

“Banking governance of external AML evidence: UK ownership, crypto benchmarking and temporal signals”

AUTHORS	Ayman Bader  Raed Alqirem Atala Alqtish Ayman Mansour Khalaf Alkhazaleh 
ARTICLE INFO	Ayman Bader, Raed Alqirem, Atala Alqtish and Ayman Mansour Khalaf Alkhazaleh (2026). Banking governance of external AML evidence: UK ownership, crypto benchmarking and temporal signals . <i>Banks and Bank Systems</i> , 21(3), 324–341. doi: 10.21511/bbs.21(3).2026.21
DOI	http://dx.doi.org/10.21511/bbs.21(3).2026.21
RELEASED ON	Wednesday, 30 September 2026
RECEIVED ON	Saturday, 30 May 2026
ACCEPTED ON	Monday, 07 September 2026
LICENSE	 This work is licensed under a Creative Commons Attribution 4.0 International License
JOURNAL	"Banks and Bank Systems"
ISSN PRINT	1816-7403
ISSN ONLINE	1991-7074
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”



NUMBER OF REFERENCES

60



NUMBER OF FIGURES

0



NUMBER OF TABLES

11

© The author(s) 2026. This publication is an open access article.



BUSINESS PERSPECTIVES



LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10,
Sumy, 40022, Ukraine
www.businessperspectives.org

Type of the article: Research Article

Received on: 30th of May, 2026

Accepted on: 7th of September, 2026

Published on: 30th of September, 2026

© Ayman Bader, Raed Alqirem, Atala Alqtish, Ayman Alkhazaleh, 2026

Ayman Bader, Ph.D., Associate Professor, Faculty of Business, Department of Accounting, Al-Zaytoonah University of Jordan, Jordan. (Corresponding author)

Raed Alqirem, Ph.D., Full Professor, Faculty of Business, Department of Business, Al-Zaytoonah University of Jordan, Jordan.

Atala Alqtish, Ph.D., Full Professor, Faculty of Business, Department of Accounting, Al-Zaytoonah University of Jordan, Jordan.

Ayman Alkhazaleh, Ph.D., Associate Professor, Faculty of Business, Department of Accounting and Finance, Middle East University, Jordan.



This is an Open Access article, distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license, which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conflict of interest statement:

Author(s) reported no conflict of interest

Ayman Bader (Jordan), Raed Alqirem (Jordan), Atala Alqtish (Jordan), Ayman Alkhazaleh (Jordan)

BANKING GOVERNANCE OF EXTERNAL AML EVIDENCE: UK OWNERSHIP, CRYPTO BENCHMARKING AND TEMPORAL SIGNALS

Abstract

Banks rely on ownership, transaction, and institutional information to support anti-money-laundering decision-making, yet these sources differ in decision proximity, uncertainty, and evidentiary strength. This multi-study article evaluates three external anti-money-laundering signals within a decision-governance framework emphasizing provenance, auditability, and bounded interpretation. Study 1 examines United Kingdom Persons with Significant Control records as ownership-verification workload signals. Of 6.2 million active companies, 745,200 (12.0%) met at least one verification-priority rule. Among flagged companies, 58.3% were assigned to multi-layered control structures, 25.0% to non-United Kingdom or foreign-linked controllers, 11.2% to high officer/controller turnover, and 5.5% to circular ownership patterns. Study 2 compares extreme gradient boosting, graph convolutional networks, Graph Sample and Aggregate, and graph attention networks on the Elliptic transaction-network benchmark. Graph attention networks recorded the highest point estimates (AUROC = 0.960; AUPRC = 0.740), followed by Graph Sample and Aggregate (0.930; 0.610), graph convolutional networks (0.910; 0.560), and extreme gradient boosting (0.820; 0.370). Study 3 identifies a reported -6.8% posterior mean deviation in the Persons with Significant Control Anomaly Index after the late-June 2020 filing-regime boundary (95% credible interval: -10.9% to -2.4%), interpreted descriptively rather than causally. Overall, the findings establish an evidentiary hierarchy for governing heterogeneous anti-money-laundering signals while preserving boundaries between decision support, adjudication, and causal inference. The empirical objects are companies, transaction nodes, and register observations rather than banks.

Keywords

banking, anti-money laundering, ownership verification, cryptocurrency, decision support, model governance

JEL Classification

G21, G28, K42, C55

INTRODUCTION

Banks occupy a central position in anti-money-laundering (AML) systems because they verify customers and beneficial owners, monitor transactions, investigate alerts, and submit suspicious activity reports. These responsibilities require institutions to transform fragmented ownership, customer, transaction, and regulatory information into proportionate review priorities. The practical and academic problem is therefore not simply whether compliance records exist, but whether externally observable data can support transparent, auditable, and limited prioritization without being interpreted as proof of wrongdoing.

The United Kingdom provides a relevant setting because it combines an international financial center with continuing reforms in corporate transparency and AML oversight. Complex ownership chains, offshore connections, professional intermediaries, digital-asset exposure,

high reporting volumes, and differences in reporting structures may complicate customer due diligence and transaction-monitoring decisions. Public disclosure and regulatory reporting can improve the information environment available to banks, yet neither necessarily resolves analytical integration challenges, defensive reporting, false-positive burdens, or weak institutional feedback.

Existing research examines ownership transparency, suspicious activity reporting, cryptocurrency analytics, and machine learning with different units and evidentiary claims. Treating them as one empirical model would create false comparability. This article instead integrates them at the level of decision governance through five criteria: source provenance, proximity to a banking decision, component-appropriate uncertainty, reproducible transformation, and a clear boundary between prioritization and adjudication. The framework asks whether each external signal can be used and audited proportionately to its evidentiary strength.

Decision proximity provides the organizing hierarchy. Study 1 is closest to customer due diligence because public ownership records can inform documentary verification, although they do not determine legal beneficial ownership or suspicion. Study 2 is more distant: Elliptic evaluates transaction-node ranking on a jurisdiction-neutral benchmark and does not measure United Kingdom bank exposure. Study 3 is contextual, describing register-level temporal behavior around a filing-regime boundary without an unaffected comparison series. This hierarchy permits joint governance analysis without promoting benchmark or aggregate evidence to bank-level proof.

1. LITERATURE REVIEW AND HYPOTHESES

The term illicit financial flows (IFFs) refers to cross-border movement of funds that are illicitly earned, transferred, hidden, and/or laundered for purposes of concealing origin, ownership, control, or destination (Mohd Padil et al., 2022). In the context of banking, IFFs are considered a permanent governance challenge owing to the banks' location in the functional borderland between legitimate finance and possible misuse of financial resources. Banks onboard their customers, verify their identity and beneficial ownership, perform transaction monitoring, file suspicious activity reports (SARs), and document audit trails for the regulators. IFFs may impact fiscal capacity, skew competition, erode trust, and fund illegal practices such as corruption, fraud, cybercrime, and terrorist financing (Jayasekara, 2021; Thakkar et al., 2024). The challenge that banks are facing thus includes not only whether there are any laundering typologies but also whether banks are capable of converting fragmentary ownership, payments, and customer risk information into meaningful review-prioritization signals.

The traditional methods of laundering include trade misinvoicing, shell companies, cash-inten-

sive businesses, nominees, fraudulent invoicing, professional facilitation, and cross-border layering (Umar, 2023). The relevance of these techniques for banks stems from the fact that they may occur during customer due diligence (CDD), enhanced due diligence (EDD), correspondent banking analysis, payment screening, and transaction monitoring risk escalation. The phenomenon of transaction laundering illustrates how illegal activity could become embedded in legitimate payment processing systems, particularly if the latter relies on isolated transaction monitoring but fails to incorporate a customer-based and behavioral approach (Saxena, 2024). Contemporary empirical evidence in banking suggests that anti-money laundering (AML) compliance failures could expose banks to substantial regulatory risks, including the recent NatWest criminal conviction in the United Kingdom that re-opened debates on the relative effectiveness of criminal prosecution versus civil penalties (Atkins, 2024).

A relevant case to consider would be the United Kingdom, whose position as a global financial center is accompanied by profound AML reforms. London draws legitimate capital, but the same depth of financial institutions can be utilized through complex corporate structures, offshore connections, professional enablers, and value

transfer through technologies (Ruggiero, 2022; Pontes et al., 2022). Additionally, data from UK regulators demonstrate that cross-jurisdictional financial crime risks remain a relevant concern for supervised financial institutions (Feridun, 2024). For banks, this scenario brings a risk-based supervision challenge as they need to evaluate complex clients, reconcile public ownership data, monitor high-volume transactions, and respond to evolving regulatory requirements. Empirical evidence on financial intelligence units reform and grey-list jurisdictions suggests that the mere existence of legal instruments cannot be considered a guarantee of disruption unless there is also enough institutional capacity, supervision quality, and intelligent information exchange (McNaughton, 2024; Tuhirirwe & Alexander, 2025).

The United Kingdom's Persons with Significant Control (PSC) register strengthens transparency by identifying persons or entities meeting specified control conditions (Companies House, 2020). For banks, PSC data are verification inputs rather than complete legal determinations of beneficial ownership. Companies House notes that a beneficial owner under the Money Laundering Regulations is not always the same as a PSC; differences therefore require contextual CDD assessment rather than automatic adverse inference (Companies House, 2023). Gilmour (2024) and Gilmour et al. (2025) likewise emphasize register interpretation and verification design. Complex chains, foreign-linked controllers, circular links, or frequent changes may increase verification effort without establishing unlawful conduct (Tiwari et al., 2025).

The SAR regime represents the second pillar of banks' AML systems. It converts the private sector's observations into financial intelligence by requiring banks and other regulated entities to report any suspicious money laundering or criminal property activity (Loh, 2021). On the one hand, the intelligence-driven approach is supposed to enhance detection and enforcement. On the other hand, the high volume of reports may lead to analytical overload if the quality of reports is inconsistent, defensive reporting incentives prevail, and feedback loop mechanisms are weak (Norton, 2024). The recent empirical research on SAR demonstrates that effectiveness

does not rely on the volume but also on the conversion of the reports into actionable intelligence (Imeny et al., 2025). The importance of this point for banks lies in the fact that the governance of the SAR process affects escalation quality, compliance workload, and the possibility for regulators to give useful feedback.

AML compliance by banks becomes increasingly dependent on transaction monitoring systems that help to detect suspicious activity and avoid an excessive number of false-positive alerts. Empirical data from the industry show that transaction monitoring is constrained by rule rigidity, lack of data integration, burden of false-positive alerts, and challenges in translating technical alerts into ready-for-investigation cases (Oztas et al., 2024). Machine learning score models were developed to prioritize the alerts and reduce time-to-SAR escalation, while false-positive suppression models aim to minimize the low-value alerts (Bakry et al., 2024; Halford et al., 2025). Customer risk classification plays a crucial role as banks need to allocate compliance resources proportionally to risk level rather than treat all customers equally (Reite et al., 2025). Additionally, empirical evidence from the banking sector suggests that organizational, technological, and institutional factors affect AML compliance behavior (Hoffman et al., 2024). However, these models require governance, validation, and human oversight.

Evidence from other digital information environments reinforces this governance problem. Al-Muntasir (2022) shows that intensive information flows can coexist with misinformation, making source verification important before external information is used in subsequent decisions. Ahmad et al. (2023) similarly find that communicators using social-media information sources must select and verify digital information before it enters organizational agendas. Oreqat (2021) further shows that perceived utility and satisfaction shape engagement with digital information systems, indicating that information availability alone does not determine how external information is used. These studies do not provide AML evidence, but they support the broader principle that external digital information requires source appraisal, contextual interpretation, and controlled incorporation into decision processes.

Cryptocurrencies amplify the existing banking AML challenge. Cryptocurrency transactions are traceable in the public ledger, yet the attribution of ownership remains difficult in the case of pseudonymity, exchanges, mixers, chain-hopping, fragmented transfers, and privacy-enhancing technologies (Atlam et al., 2024; Wang & Hsieh, 2024). Although the banks may lack direct control over blockchain networks, they have exposure through customers, virtual asset service providers, fiat on-ramps, payment flows, and correspondent relationships. Research on supervision of cryptocurrency exchanges reveals that AML obligations of virtual asset intermediaries could influence the risk environment to be screened by banks (Volten et al., 2025). Recent scholarly work in banking regulation discusses the need for banks' internal governance to change according to money laundering risks related to cryptocurrencies and value transfer enabled by technologies (Minto, 2024). Research on cyber-laundering and decentralized finance suggests that laundering patterns can be networked and adaptive rather than isolated (Wronka, 2022, 2023). Recent cross-jurisdictional evidence further indicates that crypto enforcement may produce short-run disruption while being followed by displacement toward alternative regulatory hubs, reinforcing the governance importance of coordinated cross-border oversight (Morshed, 2026).

Transaction classification can combine node attributes and relational structure. Elliptic's published node matrix already contains local features and one-hop neighbor aggregates, so a conventional learner using the full matrix is not graph-information-free. The relevant contrast is between non-message-passing models and models that learn from adjacency. Graph neural networks exploit relational structure dynamically (Kim et al., 2022), heterogeneous models extend that logic (Johannessen & Jullum, 2025), and temporal AML research stresses stability under distribution shift (Deprez et al., 2025; Zhang et al., 2026).

Artificial intelligence expands AML opportunities, but also introduces certain governance challenges. Suspicious transaction monitoring can raise legal and integrity issues since the algorithms can influence customer treatment, reporting, and auditing processes (Turksen et al., 2024).

Consequently, financial institution researchers pay special attention to explainability, reproducibility, and fairness in AML models, especially if they are aimed at detecting suspicious bank transactions (Narasapuram et al., 2026; Mazumder, 2026). Broader research in AI-based AML shows that scalable detection can enhance financial institution monitoring, but only if the output of the model remains interpretable and traceable to auditable evidence (Premkumar et al., 2026; Kungu et al., 2026). More broadly, ethical AI scholarship in business emphasizes transparency, fairness, accountability, and human oversight as governance requirements when algorithmic systems influence consequential decisions (Musa et al., 2025). Evidence from GCC logistics similarly associates AI transparency, regulatory frameworks, infrastructure readiness, and organizational alignment with sustainability-related outcomes, illustrating how AI-enabled decisions remain contingent on institutional governance (Morshed & Khrais, 2026).

AML policy is shaped by global standards, regional coordination, and national implementation. Modern frameworks emphasize risk-based supervision, ownership transparency, international cooperation, and effectiveness assessment (Jayasekara, 2021). The establishment of the European Union AML authority reflects movement toward more centralized supervision and cross-border coordination (Tosza & Voordeckers, 2024). Financial-system resilience research likewise locates AML/CFT controls within broader safeguards against systemic abuse (Korauš et al., 2024). For banking decision support, the relevant implication is narrower: an external signal is useful only when its provenance, decision role, uncertainty, and evidentiary boundary are explicit.

The literature therefore supports an evidentiary-governance view rather than a unified anti-money-laundering outcome model. Ownership records, transaction graphs, and institutional series can inform different review tasks only when provenance and purpose are explicit, uncertainty matches the empirical component, transformations are auditable, and the inferential boundary is stated. This perspective distinguishes table-level numerical traceability from complete raw-to-result replication. Accordingly, this study aims to

evaluate how heterogeneous external AML signals can be governed and interpreted proportionately across ownership verification, transaction-network benchmarking, and temporal institutional evidence, while preserving clear boundaries between decision support, adjudication, and causal inference. Accordingly, the article uses three study-specific hypotheses within one external-signal governance framework.

The hypotheses are descriptive within their own data-generating processes rather than claims of a common latent construct or pooled effect. Cross-study synthesis concerns whether the resulting evidence remains proportionate to its decision role and independently auditable.

- H1: Under the pre-specified mutually exclusive PSC workload-allocation rule, verification assignments are unevenly concentrated across the four operational categories; this hypothesis concerns allocated review workload, not raw prevalence of overlapping component flags.*
- H2: On the fixed held-out Elliptic temporal block, message-passing graph models have higher reported AUROC and AUPRC point estimates than non-message-passing XGBoost using the same published node-feature matrix; the hypothesis concerns descriptive benchmark ordering, not pairwise statistical superiority or United Kingdom banking performance.*
- H3: The post-June 2020 PSC Anomaly Index is lower than the posterior continuation of its pre-period pattern; this is a descriptive temporal hypothesis and not an intervention-effect claim.*

2. METHODS

This quantitative archival-computational multi-study design evaluates external anti-money-laundering signals using public and benchmark information rather than confidential customer files, internal alerts, suspicious activity report narratives, or supervisory records. The studies are not pooled. Each is assessed through provenance, decision

proximity, component-appropriate uncertainty, auditability, and evidentiary boundary while retaining its own unit, outcome, and inferential standard.

Empirical coverage differs across the three components. Study 1 examines UK Companies House Persons with Significant Control (PSC) records for 2017–2023, with active companies as the unit of analysis. Study 2 examines 203,769 Bitcoin transaction nodes in the Elliptic benchmark across 49 relative timesteps rather than a calendar-year panel. Study 3 analyzes monthly PSC-derived register observations for 2017–2023 around the 27 June 2020 filing-regime boundary. Thus, banks are not sampled objects in any component; they constitute the institutional decision-use context in which the external AML evidence is interpreted.

Methodologically, the present design makes the empirical unit explicit rather than treating the banking context as the observational unit. Taqa (2025), for example, defines a bank-level organizational population and analyzes 234 questionnaires from managers and related decision makers across 24 Iraqi commercial banks. By contrast, the present studies analyze UK companies, Elliptic transaction nodes, and monthly PSC register observations; banks are the intended decision-use context, not sampled units.

Study 1 evaluates company-level ownership-verification priority from Persons with Significant Control disclosures. Study 2 uses Elliptic solely as a jurisdiction-neutral transaction-ranking benchmark. Study 3 summarizes a monthly Persons with Significant Control Anomaly Index around the late-June 2020 filing-regime boundary. The article claims neither record-level linkage across studies, a common dependent variable, nor a shared causal mechanism. Its integrative object is governance of external evidence before that evidence is used in banking decisions.

The evidentiary scope is deliberately narrow. PSC observations are filed control disclosures, and the four indicators are documentation-priority rules rather than legal findings about beneficial ownership or suspicious activity. Elliptic labels have no UK-customer or UK-bank linkage. The institutional series is derived only from PSC records and

cannot identify the effect of a contemporaneous policy or administrative event. The analysis therefore concerns governance of external decision support, not realized AML effectiveness.

The workflow separates study design from submission-facing evidence. The PSC design screens, aggregates, and applies deterministic rules to a company denominator. Elliptic preserves a time-ordered development/validation/test structure with preprocessing confined to pre-test data. Study 3 constructs a monthly register-level index around the late-June 2020 filing-regime boundary. The supplementary workbook contains only the table-level quantities retained in the current manuscript and their internal consistency checks; it is not represented as a raw-data or prediction-level replication package.

Companies House PSC records provide ownership/control information (Companies House, 2019), with officer history used only for turnover. Elliptic contains 203,769 nodes, 234,355 directed links, 166 features, and 49 approximately biweekly timesteps (Weber et al., 2019). Its feature matrix comprises 94 local and 72 one-hop aggregated features. UKFIU/NCA material is contextual only. June 2020 is the Study 3 indexing boundary because temporary Companies House filing requirements took effect on 27 June 2020 and covered event-driven filings (UK Government, 2020).

The sources are empirically incompatible and never linked at record level. Table 2 harmonizes only their governance dimensions. No PSC company is mapped to Elliptic, no Elliptic label is treated as UK-bank exposure, and the PSC series validates neither other study. Cross-study synthesis con-

cerns whether each signal can be bounded and audited consistently before use. The framework therefore compares governance discipline, not empirical effect sizes.

For the PSC study, the reported analytical denominator is 6.2 million active companies. Table 4 is the principal empirical output: 745,200 companies meet at least one verification-priority rule and are allocated once to one of four mutually exclusive workload categories. The supplementary evidence matrix checks that the four reported category counts reconcile exactly to the reported flagged total and that the reported percentages are consistent with the stated denominator. Record-level assignments, linkage records, and overlap files are not part of the submission workbook.

The PSC ownership graph is

$$G_PSC = (V_C \cup V_P, E_CP), \quad (1)$$

where V_C contains companies, V_P controllers, and E_CP directed control links.

A corporate/legal controller is continued as a company node only when its recorded Companies House identifier exactly matches a retained company number; unmatched entities remain terminal. Name-only records are excluded from path continuation and cycle detection. Layering and circularity therefore use identifier-resolved links only, avoiding speculative name matching.

Four binary verification-priority rules are applied. $Layered_i = 1$ when an identifier-resolved company-to-controller path exceeds three control levels; $NonUK_i = 1$ when at least one recorded controller is associated with a non-UK jurisdiction or for-

Table 1. Data overview and AML relevance

Dataset	Coverage	Records / Nodes	Main attributes	AML relevance / verification
Companies House PSC analytical denominator	2017–2023	6.2M active companies (reported denominator)	Ownership/control fields used for verification-priority rules	External ownership-review denominator; record-level source files are not included in the submission workbook
Elliptic Bitcoin graph	49 relative timesteps	203,769 nodes; 234,355 directed edges	Transaction-node features, directed links, relative time ordering, and licit/illicit research labels	Jurisdiction-neutral ranking benchmark; not UK-bank crypto exposure
UK SAR publications and selected institutional events	2017–2023	Monthly PSC-derived series; contextual publications	PSC Anomaly Index; SAR context; 27 June 2020 filing-regime boundary	Register-level context only; boundary is non-exogenous; SAR counts are not outcomes

eign legal entity; $Circular_i = 1$ when an identifier-resolved directed path returns to an already visited company or controller; and $Turnover_i = 1$ when more than two officer or controller changes occur within the preceding 24 months. The *NonUK* label is deliberately descriptive and is not treated as an ‘offshore-risk’ classification. These rules indicate additional verification effort only and do not encode suspicion, illegality, or legal beneficial-ownership status.

$$\begin{aligned} PSC_Priority_i &= Circular_i + Turnover_i \\ &+ NonUK_i + Layered_i, \\ \text{and } PSC_Flag_i &= 1, \end{aligned} \quad (2)$$

when $PSC_Priority_i \geq 1$.

Table 4 uses a pre-specified workload hierarchy – Layered, NonUK, Turnover, Circular – so each flagged company is counted once, while all component flags remain retained. The hierarchy allocates review workload only. *H1* evaluates heterogeneity of that assigned workload and does not infer raw component prevalence from the primary counts.

The PSC analysis is descriptive and rule-based. The design includes overlap inspection, component ablation, and alternative layering and turnover thresholds; incomplete 24-month turnover history is unavailable rather than zero. Quantitative claims in the current manuscript are limited to the reported denominator, the four mutually exclusive workload counts, and the total flagged share. The submission workbook verifies their arithmetic consistency but does not substitute for company-level source records.

The Elliptic benchmark is

$$G_BTC = (V, E, X, Y). \quad (3)$$

Unlabeled nodes are excluded from supervised loss and test metrics but may remain as graph context. All models receive the same 166 features, including neighbor-aggregated information. XGBoost is therefore a non-message-passing baseline; GCN, GraphSAGE, and GAT additionally learn from adjacency. The comparison concerns learned message passing rather than a misleading ‘features versus graph’ distinction.

For node i ,

$$\hat{y}_i = f(X_i, A_i; \theta), \quad (4)$$

is the illicit-label score, with A_i used only by message-passing models. Performance is reported using AUROC and AUPRC, with emphasis on AUPRC because illicit-labelled nodes are a minority. No fixed-capacity precision metric is reported because the benchmark provides no defensible bank-specific review-capacity denominator. The reported metrics therefore evaluate threshold-free ranking only and are not customer decisions or bank alert thresholds. Related fraud-classification research likewise shows that highly imbalanced financial datasets require evaluation criteria beyond conventional accuracy (Aljawazneh & Abu Al-Ragheb, 2025).

The retained analytical specification preserves the original later-period Elliptic holdout: timesteps 35–49 are reserved for test evaluation, with timesteps 1–29 used for fitting and 30–34 for validation, early stopping, and model selection (Weber et al., 2019). Standardization and class weights are specified as fit-data operations. The current submission evidence reports only the retained AUROC and AUPRC point estimates; exact split-membership and held-out prediction files would be required for independent model rerunning.

The retained model specification uses 500-tree XGBoost with depth 4, learning rate 0.05, subsample 0.80, column sample 0.80, logistic objective, and a class-ratio weight. GCN and GraphSAGE use two layers, 64 hidden units, dropout 0.30, Adam optimization, learning rate 0.001, weighted binary cross-entropy, and early stopping; GAT uses two layers, eight first-layer attention heads, and the same optimization logic. These are the reported design settings. Table 5 retains only the reported AUROC and AUPRC point estimates; independent rerunning would require the final configuration and prediction files.

The institutional study constructs the PSC Anomaly Index from Study 1 rules. For each month, only information effective by month-end is used to determine $Layered_{it}$, $NonUK_{it}$, $Circular_{it}$, $Turnover_{it}$ and PSC_Flag_{it} . N_t contains company-months eligible for all four rules, and

$$PSC_Index_t = \frac{1}{N_t} \sum_i PSC_Flag_{it}. \quad (5)$$

Turnover requires a complete preceding 24-month history; missing look-back is not coded as zero. The index measures verification-priority prevalence, not SARs, laundering, enforcement, or bank performance.

June 2020 is an indexing boundary, not an intervention. Temporary filing rules took effect on 27 June 2020; no within-June effect is inferred, and interpretation begins with July. The conceptual univariate representation is

$$y_t = \mu_t + \tau_t + \varepsilon_t. \quad (6)$$

The current submission does not include the final state equations, priors, software configuration, monthly series, posterior draws, or dated placebos required for independent model rerunning. Table 6 therefore reports the retained posterior summary only (Gianacas et al., 2023; Liang et al., 2025).

Robustness is component-specific and is not treated as an additional empirical result unless the corresponding underlying output is available. For PSC, the manuscript retains only the main workload-allocation results. For Elliptic, Table 5 reports point metrics only; prediction-level un-

certainty and paired resampling are not reported. For Study 3, the manuscript retains the reported posterior mean deviation and credible interval only. Qualitative explainability notes and auxiliary analyses not supported by the current submission evidence are excluded from the quantitative findings.

Auditability is documented through DOI: <https://doi.org/10.5281/zenodo.21945418>. The deposited workbook contains a claim-to-evidence crosswalk, PSC reconciliation formulas, the reported Elliptic AUROC/AUPRC point estimates, the reported June 2020 PSC-index posterior summary, and a data-availability map. Only evidence used by the current manuscript is included, and the workbook does not claim that unavailable record-, prediction-, or model-level files have been independently regenerated.

3. RESULTS

Results are reported as three autonomous studies under the common external-signal governance framework. Study 1 concerns ownership-verification workload, Study 2 concerns ranking on a fixed cryptocurrency benchmark, and Study 3 concerns register-level temporal behavior. No cross-study numerical aggregation is performed, and evidence from one study is not used to validate another.

Table 2. Methods and compliance purpose

Study / decision layer	Method	Study-specific question	Output	Governance / audit boundary
Study 1: ownership verification / CDD	PSC rules; exact linkage; denominator, overlap, and sensitivity checks	How is PSC verification workload distributed under the operational rules?	PSC flag; mutually exclusive workload category; flagged share	PSC is not legal BO; hierarchy allocates workload only; overlaps retained
Study 2: transaction-ranking benchmark	XGBoost plus GCN, GraphSAGE and GAT on a fixed temporal split	How do non-message-passing and message-passing models rank the same Elliptic holdout?	AUROC; AUPRC point estimates	Same 166 features; no UK-bank linkage; retained 1-29/30-34/35-49 design; prediction-level test membership is not included in the submission evidence
Study 3: register-level temporal context	Univariate Bayesian state-space summary of monthly PSC Anomaly Index	Does the PSC-derived index deviate from pre-period continuation after the late-June 2020 boundary?	Posterior mean deviation; 95% credible interval	No unaffected control; late-June boundary is non-causal; current package traces the posterior summary but does not rerun the state-space model
Cross-study governance	Common provenance, purpose, uncertainty, reproducibility, and boundary checks	Can heterogeneous external signals be governed consistently without empirical pooling?	Study-specific evidence map	No cross-study validation, pooled score, common dependent variable, or causal synthesis

Table 3 summarizes the retained data environments without treating unavailable intermediate counts as independent findings. Persons with Significant Control records provide the United Kingdom company-level ownership-verification context, with 6.2 million active companies forming the reported denominator for Table 4. Elliptic is a public cryptocurrency benchmark with no empirical United Kingdom bank linkage. Suspicious activity report publications provide institutional background only, while the temporal outcome is constructed from Persons with Significant Control records rather than suspicious activity report counts (National Crime Agency, 2024; Feridun, 2024).

Table 4 reports mutually exclusive PSC workload assignments and evaluates H1. Overall, 745,200 active companies (12.0% of the denominator) meet at least one rule. The four workload categories reconcile exactly to that total: 58.3% of flagged companies are assigned to layered structures, 25.0% to non-UK/foreign-linked controllers, 11.2% to turnover, and 5.5% to circular structures. This pattern supports H1 as uneven workload concentration. Because assignment is hierarchical, these proportions are not raw component prevalence and do not determine legal beneficial ownership, material discrepancy, suspicious-customer status, or un-

lawful conduct (Companies House, 2023).

Table 5 evaluates H2 using the reported AUROC and AUPRC point estimates for the four benchmark models. GCN, GraphSAGE, and GAT each have higher reported values than non-message-passing XGBoost, and GAT has the highest reported AUROC (0.960) and AUPRC (0.740). Because prediction-level outputs and paired uncertainty records are not included in the submission evidence, the comparison is descriptive. It supports benchmark ordering only and does not establish statistical superiority, operational performance in United Kingdom banks, or culpability of any transaction node (Pocher et al., 2023). Comparable evidence from banking credit-risk classification also shows that deep-learning performance should be interpreted through task-specific evaluation measures and class-balancing design rather than model labels alone (Al-Shayea et al., 2025).

Table 6 reports Study 3 and evaluates H3. The retained result is a reported post-June 2020 posterior mean deviation of -6.8% , with a 95% credible interval from -10.9% to -2.4% , relative to the reported pre-period continuation. The interval is interpreted as a posterior summary from the original state-space analysis. Because the monthly series, final model configuration, and posterior draws are

Table 3. Descriptive statistics of external AML data layers

Dataset	Coverage	Key result	Interpretation
PSC Register	2017–2023	Active companies: 6.2M	Reported company denominator for Table 4 verification-priority percentages
Elliptic Bitcoin graph	49 relative timesteps	Nodes: 203,769; edges: 234,355	Jurisdiction-neutral benchmark network for transaction-node model comparison
Elliptic Bitcoin graph	49 relative timesteps	Illicit-labelled: 4,545; licit-labelled: 42,019	Minority-class benchmark composition; not UK bank exposure
Elliptic Bitcoin graph	49 relative timesteps	Illicit-labelled share among labelled nodes: 9.76%	Supports emphasis on AUPRC in the imbalanced benchmark evaluation

Table 4. PSC verification-priority signals

Primary workload category	Entities flagged	% of total companies	Interpretation
Circular ownership pattern	41,200	0.66%	Allocated after earlier categories; not raw component prevalence.
High officer/controller turnover (> 2 changes/24 months)	83,500	1.35%	Requires complete 24-month history; not misconduct evidence
Non-UK/foreign-linked controllers	186,300	3.00%	Foreign-link documentation category; not an offshore-risk finding
Multi-layered control chains (> 3 levels)	434,200	7.00%	First workload assignment under hierarchy; not raw prevalence
Total flagged entities	745,200	12.0%	At least one rule; four mutually exclusive category counts reconcile exactly to this total

Table 5. Elliptic benchmark model performance

Model (Elliptic benchmark)	AUROC	AUPRC	Input structure	Evidence interpretation
XGBoost baseline	0.820	0.370	166 features; no message passing	Reported benchmark point estimate; descriptive comparator
GCN	0.910	0.560	166 features + adjacency	Higher reported point metrics than XGBoost; no inferential claim
GraphSAGE	0.930	0.610	166 features + adjacency	Higher reported point metrics than XGBoost; no inferential claim
GAT	0.960	0.740	166 features + graph attention	Highest reported point metrics; no pairwise or UK-bank superiority claim

Note: Table 5 reports retained benchmark point estimates only. No comparative confidence interval, paired significance test, or prediction-level uncertainty is inferred from the supplementary workbook. Formal comparative inference would require the exact held-out prediction probabilities, split membership, and final model-run records.

Table 6. Exploratory PSC-index posterior temporal deviation

Indexing boundary / outcome	Posterior mean temporal deviation	95% CrI	Interpretive strength	Interpretation
Late-June 2020 filing-regime boundary / monthly PSC Anomaly Index	−6.8%	[−10.9%, −2.4%]	Traced posterior summary	Negative reported posterior deviation; retained as a descriptive summary, while the underlying state-space model is not independently rerun

Note: The reported posterior summary concerns the constructed monthly PSC verification-priority index. The 27 June 2020 boundary is documented institutional context, not an exogenous treatment date. No within-June effect is attributed. The result is descriptive and does not establish a reduction in money laundering, bank risk, or confirmed suspicious activity.

not included in the submission evidence, the paper does not claim independent rerunning of the underlying model.

Table 7 defines the verification boundary for each study. The submission workbook provides table-level arithmetic and claim consistency for the reported PSC counts, Elliptic point metrics, and June-2020 posterior summary. Record-level PSC files, Elliptic prediction-level outputs, full state-space inputs/outputs, and quantitative explainability files are not included. Accordingly, the table distinguishes what can be inspected from the

current supplement from what would be required for complete raw-to-result replication.

Table 8 preserves the three study hypotheses while aligning each conclusion with its evidence level. H1 concerns uneven PSC workload allocation; H2 concerns reported benchmark point ordering without inferential superiority; H3 concerns a reported posterior temporal deviation without causal attribution. The synthesis therefore remains decision-governance oriented rather than converting heterogeneous evidence into a common effect estimate.

Table 7. Minimum robustness and model-governance records

Domain	Validation approach	Governance relevance	Supporting documentation record	Permitted interpretive use
PSC Register	Aggregate denominator + workload reconciliation	Four category counts reconcile exactly to the reported 745,200 flagged total.	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx: PSC_RECONCILIATION	Aggregate workload audit only; record-level linkage/assignment replication not claimed.
PSC Register	Ablation and threshold checks	Underlying quantitative robustness outputs are not included in the submission evidence.	No quantitative robustness file used as independent evidence.	Do not treat robustness summaries as separate empirical findings.
Elliptic models	Fixed temporal benchmark comparison	Reported AUROC/AUPRC point metrics retained; prediction-level outputs unavailable.	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx: ELLIPTIC_METRICS	Descriptive point ordering only.

Table 7 (cont.). Minimum robustness and model-governance records

Domain	Validation approach	Governance relevance	Supporting documentation record	Permitted interpretive use
Elliptic models	Comparative uncertainty	Prediction-level paired uncertainty is unavailable.	No comparative-interval or paired-resampling evidence is used in the submission.	No statistical-superiority claim.
Institutional series	PSC-index posterior summary	Reported mean deviation and 95% credible interval retained.	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx: STUDY3_SUMMARY	Descriptive posterior summary only; no model-rerun claim.
Explainability	Explainability	Quantitative explainability outputs are not included.	No SHAP/GNNExplainer result used as quantitative evidence.	No independent explainability finding.

Table 8. Hypothesis-evidence summary

Hypothesis	Empirical result	Evidence alignment
<i>H1</i>	745,200 companies meet at least one rule; primary workload shares are 58.3% layered, 25.0% non-UK/foreign-linked, 11.2% turnover, and 5.5% circular.	Supports <i>H1</i> as uneven workload concentration; the four categories reconcile exactly to the flagged total, but no raw-overlap prevalence is inferred.
<i>H2</i>	GAT reports AUROC 0.960 and AUPRC 0.740; GCN and GraphSAGE also have higher reported point metrics than XGBoost.	Supports <i>H2</i> as descriptive benchmark ordering; no prediction-level statistical-superiority claim.
<i>H3</i>	Post-June 2020 posterior mean deviation = -6.8%; 95% CrI [-10.9%, -2.4%].	Supports <i>H3</i> as a traced descriptive posterior summary; no identified filing-regime effect and no independent state-space rerun.

Taken together, the studies form a decision-proximity hierarchy rather than one empirical system. Persons with Significant Control evidence is closest to customer due diligence because it can trigger documentary verification; Elliptic is a methodological benchmark that is informative about ranking architecture but not United Kingdom bank exposure; and the monthly index is register-level context. The same hierarchy applies to reproducibility: directly decision-adjacent evidence requires the strongest source-to-decision trace, while benchmark and contextual evidence must be prevented from acquiring operational meaning that their data cannot support (Riccardi, 2025; Ogbeide et al., 2025).

For contextual comparison, Harasis (2026) reports that data-warehousing adoption in Saudi insurance firms is associated with lower loss and combined ratios and faster claim settlement, illustrating how integrated analytical infrastructure can translate into measurable financial-sector operating outcomes. The present findings address a different question: whether external AML signals can be governed and bounded before institution-specific use, rather than estimating organizational performance effects.

4. DISCUSSION

The findings show that external anti-money-laundering (AML) evidence differs not only in analytical content but also in its proximity to banking decisions and the strength of inference it can support. Across the analyses, ownership records, transaction-network benchmarks, and institutional time-series evidence perform distinct functions. The contribution therefore lies in demonstrating how heterogeneous external signals can be incorporated into AML decision processes without assigning them evidentiary meaning beyond that supported by their underlying data.

The ownership-verification results show that workload is highly concentrated among particular ownership structures. Of the 745,200 companies meeting at least one PSC verification-priority rule, 58.3% were allocated to multi-layered structures, compared with 25.0% to non-UK or foreign-linked controllers, 11.2% to high turnover, and 5.5% to circular structures. This pattern is consistent with Gilmour (2024) and Gilmour et al. (2025), who emphasize the importance of beneficial-ownership transparency and effective verification mechanisms. The present findings extend

this literature by translating ownership complexity into an explicit workload-allocation framework rather than treating complex structures as direct indicators of misconduct. They also accord with Tiwari et al. (2025), who identify complex chains, foreign-linked controllers, and circular ownership arrangements as factors that can increase verification difficulty without independently demonstrating unlawful activity. The results therefore reinforce the distinction between verification complexity and substantive customer risk.

This distinction also complements Reite et al. (2025), who emphasize proportional allocation of AML resources across customers with different monitoring requirements. The current framework provides an operational mechanism through which such proportionality can be implemented: observable registry characteristics determine documentation priority rather than automatically generating an adverse classification. This is important because Oztas et al. (2024) show that poorly differentiated monitoring signals can increase compliance burdens without necessarily improving investigative quality. The present findings therefore suggest that the value of ownership information depends not merely on disclosure, but on how that information is converted into proportionate review priorities.

The cryptocurrency-model results provide a different comparison with prior AML analytics research. GAT records the highest reported benchmark performance, with AUROC of 0.960 and AUPRC of 0.740, followed by GraphSAGE, GCN, and XGBoost. This ordering is broadly consistent with Kim et al. (2022), who demonstrate the value of graph neural networks for exploiting relational information. However, the present results refine that interpretation because the Elliptic feature matrix supplied to XGBoost already contains one-hop aggregated features. The relevant differ-

ence therefore concerns the additional value associated with learned message passing and explicit adjacency rather than a simple contrast between graph and non-graph information. This interpretation is also consistent with Johannessen and Jullum (2025) and Deprez et al. (2025), who emphasize representation design and temporal stability in graph-based AML applications. Accordingly, the higher point estimates support the analytical relevance of graph learning without establishing universal operational superiority.

The temporal analysis likewise illustrates the importance of matching interpretation to evidentiary strength. The reported PSC Anomaly Index deviation of -6.8% , with a 95% credible interval of -10.9% to -2.4% , indicates a departure from the estimated continuation of the pre-period pattern. This interpretation is compatible with Gianacas et al. (2023), who discuss the usefulness of Bayesian structural time-series methods for evaluating temporal deviations. Nevertheless, the absence of an unaffected comparison series means that the observed change should not be interpreted as a causal effect of the filing-regime boundary. The contribution is therefore contextual rather than causal.

Taken together, the findings extend prior AML literature by showing that the principal governance issue is not simply whether external signals are informative, but how those signals are translated into banking decisions. Ownership records can support verification prioritization, graph benchmarks can inform model development, and institutional series can provide temporal context. Consistent with Premkumar et al. (2026) and Musa et al. (2025), the results reinforce the importance of interpretability, traceability, human oversight, and proportionality. The main contribution is therefore an evidentiary hierarchy that separates prioritization, analytical prediction, and causal interpretation rather than collapsing them into a single measure of AML effectiveness.

CONCLUSION

This multi-study article examined how heterogeneous external information can support anti-money-laundering decision governance without being forced into one causal model. Its unifying contribution is an evidentiary hierarchy based on provenance, decision proximity, uncertainty or sensitivity, auditability, and a boundary between prioritization and adjudication. The framework also separates table-level traceability from complete raw-to-result replication.

Three findings remain deliberately distinct. First, 745,200 active companies, equal to 12.0% of the Persons with Significant Control reporting denominator, meet at least one verification-priority rule; the four mutually exclusive workload categories reconcile exactly to that total and are unevenly concentrated, with layered cases representing 58.3% of the flagged queue. Second, the reported Elliptic point estimates are higher for the message-passing models than for non-message-passing XGBoost, with GAT AUROC 0.960 and AUPRC 0.740; this is descriptive benchmark ordering only. Third, the reported PSC-index posterior mean deviation after the late-June 2020 boundary is -6.8% with a 95% credible interval of -10.9% to -2.4% , interpreted as temporal context rather than a causal effect.

The supplementary evidence matrix strengthens transparency by restricting the submission-facing evidence to the quantities actually used in the paper and by separating table-level consistency from unavailable raw replication. Auxiliary analyses that are not supported by the current submission evidence are not presented as empirical findings.

Future work should deposit the final record-level Persons with Significant Control outputs, linkage and overlap logs, Elliptic held-out predictions and exact run metadata, and the monthly PSC-index series with final state-space configuration, posterior draws, diagnostics, and dated placebos. Institution-specific customer or transaction research would additionally require appropriate legal, governance, and confidentiality controls.

AUTHOR CONTRIBUTIONS

Conceptualization: Ayman Bader.

Data curation: Raed Alqirem Alqirem, Atala Alqtish.

Formal analysis: Ayman Bader, Raed Alqirem Alqirem, Atala Alqtish, Ayman Alkhazaleh.

Funding acquisition: Ayman Alkhazaleh.

Investigation: Ayman Bader, Atala Alqtish, Ayman Alkhazaleh.

Methodology: Ayman Bader, Raed Alqirem Alqirem, Atala Alqtish, Ayman Alkhazaleh.

Resources: Ayman Bader.

Software: Raed Alqirem Alqirem.

Supervision: Ayman Bader.

Validation: Ayman Bader, Atala Alqtish.

Visualization: Atala Alqtish, Ayman Alkhazaleh.

Writing – original draft: Raed Alqirem Alqirem, Atala Alqtish, Ayman Alkhazaleh.

Writing – reviewing & editing: Ayman Bader.

REFERENCES

- Ahmad, A. K., Nahar, H. M., & Manajreh, M. M. N. (2023). The effect of social media on shaping the agenda of the communicator in Jordanian TV channels. *Middle East Journal of Communication Studies*, 3(2), 1-40. (In Arabic). Retrieved from <https://mej.meu.edu.jo/?p=1886>
- Aljawazneh, H., & Abu Al-Ragheb, F. (2025). Fraud prediction in e-commerce transactions using TabNet with advanced data balancing techniques. *Al-Zaytoonah University Journal of Business*, 1(3). <https://doi.org/10.15849/zjib.v1i03.51>
- Al-Muntasir, M. (2022). The phenomenon of information flow from traditional and new media about the Corona pandemic from the perspective of newly graduated media professionals in Yemen. *Middle East Journal of Communication Studies*, 2(2), 7-25. (In Arabic). Retrieved from <https://mej.meu.edu.jo/?p=1781>
- Al-Shayea, Q., Saad, H., & Alaramdeen, N. (2025). Enhancing credit risk prediction using deep learning techniques. *Al-Zaytoonah University Journal of Business*, 1(2). <https://doi.org/10.15849/zjib.v1i02.43>
- Atkins, M. (2024). Should banks face criminal prosecution for breaches of Money Laundering Regulations or are civil fines effective. Analysis of the significance of the first ever criminal conviction of a bank (NatWest) for breaches of the Money Laun-

- dering Regulations. *Journal of Economic Criminology*, 6, Article 100097. <https://doi.org/10.1016/j.jeconc.2024.100097>
6. Atlam, H. F., Ekuri, N., Azad, M. A., & Lallie, H. S. (2024). Block-chain forensics: A systematic literature review of techniques, applications, challenges, and future directions. *Electronics*, 13(17), Article 3568. <https://doi.org/10.3390/electronics13173568>
 7. Bakry, A. N., Alsharkawy, A. S., Farag, M. S., & Raslan, K. R. (2024). Automatic suppression of false positive alerts in anti-money laundering systems using machine learning. *The Journal of Supercomputing*, 80(5), 6264-6284. <https://doi.org/10.1007/s11227-023-05708-z>
 8. Companies House. (2019). *Companies House data products*. GOV. UK. Retrieved from <https://www.gov.uk/guidance/companies-house-data-products>
 9. Companies House. (2020). *People with significant control (PSCs)*. GOV.UK. Retrieved from <https://www.gov.uk/guidance/people-with-significant-control-pscs>
 10. Companies House. (2023). *Report a discrepancy about a PSC or a registrable beneficial owner*. GOV. UK. Retrieved from <https://www.gov.uk/guidance/report-a-discrepancy-about-a-beneficial-owner-on-the-psc-register-by-an-obliged-entity>
 11. Deprez, B., Wei, W., Verbeke, W., Baesens, B., Mets, K., & Verdonck, T. (2025). Advances in continual graph learning for anti-money laundering systems: A comprehensive review. *WIREs Computational Statistics*, 17(3), Article e70040. <https://doi.org/10.1002/wics.70040>
 12. Feridun, M. (2024). Cross-jurisdictional financial crime risks: What can we learn from the UK regulatory data? *Journal of Financial Crime*, 31(3), 608-617. <https://doi.org/10.1108/JFC-03-2023-0044>
 13. Gianacas, C., Liu, B., Kirk, M., Di Tanna, G. L., Belcher, J., Blogg, S., & Muscatello, D. J. (2023). Bayesian structural time series, an alternative to interrupted time series in the right circumstances. *Journal of Clinical Epidemiology*, 163, 102-110. <https://doi.org/10.1016/j.jclinepi.2023.10.003>
 14. Gilmour, P. (2024). Editorial: Unmasking who controls illicit wealth through beneficial ownership transparency. *Journal of Financial Crime*, 31(4), 769-771. <https://doi.org/10.1108/JFC-07-2024-333>
 15. Gilmour, P., Pandey, D., & Goldbarsht, D. (2025). Registers of beneficial owners based on blockchain technology: Implications for the accounting profession. *Technological Forecasting and Social Change*, 214, Article 124051. <https://doi.org/10.1016/j.techfore.2025.124051>
 16. Halford, E., Gibson, I., Newfield, M., & Dhanwala, M. (2025). Developing a scoring model for managing money laundering transactions using machine learning. *Journal of Money Laundering Control*, 28(7), 30-49. <https://doi.org/10.1108/JMLC-09-2024-0152>
 17. Harasis, A. (2026). Impact of data warehousing adoption on underwriting and claims performance in Saudi insurance firms. *Insurance Markets and Companies*, 17(1), 65-77. [https://doi.org/10.21511/ins.17\(1\).2026.05](https://doi.org/10.21511/ins.17(1).2026.05)
 18. Hoffman, B. N., Okeniyi, J. O., & Samuel, S. E. (2024). Antecedents of compliance with anti-money laundering regulations in the banking sector of Ghana. *Journal of Risk and Financial Management*, 17(8), Article 373. <https://doi.org/10.3390/jrfm17080373>
 19. Imeny, V. M., Norton, S. D., Moradi, M., & Salehi, M. (2025). Suspicious activity reporting in emerging economies: The case of Iran. *Journal of Accounting in Emerging Economies*, 15(1), 201-223. <https://doi.org/10.1108/JAEE-04-2023-0101>
 20. Jayasekara, S. D. (2021). How effective are the current global standards in combating money laundering and terrorist financing? *Journal of Money Laundering Control*, 24(2), 257-267. <https://doi.org/10.1108/JMLC-05-2020-0047>
 21. Johannessen, F., & Jullum, M. (2025). Finding money launderers using heterogeneous graph neural networks. *The Journal of Finance and Data Science*, 11, Article 100175. <https://doi.org/10.1016/j.jfds.2025.100175>
 22. Kim, H., Lee, B. S., Shin, W.-Y., & Lim, S. (2022). Graph anomaly detection with graph neural networks: Current status and challenges. *IEEE Access*, 10, 111820-111829. <https://doi.org/10.1109/ACCESS.2022.3211306>
 23. Korauš, A., Jančíková, E., Gombár, M., Kurilovská, L., & Černák, F. (2024). Ensuring financial system sustainability: Combating hybrid threats through anti-money laundering and counter-terrorist financing measures. *Journal of Risk and Financial Management*, 17(2), Article 55. <https://doi.org/10.3390/jrfm17020055>
 24. Kungu, C. O., Senagi, K., & Omondi, E. (2026). Hybrid deep learning for anti-money laundering: Unsupervised detection of emerging schemes via feature fusion and explainable artificial intelligence. *Machine Learning with Applications*, 23, Article 100856. <https://doi.org/10.1016/j.mlwa.2026.100856>
 25. Liang, W., Li, Y., Liu, X., Dai, Q., & Feng, Y. (2025). AI-based Bayesian structural time series modeling for assessing PM2.5 air quality improvements during the Beijing 2022 Winter Olympics. *Atmospheric Environment*, 358, Article 121328. <https://doi.org/10.1016/j.atmosenv.2025.121328>
 26. Loh, X. (2021). Suspicious activity reports (SARs) regime: Reforming institutional culture. *Journal of Money Laundering Control*, 24(3), 514-524. <https://doi.org/10.1108/JMLC-07-2020-0078>
 27. Mazumder, P. T. (2026). Explainable and fair anti-money laundering models using a reproducible SHAP framework for financial institutions. *Discover Artificial Intelligence*, 6, Article 262. <https://doi.org/10.1007/s44163-026-00944-7>

28. McNaughton, K. J. (2024). The tale of the Baltics: Experiences, challenges, and lessons from reforming their national Financial Intelligence Units (FIUs). *Journal of Economic Criminology*, 5, Article 100080. <https://doi.org/10.1016/j.jeconc.2024.100080>
29. Minto, A. (2024). Banks' internal governance obligations vis-à-vis money laundering risks emerging from the new technology-enabled means to transfer funds or value ("crypto assets"). *Journal of Money Laundering Control*, 27(7), 43-59. <https://doi.org/10.1108/JMLC-04-2024-0068>
30. Mohd Padil, H., Kasim, E. S., Mohd Razali, F., Atan, R., & Aminullah, H. (2022). Effect of quality of governance on the relationship between illicit financial flows and economic growth. *Journal of Financial Crime*, 29(4), 1382-1395. <https://doi.org/10.1108/JFC-10-2021-0226>
31. Morshed, A. (2026). Cross-border enforcement and regulatory arbitrage in the crypto era: Evidence from the United States and the United Arab Emirates. *Research in International Business and Finance*, 91, Article 103574. <https://doi.org/10.1016/j.ribaf.2026.103574>
32. Morshed, A., & Khrais, L. T. (2026). Artificial intelligence for sustainable business through logistics: A multi-theory model of green transformation in the Gulf Cooperation Council (GCC). *Transportation Research Interdisciplinary Perspectives*, 39, Article 102209. <https://doi.org/10.1016/j.trip.2026.102209>
33. Musa, K. M., Walsh, C., Akhavi, L., & Rekhteen, N. (2025). The migration of business ethics within artificial intelligence (AI). *Al-Zaytoonah University Journal of Business*, 1(2). <https://doi.org/10.15849/zjib.v1i02.46>
34. Narasapuram, N. K., Shah, S. A. A., Shiratuddin, M. F., & Sohel, F. (2026). Explainable artificial intelligence models for detecting suspicious bank transactions. *International Journal of Machine Learning and Cybernetics*, 17, Article 111. <https://doi.org/10.1007/s13042-026-02994-w>
35. National Crime Agency. (2024). *SARs annual statistical report 2023*. UK Financial Intelligence Unit. Retrieved from <https://www.nationalcrimeagency.gov.uk/who-we-are/publications/710-sars-annual-statistical-report/file.pdf>
36. Norton, S. D. (2024). Suspicious activity reporting in the United Kingdom and the United States: Statutory obligations of auditors and optimal harvesting of information. *Journal of Money Laundering Control*, 27(3), 432-444. <https://doi.org/10.1108/JMLC-08-2023-0131>
37. Ogbeide, H., Thomson, M. E., Gonul, M. S., Onkal, D., Bhowmick, S., & Bello, A. U. (2025). Rethinking experts' perceptions in money laundering risk assessment. *European Journal on Criminal Policy and Research*, 31(4), 753-777. <https://doi.org/10.1007/s10610-024-09586-w>
38. Orekat, A. (2021). The degree of satisfaction of Facebook users about its features, usage motives and achieved gratifications: An applied study on students of the Faculty of Mass Communication at the Middle East University. *Middle East Journal of Communication Studies*, 1(1), 7-35. <https://doi.org/10.71220/2585-001-001-001>
39. Oztas, B., Cetinkaya, D., Adedoyin, F., Budka, M., Aksu, G., & Dogan, H. (2024). Transaction monitoring in anti-money laundering: A qualitative analysis and points of view from industry. *Future Generation Computer Systems*, 159, 161-171. <https://doi.org/10.1016/j.future.2024.05.027>
40. Pocher, N., Zichichi, M., Merizzi, F., Shafiq, M. Z., & Ferretti, S. (2023). Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics. *Electronic Markets*, 33, Article 37. <https://doi.org/10.1007/s12525-023-00654-3>
41. Pontes, R., Lewis, N., McFarlane, P., & Craig, P. (2022). Anti-money laundering in the United Kingdom: New directions for a more effective regime. *Journal of Money Laundering Control*, 25(2), 401-413. <https://doi.org/10.1108/JMLC-04-2021-0041>
42. Premkumar, A., Vidyavathi, K., Sowmya, V. L., Alfurhood, B. S., Irudayasamy, A., & Natarajan, R. (2026). Leveraging artificial intelligence for enhanced anti-money laundering in financial institutions. *Discover Computing*, 29, Article 240. <https://doi.org/10.1007/s10791-026-10142-9>
43. Reite, E. J., Karlsen, J., & Westgaard, E. G. (2025). Improving client risk classification with machine learning to increase anti-money laundering detection efficiency. *Journal of Money Laundering Control*, 28(1), 93-107. <https://doi.org/10.1108/JMLC-03-2024-0040>
44. Riccardi, M. (2025). A data-driven approach to measure money laundering risk and its relationship with corruption. *Journal of Economic Criminology*, 10, Article 100184. <https://doi.org/10.1016/j.jeconc.2025.100184>
45. Ruggiero, V. (2022). Sustainability and financial crime. *International Criminology*, 2(2), 143-151. <https://doi.org/10.1007/s43576-022-00052-3>
46. Saxena, C. (2024). Identifying transaction laundering red flags and strategies for risk mitigation. *Journal of Money Laundering Control*, 27(6), 1063-1077. <https://doi.org/10.1108/JMLC-11-2023-0182>
47. Taqa, S. B. A. (2025). The mediating role of remote communication on the relationship between electronic human resource management practices and organizational performance in Iraqi commercial banks. *Middle East Journal of Communication Studies*, 5(1), 1-52. (In Arabic). Retrieved from <https://mej.meu.edu.jo/?p=2068>
48. Thakkar, H., Datta, S., Bhadra, P., Dabhade, S. B., Barot, H., & Junare, S. O. (2024). Mapping the knowledge landscape of money laundering for terrorism financing: A bibliometric analysis. *Journal of Risk and Financial Management*, 17(10), Article 428. <https://doi.org/10.3390/jrfm17100428>

49. Tiwari, M., Gepp, A., & Kumar, K. (2025). Using data analytics to distinguish legitimate and illegitimate shell companies. *Journal of Economic Criminology*, 7, Article 100123. <https://doi.org/10.1016/j.jeconc.2024.100123>
50. Tosza, S., & Voordeckers, O. (2024). An anti-money laundering authority for the European Union: A new center of gravity in AML enforcement. *ERA Forum*, 25, 405-424. <https://doi.org/10.1007/s12027-024-00805-9>
51. Tuhirirwe, C., & Alexander, R. (2025). Efficacy of country legal frameworks and international guidelines in curtailing money laundering and terrorist financing activities in grey list countries: Case studies of Kenya and Uganda. *Journal of Economic Criminology*, 8, Article 100153. <https://doi.org/10.1016/j.jeconc.2025.100153>
52. Turksen, U., Benson, V., & Adamyk, B. (2024). Legal implications of automated suspicious transaction monitoring: Enhancing integrity of AI. *Journal of Banking Regulation*, 25(4), 359-377. <https://doi.org/10.1057/s41261-024-00233-2>
53. UK Government. (2020). *The Companies etc. (Filing Requirements) (Temporary Modifications) Regulations 2020*. GOV.UK. <https://www.gov.uk/government/publications/the-companies-etc-filing-requirements-temporary-modifications-regulations-2020>
54. Umar, B. (2023). Effects of trade misinvoicing on money laundering in developing economies. *Journal of Money Laundering Control*, 26(1), 60-68. <https://doi.org/10.1108/JMLC-11-2021-0124>
55. Volten, C., van Eeten, M., & van Wegberg, R. (2025). Money for nothing, supervision for a fee: Investigating the effects of the 5th Anti-Money Laundering Directive on cryptocurrency exchanges in the Netherlands. *European Journal on Criminal Policy and Research*. <https://doi.org/10.1007/s10610-025-09640-1>
56. Wang, H.-M., & Hsieh, M.-L. (2024). Cryptocurrency is new vogue: A reflection on money laundering prevention. *Security Journal*, 37(1), 25-46. <https://doi.org/10.1057/s41284-023-00366-5>
57. Weber, M., Domeniconi, G., Chen, J., Weidele, D. K. I., Bellei, C., Robinson, T., & Leiserson, C. E. (2019). *Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics* (arXiv:1908.02591). arXiv. <https://doi.org/10.48550/arXiv.1908.02591>
58. Wronka, C. (2022). "Cyber-laundering": The change of money laundering in the digital age. *Journal of Money Laundering Control*, 25(2), 330-344. <https://doi.org/10.1108/JMLC-04-2021-0035>
59. Wronka, C. (2023). Financial crime in the decentralized finance ecosystem: New challenges for compliance. *Journal of Financial Crime*, 30(1), 97-113. <https://doi.org/10.1108/JFC-09-2021-0218>
60. Zhang, Q., Liu, L., Lan, T., Chen, W., Chen, R., & Zhang, X. (2026). RACLA: Role-aware continual learning for robust AML detection. *Expert Systems with Applications*, 299, Article 130136. <https://doi.org/10.1016/j.eswa.2025.130136>

APPENDIX A. Banking AML verification and model-governance trail

Appendix Table A1 maps each reported result to the smallest evidence layer needed for inspection. The submission workbook contains only the table-level quantities retained in the current paper and identifies the underlying files that are not included. Appendix A3 therefore separates submission-facing evidence from the additional record-, prediction-, and model-level materials required for full raw-to-result replication.

Table A1. Result-to-evidence audit trail

Manuscript table	Banking AML result type	Verification evidence	Banking model-governance control	Manuscript action
Tables 1 and 3	External AML data counts and coverage	Public/source documentation plus the reported analytical denominator used in Table 4	Source roles are kept distinct; unavailable intermediate counts are not treated as independently verified from underlying records	Retain only reconciled counts; avoid implying bank-internal population coverage
Table 4	PSC ownership/control verification-workload counts	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx: PSC_RECONCILIATION	Category counts and denominator are arithmetically consistent; record-level assignments are not included	Documentation workload only; not raw prevalence, legal BO, or suspicion
Table 5	Crypto benchmark AUROC/AUPRC point ordering	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx: ELLIPTIC_METRICS	Reported point metrics are cross-mapped; prediction-level outputs and paired inference are not included	Descriptive ordering unless paired uncertainty supports more
Table 6	Late-June 2020 boundary / PSC-index temporal deviation	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx: STUDY3_SUMMARY	Reported posterior mean and credible interval are cross-mapped; the underlying model is not rerun	Report as register-level temporal context; no intervention effect
Table 7	Robustness, explainability, and model-governance records	No quantitative robustness or explainability output is used as independent submission evidence	Unsupported auxiliary summaries are excluded from the quantitative evidence path	Use as model-governance support, not separate empirical proof
Table 8	Banking-focused hypothesis summary	Derived only from Tables 4-7	Repository version and file accessibility must match the submitted manuscript package	State only the reproducibility supported by deposited and accessible materials
Repository or supplementary statement	Reproducibility statement	AML_SUBMISSION_EVIDENCE_CLEAN.xlsx	Submission supplement provides table-level evidence mapping; it is not described as a complete raw-data repository	State the documented level of reproducibility supported by the deposited materials

Table A2. Reconciled PSC count logic for banking AML interpretation

Stage	Count used in manuscript	Interpretation	Where reported
Active companies/entities after aggregation	6.2M	Reported company denominator used for Table 4 percentages	Table 1; Table 3; Table 4
Final flagged companies	745.2K companies	Reported PSC_Flag total; four mutually exclusive workload-category counts reconcile exactly to 745,200	Table 4; Table 8

Table A3. Minimum supplementary verification package

Reproducibility source / material	Verification status and required archived/deposited item	Tables supported
Submission evidence workbook	SUPPLIED: AML_SUBMISSION_EVIDENCE_CLEAN.xlsx. Contains README, evidence dashboard, PSC reconciliation, Elliptic point metrics, Study 3 posterior summary, manuscript crosswalk, and data-availability map	Tables 4-6; Appendix A

Table A3 (cont.). Minimum supplementary verification package

Reproducibility source / material	Verification status and required archived/deposited item	Tables supported
PSC table-level evidence	SUPPLIED: 6.2M denominator, four workload-category counts, 745,200 flagged total, and formula-based reconciliation. NOT INCLUDED: company-level assignments, linkage/overlap logs, and exact historical source snapshot	Table 4
Elliptic table-level evidence	SUPPLIED: reported AUROC/AUPRC point metrics for XGBoost, GCN, GraphSAGE, and GAT. NOT INCLUDED: held-out prediction probabilities, exact final split/run files, or paired-resampling outputs	Table 5
Study 3 table-level evidence	SUPPLIED: reported posterior mean deviation -0.068 and 95% CrI [-0.109, -0.024]. NOT INCLUDED: monthly PSC-index series, final state-space configuration, posterior draws, pre-period fit, or dated placebos	Table 6
Robustness / explainability boundary	No quantitative robustness or explainability output is used as an independent empirical result in the current submission. Deposit only outputs actually generated by the final analysis if such claims are retained later	Table 7
Full raw-to-result replication boundary	Not claimed. Complete replication would require the record-, prediction-, and model-level files listed above	Complete empirical replication

Table A3 distinguishes table-level supplementary evidence from complete raw-to-result replication. The current workbook documents the former for the reported values in Tables 4-6; the latter requires the record-, prediction-, and model-output files listed in the table.