

Trading with Time Series Causal Discovery: An Empirical Study

Ruijie Tang
 Business School, Imperial College London
 ruijie.tang23@imperial.ac.uk

August 30, 2024

Abstract

This study investigates the application of causal discovery algorithms in equity markets, with a focus on their potential to build investment strategies. An investment strategy was developed based on the causal structures identified by these algorithms. The performance of the strategy is evaluated based on the profitability and effectiveness in stock markets. The results indicate that causal discovery algorithms can successfully uncover actionable causal relationships in large markets, leading to profitable investment outcomes. However, the research also identifies a critical challenge: the computational complexity and scalability of these algorithms when dealing with large datasets. This challenge presents practical limitations for their application in real-world market analysis.

1 Introduction

Numerous causal discovery algorithms have been developed for time series analysis [2], and some have been applied to the stock market to identify underlying driving forces [11]. However, few studies have advanced to the next step of using these driving forces in quantitative studies to predict future stock prices and develop corresponding trading strategies. This paper aims to narrow this gap by designing causality-based trading strategies and evaluating their feasibility and effectiveness.

The major contributions of this paper include the following:

1. Applying time series causal discovery to real-world stock data
2. Developing a workflow to turn causal relations into a trading strategy
3. Backtesting the strategy on major stock markets in China and the US.

The Python and R code developed in this research, including data pre-processing scripts, accuracy evaluation tools, causal discovery interfaces, a causal graph reformatting tool, auto-regression fitting and prediction routines, trading simulators, backtesting utilities, trading performance visualizers, and execution time profiling tools, is open-sourced¹.

In this study, we mainly focus on three key questions:

- Q1 Can we quantitatively evaluate the effectiveness of a causal discovery algorithm using stock price data?
- Q2 If effectiveness can be quantified, which causal discovery algorithm performs best in driving force analysis for the market?
- Q3 Are there practical challenges when applying causal discovery to analyze stock markets, even if the methods are theoretically sound?

¹GitHub repository: <https://github.com/kaaaylaaa/causality-based-trading>

Table 1: Algorithm Overview

Algorithm	Reference	Type	Adaptation
tsFCI	[5]	Constraint-Based	Causal graph reformatting
VarLiNGAM	[13]	Noise-Based	Causal graph reformatting
TiMINo	[18]	Noise-Based	Edge direction determination

The remainder of this paper is organized as follows: Section 2 provides a brief introduction to the selected causal discovery algorithms, namely tsFCI, VarLiNGAM, and TiMINo. Section 3 outlines the detailed procedures of our workflow. Section 4 presents the experimental results along with their analysis. Finally, Section 5 offers conclusions and recommendations for future research, and answers to the research questions.

2 Background

In the traditional Granger causal inference framework, a time series Y causes time series X if, given the past of all other time series, knowing the past of Y helps in predicting future values of X [8]. Although Granger causality is a widely used tool for causal inference in time series, it is limited to pairwise relationships and can be misleading in the presence of hidden common causes [19]. Time series causal discovery algorithms offer a more robust approach by addressing these limitations and providing insights into the complex causal structures within time series data [2].

A time series causal discovery algorithm identifies potential causal relationships within a multivariate time series and represents these relationships as a directed causal graph. In this study, we focus on the “summary graph” approach [2], where each variable in the time series corresponds to a single node in the causal graph. Previous research has applied time series causal discovery techniques to uncover driving forces [11, 12]. We aim to extend this approach by predicting future stock prices based on these driving forces. With these predictions, it becomes possible to manage an investment portfolio with various trading strategies.

Despite the numerous causal discovery approaches for time series, none excel in all situations due to the variety in data distributions. In this study, we focus on three algorithms: tsFCI, VarLiNGAM, and TiMINo, as previous research suggests they are superior to other approaches when dealing with various data structures [2].

- tsFCI (time series Fast Causal Inference) [5] is based on the FCI algorithm [21]. The FCI algorithm first constructs an undirected full causal graph, removes edges using independence tests, and then orients the remaining edges according to a series of rules. This method allows for the handling of non-temporal data with hidden confounders. An extension of FCI for time series is tsFCI. Suppose we have a multivariate time series X with N observed variables and T time points. The time lag is defined as τ , and the system may contain potential hidden variables. tsFCI starts by expanding X to $(T - \tau)$ rows and $(\tau + 1)N$ columns using the sliding window approach. Then, the FCI algorithm is applied to find the causal structure by treating each component of X as an individual random variable.
- VarLiNGAM (Vector Autoregressive Linear Non-Gaussian Acyclic Model) [13] builds upon the LiNGAM algorithm [20]. The LiNGAM algorithm assumes that each variable is a linear function of its causes plus an error term e_i , which follows a non-Gaussian distribution with non-zero variance. Additionally, the error terms e_i are independent across variables, and the system is assumed to contain no latent confounders. LiNGAM begins with a structural equation model (SEM) of the form $X = \mathbf{B}X + \mathbf{e}$ where only non-temporal data is allowed, and \mathbf{B} can be permuted to a strictly lower triangular matrix with zeros on the diagonal. Solving for X yields $X = \mathbf{A}\mathbf{e}$ where $\mathbf{A} = (\mathbf{I} - \mathbf{B})^{-1}$. However, this estimation needs proper permutation and scaling before it can be used to derive \mathbf{B} . Finally, the estimated \mathbf{B} is permuted to strict lower triangularity, which contains the causal order. VarLiNGAM extract the causal structure by computing \mathbf{B}_i in $X(t) = \sum_{i=0}^{\tau} \mathbf{B}_i X(t-i) + \mathbf{e}(t)$ using LiNGAM.

Table 2: Data Overview

Dataset	Pelosi	CSI300	SP500
Market	US	US	China
No. stocks	12	98	446
Start date	30-Jul-2019	1-Sep-2009	1-Sep-2009
End date	30-Jul-2024	31-Dec-2019	31-Dec-2019
No. trading days	1259	2513	2604
No. trading days for causal discovery	1007	2010	2083
No. trading days for backtesting	252	503	521

- TiMINo (Time-varying Interactions Model for Nonlinear Observations) is another SEM-based causal discovery algorithm, which considers both nonlinear and instantaneous effects [18]. A time series X with absolutely continuous finite dimensional distributions w.r.t a product measure satisfies a TiMINo if $\exists \tau > 0$ and $\forall X \in V$, there are sets $\mathbf{PA}_0^X \subseteq V \setminus \{X\}$, $\mathbf{PA}_i^X \subseteq V$ for $1 \leq i \leq \tau$, such that for all t , $X_t = f_X((\mathbf{PA}_\tau^X)_{t-\tau}, \dots, (\mathbf{PA}_1^X)_{t-1}, (\mathbf{PA}_0^X)_t, \xi_t^X)$ where V is the set of variables in X , and \mathbf{PA}^X represents the parent nodes of X in the causal graph. ξ_t^X are required to be jointly independent over both i and t , and *i.i.d.* in t for each X . TiMINo aims to compute \mathbf{PA}_i^X for each node X based on the definition aforementioned.

3 Proposed Approach

Given a target set of stocks, this empirical study involves work in three aspects: (i) data collection and processing; (ii) causal discovery, predictions and trading; and (iii) backtesting.

3.1 Data Preparation

We selected China and the United States as two major stock markets to evaluate the effectiveness of the causal discovery algorithms. The data, sourced from [15], covers a period of 10 years (2009.09.01 - 2019.12.31) and includes prices for SP500 and CSI300 component stocks. Another dataset that captured our interest, despite its relatively small size, holds significant potential for profitability. This dataset pertains to the stock portfolio of Nancy Pelosi, who is renowned for her astute investment insights. Notably, an ETF named ‘NANC’ tracks her stock trades. We collected stock prices from her current portfolio, comprising 12 stocks, over the past five years from Yahoo Finance ².

We performed data imputation due to the substantial amount of missing values in the datasets. The imputation process involved two steps. First, linear interpolation was applied across all stocks to estimate and fill gaps where missing values occurred between known values. Second, any stock that still contained missing values after interpolation was removed from the dataset. As a result, the cleaned SP500 dataset comprises 446 stocks, while the cleaned CSI300 dataset comprises 98 stocks. The Pelosi dataset contains no missing values, so no stocks were removed. An overview of our data is in Table 2.

3.2 Causal Discovery, Prediction and Trading

The three causal discovery algorithms discussed in Section 2, namely tsFCI, VarLiNGAM and TiMINo, are employed to extract the causal structure from each dataset. Our primary focus is on identifying the causal relationships among the time series rather than the exact causal time lags. Therefore, the causal graphs generated by tsFCI and VarLiNGAM are compressed by removing the time lag attributes for subsequent analysis.

The process of using causal discovery techniques to identify driving forces for each stock is illustrated in Figure 1. From the causal graph, we can extract the set of parent nodes of stock X , denoted by \mathbf{PA}^X , which can be interpreted as the driving forces of X [11]. Then, similar to

²Yahoo Finance: <https://finance.yahoo.com/>

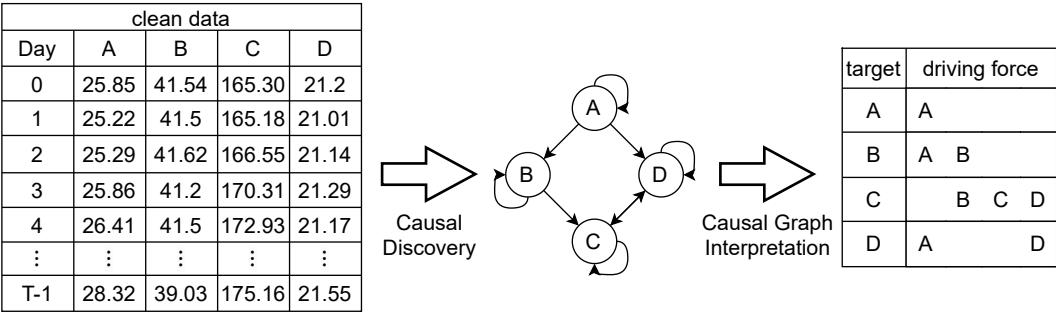


Figure 1: Use causal discovery to find driving forces

the multivariate time series model described by [22], we fit a predictive model to each stock using \mathbf{PA}^X as independent variables. The model is formulated as follows:

$$P_t^X = f_X(\mathbf{PA}_{t-1}^X, \dots, \mathbf{PA}_{t-\tau}^X) \quad (1)$$

where τ is the time lag used during causal discovery. The latest time point we use for prediction is $t - 1$ instead of t . In other words, the current value is modelled by the changes in the driving forces at previous time steps, while the present does not produce an instantaneous effect. An expanding window is employed when fitting the predictive model. The daily prediction and trading process is outlined in Figure 2. For each stock, we extract the historical values of its driving forces, fit a linear regression model on these past prices, and then make a one-step ahead prediction. Using the predicted prices, we calculate the predicted returns for the next day as follows:

$$\gamma_{t,t+1}^X = \frac{\rho_{t+1}^X - P_t^X}{P_t^X} \quad (2)$$

where ρ_{t+1}^X is the predicted price of stock X for the next day, and P_t^X is the actual stock price today. Finally, we aggregate the predictions of all stocks in the asset class for the next day and proceed with the corresponding trading actions.

Given price predictions, numerous studies have demonstrated that a long-short momentum strategy can generate significant profits [14, 23, 16, 17]. This strategy operates on the premise that stocks with strong past performance are likely to continue performing well, while those with poor past performance are expected to underperform. The cross-sectional long-short momentum strategy introduced by [14] constructs a zero-cost market-neutral portfolio by buying stocks with high past returns (winners) and selling those with low past returns (losers).

Based on this approach and integrating short-term forecasting, our trading strategy employs a dollar-neutral portfolio, ensuring that the dollar amounts of long and short positions are balanced [4, 1]. For simplicity, we assume a liquid market where short selling is permitted without incurring borrowing costs. However, given that our strategy operates on a daily basis, transaction costs cannot be ignored as they can significantly impact portfolio returns. Therefore, we impose a fixed daily transaction cost of 0.1% [7]. The strategy is implemented as follows:

First, we select an integer η as the number of winner stocks. Then, using the one-day-ahead predictions obtained through regression, we identify the η stocks with the highest (lowest) predicted returns, classifying them as winners (losers). Finally, at the end of the current trading day, we buy the winners and sell the losers. This approach allows us to construct a zero-cost, dollar-neutral portfolio with equally weighted stocks on both the long and short sides. All trading positions are closed out at the end of each day before the next trading action takes place.

3.3 Back-testing

At this stage, we use historical data to evaluate the performance of the trading strategy. The procedure follows the steps outlined in the previous two sections. To avoid look-ahead bias, we use the earliest 80% of the dataset as the training set and reserve the remaining 20% as the test set,

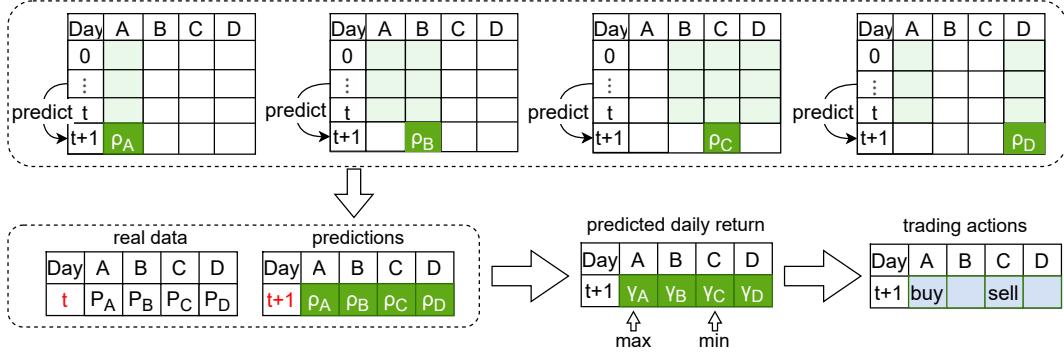


Figure 2: 1-day ahead predictions and trading actions

as detailed in Table 2. In each trading day, we run the trading strategy and compute the outcome according to the real data. In particular, the realized portfolio daily return is calculated as follows:

$$r_{t,t+1}^p = \frac{r_{t,t+1}^{L_1} + \dots + r_{t,t+1}^{L_\eta}}{\eta} - \frac{r_{t,t+1}^{S_1} + \dots + r_{t,t+1}^{S_\eta}}{\eta} - C \quad (3)$$

where C is the transaction cost; r^L, r^S represent the realized returns of the long and short positions, respectively. We use the expanding window method to continue with one-step ahead forecasts until the end of test period. Finally, the annualized portfolio return is calculated as follows [4]:

$$r_{annual}^p = (1 + r_{T_{test}}^p)^{\frac{D}{T_{test}}} - 1 \quad (4)$$

where $r_{T_{test}}^p$ represents the cumulative return over the test period, T_{test} is the length of the test period, and $D = 252$ is the number of trading days in a year.

4 Evaluation

4.1 Raw Results

Before applying the causal discovery algorithms to analyze the stock market, we validate the correctness of our setup using datasets with ground truth causal graphs, achieving the similar model accuracy scores to those reported in [2].

Table 3: Raw Result Placement

	Pelosi	CSI300	SP500
tsFCI (lags 1–2)	Figure 4	Cannot finish in 24h	Cannot finish in 24h
tsFCI (lags 3–6)	Appendix	Cannot finish in 24h	Cannot finish in 24h
TiMINo (lags 1–2)	Figure 4	Cannot finish in 24h	Cannot finish in 24h
TiMINo (lags 3–6)	Appendix	Cannot finish in 24h	Cannot finish in 24h
VarLiNGAM (lags 1–2)	Figure 4	Figure 3	Figure 3
VarLiNGAM (lag 3)	Appendix	Figure 3	Appendix
VarLiNGAM (lag 4)	Appendix	Figure 3	Memory limit exceeded
VarLiNGAM (lags 5–6)	Appendix	Appendix	Memory limit exceeded

We record the performance for a test case if the algorithm is able to finish execution within 24 hours on an Apple MacBook Air computer with Apple M2 chips and 16GB of memory. Table 3 provides a summary of our tests. Both tsFCI and TiMINo were unable to process the CSI300 and SP500 datasets due to excessive execution time. For the successful runs, we aim to test lags from 1 to 6. However, while processing the SP500 dataset with VarLiNGAM using time lags from 4 to 6, we encountered a memory limit issue. The results for VarLiNGAM on the CSI300 (lags 1 to 4)

and SP500 (lags 1 and 2) datasets are shown in Figure 3. The results for all three algorithms with the first two lags for the Pelosi dataset are shown in Figure 4. The remaining graphs are provided in the Appendix.

For ease of comparison, all returns were standardized to annualized returns. Each asset class was benchmarked against its respective index or ETF to evaluate portfolio performance. To assess the validity and necessity of causal discovery, a control portfolio was constructed using predictions based solely on self-causality.

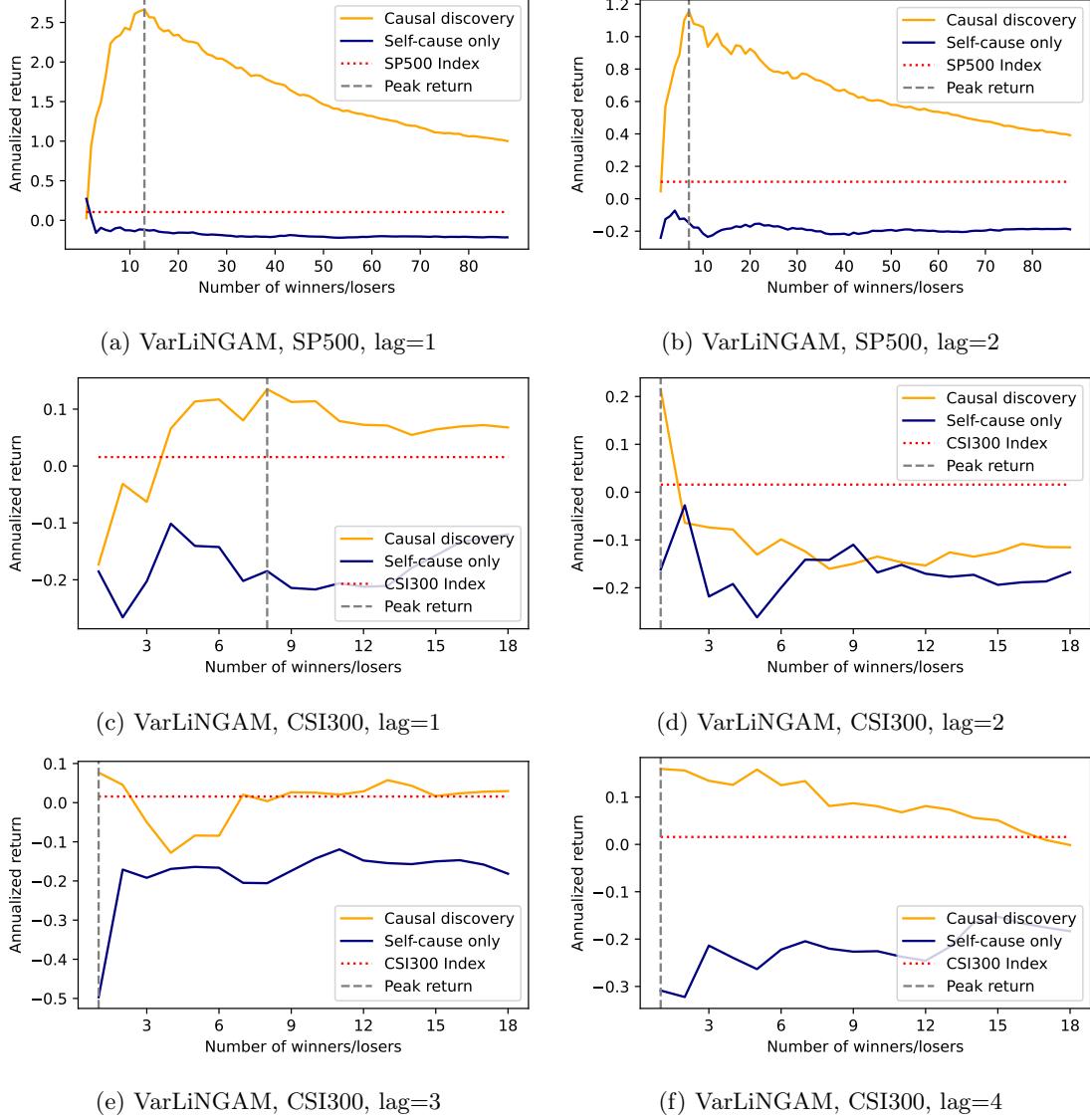


Figure 3: SP500 & CSI300 portfolio performance using VarLiNGAM
(Discussion available in Section 4.2)

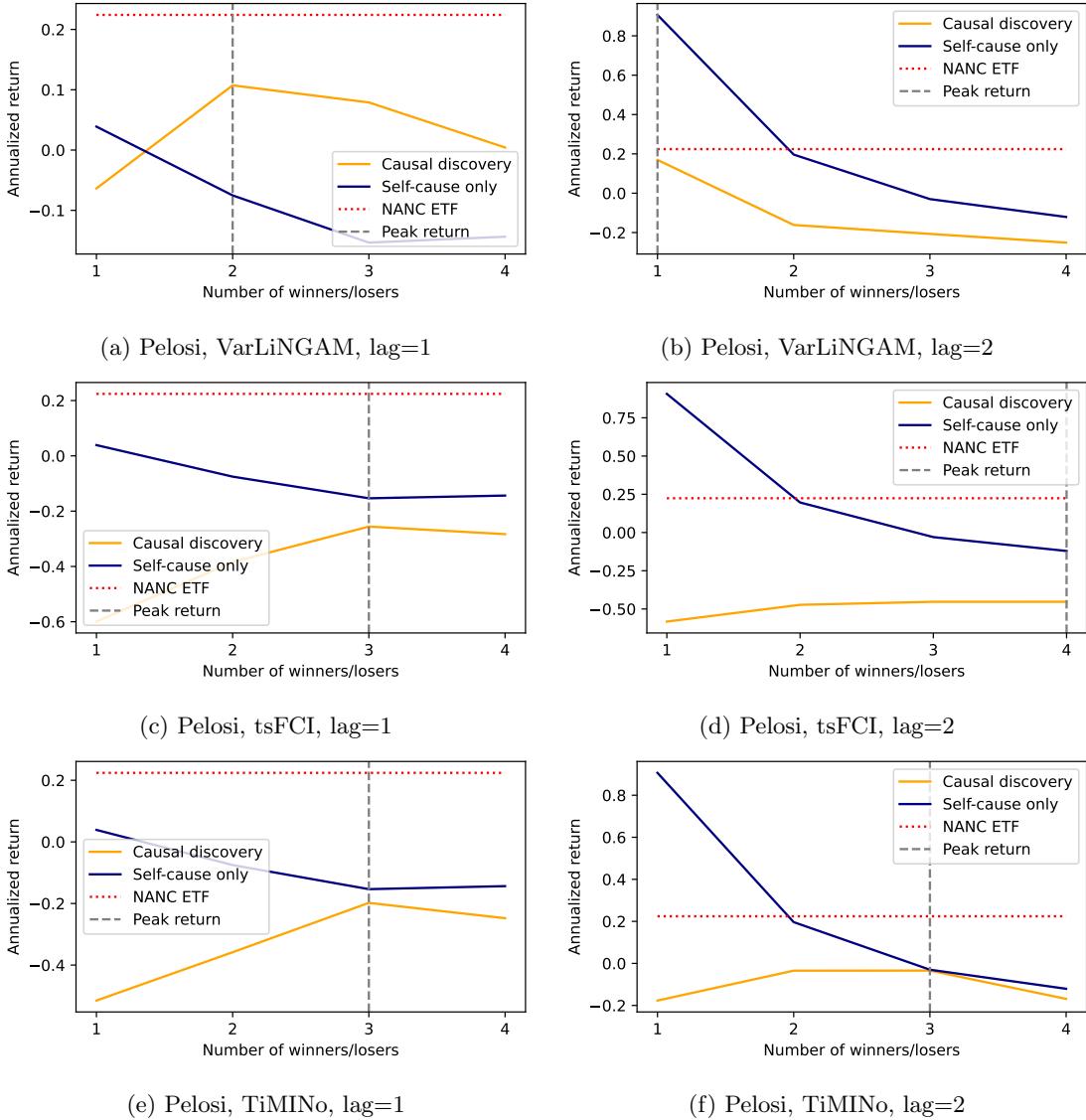


Figure 4: Pelosi portfolio performance using VarLiNGAM, tsFCI and TiMINo
(Discussion available in Section 4.2)

4.2 Discussion

Our results indicate that incorporating a causal discovery phase generally improves portfolio returns, especially with larger datasets. Among the three algorithms, VarLiNGAM is the most effective, handling large datasets well and generating the highest returns on all datasets. Major observations we obtain from the results include the following:

- 1. Causal discovery proves more effective than the self-cause-only setting.** The portfolio supported by causal discovery significantly outperforms the self-cause-only portfolio. Conversely, the self-cause-only portfolio performs poorly, providing worse performance even than the baseline index returns. In other words, relying solely on the stock's own historical data for price prediction is sub-optimal. This observation is particularly clear in Figure 3 with a large number of stocks.
- 2. Trading with causal discovery tends to be more profitable in large markets than small markets.** Comparing Figure 3 and 4, we find that the trading strategy based on causal discovery becomes more effective with the markets with more choices. A small dataset like

Pelosi hinders the causal discovery process by including too few stocks, likely excluding the true driving forces from the analysis.

3. **A smaller time lag yields higher returns.** The results from all three datasets indicate that a time lag of 1 or 2 yields the most favorable portfolio returns, with portfolio returns generally decreasing as the time lag increases. This outcome may be attributed to the increase in the number of independent variables considered in the predictive model as the time lag grows. Such an increase can lead to potential issues of overfitting or expose the limitations of the simple linear regression model chosen.
4. **The optimal number of winners ranges from 1% to 6% of the total stocks in the dataset.** In the SP500 market, the portfolio returns display a humped pattern across various time lags. As the number of winners increases, the portfolio return initially rises, reaching a peak due to diversification, which enhances returns by reducing risk. However, beyond this point, further increasing the number of winners dilutes returns, as the portfolio becomes overly diversified with too many different stocks. In contrast, analyzing the number of winners for the Pelosi dataset seems less relevant, given the poor performance of the causal discovery portfolio.
5. **VarLiNGAM achieves the highest returns among the three algorithms.** VarLiNGAM is the only method that can finish computation within 24 hours for SP500 and CSI300. Also, in the Pelosi dataset where all three methods provide valid results, VarLiNGAM still has minor advantages over tsFCI and TiMINo, especially when the lag is 1.

5 Conclusion

In conclusion, among the three tested causal discovery algorithms, VarLiNGAM is the only feasible option for analyzing the stock market with 400–500 stocks, while tsFCI and TiMINo are constrained to processing much smaller datasets. For a large collection of stocks like SP500 or CSI300 constituents, causal discovery is helpful for improving prediction accuracy and enhancing trading performance. The overall portfolio performance in the SP500 constituents surpasses that of CSI300, which is expected given the bullish overall performance of the SP500 Index. On the other hand, causal discovery proves ineffective when applied to smaller asset classes like Pelosi, where the true driving forces of its component stocks fall outside the analysis scope. Regardless of the market size, VarLiNGAM still outperforms tsFCI and TiMINo in all analysed cases, making it the best algorithm among the three. In terms of trading, it is advisable to select a smaller time lag. Specifically, for our trading strategy, setting the number of winner stocks to 1%-6% of the size of the original asset class yields the best returns. Future work of this study includes the following directions:

- F1 We explored the feasibility of developing trading strategies based on causal discovery techniques. Among the three algorithms examined, only VarLiNGAM demonstrated the capability to efficiently process large datasets. Consequently, a direction for future research is to improve the computationally efficiency and scalability of time series causal discovery techniques with respect to the data size.
- F2 Once the causal structure is established, predictions can be made based on the identified causal relationships. This study uses a simple linear regression model for the predictions, which is known to have limited predictive power. Future research should consider more advanced predictive models like Long short-term memory (LSTM) models [6, 3].

We finish this paper by answering the research questions specified in Section 1:

- A1 It is feasible to quantitatively analyze the effectiveness of a causal discovery algorithm using stock price data by constructing a trading portfolio based on the causal graph and then tracking the portfolio's returns over time.

- A2 VarLiNGAM emerged as the most effective algorithm among the three selected methods for two key reasons. First, it is capable of processing large datasets, which the other algorithms cannot handle. Second, the trading portfolio constructed using VarLiNGAM outperforms those based on the other two algorithms, generating the highest returns on the smaller Pelosi dataset. Further testing of additional algorithms is required to confirm these findings.
- A3 The primary challenge in applying causal discovery to the stock market lies in computational limitations. Among the three algorithms considered, only VarLiNGAM is capable of analyzing large datasets such as CSI300 and SP500 within 24 hours, while tsFCI and TiMINo struggle to process such extensive equity markets. Similar computational bottlenecks have been recognized and addressed in non-temporal causal discovery research [9, 10]. However, faster and more scalable time series causal discovery techniques should be developed to handle data from large markets within limited time and resources, as we discussed in Future Work F1.

References

- [1] Clifford S Asness, Tobias J Moskowitz, and Lasse Heje Pedersen. Value and momentum everywhere. *The journal of finance*, 68(3):929–985, 2013.
- [2] Charles K Assaad, Emilie Devijver, and Eric Gaussier. Survey and evaluation of causal discovery methods for time series. *Journal of Artificial Intelligence Research*, 73:767–819, 2022.
- [3] Michel Ballings, Dirk Van den Poel, Nathalie Hespeels, and Ruben Gryp. Evaluating multiple classifiers for stock price direction prediction. *Expert systems with Applications*, 42(20):7046–7056, 2015.
- [4] Ludwig B Chincarini and Daehwan Kim. *Quantitative equity portfolio management: An active approach to portfolio construction and management*. McGraw-Hill, 2006.
- [5] Doris Entner and Patrik O Hoyer. On causal discovery from time series data using fci. *Probabilistic graphical models*, 16, 2010.
- [6] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, 270(2):654–669, 2018.
- [7] Andrea Frazzini, Ronen Israel, and Tobias J Moskowitz. Trading costs. *SSRN 3229719*, 2018.
- [8] Clive WJ Granger. Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and control*, 2:329–352, 1980.
- [9] Ce Guo and Wayne Luk. Accelerating constraint-based causal discovery by shifting speed bottleneck. In *Proceedings of the 2022 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, pages 169–179, 2022.
- [10] Christopher Hagedorn, Constantin Lange, Johannes Huegle, and Rainer Schlosser. Gpu acceleration for information-theoretic constraint-based causal discovery. In *The KDD’22 Workshop on Causal Discovery*, pages 30–60. PMLR, 2022.
- [11] Biwei Huang, Kun Zhang, Jiji Zhang, Joseph Ramsey, Ruben Sanchez-Romero, Clark Glymour, and Bernhard Schölkopf. Causal discovery from heterogeneous/nonstationary data. *Journal of Machine Learning Research*, 21(89):1–53, 2020.
- [12] Biwei Huang, Kun Zhang, Jiji Zhang, Ruben Sanchez-Romero, Clark Glymour, and Bernhard Schölkopf. Behind distribution shift: Mining driving forces of changes and causal arrows. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 913–918. IEEE, 2017.

- [13] Aapo Hyvärinen, Kun Zhang, Shohei Shimizu, and Patrik O Hoyer. Estimation of a structural vector autoregression model using non-gaussianity. *Journal of Machine Learning Research*, 11(5), 2010.
- [14] Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993.
- [15] Yongli Li, Tianchen Wang, Baiqing Sun, and Chao Liu. Detecting the lead-lag effect in stock markets: definition, patterns, and investment strategies. *Financial Innovation*, 8(1):51, 2022.
- [16] Tobias J Moskowitz and Mark Grinblatt. Do industries explain momentum? *The Journal of finance*, 54(4):1249–1290, 1999.
- [17] Tobias J Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. Time series momentum. *Journal of financial economics*, 104(2):228–250, 2012.
- [18] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. Causal inference on time series using restricted structural equation models. *Advances in neural information processing systems*, 26, 2013.
- [19] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. The MIT Press, 2017.
- [20] Shohei Shimizu, Patrik O Hoyer, Aapo Hyvärinen, Antti Kerminen, and Michael Jordan. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(10), 2006.
- [21] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2001.
- [22] Raja Velu. *Algorithmic trading and quantitative strategies*. Chapman and Hall/CRC, 2020.
- [23] Tobias Wiest. Momentum: what do we know 30 years after jegadeesh and titman’s seminal paper? *Financial Markets and Portfolio Management*, 37(1):95–114, 2023.

Appendix

The following graphs are discussed in Section 4.2.

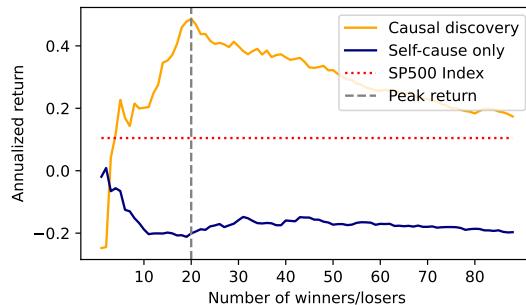


Figure 5: SP500 portfolio performance using VarLiNGAM, lag=3

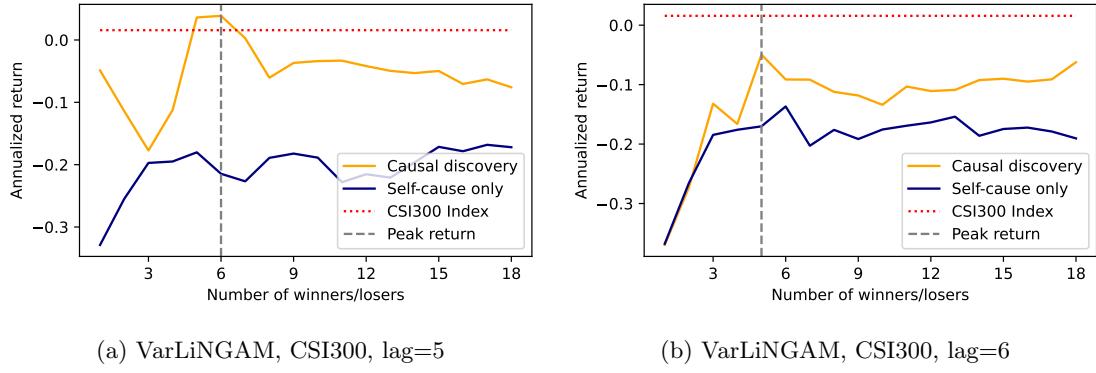


Figure 6: CSI300 portfolio performance using VarLiNGAM

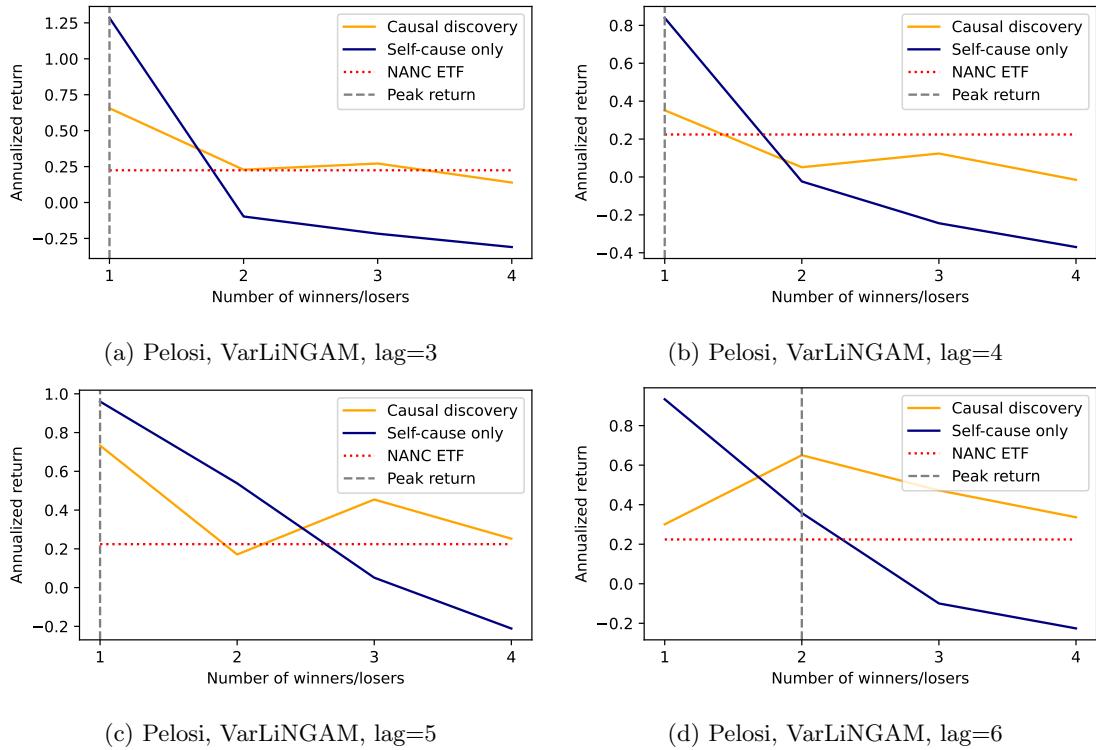


Figure 7: Pelosi portfolio performance using VarLiNGAM

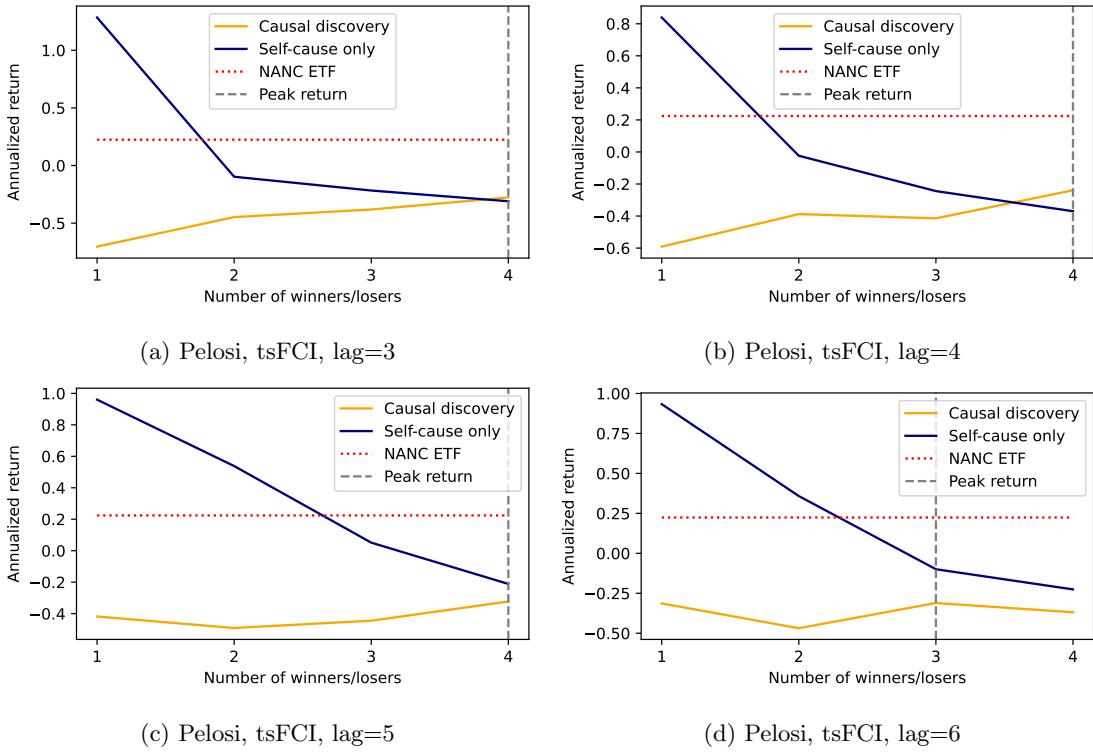


Figure 8: Pelosi portfolio performance using tsFCI

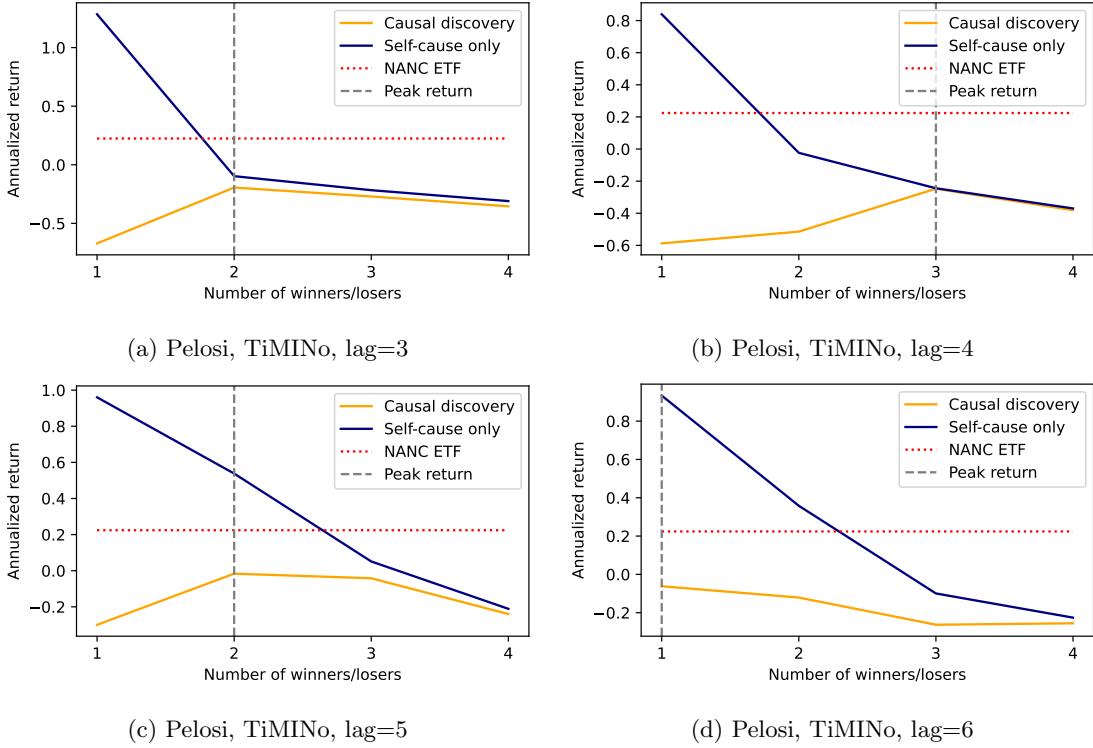


Figure 9: Pelosi portfolio performance using TiMINo