Online Appendix A: Additional descriptive tables

Table A1: Detailed survey wording, Q1

Item	**Q1: Now we are interested in knowing more about your thoughts on the function and quality of reader comments. Please take a stance on whether you agree with the following statements or not**
Q1_1	News become more interesting to follow thanks to reader comments
Q1_2	Reader comments are often of poor quality
Q1_3	Reader comments give all citizens an opportunity to express their opinion
Q1_4	Reader comments often contain interactions between different people
Q1_5	Reader comments are often dominated by a few individuals
Q1_6	Most reader comments are based on reasoning
Q1_7	Most reader comments are not based on rational arguments
Q1_8	Most reader comments refer to facts
Q1_9	Most reader comments consider potential counterarguments presented in other comments
Q1_10	Discussions in reader comments often lead to agreement on the issue
Q1_11	Discussions in reader comments are rarely respectful or polite in tone
Q1_12	People discussing in reader comments rarely show willingness to change their opinion

Translated from Swedish by authors. All items were measured on a five-point scale, ranging from Completely agree (1) to Completely disagree (5) and with an additional category for Cannot say (6).

Table A2: Detailed survey wording, Q2

Item	**Q2: What are your thoughts on how journalists and the media handle reader comments?**
Q2_1	Journalists should actively participate in the discussion that arises around an online article
Q2_2	Offensive reader comments should always be deleted
Q2_3	The option for readers to comment on articles should be available
Q2_4	In the discussion around an article, journalists should, if necessary, remind readers of the discussion rules
Q2_5	Individual readers who repeatedly comment in an offensive manner should always be banned from commenting
Q2_6	In the discussion, journalists should, if necessary, provide additional facts about the article's topic
Q2_7	In the discussion, journalists should, if necessary, provide additional facts about how the editorial team worked on the article
Q2_8	In the discussion, journalists should, if necessary, inform readers about the editorial team's general journalistic principles and guidelines

Translated from Swedish by authors. All items were measured on a five-point scale, ranging from Completely agree (1) to Completely disagree (5) and with an additional category for Cannot say (6).

Table A3: Respondents' views on the democratic quality of comments sections (%)

	Totally agree	**Partially agree**	**Neither agree nor disagree**	**Partially disagree**	**Totally disagree**	**Cannot say**
News become more interesting with comments (Q1_1)	11.0	31.9	19.4	10.8	11.7	15.1
Comments are often of poor quality (Q1_2)	15.5	33.6	24.8	9.5	2.2	14.3
Comments let all citizens express opinions (Q1_3)	29.2	39.0	11.2	6.6	3.1	10.9
Comments often contain interactions (Q1_4)	10.8	43.1	17.7	7.7	1.3	19.4
Comments are dominated by a few individuals (Q1_5)	28.5	42.1	8.0	2.7	0.5	18.2
Most comments are based on reasoning (Q1_6)	3.3	23.2	23.0	23.0	7.9	19.5
Most comments are not based on arguments (Q1_7)	12.3	36.1	22.0	9.4	2.0	18.2
Most comments refer to facts (Q1_8)	1.5	11.7	17.8	33.6	18.1	17.2
Most comments consider counterarguments (Q1_9)	5.0	22.8	20.1	20.1	12.1	19.9
Discussions often lead to agreement (Q1_10)	0.8	6.5	16.2	28.4	26.8	21.2
Discussions are rarely respectful or polite (Q1_11)	13.1	35.4	20.5	12.5	2.5	15.9
People discussing rarely change opinion (Q1_12)	35.5	36.7	9.5	2.4	0.3	15.6

n=2412–2420. Weighted data. Percentages may not total 100 due to rounding.

Table A4: Respondents' views on journalistic involvement in comments sections (%)

	Totally agree	**Partially agree**	**Neither agree nor disagree**	**Partially disagree**	**Totally disagree**	**Cannot say**
Journalists should participate in discussions (Q2_1)	14.7	31.2	22.5	13.8	5.7	12.1
Offensive comments should be deleted (Q2_2)	48.3	24.7	9.4	7.1	3.5	7.0
Option to comment should be available (Q2_3)	44.7	31.7	9.2	5.7	1.0	7.7
Journalists should remind of discussion rules (Q2_4)	47.4	29.6	8.6	2.9	2.5	9.0
Offenders should be banned (Q2_5)	36.9	32.3	10.2	6.8	3.8	10.0
Journalists should provide additional facts (Q2_6)	38.0	38.8	9.2	3.2	1.9	8.9
Journalists should inform about their work (Q2_7)	35.7	37.7	12.5	3.3	1.2	9.6
Journalists should inform about principles (Q2_8)	39.3	32.4	12.9	1.8	1.3	12.2

n=2392–2398. Weighted data. Percentages may not total 100 due to rounding.

Table A5: User engagement in comments sections

	Reading (Q3_1)		**Reacting (Q3_2)**		**Contributing (Q3_3)**	
	Freq.	**Pct.**	**Freq.**	**Pct.**	**Freq.**	**Pct.**
Daily	306	13.6	46	2.0	12	0.5
Often	449	20.0	121	5.4	30	1.3
Sometimes	609	27.1	337	15.0	148	6.6
Rarely	549	24.4	591	26.3	429	19.1
Never	326	14.5	1130	50.3	1615	71.8
Cannot say	8	0.4	22	1.0	13	0.6

n=2248. Weighted data. Percentages may not total 100 due to rounding.

Online Appendix B: Principal component analyses

Table B1: Principal component analysis: Democratic quality

	Component 1 (Toxicity)	**Component 2 (Vitality)**
People discussing rarely change opinion (Q1_12)	−0.711	
Comments are dominated by a few individuals (Q1_5)	−0.697	
Comments are often of poor quality (Q1_2)	−0.678	
Most comments are not based on arguments (Q1_7)	−0.644	
Discussions are rarely respectful or polite (Q1_11)	−0.621	
Comments often contain interactions (Q1_4)		0.683
News become more interesting with comments (Q1_1)		0.656
Most comments consider counterarguments (Q1_9)		0.647
Most comments are based on reasoning (Q1_6)		0.606
Comments let all citizens express opinions (Q1_3)		0.567
Most comments refer to facts (Q1_8)		0.529
~~Discussions often lead to agreement (Q1_10)~~	(not included in the analysis)	
Eigenvalues	3.636	1.551
Variance explained (%)	33.1	14.1
Rotated SS	2.919	2.799
Reliability (Cronbach's α)	0.74	0.71

$n = 1531$ (unweighted data). Rotated solution (oblimin rotation with Kaiser normalization). Determinant of the correlation matrix: 0.090; Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy: 0.860, with all individual values above 0.73; Bartlett's test of sphericity: approx. χ^2 (55) = 3674, $p < 0.001$. (Values below 0.4 have been suppressed from the table.)

Table B2: Principal component analysis: Journalistic involvement

	Component 1 (Participation)	**Component 2 (Moderation)**
Journalists should provide additional facts (Q2_6)	0.825	
Journalists should inform about their work (Q2_7)	0.805	
Journalists should participate in discussions (Q2_1)	0.660	
Journalists should inform about principles (Q2_8)	0.613	
Option to comment should be available (Q2_3)	0.540	
Offenders should be banned (Q2_5)		0.852
Offensive comments should be deleted (Q2_2)		0.841
~~Journalists should remind of discussion rules (Q2_4)~~	(not included in the analysis)	
Eigenvalues	2.513	1.624
Variance explained (%)	35.9	23.2
Rotated SS	2.453	1.707
Reliability (Cronbach's α)	0.72	0.70

$n = 1911$ (unweighted data). Rotated solution (oblimin rotation with Kaiser normalization). Determinant of the correlation matrix: 0.209; Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy: 0.716, with all individual values above 0.56; Bartlett's test of sphericity: approx. χ^2 (21) = 2989, $p < 0.001$. (Values below 0.4 have been suppressed from the table.)

Online Appendix C: Regression diagnostics

In Online Appendix C, we present the results of diagnostic tests for the regression models reported in the paper.

Online Appendix C.1: Model 1 through Model 4

The Durbin–Watson statistic returned values of 2.03 (Model 1), 1.89 (Model 2), 1.98 (Model 3), and 2.05 (Model 4), indicating no serious autocorrelation in the residuals. To check for heteroskedasticity, we applied the Breusch–Pagan test. It yielded p-values of 0.974 (Model 1), 0.114 (Model 2), 3.1e–05 (Model 3), and 2.2e–16 (Model 4), suggesting that heteroskedasticity may be a concern in Model 3 and Model 4. Normality of the residuals was assessed using Q–Q plots, which indicated that the residuals are approximately normally distributed in Model 1 and in Model 2, but not in Model 3 and Model 4. Multicollinearity was examined through the variance inflation factor. All values are below 10 (below 1.05), suggesting no concerns with multicollinearity. Finally, potential outliers were identified through Cook's distance. No observations exceeded the critical threshold of 1, confirming that no influential outliers significantly impacted the models. Given the issues identified in the diagnostics, particularly with heteroskedasticity and departure from normality in Model 3 and Model 4, we decided to use robust regression with an M estimator (available in the MASS package for R) in these models.

Online Appendix C.2: Model 5a through Model 7b

The assumption of proportional odds was tested using the Brant test. Some models (5b, 6a, 6b, 7b) failed the proportional odds assumption. Therefore, as a robustness check, we computed generalized ordered logit models (see Online Appendix D). Multicollinearity was examined through the variance inflation factor (using linear regression, with numeric version of dependent variable). For all models, values are far below 10 (below 1.22), suggesting no concerns with multicollinearity. Potential outliers were identified through Cook's distance. No observations exceeded the critical threshold of 1.

Online Appendix D: Robustness

In Online Appendix D, we briefly discuss the robustness of the regression models reported in the paper.

Online Appendix D.1: Model 1 through Model 4

Table D1 and Table D2 present alternative specifications of Model 1 through Model 4. In these models, the dependent variables are constructed by summing the raw item scores rather than relying on factor scores derived through the regression method. The summed raw scores are obtained by aggregating the item values that load onto the respective components. As demonstrated by Model 1′, Model 2′, Model 3′, and Model 4′, the results remain largely consistent with those of the main analyses, suggesting that the method employed to construct the dependent variables does not materially affect the substantive conclusions. Furthermore, additional analyses employing weighted data (for Model 1 through Model 4) as well as alternative robust estimation methods (for Model 3 and Model 4) produced substantively similar patterns of findings. In the interest of conciseness, these supplementary results are not presented here but are available from the authors upon request.

Table D1: Factors influencing perceived democratic quality

	Toxicity[a] **(Model 1′)**	**Vitality**[b] **(Model 2′)**
Intercept	−9.086***	−18.490***
	(0.313)	(0.409)
Age	−0.037***	0.047***
	(0.004)	(0.006)
Gender: Female	(ref.)	(ref.)
Gender: Male	0.060	0.214
	(0.157)	(0.201)
Education: Primary or secondary	(ref.)	(ref.)
Education: Post-secondary	−0.110	−1.080***
	(0.169)	(0.217)
Income: Low income household[c]	(ref.)	(ref.)
Income: Middle income household[d]	0.339	−0.751**
	(0.198)	(0.256)
Income: High income household[e]	0.677**	−1.490***
	(0.231)	(0.295)
F	15.20***	30.98***
R^2	0.04	0.09
Adjusted R^2	0.04	0.09
n	1638	1549

Unstandardized linear regression coefficients, with standard errors in parentheses. Unweighted data. *** $p < 0.001$, ** $p < 0.01$. [a]Higher values indicate greater perceived toxicity (summed raw scores, reversed). [b]Higher values indicate greater perceived vitality (summed raw scores, reversed). [c]≤3000€/mth. [d]3001–7000€/mth. [e]≥7001€/mth.

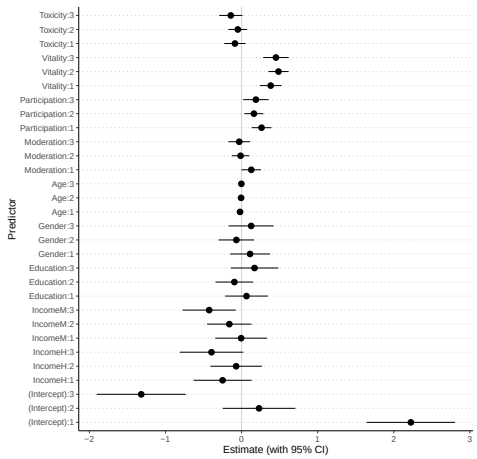
Table D2: Factors influencing attitudes toward journalistic involvement

	Participation[a] **(Model 3′)**	**Moderation**[b] **(Model 4′)**
Intercept	−10.710***	−3.806***
	(0.296)	(0.167)
Age	0.036***	0.007**
	(0.004)	(0.002)
Gender: Female	(ref.)	(ref.)
Gender: Male	−0.361*	−0.509***
	(0.147)	(0.074)
Education: Primary or secondary	(ref.)	(ref.)
Education: Post-secondary	−0.381*	0.093
	(0.154)	(0.080)
Income: Low income household[c]	(ref.)	(ref.)
Income: Middle income household[d]	−0.354*	0.173
	(0.177)	(0.092)
Income: High income household[e]	−0.523*	0.155
	(0.217)	(0.111)
F	20.51***	16.12***
Pseudo-R^2 (McFadden)	0.08	0.07
n	1823	2008

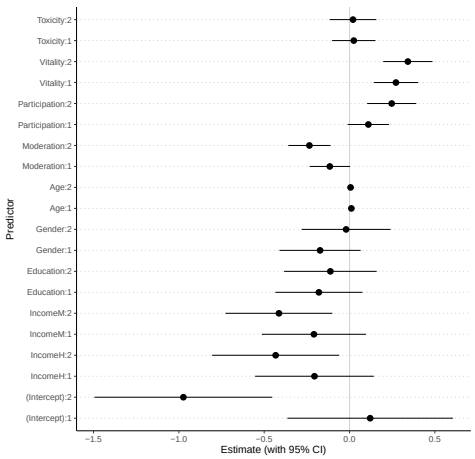
Unstandardized robust regression coefficients, with robust standard errors in parentheses. Unweighted data. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. [a]Higher values indicate more positive attitudes toward journalistic participation (summed raw scores, reversed). [b]Higher values indicate more positive attitudes toward journalistic moderation (summed raw scores, reversed). [c]≤3000€/mth. [d]3001–7000€/mth. [e]≥7001€/mth.

Online Appendix D.2: Model 5a through Model 7b

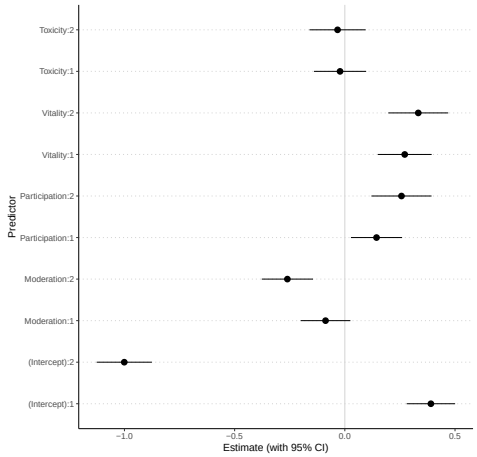
Figure D1 displays the results from generalized ordered logit models for the specifications that failed to meet the proportional odds assumption. Model 5b′ (a), Model 6a′ (b), Model 6b′ (c), and Model 7b′ (d) indicate that, overall, the substantive interpretation of the models remains consistent with earlier analyses. Moreover, supplementary analyses incorporating weighted data produced substantively equivalent results, further reinforcing the robustness of the findings. To maintain clarity and brevity, these additional results are not displayed but are available from the authors upon request.



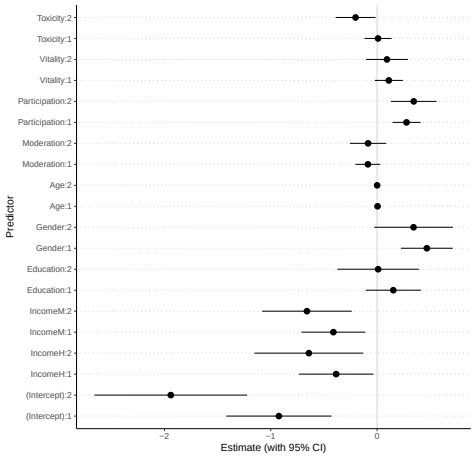
(a) Model 5b′



(b) Model 6a′



(c) Model 6b′



(d) Model 7b′

Figure D1: Generalized ordered logit models