Working Paper Series **26-08A**

Does education strengthen the perceived negative link between corruption and tax compliance? Evidence from Indonesia

Arifin Rosid

Directorate General of Taxation &
Universitas Indonesia

Does education strengthen the perceived negative link between corruption and tax compliance? Evidence from Indonesia

Arifin Rosid

Directorate General of Taxation &

Universitas Indonesia

Does education strengthen the perceived negative link between corruption and tax compliance? Evidence from Indonesia

Arifin Rosid [1, 2 a]

[1] Directorate General of Taxation, Jl. Gatot Subroto Kav 40-42, Jakarta, , Indonesia
[2] Universitas Indonesia, Jl. Prof. Dr. Sumitro Djojohadikusumo, Depok 16424, Indonesia
[a] arifin.rosid@ui.ac.id

Abstract

A pooled education gradient can be statistically significant, robust to adjustment, and yet describe no group within the sample. This study examines whether employment status modifies the association between education and corruption-related tax attitudes. Data were drawn from a 2022 cross-sectional national survey of Indonesian individual income taxpayers (N = 8,975), comprising employees, own-account entrepreneurs, and independent professionals. Respondents rated, on a four-point scale, their agreement with the statement that news of government corruption negatively affects the fulfilment of tax obligations. Responses were modelled by cumulative-logit regression with education (four levels), employment status, and their interaction as predictors. In the pooled model, the odds of stronger agreement rose by approximately 10% per additional level of education (OR = 1.101, $p < .001$), and adjustment for employment status altered the log-odds coefficient by only 1.5%. Stratified analysis, however, reversed this picture. The gradient was confined to own-account entrepreneurs (OR = 1.206, 95% CI [1.144, 1.272]) and was absent among independent professionals (OR = 1.028 [0.947, 1.115]) and employees (OR = 0.989 [0.924, 1.058]). The education-by-employment interaction was significant ($\chi^2(2) = 24.46$, $p < .001$), whereas employment status showed no main effect ($\chi^2(2) = 0.47$, $p = .792$). Equivalence testing indicated that both null gradients reflected a genuine absence of effect rather than insufficient power. Adjustment addresses confounding but remains silent about effect modification; consequently, taxpayer segmentation built on pooled gradients may mislead.

Keywords: tax compliance; corruption perceptions; ordinal regression; effect modification; equivalence testing

A. Introduction

Tax compliance depends on more than the expected-value calculus of audits and penalties. The deterrence model of Allingham and Sandmo (1972) remains foundational, but enforcement alone cannot explain the relatively high levels of compliance observed in many settings (Frey & Torgler, 2007; Slemrod, 2019; Torgler, 2007). A broader literature therefore locates compliance in tax morale, conditional cooperation, and confidence in public institutions, with stronger institutional quality and trust associated with greater willingness to comply (Frey & Torgler, 2007; Torgler, 2007). Within this framework corruption is especially corrosive, because it undermines the legitimacy of the fiscal exchange between citizens and the state. Where taxpayers perceive that public officials divert revenue for private gain, those perceptions weaken compliance attitudes and norms and reduce intentions to report income fully (Rosid et al., 2016, 2018). Corruption thus acts not only through material incentives but by eroding the moral and institutional foundations that make voluntary payment possible.

If corruption undermines compliance through a normative channel, perceptions of that link should vary with the interpretive resources taxpayers bring to public finance, and education is a plausible source of that variation. Education tends to increase knowledge of tax law, public spending, and the social purpose of taxation, which can strengthen the capacity to connect revenue collection to state performance and institutional integrity (Kornhauser, 2007; Mebratu, 2024; Trifan et al., 2023). The literature does not, however, support a simple or uniform positive gradient. Although some cross-national evidence associates university education with higher tax morale, other studies report weak, null, or negative education effects, often because more educated taxpayers are also better placed to detect administrative failure, inequity, and opportunities for noncompliance (Adem et al., 2024; Paleka et al., 2023; Williams & Krasniqi, 2017). The more defensible expectation is therefore not that education uniformly raises the belief that corruption reduces compliance, but that education conditions how taxpayers interpret corruption's implications for fairness, reciprocity, and the legitimacy of taxation. That heterogeneity remains underexamined, even though it determines whether pooled estimates of the corruption–compliance relationship are substantively meaningful.

Adjustment, confounding, and effect modification—Confounding and effect modification are distinct, yet the two are routinely conflated in applied research. Confounding is a form of bias: a third variable distorts the observed association between an exposure and an outcome, and appropriate control is intended to reduce that distortion (Inoue et al., 2024; Miettinen, 1974; Park et al., 2026). Effect modification is not a bias to be removed but a substantive finding that the association varies across levels of another variable, and in that setting a single adjusted estimate conceals meaningful heterogeneity rather than summarising it (Kenah, 2023; VanderWeele & Knol, 2014).

The practical implication is easily missed. When a pooled estimate changes little after covariate adjustment, that stability is evidence against substantial confounding, but it is uninformative about effect modification (Inoue et al., 2024). If a covariate modifies the exposure–outcome association without materially confounding it, the adjusted and unadjusted pooled estimates may be nearly identical even though the stratum-specific associations differ sharply in magnitude or direction. Robustness to adjustment is therefore fully compatible with substantial heterogeneity, and detecting that heterogeneity requires stratum-specific estimation or explicit interaction terms rather than reliance on pooled adjustment alone (Tobías et al., 2024; Tsoi & Sun, 2026; VanderWeele & Knol, 2014). The present study supplies an unusually clean empirical instance of this configuration.

Why employment status might modify the education gradient—Employment status is relevant here not because it marks a stable attitudinal difference, but because it structures how taxpayers encounter the tax system. Every respondent in this sample operates a business or practises an independent profession, yet they occupy different positions within the fiscal process. Own-account entrepreneurs bear the full demands of self-assessment: they calculate, report, and remit their own liabilities, and therefore experience the fiscal exchange with the state in a comparatively direct way (Choo et al., 2016; Muehlbacher et al., 2017; Widati et al., 2023). Independent professionals occupy a formally similar position, but professional training and institutional embeddedness may already supply settled interpretive frameworks for judging state performance, leaving less room for education to alter how the corruption–compliance relationship is understood (Ogembo, 2022). Respondents who also hold wage employment, by contrast, encounter a larger share of their taxation indirectly, through withholding or third-party reporting, which can make the act of payment less salient and weaken reflection on how revenue is used (Choo et al., 2016; Muehlbacher et al., 2017).

If that reasoning is correct, the expectation it generates is an interaction rather than a main effect. Education should sharpen the perceived link between corruption and tax compliance most strongly among taxpayers who experience the fiscal exchange unmediated, because those taxpayers are most directly exposed to the connection between payment, public finance, and state performance (Saptono & Khozen, 2023; Schächtele et al., 2022). The point matters methodologically as well as substantively: a variable can show little or no average association with an outcome and still condition the strength of another relationship across groups. Treating employment status only as a main effect would obscure the more relevant question, which is whether education matters more where taxation is directly experienced than where it is institutionally filtered.

The present study—This study analyses a single ordinal outcome—agreement with the proposition that news of government corruption negatively affects the fulfilment of tax obligations—as a function of education and employment status in a national sample of 8,975 Indonesian taxpayers. It addresses three questions. First, does agreement rise with education? Second, does it differ by employment status? Third, does employment status modify the education gradient? Two hypotheses were specified before analysis.

H1. *Agreement rises with education, because education supplies interpretive resources for connecting the conduct of the state to one's own fiscal obligations*.

H2. *Employment status modifies the education gradient, which is steepest where income is entirely self-reported and shallowest where it is subject to withholding*.

The contribution is threefold. Substantively, the paper identifies which part of the self-employed taxpayer population connects corruption to its own fiscal obligations more strongly as education rises, and shows that the connection is confined to a single group rather than being general. Methodologically, it supplies a clean empirical demonstration that stability under covariate adjustment is not evidence that a pooled estimate generalises, and that a covariate with no main effect can nonetheless be indispensable as a moderator. For administration, it shows that taxpayer segmentation built on pooled gradients can be systematically misleading. As will be seen, H1 holds in the pooled data but in only one of three strata, employment status has no main effect whatever, and H2 is supported. The combination is the point: employment status explains nothing on its own, yet it governs whether education explains anything.

B. Methods

Data and participants—The data come from a 2022 national compliance survey of Indonesian personal income taxpayers, covering all provinces. The realised sample is 8,975 respondents. All provided complete data on the outcome and on both grouping variables, so the analytic sample equals the full sample and neither imputation nor case deletion was required. The analysis frame available for this study carries no design weights; estimates are therefore unweighted and describe the realised sample. Reporting follows the quantitative reporting standards of the American Psychological Association (Appelbaum et al., 2018).

Measures—The outcome is a single item asking respondents how far they agree that news of corruption in government negatively affects the fulfilment of tax obligations, answered on a four-point scale. In the source data file the item is coded from 1 = strongly agree to 4 = strongly disagree. This paper reversed it, defining AGREE = 5 − Q30 (i.e. question number 30 in the survey question), so that 4 denotes strong agreement and a positive coefficient means more agreement. This is a presentational choice with no effect on any test statistic, but it prevents the sign confusion that reverse-coded ordinal outcomes routinely generate. Two features of the item deserve emphasis. It solicits assent to a general proposition about taxpayers rather than a confession about the respondent's own behaviour, so every inference here concerns attitudes and not conduct; and because it is a single item, no internal-consistency estimate is available and measurement error is unmodelled.

Education has four ordered levels: secondary or below, diploma, bachelor, and postgraduate. Employment status has three unordered levels: independent professional, own-account entrepreneur, and employee. Tax collection methods differ across respondent groups: employees meet most of their tax liabilities through withholding at source, whereas own-account entrepreneurs and independent professionals self-assess and directly remit a substantially larger share of their taxes. Education is modelled both as an equally spaced score and as four unordered categories, and the two representations are compared formally in Results and Discussion Section.

Statistical model—This study the following statistical model. Let $Y_i \in \{1, 2, 3, 4\}$ denote the ordinal response of respondent *i* and x_i a vector of covariates. This study fit the cumulative-logit model (Agresti, 2010; McCullagh, 1980),

*logit *P*(*Y*~i~ ≤ *j* | *x*~i~) = α~j~ − *x*~i~′*β*,   *j* = 1, 2, 3 ,*

with $\alpha_1 < \alpha_2 < \alpha_3$. Under the minus-sign convention a positive element of β shifts probability mass toward higher categories, that is, toward stronger agreement, and exp(β) is the cumulative odds ratio. The proportional-odds assumption is the restriction that β does not depend on j: a covariate shifts the whole response distribution rather than acting differently at different points of the scale.

Estimation is by maximum likelihood. Cut-points were parameterised as $\alpha_1 = \theta_1$ and $\alpha_j = \alpha_{j-1} + \exp(\theta_j)$, which enforces the ordering without constrained optimisation. Standard errors come from the observed information matrix, computed by central differences of the analytic gradient; sandwich standard errors are reported as a robustness check. Nested models are compared by likelihood-ratio tests, which are preferable to sequential sums of squares in an unbalanced design because they do not depend on the order in which terms enter.

The model sequence is as follows. M0 contains only cut-points. M1 adds employment status, M2 adds education, and M3 contains both. M4 adds the full set of interaction terms. Because education is ordered, this study also fits M3s and M4s, in which education enters as a linear score, so that the interaction can be tested with two degrees of freedom as a difference in gradients rather than six degrees of freedom as a difference in profiles. The test of H2 is the comparison of M3s with M4s.

Assumption checks and robustness—Four assumptions were examined. Proportional odds was assessed with the Brant (1990) test, both omnibus and coefficient by coefficient, supported by inspection of the three threshold-specific slopes. Where the assumption failed, the analysis was repeated in a partial proportional odds framework in which the offending coefficient is free to vary across thresholds (Peterson & Harrell, 1990; Williams, 2016). Linearity of the education score was tested by comparing the score model against the unordered four-category model. Adequacy of the asymptotic reference distribution for the key test was checked by a parametric bootstrap with 2,000 replications generated under the fitted additive model (Efron & Tibshirani, 1993), which matters here because one design cell contains only 111 respondents. Sensitivity to that cell, and to the coding of education, was examined directly.

Two further analyses address the interpretation of null results. First, equivalence testing: the 90% confidence interval for each stratum-specific gradient was compared against a smallest effect of interest set at a 25% change in the odds per level of education, that is, an odds ratio outside the interval from 0.80 to 1.25 (Lakens et al., 2018; Wasserstein & Lazar, 2016). Second, the smallest effect each stratum could detect with 80% power was computed as (z.975 + z.80) × SE = 2.802 × SE and verified by simulation at exactly that value. As a distribution-free complement to the model-based tests this study also reports the generalised Cochran–Mantel–Haenszel correlation statistic with ordinal scores, both unstratified and stratified by employment status (Cochran, 1954; Landis et al., 1978; Mantel & Haenszel, 1959).

C. Results and Discussion

Descriptive statistics—Agreement with the proposition is the majority position: 66.2% of respondents chose one of the two agreement categories, with 22.8% choosing strong agreement and only 3.7% strong disagreement. Figure 1 shows the outcome distribution and the twelve design cells. The design is unbalanced (i.e., the largest cell contains 2,464 respondents and the smallest 111) but every cell is occupied and every response category is used within every cell, so no sparse-data corrections are required.

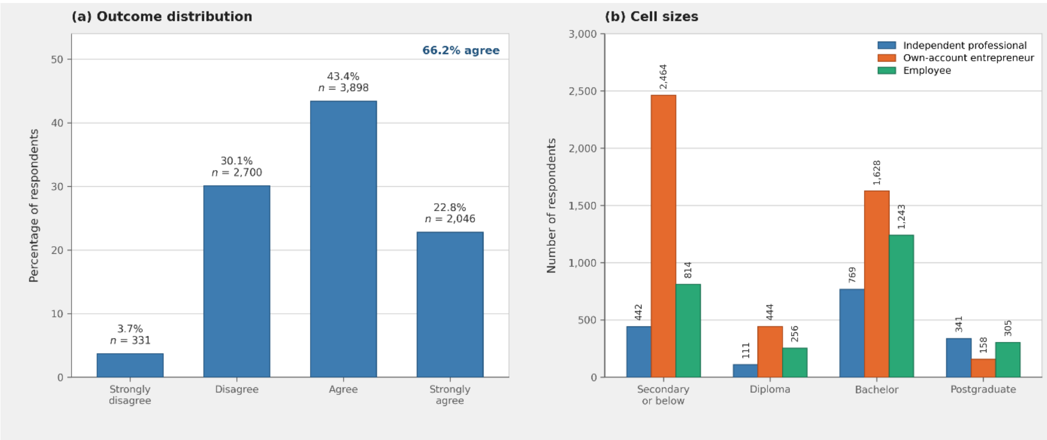

Figure 1. Outcome Distribution & Cell Sizes

Note. *N* = 8,975. The outcome is agreement that news of government corruption affects the fulfilment of tax obligations, reverse-coded so that higher categories denote stronger agreement.

Table 1. Percentage Agreeing by Education and Employment Status

Education	n = 8,975	% agreeing (3–4)	% strongly agreeing	Mean
Independent professional (n = 1,663)				
Secondary or below	442	67.42	22.17	2.862
Diploma	111	63.96	26.13	2.874
Bachelor	769	68.40	23.80	2.889
Postgraduate	341	66.86	26.10	2.894
Own-account entrepreneur (n = 4,694)				
Secondary or below	2,464	61.57	19.32	2.769
Diploma	444	66.89	24.10	2.869
Bachelor	1,628	68.92	26.84	2.927
Postgraduate	158	73.42	32.91	3.032
Employee (n = 2,618)				
Secondary or below	814	70.27	20.15	2.865
Diploma	256	66.41	21.09	2.820
Bachelor	1,243	66.61	22.85	2.859
Postgraduate	305	65.25	23.93	2.846

Note. Categories 3 and 4 are "agree" and "strongly agree". The mean treats the four-point scale as numeric and is reported for orientation only; all inference uses the ordinal model. The monotone rise across education is visible only in the own-account entrepreneur block.

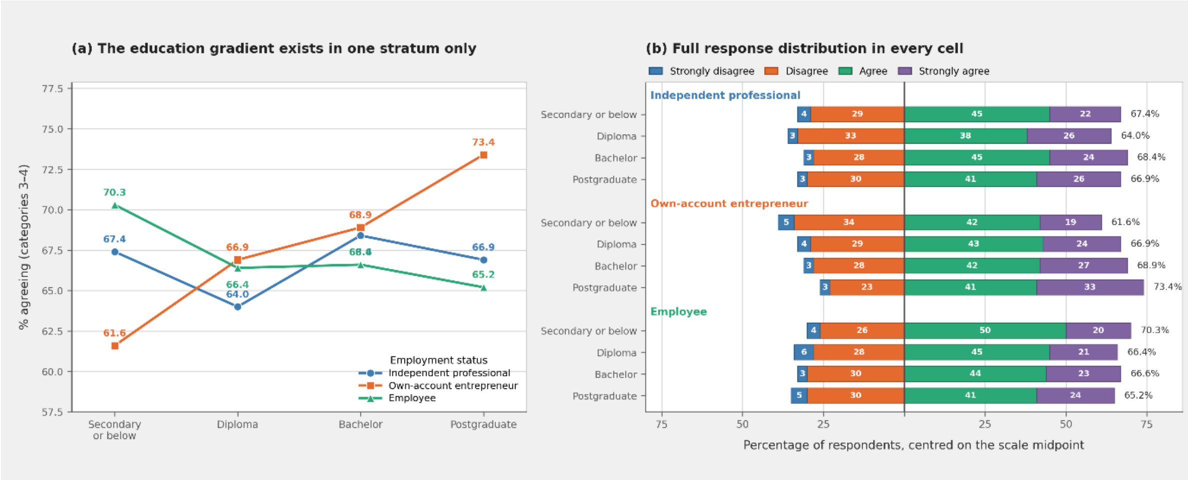

Figure 2. Observed Agreement by Education within Each Employment Stratum

Note. Panel (a) plots the percentage in the two agreement categories. Panel (b) shows the complete response distribution for each of the twelve cells as a diverging stacked bar centred on the scale midpoint, with the percentage above the midpoint at the right of each bar.

Table 1 and Figure 2 already contain the paper's central finding. Among own-account entrepreneurs the proportion agreeing climbs monotonically from 61.6% at secondary education to 73.4% at postgraduate level, a rise of 11.9 percentage points. Among independent professionals the series is flat and non-monotone, and among employees it drifts mildly downward, from 70.3% to 65.2%. The three strata are also ordered differently at different education levels, which is why no main effect of employment status emerges.

Table 2: Model Sequence and Likelihood-Ratio Tests

Model	Parameters	Log-likelihood	AIC	BIC
M0 cut-points only	4	−10,611.520	21,231.04	21,259.45
M1 + employment status	5	−10,609.882	21,229.76	21,265.28
M2 + education (4 categories)	6	−10,596.977	21,205.95	21,248.57
M3 employment + education	8	−10,596.744	21,209.49	21,266.31
M4 M3 + interaction (categorical)	14	−10,584.693	21,197.39	21,296.82
M3s employment + education score	6	−10,597.430	21,206.86	21,249.47
M4s M3s + score × employment	8	−10,585.198	21,186.40	21,243.21
Likelihood-ratio tests				
Comparison	**df**	**χ^2**	**p**	
Education | employment (M1 vs M3)	3	26.276	< .001	
Employment | education (M2 vs M3)	2	0.465	.792	
Interaction, categorical (M3 vs M4)	6	24.102	< .001	
Interaction, score (M3s vs M4s)	2	24.464	< .001	
Education trend (M1 vs M3s)	1	24.904	< .001	
Departure from linearity (M3s vs M3)	2	1.372	.504	

Note. All models are cumulative-logit fits with three cut-points. "Parameters" counts cut-points and regression coefficients. The primary test of H2 is the score interaction, M3s versus M4s. Employment status contributes nothing as a main effect yet is essential as a moderator, a combination discussed in Section 3.8.

Model sequence—Table 2 reports the model sequence. Education improves fit substantially: adding it to a model containing employment status yields $\chi^2(3) = 26.28$, $p < .001$. Employment status does not: adding it to a model already containing education yields $\chi^2(2) = 0.47$, $p = .792$, a result that in ordinary practice would justify dropping the variable. The interaction, however, is strongly supported. With education as a linear score the gradient-difference test gives $\chi^2(2) = 24.46$, $p < .001$, and with education as four categories the profile-difference test gives $\chi^2(6) = 24.10$, $p < .001$. The two-degree-of-freedom score version is the more powerful and is used as the primary test. Model M4s is the best of the seven by AIC, and also by BIC, which penalises its eight parameters heavily.

The education gradient within each employment stratum—Table 3 and Figure 3 give the result on which the paper turns. Among own-account entrepreneurs each additional level of education multiplies the odds of stronger agreement by 1.206, 95% CI [1.144, 1.272], $p < .001$. Among independent professionals the estimate is 1.028, [0.947, 1.115], $p = .509$, and among employees 0.989, [0.924, 1.058], $p = .739$. The contrast between the entrepreneur gradient and the employee gradient is unambiguous, $z = 4.53$, $p < .001$, as is the contrast with independent professionals, $z = 3.21$, $p = .001$. The two null strata do not differ from each other, $p = .471$.

It is worth being explicit about what is not claimed. The employee gradient is numerically negative but statistically indistinguishable from zero, and this work does not claim that education reduces agreement in that stratum. What the data establish is that a gradient present among own-account entrepreneurs is absent among employees and among independent professionals, and that the difference between those gradients is larger than sampling variation will account for.

Table 3: Education Gradient Estimated Separately Within Each Employment Stratum

Employment stratum	*n*	*b*	*SE*	Robust SE	OR [95% CI]	*p*
Independent professional	1,663	.0275	0.0417	0.0417	1.028 [0.947, 1.115]	.509
Own-account entrepreneur	4,694	.1876	0.0272	0.0274	1.206 [1.144, 1.272]	< .001
Employee	2,618	−.0115	0.0346	0.0343	0.989 [0.924, 1.058]	.739
Pairwise contrasts between gradients						
Independent vs Own-account		−.1601	0.0498		0.852 [0.773, 0.939]	.001
Independent vs Employee		.0391	0.0542		1.040 [0.935, 1.156]	.471
Own-account vs Employee		.1991	0.0440		1.220 [1.120, 1.330]	< .001

Note. b is the coefficient on the education score in a cumulative-logit model fitted within the stratum; OR = exp(b) is the cumulative odds ratio per one-level rise in education. Robust standard errors are sandwich estimates. For the contrasts, the odds ratio column reports the ratio of the two stratum-specific odds ratios. The omnibus test of whether the gradients differ is the interaction in Table 2.

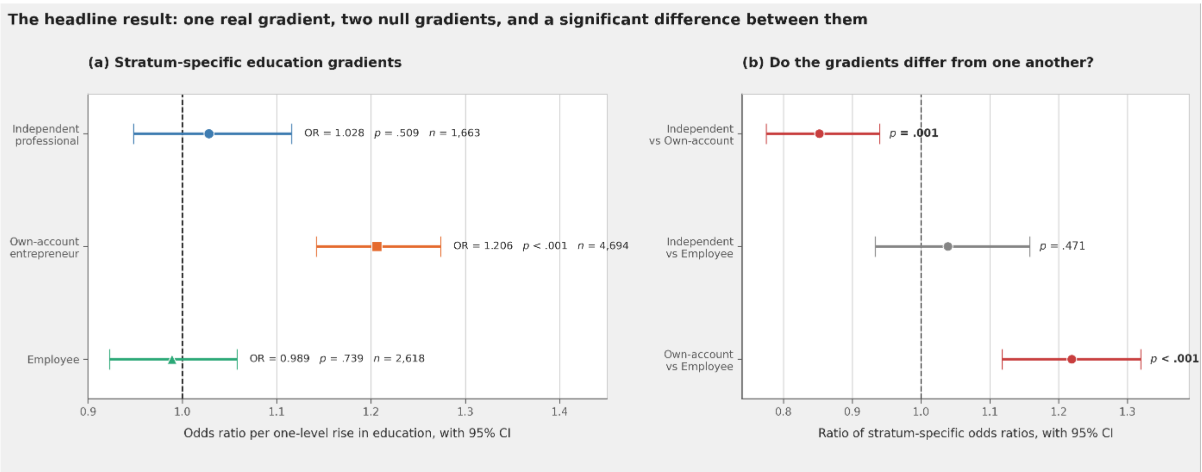

Figure 3. Stratum-specific education gradients and the contrasts between them.

Note. Panel (a) shows the odds ratio per one-level rise in education within each stratum. Panel (b) shows the ratio of stratum-specific odds ratios for each pair; an interval excluding 1 indicates that the two gradients differ.

Figure 4 translates the saturated model into response probabilities. For own-account entrepreneurs the model-implied probability of agreement rises from 61.7% at secondary education to 75.2% at postgraduate level, driven mainly by movement into the strongly agree category, whose probability rises from 19.4% to 31.1%. In the other two strata the four probabilities are almost flat across education.

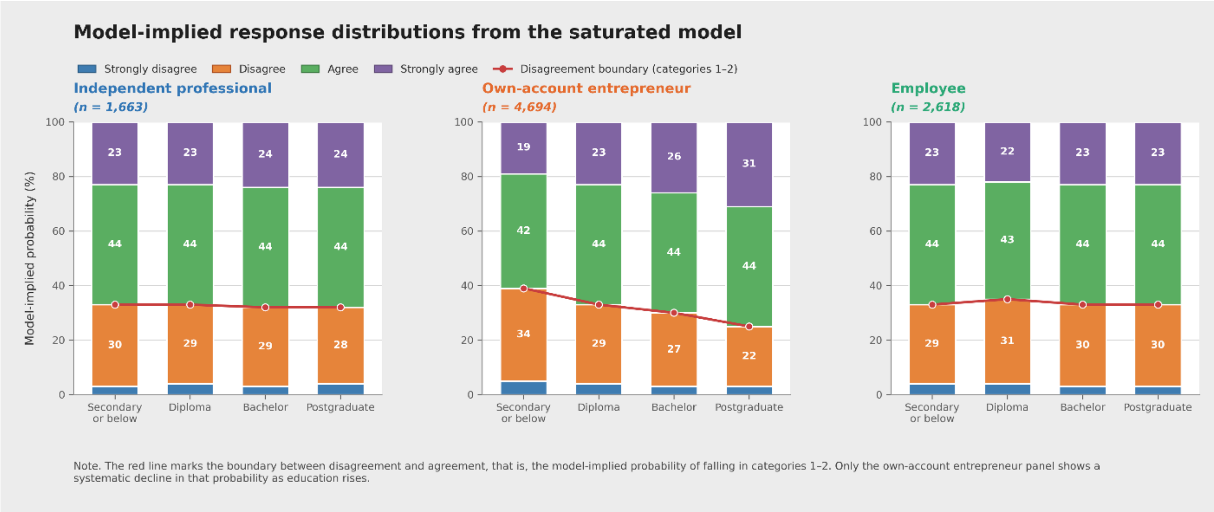

Figure 4. Model-implied response distributions from the saturated model.

Note. Probabilities are computed from model M4, which places no restriction on the twelve cell profiles beyond proportional odds.

Adjustment versus stratification—The methodological point of the paper is made by Table 4 and Figure 5. Education alone yields an odds ratio of 1.1010 per level. Adding employment status changes it to 1.0994, a shift of 1.5% on the log-odds scale. By the usual informal criterion, in which a change below about ten per cent is taken to indicate no meaningful confounding, an analyst would report the pooled gradient as robust and move on. The distribution-free Cochran–Mantel–Haenszel statistic tells the same reassuring story: $\chi^2(1) = 27.38$ unstratified and 24.83 stratified by employment status, both $p < .001$.

Yet the pooled estimate of 1.099 corresponds to no stratum in the data. The three stratum-specific odds ratios are 1.028, 1.206, and 0.989, so the pooled figure is a precision-weighted compromise between one substantial gradient and two absent ones, and it describes the experience of no taxpayer group. The reason adjustment fails to reveal this is instructive: adjustment corrects for association between the covariate and the exposure, and here the modifier operates almost orthogonally to education, so there is nothing for adjustment to correct. Effect modification is invisible to it by construction (Greenland et al., 1999; VanderWeele & Knol, 2014).

Table 4. Marginal, Adjusted, and Stratum-Specific Estimates of the Education Gradient

Estimate	b	SE	OR [95% CI]	Interpretation
Marginal (education alone)	.0962	0.0182	1.101 [1.062, 1.141]	pooled
Adjusted for employment status	.0947	0.0190	1.099 [1.059, 1.141]	pooled
Stratum-specific				
Independent professional	.0275	0.0417	1.028 [0.947, 1.115]	no gradient
Own-account entrepreneur	.1876	0.0272	1.206 [1.144, 1.272]	gradient present
Employee	−.0115	0.0346	0.989 [0.924, 1.058]	no gradient

Note. The change from the marginal to the adjusted estimate is −1.5% on the log-odds scale, which by conventional criteria indicates the absence of confounding. The stratum-specific estimates show that this reassurance is misplaced: the pooled estimate averages heterogeneous quantities.

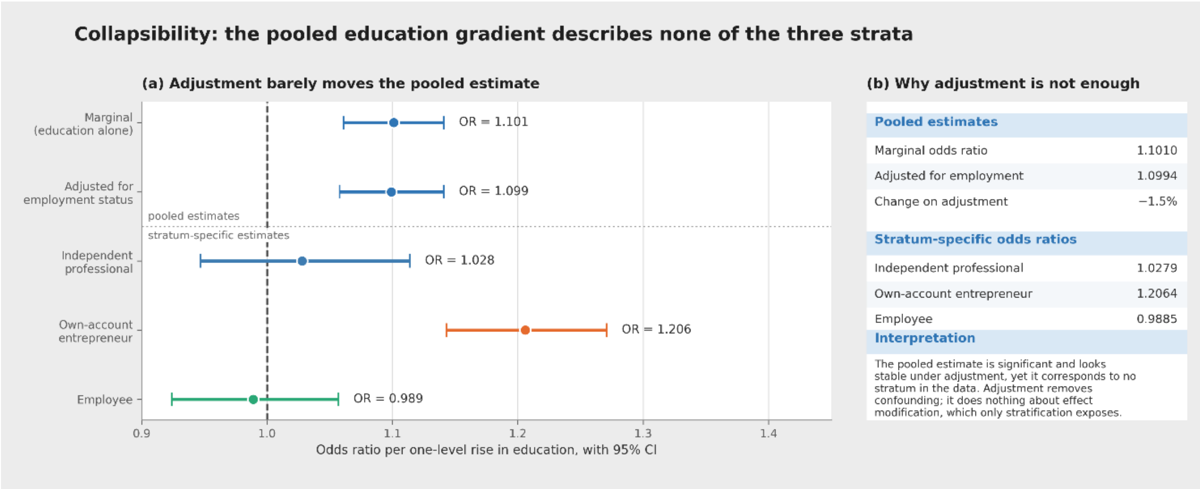

Figure 5. Adjustment does not expose effect modification.

Note. The two pooled estimates are nearly identical and both differ from every stratum-specific estimate.

Assumption checks—The proportional-odds assumption is violated. The Brant omnibus test gives $\chi^2(6) = 20.18$, $p = .003$. The coefficient-by-coefficient tests locate the violation in two places: the education score, $\chi^2(2) = 10.73$, $p = .005$, and the own-account entrepreneur indicator, $\chi^2(2) = 12.72$, $p = .002$; the independent-professional indicator is consistent with proportional odds, $\chi^2(2) = 2.76$, $p = .252$. Because the education score carries the substantive interest, this study treats its departure as the consequential one. Inspecting the threshold-specific estimates shows what this means substantively rather than merely technically. Refitting in a partial proportional odds framework, the education coefficient is .0567 at the first cumulative split, .0691 at the second, and .1342 at the third. Education thus operates most strongly at the boundary between agreement and strong agreement: it does not so much recruit taxpayers into agreeing as intensify the conviction of those already inclined to agree. Crucially, the interaction survives relaxation of the assumption. Repeating the key test in the partial proportional odds framework, with education and its interactions free to vary across thresholds, gives $\chi^2(6) = 26.71$, $p < .001$. Threshold-by-threshold binary logits confirm the pattern: among own-account entrepreneurs the odds ratio per education level is 1.133, 1.182, and 1.246 at the three successive cut-points, whereas in the other two strata every interval includes one at every cut-point.

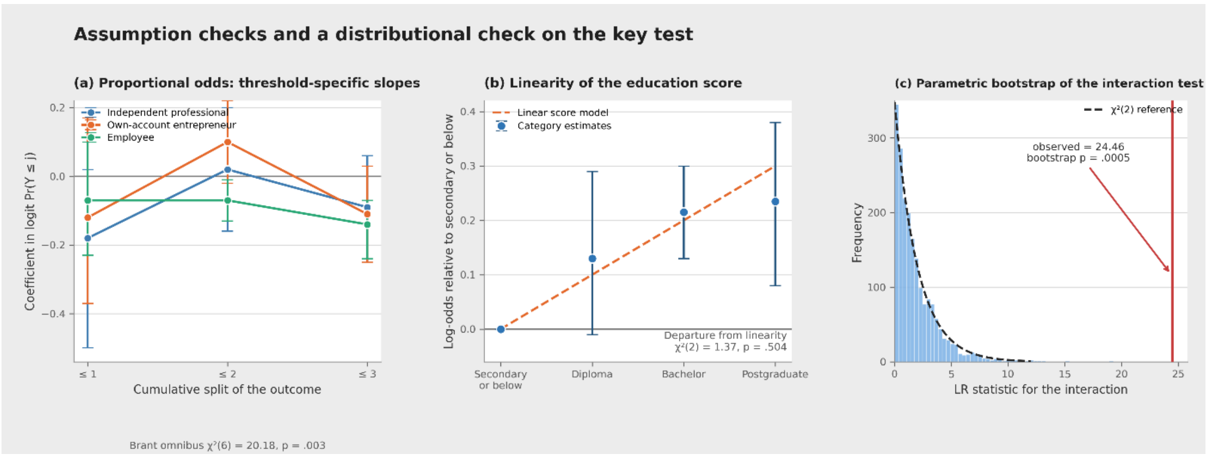

Figure 6. Assumption checks and a distributional check on the key test.

Note. In panel (a) the coefficients are for logit $P(Y \leq j)$, so the sign is reversed relative to the main text. Panel (c) compares the bootstrap null distribution of the interaction statistic against its asymptotic $\chi^2(2)$ reference; the observed value lies far outside both.

The remaining checks are satisfactory. Departure from linearity in education is negligible, $\chi^2(2) = 1.37$, $p = .504$, so the equally spaced score is an adequate representation, although the category estimates in Figure 6 suggest that the bachelor and postgraduate levels are closer together than equal spacing implies. Sandwich and model-based standard errors differ by at most 5.04% in the saturated model. The parametric bootstrap of the interaction statistic gives $p = .0005$ against an asymptotic $p = .000005$, with a bootstrap 95th percentile of 6.11 against the nominal 5.99, so the reference distribution is very slightly conservative but not materially so. The bootstrap p-value is bounded below by $1/(B + 1) = .0005$ with 2,000 replications, so it should be read as confirming that the observed statistic lies beyond the resolvable tail of the null distribution rather than as an independent estimate of that tail probability.

Are the two null gradients informative—A null result is uninterpretable without a statement of what the design could have detected. Against a smallest effect of interest of a 25% change in odds per education level, equivalence is supported in both null strata: the 90% interval for independent professionals runs from $OR = 0.960$ to 1.101, and for employees from 0.934 to 1.046, both contained within the equivalence region. For own-account entrepreneurs, correctly, equivalence is not supported. The smallest gradient detectable with 80% power is $OR = 1.124$ among independent professionals and 1.102 among employees, each verified by simulation (attained power .80 and .76 respectively over 300 replications). The observed gradients are about a quarter and an eighth of those thresholds on the log-odds scale. These strata are therefore not merely failing to show an effect; they are shown not to have one of any consequence. Because the equivalence bounds are stated on a single proportional-odds summary, they should be read as applying to the average of the threshold-specific gradients within each stratum. Figure 7 summarises both analyses.

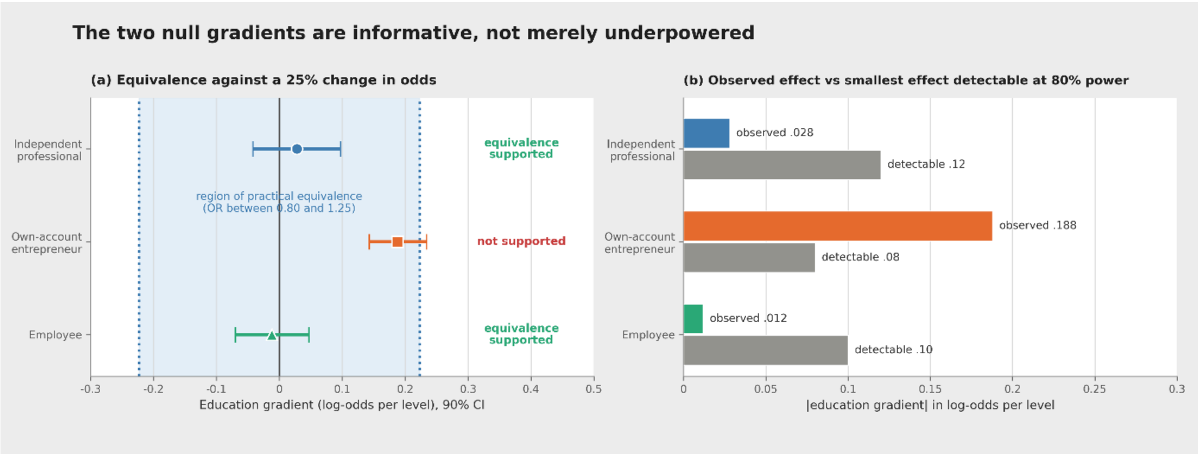

Figure 7. Equivalence testing and minimum detectable effects.

Note. The shaded region in panel (a) is the region of practical equivalence, an odds ratio between 0.80 and 1.25. In panel (b) the simulated attained power at the analytic minimum detectable effect was .76 for employees, slightly below the nominal .80, which is within Monte Carlo error for 300 replications.

Sensitivity analyses—Table 5 collects the robustness checks. The interaction is not an artefact of the smallest cells: dropping the 111 independent professionals with a diploma leaves the statistic essentially unchanged, and dropping the 158 postgraduate own-account entrepreneurs — the cell carrying the most extreme observed percentage — reduces it only from 24.46 to 21.34, still significant at $p < .001$, with the entrepreneur odds ratio barely moving from 1.206 to 1.201. Removing the entire diploma level changes nothing. Nor does the result depend on how education is scored: an equally spaced score, a rough years-of-schooling scale, and a binary tertiary-versus-not contrast all yield the same conclusion, the last necessarily with a smaller statistic because it discards information.

Table 5. Sensitivity of the Interaction to Influential Cells and to the Coding of Education

Specification	N	$\chi^2(2)$	p	Own-account OR
Excluding influential cells				
Full sample	8,975	24.464	< .001	1.206
Drop diploma × independent (*n* = 111)	8,864	24.490	< .001	1.206
Drop postgraduate × own-account (*n* = 158)	8,817	21.336	< .001	1.201
Drop the whole diploma level (*n* = 811)	8,164	24.715	< .001	1.206
Alternative coding of education				
Equally spaced 0, 1, 2, 3	8,975	24.464	< .001	—
Years of schooling 12, 15, 16, 18	8,975	24.556	< .001	—
Binary: tertiary vs not	8,975	17.697	< .001	—
Relaxing the proportional-odds assumption				
Partial proportional odds (education free)	8,975	$\chi^2(6) = 26.711$	< .001	—
Distributional check				
Parametric bootstrap, 2,000 replications	8,975	24.464	.0005	—

Note. The conclusion is unchanged across every specification. The binary coding of education yields a smaller statistic because collapsing four ordered levels into two discards information, not because the effect is fragile.

Why the gradient is confined to own-account entrepreneurs—Three findings emerge, and their combination is what makes the case worth reporting. Education predicts agreement that corruption undermines tax compliance, but only among own-account entrepreneurs, where each additional level raises the odds by about 21%. Employment status has no main effect at all. And the interaction between them is strong and survives every robustness check applied. A variable that contributes nothing on its own turns out to determine whether the other variable matters. The apparent tension with the saturated model, in which the own-account indicator is significant, is only apparent: that coefficient is a conditional simple effect at the reference education level, not a main effect averaged over education.

The observed pattern is consistent with the reasoning behind H2, although the data do not identify the mechanism directly. Own-account entrepreneurs calculate, report, and remit their own liabilities, and so experience the fiscal exchange with the state in a comparatively direct way; in that setting education may strengthen the interpretive capacity needed to connect tax payment with perceptions of how public revenue is used (Choo et al., 2016; Kurniawan, 2020; Mascagni et al., 2024). Taxpayers who also hold wage employment meet a larger share of their liability through withholding or third-party reporting, which can make the act of payment less salient and may reduce the scope for reflection on the relationship between revenue collection and state performance (Appiah et al., 2024; Choo et al., 2016). Independent professionals present the more informative null. Although they are also formally self-assessing, no comparable gradient appears in this group. One possibility is that professional training or institutional embeddedness already supplies relatively settled views about the state, so that additional education contributes little; another is that the category is more homogeneous on education-related characteristics than the observed scale captures. The present design cannot distinguish between these explanations, and this subgroup therefore warrants replication.

The partial proportional odds results add an important nuance. Education matters most at the boundary between agreement and strong agreement rather than between disagreement and agreement; in substantive terms, it seems to intensify an existing conviction more than to generate it. That reading is consistent with broader evidence that education operates through knowledge, fairness perceptions, and other conditional pathways rather than through a uniform direct effect on compliance attitudes (Akims, 2023; Appiah et al., 2024; Paleka et al., 2023; Tumoro & Pandya, 2025), and with a distribution in which most respondents already accept the corruption–compliance link, so that the variation lies in how strongly that belief is held rather than whether it is held at all.

Implications for tax administration—If the association reflects a real disposition, the practical implication is that communications about the use of public revenue will find their most receptive audience among educated own-account entrepreneurs, and will land on comparatively indifferent ground among employees regardless of their education. That is a targeting statement, and it is available only from the stratified analysis. An administration working from the pooled gradient would infer that education is a useful segmentation variable in general, which is false for two thirds of this population.

A caution belongs alongside that implication. The outcome is an attitude toward a general proposition, not a behaviour, and the literature is clear that attitudes translate into compliance behaviour imperfectly at best, particularly where opportunity rather than willingness is the binding constraint (Kleven et al., 2011; Slemrod, 2019). Nothing here licenses the inference that educated entrepreneurs who believe corruption matters actually comply less.

Limitations—Four limitations qualify the conclusions. First, the design is cross-sectional and observational, so no causal claim can be sustained; education is correlated with income, sector, region, and age, none of which is controlled here, and the deliberate restriction to three variables that makes the methodological point cleanly also makes the substantive interpretation more tentative than it would otherwise be. The absence of design weights and of a documented response rate further limits inference to the realised sample rather than to the taxpayer population.

Second, the outcome is a single item. No internal-consistency estimate is available, measurement error is unmodelled and will attenuate the gradients, and there is no way to establish that the item functions equivalently across education levels. Differential item functioning is a live alternative explanation: if more educated respondents interpret the proposition differently rather than endorsing it more strongly, the observed gradient would be a measurement artefact. Testing that would require multiple indicators of the same construct, which this instrument does not provide.

Third, the item solicits agreement with a general proposition about taxpayers rather than a self-report of behaviour or intention Fourth, the proportional-odds assumption is violated, and although the substantive conclusion survives a partial proportional odds treatment, the violation means that no single odds ratio fully summarises the education effect. The stratum-specific odds ratios in Table 3 should be read as averages across thresholds rather than as constants.

D. Conclusion

Among Indonesian taxpayers, the belief that government corruption undermines tax compliance rises with education only for those whose income is entirely self-reported. The pooled education gradient is statistically significant, barely moves under adjustment for employment status, and describes no group in the data. Employment status has no main effect and is nonetheless the variable that organises the findings. Where an analysis is to inform segmentation or targeting, the stability of a pooled estimate under adjustment is not evidence that the estimate generalises, and only stratification will show whether it does. Future work should test whether the same configuration appears for behavioural outcomes, and whether the null among independent professionals reflects settled professional views or unmeasured homogeneity within that category.

E. References

Adem, M., Desta, T., & Girma, B. (2024). Determinants of tax compliance behavior: A case study in Ethiopia. *SAGE Open*, *14*. https://doi.org/10.1177/21582440241292869

Agresti, A. (2010). *Analysis of ordinal categorical data* (2nd ed.). Wiley.

Akims, M. (2023). Tax education and tax compliance: A review of literature. *Journal of Economics*. https://doi.org/10.53819/81018102t4235

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *1*(3–4), 323–338. https://doi.org/10.1016/0047-2727(72)90010-2

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25.

Appiah, T., Domeher, D., & Agana, J. A. (2024). Tax knowledge, trust in government, and voluntary tax compliance: Insights from an emerging economy. *SAGE Open*, *14*. https://doi.org/10.1177/21582440241234757

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, *46*(4), 1171–1178.

Choo, C. Y. L., Fonseca, M. A., & Myles, G. D. (2016). Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment. *Journal of Economic Behavior and Organization*, *124*, 102–114. https://doi.org/10.1016/j.jebo.2015.09.015

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics*, *10*(4), 417–451.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.

Frey, B. S., & Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics*, *35*(1), 136–159.

Greenland, S., Robins, J. M., & Pearl, J. (1999). Confounding and collapsibility in causal inference. *Statistical Science*, *14*(1), 29–46.

Inoue, K., Sakamaki, K., Komukai, S., Ito, Y., Goto, A., & Shinozaki, T. (2024). Methodological tutorial series for epidemiological studies: Confounder selection and sensitivity analyses to unmeasured confounding from epidemiological and statistical perspectives. *Journal of Epidemiology*, *35*, 3–10. https://doi.org/10.2188/jea.je20240082

Kenah, E. (2023). Rothman diagrams: The geometry of confounding and standardization. *International Journal of Epidemiology*, *53*. https://doi.org/10.1093/ije/dyae139

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, *79*(3), 651–692.

Kornhauser, M. E. (2007). A tax morale approach to compliance: Recommendations for the IRS. *Florida Tax Review*, *8*. https://doi.org/10.5744/ftr.2007.1061

Kurniawan, D. (2020). The influence of tax education in higher education on tax knowledge and its effect on personal tax compliance. *Journal of Indonesian Economy and Business*, *35*(1), 57–72. https://doi.org/10.22146/jieb.54292

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269.

Landis, J. R., Heyman, E. R., & Koch, G. G. (1978). Average partial association in three-way contingency tables: A review and discussion of alternative tests. *International Statistical Review*, *46*(3), 237–254.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, *22*(4), 719–748.

Mascagni, G., Santoro, F., & Mukama, D. (2024). Teach to comply? Evidence from a taxpayer education program in Rwanda. *International Tax and Public Finance*, *32*, 120–162. https://doi.org/10.1007/s10797-023-09809-6

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *42*(2), 109–142.

Mebratu, A. A. (2024). Theoretical foundations of voluntary tax compliance: Evidence from a developing country. *Humanities and Social Sciences Communications*, *11*. https://doi.org/10.1057/s41599-024-02903-y

Miettinen, O. S. (1974). Confounding and effect-modification. *American Journal of Epidemiology*, *100*(5), 350–353. https://doi.org/10.1093/oxfordjournals.aje.a112044

Muehlbacher, S., Hartl, B., & Kirchler, E. (2017). Mental accounting and tax compliance. *Public Finance Review*, *45*(1), 118–139. https://doi.org/10.1177/1091142115602063

Ogembo, D. (2022). Trust or power: Improving tax compliance by self-employed professionals. *Intertax*. https://doi.org/10.54648/taxi2022012

Paleka, H., Karanović, G., & Štambuk, A. (2023). The direct and moderating effect of sociodemographic variables on tax compliance behaviour. *South East European Journal of Economics and Business*, *18*, 34–48. https://doi.org/10.2478/jeb-2023-0017

Park, E. J., Yuk, K., Jeong, J., & Lee, W. J. (2026). Confounding and the healthy worker survivor effect in studies of medical radiation workers: A systematic review of methodological approaches. *Epidemiology and Health*, *48*. https://doi.org/10.4178/epih.e2026009

Peterson, B., & Harrell, F. E. (1990). Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *39*(2), 205–217.

Rosid, A., Evans, C., & Tran-Nam, B. (2016). Do perceptions of corruption influence personal income taxpayer reporting behaviour? Evidence from Indonesia. *eJournal of Tax Research*, *14*(2).

Rosid, A., Evans, C., & Tran-Nam, B. (2018). Tax non-compliance and perceptions of corruption: Policy implications for developing countries. *Bulletin of Indonesian Economic Studies*, *54*(1), 25–60. https://doi.org/10.1080/00074918.2017.1364349

Saptono, P. B., & Khozen, I. (2023). What determines the tax compliance intention of individual taxpayers receiving COVID-19-related benefits? Insights from Indonesia. *International Journal of Sociology and Social Policy*. https://doi.org/10.1108/ijssp-02-2023-0045

Schächtele, S., Eguino, H., & Roman, S. (2022). Fiscal exchange and tax compliance: Evidence from a field experiment. *Journal of Policy Analysis and Management*. https://doi.org/10.1002/pam.22460

Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, *57*(4), 904–954. https://doi.org/10.1257/jel.20181437

Tobías, A., Kim, Y., & Madaniyazi, L. (2024). Time-stratified case-crossover studies for aggregated data in environmental epidemiology: A tutorial. *International Journal of Epidemiology*, *53*. https://doi.org/10.1093/ije/dyae020

Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis*. Edward Elgar. https://doi.org/10.4337/9781847207203

Trifan, V. A., Szentesi, S. G., Cuc, L. D., & Pantea, M. F. (2023). Assessing tax compliance behavior among Romanian taxpayers: An empirical case study. *SAGE Open*, *13*. https://doi.org/10.1177/21582440231195676

Tsoi, C., & Sun, R. (2026). Common reporting errors in subgroup analysis: A comparison of interaction and stratified regression models. *BMC Medical Research Methodology*, *26*. https://doi.org/10.1186/s12874-026-02795-3

Tumoro, D. T., & Pandya, H. (2025). Social psychological drivers and tax revenue for sustainable economic development with the mediating role of tax compliance: Evidence from Hossana city, Ethiopia. *Discover Sustainability*, *6*. https://doi.org/10.1007/s43621-025-01728-2

VanderWeele, T. J., & Knol, M. J. (2014). A tutorial on interaction. *Epidemiologic Methods*, *3*(1), 33–72. https://doi.org/10.1515/em-2013-0005

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: Context, process, and purpose. *The American Statistician*, *70*(2), 129–133.

Widati, S., Permatasari, M., & Fiana, A. (2023). Analysis of individual taxpayer's perception of tax evasion. *DEAL: International Journal of Economics and Business*, *1*(1). https://doi.org/10.37366/deal.v1i01.2685

Williams, C. C., & Krasniqi, B. (2017). Evaluating the individual- and country-level variations in tax morale: Evidence from 35 Eurasian countries. *Journal of Economic Studies*, *44*(5), 816–832. https://doi.org/10.1108/jes-09-2016-0182

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, *40*(1), 7–20.