© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722410

Internal-Control Analytics and Audit-Informed Process Improvement: A Scalable Governance Model for Growth-Stage Organizations

ASHLEY MUNASHE SHAMBARE[1], NELIA MLAMBO[2], MELODY RACHAEL CHITUKUTUKU[3], DELINE KUFANDADA[4], MUNASHE NAPHTALI MUPA[5]

[1]*Bentley University, ORCiD: 0009-0003-8458-1825*

[2]*University of Northern Iowa, ORCiD: 0009-0002-6348-8560*

[3]*American University, ORCiD: 0009-0003-4780-4540*

[4]*Illinois State University, ORCiD: 0009-0009-3675-4959*

[5]*Hult International Business School, ORCiD: 0000-0003-3509-861X*

Abstract- ***This study develops and empirically demonstrates a scalable governance model that links transaction-level internal-control analytics to audit-informed process improvement in growth-stage organizations. The analysis uses the Online Retail II transaction dataset, distributed through Kaggle and originally archived by the UCI Machine Learning Repository, comprising 1,067,371 transaction lines, 53,628 invoices, and 43 countries between December 2009 and December 2011. Rather than treating reversals, data-quality defects, and transactional irregularities as isolated bookkeeping problems, the study frames them as governance signals that can be translated into practical control redesign, escalation logic, and dashboard-based oversight. Six forward-looking control-gap indicators were constructed missing customer identifiers, duplicate lines, price outliers, quantity outliers, non-positive prices, and extreme-value invoices together with a separate reversal indicator for credit-note intensity. The study combines full-population descriptive analytics, monthly and country heatmaps, issue co-occurrence analysis, invoice-level anomaly detection, and rank-based association testing. The results show that 36.16% of all transaction lines contained at least one control gap. Missing customer identifiers (22.77%) and price outliers (15.03%) were the most prevalent issues, followed by quantity outliers (7.91%), extreme-value invoices (7.48%), and duplicate lines (6.30%). Temporal concentration was evident: December 2010 recorded the highest average control-gap score (31.42), while December 2011 exhibited the highest return-to-gross ratio (32.11%), signalling heightened pressure on reconciliation and approval processes during peak-cycle periods. Geographic hotspot analysis showed the highest composite control-gap indices among the Netherlands and the United Kingdom in the high-volume country set. A co-occurrence heatmap further revealed that missing customer identifiers and price outliers appeared together on 10.89% of all lines, indicating that master-data weakness and pricing-control weakness frequently interacted. As a secondary validation step, an Isolation Forest flagged 5.00% of invoices as anomalous; flagged invoices showed an average of 2.79 control gaps compared with 1.01 for non-flagged invoices, and anomaly scores were strongly associated with invoice gap counts (Spearman rho = 0.624, $p < 0.001$). The study contributes a governance architecture that translates exception analytics into five operational stages: data capture and standardisation, continuous exception monitoring, risk-based audit triage, control redesign with ownership, and dashboard monitoring for board and management review. The findings support the view that internal audit becomes more valuable when it moves beyond periodic sampling and instead uses full-population analytics to direct scarce remediation capacity toward the most consequential process weaknesses. For growth-stage organizations, the practical implication is clear: scalable governance does not require heavyweight bureaucracy, but it does require disciplined exception logic, threshold-based approvals, and a closed-loop link between audit findings and process owners.***

Keywords: internal control analytics; internal audit; governance; process improvement; anomaly detection; growth-stage organizations; Online Retail II

I. INTRODUCTION

Growth-stage organizations routinely scale transaction volumes faster than they scale governance routines. This creates a familiar pattern: sales and operational data expand rapidly, while control

ownership, approval thresholds, master-data discipline, and reconciliation procedures mature more slowly. The problem is not merely procedural. In the United States, small businesses account for 99.9% of firms and 45.9% of employment, which means that weak financial governance in smaller and scaling entities has system-wide consequences for jobs, resilience, and continuity of commerce (U.S. Small Business Administration Office of Advocacy, 2025). COSO has similarly emphasized that internal control should be understood not only as a compliance instrument but as a means of supporting strategic execution and sustained growth with confidence and integrity (COSO, 2026).

The professional standards environment points in the same direction. The Institute of Internal Auditors' Global Internal Audit Standards position internal audit as a board-authorised function that should communicate results clearly, strengthen governance, and monitor action plans (The IIA, 2024). The Association of Certified Fraud Examiners reports that internal audit accounted for 14% of occupational-fraud detections in its 2024 global dataset, while organizations using management review, account reconciliation, surveillance or monitoring, internal audit, and automated transaction monitoring recorded both lower losses and faster detection (ACFE, 2024). Taken together, these sources suggest that full-population control analytics should no longer be treated as optional sophistication; they are increasingly central to credible governance design.

Yet many growth-stage organizations still rely on periodic sampling, spreadsheet-based review, and reactive control repair. That approach is increasingly fragile in environments characterised by high transaction velocity, cross-border customer activity, manual pricing adjustments, return flows, and fragmented data capture. Recent research has shown that data analytics can materially improve internal-audit effectiveness, process visibility, and risk-based planning (Álvarez-Foronda et al., 2023; Duan, Vasarhelyi and Codesso, 2025). Related work has also shown that internal-control quality contributes to better accounting information quality, and that stronger internal-control components matter materially for fraud prevention in SMEs (Liu et al., 2024; Azizan and Ali, 2024; Hossain, 2025). What remains less developed is a practical bridge between transaction-level exception patterns and the design of a right-sized governance model that growth-stage organizations can operationalise without excessive administrative burden.

This study addresses that gap by using a large transactional dataset to ask a practical question: how can internal-control analytics convert audit observations into scalable process improvement and governance action? The contribution is twofold. First, the study demonstrates how a full-population transaction file can be reinterpreted as an internal-control evidence base rather than simply a revenue history. Second, it proposes an audit-informed governance cycle that links exception analytics to ownership, remediation, and monitoring. The analysis is framed for growth-stage organizations because these entities often face the sharpest tension between expansion and control maturity.

II. LITERATURE REVIEW AND CONCEPTUAL FRAMING

Internal control and internal audit are most useful when they are treated as mutually reinforcing governance capabilities rather than as separate compliance functions. Liu et al. (2024) show that higher internal-audit quality is associated with improved accounting information comparability, indicating that control strength affects not only fraud prevention but also the reliability and interpretability of reported performance. Azizan and Ali (2024), studying SMEs, find that all five core internal-control dimensions control environment, control activities, risk assessment, information and communication, and monitoring play substantial roles in preventing employee fraud. Hossain (2025) reaches a similar conclusion and places particular emphasis on segregation of duties and monitoring systems as especially effective components for fraud prevention in SMEs.

The role architecture around internal control is also changing. Mupa et al. (2024) argue that management accountants increasingly contribute to risk management and internal-control design in

environments facing strategic, technological, and regulatory complexity. Netshifhefhe et al. (2024) extend this argument by showing that integrating internal auditing and legal compliance can yield a more complete risk profile, improve transparency, and strengthen fraud prevention. Adebiyi, Nwokedi and Mupa (2025) further show that internal controls are central to the sustainability of small and medium-sized enterprises, especially where resource constraints make weak cash-flow discipline and ad hoc decision-making particularly costly.

The analytics literature has moved beyond arguing that data tools are useful and has begun to specify how they can be embedded into audit work. Álvarez-Foronda et al. (2023) propose a practical implementation model for internal-audit analytics built around the identification of need, cost-benefit analysis, implementation decisions, and follow-up monitoring. Jans and Eulerich (2022) show that process mining is valuable because it reconstructs how processes are actually executed and enables auditors to evaluate whether underlying process controls are functioning as intended. Duan, Vasarhelyi and Codesso (2025) advance this logic by integrating process mining and machine learning into a model that identifies control weaknesses, missing controls, and anomalous areas at the transaction level. In a related risk-based framing, Taanisa et al. (2026) argue that audit analytics can strengthen SME financial reporting by moving anomaly detection earlier in the control cycle.

A common implication across these studies is that high-value internal audit increasingly depends on full-population review, exception prioritisation, and a clear link between risk signals and management action. What is still needed, however, is a governance model suited to organizations that are large enough to need structured control analytics but not so mature that they can sustain a heavyweight assurance bureaucracy. This study therefore adopts a design logic aimed at scalability: identify a limited set of interpretable control-gap indicators, map their temporal and geographic concentration, validate their clustering at invoice level, and translate the findings into a governance operating model that finance, internal audit, and operational leadership can jointly own.

III. MATERIALS AND METHODS

The study used the Online Retail II dataset, distributed through Kaggle and originally archived by the UCI Machine Learning Repository (Kaggle, n.d.; UCI Machine Learning Repository, 2019). The dataset contains all transactions recorded by a UK-based non-store online retailer between 1 December 2009 and 9 December 2011. Many customers are wholesalers, making the dataset particularly suitable for examining the kinds of transaction-volume, customer-traceability, and cross-border processing challenges common in growth-stage organizations.

Both source worksheets (2009–2010 and 2010–2011) were combined into a single analytical file. Data types were standardised, invoice dates were converted to time stamps, and derived fields were created for invoice month, gross line amount, absolute transaction amount, invoice-level totals, and country-level activity. Importantly, cancellation and reversal lines were retained rather than dropped, because they represent useful control signals for audit analysis. This full-population approach produced 1,067,371 transaction lines and 53,628 unique invoices.

Six forward-looking control-gap indicators were then constructed. First, missing customer identifier flagged lines with no customer ID, representing weakened traceability for downstream review. Second, duplicate line flagged repeated invoice-stock-quantity-price-date combinations, which may reflect order-entry or interface duplication. Third, price outlier flagged positive selling prices falling outside stock-code-specific interquartile thresholds. Fourth, quantity outlier applied the same logic to positive quantities. Fifth, non-positive price captured zero-priced or negatively priced lines. Sixth, extreme-value invoice flagged invoice totals above the 99th percentile of absolute invoice value, representing transactions requiring stronger approval discipline. Credit-note or reversal invoices were tracked separately, because reversals often reflect the

downstream manifestation of control breakdowns rather than the original control gaps themselves.
A composite control-gap score was used to prioritise issues at line level:

$$CGS_i = 25(MCI_i) + 20(DUP_i) + 20(PO_i) + 15(QO_i) + 10(NPP_i) + 10(EVI_i)$$

where MCI denotes missing customer identifier, DUP duplicate line, PO price outlier, QO quantity outlier, NPP non-positive price, and EVI extreme-value invoice. Each component takes the value 1 if present and 0 otherwise. The weights were set to reflect practical audit severity: identifiers and duplication received relatively high weight because they directly affect traceability and record integrity, while quantity and pricing anomalies were treated as operationally material but somewhat more contextual.

The analysis proceeded in four layers. First, descriptive statistics established the prevalence and financial significance of control signals. Second, monthly and country heatmaps were used to identify temporal concentration and geographic hotspots. Third, an issue co-occurrence matrix was used to assess whether control weaknesses clustered rather than appearing independently. Fourth, invoice-level anomaly validation was conducted using an Isolation Forest trained on invoice features including duplication, outlier flags, missing identifiers, non-positive prices, line counts, and invoice totals. A contamination rate of 5% was selected to isolate the most unusual invoices without over-labeling ordinary operational variability. Finally, Spearman rank correlation was used to test the association between anomaly scores and invoice gap counts.

IV. RESULTS

Table 1 summarises the analytical profile of the dataset. The file spans 43 countries, 5,305 stock codes, and positive sales value of £20.97 million, against absolute return and credit-note value of £1.53 million. Net transaction value across the observation period was £19.29 million. Most importantly from a governance perspective, 36.16% of all transaction lines contained at least one control gap. This immediately indicates that control analytics in this setting is not about rare anomalies alone; it is about prioritising a large exception surface in a disciplined way.

Table 1. Analytical profile of the study dataset

Metric	Value
Observation period	01 Dec 2009 to 09 Dec 2011
Transaction lines analysed	1,067,371
Unique invoices	53,628
Unique stock codes	5,305
Countries	43
Positive sales value (£)	20,972,594.57
Absolute credit-note / return value (£)	1,527,415.00
Net transaction value (£)	19,287,250.57
Share of lines with at least one control gap (%)	36.16

Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Table 2 shows the prevalence of specific control indicators. Missing customer identifiers were the dominant weakness at 22.77% of all lines, followed by price outliers at 15.03%. Quantity outliers (7.91%), extreme-value invoices (7.48%), and duplicate lines (6.30%) were also material. Credit-note lines accounted for 1.83% of all observations, while non-positive prices represented 0.58%. The pattern matters because these signals correspond to distinct governance concerns: traceability, pricing discipline, dispatch or ordering discipline, approval thresholds, data integrity, and reversal management.

Table 2. Control indicators, audit rationale and prevalence

Control indicator	Audit rationale	Lines flagged	Share of lines (%)
Missing customer identifier	Weakens traceability, reconciliation, customer-level review, and accountability.	243,007	22.77

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722410

Price outlier by stock code	Signals pricing override, master-data error, or approval breakdown.	160,426	15.03
Quantity outlier by stock code	Signals dispatch, ordering, or return-authorisation irregularity.	84,436	7.91
Duplicate transaction line	May indicate duplicate capture, interface error, or posting redundancy.	67,246	6.30
Credit note / reversal invoice	Represents downstream correction activity requiring root-cause review.	19,494	1.83
Non-positive selling price	Signals data-quality weakness, exceptional pricing treatment, or misposting.	6,207	0.58
Extreme-value invoice	Requires higher approval discipline because exposure is unusually concentrated.	79,868	7.48

Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Temporal concentration is evident in Figure 1 and Figure 2. The monthly line series shows that control-gap intensity does not move mechanically with revenue; instead, it spikes during specific operational periods. December 2010 produced the highest average control-gap score (31.42), while December 2011 recorded the highest return-to-gross ratio (32.11%), indicating substantial year-end reversal or correction activity. The month-by-issue heatmap shows that duplicate lines surged sharply in December 2010, while missing customer identifiers were persistently elevated in several periods, particularly December 2009, December 2010, January 2011, July 2011, and December 2011. The broad implication is that peak-cycle activity and closing-period pressure amplify the need for stronger validation, approval, and review controls.

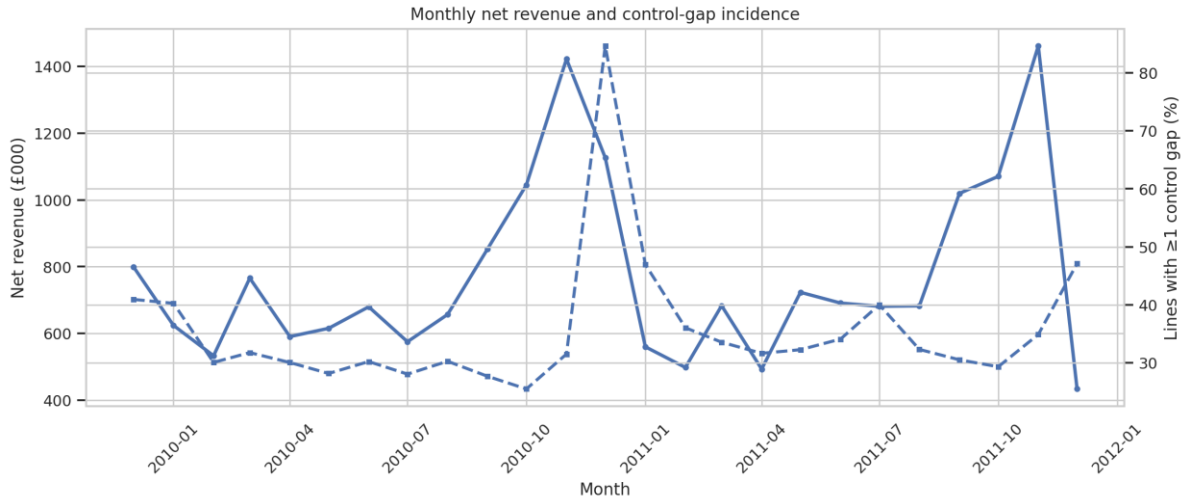

Figure 1. Monthly net revenue and control-gap incidence. Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722410

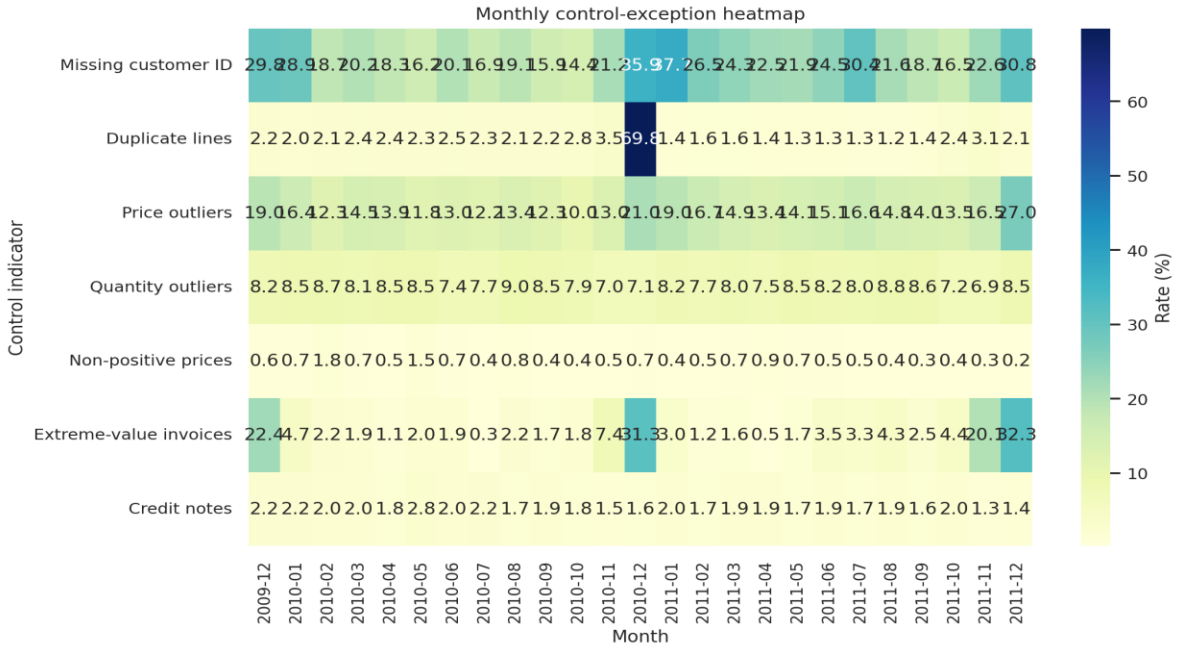

Figure 2. Monthly control-exception heatmap. Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Table 3. Highest-risk months by average control-gap score

Month	Average gap score	Any control-gap rate (%)	Credit-note rate (%)	Return-to-gross (%)	Net revenue (£)
2010-12	31.42	84.66	1.63	10.78	1,126,445.47
2011-12	18.06	47.11	1.41	32.11	433,686.01
2009-12	15.20	40.93	2.24	3.13	799,847.11
2011-01	15.06	46.96	1.99	19.00	560,000.26
2011-07	12.73	40.01	1.73	5.27	681,300.11

Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Country-level results reinforce the hotspot logic. Among countries with at least 1,000 lines, the highest control-gap indices were recorded for the Netherlands and the United Kingdom, followed by Australia, Sweden, and Norway. The Netherlands showed particularly intense quantity outlier behaviour, while the United Kingdom combined very high transaction volume with elevated duplication and missing-customer frequency. Norway, although smaller in volume, recorded a notably high return-to-gross ratio, signalling that even lower-volume

geographies can warrant targeted review when reversal intensity is high. Figure 3 and Table 4 therefore suggest that governance in growth-stage organizations should not assume that the highest-volume channel is the only risk hotspot; moderate-volume channels with concentrated exception patterns may require disproportionate audit attention.

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722410

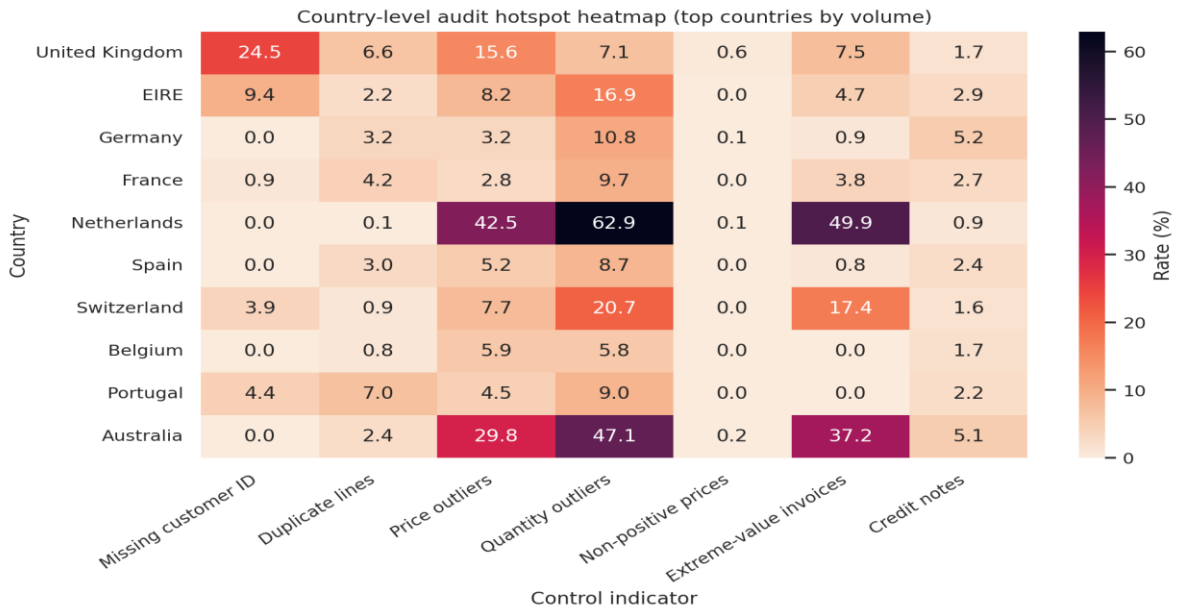

Figure 3. Country-level audit hotspot heatmap. Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Table 4. Country hotspots in the high-volume country set

Country	Lines	Credit-note rate (%)	Missing customer ID (%)	Duplicate-line rate (%)	Control-gap index
Netherlands	5,140	0.91	0.00	0.12	17.99
United Kingdom	981,330	1.70	24.46	6.58	17.91
Australia	1,913	5.12	0.00	2.40	13.53
Sweden	1,364	1.91	1.39	0.29	9.97
Norway	1,455	1.24	0.00	20.21	9.04
EIRE	17,866	2.87	9.35	2.23	7.58
Switzerland	3,189	1.63	3.92	0.88	6.01
Portugal	2,620	2.21	4.43	6.98	5.00
France	14,330	2.71	0.89	4.24	3.22
Finland	1,049	1.62	0.00	0.00	3.10

Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

The co-occurrence heatmap in Figure 4 shows that control weaknesses often interact. Missing customer identifiers and price outliers appeared together on 10.89% of all lines, making this the strongest paired signal in the dataset. Missing customer identifiers also co-occurred with extreme-value invoices on 6.10% of lines, while price and quantity outliers overlapped on 3.17%. These combinations are operationally important because they indicate that control failures are not isolated. When a high-value invoice also lacks strong customer traceability or

pricing discipline, the governance response should escalate beyond routine reconciliation toward focused review, approval redesign, and ownership assignment.

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722410

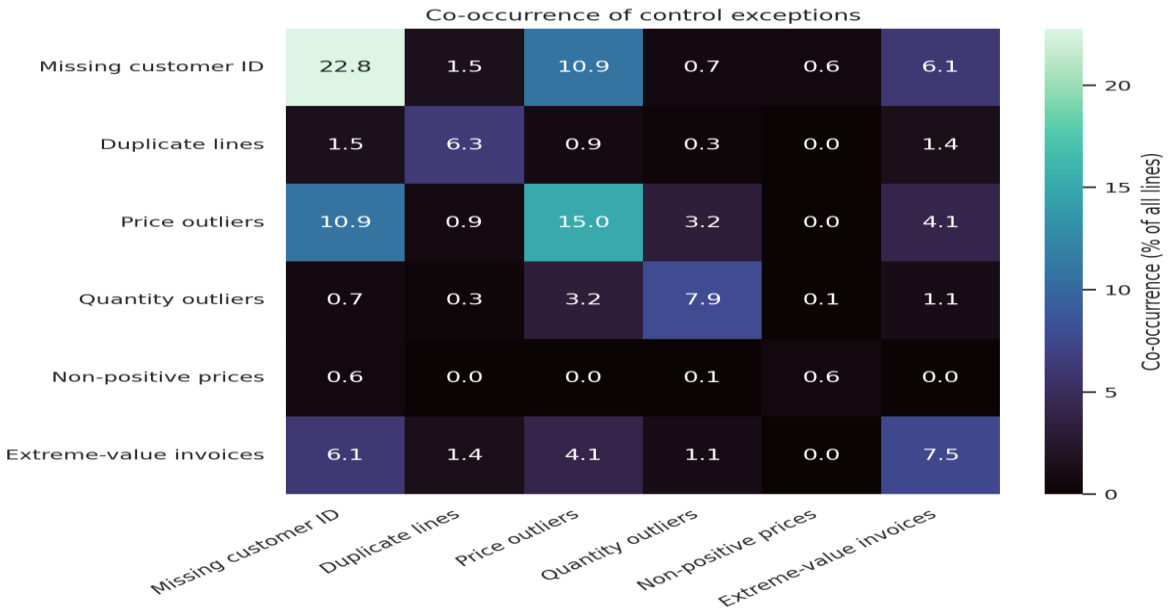

Figure 4. Co-occurrence of control exceptions. Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Invoice-level anomaly validation provides a useful secondary check. The Isolation Forest flagged 2,682 invoices, or exactly 5.00% of the invoice population, as anomalous. Flagged invoices showed a mean control-gap count of 2.79, compared with 1.01 among non-flagged invoices. The mean anomaly score for flagged invoices was 0.661, versus 0.463 across the full invoice population. The relationship between anomaly score and invoice gap count was strongly positive (Spearman rho = 0.624, $p < 0.001$), as illustrated in Figure 5 and summarised in Table 5. This result does not prove fraud or misstatement; rather, it shows that the control-gap architecture captures structurally unusual invoice behaviour in a way that is highly relevant to risk-based audit triage.

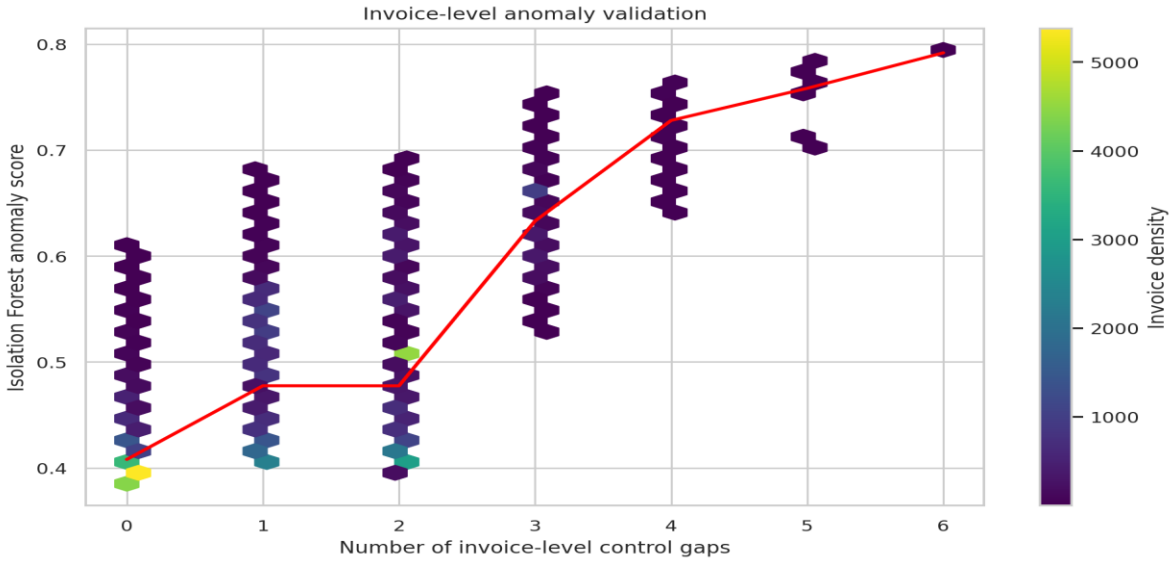

Figure 5. Invoice-level anomaly validation. Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

Table 5. Invoice-level anomaly validation summary

Validation metric	Result
Invoices analysed	53,628
Invoices flagged anomalous	2,682
Anomaly share (%)	5.00
Mean anomaly score (all invoices)	0.463
Mean anomaly score (flagged invoices)	0.661
Mean control-gap count (flagged invoices)	2.79
Mean control-gap count (non-flagged invoices)	1.01
Spearman correlation: anomaly score vs. gap count	rho = 0.624, p < 0.001

Source: Analysis of the Online Retail II dataset (Kaggle, n.d.; UCI Machine Learning Repository, 2019).

V. DISCUSSION

The findings support a simple but important proposition: growth-stage organizations do not fail governance because they lack policies on paper; they fail because transactional complexity outruns the design and monitoring of everyday controls. The prevalence of missing customer identifiers shows how basic master-data weakness can weaken the entire downstream control chain, from reconciliation and exception investigation to credit-note review and customer-level profitability analysis. The prominence of pricing and quantity outliers suggests that operational variance and manual overrides need more formal tolerance logic. The clustering of issues in peak-cycle months suggests that capacity strain, deadline pressure, and uneven supervision are central control risks during growth.

These results are consistent with prior literature. Survey-based evidence has shown that strong internal-control components reduce fraud exposure in SMEs (Azizan and Ali, 2024; Hossain, 2025), while internal-audit quality contributes to more reliable accounting information (Liu et al., 2024). The present study extends that literature by showing how those principles appear in a large transaction population. The study also aligns with the data-analytics literature, which has argued that internal audit should move toward full-population, technology-enabled review and continuous follow-up (Álvarez-Foronda et al., 2023; Jans and Eulerich, 2022; Duan, Vasarhelyi and Codesso, 2025). Rather than viewing exception analytics as an isolated audit technique, the evidence here supports treating it as a governance operating system.

For practice, the most useful output is not the detection of anomalies per se, but the conversion of anomalies into process redesign. Figure 6 presents the governance cycle proposed in this study, while Table 6 translates the main audit signals into concrete control actions. The model is intentionally right-sized for growth-stage organizations. It starts with structured data capture and standardisation, because no analytics layer can compensate for weak underlying fields. It then moves to continuous exception monitoring, risk-based audit triage, control redesign with named owners, and dashboard monitoring for management and board review. This design is consistent with COSO's view of internal control as a support for strategy and sustained growth (COSO, 2026) and with the IIA's emphasis on communication, risk-based assurance, and monitoring of action plans (The IIA, 2024).

VI. MANAGERIAL IMPLICATIONS AND GOVERNANCE MODEL

Three managerial implications follow. First, organizations should treat master-data completeness as a control issue, not just an administrative inconvenience. Second, approval architecture should become risk-weighted: extreme-value invoices, unusual prices, and concentrated reversals should trigger higher levels of review automatically. Third, internal audit and finance should share a common exception dashboard so that audit findings feed directly into process ownership, KPI monitoring, and board reporting. The literature increasingly recognises this convergence across management accounting, internal audit, compliance, and analytics (Mupa et al., 2024; Netshifhefhe et al., 2024; Taanisa et al., 2026).

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722410

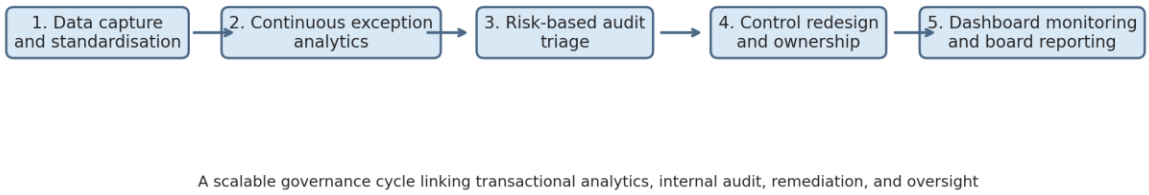

Figure 6. Proposed audit-informed governance cycle for growth-stage organizations. Source: Developed from the study findings.

Table 6. Audit-informed control redesign matrix

Audit signal	Process improvement response	Primary owner	Monitoring KPI
Missing customer identifier	Make customer ID mandatory for account-linked transactions; block posting until validated; add exception log review	Finance operations / sales administration	% lines without customer ID
Duplicate transaction line	Deploy duplicate-match logic at order-entry and posting stage; review recurring source systems	Finance systems owner	Duplicate-line rate
Price outlier	Introduce tolerance bands and secondary approval for manual price overrides	Commercial finance	Price-outlier rate
Quantity outlier	Apply variance thresholds and dual approval for unusual volumes or returns	Operations / warehouse control	Quantity-outlier rate
Credit-note cluster	Require reason codes, matched authorisation, and periodic reversal analytics	Financial controller	Return-to-gross ratio
Extreme-value invoice	Escalate 99th-percentile invoices to dual approval and post-review testing	CFO / delegated approver	Extreme-value invoice rate

Source: Developed from the study findings and informed by COSO (2026), The IIA (2024), Álvarez-Foronda et al. (2023), Duan, Vasarhelyi and Codesso (2025), and Taanisa et al. (2026).

VII. LIMITATIONS AND FUTURE RESEARCH

The study has limitations. The dataset reflects a UK-based online retailer and therefore does not directly represent every industry or governance setting. Credit notes and quantity reversals are not automatically control failures; some represent legitimate commercial returns. The control-gap score is a practical weighting device rather than a universal standard, and alternative weighting schemes could be tested in future research. Even so, the dataset remains highly valuable for governance analysis because it captures a real, high-volume transaction environment with meaningful variation across time, value, and geography.

VIII. CONCLUSION

This study demonstrates that transaction-level internal-control analytics can be translated into a practical governance model for growth-stage organizations. Using more than one million transaction lines, the analysis showed that control weaknesses in traceability, duplication, pricing, quantity variance, and approval concentration are sufficiently prevalent and patterned to justify continuous, full-population review. Control gaps clustered temporally and geographically, co-occurred in meaningful ways, and were strongly aligned with invoice-level anomaly scores. The resulting governance model is deliberately operational:

standardise data, monitor exceptions continuously, triage risk through internal audit, assign remediation ownership, and report through dashboard logic. For growth-stage organizations seeking discipline without bureaucracy, that is the central lesson scalable governance begins with interpretable transaction analytics and becomes durable only when audit findings are converted into managed process change.

REFERENCES

[1] ACFE (2024) Occupational Fraud 2024: A Report to the Nations. Austin, TX: Association of Certified Fraud Examiners.

[2] Adebiyi, O., Nwokedi, A.O. and Mupa, M.N. (2025) 'An Analysis of Financial Strategies, and Internal Controls for the Sustainability of SMMEs in the United States', Iconic Research and Engineering Journals, 8(7), pp. 340-356.

[3] Alqudah, H., Amran, N.A., Hassan, H., Lutfi, A., Alessa, N., Alrawad, M. and Almaiah, M.A. (2023) 'Examining the critical factors of internal audit effectiveness from internal auditors' perspective: Moderating role of extrinsic rewards', Heliyon, 9(10), e20497. doi: 10.1016/j.heliyon.2023.e20497.

[4] Azizan, A.F. and Ali, M.M. (2024) 'The Effect of Internal Control Mechanism Towards Fraud Prevention in Small and Medium Enterprises', Indonesian Journal of Sustainability Accounting and Management, 8(1), pp. 178-188. doi: 10.28992/ijsam.v8i1.843.

[5] COSO (2026) Internal control. Committee of Sponsoring Organizations of the Treadway Commission. Available at: https://www.coso.org/internal-control (Accessed: 21 June 2026).

[6] Duan, H.K., Vasarhelyi, M.A. and Codesso, M. (2025) 'Integrating Process Mining and Machine Learning for Advanced Internal Control Evaluation in Auditing', Journal of Information Systems, 39(1), pp. 55-75. doi: 10.2308/ISYS-2022-028.

[7] Hossain, M.Z. (2025) 'Effectiveness of Internal Control Systems in Preventing Financial Fraud in SMEs', SSRN Electronic Journal. doi: 10.2139/ssrn.5255489.

[8] Jans, M. and Eulerich, M. (2022) 'Process Mining for Financial Auditing', in van der Aalst, W. et al. (eds.) Process Mining Handbook. Lecture Notes in Business Information Processing, vol. 448. Cham: Springer, pp. 445-467. doi: 10.1007/978-3-031-08848-3_15.

[9] Kaggle (n.d.) Online Retail II UCI. Available at: https://www.kaggle.com/datasets/mashlyn/online-retail-ii-uci (Accessed: 21 June 2026).

[10] Liu, G., Wang, J., Sun, Y., Guo, J. and Zhao, Y. (2024) 'Internal audit quality and accounting information comparability: Evidence from China', PLOS ONE, 19(10), e0310959. doi: 10.1371/journal.pone.0310959.

[11] Mupa, M.N., Chiganze, F.R., Mpofu, T.I., Mangeya, R. and Mubvuta, M. (2024) 'The Evolving Role of Management Accountants in Risk Management and Internal Controls in the Energy Sector', Iconic Research and Engineering Journals, 8(2), pp. 859-881.

[12] Netshifhefhe, K., Netshifhefhe, M.V., Mupa, M.N. and Murapa, K.A. (2024) 'Integrating Internal Auditing and Legal Compliance: A Strategic Approach to Risk Management', Iconic Research and Engineering Journals, 8(4), pp. 446-465.

[13] Taanisa, T., Mukwata, N.A., Sydney, J., Ganyani, L., Maturure, R.N., Chingezi, E., Yelduora, P.G. and Mupa, M.N. (2026) 'AI-Enabled Audit Analytics for SME Financial Reporting and Anomaly Detection: A Risk-Based Framework for Early Irregularity Identification and Control Strengthening', World Journal of Advanced Research and Reviews, 30(3), pp. 1113-1126. doi: 10.30574/wjarr.2026.30.3.1596.

[14] The IIA (2024) Global Internal Audit Standards. Lake Mary, FL: The Institute of Internal Auditors.

[15] U.S. Small Business Administration Office of Advocacy (2025) 2025 Small Business

Profile: United States. Washington, DC: U.S. Small Business Administration.

[16] UCI Machine Learning Repository (2019) Online Retail II. Available at: https://archive.ics.uci.edu/ml/datasets/online+retail+II (Accessed: 21 June 2026).

[17] Álvarez-Foronda, R., Ruiz-Martín, C., Laviada, A.F. and Núñez-Carballosa, A. (2023) 'Implementation model of data analytics as a tool for improving internal audit processes', Frontiers in Psychology, 14, 1140972. doi: 10.3389/fpsyg.2023.1140972.