.2
	Insufficient	88	49.2
	Not available	2	1.1
	Not recorded	1	0.6
Water continuity	Year-round	99	55.3
	Seasonal	79	44.1
	Often dry	1	0.6
Hazard exposure	None reported	133	74.3
	Flood-prone	30	16.8

Characteristic	Category	n	% or mean (SD)
	Drought-prone	9	5.0
	Landslide-prone	6	3.4
	Erosion-prone	1	0.6
Cropping index	IP100 (one season)	19	10.6
	IP200 (two seasons)	156	87.2
	IP300 (three seasons)	4	2.2
Dominant variety	Local only	156	87.2
	Improved (incl. mixed with local)	23	12.8
Fertilizer access	Easy	131	73.2
	Adequate	23	12.8
	Difficult	17	9.5
	Very difficult	5	2.8
	Not recorded	3	1.7
Machinery access	Adequate	6	3.4
	Insufficient	48	26.8
	None	122	68.2
	Not recorded	3	1.7
Annual mean productivity (t ha^{-1})	Mean (SD)	179	2.67 (1.75)
	Median (IQR)		2.00 (1.05–4.00)

Note: IQR = Interquartile Range. Percentages are for the 179 groups.

Distribution of Extension Contact and Productivity

Reported extension contact was high and concentrated in the upper categories (Table 3). Routine contact with extension workers was reported by 88.3% of groups, occasional contact by 11.2%, and rare contact by a single group (0.6%), and 92.7% described mentoring as very active. Meeting regularity varied more, with seasonal meetings the most common (65.4%). The three items were weakly inter-correlated (Spearman ρ=0.09–0.25). Productivity did not differ by meeting regularity (Kruskal–Wallis p=0.156) or mentoring activeness (p=0.327) but differed by contact frequency (p<0.001): groups with routine contact reported a mean of 2.48 t ha^{-1} (median 2.00), compared with 4.06 t ha^{-1} (median 4.50) among the 21 groups with rare or occasional contact (Hodges–Lehmann difference −1.52 t ha^{-1}; 95% CI −2.50 to −0.59). Only contact frequency was correlated with productivity (ρ=−0.27; 95% CI −0.40 to −0.13).

The association was not uniform across the productivity distribution. The share of groups with routine contact was 95.6%, 93.3%, and 95.5% in the first three productivity quartiles but 68.9% in the highest quartile (≥4.0 t ha^{-1}) (χ^2=21.94, p<0.001; Cochran–Armitage trend p<0.001). Therefore, the negative association arises mainly because groups with less frequent contact were over-represented among the highest-yielding groups, rather than from a gradual decline across the distribution. Mean productivity differed among subdistricts (Kruskal–Wallis p=0.005), and rare or occasional contact was concentrated in a few subdistricts (six groups each in Buntu

Malangka and Mamasa, two each in Aralle, Tabulahan, and Tanduk Kalua, and one each in Balla, Nosu, and Sesena Padang; Appendix Table A2). However, at the subdistrict level, the share of groups with routine contact was not correlated with mean productivity (ρ=0.05, p=0.854; 14 subdistricts with at least four groups), so the group-level association is not simply a between-subdistrict pattern. Productivity differed little among land typologies (rainfed 2.64, semi-technical 2.55, and technical irrigation 2.81 t ha^{-1}; p=0.805).

Table 3. Distribution of extension-contact indicators and annual mean productivity by category (n=179)

Indicator	Category	n (%)	Productivity, mean (SD)	Median	Test; Spearman ρ (95% CI)
Regularity of extension meetings (PP1)	Incidental	9 (5.0)	3.49 (1.53)	4.00	KW H=3.71; p=0.156
	Seasonal	117 (65.4)	2.60 (1.89)	2.00	ρ=−0.004 (−0.143 to 0.136)
	Routine	53 (29.6)	2.67 (1.43)	2.75	
Contact frequency with extension workers (PP2)	Rare	1 (0.6)	10.00 (–)	10.00	KW H=14.07; p<0.001
	Occasional	20 (11.2)	3.76 (1.47)	4.50	ρ=−0.273 (−0.401 to −0.133)
	Routine	158 (88.3)	2.48 (1.64)	2.00	
Mentoring activeness (PP3)	Fairly active	13 (7.3)	3.03 (1.52)	3.00	KW H=0.96; p=0.327
	Very active	166 (92.7)	2.64 (1.77)	2.00	ρ=−0.073 (−0.204 to 0.063)
PP2, collapsed	Rare or occasional	21 (11.7)	4.06 (1.98)	4.50	Mann–Whitney p<0.001
	Routine	158 (88.3)	2.48 (1.64)	2.00	HL difference −1.52 (−2.50 to −0.59)

Note: Productivity in t ha^{-1} GKP. KW = Kruskal–Wallis test; HL = Hodges–Lehmann median difference (routine minus rare/occasional); CIs from 10,000 bootstrap resamples.

Contact frequency was entangled with land, water, and technology conditions. None of the groups with rare or occasional contact farmed technically irrigated fields, compared with 30.8% of routine-contact groups (p=0.002), and fewer reported adequate water (23.8% vs 52.9%; p=0.019). Groups with less frequent contact had flatter topography (mean score 3.62 vs 2.79; p<0.001) but poorer drainage (1.06 vs 1.94; p<0.001), more often used improved varieties (33.3% vs 10.1%; p=0.008) and reported better access to machinery (1.71 vs 1.29; p<0.001). They did not differ significantly in rainfed share, water continuity, hazard exposure, cropping index, fertilizer access, or chairperson and group characteristics (all p≥0.108).

Measurement Assessment of the Original Latent Specification

Re-estimation of the original reflective model reproduced earlier estimates exactly and confirmed the measurement problems (Table 4). The Extension Training construct was dominated by PP2 (loading 0.973; 95% CI 0.813 to 0.998), whereas the bootstrap CIs for the loadings of PP1 (0.383; −0.013 to 0.644) and PP3 (0.373; −0.090 to 0.736) included zero. Internal consistency and convergent validity were inadequate (α=0.399; ρ_A=0.157; ρ_c=0.629; AVE=0.411). The very high reliability of the productivity construct (α=0.994; ρ_c=0.996; AVE=0.989; loadings

0.991–1.000) was an artifact of deterministic overlap: the mean-productivity indicator equalled the arithmetic mean of the two seasonal indicators in all 161 groups, and the two seasonal values were identical in 71.4% of groups. The Fornell–Larcker criterion (√AVE of X=0.641 > |r|=0.362) and HTMT (0.399; bootstrap 95th percentile 0.620) show only that the two constructs are empirically separable; they do not establish that extension exposure was measured adequately. The latent path from this specification (β=−0.362; 95% CI −0.548 to −0.187) is therefore not interpreted substantively, and the analyses below use the observed indicators and a single productivity outcome.

Table 4. Measurement properties of the original reflective PLS-SEM specification (n=161)

Construct	Indicator	Loading	95% bootstrap CI	p	Construct-level statistics
Extension Training (X)	PP1 meeting regularity	0.383	−0.013 to 0.644	0.022	α=0.399; ρ_A=0.157
	PP2 contact frequency	0.973	0.813 to 0.998	<0.001	ρ_c=0.629; AVE=0.411
	PP3 mentoring activeness	0.373	−0.090 to 0.736	0.080	√AVE=0.641
Rice Productivity (Y)	PR1 mean productivity	1.000	1.000 to 1.000	<0.001	α=0.994; ρ_c=0.996
	PR2 first season	0.992	0.985 to 0.996	<0.001	AVE=0.989
	PR3 second season	0.991	0.984 to 0.996	<0.001	PR1 = (PR2+PR3)/2 in 161/161 groups
Structural path	X → Y	β=−0.362	−0.548 to −0.187	<0.001	R^2=0.131; HTMT=0.399 (95th percentile 0.620)

Source: Re-analysis of the earlier model; 10,000 bootstrap resamples, percentile CIs, construct-level sign alignment. p-values from the bootstrap standard errors.

Associations Between Extension Contact and Productivity

In Model M1, routine contact with extension workers was associated with 1.59 t ha^{-1} lower annual productivity than rare or occasional contact (95% CI −2.65 to −0.54; p=0.003; bootstrap BCa CI −2.72 to −0.66), holding meeting regularity and mentoring activeness constant (Table 5). Neither incidental nor routine meetings (relative to seasonal meetings) nor very active mentoring were associated with productivity (joint Wald test F=0.36, p=0.780). Contact frequency alone (M2) yielded a nearly identical estimate (−1.58 (−2.49 to −0.67); p<0.001), corresponding to a standardized association of r=−0.29. The extension-contact items jointly explained 9.0% of the variance in productivity (adjusted R^2=0.069), and the partial effect size of contact frequency was f^2=0.082. The small share of explained variance is expected because productivity also depends on land area, fertilizer use, variety, rainfall, and pest pressure (Andrias, 2018). Residuals were right-skewed (Shapiro–Wilk W=0.949, p<0.001) without evidence of heteroscedasticity (Breusch–Pagan p=0.383), supporting the robust and bootstrap intervals and the log-scale sensitivity check. Because f^2 for a single term is a deterministic function of the change in R^2, it is reported only as a descriptive effect size and not as independent evidence for the association.

Table 5. Regression models of annual mean productivity (t ha^{-1}) on extension-contact indicators

Term	M1: all indicators	M2: contact frequency only	M3: adjusted (exploratory)	M4: sub-district fixed effects
Intercept	3.95 (2.96 to 4.93) p<0.001	4.06 (3.19 to 4.93) p<0.001	1.27 (−2.62 to 5.15) p=0.520	4.88 (3.83 to 5.94) p<0.001
PP1: Incidental meetings (ref. seasonal)	0.40 (−0.99 to 1.79) p=0.573	–	0.94 (−0.35 to 2.22) p=0.152	0.41 (−1.00 to 1.82) p=0.568
PP1: Routine meetings (ref. seasonal)	0.25 (−0.29 to 0.79) p=0.361	–	0.54 (−0.14 to 1.22) p=0.120	0.32 (−0.30 to 0.93) p=0.313
PP2: routine contact (ref. rare/occasional)	−1.59 (−2.65 to −0.54) p=0.003	−1.58 (−2.49 to −0.67) p<0.001	−0.88 (−2.05 to 0.29) p=0.141	−1.61 (−2.79 to −0.42) p=0.008
PP3: Very active mentoring (ref. fairly active)	0.03 (−0.89 to 0.96) p=0.941	–	−0.38 (−1.46 to 0.71) p=0.496	−0.21 (−1.02 to 0.60) p=0.608
Covariates in M3				
Semi-technical irrigation (ref. rainfed)	–	–	0.08 (−0.60 to 0.75) p=0.824	–
Technical irrigation (ref. rainfed)	–	–	0.77 (−0.14 to 1.67) p=0.097	–
Topography score (1–4)	–	–	0.19 (−0.09 to 0.46) p=0.179	–
Drainage score (1–3)	–	–	−0.02 (−0.45 to 0.42) p=0.945	–
Adequate water supply	–	–	−0.38 (−0.94 to 0.18) p=0.184	–
Year-round water availability	–	–	0.47 (−0.16 to 1.09) p=0.146	–
Hazard-prone (flood/drought/landslide)	–	–	0.56 (−0.07 to 1.19) p=0.083	–
Two or more seasons per year (ref. IP100)	–	–	−0.07 (−1.05 to 0.91) p=0.887	–
Improved variety used	–	–	0.46 (−0.42 to 1.35) p=0.303	–
Fertilizer access score (1–4)	–	–	−0.34 (−0.67 to −0.01) p=0.045	–
Machinery access score (1–3)	–	–	0.73 (0.13 to 1.32) p=0.018	–
Chairperson age (years)	–	–	0.02 (−0.01 to 0.06) p=0.189	–
Education score (1–5)	–	–	0.00 (−0.30 to 0.29) p=0.994	–
Active members (n)	–	–	0.02 (−0.02 to 0.06) p=0.315	–
Cultivated area per member (ha)	–	–	−0.34 (−1.38 to 0.70) p=0.517	–
Sub-district fixed effects	No	No	No	Yes (16 dummies)
n	179	179	161	179
R^2 (adjusted R^2)	0.090 (0.069)	0.084 (0.079)	0.246 (0.145)	0.262 (0.169)

Coefficients are unstandardized (t ha^{-1}) with 95% CIs in parentheses; HC3 standard errors for M1–M3 and HC1 for M4. Bootstrap (10,000 resamples) 95% CIs for PP2 routine contact: M1 percentile −2.69 to −0.65, BCa −2.72 to −0.66; M2 percentile −2.48 to −0.70; and M3 percentile −2.01 to 0.18. M3 used groups with complete covariates.

Adding sub-district fixed effects (M4) left the estimate essentially unchanged (−1.61 (−2.79 to −0.42); p=0.008), although sub-district membership explained additional variation (R^2 increased from 0.090 to 0.262; F=2.30, p=0.005). In the exploratory adjusted model (M3; n=161), the coefficient attenuated to −0.88 (−2.05 to 0.29) (p=0.141), 27.8% smaller than the unadjusted estimate in the same subsample (−1.21 (−2.14 to −0.28); p=0.011). Block-wise adjustment showed that agroecological covariates (−1.32; p=0.011) and chairperson and group characteristics (−1.31; p=0.010) left the association essentially unchanged, whereas technology-access covariates reduced it to −0.73 (95% CI −1.80 to 0.34; p=0.180). Individually, machinery access and improved-variety use reduced the coefficient by 30.9% and 23.5%, respectively. In M3, machinery access was positively associated with productivity (0.73 (0.13 to 1.32); p=0.018), whereas the fertilizer-access score was negatively associated (−0.34 (−0.67 to −0.01); p=0.045), a counterintuitive result that probably reflects the coarse self-reported access measure and residual confounding. The variance inflation factors were ≤2.33.

Robustness to Alternative Specifications

The negative association between routine contact and productivity was stable across alternative specifications of both extension exposure and productivity (Table 6). This persisted after excluding the single most influential group (10.0 t ha^{-1} with rare contact; Cook's D=0.164) and after excluding all groups with Cook's D>4/n; on the log scale, routine contact was associated with 44.6% lower productivity (95% CI 21.6% to 60.8%). It was also evident for first-season and second-season yields analyzed separately and in the season-stacked model, in the original 161-group complete-case sample, after recovering probable kg ha^{-1} entry errors (n=187), with the ordinal items entered as linear scores, and with standard errors clustered by sub-district (wild cluster bootstrap p=0.038). Leave-one-sub-district-out estimates ranged from −1.81 to −1.36 (all p≤0.033). All PLS-SEM specifications with single-item annual productivity also yielded negative paths with CIs excluding zero. In the composite (Mode B) specification, only PP2 had an outer weight distinguishable from zero (weight 1.03; 95% CI 0.72 to 1.10; PP1 −0.13, 95% CI −0.59 to 0.35; PP3 −0.04, 95% CI −0.39 to 0.45; VIF ≤1.12), confirming that the latent extension score is in practice a contact-frequency score. The only specification in which the association lost statistical significance was the exploratory covariate-adjusted model.

Table 6. Sensitivity of the association between routine extension contact and productivity to alternative specifications

Specification	n	Estimate (95% CI)	p
A. Regression; coefficient for PP2 routine contact (t ha^{-1})			
Main model (M1)	179	−1.59 (−2.65 to −0.54)	0.003
Excluding the most influential group	178	−1.24 (−2.07 to −0.42)	0.003
Excluding all groups with Cook's D > 4/n	170	−2.17 (−2.57 to −1.77)	<0.001
Log productivity (coefficient on log scale)	179	−0.590 (−0.937 to −0.243)	<0.001
First-season yield (MT1) as outcome	179	−1.58 (−2.63 to −0.52)	0.004

Specification	n	Estimate (95% CI)	p
Second-season yield (MT2) as outcome	161	−1.82 (−2.91 to −0.74)	0.001
Season-stacked MT1 and MT2 (clustered by group; observations)	340	−1.69 (−2.66 to −0.72)	<0.001
Original complete-case sample	161	−1.81 (−2.87 to −0.74)	0.001
Recovering probable kg ha^{-1} entry errors	187	−1.66 (−2.70 to −0.61)	0.002
Ordinal items as linear scores (per PP2 category)	179	−1.79 (−2.96 to −0.62)	0.003
Sub-district fixed effects (M4)	179	−1.61 (−2.79 to −0.42)	0.008
Clustered by sub-district (wild cluster bootstrap p)	179	−1.59 (−2.47 to −0.72)	0.038
Adjusted for covariates (M3, exploratory)	161	−0.88 (−2.05 to 0.29)	0.141
Leave-one-sub-district-out (range of 17 estimates)	150–178	−1.81 to −1.36	≤0.033
B. PLS-SEM; standardized path coefficient β			
Original reflective model (3-indicator productivity)	161	−0.362 (−0.548 to −0.187)	<0.001
Composite Mode B extension contact → annual productivity	179	−0.342 (−0.535 to −0.163)	<0.001
Reflective Mode A extension contact → annual productivity	179	−0.325 (−0.513 to −0.157)	<0.001
PP2 (score 1–3) as single indicator → annual productivity	179	−0.339 (−0.519 to −0.135)	<0.001

Regression estimates use HC3 95% CIs unless stated; clustered models use CR1 standard errors. PLS-SEM CIs are percentile intervals from 10,000 bootstrap resamples. Leave-one-subdistrict-out sample sizes vary with the omitted subdistrict.

Heterogeneity Across Land and Water Categories

The results of exploratory heterogeneity analyses are summarized in Table 7. The association was evident among semi-technically irrigated groups (−1.79 (−2.63 to −0.95); p<0.001) and of a similar magnitude but imprecise among rainfed groups (−1.73 (−5.29 to 1.83); only five groups with rare or occasional contact); it could not be estimated among technically irrigated groups, all of which reported routine contact. Estimates were similar for groups with adequate (−1.48) and insufficient (−1.54) water supply, and larger for groups with year-round (−2.10) than seasonal (−0.70) water, but none of the interaction terms was statistically significant (p=0.109–0.973). Quantile regression showed negative differences at the median (−2.50; 95% CI −3.00 to −1.00) and at the 75th percentile (−1.50; −2.00 to −0.50), with a less precise estimate at the 25th percentile (−2.50; −3.50 to 0.30). Because few groups reported less than routine contact, these analyses have limited power and cannot establish whether the association differs across land or water categories.

Table 7. Exploratory heterogeneity in the association between routine extension contact and productivity

Subgroup or model	n (rare/occasional contact)	Estimate (95% CI)	p
A. Stratified models (PP2 routine vs rare/occasional)			
Rainfed lowland	50 (5)	−1.73 (−5.29 to 1.83)	0.335
Semi-technical irrigation	78 (15)	−1.79 (−2.63 to −0.95)	<0.001
Technical irrigation	48 (0)	Not estimable (no variation)	–
Adequate water supply	88 (5)	−1.48 (−3.09 to 0.13)	0.071

Subgroup or model	n (rare/occasional contact)	Estimate (95% CI)	p
Insufficient or no water supply	90 (16)	−1.54 (−2.68 to −0.39)	0.009
Year-round water availability	99 (13)	−2.10 (−3.30 to −0.91)	<0.001
Seasonal or often-dry water	80 (8)	−0.70 (−1.96 to 0.56)	0.275
Local varieties only	156 (14)	−1.66 (−2.97 to −0.34)	0.014
B. Interaction terms			
PP2 routine × semi-technical (vs rainfed)	128	−0.06 (−3.67 to 3.54)	0.973
PP2 routine × adequate water	178	0.05 (−1.91 to 2.02)	0.957
PP2 routine × year-round water	179	−1.41 (−3.13 to 0.32)	0.109
C. Quantile regression (M1 covariates)			
25th percentile	179	−2.50 (−3.50 to 0.30)	0.059
50th percentile	179	−2.50 (−3.00 to −1.00)	<0.001
75th percentile	179	−1.50 (−2.00 to −0.50)	0.003

Stratified and interaction models used HC3 95% CIs; quantile regression CIs were percentile intervals from 2,000 bootstrap resamples. The estimates were unstandardized (t ha^{-1}).

Predictive Assessment

The out-of-sample predictive performance was low. In repeated 10-fold cross-validation, contact frequency alone reduced the RMSE from 1.761 t ha^{-1} (naïve mean benchmark) to 1.698 t ha−1 and the MAE from 1.482 to 1.388 t ha^{-1} ($Q^2_{predict}$=0.070); the full M1 model performed slightly worse (RMSE 1.721; $Q^2_{predict}$=0.044). In the covariate subsample (n=161), adding covariates did not improve the prediction ($Q^2_{predict}$=0.023 for the extension items and 0.010 for M3). For the original PLS-SEM model, $Q^2_{predict}$ for the three productivity indicators was 0.092–0.098, but PLS prediction errors exceeded those of the linear-model benchmark for all three indicators (RMSE 1.707–1.725 vs. 1.701–1.723), indicating low predictive power (Shmueli et al., 2019). Therefore, a positive in-sample Q^2 does not demonstrate generalizable predictive accuracy.

Discussion

The revised analysis shows that the negative relationship reported in the earlier version of this study is best described as a negative cross-sectional association between one component of extension exposure routine contact with extension workers and chairperson-reported rice productivity. It is not a property of a well-measured extension training construct: meeting regularity and mentoring activeness were unrelated to productivity, and the latent score was essentially a contact-frequency score. The association was robust to the measurement and outcome problems identified in the original specification, persisted within sub-districts, and was not driven by any single sub-district, but it was concentrated among the highest-yielding groups and attenuated after adjustment for technology access. The data are therefore inconsistent with H1 and compatible with H2, but they do not show that extension reduces productivity.

Competing explanations for the negative association

At least four explanations are consistent with this pattern, and each should be regarded as a hypothesis rather than a finding. *(a) Need-based targeting.* Extension workers may visit groups with production problems more often, so that low productivity leads to more contact rather than the reverse. The survey did not record targeting rules, baseline productivity, or the allocation of extension workers; therefore, this mechanism cannot be tested. The descriptive evidence is mixed: routine-contact groups were more often located in technical irrigation schemes and more often reported adequate water, but they also had steeper topography, less access to machinery, and less use of improved varieties. Moreover, directing extension toward lower-performing groups may be a desirable policy; what is missing is evidence on whether those groups subsequently improve relative to a credible counterfactual. *(b) Poorly matched extension content.* Advice may not fit highland non-swamp constraints such as poor drainage, sloping fields, and reliance on local varieties, so that frequent contact does not translate into yield gains. Extension content was not recorded, so this possibility cannot be assessed. *(c) Missing complementary inputs.* Groups may be unable to implement recommendations without machinery, improved seeds, or reliable water. The attenuation of the association after adjustment for machinery access and improved-variety use is consistent with the possibility that groups with better access to technology achieve higher yields while having less need for, or less contact with, extension workers because technology access may itself be influenced by extension; this pattern does not identify the direction of influence. *(d) Unmeasured agroecological differences.* Soil fertility, rainfall, elevation, and pest pressure were not recorded and may differ between groups with more and less frequent contact, even within subdistricts. In addition, the association may partly reflect measurement because contact frequency was recorded in coarse verbal categories without a reference period, and yields were recall-based reports from a single informant.

The non-swamp context

The non-swamp context is central to weighing these explanations. In Mamasa, non-swamp rice is lowland *sawah* cultivated under highland conditions: half of the groups reported insufficient water, about two in five reported poor drainage or undulating to steep fields, nearly nine in ten relied on local varieties, and two thirds had no access to machinery. In such rainfed and partially irrigated systems, yields are strongly constrained by water availability, drainage, and soil fertility, and by the technology gap between local and improved varieties (Haefele et al., 2014; Takahashi et al., 2020), and extension contact alone cannot relax these constraints. If contact is more intensive where constraints are most binding, a negative cross-sectional association can arise even when extension raises yields relative to what those groups would otherwise achieve (Khandker et al., 2010). Conversely, the absence of any technically irrigated group with infrequent contact indicates that contact frequency is partly structured by irrigation infrastructure and programme presence, which limits comparisons across land categories. Blanket extension recommendations are therefore unlikely to be appropriate across these heterogeneous systems.

Extension contact versus effective extension

The findings also underline the difference between extension contact and effective extension. The measured items describe how often groups meet or interact with extension

workers, not what is taught, whether practices are demonstrated, whether follow-up occurs, or whether recommendations are adopted. High contact frequency can coexist with low relevance, insufficient technical content, lack of demonstration, weak follow-up, or an inability to adopt recommendations (Anderson & Feder, 2004). The near-universal reporting of routine contact (88.3%) and very active mentoring (92.7%) also suggests ceiling effects or socially desirable reporting by chairpersons, which limits the ability of these items to distinguish more effective from less effective extension. Practical recommendations for extension agencies, therefore, require measures of extension quality and adoption, not contact frequency alone.

Comparison with previous studies

Differences from earlier Indonesian studies are better explained by differences in population, agroecology, intervention content, and outcome measurement than by statistical technique. Hernalius et al. (2018) examined irrigated rice fields in Sukabumi, whereas most groups in the present study farmed semi-technically irrigated or rainfed highland fields. Latif et al. (2022), Jamil et al. (2023), and Salam et al. (2024) examined the role or effectiveness of extension using farmer-level survey data, whereas the present study used group-level contact frequency reported by chairpersons. The outcomes of those studies ranged from extension effectiveness to productivity, whereas the outcome here is a recall-based group yield. The present results therefore complement, rather than contradict, evidence that well-designed extension can be beneficial, and they are consistent with international reviews reporting heterogeneous extension impacts (Birkhaeuser et al., 1991; Takahashi et al., 2020).

Limitations

Several limitations apply. First, the design is cross-sectional, and extension and yield items were recorded in the same interview, so temporal ordering and causal direction cannot be established. Second, all variables were reported by a single key informant; yields were recall-based GKP estimates without moisture standardization or crop-cut verification, which are prone to reporting error (Gourlay et al., 2019); 12 implausible yield records were excluded; and the extension items were coarse categories without a reference period. Third, only 21 groups reported less than routine contact, which limits precision, particularly in subgroup analyses, and makes the estimates sensitive to a small number of high-yielding groups. Fourth, the covariates available for adjustment were coarse, and important determinants such as soil fertility, rainfall, input rates, and pest pressure were not measured. Fifth, the heterogeneity analyses involved multiple comparisons and are exploratory. Finally, the survey instrument was designed for land-optimization mapping rather than for measuring extension quality.

CONCLUSIONS

Among 179 non-swamp farmer groups in all 17 subdistricts of Mamasa Regency, this study identified a negative cross-sectional association between measured extension contact and rice productivity. The association was specific to the frequency of contact with extension workers: groups reporting routine contact reported 1.59 t ha^{-1} lower annual productivity than groups with

rare or occasional contact (95% CI −2.65 to −0.54), whereas meeting regularity and mentoring activeness were not associated with productivity. This result describes an association and does not show that extension reduces productivity.

The weak measurement properties of the original Extension Training construct (ρ_c=0.629; AVE=0.411; two of three loadings with CIs including zero) and the redundancy of its productivity indicators materially limit the interpretation of the latent structural path reported in the earlier version. When extension exposure was represented by observed indicators and productivity by a single annual outcome, the negative association was stable across alternative specifications, influential-case exclusions, season-specific outcomes, and sub-district fixed effects. However, it attenuated by 28% and was no longer statistically significant after exploratory adjustment for agroecological, technology-access, and group covariates, and its out-of-sample predictive value was low. Therefore, the finding is robust to measurement choices but not demonstrably independent of differences in access to machinery and improved varieties.

Accordingly, the results do not support the judgement that current extension targeting is inappropriate. The strongest implication is the need for context-specific extension packages for heterogeneous non-swamp rice systems — differentiated by water regime, drainage, topography, and access to improved seed and machinery — together with evaluation designs capable of separating selection into extension from subsequent productivity outcomes, so that agencies can assess how targeting decisions, the relevance of extension content, and complementary constraints jointly shape outcomes. Agriculture offices and extension agencies can support this by documenting the criteria used to allocate extension contact, collecting baseline productivity and agroecological information for each farmer group, and monitoring extension outcomes over time.

Future research should measure not only land characteristics but also extension quality, content, and timing, adoption of recommended practices, access to complementary inputs, and market and institutional conditions, using crop-cut or otherwise verified yields within panel or quasi-experimental designs. These variables are needed to explain why extension contact does or does not translate into productivity gains in highland non-swamp rice systems.

ACKNOWLEDGMENT

This research was funded by the Institute for Research and Community Service, Hasanuddin University (LPPM-UH), under contract number 03285/UN4.1.7/PT.01.03/2026. The authors thank LPPM Hasanuddin University as the institution responsible for the program, the farmer group chairpersons who participated in the survey, and all parties who assisted in the conduct of this research.

REFERENCES

Anderson JR & Feder G. 2004. Agricultural extension: Good intentions and hard realities. *The World Bank Research Observer*. 19(1):41-60. https://doi.org/10.1093/wbro/lkh013

Andrias. 2018. Pengaruh luas lahan terhadap produksi dan pendapatan usahatani padi sawah (Suatu kasus di Desa Jelat Kecamatan Baregbeg Kabupaten Ciamis). *Jurnal Ilmiah Mahasiswa AGROINFO GALUH*. 5(1):785-793.

Austin PC. 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*. 28(25):3083-3107. https://doi.org/10.1002/sim.3697

[BBPP Ketindan] Balai Besar Pelatihan Pertanian Ketindan. 2025. Melalui SID, optimasi lahan non rawa siap diwujudkan untuk pertanian berkelanjutan. Available at: https://bbppketindan.bppsdmp.pertanian.go.id/blog/post/melalui-sid-optimasi-lahan-non-rawa-siap-diwujudkan-untuk-pertanian-berkelanjutan [accessed 11 September 2026].

Birkhaeuser D, Evenson RE & Feder G. 1991. The economic impact of agricultural extension: A review. *Economic Development and Cultural Change*. 39(3):607-650. https://doi.org/10.1086/451893

[BPS] Badan Pusat Statistik Kabupaten Mamasa. 2023. Pada 2023, luas panen padi di Sulawesi Barat diperkirakan sebesar 59,12 ribu hektare dengan produksi padi sekitar 294,03 ribu ton gabah kering giling (GKG). Available at: https://mamasakab.bps.go.id/id/pressrelease/2023/11/02/684/ [accessed 16 July 2026].

Cameron AC, Gelbach JB & Miller DL. 2008. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*. 90(3):414-427. https://doi.org/10.1162/rest.90.3.414

Diamantopoulos A & Winklhofer HM. 2001. Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*. 38(2):269-277. https://doi.org/10.1509/jmkr.38.2.269.18845

Efron B & Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. New York (US): Chapman & Hall.

Feder G, Murgai R & Quizon JB. 2004. Sending farmers back to school: The impact of farmer field schools in Indonesia. *Review of Agricultural Economics*. 26(1):45-62. https://doi.org/10.1111/j.1467-9353.2003.00161.x

Fornell C & Larcker DF. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*. 18(1):39-50. https://doi.org/10.2307/3151312

Gourlay S, Kilic T & Lobell DB. 2019. A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale-productivity relationship in Uganda.

Journal of Development Economics. 141:102376. https://doi.org/10.1016/j.jdeveco.2019.102376

Haefele SM, Nelson A & Hijmans RJ. 2014. Soil quality and constraints in global rice production. *Geoderma*. 235-236:250-259. https://doi.org/10.1016/j.geoderma.2014.07.019

Hair JF, Risher JJ, Sarstedt M & Ringle CM. 2019. When to use and how to report the results of PLS-SEM. *European Business Review*. 31(1):2-24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler J, Ringle CM & Sarstedt M. 2015. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*. 43(1):115-135. https://doi.org/10.1007/s11747-014-0403-8

Hernalius LA, Sumardjo S & Hamzah H. 2018. Pengaruh penyuluhan pertanian terhadap tingkat produktivitas padi sawah. *Jurnal Sains Komunikasi dan Pengembangan Masyarakat (JSKPM)*. 2(3):279-288. https://doi.org/10.29244/jskpm.2.3.279-288

Hidayati F, Syahni R, Suliansyah I & Tanjung HB. 2025. Model teoritis adopsi inovasi pertanian: Integrasi pendekatan perilaku petani. *Rawa Sains: Jurnal Sains STIPER Amuntai*. 15(1):70-79.

Jamil MH, Tika N, Fudjaja L, Tenriawaru AN, Salam M, Ridwan M, Muslim AI & Chand NV. 2023. Effectiveness of agricultural extension on paddy rice farmer's Baubau City, Southeast Sulawesi, Indonesia. *Sustainability*. 15(4):3773. https://doi.org/10.3390/su15043773

Jarvis CB, MacKenzie SB & Podsakoff PM. 2003. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*. 30(2):199-218. https://doi.org/10.1086/376806

Khandker SR, Koolwal GB & Samad HA. 2010. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington DC (US): The World Bank. https://doi.org/10.1596/978-0-8213-8028-4

Kock N & Hadaya P. 2018. Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*. 28(1):227-261. https://doi.org/10.1111/isj.12131

Koenker R & Bassett G. 1978. Regression quantiles. *Econometrica*. 46(1):33-50. https://doi.org/10.2307/1913643

Kumar N, Stern LW & Anderson JC. 1993. Conducting interorganizational research using key informants. *Academy of Management Journal*. 36(6):1633-1651. https://doi.org/10.5465/256824

Latif A, Ilsan M & Rosada I. 2022. Hubungan peran penyuluh pertanian terhadap produktivitas petani padi. *Wiratani: Jurnal Ilmiah Agribisnis*. 5(1):1-9.

Long JS & Ervin LH. 2000. Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*. 54(3):217-224. https://doi.org/10.1080/00031305.2000.10474549

Podsakoff PM, MacKenzie SB, Lee JY & Podsakoff NP. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*. 88(5):879-903. https://doi.org/10.1037/0021-9010.88.5.879

Rogers EM. 2003. *Diffusion of Innovations*. 5th ed. New York (US): Free Press.

Salam M, Jamil MH, Tenriawaru AN, Kamarulzaman NH, Syam SH, Rahmadanih, Ramadhani A, et al. 2024. The effectiveness of agricultural extension in rice farming in Bantaeng Regency, Indonesia: Employing structural equation modeling in search for the effective ways in educating farmers. *Journal of Agriculture and Food Research*. 18:101487. https://doi.org/10.1016/j.jafr.2024.101487

Shmueli G, Sarstedt M, Hair JF, Cheah JH, Ting H, Vaithilingam S & Ringle CM. 2019. Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*. 53(11):2322-2347. https://doi.org/10.1108/EJM-02-2019-0189

Takahashi K, Muraoka R & Otsuka K. 2020. Technology adoption, impact, and extension in developing countries' agriculture: A review of the recent literature. *Agricultural Economics*. 51(1):31-45. https://doi.org/10.1111/agec.12539

Webb MD. 2023. Reworking wild bootstrap-based inference for clustered errors. *Canadian Journal of Economics*. 56(3):839-858. https://doi.org/10.1111/caje.12661

APPENDIX

Table A1. Comparison of retained and excluded farmer groups (missing-data assessment)

Characteristic	Retained (n=179)	Excluded (n=16)	SMD	p
PP1 routine meetings (%)	29.6	25.0	0.10	0.783
PP2 routine extension contact (%)	88.3	86.7	0.05	0.693
PP3 very active mentoring (%)	92.7	93.8	−0.04	1.000
Rainfed lowland (%)	28.4	40.0	−0.25	0.380
Technical irrigation (%)	27.3	40.0	−0.27	0.370
Adequate water (%)	49.4	75.0	−0.55	0.067
Year-round water (%)	55.3	50.0	0.11	0.795
Single season, IP100 (%)	10.6	18.8	−0.23	0.399
Hazard-prone (%)	25.7	62.5	−0.80	0.004
Topography score (1-4)	2.89 (0.93)	2.94 (1.06)	−0.05	0.824
Drainage score (1-3)	1.84 (0.83)	2.00 (0.93)	−0.18	0.522
Chairperson age (years)	48.42 (11.15)	47.88 (7.41)	0.06	0.694
Education score (1-5)	3.63 (1.06)	3.44 (1.03)	0.19	0.371
Active members	24.26 (5.89)	23.62 (4.51)	0.12	0.791
Area per member (ha)	0.33 (0.27)	0.38 (0.23)	−0.21	0.225
Fertilizer access (1-4)	3.59 (0.78)	3.50 (0.73)	0.12	0.391

Values are percentages for binary variables and mean (SD) for ordinal and continuous variables. SMD, standardized mean difference; p from Fisher's exact test (binary) or Mann–Whitney test (other). The denominators vary slightly with item non-response.

The retained and excluded groups had nearly identical extension-contact profiles ($|\text{SMD}| \leq 0.10$). The excluded groups were more often hazard-prone (62.5% vs 25.7%; $p=0.004$) and more often reported adequate water (75.0% vs 49.4%; $p=0.067$); other differences were small ($|\text{SMD}| \leq 0.27$). Because exclusion arose from implausible or missing yield records rather than from extension status, and excluded groups had similar extension contact, substantial bias in the contact–productivity association is unlikely, although the underrepresentation of hazard-prone groups should be noted. By contrast, the 18 single-season groups excluded by the earlier complete-case rule differed strongly from the other 161 groups in meeting regularity (88.9% vs 23.0% routine; $p<0.001$) and water adequacy (88.9% vs 45.0%; $p<0.001$), although not in productivity (2.83 vs 2.65 t ha^{-1}; $p=0.542$) or contact frequency (88.9% vs 88.2% routine).

Table A2. Sample distribution, productivity and extension contact by sub-district (n=179)

Sub-district	n	Productivity, mean (SD)	Routine contact (PP2), %	Routine meetings (PP1), %	Rainfed, %	Adequate water, %
Tanduk Kalua	29	1.87 (1.52)	93	21	0	75
Mamasa	25	2.90 (2.41)	76	12	24	48
Rantebulahan Timur	24	2.31 (0.92)	100	46	42	21
Tawalian	18	3.49 (1.96)	100	39	33	17
Balla	14	2.02 (1.46)	93	7	43	50
Buntu Malangka	12	3.79 (1.10)	50	25	17	42
Nosu	11	2.08 (1.16)	91	100	18	100

Sub-district	n	Productivity, mean (SD)	Routine contact (PP2), %	Routine meetings (PP1), %	Rainfed, %	Adequate water, %
Sesena Padang	9	1.67 (1.33)	89	11	14	67
Mehalaan	8	3.19 (1.89)	100	25	62	38
Messawa	7	3.98 (2.06)	100	14	43	57
Aralle	4	4.00 (0.81)	50	25	0	25
Pana	4	3.60 (0.00)	100	100	0	100
Sumarorong	4	2.41 (1.91)	100	0	75	25
Tabulahan	4	2.31 (0.90)	50	25	75	25
Tabang	3	2.47 (2.98)	100	0	100	100
Bambang	2	4.50 (2.12)	100	0	50	0
Mambi	1	1.50 (–)	100	100	0	100
All sub-districts	**179**	**2.67 (1.75)**	**88**			

Rainfed and adequate-water percentages exclude groups with unspecified typology or missing water data. Kruskal–Wallis test for productivity across sub-districts: H=34.33, p=0.005.