Strategic legitimacy signals and journal indexing: Declarative ethical visibility in a semi-peripheral publishing system

Kemal Yayla
Department of Information and Document Management,
Faculty of Social Sciences and Humanities,
Izmir Kâtip Çelebi University, Izmir, TURKIYE
e-mail: kemal.yayla@ikcu.edu.tr
ORCID ID: 0000-0001-9064-611X

ABSTRACT

This study examines how ethical policy adoption functions as a strategic legitimacy mechanism in semi-peripheral publishing environments, focusing on 2,592 open-access journals hosted on Türkiye's DergiPark platform. Drawing on institutional isomorphism theory and the framework of strategic legitimacy, the research investigates whether declarative ethical visibility, the publicly accessible presence of ethics policy statements, predicts international indexing outcomes after controlling for institutional maturity and platform prominence. Using nested logistic regression models with extensive robustness checks across four classification algorithms and five resampling strategies, the analysis reports three principal findings. First, ethical signalling operates through "threshold logic": journals that reference international governance bodies have 31% higher odds of indexing (OR = 1.31, p < 0.001), yet cumulative policy intensity contributes negligibly, revealing an "Intensity Paradox." Second, a decoupling of declarative signals from operational practice is evident, as observable ethics manifestations (erratum publication) show no association with indexing (OR = 1.04, p > 0.05). Third, while ethical signalling confers a modest advantage (a 6–10 percentage-point increase), it remains structurally subordinate to the maturity (OR = 3.02) and visibility (OR = 4.05) determinants. These patterns are consistent with the view that ethics policies primarily function as symbolic compliance mechanisms that reduce exclusion risk without guaranteeing selection. The findings highlight critical vulnerabilities in gatekeeping regimes that rely on easily mimicked declarative signals, underscoring the need for process-oriented evaluation frameworks that assess operational practice alongside policy presence.

Keywords: Publication ethics; Strategic legitimacy; Institutional isomorphism; Journal indexing; Gatekeeping mechanisms; Research integrity.

INTRODUCTION

The advancement of scientific knowledge is fundamentally sustained by systematic inquiry, rigorous peer evaluation, and transparent dissemination mechanisms. As a core pillar of this process, academic publishing serves not only as a channel for knowledge transmission but also as a normative system that ensures the reliability and accountability of scholarly outputs (Hackett & Kelly, 2020). In parallel, publication ethics has become a rapidly expanding domain of scholarly attention and infrastructural investment, driven by increasing integrity

threats – most prominently industrialised misconduct and paper mills – which have intensified scrutiny of journals' editorial governance and publicly documented policies (Zhang et al., 2024).

Substantial empirical research further indicates that academic publishing has become increasingly intertwined with strategic and institutional performance imperatives. Van Dalen and Henkens (2012) conceptualise scholarly outputs as the instrumentalisation of publication, frequently enhancing symbolic capital within competitive hierarchies. Within this evaluative environment, ethical governance has shifted from a largely implicit professional expectation to an explicit legitimacy signal. A defining feature of illegitimate publishing, for example, is deviation from best editorial and publication practices and a lack of transparency (Grudniewicz et al., 2019). In turn, "transparency" and adherence to "best practices" have become central differentiators in contemporary boundary work between reputable and questionable journals, not least because such signals are increasingly legible to evaluative infrastructures.

These pressures are particularly acute in semi-peripheral academic systems, where reputational uncertainty and asymmetries in global knowledge circulation remain structurally persistent. Research on the Global South indicates that researchers in these regions often prioritise journal reputation and indexing status over open access, partly to avoid the stigma of predatory publishing (Nobes & Harris, 2023). As Marginson (2022) explains, this tendency reflects the structural hegemony of the global science system, in which Euro-American standards define the 'centre', compelling institutions in the semi-periphery to emulate these forms to gain validity. For journals operating in such contexts, aligning with internationally recognised transparency norms – such as the Principles of Transparency and Best Practice in Scholarly Publishing (Committee on Publication Ethics et al., 2022) - is not merely an administrative or symbolic task, but a strategic response to evaluative expectations that are increasingly standardised across the publishing field (Teixeira Da Silva & Moussa, 2024).

Türkiye represents a strategically significant case study in this context. Its academic publishing ecosystem has grown rapidly through the state-supported DergiPark platform (https://dergipark.org.tr/en/), which hosts nearly 3,000 journals and requires all hosted journals to be open access (Tamer et al., 2020). However, DergiPark's scale does not automatically result in proportional international indexing outcomes, raising a focused research question: to what extent is the declarative visibility of ethical policies associated with a journal's international indexing outcome?

In this study, international indexing is defined as inclusion in major global indexing systems that apply formal editorial selection and policy requirements, such as Scopus and the Web of Science Core Collection. Clarivate's Web of Science editorial selection process uses a single set of criteria to select journals for "editorial rigour and best practice". Both Clarivate and Scopus require candidate journals to provide a clear, publicly available statement of publication ethics and malpractice, considering the presence and accessibility of such statements an essential condition in the evaluation process (Lovakov & Teixeira Da Silva, 2025). These requirements are significant for semi-peripheral journals because they establish legitimacy partly through the public availability of governance signals, not solely through internal enforcement.

To ensure conceptual precision and avoid conflating policy display with policy enforcement, this study reframes "ethical policy adoption" as declarative ethical visibility – that is, the

publicly accessible presence of ethical policy statements and related governance signals on a journal's website. Drawing on the account of Long and Driscoll (2007) of ethical codes as instruments of "strategic legitimacy," this study posits that the public display of ethical policies can function as a quantifiable signal of legitimacy: a rational alignment mechanism through which journals signal conformity with widely institutionalised expectations. This framing treats declarative visibility as a signal of institutional isomorphism, not as evidence of operational compliance.

Against this theoretical backdrop, this study contributes to the scholarly communication literature by systematically quantifying the strategic value of these ethical signals in a semi-peripheral publishing ecosystem. Using nested logistic regression models on a comprehensive dataset of 2,592 journals, the study aims to disentangle the independent predictive power of declarative ethics from structural determinants, such as institutional maturity and platform visibility. Ultimately, this analysis seeks to determine whether the adoption of global ethical scripts serves as an effective mechanism for international integration or remains a symbolic performance detached from gatekeeping outcomes.

LITERATURE REVIEW

Academic publishing serves as an institutional mechanism to reduce information asymmetry between knowledge producers and consumers (Connelly et al., 2011). In this asymmetric market, peer-reviewed publications act as costly signals of otherwise unobservable quality, which, in principle, distinguish rigorous scholarship from lower-quality work (Bagues et al., 2019). The credibility of this signalling regime has historically depended on a baseline assumption of good faith participation by authors, reviewers, editors, and publishers (Resnik & Dinse, 2012). Recent evidence suggests that when this assumption is consistently violated, the publication system not only experiences occasional problems but also faces a broader legitimacy crisis (Richardson et al., 2025). Meanwhile, increasing publication pressures (Van Dalen & Henkens, 2012) have stretched the capacity of conventional peer review to validate signals at scale, necessitating an examination of how journals construct, display, and operationalise trustworthiness under contemporary incentive structures (Siler, 2020).

Organisational legitimacy is a central analytical lens for understanding these dynamics. Defined as a generalised perception or assumption that an entity's actions are desirable, proper, or appropriate within a socially constructed system of norms, values, beliefs, and definitions (Suchman, 1995), legitimacy can take three distinct forms. Long and Driscoll (2007) distinguish between moral legitimacy, arising from conformance to social values and obligations; cognitive legitimacy, stemming from the adoption of structures and behaviours that appear orthodox and taken for granted; and strategic legitimacy, which is instrumental – organisations use symbols and communication to shape perceptions of what is right and wrong. In the competitive world of academic publishing, journals often seek a balance between professional logics (e.g., knowledge contribution and disciplinary stewardship) and market logics such as financial viability, growth, and profit (Siler, 2020).

This distinction is significant because it clarifies why ethical governance may devolve into performative actions. Organisations may adopt ethical codes as facades – visible markers of conformity that provide reputational benefits without necessarily ensuring equivalent enforcement. Long and Driscoll (2007) describe this process as a form of institutional isomorphism: conformity to externally legitimised templates that prioritise symbolic alignment over substantive transformation. In scholarly publishing, the mere presence of

ethical policies on journal websites can therefore signal declarative ethical visibility rather than demonstrable operational control. Bryce et al. (2020) characterises this phenomenon as 'tick-box ethics', where regulatory compliance becomes a bureaucratic performance detached from genuine moral reflection. In this framework, journals may adopt policies not to guide conduct, but to satisfy the procedural checklists of external evaluators. Building on this insight, the present study treats the adoption of ethics policies – particularly in semi-peripheral ecosystems – as a rational signalling strategy shaped by global evaluative norms.

Quantitative indicators dominate the evaluative regime, further reinforcing strategic legitimacy. Reliance on metrics such as publication counts, journal rank, and the Impact Factor has reduced complex quality judgements to simplified, high-stakes proxies, enabling what is often described as a “metric culture” (Müller & de Rijcke, 2017; Niles et al., 2020). Researchers often perceive that career advancement depends on publishing a high volume of papers in prestigious outlets, even though these proxies imperfectly reflect epistemic contribution (Niles et al., 2020).

Under these conditions, the cost of failure rises while the perceived payoff of reputational signals increases, creating fertile ground for misconduct and strategic rule-bending. Empirical evidence from semi-peripheral science systems – including Russia and India – links high-powered incentives and rigid performance targets to higher retraction rates and integrity violations (Le Maux et al., 2019; Lovakov & Teixeira Da Silva, 2025). Sharma (2024), for example, attributes a significant share of retractions to fake peer review and plagiarism, illustrating how metric-driven incentives can erode research integrity by treating publication as an outcome to be achieved rather than a contribution to be validated.

As credibility signals become more valuable, they also become increasingly commodified. This commodification has shifted the integrity threat landscape from individual deviance to industrialised and coordinated fraud. Recent literature identifies “paper mills” as sophisticated operations that broker authorship, manufacture manuscripts, and manipulate editorial workflows to secure publication outcomes (Richardson et al., 2025). Importantly, such practices do not merely exploit weak journals; they actively reconfigure the conditions under which legitimacy signals are generated and recognised.

At the same time, the conventional binary between “legitimate” and “predatory” outlets has become analytically inadequate. Siler (2020) describes a spectrum of "grey" publishing, where journals may exhibit some features of peer review yet rely on aggressive solicitation and high acceptance rates. Predatory journals can be identified by systematic opacity and misrepresentation (Grudniewicz et al., 2019). In contrast, grey journals often imitate the outward signs of legitimacy – editorial boards, policy pages, and ethics declarations – making it more difficult to assess governance quality from surface-level signals alone. In this mimetic environment, distinguishing the visibility of ethics from the capacity for integrity enforcement becomes increasingly important.

In response to escalating integrity threats, transnational normative bodies – including the Committee on Publication Ethics (COPE) and the International Committee of Medical Journal Editors (ICMJE) – have assumed a more prominent role as producers of governance templates. These bodies provide procedural scripts for authorship standards, transparency expectations, and the management of misconduct. However, the effectiveness of voluntary governance remains contested. As Wessels and Visagie (2017) argue in the context of open science, successful implementation of such standards requires more than mandates; it demands resources and infrastructure to translate policy into practice. Moussa and Teixeira

da Silva (2023) note that formal affiliation with ethical initiatives (e.g., COPE membership) does not guarantee compliance, suggesting that ethical branding may serve as a reputational signal rather than a reliable indicator of enforcement capability (Teixeira da Silva & Moussa, 2024).

The juxtaposition of global governance scripts with local indexing regimes can produce a critical divergence. COPE's framework increasingly emphasises process integrity in the face of coordinated manipulation, for example, by outlining how third parties may secure acceptance for a fee by subverting the peer-review system. This approach operationalises integrity as a forensic capacity to detect and disrupt manipulation at the workflow level (Systematic Manipulation of the Publication Process, 2019).

In contrast, national indexing systems often prioritise compliance logics focused on subject protection and administrative permissibility. For example, the ethical criteria of TR Dizin (https://trdizin.gov.tr/) – Türkiye's national citation index – emphasise bureaucratic requirements, such as ensuring that ethics committee approval has been obtained for research involving humans or animals and that this approval is explicitly documented in the manuscript with details (date, number, etc.). At the journal level, TR Dizin requires the journal website state its ethical principles and include references to national and international standards, such as those of ICMJE and COPE (TR Dizin, 2022). Together with Clarivate (https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/web-of-science-core-collection/editorial-selection-process/journal-evaluation-process-selection-criteria/) and Scopus (https://www.elsevier.com/products/scopus/content/content-policy-and-selection#1-journals%E2%80%AF) guidelines, this indicates a shared gatekeeping logic in which ethical governance is assessed not only through editorial practice but also through the public visibility and accessibility of ethics declarations; in particular, when journals invoke third-party principles, they are expected to provide direct access to these documents (via on-page text or a functioning link). These differences create a semi-peripheral "compliance gap": journals can achieve high declarative visibility by meeting locally enforceable documentation requirements (e.g., permissions and approvals), while remaining structurally vulnerable to the sophisticated, networked manipulation problems targeted by global gatekeepers.

This divergence illustrates what Marginson (2022) describes as the tension between the 'global' scientific field and 'national' regulatory containers. While global gatekeepers demand resilience against industrialised fraud, national mechanisms often incentivise what Wessels and Visagie (2017) describe as compliance without cultural integration – adopting the 'mandates' of open science and ethics without the necessary infrastructure or incentives for effective enforcement. In such contexts, ethical visibility may prioritise indexation eligibility over robust, process-oriented integrity control, complicating the inference that "policy presence" equates to "governance effectiveness".

Indexing databases serve as primary "signal receivers" and gatekeepers in the publishing ecosystem. Inclusion in major indices (e.g., Scopus, Web of Science) confers cognitive legitimacy by transforming journal status into an externally validated credential (Clarivate, n.d.; Elsevier, n.d.). For journals in semi-peripheral environments, indexation is not merely a dissemination channel but an institutional threshold that shapes incentives for policy adoption and reputational signalling. As indexing bodies increasingly embed "best practice" requirements into selection criteria, isomorphism pressures intensify and journals face strong incentives to standardise outward-facing governance artefacts (e.g., ethics pages,

misconduct procedures, authorship policies) even when enforcement capacity does not scale accordingly (Asubiaro et al., 2024; Jiang et al., 2025; Tutuncu & Nasir, 2025). In this context, local indices such as TR Dizin may serve as intermediate gatekeepers that shape the form of compliance, particularly by prioritising explicit permission and approval documentation, while global indices increasingly emphasise integrity vulnerabilities related to systematic manipulation and reviewer identity verification. The result is a layered governance environment in which journals may optimise for the signals most salient to the relevant gatekeeper rather than for comprehensive integrity controls across the full spectrum of manipulation.

Existing scholarship provides robust documentation of ethical governance failures through retraction analyses (Hesselmann et al., 2017; Sharma, 2024; Wessels & Visagie, 2017) and typologies of predatory or deceptive publishing (Grudniewicz et al., 2019). However, fewer studies empirically examine ethics policy adoption as a strategic antecedent to indexing outcomes, particularly within non-Western, open-access publishing ecosystems where local and global governance scripts may diverge. Accordingly, a key unresolved question concerns how declarative ethical visibility – the public display and formalisation of ethics policies – relates to indexation success under multi-layer gatekeeping.

This study addresses that gap by analysing the relationship between comprehensive ethics policy adoption and indexing outcomes, framing policy adoption not necessarily as evidence of enforcement but as a strategic legitimacy signal shaped by isomorphism pressures and differentiated compliance logics across local and global governance regimes.

METHODS

This study examines the relationship between the declarative ethical policy frameworks of academic journals and their inclusion in citation indexes. Rather than assessing the substantive quality or enforcement of ethical practices, the research is based on the theoretical framework of "Declarative Ethics Visibility." This concept suggests that explicit references to national and international regulatory bodies in journal policies serve as observable signals of conformity to global open science norms. These citations are treated as proxies for institutional legitimacy, rather than direct indicators of enforcement fidelity. Within this framework, a descriptive-analytical research design is used to model the relationship between ethical signalling and the likelihood of journal indexation. Three research questions guide the analysis:

- To what extent is declarative ethical visibility associated with the likelihood of international indexing among open-access journals in a semi-peripheral publishing ecosystem?
- Does the relationship between ethical signalling and indexing outcomes follow a threshold logic, where binary policy presence matters more than cumulative policy intensity?
- How do declarative ethics signals compare with observable operational manifestations and structural determinants (namely institutional maturity and platform visibility) in predicting indexing outcomes?

Sample selection

The sample includes all active academic journals hosted on DergiPark, Türkiye's national open-access scholarly publishing platform. DergiPark provides web hosting and a comprehensive editorial management infrastructure, including submission, peer review, and publication workflows (Aslan, 2019). It also provides a no-fee DOI service for eligible journals; however, DOI assignment begins only after entry into the system, with no retroactive DOI minting (Tuglular et al., 2022).

To ensure methodological rigour and data integrity, journals with inactive URLs or inaccessible policy pages were systematically excluded from the dataset during the initial screening process. Data were collected over five days (2–7 February 2024) to minimise temporal variation in journal policies. Journals with inactive URLs or inaccessible policy pages were excluded during the initial screening. A deterministic, rule-based algorithm was used for data extraction to ensure reproducibility and transparency. A custom Python script used the requests and BeautifulSoup libraries to scrape and parse journal policy texts. A keyword-matching algorithm then identified two types of ethical markers: local regulatory references (such as "ULAKBİM," "YÖK," and national intellectual property laws) and global governance bodies (such as "COPE," "ICMJE," "OASPA," "WAME," "CSE," and "DOAJ"). This regex-based approach eliminates observer bias and ensures replicable measurement of explicitly stated affiliations.

Variable operationalisation

The dependent variable, Indexing Status, was defined as a binary outcome indicating whether a journal is included in at least one primary academic index. In line with standard bibliometric practice, "indexed" refers to inclusion in TR Dizin (Türkiye's national citation index, treated here as a national gatekeeper), the Emerging Sources Citation Index (ESCI), Scopus, the Science Citation Index Expanded (SCI-E), the Social Sciences Citation Index (SSCI), or the Arts & Humanities Citation Index (AHCI). Journals indexed only in discipline-specific or lower-tier databases were coded as non-indexed, ensuring that the dependent variable reflects international recognition through major gatekeeping systems rather than minimal listing.

To measure the extent and structure of ethical signalling, three complementary metrics were developed, based on established bibliometric literature. First, Global Ethics Intensity was calculated as a count variable representing the number of distinct international governance bodies explicitly cited in a journal's ethics policy statement. The algorithm searches for six major organisations: COPE, ICMJE, OASPA, WAME, CSE, and DOAJ. Following Fernandes et al. (2011), who used similar citation-counting approaches to measure alignment with the editorial ecosystem, this variable serves as a proxy for the degree of integration with global scholarly publishing norms. Second, Declarative Ethics Breadth was defined as a categorical variable classifying journal by their reference scope – citing only local regulations, only global frameworks, both, or neither (Broga et al., 2014). Previously conceptualised as "Ethical Diversity" (Fernandes et al., 2011), these variable measures the symbolic adoption of dual legitimacies and captures journals' strategic positioning within overlapping governance regimes. Third, Alignment Indicators consist of two binary variables denoting whether a journal references at least one local regulatory body and/or one international organisation. These indicators isolate the distinct effects of local and global alignment strategies.

Several control variables were included to account for structural advantages and confounding factors. Institutional maturity was measured using two indicators: journal age

and lifecycle stage. Journal age was calculated as the number of years since founding and log-transformed to address the right-skewed distribution. This variable controls for accumulated reputational capital and institutional stability, as older journals often have established networks and editorial expertise that facilitate their inclusion in indexes (Bean & Bernardi, 2005; Gu & Blackmore, 2017). Additionally, a journal lifecycle stage variable was derived from journal age (classified as New, Evolving, Well-developed, or Established) to capture non-linear maturity effects beyond continuous age, recognising that journal development follows distinct growth phases rather than linear increments (Mabe & Amin, 2001; Gu & Blackmore, 2017).

Operational Practice was measured by Annual Issue Frequency, defined as the average number of issues published per year. This metric reflects publication intensity and editorial capacity (Huang, 2016; Demeter et al., 2022), as journals with high-frequency publication schedules typically possess stronger organisational infrastructure. Erratum publication data served as a proxy for Observable Ethics Manifestation, with explicit emphasis on editorial vigilance and post-publication correction practices (Teixeira da Silva & Dobránszki, 2017). Margalida and Colomer (2015) argue that erratum publication demonstrates editorial transparency rather than negligence, supporting the distinction between these variables as declarative commitments and observable corrective actions. Journals without erratum records were coded as zero, acknowledging the difference between "confirmed zero errata" and data unavailability. From this, an Erratum Presence binary indicator and a log-transformed Erratum Intensity count were derived.

Platform Visibility was measured using log-transformed Views from DergiPark analytics. This variable captures algorithmic prominence and readership engagement, helping to prevent traffic-driven metrics from obscuring indexing status (Moradzadeh et al., 2023). Download counts were considered but excluded because they were nearly perfectly correlated with view counts. This approach was adopted to avoid severe multicollinearity, consistent with best practices in regression modelling (O'Brien, 2007). Visibility metrics were introduced in the final stage of modelling to address potential post-treatment bias, as indexed journals may attract higher traffic due to database exposure.

Statistical modelling

The analytical strategy aimed to balance coefficient-level interpretability with methodological robustness, employing a two-phase approach that combined parametric inference with algorithmic validation. In the primary analytical phase, nested logistic regression models were estimated following a 'Theory-Driven Pathway' grounded in Signalling Theory (Connelly et al., 2011; Tiokhin et al., 2021). Rather than arbitrarily including variables, this hierarchical framework dictated the sequential entry of predictor blocks to reflect the theoretical accumulation of legitimacy within the publishing ecosystem.

The modelling sequence began by establishing a structural baseline (Block 1) using Journal Age and Frequency. This specification controlled for the screening role of established reputation signals in information-rich environments (Van Dalen & Henkens, 2005). Next, declarative ethics signals (Block 2) were introduced to quantify the incremental predictive value of strategic signalling. Consistent with DiMaggio and Powell (1983), this block isolated variance attributable to institutional isomorphism and symbolic compliance, distinct from the structural advantages modelled in the first block. The sequence then examined observable manifestations (Block 3) to differentiate between declarative commitments and operational practice. It concluded with platform visibility (Block 4) to address potential

endogeneity, ensuring that the estimated effects of ethical signals were robust to algorithmic prominence.

To benchmark predictive performance and assess robustness to sequential ordering, alternative specifications were also estimated, including a predictive pathway (Maturity → Visibility → Ethics) and a conservative pathway that excluded visibility entirely to provide lower-bound estimates unconfounded by reverse causality. Model fit was evaluated using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), McFadden's Pseudo-R^2, and Likelihood Ratio Tests.

To explore potential nonlinear relationships and validate findings across modelling approaches, four classification algorithms were employed: Logistic Regression, Decision Tree, Random Forest, and XGBoost (Extreme Gradient Boosting). This multi-algorithm strategy addressed the possibility that observed associations were artefacts of linear model assumptions (Breiman, 2001). Because the dataset had an uneven number of classes, sensitivity analyses were conducted using four resampling methods: SMOTE (Synthetic Minority Over-sampling Technique) (Chawla et al., 2002), ADASYN (Adaptive Synthetic Sampling) (He et al., 2008), Random Undersampling, and a baseline with no resampling.

All models were validated using 5-fold stratified cross-validation to prevent overfitting and ensure generalisability. Performance was evaluated using both ROC-AUC (Area Under the Receiver Operating Characteristic Curve) (Fawcett, 2006) and PR-AUC (Precision-Recall Area Under the Curve) (Davis & Goadrich, 2006), the latter providing a more informative assessment for imbalanced datasets. This dual-phase strategy confirms that observed patterns generalise across multiple algorithmic specifications and resampling conditions.

RESULTS

Sample characteristics

The analytical sample comprises 2,592 journals from the DergiPark platform, of which 754 (29.1%) were indexed in at least one major academic database (TR Dizin, ESCI, Scopus, SCI-E, SSCI, or AHCI), while 1,838 (70.9%) were unindexed. This 71:29 distribution reflects substantial filtering in semi-peripheral publishing environments and informed the decision to use both standard logistic regression and SMOTE-based sensitivity analyses.

Table 1 presents descriptive statistics and bivariate associations between predictor variables and indexing status. Point-biserial correlations indicate that journal age ($r = 0.407$, Cohen's $d = 0.980$) and platform visibility (log_views: $r = 0.356$, Cohen's $d = 0.838$) have significant effects, while declarative ethics variables show small-to-medium associations (global_ethics_intensity: $r = 0.118$, Cohen's $d = 0.261$). Ethics variables correlate weakly with maturity and visibility metrics ($r < 0.25$), indicating that ethics adoption patterns operate independently of structural advantages. These supports treating them as strategic legitimacy signals rather than mere proxies for journal age or platform prominence.

Table 1: Descriptive statistics and bivariate associations with indexing status

Variable	Mean	SD	r_pb	Cohen's d	Effect Size
global_ethics_intensity	0.78	0.53	0.118***	0.261	Small-Medium
erratum_intensity	0.15	0.36	0.123***	0.272	Small-Medium
log_journal_age	2.35	0.77	0.407***	0.980	Large
issues_per_year	2.44	0.87	0.171***	0.382	Small-Medium
log_views	9.94	3.38	0.356***	0.838	Large

Note: r_pb = point-biserial correlation. *** $p < 0.001$. Effect sizes follow Cohen's conventions (small d ≈ 0.2, medium d ≈ 0.5, large d ≈ 0.8). Variables shown are those included in final models; log_downloads and engagement_ratio excluded due to multicollinearity with log_views (r = 0.99 and r = 0.67 respectively).

Declarative ethical visibility and indexing status

Examination of declarative ethics patterns reveals systematic differences by indexing status. Among journals citing both local and global ethics sources (ethics breadth), 38.5% were indexed, compared with 30.3% citing only global sources ($\chi^2 = 12.13$, $p = 0.002$, Cramér's V = 0.076). This pattern supports the view that journals signalling dual legitimacy by satisfying both national compliance requirements and international governance norms position themselves favourably across layered gatekeeping systems. Journal lifecycle stage shows even steeper gradients: indexing rates rise from 3.8% (new journals, 0–1 years) to 57.5% (established journals, over 15 years), underscoring cumulative advantage mechanisms in which institutional longevity confers reputational capital that newer journals cannot replicate through policy adoption alone ($\chi^2 = 353.96$, $p < 0.001$, Cramér's V = 0.370, as shown in Figure 1).

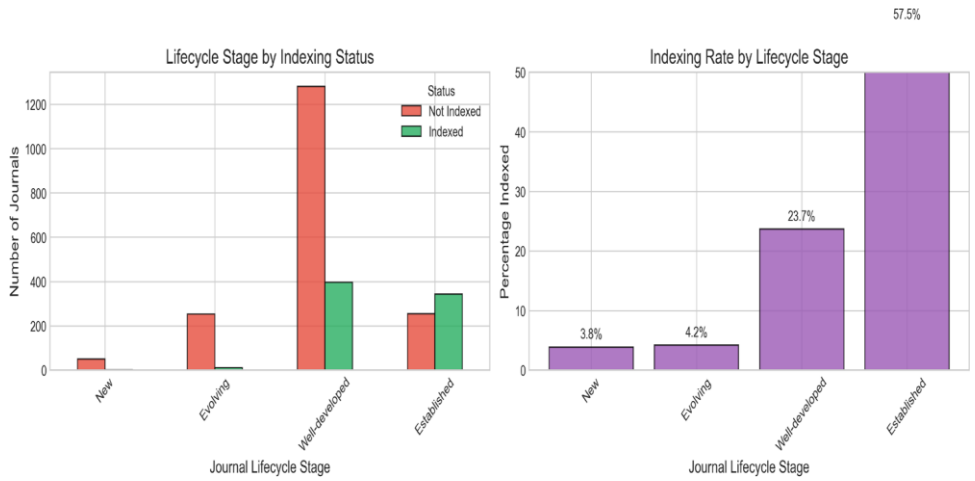

Figure 1: Indexing rate by journal lifecycle stage (N = 2,592)

The empirical analysis used a series of nested logistic regression models, structured along a 'Theory-Driven Pathway', to isolate the independent predictive contribution of declarative ethics signals from structural confounders. This hierarchical sequence comprised four theoretically distinct blocks: (M1) an Institutional Maturity Baseline to account for accumulated reputational capital; (M2) Declarative Ethics Signals to quantify the strategic legitimacy effect; (M3) Observable Ethics Manifestations to test operational decoupling; and (M4) Platform Visibility to control for potential endogeneity. This stepwise architecture isolates variance attributable to ethical signalling from that explained by institutional

longevity or post-indexation traffic amplification. Feature standardisation was applied across all models to ensure coefficient comparability.

Table 2 shows incremental improvements in model fit and predictive accuracy across the nested sequence. The baseline Structural Maturity model (M1) yielded a cross-validated ROC-AUC of 0.778 [95% CI: 0.761, 0.797], indicating that institutional age and publication frequency alone account for a substantial share of the variance (Pseudo-R^2 = 0.150). Introducing Declarative Ethics Signals in Model M2 produced a statistically significant improvement in model fit (Likelihood Ratio χ^2 = 109.63, df = 3, $p < 0.001$), raising the ROC-AUC to 0.795 and Pseudo-R^2 to 0.185. This gain emerged after controlling for maturity, confirming that ethical signals convey independent legitimacy information orthogonal to institutional longevity. In contrast, adding Observable Ethics Manifestations (M3) provided only marginal explanatory gain (LR χ^2 = 7.16, p = 0.028), supporting the premise that gatekeepers prioritise declarative governance frameworks over retrospective corrective actions. Finally, including Platform Visibility (M4) produced the most pronounced incremental gain (LR χ^2 = 233.20, $p < 0.001$), elevating the ROC-AUC to 0.842 (Pseudo-R^2 = 0.262), underscoring the dominant – albeit potentially endogenous – role of digital prominence in the indexing process.

Table 2: Nested logistic regression models and likelihood ratio tests

Model	Features (n)	CV AUC [95% CI]	Pseudo-R^2	AIC	LR χ^2	df	p-value
M1: Maturity	3	0.778 [0.761, 0.797]	0.150	2664.7	—	—	—
M2: +Declarative Ethics	6	0.795 [0.779, 0.814]	0.185	2561.1	109.63	3	<0.001***
M3: +Observable Ethics	8	0.796 [0.781, 0.816]	0.187	2557.9	7.16	2	0.028*
M4: +Visibility	9	0.842 [0.826, 0.859]	0.262	2326.7	233.20	1	<0.001***

Note: CV AUC = cross-validated ROC-AUC from 5-fold stratified validation. LR χ^2 = likelihood ratio chi-squared testing improvement vs. previous model. *** $p < 0.001$, * $p < 0.05$.

Table 3 presents detailed coefficients for the final model (M4), which includes all predictor blocks. Visibility (log_views: β = 1.400, OR = 4.05, 95% CI [2.36, 7.46]) was the strongest predictor, followed by journal age (β = 1.105, OR = 3.02, 95% CI [2.38, 3.93]). Both global (β = 0.271, OR = 1.31, 95% CI [1.15, 1.49]) and local (β = 0.236, OR = 1.27, 95% CI [1.18, 1.37]) ethical references were statistically significant among declarative ethics variables. Thus, at mean predictor values (baseline probability ≈ 29%), the probability of indexing increased by about 6–10 percentage points.

Table 3: Final model coefficients with confidence intervals (Model M4)

Feature	β	Odds ratio	95% CI	p-value
log_views	1.400	4.05	[2.36, 7.46]	<0.001***
log_journal_age	1.105	3.02	[2.38, 3.93]	<0.001***
has_global_ethics	0.271	1.31	[1.15, 1.49]	<0.001***
has_local_ethics	0.236	1.27	[1.18, 1.37]	<0.01**
global_ethics_intensity	0.110	1.12	[1.00, 1.28]	0.056
lifecycle_numeric	-0.135	0.87	[0.71, 1.06]	0.167
has_erratum	0.038	1.04	[0.96, 1.12]	0.312

Note: All features standardized. *** $p < 0.001$, ** $p < 0.01$. Coefficients represent standardized effects.

Threshold logic versus policy intensity

Notably, global ethics intensity (the number of distinct international organisations cited) had a non-significant coefficient ($\beta = 0.110$, OR = 1.12, $p > 0.05$), indicating that among journals citing at least one organisation, citing additional organisations provides minimal marginal benefit. This pattern supports a threshold signalling model: gatekeepers distinguish journals that signal awareness of international governance norms from those that do not, but do not reward cumulative adoption diversity proportionally. Erratum presence ($\beta = 0.038$, OR = 1.04, $p > 0.05$) contributed negligible explanatory power, suggesting that observable manifestations of editorial vigilance do not independently predict indexing after controlling for declarative signals.

Examining coefficient stability across nested models shows that the effects of declarative ethics remain significant, with only slight reductions when visibility is added (has_global_ethics: $\beta = 0.430$ in M2 $\rightarrow$ 0.183 in M4; has_local_ethics: $\beta = 0.262 \rightarrow 0.244$). This suggests that ethics signals operate through pathways only partially linked to platform prominence, rather than solely to high traffic. Alternative model pathways – Predictive (visibility introduced before ethics) and Conservative (visibility excluded entirely)—confirm that declarative ethics retain independent significance regardless of variable sequencing (see Table 4).

Robustness and validation analyses

Extensive robustness checks across four dimensions confirm that the observed associations generalise beyond specific modelling choices. First, 5-fold stratified cross-validation produced a mean ROC-AUC of 0.839 ± 0.012 and a Precision-Recall AUC of 0.653 ± 0.027, providing a more conservative estimate that accounts for class imbalance. Second, benchmarking logistic regression against three alternative classifiers (Decision Tree, Random Forest, and XGBoost) confirmed that the parametric model outperformed all alternatives (logistic AUC = 0.838 vs 0.799–0.815 for ensemble methods), indicating that relationships between predictors and indexing outcomes are approximately linear on the logit scale rather than reflecting complex nonlinear interactions (see Table 5).

Resampling strategy sensitivity analysis evaluated five approaches to address the 71:29 class imbalance: no resampling (baseline), SMOTE, ADASYN, random undersampling, and SMOTE-Tomek. ROC-AUC remained remarkably stable across strategies (0.821–0.830, $\Delta = 0.009$), confirming that the observed associations are not artefacts of class balance handling (see Table 6). Feature set ablation studies showed that declarative ethics variables provide independent information: models using only declarative ethics achieved barely above-chance performance (AUC = 0.595), while adding ethics signals to control variables (maturity + visibility) produced a significant improvement ($\Delta AUC = +0.013$, $p < 0.01$), confirming that ethics signals capture variance orthogonal to structural advantages (see Table 7). A random-seed stability analysis across 10 cross-validation initialisations yielded a mean AUC of 0.834 ± 0.003 (relative variation < 0.4%), indicating that results are robust to data partitioning choices.

Table 4: Alternative model pathways

Panel A: Conservative Pathway (Visibility Excluded)			
Model	**Feature**	**β**	**OR**
M1_Maturity	log_journal_age	0.904	2.469
M1_Maturity	lifecycle_numeric	0.081	1.084
M1_Maturity	issues_per_year	0.202	1.224
M2_Declarative_Ethics	log_journal_age	0.986	2.682
M2_Declarative_Ethics	has_local_ethics	0.262	1.300
M2_Declarative_Ethics	has_global_ethics	0.430	1.537
M2_Declarative_Ethics	global_ethics_intensity	-0.008	0.992
M3_Observable_Ethics	log_journal_age	0.974	2.649
M3_Observable_Ethics	has_local_ethics	0.258	1.294
M3_Observable_Ethics	has_global_ethics	0.423	1.526
M3_Observable_Ethics	has_erratum	0.049	1.051
Panel B: Predictive Pathway (Visibility → Ethics)			
Model	**Feature**	**β**	**OR**
M1_Maturity	log_journal_age	0.904	2.469
M2_Visibility	log_journal_age	0.819	2.269
M2_Visibility	log_views	1.415	4.116
M3_Declarative_Ethics	log_journal_age	0.857	2.357
M3_Declarative_Ethics	log_views	1.331	3.784
M3_Declarative_Ethics	has_local_ethics	0.244	1.276
M3_Declarative_Ethics	has_global_ethics	0.184	1.202
M4_Observable_Ethics	log_views	1.325	3.762
M4_Observable_Ethics	has_local_ethics	0.244	1.276
M4_Observable_Ethics	has_erratum	-0.013	0.987

Note: β = unstandardised logistic regression coefficient; OR = odds ratio. Conservative pathway excludes log_views entirely. Predictive pathway introduces log_views in M2, followed by declarative ethics in M3.

Table 5: Algorithm comparison (5-fold cross-validation)

Algorithm	**Accuracy**	**Precision**	**Recall**	**F1**	**ROC-AUC**
Logistic Regression	0.763	0.567	0.796	0.661	0.838
Decision Tree	0.703	0.491	0.599	0.539	0.671
Random Forest	0.761	0.579	0.664	0.618	0.815
XGBoost	0.753	0.569	0.634	0.599	0.799

Note: All metrics are mean values from 5-fold stratified cross-validation. ROC-AUC = Area Under the Receiver Operating Characteristic Curve.

Table 6: Resampling sensitivity analysis

Strategy	Accuracy	Precision	Recall	F1	ROC-AUC
None (baseline)	0.785	0.647	0.574	0.607	0.827
SMOTE	0.774	0.592	0.719	0.649	0.829
ADASYN	0.769	0.580	0.749	0.653	0.830
Undersampling	0.759	0.563	0.765	0.648	0.821
SMOTE-Tomek	0.781	0.603	0.735	0.661	0.829

Note: SMOTE = Synthetic Minority Over-sampling Technique; ADASYN = Adaptive Synthetic Sampling. All metrics from 5-fold stratified cross-validation.

Table 7: Feature set ablation

Feature Set	N	Accuracy	Precision	Recall	F1	ROC-AUC
Declarative only	3	0.497	0.327	0.678	0.432	0.595
Controls only	4	0.755	0.562	0.728	0.634	0.816
Core	8	0.774	0.592	0.719	0.649	0.829
Extended	11	0.782	0.609	0.704	0.653	0.833

Note: Declarative only = has_global_ethics, has_local_ethics, global_ethics_intensity. Controls only = log_journal_age, lifecycle_numeric, issues_per_year, log_views. Core = controls + declarative ethics + erratum. Extended = core + additional interaction terms.

Classification performance and practical significance

The final model achieved a test set accuracy of 0.784, with precision of 0.592, recall of 0.834, F1 score of 0.692, and ROC-AUC of 0.872. The confusion matrix revealed asymmetric error patterns: false positives (n = 87, 23.6% of non-indexed) substantially outnumbered false negatives (n = 25, 16.6% of indexed), reflecting the precision-recall trade-off inherent in SMOTE-based minority class detection (see Figure 2).

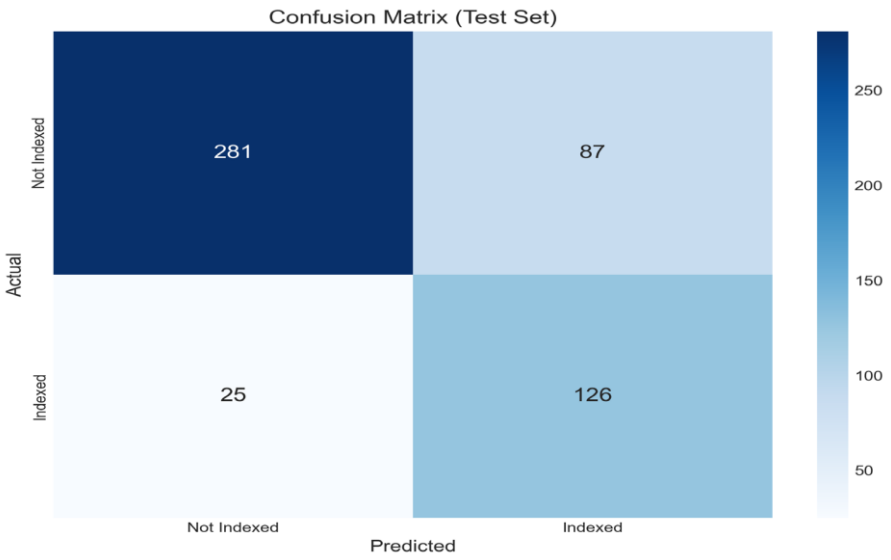

Figure 2: Confusion matrix for the final model on the held-out test set

Declarative signals, operational manifestations, and structural determinants

Figure 3 shows the evolution of the coefficients across the nested model sequence, illustrating three key patterns. First, log_journal_age remains remarkably stable (β ≈ 0.85–0.90 across all models), confirming that maturity effects persist independently of ethics adoption and visibility. Second, the coefficients for declarative ethics (has_global_ethics, has_local_ethics) remain relatively stable from M2 to M4, with only a slight decrease when visibility is added. This persistence indicates that ethics signals operate through pathways partially independent of platform prominence – consistent with the strategic legitimacy interpretation that declarative visibility acts as a direct signal to gatekeepers rather than merely correlating with traffic advantages. Third, global_ethics_intensity has coefficients very close to zero across all specifications (β ≈ -0.01 to 0.06), supporting the idea that binary presence matters, but cumulative diversity does not.

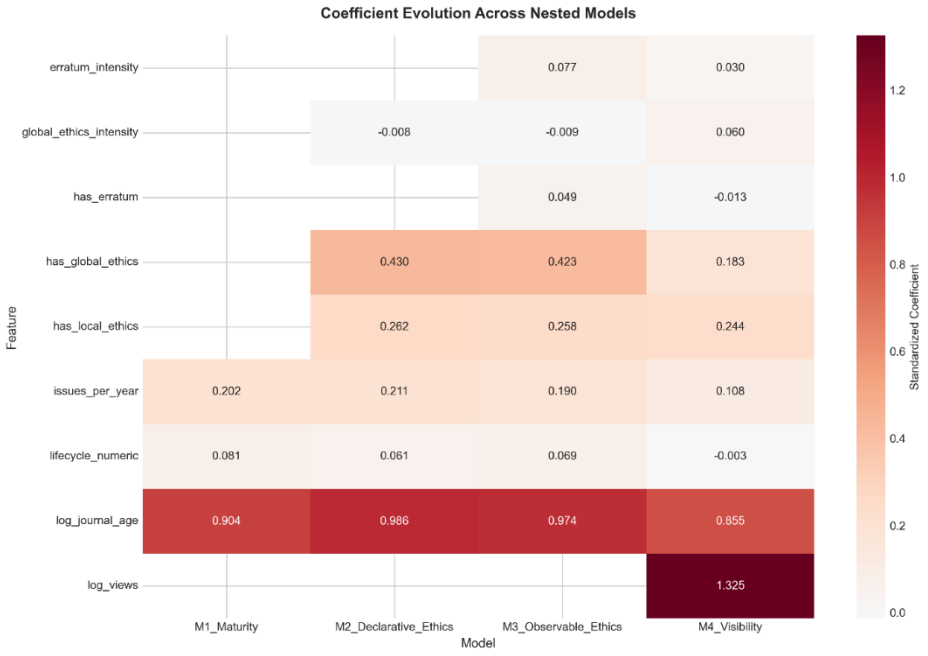

Figure 3: Coefficient evolution across nested models

XGBoost feature importance rankings (see Figure 4) support the logistic regression hierarchy: visibility metrics are dominant (75.7% combined importance), followed by journal age (9.8%), with declarative ethics variables together contributing 15.6%. The consistency between parametric coefficients and non-parametric importance scores (Spearman $\rho = 0.86$, $p < 0.01$) indicates that the hierarchical structure of predictive contributions is robust to algorithmic choice.

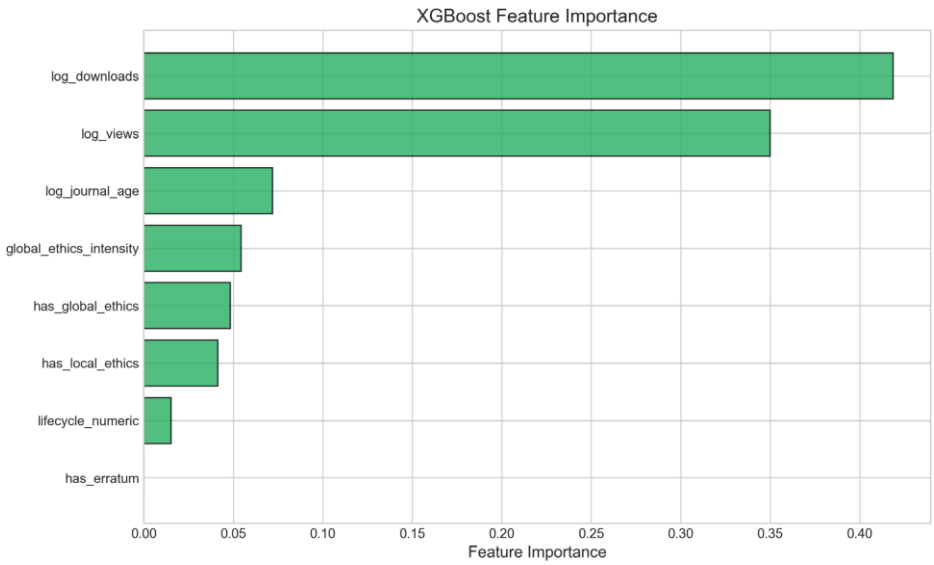

Figure 4: Ranking of XGBoost feature importance scores

DISCUSSION

Thiis study examined whether the adoption of ethical policies in Türkiye's open-access academic publishing ecosystem serves as a normative commitment to integrity or as a strategic signal of legitimacy. The empirical results, robust across multiple algorithmic specifications, support a "Strategic Legitimacy" interpretation. The study finds that declarative ethical visibility is a significant predictor of indexation; however, this relationship follows a threshold logic rather than continuous engagement and is essentially independent of operational enforcement. These findings highlight the structural tensions inherent in semi-peripheral publishing systems.

A primary theoretical contribution of this study is the identification of an "Intensity Paradox." The models show that the binary presence of ethical references (global or local) is a significant predictor of indexation (OR $\approx$ 1.30), whereas the cumulative intensity of these signals provides only marginal benefit (OR = 1.12, ns). This empirical pattern supports Tiokhin et al.'s (2021) evolutionary framework of "honest signalling" in competitive academic markets. Tiokhin et al. (2021) argue that as "competition for attention" intensifies, actors evolve conspicuous signals optimised for legibility rather than nuance.

In the DergiPark ecosystem, journals appear to follow what Siler (2020) describes as "mimicry" in the "grey zones" of publishing. By displaying visible symbols of legitimacy – such as a COPE reference or a standard ethics statement – journals satisfy gatekeepers' heuristic screening criteria without necessarily investing in substantive compliance. This heuristic screening likely arises from the institutionalisation of criteria identified by Astaneh and Masoumi (2018). In their seminal analysis, they observed that Scopus was unique in requiring explicit publication ethics statements aligned with COPE or ICMJE guidelines, hypothesising that such mandates would compel journals to adhere to ethical standards. However, our findings reveal an unintended consequence of this systemic shift: while the mandate succeeded in universalising the declarative presence of these policies (creating the threshold), it incentivised journals to satisfy the algorithm's binary check rather than distinguishing between substantive policy depth and superficial adoption.

The null association between indexing outcomes and observable ethical manifestations (erratum publication) reveals a critical "compliance gap." Despite the growing threat of "industrialised fraud" noted by Richardson et al. (2025), which demands rigorous post-publication correction mechanisms, our results indicate that gatekeepers do not penalise journals for a lack of observable corrections. This finding confirms the prevalence of "tick-box ethics," where regulatory compliance becomes a bureaucratic performance detached from implementation fidelity (Bryce et al., 2020). Turkish journals face dual isomorphic pressures: they must satisfy local bureaucratic requirements (TR Dizin's documentation focus) and global normative expectations (COPE alignment). While they achieve high declarative visibility to navigate these pressures, the disconnect from erratum data suggests that this compliance remains largely performative. As Le Maux et al. (2019) argue, when incentives are misaligned, actors prioritise the metrics that are rewarded (indexing status) over those that are costly (policing misconduct). This decoupling is not exclusive to the semi-periphery. Aubert Bonn et al. (2017) analysed research integrity guidance at top European universities (LERU) and found significant variability in accessibility and content, suggesting that policies often reflect institutional preferences rather than a uniform enforcement standard. However, a critical distinction emerges in the Turkish context. While European institutions may develop diverse guidelines to foster internal research culture, Turkish journals appear to adopt standardised ethics statements primarily as a survival mechanism to navigate global gatekeeping thresholds.

While ethical signalling offers a statistically significant advantage, it remains secondary to the structural determinants of Institutional Maturity (OR = 3.02) and Platform Visibility (OR = 4.05). Marginson's (2022) analysis of the global science system, in which Euro-American dominance permanently constrains emerging countries, aligns with this hierarchy. For journals in the semi-periphery, ethical policies serve as a necessary "entry ticket" – a mechanism of institutional isomorphism that enables them to gain legitimacy within a system defined by the 'centre.' However, these signals cannot fully compensate for the cumulative reputational capital represented by journal age or the algorithmic prominence driven by traffic. This supports Berghauser et al.'s (2025) observation that researchers and institutions in these regions often lack the structural support to move beyond mere mimicry and develop autonomous, robust integrity cultures.

The robustness of these findings across diverse algorithms confirms that strategic signalling is a systemic feature of the current indexing landscape. However, this stability results in a systemic vulnerability. If indexing systems continue to rely on declarative signals (threshold logic) rather than operational indicators, they risk falling victim to the "mimicry" that Grudniewicz et al. (2019) warned about. As Teixeira da Silva and Moussa (2024) critique, when "Best Practice" principles are commodified into branding tools without verification, they lose their screening utility. Therefore, for gatekeepers (Scopus, WoS, TR Dizin), these findings suggest an urgent need to transition from evaluating declarative visibility (what journals say) to operational practice (what journals do). Without this shift, the "metric tide" will continue to incentivise the construction of flawless ethical facades that mask systemic vulnerabilities to fraud.

Interpreting these findings is subject to several limitations. First, the cross-sectional design precludes definitive causal inference. Although coefficient stability across nested models suggests a robust structural relationship, reverse causality – where indexation requirements drive policy adoption – cannot be ruled out without longitudinal data. Second, operationalising "declarative visibility" through automated keyword detection captures the presence of signals rather than implementation fidelity. This proxy is useful for measuring

signalling intent; however, it remains vulnerable to "adversarial mimicry", in which journals use compliance language to appease algorithmic scrapers without enforcing it. Third, despite controlling for institutional maturity and visibility, unobserved heterogeneity (e.g., informal editorial networks or disciplinary prestige) may partially confound the observed associations. Finally, the focus on Türkiye's DergiPark ecosystem is a boundary condition. While the identified mechanisms of strategic legitimacy and dual compliance likely extend to other semi-peripheral contexts, the magnitudes of their effects may vary across national regulatory frameworks. Future scholarship employing quasi-experimental designs is needed to validate the universality of this "threshold logic" across diverse geographies.

CONCLUSIONS

This study demonstrates that declarative ethical visibility in Türkiye's open-access publishing ecosystem primarily serves as a strategic legitimacy signal, operating through a "threshold logic" in which the binary presence of a policy reduces exclusion risk without rewarding policy sophistication. The empirical separation of declarative signals from observable manifestations substantiates that ethics adoption currently functions as symbolic compliance rather than operational enforcement. Furthermore, the modest magnitude of these effects relative to structural advantages underscores the persistent centrality of cumulative reputational capital in semi-peripheral gatekeeping.

These repeatedly tested patterns show that evaluative infrastructures have systemic weaknesses when they place greater value on mimicked declarative signals than on process-oriented capacity. Addressing these vulnerabilities requires fundamental reforms: moving from policy-presence checklists to behavioural indicators, implementing tiered evaluation frameworks, and recognising the limits of declarative transparency. While the identified mechanisms of strategic legitimacy and threshold signalling likely extend to other semi-peripheral publishing contexts (e.g., Latin America, Southeast Asia, Eastern Europe), the magnitudes of the effects may vary across national regulatory frameworks, language policies, and infrastructure maturity. Comparative studies across multiple national platforms would help establish the boundary conditions of these findings. Ultimately, closing the compliance gap requires aligning the legitimacy signals gatekeepers evaluate with the operational capacities that substantively ensure research integrity – a shift vital to driving authentic governance transformation in globally stratified publishing systems.

ACKNOWLEDGMENTS

This study has not received any special grants from public, commercial, or non-profit organisations.

The author acknowledges the use of Claude (Anthropic), an AI-assisted writing tool, in preparing this manuscript. Claude was used solely to improve the clarity, language, and grammatical accuracy of the text. The author remains fully responsible for all content and confirms that no AI-generated text was accepted without critical human review, in accordance with COPE guidelines and the ethical standards of the journal.

CONFLICT OF INTEREST

The author has no relevant competing interests to declare in relation to the content of this article.

ETHICAL APPROVAL AND INFORMED CONSENT STATEMENTS

This study relies exclusively on publicly available secondary data sources (journal websites and open-access databases) and did not involve human participants, identifiable personal data, or experimental procedures. Therefore, formal ethical review was not required.

REFERENCES

Aslan, A. (2019). *DergiPark akademik. Acta Medica Alanya, 3*(3), 205–206. https://doi.org/10.30565/medalanya.632462.

Astaneh, B., & Masoumi, S. (2018). From paper to practice; indexing systems and ethical standards. *Science and Engineering Ethics*, *24*(2), 647–654. https://doi.org/10.1007/s11948-017-9899-x.

Asubiaro, T., Onaolapo, S., & Mills, D. (2024). Regional disparities in Web of Science and Scopus journal coverage. *Scientometrics, 129*(3), 1469–1491. https://doi.org/10.1007/s11192-024-04948-x.

Aubert Bonn, N., Godecharle, S., & Dierickx, K. (2017). European universities' guidance on research integrity and misconduct: Accessibility, approaches, and content. *Journal of Empirical Research on Human Research Ethics*, *12*(1), 33–44. https://doi.org/10.1177/1556264616688980.

Bagues, M., Sylos-Labini, M., & Zinovyeva, N. (2019). A walk on the wild side: 'Predatory' journals and information asymmetries in scientific evaluations. *Research Policy*, *48*(2), 462–477. https://doi.org/10.1016/j.respol.2018.04.013.

Berghaeuser, H., Prass, M., & Lindner, R. (2025). Researchers' practice and perception of research ethics and the role of institutional support: Insights from a pan-European researcher survey. *Journal of Academic Ethics*. *23*, 717–738 https://doi.org/10.1007/s10805-024-09590-z.

Breiman, L. (2001). Random forests. Machine Learning, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Broga, M., Mijaljica, G., Waligora, M., Keis, A., & Marusic, A. (2014). Publication ethics in biomedical journals from countries in central and eastern Europe. *Science and Engineering Ethics*, *20*(1), 99–109. https://doi.org/10.1007/s11948-013-9431-x.

Bryce, C., Dowling, M., & Lucey, B. (2020). The journal quality perception gap. *Research Policy, 49*(5), 103957. https://doi.org/10.1016/j.respol.2020.103957.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953.

Clarivate. (n.d.). *Web of Science core collection: Editorial selection process.* https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/web-of-science-core-collection/editorial-selection-process/journal-evaluation-process-selection-criteria/.

Committee on Publication Ethics, Directory of Open Access Journals, Open Access Scholarly Publishing Association, & World Association of Medical Editors. (2022). *Principles of*

transparency and best practice in scholarly publishing. https://doi.org/10.24318/cope.2019.1.12

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67. https://doi.org/10.1177/0149206310388419.

Davis, J., & Goadrich, M. (2006). *The relationship between precision-recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning* (pp. 233–240). ACM. https://doi.org/10.1145/1143844.1143874.

Demeter, M., Pelle, V., Mikulás, G., & Goyanes, M. (2022). Higher quantity, higher quality? Current publication trends of the most productive journal authors on the field of communication studies. *Publishing Research Quarterly*, *38*(3), 445–464. https://doi.org/10.1007/s12109-022-09893-2.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147–160. https://doi.org/10.2307/2095101.

Elsevier. (n.d.). *Scopus content policy and selection.* https://www.elsevier.com/products/scopus/content/content-policy-and-selection.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Fernandes, M. R., Antonelli Monteiro Queiroz, M. C. C., De Moraes, M. R., Barbosa, M. A., & Lima Sousa, A. L. (2011). Ethical standards adopted by Brazilian journals of medical specialties. *Revista da Associação Médica Brasileira (English Edition), 57*(3), 264–268. https://doi.org/10.1016/S2255-4823(11)70057-1.

Grudniewicz, A., Moher, D., Cobey, K. D., Bryson, G. L., Cukier, S., Allen, K., Ardern, C., Balcom, L., Barros, T., Berger, M., Ciro, J. B., Cugusi, L., Donaldson, M. R., Egger, M., Graham, I. D., Hodgkinson, M., Khan, K. M., Mabizela, M., Manca, A., … Lalu, M. M. (2019). *Predatory journals: No definition, no defence. Nature*, *576*(7786), 210–212. https://doi.org/10.1038/d41586-019-03759-y.

Gu, X., & Blackmore, K. (2017). Characterisation of academic journals in the digital age. *Scientometrics, 110*(3), 1333–1350. https://doi.org/10.1007/s11192-016-2219-4.

Hackett, R., & Kelly, S. (2020). Publishing ethics in the era of paper mills. *Biology Open, 9*(10), bio056556. https://doi.org/10.1242/bio.056556.

He, H., Yang, B., Garcia, E. A., & Li, S. (2008). *ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In 2008 IEEE International Joint Conference on Neural Networks* (pp. 1322–1328). IEEE. https://doi.org/10.1109/IJCNN.2008.4633969.

Hesselmann, F., Graf, V., Schmidt, M., & Reinhart, M. (2017). The visibility of scientific misconduct: A review of the literature on retracted journal articles. *Current Sociology*, *65*(6), 814–845. https://doi.org/10.1177/0011392116663807.

Huang, D. (2016). Positive correlation between quality and quantity in academic journals. *Journal of Informetrics*, *10*(2), 329–335. https://doi.org/10.1016/j.joi.2016.02.002.

Jiang, X., Li, M., & Zhang, Z. (2025). 'Bridging borders in academia': The motivations, barriers, and strategies of Global North scholars in Chinese higher education research. *Higher Education*. https://doi.org/10.1007/s10734-025-01558-1.

Le Maux, B., Necker, S., & Rocaboy, Y. (2019). Cheat or perish? A theory of scientific customs. *Research Policy, 48*(9), 103792. https://doi.org/10.1016/j.respol.2019.05.001.

Long, B. S., & Driscoll, C. (2007). Codes of ethics and the pursuit of organizational legitimacy: Theoretical and empirical contributions. *Journal of Business Ethics*, *77*(2), 173–189. https://doi.org/10.1007/s10551-006-9307-y.

Lovakov, A., & Teixeira Da Silva, J. A. (2025). Scientometric indicators in research evaluation and research misconduct: Analysis of the Russian university excellence initiative. *Scientometrics, 130*, 1813–1829. https://doi.org/10.1007/s11192-025-05269-3.

Mabe, M., & Amin, M. (2001). Growth dynamics of scholarly and scientific journals. *Scientometrics, 51*(1), 147–162. https://doi.org/10.1023/A:1010520913124

Margalida, A., & Colomer, M. À. (2015). *Mistake index as a surrogate of quality in scientific manuscripts. Proceedings of the National Academy of Sciences, 112*(13), E1511. https://doi.org/10.1073/pnas.1500322112.

Marginson, S. (2022). What drives global science? The four competing narratives. *Studies in Higher Education*, *47*(8), 1566–1584. https://doi.org/10.1080/03075079.2021.1942822.

Moussa, S., & Teixeira Da Silva, J. A. (2023). Testing the robustness of COPE's characterization of predatory publishing on a COPE member publisher (Academic and Business Research Institute). *Publishing Research Quarterly, 39*(4), 337–367. https://doi.org/10.1007/s12109-023-09967-9.

Müller, R., & de Rijcke, S. (2017). Thinking with indicators: Exploring the epistemic impacts of academic performance indicators in the life sciences. *Research Evaluation, 26*(3), 157–168. https://doi.org/10.1093/reseval/rvx023.

Niles, M. T., Schimanski, L. A., McKiernan, E. C., & Alperin, J. P. (2020). Why we publish where we do: Faculty publishing values and their relationship to review, promotion and tenure expectations. *PLOS ONE*, *15*(3), e0228914. https://doi.org/10.1371/journal.pone.0228914.

Nobes, A., & Harris, S. (2023). Open access in low- and middle-income countries: Attitudes and experiences of researchers. *Emerald Open Research*, *1*(3). https://doi.org/10.1108/EOR-03-2023-0006.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6.

Resnik, D. B., & Dinse, G. E. (2012). Do U.S. research institutions meet or exceed federal mandates for instruction in responsible conduct of research? *A national survey. Academic Medicine*, *87*(9), 1237–1242. https://doi.org/10.1097/ACM.0b013e318260fe5c.

Richardson, R. A. K., Hong, S. S., Byrne, J. A., Stoeger, T., & Amaral, L. A. N. (2025). *The entities enabling scientific fraud at scale are large, resilient, and growing rapidly. Proceedings of the National Academy of Sciences, 122*(32), e2420092122. https://doi.org/10.1073/pnas.2420092122.

Sharma, K. (2024). Over two decades of scientific misconduct in India: Retraction reasons and journal quality among inter-country and intra-country institutional collaboration. *Scientometrics*, *129*(12), 7735–7757. https://doi.org/10.1007/s11192-024-05192-z.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *The Academy of Management Review, 20*(3), 571-610. https://doi.org/10.2307/258788

Siler, K. (2020). Demarcating spectrums of predatory publishing: Economic and institutional sources of academic legitimacy. *Journal of the Association for Information Science and Technology, 71*(11), 1386–1401. https://doi.org/10.1002/asi.24339.

Tamer, H. Y., Övgün, B., & Yalçintaş, A. (2020). Akademik büyük veri ve bilimsel bilgi üretimi: DergiPark örneği. *Ankara Üniversitesi Sosyal Bilimler Dergisi*, *11*(1), 93. https://doi.org/10.33537/sobild.2020.11.1.10.

Teixeira da Silva, J. A., & Dobránszki, J. (2017). Notices and policies for retractions, expressions of concern, errata and corrigenda: Their importance, content, and context. *Science and Engineering Ethics, 23*(2), 521–554. https://doi.org/10.1007/s11948-016-9769-y.

Teixeira Da Silva, J. A., & Moussa, S. (2024). The COPE/DOAJ/OASPA/WAME principles of transparency and best practice in scholarly publishing: A critical analysis. *Ethics in Progress*, *15*(1), 130–154. https://doi.org/10.14746/eip.2024.1.7.

Tiokhin, L., Panchanathan, K., Lakens, D., Vazire, S., Morgan, T., & Zollman, K. (2021). Honest signaling in academic publishing. *PLOS ONE*, *16*(2), e0246675. https://doi.org/10.1371/journal.pone.0246675.

TR Dizin. (2022). *TR Dizin etik ilkeleri akış şeması.* https://trdizin.gov.tr/wp-content/uploads/2022/04/TRDizin_etik_ilkeleri_akis_semasi.pdf.

Tuglular, T., Gurdal, G., Can, G. K., & Ozdemirden, A. S. (2022). Repository landscape in Turkiye and GCRIS: The first national research information system. *Procedia Computer Science, 211*, 222–229. https://doi.org/10.1016/j.procs.2022.10.195.

Tutuncu, L., & Nasir, H. (2025). Publishing local articles and book chapters: A gamer's shortcut to promotion in a metric-based academic system. *Higher Education Policy*. https://doi.org/10.1057/s41307-025-00403-5.

Van Dalen, H. P., & Henkens, K. (2005). Signals in science: On the importance of signaling in gaining attention in science. *Scientometrics, 64*(2), 209–233. https://doi.org/10.1007/s11192-005-0248-5.

Van Dalen, H. P., & Henkens, K. (2012). Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. *Journal of the American Society for Information Science and Technology*, *63*(7), 1282–1293. https://doi.org/10.1002/asi.22636.

Wessels, J. S., & Visagie, R. G. (2017). The eligibility of public administration research for ethics review: A case study of two international peer-reviewed journals. *International Review of Administrative Sciences*, *83*(1_suppl), 156–176. https://doi.org/10.1177/0020852315585949.

Zhang, M., Xu, J., Xu, C., Zheng, Q., Liu, M., Zhang, J., Fu, H., Qi, W., Zhang, J., & Tian, J. (2024). Bibliometric review and mapping analysis of publication ethics research. *Ethics & Behavior*, *34*(6), 397–409. https://doi.org/10.1080/10508422.2024.2306134.