



A Comprehensive Analysis Of Embezzlement Practices For Small And Medium-Sized Enterprises In Vietnam

Manh-Hung Ha^{1,*}, Duy-Thuc Nguyen¹, Zhang Yuemei²

Thi-Minh-Ngoc Luu², Thanh-Thao Bui¹, Thuy-Huyen Le¹

¹ Faculty of Applied Sciences, International School, Vietnam National University, Hanoi, Vietnam

² Faculty of Economics and Management, International School, Vietnam National University, Hanoi, Vietnam

*Corresponding author: hungmh@vnu.edu.vn

Abstract

Embezzlement has become one of the most pressing issues in Vietnam recently, leaving significant repercussions across various fields, such as economics, politics, and education. To find out the causes and solutions to minimize the frequency and consequences of this issue, we delve into the effects of selected factors on embezzlement behaviors within Vietnamese Small and Medium-sized Enterprises (SMEs) by utilizing an extensive survey on diverse demographics and conducting an in-depth analysis of critical factors influencing embezzlement practices based on weighted linear regression. Particularly, this study embezzlement habits, giving significant insights through a variety of methods based on different factors such as Cronbach's Alpha emphasizes the high dependability of measurable parameters to aspects that contribute significantly to our knowledge of embezzlement dynamics. The Spearman's rank correlation coefficients reveal significant connections between embezzlement practices and seven independent factors. Notably, embezzlement methods and reasons have substantial and modest relationships, respectively. The multiple linear regression model, with a high R-squared value, yields useful prediction information. Embezzlement practices vary significantly by gender and age group. Middle-aged ladies with significant expertise make up the bulk of responders, highlighting the widespread prevalence of embezzlement across demographics. Other personal characteristics had no significant effect on embezzlement practices. The study concludes with recommendations for businesses to implement tailored preventive measures, focusing on strict management and employee education, to reduce embezzlement. The findings demonstrate significant correlations between embezzlement activities and seven principal variables, notably the methods and motivations behind these acts.

Research purpose:

The study aims to explore the causes and solutions for minimizing embezzlement in Vietnamese Small and Medium-sized Enterprises (SMEs). Specifically, it seeks to understand the factors that influence embezzlement behaviors, providing insights into their impact on businesses and society.

Research motivation:

The rising occurrence of embezzlement in Vietnam across various sectors, including economics, politics, and education, has created a need for research on the causes of such activities. Understanding these factors can help develop effective preventive measures for organizations to mitigate the damaging consequences of embezzlement.

Research design, approach, and method:

The study employed a survey-based approach, targeting diverse demographic groups within Vietnamese SMEs. Using statistical methods, such as Cronbach's Alpha for reliability testing and Spearman's rank correlation for determining relationships between variables, a multiple linear regression model was developed to predict embezzlement behaviors. The high R-squared value of the model suggests robust predictive power regarding factors influencing embezzlement.

Main findings:

Embezzlement practices are significantly associated with seven independent factors. The study reveals a strong link between the methods and motivations behind embezzlement. The research emphasizes the role of preventive measures, such as strict management practices and employee education, to address this issue effectively.

© The Author(s) 2025

D. Nguyen Van et al. (eds.), *Proceedings of the International Conference on Emerging Challenges: Sustainable Strategies in the Data-driven Economy (ICECH 2024)*, Advances in Economics, Business and Management Research 320,

https://doi.org/10.2991/978-94-6463-694-9_38

Practical/managerial implications:

The findings suggest that businesses should focus on tailored preventive strategies, including improved management oversight and targeted employee training, to reduce the incidence of embezzlement. Implementing these measures can lead to better protection of company assets and reduce financial losses caused by unethical behaviors.

Keywords: Embezzlement Behaviors; Linear regression; Enterprises; Descriptive Statistics; Generalized Method of Moments.

1. INTRODUCTION

Embezzlement, the act of fraudulently converting public property into private property, appears in many forms, with different levels, nature, and severity depending on each act by Forbes Advisor (2024) (*What Is Embezzlement? Legal Definition, Types & Examples – Forbes Advisor*, n.d.). It can be a daily act in stores, the level can be just taking a few items of small value or it can take place on a large scale, by people holding positions and powers in companies and organizations have the authority to commit acts of embezzlement with assets of great value. This is an act prohibited by law because it harms public property and public interests.

The embezzlement situation in Vietnam is a pressing issue, causing significant losses to businesses, organizations, and the national economy. Recently, the Minister of Justice signed a decision approving and announcing the top 10 notable events of 2023 in the judicial sector, including "the highest-ever results in civil enforcement regarding matters, money, and asset recovery in corruption and economic cases." Specifically, in 2023, 574,819 cases were resolved, settling over 89 trillion VND, an increase of over 14 trillion VND compared to 2022 (with asset recovery in economic and corruption criminal cases increasing by over 4,415 trillion VND). This does not even account for the cases yet to be enforced, which have already significantly exceeded expectations.

ASIA PACIFIC

VIETNAM

Share how your country is doing



Score

42/100 [What does the CPI score mean?](#)

Rank

77/180

Score change

↑ +3 since 2021

Score changes 2012 - 2022

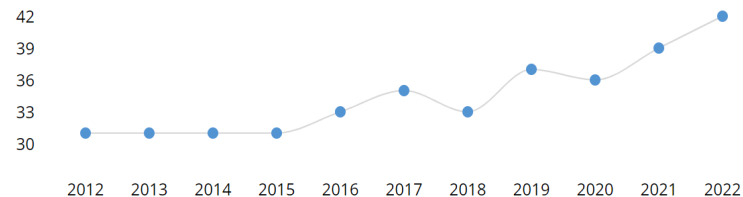


Fig. 1. Vietnam's CPI index

Figure 1 reveal the report released by Transparency International (TI) reveals the 2022 Corruption Perceptions Index (CPI) - a metric for assessing countries based on the perceived level of government corruption, showing that Vietnam is one of six countries in the Asia-Pacific region that have made significant progress by Alfaro, Emigdio (Alfaro, 2022) and Transparency (2022 *Corruption Perceptions Index*, 2023). A country's score can range from 0 to 100, with 0 being the highest level of corruption and 100 being considered the 'cleanest'. With 42 points, Viet Nam has increased by 9 points since 2018. However, 2023 was a tumultuous year in Vietnam with quite a few scandalous incidents, such as: A recent prominent case, the 'rescue flight' project, saw the Ministry of Public Security initiate proceedings against a total of 38 suspects, primarily officials from 9 ministries, local governments, and related businesses (*The Defendants Received the Most Bribes in the "rescue Flight" Case - Vietnam.Vn*, 2023). During the 2020-2021 period, when the COVID-19 pandemic was rampant both domestically and internationally, the demand for return tickets to the country was very high. These tickets were extremely valuable, leading some officials to abuse their positions for personal gain and accept bribes. The number of bribery incidents exceeded 500, totaling nearly 165 billion VND. Particularly notable is the recent case on November 18, 2023, where Ms. Truong My Lan - the chairwoman of Van Thinh Phat Group - was recommended for prosecution by the Investigative Agency of the Ministry of

Public Security (C03) on three charges: Bribery, violation of banking regulations, and embezzlement, with the embezzled amount exceeding 304 trillion VND (*Ms. Truong My Lan Manipulated and Appropriated Money from SCB Bank - Vietnam. Vn*, 2023). Therefore, the most recent report from Transparency International (TI) released in early February 2024 disclosed that Vietnam's Corruption Perceptions Index (CPI) for 2023 had decreased by one point compared to 2022, indicating that the state of corruption has become difficult to control by Transparency (*2023 Corruption Perceptions Index - Explore Vietnam's Results*, 2024).

We carried out this research to find out the characteristics and capabilities of a person who can commit embezzlement by developing a regression model. Thereby helping businesses establish strict inspection and monitoring processes to control fraud risks, which are necessary to protect the sustainable development and success of business enterprises in Vietnam. we also partly provide ways to help businesses have plans and strategies to prevent this behavior.

2. LITERATURE REVIEW

According to the 2018 survey report "Global Economic Crime and Fraud Survey: Vietnam Perspective" (PricewaterhouseCoopers, 2018), 52% of businesses in Vietnam faced fraud in the past two years. Among these, asset embezzlement is the most common type of fraud, accounting for 40%. This demonstrates the significant threat that fraud poses to Vietnamese businesses, particularly in the SME sector.

The other study "Embezzlement and Operational Efficiency of Businesses: New Econometric Evidence from the Analysis of Small and Medium Enterprises in Vietnam" (Van et al., 2017). measures corruption by whether a business has to pay any informal fees (measured by a dummy variable) and the degree of corruption by the value of informal payments relative to revenue (measured as a percentage); and business efficiency by the value added. This research quantifies the impact of corruption, the level of corruption, and the type of corruption on the financial performance of companies (measured by ROA: return on assets). The authors argue that using a dummy variable to measure the level of corruption does not fully reflect the extent of corruption, so they use both dummy and continuous variables. Initially, the authors used OLS estimation, but the results were biased compared to those from other studies due to the inability to control for unobserved factors, so they proceeded to use the GMM method to address endogeneity issues for all independent variables, which made the impact of corruption on financial performance insignificant after controlling for unobservable and endogenous factors. This finding supports the institutional theory perspective and reflects the reality that corruption is prevalent in Vietnamese enterprises and therefore, participating in corrupt activities does not provide financial efficiency for the businesses. Furthermore, when the authors replaced the corruption dummy variable with a continuous variable measuring the level of corruption, they found that the degree of corruption negatively affects the financial operations of companies regardless of the model used. After analyzing the regression of the impact of corruption on financial efficiency, the study found that the level of bribery (measured as a percentage of the value of the bribe relative to revenue) at the company and provincial levels negatively affects the productivity of small and medium enterprises (SMEs) in Vietnam. The results show that the impact of different types of corruption varies, some forms of corruption (e.g., paying informal fees for certain sectors or procedures) do not affect the financial operations of the company, while others (e.g., paying fees for licenses and government contracts) have a significant negative impact. In addition, the results of the model have shown that the relationship between corruption and company financial performance should be analyzed in a dynamic model for a more in-depth look at this issue. After reading and analyzing this study, we find it to be quite comprehensive, meticulous, and realistic. The study has provided the most suitable and optimal method to solve the problem of endogeneity that remains in other studies. Moreover, the authors provide specific details on the selection process of the dummy variable and explain why they chose that approach, compared with other measurement methods to help the reader understand the advantages and disadvantages of their method. However, there are still limitations such as the difficulty in accurately measuring the level of corruption and the type of corruption, the lack of uniformity in measurement, which reduces the reliability of the results. Besides, the possibility of sample selection bias is also an issue that can affect the expressiveness of the results. If the businesses that self-select to participate in the study are different, the results may not be applicable to all businesses in Vietnam.

Research the effect of internal control on asset misappropriation: the case of Vietnam. Business and Economic Horizons (Ha Le & Tran, 2018) was conducted to examine the effects of internal control systems on asset misappropriation behaviors in Vietnamese enterprises through 279 questionnaires, assumptions collected from internal employees in Vietnamese companies. The study used SPSS software to process the questions, test reliability with variables using Cronbach's Alpha scale, and filter out invalid variables with SPSS 20. Then, to test the research hypotheses, the article conducted two analyses: Pearson correlation analysis to test the linear correlation between variables; and multiple regression analysis to measure the impact of independent variables on the dependent variable. These are common and powerful statistical methods. Through this research method, the author has shown that the control environment has the greatest impact on fraud control in businesses, especially when it includes the integrity and ethical values of the organization. Control activities, such as task separation and delegation,

also play an important role in dealing with the possibility of fraud. On the other hand, information and communication play an equally important role in controlling fraud, helping to reduce the chance of committing fraud by ensuring a smooth flow of information between different parts of the enterprise. Conversely, monitoring and risk assessment are seen as having less impact on fraud control, possibly due to the lack of attention of Vietnamese enterprises to risk management and control supervision. This may be particularly applicable to small and medium-sized enterprises, which may lack the resources and experience to face fraud risks. This study contributes to the theory on the relationship between internal control and the detection and prevention of asset misappropriation in enterprises. It shows that Vietnamese enterprises need to make further efforts to build a strong control environment in their organizations. After careful reading and analysis, we have identified some limitations in this study. Firstly, the author admits that 'Data compilation is not absolutely comprehensive,' a lack of data integrity can affect the accuracy of results and the ability to generalize. Furthermore, the study directly inputs the sample and data into the analysis model but does not present the specific data collection process and its rationale, this lack of information can reduce the transparency and meticulousness of the article. Besides, the study has not clearly stated the limitations of the regression model, such as the causality relationship. In summary, these limitations help identify the strengths and weaknesses of the research and form the basis for our study.

3. METHODOLOGY

3.1 Linear Regression and Weight Estimation Chain Simulations

The field of statistics has experienced notable progress, particularly in the application of Linear Regression model. This methodology has become increasingly popular due to its methodological robustness and empirical effectiveness. Additionally, advancements in software have revolutionized econometrics, providing researchers with superior analytical capabilities compared to traditional methods (Yu et al., 2014).

Linear Regression (Jagdeesh, 2023) serves the purpose of determining the linear relationship between a dependent variable and one or more independent variables. This approach has been widely applied in diverse fields. M. Begum et al anticipated long-term software fault by considering the correlation between relative error and testing days (Begum et al., 2023). Meanwhile, K. Borna et al emulated linear prediction methods, neural networks and fuzzy logic to forecast the price of electricity (Borna & Palizdar, 2016). The evaluation of correlation between attributes is also provided in the research about predicting student performance (Kalaivani K et al., 2021).

Linear regression offers several advantages, including the ability to approximate complex models and handle data sets with insufficient sample sizes. Moreover, it applies to all parametric models, providing a versatile tool for statistical analysis. However, challenges exist, where researchers often struggle with parameter pre-selection. A cautious approach involves opting for presets that are either data-dependent or devoid of informative content, which differs from the typical approach in linear regression (Dao-de, 2000). Despite their potential benefits, presets lacking informative cues may prove inadequate in resolving common regression issues, such as the problem of separation.

To address these challenges, researchers have employed weight estimation to discern weaker information priorities for the regression coefficients in linear regression by assigning weights to individual cases in the dataset (Meermeyer, 2015). This allows for a more nuanced understanding of the relationships between variables and aids in improving the accuracy of the models. Within our framework of weight estimation modeling, embezzlement behavior can be comprehensively considered at various levels. This highlights the practical applications of these statistical methods in real-world scenarios, where complex factors need to be accounted for in data analysis.

3.2 Model specification

Based on earlier studies about embezzlement correlations by Van, Huong Vu, et al (2017), the authors specified a linear regression model to analyze Embezzlement Behavior (EB) as follows:

$$EB_i = \alpha + \beta_1 DW_i + \beta_2 CS_i + \beta_3 EM_i + \beta_4 AO_i + \beta_5 RM_i + \beta_6 EE_i + \varepsilon_i \quad (1)$$

where EB_i presenting the index of Embezzlement Behavior of the person i while DW_i , CS_i , EM_i , AO_i , RM_i , EE_i respectively denoting for the index of Reasons of Embezzlement, Assessment from Others, Difficulties at Work, Confidence Score, Rigorous Management and Embezzlement Methods of the person i .

Weighted linear regression (Vaghefi, 2021) is a generalization of linear regression where the covariance matrix of errors is incorporated in the model. In weighted linear regression, each observation i has an associated weight w_i and the objective function is to minimize the weighted sum of squared residuals (WSSR):

$$WSSR = (y - X\beta)^T W (y - X\beta) \quad (2)$$

The maximum likelihood estimation of weighted linear regression can be derived as:

$$\beta = (X^T W X)^{-1} X^T W y \quad (3)$$

where X is the design matrix that contain the independent variables and a column of ones for the intercept, y is the dependent variable vector and W is the diagonal matrix with weight w_i .

3.3 Hypothesis Development

Dependent variables: In this study, the dependent variable is Embezzlement (EB), which is identified as the act of fraudulently converting public property into private property. Embezzlement can take various forms with different levels of severity, and it is considered illegal as it harms public property and interests.

Table 1. Abbreviation of variables

Variables	Abbreviation
Embezzlement Behaviors	EB
Reasons of Embezzlement	RE
Assessment from Others	AO
Difficulties at Work	DW
Confidence Score	CS
Rigorous Management	RM
Embezzlement Methods	EM

Independent variables: The research focuses on exploring the characteristics and capabilities of individuals engaging in embezzlement by considering independent variables which are presented in Table 1. From the above discussion, the following hypotheses can be formulated:

Hypothesis H₁: There is no significant relationship between Reason of Embezzlement (RE) on Embezzlement Behavior (EB).

Hypothesis H₂: There is no significant relationship between Difficulties at Work (DW) and Embezzlement Behavior (EB).

Hypothesis H₃: There is no significant relationship between Confidence Score (CS) and Embezzlement Behavior (EB).

Hypothesis H₄: There is no significant relationship between Embezzlement Methods (EM) and Embezzlement Behavior (EB).

Hypothesis H₅: There is no significant relationship between Assessment from Others (AO) on Embezzlement Behavior (EB).

Hypothesis H₆: There is no significant relationship between Rigorous Management (RM) on Embezzlement Behavior (EB).

Hypothesis H₇: There is no significant relationship between Education about Embezzlement (EE) on Embezzlement Behavior (EB).

3.4 Data

The dataset in this study are gathered from surveys conducted anonymously to people who left their previous jobs within small and medium-sized enterprises (SMEs).

Data manipulation techniques involve addressing issues such as missing data, validating data reliability, and standardizing data to enhance its effectiveness in statistical methodologies. Through the amalgamation of data from reliable sources, our objective is to construct a robust statistical model capable of elucidating and predicting the extent of embezzlement based on the identified research factors.

This tailored approach ensures that the statistical model is finely tuned to the dynamics of embezzlement within the context of small and medium-sized enterprises, providing insights that are directly applicable and pertinent to this specific business segment.

Data preprocessing: In the comprehensive analysis of survey responses from Vietnamese small and medium-sized enterprises (SMEs), a rigorous data cleaning and transformation process was meticulously implemented to ensure the dataset's integrity and consistency. The initial steps involved loading the data into a pandas DataFrame, delete Timestamp information,

remove the outliers labeled 'PhD' and address missing values by uniformly replacing them with the term 'Marketing.' Duplicate rows were subsequently removed to enhance data accuracy. Recognizing the challenges posed by lengthy and complex questions affecting our analysis and model execution, we've replaced them with concise symbols or abbreviations, as detailed in the table below. This strategic simplification aims to streamline our workflow, reducing cognitive load and enhancing efficiency during both analysis and model runs. In conclusion, these stringent data cleaning and transformation procedures have fortified the survey dataset, establishing a robust foundation for deriving meaningful insights into the practices and perspectives of Vietnamese SMEs. The standardized dataset provides a reliable basis for nuanced and insightful analyses, contributing to the final report's depth and credibility.

3.5 Tools use for this project

We can use a variety of tools to conduct our research on embezzlement in SMEs. Some of the most common tools we use in this project include:

- + SPSS: A statistical software package that can be used for data analysis.
- + Tableau: A data visualization tool that can be used to create charts and graphs of our data.
- + Python: A programming language that can be used for data analysis and automation.

SPSS was used to: SPSS was crucial in the analysis of embezzlement data, translating categorical data to numerical format, improving data comprehension. This first phase was designed to allow for a more in-depth examination of variable connections. Following that, SPSS was used for descriptive statistical analysis, computing means and standard deviations to provide a full overview of the embezzlement data's central tendency and dispersion. SPSS was used to conduct correlation and regression analysis to further investigate the linkages within the data. Correlation analysis revealed variable associations, but regression analysis revealed predictive links. The correlation and regression analysis produced a model that explained the data variability. Notably, RE, DW, CS, and EM had positive associations with embezzlement (EB), whereas AO, RM, and EE had negative correlations with EB.

Tableau was used to: Create visualizations of the embezzlement data. This was done to communicate the results of the research to a wider audience. For example, Tableau was used to create bar charts, line charts, and scatter plots to show the relationship between different factors and embezzlement. Tableau was used to create visualizations of the embezzlement data. This was done to communicate the results of the research to a wider audience. For example, Tableau was used to create bar charts, line charts, and scatter plots to show the relationship between different factors and embezzlement.

Python was used to: Python was used to be instrumental in streamlining data preprocessing tasks, ensuring the quality and readiness of data for subsequent in-depth analysis in SPSS. Key preprocessing steps executed in Python are included. Python scripts seamlessly extracted relevant data from spreadsheets. Python's robust libraries effectively addressed data quality issues, including identifying and handling missing values, correcting inconsistencies and errors, and formatting data into consistent structures. Python facilitates essential data transformations to optimize it for analysis, such as creating new variables or features, normalizing or scaling data, and handling categorical variables appropriately. SPSS then took center stage in conducting the in-depth statistical analysis, leveraging its specialized capabilities to uncover meaningful insights from the meticulously prepared data.

4. RESULTS AND DISCUSSION

Out of the 276 individuals surveyed, the bulk of respondents comprised females, making up 77.5% of the total, while males constituted a smaller proportion at approximately 22.5%. The largest demographic slice falls within the age group of over 41 years, accounting for roughly 36.7%, closely followed by those under 30, comprising 33.5%. The age ranges of 31-35 and 36-40 each represent just under 20% of the respondents. Over half of the surveyed individuals possess over 6 years of working experience, with the remaining portion having less than 6 years. The majority of respondents, totaling 50.5%, are employed in small companies employing 5-30 people. The primary fields of employment include Service (35.6%), Business (16.4%), Production (14.9%), and Education (13.8%), with other sectors each registering below 10%. Employment distribution across departments appears relatively balanced, without significant concentration in any specific area. The most prevalent job title among respondents is Employee/Worker, encompassing 84.7% of the surveyed population. Nearly 70% of respondents hold a Bachelor's degree, indicating a relatively high level of education among the sample.

Table 2. Descriptive statistics and reliability of the dataset

<i>Variables</i>	<i>N</i>	<i>Items</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>Cronbach's alpha</i>
Embezzlement's behaviors (EB)	276	9	2.12	1.12	

Reasons (RE)	276	10	2.85	1.1	0.931
Assessment from Others (AO)	276	8	3.15	1.15	0.943
Difficulties of work (DW)	276	9	2.89	1.09	0.94
Confidence (CS)	276	6	3.17	1.12	0.949
Rigorous Management (RM)	276	7	3.3	1.17	0.948
Embezzlement methods (EM)	276	4	2.29	1.33	0.935
Education about Embezzlement (EE)	276	8	3.27	1.23	0.97

4.1 Reliability and Validity

The dataset for researching embezzlement behaviors comprises 8 parent factors, each comprising 4-10 observed child variables. Given the considerable number of child variables aimed at delineating the characteristics of the parent factors, assessing their reliability, validity, and coherence becomes crucial, as not all child variables may effectively capture the essence of the parent factor.

To address this challenge, Cronbach's Alpha scale reliability testing tool (Vaske et al., 2016), is employed. This statistical test gauges the degree of correlation among observed variables within the same factor. A higher correlation level signifies greater internal consistency, leading to a higher Cronbach's Alpha coefficient. This coefficient serves to identify which observed variables within a factor contribute meaningfully to the measurement of the factor concept, discerning them from variables that may not align closely with the intended construct. According to research, a good scale will have a Cronbach's Alpha value of 0.6 or higher. Table 2 showed extremely high reliability with results ranging from 0.931 to 0.97. This has shown that the observed variables are closely correlated and meaningful in explaining the parent factor.

For the variables showing the level of embezzlement behavior, the average is about 2.12, meaning the values will mainly be 1(never) and 2(rarely); the Standard Deviation value is 1.12, and the concentration level is relatively high. These are the two lowest Mean and SD values when measuring factors. In general, other variables to evaluate the causes, methods, and difficulties in the working process... the average value is much higher, and the SD is also relatively low, reflecting relatively reliable results for the article survey. For the Reasons factor, it reached a relatively high average of 2.85 with answers concentrated on 2(Disagree), 3(Neutral), and 4(Agree). For the Rigorous Management factor, it reached the highest average with a value of 3.3 with answers concentrated on 3(Neutral), 4(Agree), and 5(Completely Agree).

4.2 Results of Weighted linear regression

Table 1. The correlations between Embezzlement behaviors with independent factors

x	EB	RE	AO	DW	CS	RM	EM	EE
EB	1							
RE	.570**	1						
AO	.338**	.729**	1					
DW	.481**	.732**	.746**	1				
CS	.385**	.653**	.766**	.701**	1			
RM	.266**	.579**	.745**	.597**	.753**	1		
EM	.667**	.495**	.280**	.500**	.343**	.235**	1	
EE	.252**	.563**	.695**	.558**	.710**	.765**	.229**	1

Applied for the non-parametric measure of rank correlation, the spearman's rho (Gupta et al, 2022) is used for testing the hypothesis. The result from Table 3 demonstrates significant relationships between EB and the seven independent variables. In particular, EM displays a strong relationship while RE and DW show moderate correlations. The AO, CS, RM, and EE only indicate a weak relationship with misappropriation actions.

The model summary of multiple linear regression model in Table 4 shows an R-squared of 0.591 which means the model can explain 59.1% of the variability in the data. The p-value smaller than 0.05 demonstrates that the model is statistically significant.

Table 2. Summary of multiple linear regression model

<i>Model</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>	<i>Std Error of the Estimate</i>	<i>Change Statistic</i>				
				<i>R-Squared Change</i>	<i>F Change</i>	<i>df1</i>	<i>df2</i>	<i>Sig. F Change</i>
Linear	.591	.580	.72.628	.591	55.105	7	267	.000

** Correlation is significant at the 0.01 level (2-tailed).

Table 3. Summary of weighted linear regression

<i>Multiple R</i>	<i>R-Squared</i>	<i>Adjusted R-Squared</i>	<i>Std. Error of the Estimate</i>	<i>Log-likelihood Function Value</i>
0.794	0.63	0.621	0.184	-249.906

Table 5 show that the R-squared can be improved to 0.63 by applying weighted linear regression. That increases the explanatory power of the model which regards the variance in the variable embezzlement behavior. It also leads to better predictive accuracy for reliable decision-making.

According to the Table 6, the weight of each variable has certainly changed. The model's coefficient also ranges from -0.5 to 0.5, with the RE variable having the largest value of 0.418 and the AO variable having the smallest value of -0.51. However, it can be seen that in the resulting table, there are only 2 variables: RE and EM have statistical significance (Sig < 0.05), and the rest have Sig > 0.05 so those variables have slight meaning in the model. Along with that, the coefficients of the RE and EM are also the highest, which is strongly influenced by the dependency. RE (Reasons of Embezzlement) has the most influence, when this variable increases by one unit, the Embezzlement Behaviors increases by 0.418 units. It can be understood that the more it is agreed with the cause of the embezzlement, the more likely they perform the acts of the embezzlement and the more often. EM (Embezzlement Methods) has a little less influence, when an increase in this unit in this variable, dependent variables increased by 0.339, that is, the more agreed with the methods of embezzlement, the more ability to cause embezzlement behaviors. In particular, the R-squared value of 0.63 higher than the initial result of 0.59 demonstrates that the adjusted coefficients build the fitter model with the dataset.

Table 4. Coefficients of variables in the weighted linear regression model

	<i>Unstandardized Coefficients</i>		<i>Standardized Coefficients</i>		<i>t</i>	<i>Sig</i>
	<i>B</i>	<i>Std.Error</i>	<i>Beta</i>	<i>Std.Error</i>		
(Constant)	.336	.062			5.402	.000
RE	.418	.053	.521	.066	7.947	.000
AO	-.051	.051	-.087	.086	-1.008	.314
DW	.021	.057	.030	.083	.364	.716
CS	.028	.051	.048	.089	.539	.590
RM	.004	.049	.007	.093	.078	.938
EM	.339	.037	.444	.048	9.289	.000
EE	-0.64	.042	-.127	.083	-1.519	.130

Recommendation: The project makes a number of concrete recommendations targeted at reducing embezzlement and creating a safer corporate environment. To address the projected discrepancies in embezzlement behaviors among genders and age groups, businesses should develop and execute preventive measures targeted to certain demographics. Customized awareness campaigns can improve comprehension and encourage ethical conduct. Organizations can focus on improving rigorous management processes to strengthen their defenses against embezzlement. This entails establishing clear roles and responsibilities, putting in place comprehensive control systems, and ensuring that policies and procedures are effectively communicated throughout the business. Investing in education on embezzlement is critical for boosting staff awareness. Prioritize training programs based on criteria with high association, such as Embezzlement Methods (EM) and Reasons (RE). Tailored instructional programs can help greatly reduce embezzlement incidences. Implementing a system of continuous monitoring and inspection is critical for detecting and deterring embezzlement in its early stages. Regular scrutiny of work processes, financial transactions, and internal controls can serve as a proactive measure to safeguard organizational assets. Encourage more project on the extra elements or circumstances that may lead to embezzlement practices. Understanding the evolving landscape of embezzlement risks enables organizations to adapt their preventive strategies to emerging challenges. Conduct follow-up project to assess the efficiency of the adopted preventative measures. Regular evaluation and adjustment of strategies based on empirical evidence ensures that embezzlement prevention remains dynamic and responsive over time.

The main limitation of the research: The research focuses on embezzlement behaviors, utilizing a reliable dataset and comprehensive analysis to identify causes and demographic patterns. Potential biases in self-reported data and the limited universality of findings are acknowledged, suggesting the influence of demographics and cultural factors on embezzlement variations. Consideration of geopolitical factors and industry-specific dynamics is essential for understanding and preventing embezzlement. The use of advanced methodologies like the Generalized Method of Moments (GMM) enhances precision and impartiality in correlating variables, ensuring robust analysis and reducing biases in estimations.

Future work: Future project initiatives should project the incorporation of sophisticated approaches like Generalized Method of Moments (GMM) to solve potential endogeneity difficulties. GMM provides a comprehensive strategy to handle endogeneity, enabling more accurate and unbiased estimations of the correlations between variables. To increase the credibility of project findings, it is critical to maintain a high level of data dependability. Projecters should prioritize data integrity and consistency throughout the data collection process. Rigorous validation procedures, adherence to defined protocols, and clear projecting standards can add to the dependability of project findings. An in-depth examination of variable sources is advised to get a full grasp of how independent and dependent factors impact lymphatic function and performance. Projecters should go into the nuances of variable selection, including the theoretical underpinnings and empirical evidence justifying their inclusion in the project. This strategy ensures a more nuanced interpretation of outcomes and helps to the overall validity of the project.

5. CONCLUSION AND POLICYIMPLICATIONS

In conclusion, our study has made significant strides in understanding embezzlement behaviors. The foundation of a high-reliability dataset facilitated a comprehensive analysis of the causes of embezzlement. Notably, our findings highlight variations in embezzlement practices based on gender and age, emphasizing a higher susceptibility among males and

individuals aged 30–40. The application of correlation and multiple linear regression models revealed compelling relationships between embezzlement and factors such as methods, reasons, and awareness. These insights contribute to a more nuanced understanding of embezzlement dynamics and provide a valuable foundation for future research and preventive measures in addressing white-collar crime.

To enhance organizational integrity and safeguard against embezzlement, implementing a system of continuous monitoring is essential. This involves the establishment of rigorous management processes, ensuring ongoing scrutiny of financial activities to detect and address irregularities promptly. Concurrently, fostering awareness among employees about embezzlement is crucial. By providing targeted training programs, employees can become vigilant, recognizing signs and understanding preventive measures. Additionally, recognizing and rewarding outstanding employee performance, whether during holidays or for exceptional achievements, not only boosts motivation but also alleviates financial burdens. This multifaceted approach contributes to a more secure and resilient organizational environment.

6. ACKNOWLEDGMENTS

This research is funded by International School, Vietnam National University, Hanoi (VNU-IS).

7. REFERENCES

- 2022 *Corruption Perceptions Index: Explore the results*. (2023, January 31). Transparency.Org. <https://www.transparency.org/en/cpi/2022>
- 2023 *Corruption Perceptions Index—Explore Vietnam’s results*. (2024, January 30). Transparency.Org. <https://www.transparency.org/en/cpi/2023>
- Alfaro, E. (2022). Understanding the Corruption Perceptions Index. *Springer Books*, 233–270.
- Begum, M., Rony, J. H., Islam, M. R., & Uddin, J. (2023). Long-Term Software Fault Prediction Model with Linear Regression and Data Transformation. *Journal of Information Systems and Telecommunication (JIST)*, 3(43), 222. <https://doi.org/10.61186/jist.36585.11.43.222>
- Borna, K., & Palizdar, S. (2016). Short Time Price Forecasting for Electricity Market Based on Hybrid Fuzzy Wavelet Transform and Bacteria Foraging Algorithm. *Journal of Information Systems and Telecommunication (JIST)*, 4(16), 1. <https://doi.org/10.7508/jist.2016.04.001>
- Dao-de, S. (2000). Selection of the linear regression model according to the parameter estimation. *Wuhan University Journal of Natural Sciences*, 5(4), 400–405. <https://doi.org/10.1007/BF02850764>
- Ha Le, T. T., & Tran, M. D. (Eds.). (2018). The effect of internal control on asset misappropriation: The case of Vietnam. *Business and Economic Horizons (BEH)*. <https://doi.org/10.22004/ag.econ.287239>
- Jagdeesh. (2023, August 31). Linear Regression Algorithm—Applications and Concepts of Linear Algebra Using the Linear Regression Algorithm. *Machine Learning Plus*. <https://www.machinelearningplus.com/linear-algebra/linear-regression-algorithm/>
- K, K., K, U., A, S., & Kumar, A. (2021). Predicting Student Performance for Early Intervention using Classification Algorithms in Machine Learning. *Journal of Information Systems and Telecommunication (JIST)*, 4(36), 226. <https://doi.org/10.52547/jist.9.36.226>
- Meermeyer, M. (2015). Weighted linear regression models with fixed weights and spherical disturbances. *Computational Statistics*, 30(4), 929–955. <https://doi.org/10.1007/s00180-015-0572-z>
- Ms. Truong My Lan manipulated and appropriated money from SCB Bank—Vietnam.vn. (2023, November 18). <https://www.vietnam.vn/en/ba-truong-my-lan-thao-tung-chiem-doat-tien-ngan-hang-scb/>
- PricewaterhouseCoopers. (n.d.). *Global Economic Crime and Fraud Survey 2018: Vietnam Perspectives*. PwC. Retrieved June 5, 2024, from <https://www.pwc.com/vn/en/publications/vietnam-publications/economic-crime-fraud-survey-2018.html>

- Spearman's Rank Correlation: The Definitive Guide To Understand* | Simplilearn. (n.d.). Simplilearn.Com. Retrieved June 5, 2024, from <https://www.simplilearn.com/tutorials/statistics-tutorial/spearmans-rank-correlation>
- The defendants received the most bribes in the “rescue flight” case—Vietnam.vn.* (2023, July 11). <https://www.vietnam.vn/en/nhung-bi-cai-nhan-hoi-lo-nhieu-nhat-trong-vu-chuyen-bay-giai-cuu/>
- Vaghefi, R. (2021, February 6). *Weighted Linear Regression*. Medium. <https://towardsdatascience.com/weighted-linear-regression-2ef23b12a6d7>
- Van, H. V., Tran, T. Q., Nguyen, T. V., & Steven, L. (n.d.). *Corruption and Firm Financial Performance: New Evidence from Vietnam*.
- Vaske, J., Beaman, J., & Sponarski, C. (2016). Rethinking Internal Consistency in Cronbach's Alpha. *Leisure Sciences*, 39, 1–11. <https://doi.org/10.1080/01490400.2015.1127189>
- What Is Embezzlement? Legal Definition, Types & Examples – Forbes Advisor.* (n.d.). Retrieved June 2, 2024, from <https://www.forbes.com/advisor/legal/criminal-defense/embezzlement/>
- Yu, C., Yao, W., & Bai, X. (2014). *Robust Linear Regression: A Review and Comparison* (arXiv:1404.6274). arXiv. <http://arxiv.org/abs/1404.6274>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

