

A Bayesian Network Analysis of Calendar Effects in the Colombo Stock Exchange

H. K. R. Rathnaweera*  and Rajitha M. Silva 

Department of Statistics, University of Sri Jayewardenepura

*Corresponding author: rashminirathnaweera@gmail.com

Received: 26th June 2023/ Revised: 17th September 2023/ Published: 30th December 2023

©IApstat-SL2023

ABSTRACT

This study applies Bayesian Network analysis to examine the probabilistic causal relationship between calendar effects and stock market anomalies in the Colombo Stock Exchange. While prior research has explored the existence of Calendar Anomalies in the Colombo Stock Exchange, few studies have examined the underlying cause-and-effect relationship between these anomalies and their associated probabilities. This study employs a Bayesian Network model using market data from 2007 to 2020 to investigate this relationship. The results indicate that calendar effects are prevalent in the market, and the analysis identifies a probabilistic causal relationship between abnormal market returns and Day-of-the-Week and Turn-of-the-Month calendar anomalies. The findings of this study enable investors to time their trades by assigning probabilities to positive or negative market returns on specific trading days, maximizing their returns and improving the efficiency of their trades in the Colombo Stock Exchange.

Keywords: Probabilistic Causal Relationship; Stock Market Anomalies; Abnormal Market Returns

1 Introduction

The stock market is a primary place where public companies issue secondary shares to raise capital, thereby ensuring economic expansion and providing opportunities for investors to earn high returns on their investments. However, it is still a much-debated topic among researchers and economists to what extent an individual investor can abnormally earn such profits. In a research article titled "Trading is Hazardous to Your Wealth," (Barber and Odean, 2000)

the authors present empirical evidence supporting the view that overconfidence leads to excessive trading, and the timing of trades hurts investors. Few investors try to perceive a cause-effect relationship between the timing of their trades and the corresponding returns. In statistical terms, such models are known as causal models. This study focuses on building a causal model to understand the stock market anomalies of the Colombo Stock Exchange (CSE).

As country's only licensed stock exchange, the Colombo Stock Exchange has demonstrated consistent growth for nearly four decades, positioning itself as a prominent global stock market. It has received recognition for its achievements from various reputable sources such as the esteemed title of Best Performing Stock Exchange in the World for the month of January, 2021, as indicated in the annual report by the Colombo Stock Exchange (2021). According to the annual report of the Colombo Stock Exchange (2022), the CSE mobilizes capital flows amounting to 686 billion Rupees. This substantial figure includes a notable contribution of 52 billion Rupees from foreign capital flows. Such influx of foreign capital has played a crucial role in fostering a dynamic stock market environment for the stakeholders involved. Despite its global significance, there is a paucity of literature available on the behavioral analysis of CSE. Therefore, it is essential to explore economic theories and postulates that are relevant to the stock exchange to comprehend the nature and behavioral patterns of CSE. This understanding would enable investors, both local and global, to adjust their investing strategies effectively and maximize their returns. One such crucial theory is the Efficient Market Hypothesis (EMH). EMH is a fundamental theory in economics that posits that share prices reflect all available information, resulting in properly priced securities and an efficient market. This hypothesis is also referred to as the efficient market theory. Fama (1970) classified market efficiency into three forms based on the level of relevant information used: weak form, semi-strong form, and strong form. However, market anomalies, including fundamental, technical, and calendar anomalies, are instances that defy the assumptions of the Efficient Market Hypothesis (EMH), as mentioned in Kaur (2019). Calendar anomalies, which are market anomalies associated with various time periods of the calendar, are the focus of this study.

Numerous studies have documented evidence of seasonality in stock returns, including day-of-the-week, month-of-the-year, turn-of-the-year, turn-of-the-month, and holiday effects. Yatiwella and De Silva (2011) conducted a study on the day-of-the-week effect in the Colombo Stock Exchange (CSE) using an Ordinary Least Square (OLS) regression model. The study revealed the existence significant positive returns on Fridays and negative returns on Mondays and Tuesdays. Jahfer (2015) employed a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and obtained comparable results in relation to Yatiwella and De Silva (2011). Deyshappriya (2014) also used the GARCH model to explore the month-of-the-year effect in the

CSE and discovered significantly positive returns in January and significantly negative returns in December.

The month-of-the-year effect was also examined in many foreign stock markets using an OLS regression model. Cheung and Andrew Coutts (1999) explored the Indian and Chinese stock markets whilst Tan and Tat (1998) explored Singapore market. They observed a significant positive return in January in the Singapore market, but the effect was found to be absent in the Indian and Chinese markets. Agarwal et al. (2019) investigated the holiday effect in the Indian and Singapore stock markets, employing both an OLS regression model and a non-parametric Paired sample t-test. They found that the stock markets exhibited higher returns after holidays (post-holiday effect). Regarding the Singapore stock exchange, Tan and Tat (1998) examined the turn-of-the-month effect, discovering that trading days during this period showed three times higher returns compared to other trading days. Shakila et al. (2017) explored the semi-monthly effect using the parametric Mann Whitney U test, finding that stocks earned a positive average return in the beginning and first half of the month, but had a zero average return in the second half of the month.

Hence, it can be seen that these calendar effects have been explored and identified using statistical methods and techniques such as Ordinary Least Square (OLS) Regression, Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, as well as various parametric and non-parametric techniques. However, a significant drawback of these methods is that while they may fit well and satisfy all model assumptions, they do not necessarily indicate a causal relationship among the variables. Therefore, investors should exercise caution when basing their investment decisions solely on the results of such models.

In this paper, we employ a Bayesian network analysis to identify and examine calendar effects in the Colombo Stock Exchange (CSE). Bayesian networks offer a unique advantage in that they allow for the representation of complex relationships and dependencies between variables, including causal relationships. This method offers a robust and reliable approach to analyzing calendar effects, enabling investors to adjust their investment strategies more effectively and maximize their returns. Bayesian Networks (BN) provides a probabilistic graphical model that represents relationships between variables in a dataset using a directed graph. This approach can be used for anomaly detection in stock markets by establishing data-driven causality. However, there is a limited amount of literature on the use of Bayesian Networks for analyzing stock market anomalies in the context of Asian markets, including Sri Lanka. Therefore, this study attempts to fill this gap by modeling stock market anomalies of the Colombo Stock Exchange using Bayesian Networks.

The remainder of this research paper is organized as follows. Section 2 outlines the methodology used in this study, including data collection, data analysis, and model development. Section 3 presents the results of the Bayesian

Network analysis and discusses the implications of the findings. Finally, section 4 concludes the study and provides recommendations for future research.

2 Methodology

2.1 Data Collection

This section outlines the process of data handling for constructing a Bayesian Network to model stock market anomalies of the Colombo Stock Exchange. The study utilized market data from a 14-year period, beginning on January 2, 2007, and ending on December 31, 2020. Data has been limited to no later than the end of 2020 to avoid the economic abnormalities that was caused due to the second wave of COVID-19 which impacted Sri Lanka at the latter part of the same year, in turn resulting in multiple curfews and lockdowns within the country during 2021. Accordingly, CSE being at the forefront of the Sri Lankan capital market was not able to fully function conventionally. Hence the particular year has been ommitted from the sample.

The daily All-Share Price Index (ASPI) and Total Returns Index (TRI) were obtained from the Colombo Stock Exchange's data library upon request. The ASPI serves as a broad market index that measures the overall market's movements, while the TRI reflects returns due to both price changes and dividend income.

In addition to the ASPI and TRI, the study required further data. The 364-day Treasury-Bill Rates were obtained from the Central Bank of Sri Lanka's economic data library. The study also required data on selected calendar anomalies, such as the dates of presidential, local, and general elections held within Sri Lanka from 2007 to 2020, which were obtained from the official website of the Parliament of Sri Lanka. Furthermore, the dates of public holidays within the study period were collected from a verified website called Portal Seven.

2.2 Data Transformation and Coding of Variables

The study requires data transformation to create variables for testing the presence of selected Calendar Anomalies within the CSE, including the Day-of-the-Week (DOW) effect, Month-of-the-Year (MOY) effect, Turn-of-the-Month (TOM) effect, Turn-of-the-Year (TOY) effect, Holiday effect, Effect of Election days, and the unprecedented Black Friday (BF) effect. The DOW and MOY variables can be created based on the corresponding trading day by coding it as one or zero depending on whether it belongs to a particular day of the week or month of the year. For instance, Monday to Friday are coded as 1 to 5, respectively, while January to December are coded as 1 to 12, respectively. The variables TOM and TOY are categorized based on the first day of a new month or year, 3 days before the first day, one week before

the first day, 3 days after the first day, and 1 week after the first day, while all other remaining days are coded as normal trading days.

Regarding the BF effect, it occurs on the fourth Friday of November unless 1st of November is a Friday. ShopperTrak (2010) predicts that Black Friday (Nov. 26) will be one of the strongest sales days of the holiday shopping season.. Thus, the variable BF is created by coding it as one if the date is a Friday falling between the 23rd and 29th of November or if it is a Friday falling on the 1st of November; otherwise, it is coded as zero. Finally, the variables related to holidays and elections are coded by mapping the trading days with the respective dates.

2.2.1 Market Returns

The calculation of daily stock returns for the entire study period is performed using the following formula:

$$R_t = \ln \left(\frac{P_t}{P_{t-1}} \right) \quad (1)$$

Here, R_t denotes the daily stock return at a specific time point t , P_t represents the ASPI value at time point t , and P_{t-1} represents the ASPI value at the immediate preceding time point $t-1$.

2.3 Missing Value Imputation

Due to the presence of missing values in the 364-day Treasury-Bill (T-bill) interest rates, it is necessary to perform imputation before conducting advanced analysis. The Last Value Carried Forward (LVCF) method was employed to impute these missing values. According to this method, if the interest rate for a specific time point t is missing, it is replaced with the interest rate from the previous time point, denoted as $\hat{P}_t = P_{t-1}$. In cases where the interest rate at time $t-1$ is also missing, the value is carried forward from an earlier time point.

2.4 Testing the Weak Form of Market Efficiency

The notion of market efficiency is a dynamic concept that has been extensively explored in literature, highlighting the possibility of inefficient markets transitioning to efficiency over time. The presence of calendar anomalies relies entirely on the existence of inefficiencies within a particular market. Therefore, it is crucial to first examine the weak form efficiency in stock markets before proceeding to test for semi-strong form or strong form efficiency. According to Fernando and Gunasekara (2018), a market cannot be considered highly efficient unless it successfully passes the test of weak form efficiency.

This implies that assessing the weak form efficiency is a fundamental prerequisite for determining the overall level of efficiency in a market. Hence, it is crucial to test the existence or non-existence of weak form of efficiency of the current CSE. The Random Walk Hypothesis (RWH) is being evaluated through the Runs test, Bartels Rank test, and Lo and MacKinlay Variance Ratio test to confirm the presence of market efficiency.

2.5 Capital Asset Pricing Model

The Bayesian Network construction begins with the identification of potential outliers using the Capital Asset Pricing Model (CAPM). The CAPM model establishes a relationship between the expected return and risk of investing in a security. The expected return of an asset is determined by a linear function that incorporates the risk-free rate of return (R_f), the return on the market portfolio (R_m), and the systematic risk measure (β), and can be expressed as

$$E(R) = R_f + \beta[E(R_m - R_f)] \quad (2)$$

The systematic risk represents the stock's volatility of returns relative to the overall market, while the market risk premium ($[E(R_m - R_f)]$) indicates the excess return above the risk-free rate. The CAPM model demonstrates that the expected return on an asset is the sum of the risk-free return and a risk premium based on the security's systematic risk.

The risk-free rate represents the return offered by a zero-risk investment, typically approximated by the return on short-dated government bonds. In this study, the 364 days Treasury Bill Rate is used as a proxy variable for the risk-free rate, while the Total Return Index (TRI) serves as an observable proxy for the market portfolio return. The Capital Asset Pricing Model (CAPM) incorporates an additional coefficient, denoted as α , which indicates whether an asset outperforms or underperforms the expected return predicted by the CAPM. The model can be represented as

$$E(R) - R_f = \alpha + \beta[E(R_m - R_f)] \quad (3)$$

By fitting a regression model and calculating the standard residuals, potential outliers are identified as data points with standard residuals exceeding one.

2.6 Bayesian Regression

The next step involves labeling outlier observations as positive or negative anomalies by assuming that the excess return of a given share at different time points can be more accurately described by parallel regression lines. This approach allows modeling outliers as a shift in the regression mean, which is

accomplished by fitting a Bayesian Regression model to the potential outliers identified through the CAPM.

The Bayesian regression model incorporates subjective priors for the coefficients of the regression equation, enabling the integration of prior beliefs and observed data to derive the posterior distribution of the coefficients. Further details regarding the selection of priors are provided in Section 3.

2.7 Constructing Bayesian Network

Preceding the data transformation and preliminary analysis, provided that the specified conditions are met, the data can be input into a Bayesian Network algorithm to generate the corresponding network structure. These conditions include the absence of weak efficiency within CSE, no missing values in the data, all variables being coded as discrete values, and stock returns being categorized as positive or negative anomalies or normal observations.

For the construction of the Bayesian Network, the R programming language is employed. Specifically, the study utilizes the bnlearn package in R, which offers a range of learning algorithms. The PC Stable algorithm is specifically employed in this study to facilitate the process of fitting the Bayesian Network.

3 Results and Discussion

The dataset used in the analysis covers the period from January 2, 2007, to December 31, 2020, with a total of 3,331 daily time periods. The missing values for the Treasury Bill interest rate were imputed using the Last Value Carried Forward method. Furthermore, the rates were provided in terms of a 364-day period, accordingly the daily rate was calculated assuming a straight-line basis, where the interest gain is evenly distributed over the 364-day period. The stock market returns were plotted against time as shown in Figure 1 and it is evident that there are sudden increases and decreases across the period thus suggesting the presence of significant anomalous stock returns in the market.

The results of the Runs test and Bartels Rank test are as shown in Table 1 where the p-value is less than 0.05 and hence the null hypothesis of the existence of the random walk process is rejected for both tests leading to the conclusion that the Weak Form of efficiency does not exist within the CSE. Furthermore, Table 2 represents the results of the Lo-MacKinlay Variance Ratio test where $M1$ is the resulting test statistic assuming variance is constant and $M2$ is the resulting test statistic assuming variance is not constant. It can be seen that all test statistic values with respect to all lags are greater than 1.96 (The standard normal distribution value at 5% level of significance). Hence,

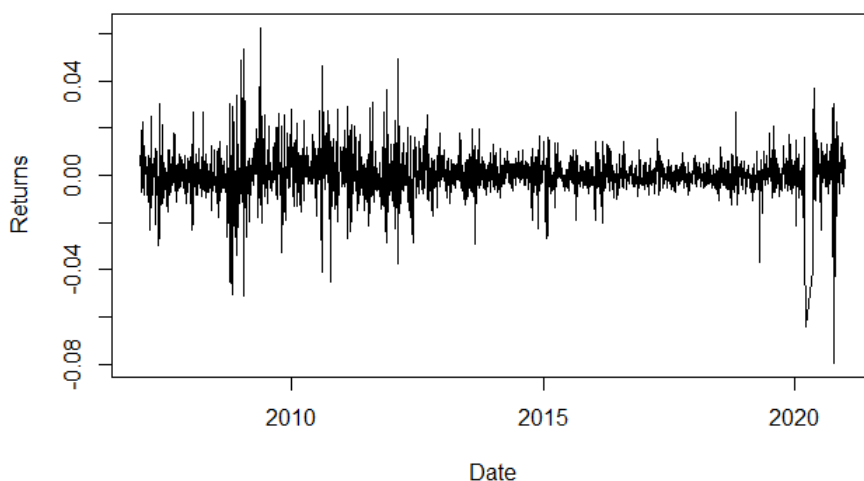


Fig. 1: Distribution of Total Returns Index: 2007 to 2010

the null hypothesis is rejected substantiating that the Weak Form of efficiency does not exist within the CSE.

Test Name	Test Statistic	p-value	Conclusion
Runs Test	-10.087	0.000	The process is not random
Bartels Rank Test	-14.404	0.000	The process is not random

Table 1: Results of Randomness Test

Lag (p)	M1	M2	Conclusion
2	13.04	6.83	The process is not random walk at lag 2
5	15.82	8.14	The process is not random walk at lag 5
10	15.25	8.22	The process is not random walk at lag 10

Table 2: Results of Variance Ratio Test

Subsequent to the data exploration and assessing the absence of weak form efficiency within the CSE, empirical evidence for the possibility of existence of calendar anomalies was established. The CAPM model was then fitted by regressing the excess return on stock with regards to the excess return on market portfolio resulting in statistically significant intercept and slope parameters at 5% level of significance as shown in Table 3. The standard residuals of each observation was calculated and 541 observations out of 3331 was identified as potential outliers. The identified 541 observations are hence further categorized as positive anomalies and negative anomalies by means of Bayesian regression.

Parameter	Coefficient	t value	p-value
Intercept	4.47×10^{-2}	-59.63	0.00
Slope	2.99×10^{-6}	28.71	0.00

Table 3: Results of CAPM model

The priors defined with respect to the Bayesian regression model fitted to the potential outliers are as shown below whereby systematic risk and risk adjusted performance of an asset are assumed to be distributed Normally hence enabling the parameters to take a positive or a negative value. The prior for standard deviation is defined in terms of precision that is the reciprocal of the standard deviation and assumed to follow a Gamma distribution.

Priors: $\alpha \sim N(0, 0.001)$, $\beta \sim N(0, 0.001)$, $\tau \sim Gamma(0.1, 0.1)$

The distribution of the α values (intercepts) of set of parallel regression lines obtained through the posterior simulated with respect to the potential outliers are as shown in Figure 2. The α values range from a minimum of -1.05 to a maximum of 0.93. The median of the α values was calculated as -0.07. Accordingly, given that the deviation of the α from the median of the α values is negative the return on that particular trading day is categorized as a negative anomaly else as a positive anomaly. Hence, 271 were identified as positive anomalies and 270 was identified as negative anomalies out of the total 541 potential outliers.

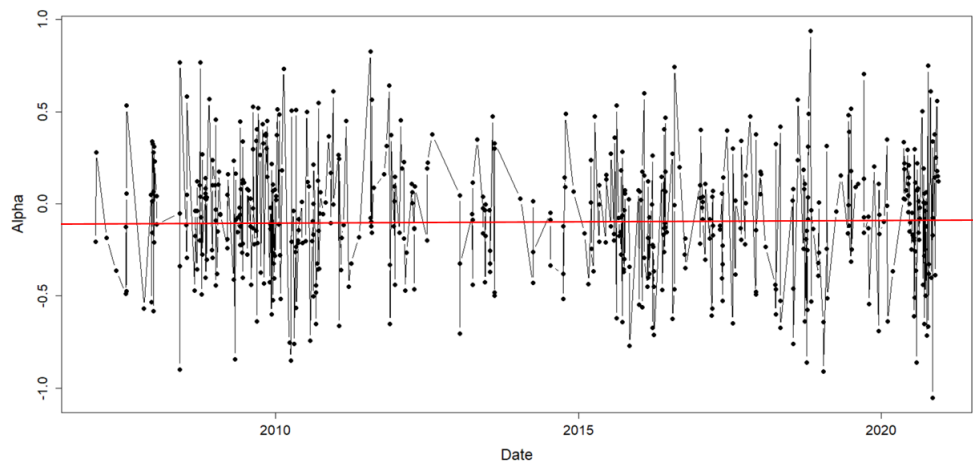


Fig. 2: Distribution of α

Thereafter, the data consisting of the outlier categorization and the categorical calendar effect variables were fed to a PC stable algorithm and the resulting Bayesian Network (BN) is as shown in Figure 3. The directed arcs from DOW (Day-of-the-Week) and TOM (Turn-of-the-Month) to Abnormal Returns indicate that only DOW and TOM have a causal relationship with Abnormal Returns, while the other calendar anomalies are not present in the

study's domain. This probabilistic dependence, representing cause-effect relationships, can be expressed through Conditional Probability Tables (CPTs). The CPT table for the obtained BN provides the probability of earning different types of returns on any given day of the year. For example, if a trading day is a Monday falling on the first day of the month, there is an approximately 30% probability of experiencing negative returns. The probability table also enables the comparison of probabilities between positive and negative returns. For instance, the probability of earning negative returns is twice as high as earning positive returns when the trading day is a Wednesday falling on the first day of the month.

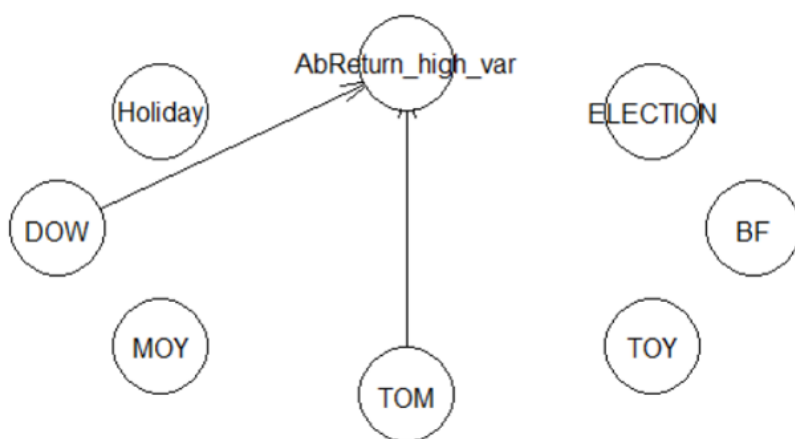


Fig. 3: Bayesian Network Model for Stock Returns of CSE

			Day of week				
			Monday	Tuesday	Wednesday	Thursday	Friday
Turn of Month	First day	Normal	0.513	0.586	0.333	0.421	0.683
		Abnormal Positive	0.210	0.189	0.252	0.271	0.158
		Abnormal Negative	0.277	0.225	0.414	0.308	0.158
	3 days before	Normal	0.741	0.790	0.785	0.607	0.749
		Abnormal Positive	0.129	0.105	0.093	0.228	0.096
		Abnormal Negative	0.129	0.105	0.122	0.165	0.155
	1 week before	Normal	0.779	0.790	0.711	0.711	0.733
		Abnormal Positive	0.125	0.105	0.129	0.099	0.102
		Abnormal Negative	0.096	0.105	0.160	0.190	0.165
	3 days after	Normal	0.575	0.541	0.544	0.612	0.648
		Abnormal Positive	0.196	0.244	0.228	0.179	0.109
		Abnormal Negative	0.228	0.215	0.228	0.208	0.243
	1 week after	Normal	0.648	0.725	0.711	0.680	0.772
		Abnormal Positive	0.176	0.170	0.129	0.160	0.099
		Abnormal Negative	0.176	0.105	0.160	0.160	0.129
	Others	Normal	0.841	0.838	0.804	0.793	0.898
		Abnormal Positive	0.076	0.087	0.107	0.110	0.050
		Abnormal Negative	0.083	0.075	0.088	0.097	0.052

Fig. 4: Conditional Probability Table

4 Conclusion

Empirical studies have extensively examined the weak form efficient Market Hypothesis and calendar anomalies in the context of the Colombo Stock Exchange (CSE). These studies have demonstrated the absence of market efficiency and the existence of various seasonality patterns in stock returns. However, they have not established a causal probabilistic relationship between these seasonality patterns and stock market returns. This study aims to fill this gap by modeling such a probabilistic relationship using CSE market data spanning a 14-year period. The findings of the study confirm the absence of weak form efficiency in the CSE, indicating the presence of market anomalies. This aligns with previous research that consistently provides evidence against weak form efficiency over the years. Therefore, the applicability of the dynamic market concept is questionable in the case of the CSE. The market's inefficiency allows for further exploration of calendar anomalies.

The Bayesian Network (BN) reveals a causal relationship between abnormal market returns and the Day-of-the-Week and Turn-of-the-Month variables. The Conditional Probability Table (CPT) derived from the BN provides probabilities of earning different types of returns on a given trading day, including normal returns, abnormal positive returns, and abnormal negative returns thus enabling investors to strategically plan and time their trades to maximize their returns. Notably, the CPT shows a high probability (41%) of experiencing abnormal returns on Wednesdays that fall on the first day of the month. This suggests that investors tend to make poor investment decisions on such days due to a combination of high tendency for instantaneous spending (Hastings and Washington, 2010) and stress and exhaustion (Miura, 2018).

5 Recommendations

Bayesian Networks are capable of capturing causality in complex domains and facilitate effective communication between statisticians and non-statisticians due to their intuitive interface. However, a limitation of BNs is the requirement for accurate prior probabilities, as inaccurate priors can lead to misleading results. In this study, the BN was learned entirely through an algorithm, but incorporating accurate expert knowledge can further enhance the results. Additionally, the use of discrete variables in the BN may result in information loss, and future improvements could explore techniques for handling continuous variables.

It is important to note that the study was conducted using data from a conventional time period and may not directly generalize to unconventional situations that cause economic instability, such as political instability, hyperinflation, or currency crises. However, the model can be enhanced to account for such crisis situations by considering the derived probabilities from the

Conditional Probability Table as prior probabilities and incorporating external factors that may impact stock returns during crises based on expert judgment. The developed model is just the beginning, and there are multiple approaches to further improve it and account for various possibilities.

Acknowledgement

We would like to express our sincere gratitude to all the academic and non-academic staff of the Statistics Department of University of Sri Jayewardenepura.

References

- Agarwal, A., Dahiya, K. L., and Gupta, P. (2019). Holiday effect in select asian stock markets *Business Analyst*, 40(2), 3-18.
- Barber, B. M. and Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance*, 55(2), 703-806.
- Cheung, K. C. and Andrew Coutts, J. (1999). The january effect and monthly seasonality in the hang seng index: 1985-97 *Applied Economics Letters*, 6(2), 121-123.
- Colombo Stock Exchange. (2021). Annual Report 2021.
Retrieved from https://cdn.cse.lk/pdf/CSE_Annual_Report_2021.pdf
- Colombo Stock Exchange. (2022). Annual Report 2022.
Retrieved from https://cdn.cse.lk/pdf/CSE_Annual_Report_2022.pdf
- Deyshappriya, N. R. (2014). An empirical investigation on stock market anomalies: The evidence from colombo stock exchange in sri lanka. *International Journal of Economics and Finance*, 6(3), 177-187.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.
- Fernando, P. N. D., & Gunasekara, A. L. (2018). Is the market efficiency static or dynamic—evidence from Colombo Stock Exchange (CSE). *Kelaniya Journal of Management*, 7(1), 13-25.
- Hastings, J. and Washington, E. (2010). The first of the month effect: consumer behavior and store responses. *American economic Journal: economic policy*, 2(2), 142-162.

- Jahfer, A. (2015). Calendar effects of the Colombo stock market. *Journal of Management*, 12(2), 121-132.
- Kaur, G. (2019). January Anomaly: A Myth or Reality?. *Business Analyst*, 39(2), 169-190.
- Miura, A. (2018). Get psyched: Why wednesday is the worst (give monday a break). *Business Analyst*, 39(2), 169-190.
- Shakila, B., Pinto, P., and Hawaldar, I. T. (2017). Semi-monthly effect in stock returns: new evidence from Bombay stock exchange *Investment Management and Financial Innovations*, 14(3), 160-172.
- ShopperTrak. (2010, November 16). ShopperTrak Reports Positive Response to Early Holiday Promotions Boosts Projections for 2010 Holiday Season [Press release]. Retrieved from <http://web.archive.org/web/20101129122823>
- Tan, R. S. K. and Tat, W. N. (1998). The diminishing calendar anomalies in the stock exchange of Singapore. *Applied Financial Economics*, 8(2), 119-125.
- Yatiwella, W. B. and De Silva, J. (2011). Calendar anomalies: Evidence from the Colombo stock exchange. *Economics, Management and financial markets*, 6(2), 84-105.