Data Science in Finance and Economics

https://www.aimspress.com/journal/dsfe

DSFE, 5(1): 1–18.
DOI:10.3934/DSFE.2025001
Received: 08 October 2024
Revised: 25 December 2024
Accepted: 10 January 2025
Published: 14 January 2025

Research article

Discovering causal relationships among financial variables associated with firm value using a dynamic Bayesian network

Ji Young Choi, Chae Young Lee and Man-Suk Oh*

Department of Statistics, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760,Korea

* **Correspondence:** Email: msoh@ewha.ac.kr; Tel: +8232772374.

Abstract: This study investigated the causal relationships among financial variables associated with firm value using a Causal Dynamic Bayesian Network (CDBN), which is an extension of the basic Bayesian network that captures both temporal and contemporaneous causal relationships. The CDBN model was constructed using a panel dataset of listed manufacturing companies in Korea over a 14-year period (2009–2022). By visualizing the interactions between financial factors, the model makes it easy to understand their dynamic and instantaneous relationships, offering valuable insights into corporate finance. Key findings in the model include evidence of autocorrelation in all dynamic variables, a lagged feedback loop between the intangible assets ratio and firm value, the widespread impact of the COVID-19 pandemic on the financial sector, and important causal relationships involving key financial metrics such as the fixed assets ratio, firm value, and return on assets ratio.

Keywords: firm value; dynamic Bayesian network; causality; panel data; lagged effect

JEL Codes: C14, C33, C45, G32

1. Introduction

Firm value, also known as corporate valuation or enterprise value, represents the total economic worth of a business entity, encompassing both equity and debt components. It serves as a critical measure to assess a company's financial health, growth potential, and overall performance (Dwicahyani et al., 2022; Suzan et al., 2023). Firm value also functions as a benchmark to evaluate investment opportunities, inform strategic decisions, conduct mergers and acquisitions, and appraise overall business outcomes.

The determination of firm value depends on various factors, including a company's future cash flows, growth prospects, risk profile, and prevailing market conditions. These factors often interact in complex ways, making their interplay crucial for understanding and predicting firm value. A comprehensive analysis that accounts for these interactions empowers investors and management to optimize firm value and make informed, strategic decisions essential for effective management.

Regression models have been widely employed in studies examining firm value (Cheng et al., 2010; Sunarsih et al., 2019; Hirdinis, 2019; Ispriyahadi and Abdullah, 2021; Siregar et al., 2023). However, these models typically rely on predefined functional forms and require the prior specification of both response and explanatory variables. Additionally, regression models often face challenges in identifying complex association terms among explanatory variables. Moreover, uncovering causal relationships using regression models is inherently challenging without relying on strong assumptions, such as the presence of randomized trials or the use of instrumental variables.

Bayesian Networks (BNs) have recently gained significant attention as a powerful analytical method for handling highly complex interdependencies across various disciplines, including economics, medicine, biology, and environmental science. A BN is a directed probabilistic graph that connects related variables with edges, where these connections represent conditional dependencies between the variables (Pearl, 1988). By employing graphical structures, BNs provide a visual representation of dependencies, enabling intuitive interpretation and facilitating a deeper understanding of underlying dynamics. Foundational work on BNs is attributed to Pearl (1988), and recent works that provide comprehensive summaries of BNs and their applications include Friedman et al. (1997), Koski and Noble (2011), and Scutari and Denis (2021).

BNs provide significant advantages over traditional regression methods, particularly in complex and uncertain contexts such as firm value analysis. First, BNs do not require predefined functional forms, allowing them to uncover interdependencies and key interactions among variables directly from the data. Second, BNs can provide a formal framework for causal reasoning, enabling deeper insights into how changes in one variable can propagate through the network to influence firm value. This capability is crucial for realistic scenario analyses, allowing analysts to test various potential conditions and observe their effects on firm value comprehensively. Third, BNs can incorporate expert knowledge, such as industry-specific risks or forward-looking insights. This is especially valuable when historical data is limited or when market conditions are expected to deviate from past trends (Madden, 2009; Liu and Motoda, 2012; Oh et al., 2022; Jo et al., 2023). By integrating data-driven insights with expert knowledge, BNs provide a flexible and insightful framework for firm value analysis, supporting more accurate and resilient valuations in complex financial contexts. Finally, BNs handle missing data naturally through probabilistic reasoning, enabling robust analysis even in incomplete datasets. In contrast, regression models typically address missing data through imputation techniques or by excluding incomplete cases, which can introduce bias or reduce the sample size. This inherent advantage of BNs ensures a more reliable analysis in data-constrained scenarios.

Given these advantages, BNs have gained considerable attention in recent years for their application in firm value analysis (Ali and Anis, 2012; Sun, 2015; Sun and Park, 2017; Teles et al., 2020; Cao et al., 2022; Chan et al., 2023). However, despite their growing prominence, most existing research overlooks the temporal dependencies and causal structures in this domain.

Firm value is influenced by factors that are typically observed over consecutive time periods and often exhibit temporal dependencies. In other words, the value at a given time point is frequently affected by values from preceding periods (Campbell et al., 1997; Tsay, 2010; Enders, 2015). Considering these temporal dependencies is essential for understanding how these factors evolve and interact dynamically over time.

Causal discovery offers significant advantages in understanding complex systems by identifying underlying cause-and-effect relationships rather than relying solely on correlations. This enables better prediction, explanation, and intervention strategies. By revealing the mechanisms driving observed data, causal discovery enhances decision-making, supports the development of effective policies, and enables targeted interventions. Researchers are particularly interested in identifying causal structures from observational data rather than experimental data, because observational data often comes from naturally occurring events or retrospective studies, while experimental data is often difficult to obtain in real-world settings.

This study investigated the *causal* relationships among factors influencing firm value, while accounting for *temporal* dependencies. To achieve this, a Causal Dynamic Bayesian Network (CDBN) model was developed using panel data from listed manufacturing companies in Korea, spanning the period from 2009 to 2022.

Dynamic Bayesian Networks (DBNs) build on traditional BNs by incorporating temporal dependencies, enabling the identification of both contemporaneous and temporal relationships among variables. This allows DBNs to capture interactions at a single time point as well as the evolution of these relationships over time (Dagum et al., 1992). CDBNs further extend the DBN framework by explicitly modeling causal relationships. In CDBNs, edges represent cause-and-effect relationships, facilitating the identification of both within-time and between-time causal pathways.

The remainder of this paper is organized as follows: Section 2 describes the data used in this study. Section 3 outlines the theoretical foundation and structure of the CDBN. Section 4 details the methodology for constructing the CDBN model. Section 5 presents the results, including the final CDBN model and the identified causal relationships. Section 6 concludes the study, and Section 7 discusses its limitations and suggests directions for future research.

2. Data

The data used in this study is a panel dataset comprising 11 variables from 227 listed manufacturing companies in Korea, observed over 14 years (2009–2022). Annual financial statement data and market capitalization were sourced from DeepSearch, a corporate data platform (https://www.deepsearch.com/). Foreign ownership percentages were extracted from the Korea Exchange (KRX) Information Data System (http://data.krx.co.kr/contents/MDC/MAIN/main/index.cmd), capturing values from the last day of trading each year.

Among the various approaches to measuring firm value, Tobin's Q (Tobin, 1969) is the most widely used metric. Tobin's Q is defined as the ratio of a company's market value to its replacement cost or book value. A Tobin's Q greater than 1 indicates that the company is overvalued, while a value less than 1 suggests that it is undervalued. In this study, firm value was measured by considering the market value of a company as the sum of its market capitalization and total debt. Tobin's Q was then calculated as the ratio of this sum to the company's total assets.

Financial variables related to firm value were selected based on the previous research (Francis and Schipper, 1999; Pástor and Pietro, 2003; Rountree et al., 2008; Sun and Park, 2017). The selected variables are Sales growth rate (Growth), Return on assets ratio (ROA), Tangible assets ratio (TA), Company leverage (LEV), Liquidity (LIQ), Fixed assets to long-term capital ratio (FA), Firm size (Size), Foreign ownership ratio (FOR), Intangible assets ratio (ITA), and Depreciation charge (DC). These variables reflect a company's sales growth, profitability, stability and liquidity, capital structure, and other characteristics.

Additionally, numerous studies have highlighted the significant impact of the Coronavirus Disease 2019 (COVID-19) on the financial sector (Afrina et al., 2020; Hertati et al., 2020; Mishra and Mishra, 2020). To incorporate the effects of COVID-19 in the analysis, a dummy variable, COVID-19, was introduced to differentiate between the pre-COVID period (2009–2019) and the post-COVID period (2020–2022).

The final dataset comprises observations for 11 variables across 227 companies over 14 years, along with the COVID-19 dummy variable. Table 1 provides detailed descriptions, means, and standard deviations of these variables.

Table 1. Descriptions and summary statistics of the variables.

Name	Description	Formula	Mean	SD
COVID-19	Before/After COVID-19	0=(Year<=2019), 1= (Year>=2020)	-	-
Size	Logarithm of Total Assets	ln(Total Assets)	26.7294	1.4735
TobinsQ	Tobin's Q	(Market Capitalization+Total Debt)/Total Assets	1.1002	0.8804
DC	Depreciation Charge	Depreciation/Total Assets	0.0032	0.0050
LEV	Leverage	Total Debt/Total Capital	0.9366	1.4763
LIQ	Liquidity	Cash&Cash Equivalents/Total Assets	0.0535	0.0577
TA	Tangible Assets Ratio	Total Tangible Assets/Total Assets	0.3226	0.1692
ITA	Intangible Assets Ratio	Total Intangible Assets/Total Assets	0.015	0.036
Growth	Growth Rate of Sales	(Current Sales - Previous Sales)/Previous Sales	0.0007	0.0039
ROA	Return on Assets Ratio	Operating Profit/Total Assets	0.0346	0.0689
FOR	Foreign Ownership	Number of Foreign Shares/Total Number of Shares	0.0971	0.1282
FA	Fixed Assets Ratio	Fixed Assets/Long-term Capital	0.8495	0.3396

3. Causal dynamic bayesian network

A BN is a machine learning technique that models relationships between variables as a probabilistic graph, consisting of nodes and directed acyclic edges. These edges represent dependencies between variables, based on the concept of conditional independence (Pearl, 1988). In a Directed Acyclic Graph (DAG), each node corresponds to a variable, while the edges illustrate the dependencies, with their presence and direction determined by conditional probabilities.

The starting point of a directed edge is referred to as the parent node, and the receiving end is called the child node. If the i-th node in a BN represents variable X_i for $i = 1, ..., n$, the parent set $Pa(X_i)$ of X_i is the set of variables from which edges originate and point towards X_i. Then the joint probability distribution of $\mathbf{X} = (X_1, ..., X_n)$ can be represented by Equation (**??**) according to the product rule of probability.

$$P(X_1, ..., X_n) = \prod_{i=1}^{n} P(X_i|Pa(X_i)) \tag{1}$$

BNs use Equation (**??**) to visually display the relationships between variables through DAGs, making it easy to interpret the relationships.

A DBN is an extension of a traditional (static) BN to encounter temporal dependencies in dynamic variables observed over time. Developed by Dagum et al. (1992), DBNs have been applied to various time series data, including those by Çambaşı et al. (2019), Chang et al. (2023), and Lee and Kwon (2023).

To effectively model temporal dependencies, DBNs rely on several key assumptions: 1) Discrete Time Slices: Data is observed at distinct, discrete time points, enabling the analysis of temporal patterns and trends; 2) Markov Property: Variables at each time point are conditionally independent of all earlier time points, given the variables at the immediately preceding time point. This assumption simplifies the network structure and computational requirements; 3) Stationarity: The relationships among variables within a time slice and the transition probabilities between consecutive time slices are assumed to remain constant over time, further streamlining the model's complexity.

Under the aforementioned assumptions, a DBN comprises two components: a static BN (B_0) and a transition network ($\vec{B}$) (Li et al., 2013). The static network B_0 represents intra-slice dependencies, capturing contemporaneous relationships among variables within the same time slice t. The transition network $\vec{B}$ models inter-slice dependencies, describing how variables at time $t-1$ influence those at time t and capturing temporal (lagged) effects. Together, these components allow a DBN to effectively model both contemporaneous and temporal dependencies.

A CDBN builds on the DBN framework by explicitly modeling causal relationships. In a CDBN, edges represent causal mechanisms rather than mere statistical dependencies, integrating causal discovery into the DBN structure. This enhancement allows for a deeper understanding of how changes in one variable causally affect others, offering clear explanations for observed outcomes and improving the model's interpretability.

4. Construction of a causal dynamic bayesian network model

This section describes the process of constructing the CDBN model using panel data from listed manufacturing companies in Korea.

To infer the structure of the CDBN from the data, the study employed the PC algorithm, originally developed by Spirtes and Glymour (1991) and Spirtes et al. (2000). The PC algorithm is a constraint-based method widely used to uncover the causal structure of a BN.

The key assumptions of the PC algorithm are as follows: i) All relevant variables are included in the dataset, ensuring that there are no hidden or unmeasured confounding variables that influence relationships among the observed variables (causal sufficiency); ii) Given the structure of the causal graph, each variable is conditionally independent of its non-descendants, given its parents (causal Markov condition); iii) the observed independencies in the data must align with the independencies implied by the underlying causal structure (faithfulness).

A brief outline of the PC algorithm is as follows (Spirtes and Glymour, 1991; Spirtes et al., 2000):

- Start with a fully connected graph: Start with an undirected graph where each node represents a variable, and every pair of nodes is connected.
- Skeleton formation: Iteratively test each pair of nodes for conditional independence, given various subsets of adjacent nodes (conditioning sets). Remove an edge between two nodes if they are conditionally independent. After testing all possible pairs, the remaining undirected graph (skeleton) represents the dependencies between nodes.
- Collider orientation (V-Structure): For any three nodes X, Z, and Y, where X and $Y$are not directly connected but share a common neighbor Z (i.e., the structure $X-Z-Y$), if X and Y are conditionally

independent given any subset that does not include Z, then Z is a collider. The edges are directed as $X \rightarrow Z \leftarrow Y$, forming a V-structure.

- Propagate edge directions: Orient the remaining edges based on constraints to maintain a DAG structure and prevent the formation of new colliders. Iterate until no further orientations are possible without creating cycles or violating the conditional independencies inferred from the data.

The data used in this study includes a dummy variable, COVID-19, which indicates the outbreak of the COVID-19 pandemic along with a panel data consisting of observations for 11 variables from 227 companies in Korea over 14 years (2009–2022). To construct a CDBN for this panel data, the data was reshaped to ensure that observations for each dynamic variable were available at two consecutive time points, t_0 and t_1, for each company (Scutari et al., 2024).

Let X_{ij} be the jth variable for the ith company, with 14 observations corresponding to the years 2009-2022. For each X_{ij}, new variables X_{ijt_0} and X_{ijt_1} were created: X_{ijt_0} is a collection of observations of X_{ij} from the years 2009-2021 and X_{ijt_1} from the years 2010–2022. This restructuring aligns each observation of X_{ij} at t_0 with the corresponding observation at t_1, The new data has variables at time t_0 and t_1, enabling the construction of edges from the previous time point to the current time point. The PC algorithm can then be applied to both the static (intra-slice) and transition (inter-slice) components of the DBN, facilitating the identification of causal relationships within and across time slices.

To ensure the validity of the causal structure, the following constraints were applied during the construction of the causal network: i) Edges from any financial variables to COVID-19 were prohibited, as financial variables cannot causally influence the occurrence of COVID-19; ii) To maintain temporal causality, edges from t_1 to t_0 were disallowed, ensuring that future events do not exert any influence on past events; iii) Variables at time t_0 were not permitted to influence each other, reflecting an assumption of no contemporaneous causation within this timeframe; iv) Variables at time t_1 were allowed to influence each other, accounting for potential contemporaneous causation in this timeframe.

The PC-Stable algorithm (Colombo and Maathuis, 2014), a robust variation of the original PC algorithm, was used to improve the consistency and reliability of structure learning (Kalisch et al., 2024). The algorithm was implemented using the `pc.stable` function from the R package *bnlearn* (Scutari, 2023).

A single dataset does not provide sufficient information on the reliability of the relationships between nodes in the network. To address this limitation, bootstrap aggregation (bagging) was applied. For each bootstrap sample, 181 companies (approximately 80% of the dataset) were randomly selected, with replacement, and the variable order was shuffled to increase sample variability (Scutari et al., 2024). The bootstrap sampling was repeated 1500 times and the CDBN structure was learned from each of the samples. From the resulting 1,500 CDBN networks, edge strength and direction were computed for each edge. The edge strength is the relative frequency of the edge's presence across the networks, regardless of direction. The edge direction is the relative frequency of the edge's specific direction across the networks. Both metrics range from 0 to 1, with values close to 1 indicating strong support for the presence or direction of the edge (Friedman et al., 1999).

Using model averaging, a consensus network was constructed by including only edges with edge strengths exceeding a specified threshold to eliminate weak or spurious connections and retain only robust relationships. Higher thresholds correspond to sparser networks. Following the approach by Scutari and Nagarajan (2013), the threshold was determined as the edge strength significance level that minimized the L1-norm between the cumulative density function (CDF) of the observed edge strengths and its asymptotic counterpart. For this study, a threshold of approximately 0.5 was applied.

The final consensus network, created using the `averaged.network` function in the R package *bnlearn*, includes robust causal relationships supported by the data. The edge strengths and directions for the final CDBN are presented in Table 2.

Table 2. Edge strengths and directions in the final CDBN.

inter-year				intra-year			
From($t-1$)	To(t)	Strength	Direction	From	To	Strength	Direction
Size	Size	1.000	1.000	COVID-19	TobinsQ	0.925	1.000
TobinsQ	TobinsQ	1.000	1.000	COVID-19	DC	0.987	1.000
TobinsQ	ITA	0.829	1.000	COVID-19	LEV	0.610	1.000
DC	DC	1.000	1.000	COVID-19	LIQ	0.582	1.000
LEV	LEV	1.000	1.000	COVID-19	ITA	0.839	1.000
LIQ	LIQ	1.000	1.000	COVID-19	Growth	0.962	1.000
TA	TA	1.000	1.000	COVID-19	ROA	0.697	1.000
ITA	TobinsQ	0.713	1.000	COVID-19	FA	0.535	1.000
ITA	ITA	1.000	1.000	TobinsQ	LIQ	0.670	0.555
Growth	Growth	0.825	1.000	LIQ	FA	1.000	0.685
ROA	ITA	0.553	1.000	TA	FA	0.975	0.884
ROA	ROA	1.000	1.000	Growth	ROA	0.920	0.731
FOR	FOR	1.000	1.000	ROA	DC	0.783	0.512
FA	FA	1.000	1.000	FOR	ROA	0.636	0.847
				LEV	FA	1.000	0.819

The flow chart of the CDBN model development process is shown in Figure 1.

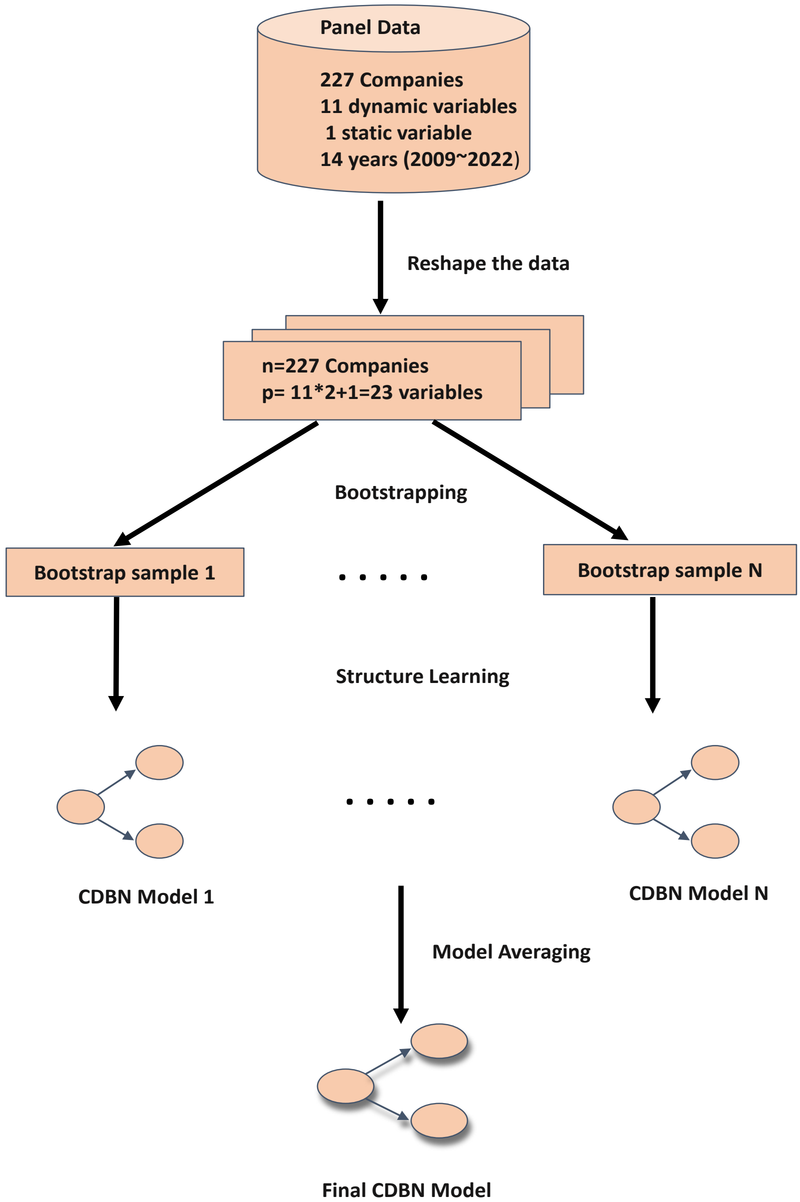

Figure 1. Flow chart of the CDBN model development process.

5. Results

The constructed CDBN model, displayed in Figure 2, consists of (a) inter-year (temporal) causal relationships and (b) intra-year (contemporaneous) causal relationships. In (a), autocorrelation effects are represented by dashed lines to distinguish lagged effects between different variables. In (b), two edges, one between LIQ and TobinsQ and another between ROA and DC, have edge directions close to 0.5. This indicates that, while these pairs of variables are related, the cause-and-effect relationships between them are ambiguous. To represent this uncertainty in directionality, these edges are depicted as blue bi-directional lines.

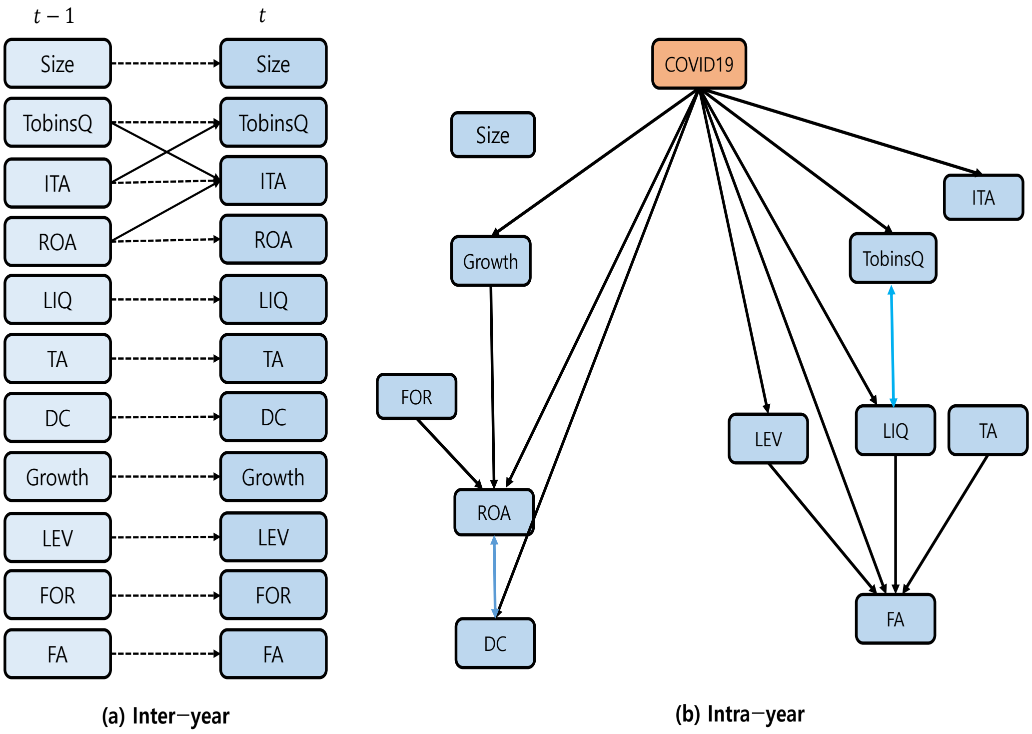

Figure 2. The Causal Dynamic Bayesian Network structure of financial variables relevant to firm value for listed manufacturing companies in Korea consists of (a) inter-year (temporal) causal relationships and (b) intra-year (contemporaneous) causal relationships. In (a), autocorrelation effects are represented by dashed lines, while in (b), edges with directionality close to 0.5 are displayed as blue bi-directional lines.

Figure 2 highlights that all dynamic variables exhibit autocorrelation effects. It is well known that in most financial variables, past values tend to influence future values due to market inertia, information dissemination processes, and investor reactions.

Notably, a lagged feedback loop exists between ITA and TobinsQ: ITA positively influences TobinsQ in subsequent year, which, in turn, elevates ITA at the following year, thereby creating a virtuous cycle. ITA is a non-physical asset that provides value to a company, such as research and development (investments in innovation and new product development), intellectual property (patents, trademarks, copyrights), goodwill, brand recognition, and proprietary technologies.

The lagged feedback loop between intangible assets and TobinsQ may arise because of the delayed impact of intangible investments on both financial statements and market value. When firms invest in intangible assets, accounting conservatism requires these costs to be expensed immediately rather than capitalized. This keeps book value low and initially suppresses TobinsQ (Chen and Srinivasan, 2024; McNichols et al., 2014). Over time, as these intangible investments begin to generate measurable returns - such as increased revenues, competitive advantage, or market share - the firm's market value rises, which eventually lifts TobinsQ. This adjustment occurs gradually, introducing a temporal lag between the time of investment and the point at which TobinsQ reflects the full value of these intangibles.

An increase in TobinsQ, due to the market recognizing the value of intangible assets, may prompt companies to focus on intangible growth, leading to further improvements in intangible assets. This cycle can create a feedback effect where higher TobinsQ stimulates investment in intangibles, which, in turn, could elevate TobinsQ further.

In"New Economy" firms - companies focused on technology, digital services, biotech, and other sectors driven by intangible assets - the lagged feedback effect between intangible assets and TobinsQ can be pronounced. This is because these firms often rely heavily on intangible assets, which are underreported in financial statements (Core et al., 2003).

The lagged feedback loop between ITA and TobinsQ suggests that the market often takes time to recognize the value of intangibles, impacting investment strategies for various stakeholders. For investors, this effect presents an opportunity to identify and invest in undervalued firms with significant intangible assets, anticipating future gains as market recognition catches up. Corporate managers can use this insight to justify strategic investments in intangibles such as R&D and brand-building, even if these do not immediately boost market value, while enhancing transparency around these assets to reduce the recognition lag. Policymakers, recognizing the limitations of current accounting standards, may consider reforms to improve intangible asset reporting and introduce incentives that encourage firms to invest in intangibles, fostering innovation and supporting knowledge-driven sectors. This understanding promotes a balanced approach, focusing on sustainable long-term growth rather than short-term gains.

Valuation of the intangible assets has been a widespread topic of interest and many studies revealed the positive effect of intangible assets to firm value. See Nagaraja and Vinay (2016), Ocak and Findik (2019), Tsai et al. (2012), Glova and Mrázková (2018), Mohammed and Al-Ani (2020), among others. The lagged feedback loop between ITA and TobinsQ found in this study goes in line with the results of previous research.

Additionally, the study reveals a lagged effect of ROA on ITA. A higher ROA in the current year leads to an increase in ITA in the following year. This indicates that as companies experience better asset utilization and profitability, they are more likely to reinvest those returns into intangible assets after a period of financial evaluation and strategic decision-making.

The intra-year structure, depicted in Figure 2(b), underscores the broad impact of COVID-19 on the financial sector. COVID-19 directly or indirectly influences all variables except Size, FOR, and TA. Notably, the distributions of Size, FOR, and TA remained nearly identical between the pre-COVID and post-COVID periods. Furthermore, these variables exhibit high autocorrelation coefficients—0.995 for Size, 0.964 for FOR, and 0.942 for TA—indicating that their values from the previous year almost entirely determine their current-year values. As a result, given their previous-year values, these variables need no other parent nodes in the model, with their only parent being their respective value from the previous year.

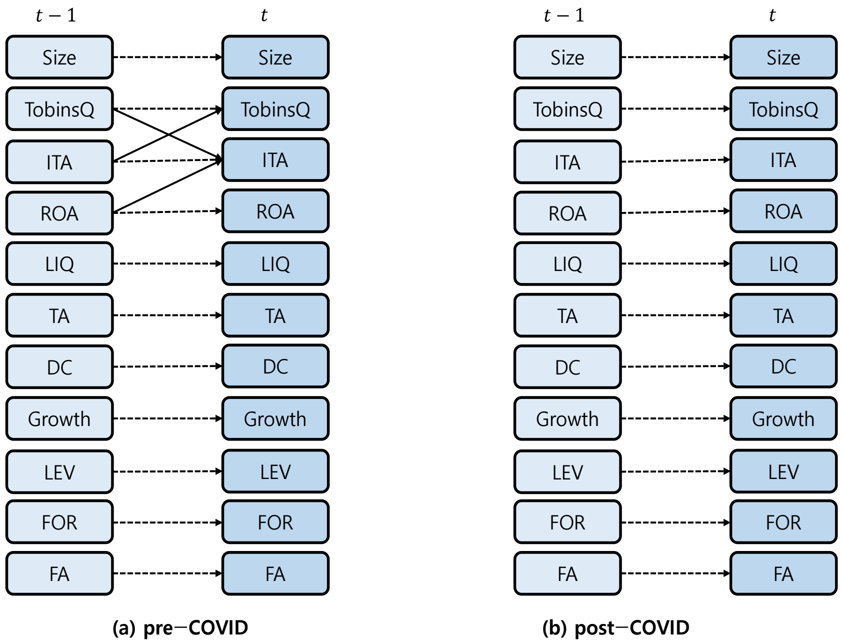

Figure 3. The inter-year (temporal) structures of the pre-COVID and post-COVID CDBN models. The autocorrelation effects are represented by dashed lines.

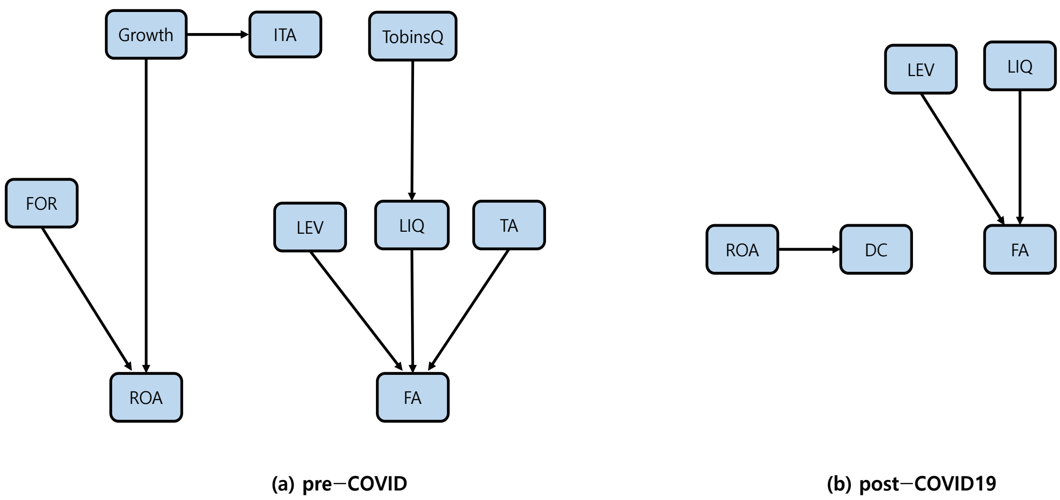

Figure 4. The intra-year (contemporaneous) structures of the pre-COVID and post-COVID CDBN models. Isolated nodes are omitted.

To examine the impact of COVID-19 in greater detail, particularly its influence on the causal network structure among financial variables, the dataset was divided into pre-COVID (2009–2019) and post-COVID (2020–2022) subsets, and separate CDBN models were constructed (Figures 3 and 4). Figure 3 illustrates the inter-year structures of the pre-COVID and post-COVID CDBN models, while Figure 4 depicts their intra-year structures. Figure 3 illustrates that, while autocorrelation effects are maintained, the lagged feedback loop between ITA and TobinsQ, as well as the lagged impact of ROA on ITA, are absent in the post-COVID model. Similarly, Figure 4 shows that only the edges from LEV to FA and LIQ to FA persist in the post-COVID model. These findings suggest that most of the previously observed relationships have disappeared during the pandemic, highlighting the significant disruption to existing financial structures caused by COVID-19.

One notable new relationship in Figure 4 is the influence of ROA on DC during the pandemic. This likely reflects firms' strategic adjustments to their depreciation practices in response to declining profitability. For instance, firms experiencing lower profitability may have accelerated depreciation to reduce taxable income or align asset book values with diminished expectations for their utility in an uncertain economic environment.

The contemporaneous relationships illustrated in Figure 2(b) identify two distinct groups of variables—(LIQ, LEV, FA, TA, TobinsQ) and (ROA, Growth, FOR, DC)—that are linked exclusively through COVID-19. This indicates that the two groups are conditionally independent, given the influence of COVID-19 in the same year. The first group comprises key financial variables—LEV, LIQ, TA, FA, and TobinsQ—that provide critical insights into a firm's financial stability and operational efficiency. The edge directions in Table 2 offer strong evidence that LEV and TA influence FA, and moderate evidence that LIQ influences FA. Additionally, while TobinsQ and LIQ are related, the causal direction between the two remains unclear, as the edge direction is approximately 0.5. In the second group, Growth and ROA capture performance and profitability, while DC and FOR influence financial statements and investor perceptions. The edge directions in Table 2 provide moderate evidence that Growth and FOR influence ROA. However, the causal relationship between ROA and DC remains ambiguous, as the edge direction is also close to 0.5.

Finally, Size (the logarithm of total assets) is not connected to any other variables. In this study, all asset-related variables are expressed as ratios relative to total assets, inherently capturing the influence of total assets in their values. The correlation coefficients between Size and other variables are close to zero, except for FOR, which has a correlation coefficient of approximately 0.5 with Size. However, both Size and FOR show strong autocorrelation effects, with autocorrelation coefficients close to 1. The dominant influence of their previous-year values mitigates the connection between Size and FOR in the current year.

6. Conclusions

We have developed a CDBN model to uncover and visualize the temporal and contemporaneous causal relationships among financial variables associated with firm value, using data from listed manufacturing companies in Korea. The CDBN structure revealed strong evidence of autocorrelation across all dynamic variables, as well as a lagged feedback loop between intangible assets (ITA) and TobinsQ (TobinsQ), and a lagged effect of return on assets (ROA) on ITA.

The model also highlighted the extensive impact of the COVID-19 pandemic on the financial sector;

key factors influencing important financial metrics such as FA, TobinsQ, and ROA; and two distinct groups of variables—(LIQ, LEV, FA, ITA, TA, TobinsQ) and (ROA, Growth, FOR, DC)—connected only through COVID-19 and the previous year's ROA. These findings offer a clear understanding of the dynamic and instantaneous interactions between financial variables, shedding light on the complex relationships.

The primary focus of this study was to explore the interrelationships between variables associated with firm value, rather than building a predictive model for firm value. Accordingly, we did not distinguish between dependent (target) and independent (explanatory) variables. Instead, we examined the joint causal dynamics among all variables. This approach offers great flexibility, as any variable can be treated as a target or explanatory variable depending on the context of the analysis. For example, if TobinsQ is the variable of interest, the effects of its ancestor nodes, which directly or indirectly influence TobinsQ (such as COVID-19; previous year's ITA and TobinsQ), can be inferred using the CDBN model. Likewise, if ROA is the focus, the same CDBN model also allows inference on the effects of its ancestor nodes (COVID-19; previous year's ROA; Growth; FOR).

The process of constructing the CDBN model proposed in this study can be tailored for specific inference tasks by selectively blocking or enforcing certain edges within the network. If the goal is to assess the influence of all other variables on a specific variable, that variable can be designated as the target, while the remaining variables serve as explanatory variables. In this setup, the target variable is treated as an end node with no child nodes, and the explanatory variables are considered as potential parent nodes. As an illustration, we constructed a CDBN with TobinsQ as the target variable and all other variables as explanatory variables, while blocking any edges from TobinsQ to other variables in addition to the edge direction constraints described in Section 4. Using the same dataset, the resulting CDBN model is identical to the one shown in Figure 2, except that the direction of the edge between LIQ and TobinsQ is reversed, pointing from TobinsQ to LIQ.

7. Limitations and directions for future research

To the best of our knowledge, the CDBN model is the first to assess both temporal and contemporaneous causal relationships between financial variables related to firm value. However, it is essential to acknowledge several limitations and caveats associated with this research. First, we advise readers against over-generalizing our findings, as the dataset used pertains to the manufacturing sector in Korea. Although the CDBN model proposed here can be applied to datasets from other countries, contexts, or sectors, the resulting causal relationships may vary depending on the financial environments from which these datasets were derived. Second, the inferred causal relationships in this study are valid only under the assumption that all relevant causes are represented within the graph and that no unmeasured confounders exist to explain the relationships between any two observed variables.

Future research could focus on expanding the dataset to include multiple countries and sectors, facilitating a more comprehensive analysis of causal network structures related to firm value and its determinants. Another promising direction is to incorporate a broader range of firm characteristics, as suggested by Green et al. (2017), along with key macroeconomic variables such as interest rates, exchange rates, and unemployment rates.

To assess the predictive performance of the CDBN model relative to a standard regression model for panel data, we used data from 2015 to 2021 to build both models and predicted TobinsQ for 2022. The CDBN model achieved a mean squared prediction error of 0.184, which was significantly lower than the

0.379 of the regression model. Although this is a relatively simple comparison, it highlights the CDBN model's superior predictive accuracy, likely due to its ability to capture essential interactions within the data. A more thorough evaluation of the CDBN model's performance, particularly in comparison with Structural Equation Models and other machine learning methods, is recommended as a valuable direction for future research.

Author contributions

M.O.: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing.
J.C., C.L.: Data Curation, Software, Formal analysis, Writing - Original Draft.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

We are grateful for insightful comments from the Editor and two anonymous reviewers, which helped us improve the paper. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2022R1A2C1091256).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Afrina T, Beg TH, Zayed NM, et al. (2020) An analysis of the effects of corona virus (COVID-19) on international financial derivatives market. *Indian J Financ Bank* 4: 93–98. https://doi.org/10.46281/ijfb.v4i2.757

Ali AM, Anis J (2012) CEO emotional bias and dividend policy: Bayesian network method. *Bus Econ Horizons* 7: 1–18. Available from: https://www.ceeol.com/search/article-detail?id=7652

Campbell JY, Lo AW, MacKinlay AC (1997) The Econometrics of Financial Markets. Princeton University Press. https://doi.org/10.1515/9781400830213

Cao Y, Liu X, Zhai J (2022) A two-stage Bayesian network model for corporate bankruptcy prediction. *Int J Financ Econ* 27: 455–472. https://doi.org/10.1002/ijfe.2162

Çambaşı H, Kuru Ö, Amasyalı MF, et al. (2019) Comparison of Dynamic Bayesian Network Tools. *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)*: 1–6. https://doi.org/10.1109/ASYU48272.2019.8946390

Chan LS, Chu AM, So MK (2023) A moving-window bayesian network model for assessing systemic risk in financial markets. *PloS one* 18. https://doi.org/10.1371/journal.pone.0279888

Chang J, Bai Y, Xue J, et al. (2023) Dynamic Bayesian networks with application in environmental modeling and management: A review. *Environ Modell Softw*. https://doi.org/10.1016/j.envsoft.2023.105835

Chen W, Srinivasan S (2024) Going Digital: Implications for Firm Value and Performance. *Rev Account Stud* 29: 1619–1665. Available from: https://link.springer.com/article/10.1007/s11142-023-09753-0

Cheng YS, Liu YP, Chien CY (2010) Capital structure and firm value in China: A panel threshold regression analysis. *S Afr J Bus Manag* 4: 2500–2507. Available from: https://academicjournals.org/journal/AJBM/article-stat/B20594B27420

Colombo D, Maathuis MH (2014) Order-Independent Constraint-Based Causal Structure Learning. *J. Mach. Learn. Res* 15:: 3741-3782. Available from: https://jmlr.org/papers/volume15/colombo14a/colombo14a.pdf.

Core JE, Guay WR, Van Buskirk A (2003) Market valuations in the new economy: An investigation of what has changed. *J Account Econ* 34: 43–67. https://doi.org/10.1016/S0165-4101(02)00087-3

Dagum P, Galper A, Horvitz E (1992) Dynamic network models for forecasting. *Uncertainty Artif Intell*: 41–48. https://doi.org/10.1016/B978-1-4832-8287-9.50010-4

Dwicahyani D, Rate PV, Jan ABH (2022) The Effect of Leverage, Profitability, Company Size, Managerial Ownership And Institutional Ownership On The Value Of Non-Cyclicals. *J EBMA* 10: 275–286. Available from: https://ejournal.unsrat.ac.id/v3/index.php/emba/article/view/43790

Enders W (2015) *Applied econometric time series fourth edition*. John Wiley & Sons, New York, 30.

Francis J, Schipper K (1999) Have financial statements lost their relevance? *J Account Res* 37: 319–352. https://doi.org/10.2307/2491412

Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. *Mach Learn* 29: 131–163. https://doi.org/10.1023/A:1007465528199

Friedman N, Goldszmidt M, Wyner A (1999) Data Analysis with Bayesian Networks: A Bootstrap Approach. *In Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence*, 196-205. https://doi.org/10.48550/arXiv.1301.6695

Glova J, Mrázková S (2018) Impact of intangibles on firm value: An empirical evidence from European public companies. *Ekonomický časopis* 66: 665–680. Available from: https://api.semanticscholar.org/CorpusID:210916344

Green J, Hand JRM, Zhang XF (2017) The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. *Rev Financ Stud* 30: 4389–4436. Available from: https://academic.oup.com/rfs/article-abstract/30/12/4389/3091648

Hertati L, Widiyanti M, Desfitrina D, et al. (2020) The effects of economic crisis on business finance. *Int J Econ Financ Issues* 10: 236–244. https://doi.org/10.32479/ijefi.9928

Hirdinis M (2019) Capital structure and firm size on firm value moderated by profitability. *Int J Econo Bus Admin*. Available from: https://www.um.edu.mt/library/oar/handle/123456789/43966

Ispriyahadi H, Abdulah B (2021) Analysis of The Effect of Profitability, Leverage and Firm Size on Firm Value. *J Bus Manage Account* 3: 64–80. Available from: https://www.neliti.com/publications/432230/analysis-of-the-effect-of-profitability-leverage-and-firm-size-on-firm-value

Jo M, Oh R, Oh MS (2023) Prediction of PM10 concentration in Seoul, Korea using Bayesian network. *Commun Stat Appl Met* 30: 517–530. https://doi.org/10.29220/CSAM.2023.30.5.517

Kalisch M, Hauser A, Maechler M, et al. (2024) Package 'pcalg'. Available from: https://cran.r-project.org/web/packages/pcalg/pcalg.pdf

Koski T, Noble J (2011) *Bayesian networks: an introduction.* John Wiley & Sons. https://doi.org/10.1002/9780470684023

Lee D, Kwon K (2023) Dynamic Bayesian network model for comprehensive risk analysis of fatigue-critical structural details. *Reliab Eng Syst Safe* 229. https://doi.org/10.1016/j.ress.2022.108834

Li Z, Yang J, Tan S (2013) Systematically discovering dependence structure of global stock markets using dynamic Bayesian network. *J Computat Inf Syst* 9: 7215–7226. Available from: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=dcbc4c3bca5e71f8ed1d645db5c5e76ff6bba0db

Liu H, Motoda H (2012) *Feature selection for knowledge discovery and data mining.* Springer science & business media. https://doi.org/10.1007/978-1-4615-5689-3

Madden MG (2009) On the classification performance of TAN and general Bayesian networks. *Knowl-Based Syst* 22: 489–495. https://doi.org/10.1016/j.knosys.2008.10.006

McNichols M, Rajan M, Reichelstein S (2014) Conservatism correction for the market-to-book ratio and tobin's q. *Rev Account Stud* 19: 1393–1435. https://doi.org/10.1007/s11142-013-9275-2

Mishra PK, Mishra SK (2020) Corona pandemic and stock market behaviour: Empirical insights from selected Asian countries. *Millenn Asia* 11: 341–365. https://doi.org/10.1177/0976399620952354

Mohammed ZO, Al Ani MK. (2020) The effect of intangible assets, financial performance and financial policies on the firm value: Evidence from Omani industrial sector. *Contemp Econ* 2020: 379–391. https://doi.org/10.5709/ce.1897-9254.411

Nagaraja N, Vinay N (2016) The effect of intangible assets on the firm value. *Int J Eng Manag Res* 6: 307–315. Available from: https://www.indianjournals.com/ijor.aspx?target=ijor:ijemr&volume=6&issue=1&article=053

Ocak M, Fındık D (2019) The impact of intangible assets and sub-components of intangible assets on sustainable growth and firm value: evidence from Turkish listed firms. *Sustainability* 11: 5359. https://doi.org/10.3390/su11195359

Oh R, Lee HK, Pak YK, et al. (2022) An interactive online app for predicting diabetes via machine learning from environment-polluting chemical exposure data. *Int J Environ Res Public Health* 19: 5800. https://doi.org/10.3390/ijerph19105800

Pástor Ľ, Pietro V (2003) Stock valuation and learning about profitability. *J Financ* 58: 1749–1789. https://doi.org/10.1111/1540-6261.00587

Pearl J (1988) Probabilistic Reasoning in Intelligent Systems; Network of Plausible Inference. *Morgan Kaufmann*. Available from: https://www.sciencedirect.com/book/9780080514895/probabilistic-reasoning-in-intelligent-systems

Rountree B, Weston JP, Allayannis G (2008) Do investors value smooth performance? *J Financ Econ* 90: 237–251. https://doi.org/10.1016/j.jfineco.2008.02.002

Scutari M (2023) bnlearn: Bayesian Network Structure Learning, Parameter Learning and Inference. *R package version 4.8*. Available from: https://CRAN.R-project.org/package=bnlearn

Scutari M, Denis JB (2021) Bayesian networks: with examples in R. *Chapman and Hall/CRC*. https://doi.org/10.1201/9780429347436

Scutari M, Kerob D, Salah S (2024) Inferring skin–brain–skin connections from infodemiology data using dynamic Bayesian networks. *Sci Rep* 14: 10266. https://doi.org/10.1038/s41598-024-60937-3

Scutari M, Nagarajan R (2013) Identifying significant edges in graphical models of molecular networks. *Artif Intell Med* 57: 207–217. https://doi.org/10.1016/j.artmed.2012.12.006

Siregar SD, Toni N, Ariesa Y (2023) Impact of dividend policy, capital structure, and profitability on consumer goods firm value: Role of firm size (2013–2022). *J Econ Bus Lett* 3: 38–48. https://doi.org/10.55942/jebl.v3i4.234

Spirtes P, Glymour C (1991) An algorithm for fast recovery of sparse causal graphs. *Soc Sci Comput Rev* 9: 62–72. https://doi.org/10.1177/089443939100900106

Spirtes P, Glymour C, Scheines R, et al. (2000) Constructing Bayesian network models of gene expression networks from microarray data. *Carnegie Mellon University*. https://doi.org/10.1184/R1/6491291.v1

Sun EJ (2015) Bayesian Network Analysis for the Dynamic Prediction of Financial Performance Using Corporate Social Responsibility Activities. *Manageme Inf Syst Rev* 34: 71–92. https://doi.org/10.29214/damis.2015.34.5.004

Sun EJ, Park SJ (2017) The relationship between chaebol and firm value using Bayesian network. *J Appl Econ Bus Res* 33. Available from: https://journals.klalliance.org/index.php/JABR/article/view/359

Sunarsih NM, Dewi NPS, Kireina MNNA (2019) Analysis of factors effecting the firm value factors that effect the firm value. *Int J Appl Bus Int Manage* 4: 94–103. Available from: https://ejournal.aibpmjournals.com/index.php/IJABIM/article/view/687

Suzan L, Ramadhani NI (2023) Firm Value Factors: The Effect Of Intellectual Capital, Managerial Ownership, And Profitability. *J Akuntansi* 27: 401–420. https://doi.org/10.24912/ja.v27i3.1487

Teles G, Rodrigues JJPC, Rabê RA, et al. (2020) Artificial neural network and Bayesian network models for credit risk prediction. *J Artif Intell Syst* 2: 118–132. https://doi.org/10.33969/AIS.2020.21008

Tobin J (1969) A general equilibrium approach to monetary theory. *J Money Credit Bank* 1: 15–29. https://doi.org/10.2307/1991374

Tsai CF, Lu YH, Yen DC (2012) Determinants of intangible assets value: The data mining approach. *Knowl-Based Syst* 31: 67–77. https://doi.org/10.1016/j.knosys.2012.02.007

Tsay RS (2010) *Analysis of Financial Time Series*. John Wiley & Sons.