Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: https://www.science-gate.com/IJAAS.html

Uncovering hidden patterns: Predicting money laundering risk using financial data

Altan-Erdene Batbayar *, Badrakh Battsengel

Business School, National University of Mongolia, Ulaanbaatar, Mongolia

ARTICLE INFO

Article history:
Received 26 October 2025
Received in revised form
13 April 2026
Accepted 20 April 2026

Keywords:
Money laundering detection
Financial statement analysis
Probit regression model
Emerging markets
Illicit financial behavior

ABSTRACT

This study examines whether firm-level financial statement data can be used to identify companies involved in money laundering. Using Mongolia as a case study of a resource-rich emerging economy with institutional constraints, the research is based on a balanced panel dataset of 118 companies over the period 2013–2022. A probit regression model is applied to evaluate the predictive ability of key financial indicators, including firm size, leverage, inventory, profitability, and social insurance contributions, in detecting confirmed cases of money laundering. The results show that firms without inventory or social insurance payments, with smaller asset size, higher leverage, and consecutive net losses, are significantly more likely to be associated with illicit financial activities. These findings remain consistent across alternative model specifications and demonstrate the usefulness of accounting-based indicators as early warning tools for anti-money-laundering enforcement. This study provides one of the first empirical applications of financial statement analysis to predict money laundering risk in a developing market context. The results also offer practical implications for regulators by suggesting that automated red-flag systems based on publicly available financial reports can improve monitoring in settings with limited manual oversight. Overall, the study contributes to the literature on anti-money laundering and supports the development of data-driven compliance approaches.

1. Introduction

Money laundering has become one of the most pressing challenges for modern financial systems, as increasingly sophisticated techniques enable criminals to integrate illicit proceeds into the legitimate economy. Although the concept dates to early cash-intensive businesses used by organized crime groups, today money laundering is recognized as a systemic global threat that undermines financial integrity, distorts market competition, and weakens public trust in institutions. International organizations estimate that between 2 and 5 percent of global GDP is laundered annually, indicating the magnitude and pervasiveness of this phenomenon.

Money laundering is commonly defined as the process by which illegally obtained funds are concealed, converted, or transferred to disguise their illicit origin. Consistent with international conventions, Mongolia's Law on Combating Money Laundering and Terrorist Financing identifies the acquisition, possession, use, conversion, or concealment of criminal proceeds as money-laundering activities. In recent years, Mongolia has strengthened its regulatory and institutional framework in response to increasing scrutiny by the Financial Action Task Force (FATF) and the growing recognition that the country is exposed to money-laundering risks due to its resource-dependent economy, extensive cash-based transactions, and rapidly expanding corporate sector. Investigations and supervisory reports have revealed that not only financial institutions but also registered corporations may be involved in suspicious transactions.

Most existing anti-money-laundering (AML) research focuses on transaction-level monitoring in the banking sector, using suspicious transaction reports or network-based analysis. However, considerably less attention has been paid to firms' financial statements as potential sources of information for detecting money laundering risk. Corporate financial reporting may contain indirect

* Corresponding Author.
Email Address: altanerdene.b@num.edu.mn (A.-E. Batbayar)
https://doi.org/10.21833/ijaas.2026.04.017
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-3255-4580

signals of illicit behavior, such as abnormal leverage, weak payroll reporting, unusual inventory patterns, or persistent losses. Nevertheless, empirical evidence on whether such accounting indicators are systematically associated with firms suspected of money laundering remains limited, particularly in emerging economies. This constitutes an important research gap that this study seeks to address.

This study examines whether firm-level financial statement indicators can be used to discriminate between firms suspected of money laundering and normal firms in Mongolia. We compare companies that have been officially investigated for suspicious activities with other listed firms and large enterprises. Drawing on theories of financial fraud and crime opportunity, we argue that companies involved in illicit financial activities are likely to display distinct financial characteristics, which can be captured through accounting ratios and structural indicators.

The main contributions of this study are threefold. First, it extends the AML literature beyond transaction-based approaches by introducing financial-statement-based predictors of money-laundering risk. Second, it provides novel evidence from Mongolia, an emerging market where institutional reforms and AML enforcement have been rapidly evolving, but empirical studies remain scarce. Third, the paper develops an empirical model that may support regulators, auditors, and policymakers in the early identification of firms exhibiting suspicious financial patterns.

The remainder of this paper is structured as follows. Section 2 reviews related theoretical and empirical literature. Section 3 describes the data and methodology. Section 4 presents and discusses the empirical findings, and Section 5 concludes with policy implications and directions for future research.

2. Literature review

The literature on money laundering (ML) and financial crime has evolved along two closely related but often separately developed streams: (i) studies that conceptualize ML as an economic, institutional, and policy problem, and (ii) studies that develop methods to detect illicit financial activities using transaction-level, accounting, and data-driven approaches.

Early conceptual and policy-oriented studies emphasize the systemic nature of ML and the limitations of existing regulatory frameworks. Levi and Reuter (2006) provided a foundational analysis of the dynamics of ML, enforcement challenges, and policy trade-offs, arguing that although AML regimes are indispensable, they must be continuously adapted to remain efficient, fair, and proportionate in a globalized financial system. Ferwerda (2009) complemented this perspective by employing economic modeling and empirical analysis to examine whether AML policies reduce crime, concluding that their effectiveness depends critically on institutional quality, consistent enforcement, and international coordination. Schneider (2010) further broadens the scope of the problem by showing that, beyond clearly illegal activities such as drug trafficking and corruption, quasi-legal practices related to income concealment also constitute a substantial part of the shadow economy. Unger (2009) traced the transformation of money laundering from a niche concern into a central global policy issue linked to organized crime, terrorism, and financial instability, and argues for more cost-effective and intelligence-driven AML enforcement strategies.

Efforts to quantify the scale and geography of ML represent another important strand of this literature. Walker and Unger (2013) introduced the Walker Gravity Model, the first systematic quantitative framework for estimating global ML flows. Their estimates suggest that approximately 2.7% of global GDP, about USD 1.6 trillion in 2009, may be laundered annually, with major destination countries including large financial centers and secrecy jurisdictions, and major source countries being high-corruption and drug-producing economies. This work remains a seminal contribution and has inspired subsequent research that has extended the approach to new domains, such as digital finance and emerging laundering technologies.

Parallel to these macro-level perspectives, a growing body of research has focused on operational and transaction-level detection of ML. Zhang et al. (2003) proposed the Link Discovery based on Correlation Analysis (LDCA) method to analyze investigative data and uncover hidden criminal networks, demonstrating that correlation-based techniques can successfully identify members of organized crime groups. In the area of trade-based ML, Zdanowicz (2009) showed how discrepancies between declared trade prices and market benchmarks can signal suspicious transactions, providing concrete examples of systematic price manipulation in international trade. Unger (2013) applied this approach to the real estate sector, illustrating how asset markets can be used as laundering vehicles, while Bidabad (2017) suggested comparing firms' banking transactions with value-added tax (VAT) reports to detect inconsistencies indicative of illicit activity.

Another influential stream of work focuses on rule-based and expert-driven monitoring systems used in financial institutions. Visser and Yazdiha (2020) described traditional threshold-based transaction monitoring frameworks, in which transactions exceeding predefined limits trigger alerts that are reviewed by successive layers of analysts. While such systems are widely used in practice, they are criticized for generating excessive false positives and lacking adaptive, data-driven intelligence. To address these limitations, Khan et al. (2013) proposed a Rule-Based Bayesian Network approach that assigns behavioral scores to clients and flags abnormal deviations from historical patterns, demonstrating how probabilistic reasoning can enhance AML surveillance systems.

With the increasing availability of large datasets and computational power, recent literature has shifted toward data mining, machine learning, and artificial intelligence (AI) approaches. Ngai et al. (2011) provided an early and influential review of data mining applications in financial fraud detection, noting that although such methods are widely used in insurance and credit card fraud, their application to ML remains comparatively limited. More recent review studies confirm both the rapid growth and the fragmentation of this research field. Tiwari et al. (2020) analyzed 136 studies and classified the literature into typologies, regulatory frameworks, detection technologies, sectoral vulnerabilities, and economic impacts, calling for more empirically grounded and operationally relevant research. Al-Hashedi and Magalingam (2021), in a review of 75 studies, show that support vector machines and other machine learning techniques dominate fraud detection research, with a strong focus on the banking and insurance sectors. Using bibliometric and topic modeling tools, Soltani et al. (2023) documented the growing importance of AI-based methods while highlighting persistent challenges related to transparency, real-time implementation, and regulatory alignment.

Several studies further elaborate on the technical and institutional challenges of AI-based AML systems. West and Bhattacharya (2016) argued that hybrid systems combining accuracy, flexibility, and interpretability are essential for effective financial fraud detection. Goecks et al. (2022) emphasized structural obstacles, such as data imbalance, evolving laundering strategies, and regulatory constraints, and stressed the importance of explainable, legally defensible AI models. Jiao (2023) highlights the transformative role of big data analytics in AML compliance but also points to the need for stronger data governance and regulatory coherence. Hernandez Aros et al. (2024) and Zhu et al. (2024) showed that ensemble and deep learning methods can significantly improve detection performance. However, they require high-quality labeled data and careful attention to model interpretability.

In parallel with transaction-level and AI-based approaches, a long-standing literature in accounting and finance has developed models to detect financial statement manipulation and corporate distress using financial ratios. Dalnial et al. (2014), along with earlier studies by Persons (1995) and Nia (2015), emphasize the usefulness of leverage, profitability, liquidity, and efficiency ratios in identifying fraudulent financial reporting. The most prominent contribution in this area is the Beneish M-Score model (Beneish, 1999), which uses a set of financial indicators to detect earnings manipulation and has been shown to classify manipulators in 58-76% of cases correctly. Subsequent empirical tests yield mixed but generally supportive evidence: Mohamad Kamal et al. (2016) reported high classification accuracy in the Malaysian context, while Mehta and Bhavani (2017) showed that the model failed to detect the Toshiba scandal, and Hołda (2020) reported perfect classification in a small sample from the Warsaw Stock Exchange. These findings suggest that while accounting-based models are powerful screening tools, their performance may depend on institutional and reporting environments.

Closely related is the literature on corporate failure prediction, pioneered by Altman's (1968) Z-score model, which remains widely used to assess financial distress and firm stability. This line of research demonstrates that financial statement data contains systematic signals about abnormal corporate behavior and risk.

In the Mongolian context, Batbayar et al. (2015) develop a localized M-score model using data from 331 firms and show that it outperforms several international benchmarks. This study confirms the relevance of accounting-based detection tools in emerging markets with limited enforcement capacity.

Taken together, the literature reveals two important gaps. First, despite extensive research on transaction-level AML systems and on financial statement fraud detection, there is relatively little empirical work that explicitly links firm-level financial statements to confirmed cases of money laundering. Second, most existing studies focus either on advanced economies or on banking transaction data, leaving emerging markets and non-financial corporate entities underexplored. This study seeks to bridge these gaps by systematically examining whether publicly available financial statement indicators can be used to predict firm-level involvement in money laundering in an emerging economic context.

3. Data and methodology

To investigate the financial characteristics associated with corporate involvement in ML, this study uses firm-level panel data from 2013 to 2022. We obtained financial statements of companies listed on the Mongolian Stock Exchange and enterprises included in the annual Top-100 ranking published by the Mongolian National Chamber of Commerce and Industry from the E-Balance system of the Ministry of Finance of Mongolia. In total, 118 firms are included in the sample. We collected information on companies officially identified or under investigation for ML from the Independent Authority Against Corruption (IAAC) of Mongolia. Based on this information, a binary dependent variable is constructed, which takes the value of 1 if a firm has been implicated in money laundering and 0 otherwise.

The set of variables used in the empirical analysis, together with their definitions and data sources, is reported in Table 1. The explanatory variables capture firm size, leverage, profitability, and several indicators of operational irregularities that may signal abnormal or illicit behavior. The natural logarithm of total assets serves as a proxy for firm size. Financial leverage is measured in two ways: a continuous ratio of total liabilities to total assets, and a dummy variable equal to 1 if total liabilities exceed total assets. Profitability is proxied by return on assets (ROA). Additional controls include the market-to-book ratio, an adjusted asset utilization ratio, and the year-to-year change in net income. Two further dummy variables are constructed to identify firms that report neither inventory nor social insurance payments, and firms that report losses in two consecutive years. These conditions are interpreted as potential indicators of abnormal reporting or weak operational transparency. The final dataset consists of 1,038 firm-year observations, of which 49 observations are associated with firms identified as being involved in ML. Summary statistics for all variables, reported separately for ML and non-ML firms, are presented in Table 2.

The descriptive statistics indicate systematic differences between the two groups. On average, firms associated with ML tend to be smaller, less profitable, and more highly leveraged than firms not implicated in such activities. These univariate differences provide preliminary evidence that financial statement information may be informative for identifying firms exposed to ML risk.

Table 1: Variable description

Variable	Description	Source
dml	Dummy variable indicating money laundering status (1 = yes, 0 = no)	IAAC
$invsi$	Dummy variable: 1 if the firm has no inventory and no social insurance payments	
$size$	Logarithm of total assets	
$lev1$	Total liabilities divided by total assets	
$lev2$	Dummy: 1 if total liabilities exceed total assets	
roa	Return on assets, net income/total assets	E-Balance
$mvtb$	Market-to-book ratio, total equity/total liabilities	
$futl$	Funds provided by operations divided by total liabilities	
$intwo$	Dummy: 1 if the firm had losses for two consecutive years	
$chin$	Change in net income relative to the absolute sum of net income in two consecutive years	

IAAC: Independent Authority Against Corruption; Mongolia; E-Balance: E-Balance system, Ministry of Finance of Mongolia

Table 2: Summary statistics

Variable	Money laundering	Observation	Mean	SD	Min	Max
	All	1,038	0.0472	0.2122	0.0000	1.0000
dml	No	989	0.0000	0.0000	0.0000	0.0000
	Yes	49	1.0000	0.0000	1.0000	1.0000
	All	1,038	0.0414	0.1994	0.0000	1.0000
$invsi$	No	989	0.0303	0.1716	0.0000	1.0000
	Yes	49	0.2653	0.4461	0.0000	1.0000
	All	1,038	17.1604	2.5295	3.8133	23.5843
$size$	No	989	17.2731	2.4536	3.8133	23.5843
	Yes	49	14.8848	2.9619	8.1691	19.3711
	All	1,038	0.5491	0.4294	0.0000	2.0119
$lev1$	No	989	0.5333	0.4152	0.0000	2.0119
	Yes	49	0.8680	0.5706	0.0057	2.0119
	All	1,038	0.1272	0.3333	0.0000	1.0000
$lev2$	No	989	0.1082	0.3108	0.0000	1.0000
	Yes	49	0.5102	0.5051	0.0000	1.0000
	All	1,038	0.0563	0.1690	-0.3581	1.7159
roa	No	989	0.0589	0.1669	-0.3581	1.7159
	Yes	49	0.0048	0.2032	-0.3581	0.5099
	All	1,029	6.0206	15.6094	-0.9958	90.6678
$mvtb$	No	980	5.9798	15.4116	-0.9958	90.6678
	Yes	49	6.8365	19.3173	-0.8855	90.6678
	All	1,029	0.8953	2.7068	-0.7126	20.7362
$futl$	No	980	0.8952	2.7113	-0.7126	20.7362
	Yes	49	0.8977	2.6426	-0.7126	10.5485
	All	1,038	0.1291	0.3355	0.0000	1.0000
$intwo$	No	989	0.1223	0.3279	0.0000	1.0000
	Yes	49	0.2653	0.4461	0.0000	1.0000
	All	899	0.0623	0.5992	-1.0000	1.0000
$chin$	No	862	0.0659	0.5983	-1.0000	1.0000
	Yes	37	-0.0220	0.6218	-1.0000	1.0000

To formally examine the relationship between firm-level financial characteristics and the likelihood of involvement in ML, this study estimates a probit regression model.

The probit specification is appropriate given the binary nature of the dependent variable and allows for non-linear effects of explanatory variables on the probability of money laundering involvement. The empirical model is specified as follows:

$$\Pr(dml_{it} = 1) = \Phi(\beta_0 + \beta_X X_{it} + \varepsilon_{it})$$

where, $\Pr(dml_{it} = 1)$ denotes the probability that firm i in year t is identified as being involved in ML; Φ is the cumulative distribution function of the standard normal distribution; X is the vector of firm-level explanatory variables (including size, leverage, profitability, and operational indicators), and ε is the normally distributed error term. The probit model is estimated to use all available firm-year observations.

To facilitate economic interpretation, marginal effects are computed to assess how changes in each explanatory variable affect the predicted probability of involvement in ML. All financial variables are derived from audited financial statements submitted to the Ministry of Finance through the E-Balance system.

4. Empirical findings

Before estimating the probit model, pairwise correlations between the key variables were computed to assess the direction and strength of their linear relationships. The correlation matrix is reported in Table 3. As expected, the dummy variable for ML shows a positive and statistically significant correlation with several financial risk indicators, including $invsi$, $lev1$, and $lev2$. In contrast, firm size ($size$) and profitability (roa) are negatively correlated with ML, suggesting that smaller and less profitable firms are more likely to be involved in suspicious financial activities.

It is also evident that the continuous leverage ratio ($lev1$) and the high-leverage dummy variable ($lev2$) are strongly correlated, as both capture related aspects of firms' capital structure. To avoid multicollinearity and to examine the robustness of the leverage effect, these two variables are not included simultaneously in the same specification. Instead, $lev1$ is included in Models 1, 3, and 5, while $lev2$ is included in Models 2, 4, and 6. This strategy allows us to test whether the impact of leverage on money laundering risk remains stable across alternative operationalizations of financial leverage. The remaining explanatory variables exhibit moderate to low correlations, suggesting that multicollinearity is unlikely to be a major concern. Overall, the correlation results support the theoretical expectations and justify the selection of variables included in the regression models. The results of the probit regression analysis are reported in Table 4. The conditional marginal effects computed from the probit models are presented in Table 5.

Table 3: Correlation matrix

	dml	invsi	size	lev1	lev2	roa	mvtb	futl	intwo
invsi	0.2501***								
size	-0.2003***	-0.2981***							
lev1	0.1654***	0.0508	0.0395						
lev2	0.2559***	0.1093***	-0.0970***	0.7193***					
roa	-0.0678**	-0.0940***	0.0765**	-0.4338***	-0.3848***				
mvtb	0.0117	0.0905***	-0.1995***	-0.4355***	-0.1541***	0.0477			
futl	0.0002	0.0273	-0.2084***	-0.3442***	-0.1357***	0.4188***	0.6374***		
intwo	0.0904***	0.0641**	-0.0546*	0.2987***	0.3532***	-0.3570***	-0.0025	-0.1379***	
chin	-0.0292	-0.0063	0.0049	-0.0379	-0.0757**	0.3200***	0.0145	0.1133***	-0.0376

***, **, and *: Statistical significance at the 1%, 5%, and 10% levels, respectively

Table 4: Probit regression results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$invsi$	0.8911***	0.8532***	0.9223***	0.8911***	1.0426***	1.0286***
$size$	-0.0993***	-0.0910***	-0.1010***	-0.0923***	-0.1595***	-0.1572***
$lev1$	0.5669***		0.6969***		0.6836***	
$lev2$		0.9421***		1.1072***		1.1195***
roa			0.6271	0.7454*	2.2375***	2.2151***
$mvtb$					0.0127*	0.0115*
$futl$					-0.0644	-0.0775*
$intwo$					0.4563*	0.2465
$chin$					-0.2225	-0.1658
Constant	-0.5018	-0.5020	-0.5923	-0.5619	0.1604	0.3487
Observation	1,038	1,038	1,038	1,038	893	893
Pseudo R^2	0.1700	0.2077	0.1751	0.2161	0.2150	0.2507
χ^2 statistics	67.1410	81.9964	69.1358	85.3142	66.2138	77.2153
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
AIC	335.7253	320.8699	365.7608	319.5520	259.8230	248.8215
BIC	355.5055	340.6501	360.4557	344.2773	302.9743	291.9728

***, **, and *: Statistical significance at the 1%, 5%, and 10% levels, respectively

Table 5: Conditional marginal effects from the probit model

Variable	Model1	Model2	Model3	Model4	Model5	Model6
invsi	0.0632 ***	0.0551 ***	0.0641 ***	0.0556 ***	0.0557 ***	0.0488 ***
size	-0.0070 ***	-0.0059 ***	-0.0070 ***	-0.0058 ***	-0.0085 ***	-0.0075 ***

Marginal effects are evaluated at the mean of covariates; All reported estimates are statistically significant at the 1% level (***)

Across all models, the variable $invsi$, which flags companies that report neither inventory nor social insurance contributions, shows a consistently positive and statistically significant association with ML involvement. Based on the conditional marginal effects, companies without inventory and social insurance payments are approximately 4.9-6.4 percentage points more likely to be associated with ML than companies that report such items (Table 5). This estimate reflects the discrete change in predicted probability when the dummy variable $invsi$ shifts from 0 to 1, holding all other covariates constant. These findings suggest that firms exhibiting operational opacity or irregular compliance behavior face a substantially higher risk of engaging in illicit financial activities.

Firm size, measured by the logarithm of total assets ($size$), is negatively associated with ML across all models. The marginal effect estimates indicate that a one-unit increase in firm size reduces the probability of ML by approximately 0.6-0.9 percentage points (Table 5), highlighting that smaller companies are more susceptible to illicit practices, possibly due to weaker internal controls, reduced regulatory oversight, or more informal operational structures.

Leverage indicators also emerge as important predictors. Both the continuous leverage ratio ($lev1$) and the dummy variable for excessive leverage ($lev2$) exhibit positive and statistically significant coefficients, implying that highly indebted firms have a greater probability of engaging in ML. This may reflect the tendency of financially distressed firms to conceal financial weaknesses or to rely on illicit funds to meet obligations.

Profitability, measured by return on assets (roa), becomes statistically significant in the extended models (Models 5 and 6), suggesting that firms with weak or manipulated earnings performance are more likely to be associated with ML activities. Other financial indicators, such as the market-to-book ratio ($mvtb$) and the adjusted asset utilization ratio ($futl$), contribute marginally to the explanatory power of the model in later specifications, though with smaller effect sizes. The dummy variable for two consecutive years of losses ($intwo$) and the change in net income ($chin$) also appear with expected signs, although their statistical significance is model-dependent. Overall, the probit model demonstrates reasonable explanatory power, with pseudo-R-squared values ranging from 0.17 to 0.25 across specifications. The model fit improves notably with the inclusion of additional financial indicators in later models, as reflected by reductions in the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). These results support the view that ML risk can be partly predicted using firm-level financial statement information.

To assess the robustness of the findings, a complementary set of estimations using a logit specification was conducted. The conditional marginal effects from the logit models are presented in Table 6.

Overall, the signs, magnitudes, and statistical significance of the coefficients in the logit models are highly consistent with those obtained from the probit specification, confirming the stability of the main findings. Among the estimated models, the probit specification demonstrates slightly better model fit, with lower AIC and BIC values and slightly higher pseudo-R squared values, particularly in the fully specified Model 6. Therefore, the probit model is retained as the primary specification for interpretation.

Like the probit results, the variable $invsi$ remains positively and significantly associated with ML involvement across all logit specifications. Based on the conditional marginal effects from the logit model (Table 7), companies without inventory and social insurance payments are approximately 3.9-5.7 percentage points more likely to be involved in ML compared to compliant companies.

The marginal effect for company size ($size$) also remains consistent in direction and significance. A one-unit increase in firm size is associated with an approximately 0.4-0.7 percentage point decrease in the probability of ML involvement (Table 7). These findings reinforce the conclusion that smaller firms, often facing weaker governance structures and lower regulatory scrutiny, are more vulnerable to illicit financial activities.

Leverage-related variables ($lev1$ and $lev2$) again appear as strong predictors, both exhibiting positive and statistically significant effects. The profitability measure (roa) remains significant in the extended specifications, while other control variables preserve the same directional effects observed in the probit models. The consistency of results across the two estimation frameworks strengthens the credibility of empirical findings and supports the conclusion that company-level financial indicators can be effectively used to identify firms with elevated ML risk.

5. Conclusions

This study investigated whether firm-level financial statement information can be systematically used to identify companies involved in money laundering, using a unique panel dataset of 118 Mongolian enterprises over the period 2013-2022. By integrating audited financial reports with verified investigation records from the Independent

Authority Against Corruption, the study provides one of the first empirical attempts to operationalize accounting-based indicators for money laundering risk detection in an emerging market context.

Table 6: Logit regression results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$invsi$	1.7665 ***	1.6611 ***	1.8151 ***	1.7058 ***	2.0620 ***	1.9514 ***
$size$	-0.1749 ***	-0.1568 ***	-0.1760 ***	-0.1582 ***	-0.3281 ***	-0.3083 ***
$lev1$	1.1585 ***		1.3862 ***		1.3086 ***	
$lev2$		1.9444 ***		2.2547 ***		2.3020 ***
roa			1.0943	1.3902 *	4.1435 ***	4.3670 ***
$mvtb$					0.0256 *	0.0254 **
$futl$					-0.1295	-0.1524 *
$intwo$					0.8742 *	0.5317
$chin$					-0.4820	-0.4347
Constant	-1.0911	-1.1500	-1.2761	-1.2805	0.8204	0.8198
Observation	1038	1038	1038	1038	893	893
Pseudo R^2	0.1580	0.1997	0.1617	0.2071	0.2061	0.2454
χ^2 statistics	62.3825	78.8628	63.8307	81.7926	63.4956	75.5811
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
AIC	340.4837	324.0034	341.0355	323.0736	262.5413	250.4558
BIC	360.2639	343.7836	365.7608	347.7989	305.6925	293.6070

***, **, and *: Statistical significance at the 1%, 5%, and 10% levels, respectively

Table 7: Conditional marginal effects from the logit model

Variable	Model1	Model2	Model3	Model4	Model5	Model6
$invsi$	0.0561 ***	0.0462 ***	0.0568 ***	0.0459 ***	0.0471 ***	0.0385 ***
$size$	-0.0056 ***	-0.0044 ***	-0.0055 ***	-0.0043 ***	-0.0075 ***	-0.0061 ***

Marginal effects are evaluated at the mean of covariates; All reported estimates are statistically significant at the 1% level (***)

The empirical results from the probit models demonstrate that firms exhibiting operational opacity and financial distress, particularly those without inventory and social insurance payments, with smaller asset size, higher leverage, and consecutive losses, are significantly more likely to be associated with money laundering activities. These findings are consistent with theoretical perspectives from the financial fraud and crime opportunity literature, which argue that weak internal controls, limited transparency, and economic pressure create fertile conditions for illicit financial behavior. They also align with prior studies emphasizing abnormal financial structures and compliance irregularities as early warning signals of corporate misconduct.

The robustness of these results is further confirmed by alternative logit specifications, which yield consistent signs, significance levels, and marginal effect patterns. The superior fit statistics of the probit models support their use as the main empirical specification, reinforcing the stability and reliability of the proposed framework.

From a policy and regulatory perspective, this study offers critical practical implications. The results suggest that financial statement-based screening models can serve as a complementary tool to traditional transaction-monitoring systems, particularly in resource-constrained regulatory environments. Supervisory authorities such as the Financial Regulatory Commission and the General Authority for Taxation could incorporate these indicators into risk-based supervision, enabling earlier detection of high-risk firms and more targeted allocation of investigative resources.

Several limitations should be acknowledged. First, although the dataset is unique, the number of confirmed cases remains relatively small, and some firms are still under investigation, which may affect classification accuracy. Second, the analysis relies primarily on accounting data and does not incorporate transaction-level information, ownership networks, or audit quality variables. Future research could address these limitations by integrating micro-level transaction data, governance indicators, and network-based measures to develop more comprehensive hybrid detection models.

Despite these constraints, this study provides a foundational step toward a more data-driven, scalable, and cost-effective approach to anti-money laundering enforcement. More broadly, it demonstrates that corporate financial statements, traditionally used for performance evaluation, can also serve as valuable instruments for financial crime surveillance, not only in Mongolia but also in other emerging economies facing similar institutional and enforcement challenges.

List of abbreviations

AI	Artificial intelligence
AIC	Akaike information criterion
AML	Anti-money laundering
BIC	Bayesian information criterion
chin	Change in net income relative to the absolute sum

	of net income in two consecutive years
dml	Dummy variable indicating money laundering status (1 = yes, 0 = no)
FATF	Financial Action Task Force
futl	Funds provided by operations divided by total liabilities
GDP	Gross domestic product
IAAC	Independent authority against corruption
intwo	Dummy variable equal to 1 if the firm reports losses for two consecutive years
invsi	Dummy variable equal to 1 if the firm has no inventory and no social insurance payments
LDCA	Link discovery based on correlation analysis
lev1	Total liabilities divided by total assets
lev2	Dummy variable equal to 1 if total liabilities exceed total assets
ML	Money laundering
mvtb	Market-to-book ratio (total equity divided by total liabilities)
roa	Return on assets (net income divided by total assets)
SD	Standard deviation
size	Logarithm of total assets
VAT	Value-added tax

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Al-Hashedi KG and Magalingam P (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 40: 100402. https://doi.org/10.1016/j.cosrev.2021.100402

Altman EI (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, 23(4): 589-609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Batbayar AE, Boldbaatar M, and Enkh-Amgalan T (2015). Corporate bankruptcy prediction model in Mongolia. In the 13th Conference of International Federation of East Asian Management Associations in Ulaanbaatar, IFEAMA SPSCP, 5: 136-145.

Beneish MD (1999). The detection of earnings manipulation. Financial Analysts Journal, 55(5): 24-36. https://doi.org/10.2469/faj.v55.n5.2296

Bidabad B (2017). Money laundering detection system (MLD)(a complementary system of Rastin banking). Journal of Money Laundering Control, 20(4): 354-366. https://doi.org/10.1108/JMLC-04-2016-0016

Dalnial H, Kamaluddin A, Sanusi ZM, and Khairuddin KS (2014). Detecting fraudulent financial reporting through financial statement analysis. Journal of Advanced Management Science, 2(1): 17-22. https://doi.org/10.12720/joams.2.1.17-22

Ferwerda J (2009). The economics of crime and money laundering: does anti-money laundering policy reduce crime? Review of Law and Economics, 5(2): 903-929. https://doi.org/10.2202/1555-5879.1421

Goecks LS, Korzenowski AL, Gonçalves Terra Neto P, de Souza DL, and Mareth T (2022). Anti-money laundering and financial fraud detection: A systematic literature review. Intelligent Systems in Accounting, Finance and Management, 29(2): 71-85. https://doi.org/10.1002/isaf.1509

Hernandez Aros L, Bustamante Molano LX, Gutierrez-Portela F, Moreno Hernandez JJ, and Rodríguez Barrero MS (2024). Financial fraud detection through the application of machine learning techniques: A literature review. Humanities and Social Sciences Communications, 11(1): 1-22. https://doi.org/10.1057/s41599-024-03606-0

Hołda A (2020). Using the Beneish M-score model: Evidence from non-financial companies listed on the Warsaw Stock Exchange. Investment Management & Financial Innovations, 17(4): 389-401. https://doi.org/10.21511/imfi.17(4).2020.33

Jiao M (2023). Big data analytics for anti-money laundering compliance in the banking industry. Highlights in Science, Engineering and Technology, 49: 302-309. https://doi.org/10.54097/hset.v49i.8522

Khan NS, Larik AS, Rajput Q, and Haider S (2013). A Bayesian approach for suspicious financial activity reporting. International Journal of Computers and Applications, 35(4): 181-187. https://doi.org/10.2316/Journal.202.2013.4.202-3864

Levi M and Reuter P (2006). Money laundering. Crime and Justice, 34(1): 289-375. https://doi.org/10.1086/501508

Mehta A and Bhavani G (2017). Application of forensic tools to detect fraud: The case of Toshiba. Journal of Forensic and Investigative Accounting, 9(1): 692-710.

Mohamad Kamal ME, Md Salleh MF, and Ahmad A (2016). Detecting financial statement fraud by Malaysian public listed companies: The reliability of the Beneish M-Score model. Jurnal Pengurusan, 46: 23-32. https://doi.org/10.17576/pengurusan-2016-46-03

Ngai EW, Hu Y, Wong YH, Chen Y, and Sun X (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision Support Systems, 50(3): 559-569. https://doi.org/10.1016/j.dss.2010.08.006

Nia SH (2015). Financial ratios between fraudulent and non-fraudulent firms: Evidence from Tehran Stock Exchange. Journal of Accounting and Taxation, 7(3): 38-44. https://doi.org/10.5897/JAT2014.0166

Persons OS (1995). Using financial statement data to identify factors associated with fraudulent financial reporting. Journal of Applied Business Research, 11(3): 38-46. https://doi.org/10.19030/jabr.v11i3.5858

Schneider F (2010). Turnover of organized crime and money laundering: Some preliminary empirical findings. Public Choice, 144: 473-486. https://doi.org/10.1007/s11127-010-9676-8

Soltani M, Kythreotis A, and Roshanpoor A (2023). Two decades of financial statement fraud detection literature review; combination of bibliometric analysis and topic modeling approach. Journal of Financial Crime, 30(5): 1367-1388. https://doi.org/10.1108/JFC-09-2022-0227

Tiwari M, Gepp A, and Kumar K (2020). A review of money laundering literature: The state of research in key areas. Pacific Accounting Review, 32(2): 271-303. https://doi.org/10.1108/PAR-06-2019-0065

Unger B (2009). Money laundering-a newly emerging topic on the international agenda. Review of Law & Economics, 5(2): 807-819. https://doi.org/10.2202/1555-5879.1417

Unger B (2013). Can money laundering decrease? Public Finance Review, 41(5): 658-676. https://doi.org/10.1177/1091142113483353

Visser F and Yazdiha A (2020). Detection of money laundering transaction network structures and typologies using machine learning techniques. M.Sc. Thesis, Erasmus School of Economics, Rotterdam, Netherlands.

Walker J and Unger B (2013). Measuring global money laundering: The 'Walker Gravity Model'. In: Unger B and Van der Linde D (Eds.), Research handbook on money laundering:

159-171. Edward Elgar Publishing, Cheltenham, UK. https://doi.org/10.4337/9780857934000.00023

West J and Bhattacharya M (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57: 47-66. https://doi.org/10.1016/j.cose.2015.09.005

Zdanowicz JS (2009). Trade-based money laundering and terrorist financing. Review of Law & Economics, 5(2): 855-878. https://doi.org/10.2202/1555-5879.1419

Zhang Z, Salerno JJ, and Yu PS (2003). Applying data mining in investigating money laundering crimes. In the Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Washington D.C., USA: 747-752. https://doi.org/10.1145/956750.956851 **PMid:12897971**

Zhu S, Wu H, Ngai EW, Ren J, He D, Ma T, and Li Y (2024). A financial fraud prediction framework based on stacking ensemble learning. Systems, 12(12): 588. https://doi.org/10.3390/systems12120588