International Journal of Web of Multidisciplinary Studies
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
website: *https://ijwos.com*
Volume 03, Issue 11, November 2026
OPEN ACCESS
E-ISSN : 3049-2424
DOI: https://doi.org/10.71366/ijwos

Forensic Auditing and Occupational Fraud Detection: Uncovering Anomalies Using Multivariate Benford's Law and Ratio Analysis

Manasa Mohan Deshpande[1], Niharika B[1], and Dr. Chandrashekar C[2]
[1] *B.Com (Accounting and Taxation), Presidency School of Commerce, Presidency University, Bengaluru, Karnataka, India*
[2] *Assistant Professor in Finance and Taxation, Presidency School of Commerce, Presidency University, Bengaluru, Karnataka, India*

Article Info

Article History:
Published: 13 Nov 2026

Publication Issue:
Volume 03, Issue 11
November 2026

Page Number:
179-186

Corresponding Author:
Niharika B

Abstract:
Occupational fraud poses serious operational and financial risks to businesses, but standard digit-based audit tests often get distorted by sample size and overall ledger structure. This study looks at whether combining several positions of Benford's Law—first-digit, second-digit, and first-two-digit pattern tests—with financial ratio analysis can improve how auditors spot workplace fraud in Indian companies. Because real corporate ledgers with confirmed fraud are rarely available for public research, we generated a synthetic Monte Carlo dataset representing 200 medium-sized companies (160 normal entities and 40 fraud-affected entities covering billing, expense-reimbursement, and payroll schemes). Each company dataset includes 400 transaction entries, with an additional benchmark group of 100 normal companies used for baseline calibration. We analyzed the simulated data using statistical tests including Welch t-tests, MANOVA, ANOVA with Tukey post-hoc comparisons, chi-square tests, and cross-validated logistic regression. The results showed that fraud-affected ledgers deviate significantly more from natural Benford distributions than clean ledgers (MAD = .021 vs. .012; $t(45.6) = 8.64$, $p < .001$, $d = 2.09$). The composite multivariate Benford distance varied across fraud types, identifying billing and payroll schemes much more effectively than expense reimbursement. A calibrated Benford alert threshold strongly correlated with fraud presence (odds ratio = 17.5). Combining Benford digit checks with ratio analysis yielded an overall classification performance (cross-validated AUC) of .912, though its performance boost over the strongest single test remained modest. Overall, these findings highlight that combining digit patterns with financial ratio screening creates a practical, reliable setup for prioritizing audit targets before launching full investigations.
Keywords: Forensic auditing; occupational fraud; Benford's Law; ratio analysis; multivariate anomaly detection

1. Introduction

Occupational fraud—where employees misappropriate company resources—harms financial health, drains profits, and undermines trust in internal governance controls. According to recent industry findings by the Association of Certified Fraud Examiners (ACFE, 2026), workplace asset misappropriation remains one of the most common and costly risks businesses face. In India, corporate fraud concerns and strict reporting requirements under Section 143(12) of the Companies Act, 2013 place clear responsibility on auditors to spot and report suspicious activities early (Gupta & Gupta, 2015).

Historically, fraud discovery has relied heavily on anonymous tips, whistleblower reports, or routine sample testing. Analytical audit procedures offer a valuable addition by scanning entire financial ledgers at once. Benford's Law, first noticed by Newcomb (1881) and later detailed by Benford (1938), describes how leading digits naturally occur in non-constrained numerical datasets. Under this law, the probability that a number starts with digit d is $\log_{10}(1 + 1/d)$,

Copyright © 2026 The Author(s) : This is an open access article under the CC BY license (http://creativecommons.org/licenses/by-nc/4.0/)

meaning the number 1 appears as a leading digit far more frequently than 9. Nigrini (1996, 2012) adapted this statistical property into practical accounting tools, and subsequent research extended digit testing to corporate financial reporting (Amiram et al., 2015).

However, Benford's Law is not a foolproof standalone test. It performs poorly on assigned identifiers, small datasets, or highly constrained transaction limits, and standard significance tests can produce false alarms in very large sample sizes (Durtschi et al., 2004; Cleary & Thibodeau, 2005). Financial ratios provide a different perspective, though models designed for financial statement manipulation do not always catch everyday employee disbursement schemes (Beneish, 1999). This study evaluates whether combining first-digit, second-digit, and first-two-digit deviations into a single Mahalanobis distance measure—used alongside financial ratios—improves fraud detection in corporate disbursement ledgers. The primary goal is to help internal and external auditors flag high-risk transactions for detailed review without generating excessive false positives.

2. Literature Review

2.1 Occupational fraud and behavioral patterns

Cressey (1953) established the classic Fraud Triangle, explaining that financial fraud requires pressure, opportunity, and rationalization. Wolfe and Hermanson (2004) later expanded this concept into the Fraud Diamond by incorporating employee capability. Dorminey et al. (2012) further synthesized these framework developments. In accounting environments, these behavioral drivers leave distinct numerical traces: employees working around approval thresholds might repeatedly split or round purchase orders, while fake expense claims or ghost payroll entries tend to create repetitive, artificial transaction values. Gupta and Gupta (2015) highlighted that oversight gaps and inadequate due diligence in Indian corporate settings increase exposure to these schemes.

2.2 Benford's Law in forensic auditing

The mathematical foundation of Benford's Law rests on scale invariance (Pinkham, 1961) and the convergence properties of mixed numerical distributions (Hill, 1995). Nigrini (1996, 2012, 2020), Nigrini and Mittermaier (1997), and Drake and Nigrini (2000) introduced digit analysis techniques into tax auditing and corporate forensic accounting. Amiram et al. (2015) demonstrated how broad accounting data aligns with these digit expectations. Studies across other disciplines, such as Cho and Gaines (2007), Rauch et al. (2011), and Diekmann (2007), show that inspecting second and combined digit positions provides crucial clues when primary leading digits are masked or altered.

2.3 Methodological challenges and refinements

A known drawback of Benford testing is its sensitivity to dataset structure. As Durtschi et al. (2004) observed, assigned numbers, sequential IDs, or fixed price points naturally violate standard Benford distributions. Moreover, Cleary and Thibodeau (2005) warned that massive sample sizes can make tiny, irrelevant deviations statistically significant, raising Type I error rates. To address this, researchers like Morrow (2014) and Barney and Schulzke (2016) recommend using Mean Absolute Deviation (MAD) effect sizes and benchmark-calibrated thresholds rather than relying solely on raw p-values.

2.4 Ratio analysis and integrated screening

Beneish (1999) created an eight-variable M-Score model to flag earnings manipulation using financial metrics like receivables growth, asset quality, and margin shifts. Dechow et al. (2011) similarly demonstrated that combining financial and operational metrics improves misstatement prediction. While these models target executive-level accounting adjustments, combining ratio metrics with transaction-level digit testing offers a more complete setup for catching lower-level employee disbursement fraud.

3. Research Gaps and Objectives

While existing research covers both digit testing and financial ratios, four key gaps remain: (1) Most forensic studies look at individual digit positions separately rather than evaluating them together; (2) Digit analysis and financial ratios are rarely tested within the same evaluation framework; (3) The relative detectability of different fraud schemes (billing vs. expense claims vs. ghost payroll) is not widely compared; and (4) Obtaining real corporate ledgers with verified fraud is difficult, making transparent simulation essential for testing new detection workflows.

To address these gaps, this study answers five central research questions: RQ1: Do fraud-affected ledgers exhibit higher overall Benford deviation than normal ledgers at the company level? RQ2: Does combining first-digit, second-

digit, and first-two-digit measures yield a stronger detection signal than looking at any single digit alone? RQ3: Does detection accuracy vary across billing, expense, and payroll fraud schemes? RQ4: Does combining Benford metrics with financial ratios improve overall separation between clean and fraud-affected companies? RQ5: How well does a benchmark-calibrated flag control false alarms compared to standard fixed thresholds?

4. Research Methodology

4.1 Dataset generation and experimental design

We used a quantitative simulation design to evaluate the proposed detection methodology under controlled conditions. The simulated dataset represents 200 mid-sized private companies across four Indian business sectors: manufacturing, IT/services, retail/trading, and construction (50 companies per sector). The total sample consists of 160 normal (clean) entities and 40 fraud-affected entities, split across three specific schemes: 18 billing cases, 12 expense reimbursement cases, and 10 payroll cases. Each company ledger contains 400 disbursement records, creating a primary dataset of 80,000 transactions. A separate benchmark panel of 100 clean companies (40,000 transactions) was created exclusively to calibrate distance metrics and establish baseline reference distributions.

4.2 Simulation rules and fraud schemes

Normal ledger amounts were drawn from lognormal distributions tailored to each company, with a 3% rate of routine round-number payments. For fraud-affected companies, between 8% and 22% of total transaction entries were modified according to specific fraudulent behavior patterns:

Group	Sample Size (n)	Data Generation & Scheme Rules
Clean Entities	160	$\log_{10}(\text{amount}) \sim \text{Normal}(\mu, \sigma)$; $\mu \sim U(3.9, 4.6)$, $\sigma \sim U(0.85, 1.25)$; 3% routine round figures.
Billing Scheme	18	60% of fake invoices set between ₹44,000–₹49,990 (just under the ₹50,000 limit); 40% round amounts (₹25,000–₹200,000).
Expense Claims	12	Exaggerated claims ₹100–₹99,999; preferred leading digits 5–8; 50% rounded to nearest ten.
Payroll Scheme	10	Ghost employee salaries creating 3–6 distinct repeating payment figures; $\log_{10}$ mean of 4.4.
Calibration Panel	100	Separate panel of normal companies used to calculate reference vectors and covariance matrices.

4.3 Measurement variables

Expected Benford probabilities were calculated using standard formulas: $P(D_1=d) = \log_{10}(1 + 1/d)$ for first digits, $P(D_2=d) = \sum_k \log_{10}(1 + 1/(10k + d))$ for second digits, and $P(D_1D_2=dd) = \log_{10}(1 + 1/dd)$ for first-two digits. Deviation from expected patterns was evaluated using Mean Absolute Deviation (MAD): $MAD = (1/K) \sum |AP_k - EP_k|$, yielding MAD1, MAD2, and MAD12 for each entity.

To evaluate combined digit shifts, we calculated a composite Mahalanobis distance: $D^2 = (x - \mu_0)^T \Sigma_0^{-1} (x - \mu_0)$, where x is the vector containing [MAD1, MAD2, MAD12], and μ_0 and Σ_0 represent the mean vector and covariance matrix estimated from the 100-entity clean calibration panel. The calibrated flag threshold was set at $D^2 > 7.19$ (corresponding to the 95th percentile of the clean benchmark panel). Financial ratio performance was summarized using eight Beneish-type indices and the overall composite M-Score.

5. Data Analysis and Results

5.1 Descriptive summary

Summary metrics showed clear differences in digit deviations and financial ratios between clean and fraud-affected companies. Billing schemes produced elevated M-Scores (mean = -1.68) compared to normal ledgers (mean = -2.26), driven primarily by shifts in receivables and sales growth indicators.

Metric	Clean (n=160)	Billing (n=18)	Expense (n=12)	Payroll (n=10)
MAD1	.012 (.004)	.024 (.005)	.017 (.006)	.020 (.007)
MAD2	.013 (.003)	.019 (.004)	.012 (.004)	.022 (.008)
MAD12	.0040 (.0004)	.0052 (.0006)	.0042 (.0004)	.0056 (.0006)
Mahalanobis D^2	3.77 (3.23)	27.53 (16.50)	7.57 (8.44)	35.30 (34.13)
DSRI	1.03 (0.13)	1.30 (0.15)	1.11 (0.08)	1.22 (0.18)
SGI	1.13 (0.14)	1.44 (0.18)	1.22 (0.14)	1.32 (0.20)
AQI	1.03 (0.14)	1.12 (0.18)	1.10 (0.13)	1.03 (0.16)
TATA	.018 (.036)	.011 (.039)	.035 (.035)	.029 (.048)
M-score	-2.26 (0.26)	-1.68 (0.33)	-1.99 (0.19)	-1.85 (0.40)

5.2 Leading-digit tests (H1)

At the pooled level, clean ledgers (64,000 total records) yielded an overall MAD of .003 ($\chi^2(8) = 49.80$, $p < .001$). Fraud-affected ledgers (16,000 records) recorded a higher MAD of .011 ($\chi^2(8) = 342.63$, $p < .001$). The fraud ledgers showed a notable spike in leading digit 4 (13.6% observed vs. 9.7% expected), reflecting invoices structured just below the internal ₹50,000 approval limit. These results support Hypothesis 1, demonstrating that fraudulent entries alter natural digit frequencies.

Ledger Group	Total Records	Chi-Square $\chi^2(8)$	p-value	MAD	Conformity Level
Clean Ledgers	64,000	49.80	< .001	.003	Close conformity (< .006)
Fraud Ledgers	16,000	342.63	< .001	.011	Acceptable conformity (.006–.012)

5.3 Entity-level comparison (H2)

Comparing metrics at the company level revealed significant differences between normal and fraud-affected businesses across all individual and combined digit metrics. A overall MANOVA test confirmed significant group separation across the combined digit measures (Wilks' $\Lambda = .546$, $F(3, 196) = 54.35$, $p < .001$), fully supporting Hypothesis 2.

Metric	Fraud Mean (SD)	Clean Mean (SD)	t-statistic (df)	p-value	Effect Size (d)
MAD1	.021 (.006)	.012 (.004)	8.64 (45.6)	< .001	2.09
MAD2	.018 (.006)	.013 (.003)	5.11 (45.3)	< .001	1.25
MAD12	.0050 (.0008)	.0040 (.0004)	7.80 (43.0)	< .001	2.11
ln D^2	2.60 (1.23)	0.98 (0.90)	7.81 (49.9)	< .001	1.66
M-score	-1.81 (0.33)	-2.26 (0.26)	7.87 (51.6)	< .001	1.61

5.4 Typology comparison (H3)

An ANOVA test evaluating composite distance (ln D^2) across fraud types revealed significant differences ($F(3, 196) = 43.19$, $p < .001$, $\eta^2 = .40$). Tukey post-hoc tests demonstrated that billing and payroll schemes created large, easily detectable shifts from clean ledgers. In contrast, expense reimbursement claims were harder to spot using the composite distance alone ($p = .400$ vs. clean), though individual MAD1 testing remained effective. This confirms Hypothesis 3, showing that detection success varies depending on the underlying fraud scheme.

Group Comparison	Diff in ln D^2	p-value (D^2)	Diff in MAD1	p-value (MAD1)
Clean vs. Billing	-2.13	< .001	-.0124	< .001
Clean vs. Expense	-0.43	.400	-.0050	< .001
Clean vs. Payroll	-2.14	< .001	-.0085	< .001
Billing vs. Expense	1.70	< .001	.0073	< .001
Billing vs. Payroll	-0.01	1.000	.0038	.096
Expense vs. Payroll	-1.71	< .001	-.0035	.213

5.5 Classification performance and combined modeling (H4 & H5)

Using the benchmark-calibrated threshold ($D^2 > 7.19$), Benford flag status was strongly associated with actual fraud presence ($\chi^2(1) = 60.32$, $p < .001$, $\varphi = .55$, Odds Ratio = 17.5), supporting Hypothesis 4. The calibrated threshold

achieved 89.4% specificity and a 61.4% positive predictive value (PPV), significantly reducing false alarms compared to a traditional fixed threshold (MAD1 > .015), which yielded an 82.5% specificity and a 52.5% PPV.

Flagging Method	**Chi-Square $\chi^2(1)$**	**Phi (φ)**	**Sensitivity**	**Specificity**	**PPV**
Calibrated ($D^2 > 7.19$)	60.32	.55	67.5%	89.4%	61.4%
Fixed Rule (MAD1 > .015)	55.39	.53	77.5%	82.5%	52.5%

Finally, we evaluated a combined logistic regression model incorporating both multivariate digit distance ($\ln D^2$) and financial ratio performance (M-Score). The combined model achieved an overall cross-validated AUC of .912 (95% CI [.843, .968]). While the combined model delivered the highest overall discrimination, its boost over the single best standalone indicator (MAD1 at .882) fell within the statistical margin of overlap. Thus, Hypothesis 5 is partially supported.

Model Version	**Predictor Metric**	**Odds Ratio [95% CI]**	**CV AUC [95% CI]**
Standalone Digit	MAD1	—	.882 [.811, .941]
Multivariate Digit	$\ln D^2$	5.80 [2.88, 11.69]	.840 [.747, .921]
Financial Ratio	M-Score	5.71 [2.68, 12.17]	.853 [.783, .916]
Combined Model	$\ln D^2$ + M-Score	—	.912 [.843, .968]

Key Analytical Finding: Combining digit pattern analysis with financial ratios significantly improves detection efficiency, achieving a .912 AUC while keeping false positive flags under 11%.

6. Discussion and Practical Implications

The results demonstrate that evaluating digit patterns across multiple positions provides a clearer diagnostic signal than inspecting first digits in isolation. Billing and payroll fraud schemes created strong, distinct anomaly footprints because approval workarounds and ghost-employee payments concentrate specific payment ranges or repeat exact numerical values. Conversely, expense claim fraud proved more subtle, showing that low-value, distributed entries can partially blend into natural transaction distributions.

For practicing auditors and forensic accountants in India, these findings offer several practical insights: First, Benford testing works best as an automated triage tool to filter large ledger datasets down to manageable risk pools before starting manual reviews. Second, audit teams should avoid static cut-offs and instead calibrate alert thresholds against clean historical ledgers or sector benchmarks to minimize costly false alarms. Third, analytical flags must be treated as indicators for follow-up testing rather than conclusive evidence of wrongdoing.

7. Limitations and Future Research

This study has a few operational limitations. First, the analysis relies on synthetic dataset modeling designed to evaluate methodology rather than estimate real-world fraud rates. Second, real-world fraudsters may deliberately structure transactions to bypass standard Benford distributions if they know audit tests are active. Third, the Beneish M-Score was originally designed for executive financial statement fraud rather than lower-level disbursement schemes, requiring careful adaptation.

Future research should focus on testing this combined framework on actual corporate ledgers provided by external audit partners. In addition, researchers can explore machine learning models, duplicate transaction matching, and cost-sensitive threshold tuning to further improve forensic workflows.

8. Conclusion

This study demonstrates that combining multivariate Benford digit metrics with financial ratio analysis creates a robust setup for detecting occupational fraud in corporate disbursement ledgers. By evaluating multiple digit positions jointly and calibrating detection thresholds against clean peer benchmarks, audit teams can flag high-risk transactions with greater precision while keeping false alarms low. When integrated into broader forensic workflows, this combined screening model provides practical support for auditors seeking to satisfy reporting responsibilities and protect organizational assets.

Declarations

• Ethical Considerations: This study used computer-generated synthetic data and did not involve human participants, personal data, or confidential company records.

• Conflict of Interest: The authors declare that they have no competing financial or personal interests that could have influenced this work.

• Funding Details: No specific grant or financial support was received from public, commercial, or non-profit sectors for this research.

Use of AI: AI-based tools were used to assist with language refinement and document presentation. The authors reviewed the final manuscript and take responsibility for its content.

References

1. Amiram, D., Bozanic, Z., & Rouen, E. (2015). Financial statement errors: Evidence from the distributional properties of financial statement numbers. Review of Accounting Studies, 20(4), 1540–1593.
2. Association of Certified Fraud Examiners. (2026). Occupational fraud 2026: A report to the nations. ACFE.
3. Barney, B. J., & Schulzke, K. S. (2016). Moderating 'cry wolf' events with excess MAD in Benford's law research and practice. Journal of Forensic Accounting Research, 1(1), A66–A90.
4. Beneish, M. D. (1999). The detection of earnings manipulation. Financial Analysts Journal, 55(5), 24–36.
5. Benford, F. (1938). The law of anomalous numbers. Proceedings of the American Philosophical Society, 78(4), 551–572.
6. Cho, W. K. T., & Gaines, B. J. (2007). Breaking the (Benford) law: Statistical fraud detection in campaign finance. The American Statistician, 61(3), 218–223.
7. Cleary, R., & Thibodeau, J. C. (2005). Applying digital analysis using Benford's law to detect fraud: The dangers of type I errors. Auditing: A Journal of Practice & Theory, 24(1), 77–81.
8. Cressey, D. R. (1953). Other people's money: A study in the social psychology of embezzlement. Free Press.
9. Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements. Contemporary Accounting Research, 28(1), 17–82.
10. Diekmann, A. (2007). Not the first digit! Using Benford's law to detect fraudulent scientific data. Journal of Applied Statistics, 34(3), 321–329.
11. Dorminey, J., Fleming, A. S., Kranacher, M.-J., & Riley, R. A., Jr. (2012). The evolution of fraud theory. Issues in Accounting Education, 27(2), 555–579.
12. Drake, P. D., & Nigrini, M. J. (2000). Computer assisted analytical procedures using Benford's law. Journal of Accounting Education, 18(2), 127–146.
13. Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. Journal of Forensic Accounting, 5(1), 17–34.
14. Gupta, P. K., & Gupta, S. (2015). Corporate frauds in India, perceptions and emerging issues. Journal of Financial Crime, 22(1), 79–103.
15. Hill, T. P. (1995). A statistical derivation of the significant-digit law. Statistical Science, 10(4), 354–363.
16. Morrow, J. (2014). Benford's law, families of distributions and a test basis (CEP Discussion Paper No. 1291). London School of Economics and Political Science.

17. Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. American Journal of Mathematics, 4(1), 39–40.
18. Nigrini, M. J. (1996). A taxpayer compliance application of Benford's law. Journal of the American Taxation Association, 18(1), 72–91.
19. Nigrini, M. J. (2012). Benford's law: Applications for forensic accounting, auditing, and fraud detection. Wiley.
20. Nigrini, M. J. (2020). Forensic analytics: Methods and techniques for forensic accounting investigations (2nd ed.). Wiley.
21. Pinkham, R. S. (1961). On the distribution of first significant digits. The Annals of Mathematical Statistics, 32(4), 1223–1230.
22. Rauch, B., Göttsche, M., Brähler, G., & Engel, S. (2011). Fact and fiction in EU-governmental economic data. German Economic Review, 12(3), 243–255.
23. Wolfe, D. T., & Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. The CPA Journal, 74(12), 38–42.