Causal Discovery of Volatility Spillovers in Global Markets via Interpretable Machine Learning with Walk-Forward Robustness Guarantees

Zachary J. Wood
School of Computing, Clemson University, Clemson, SC, USA.
zacharywood667@clemson.edu

Sanjay Venkataraman
Department of Computer Science and Engineering, University of Nevada, Reno, Reno, NV, USA.
sanjay.venkataraman900@unr.edu

Pankaj R. Saini
Department of Computer Science, Colorado State University, Fort Collins, CO, USA.
pankaj.saini800@colostate.edu

Ravi L. Singh
Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA.
singh899@missouri.edu

Abstract

Understanding the propagation of financial volatility across global markets is a fundamental challenge for systemic risk assessment, portfolio construction, and regulatory oversight. Traditional econometric approaches, such as Granger causality tests and vector autoregressive models, impose strong linearity and stationarity assumptions that fail to capture the nonlinear, time-varying nature of modern financial interconnectedness. This paper introduces a framework for causal discovery of volatility spillovers that integrates interpretable machine learning with rigorous walk-forward robustness guarantees. We develop a system architecture that combines constraint-based causal structure learning with attention-based neural network components, enabling the extraction of directed acyclic graphs representing volatility transmission channels. The walk-forward validation regime ensures that all causal estimates are generated using only information available at the time of inference, thereby avoiding look-ahead bias and producing out-of-sample stability guarantees. We discuss the structural trade-offs between model complexity, interpretability, and computational feasibility in large-scale deployment across multiple asset classes and geographies. The paper further examines governance implications, including the need for algorithmic transparency in systemic risk monitoring, fairness considerations in cross-border spillover estimation, and policy recommendations for regulatory adoption of interpretable causal models. Our analysis highlights the importance of combining causal discovery with robustness guarantees to build trustworthy decision-support infrastructures for global financial stability.

Keywords

causal discovery, volatility spillovers, interpretable machine learning, walk-forward validation, systemic risk, financial infrastructure, robustness guarantees.

1. Introduction

The increasing integration of global financial markets has intensified the transmission of shocks across borders, asset classes, and time scales. Volatility spillovers, defined as the propagation of second-moment shocks from one market to another, are central to understanding systemic risk, contagion dynamics, and portfolio diversification benefits [1]. Traditional approaches to modeling spillovers, such as the Diebold–Yilmaz connectedness measure based on variance decompositions from vector autoregressions, rely on linear reduced-form frameworks that assume stable causal structures over the estimation window [2]. These methods are ill-suited to capture the nonlinear dependencies, regime shifts, and asymmetric responses that characterize modern financial systems. Moreover, they often fail to distinguish between mere correlation and genuine causal mechanisms, a distinction that is critical for policy interventions and risk management.

Recent advances in machine learning have offered new tools for modeling financial time series, including deep learning architectures that can capture complex temporal patterns [3]. However, the black-box nature of many such models undermines their utility in contexts where interpretability and causal understanding are paramount, such as regulatory stress testing or central bank monitoring [4]. Causal discovery, a subfield of artificial intelligence concerned with inferring causal relationships from observational data, provides a principled alternative. By combining graph-based representations of structural equations with statistical tests for conditional independence, causal discovery algorithms can produce directed graphs that encode hypotheses about the underlying data-generating process [5]. When applied to volatility spillovers, these methods hold the promise of revealing not only which markets transmit volatility to others, but also the direction and conditional nature of those transmissions.

Yet the deployment of causal discovery in financial applications faces two significant obstacles. First, the non-stationarity and evolving nature of financial markets means that any static causal graph estimated from historical data may become invalid in a new market regime. Second, the computational complexity of exact causal discovery over a large number of variables can be prohibitive, necessitating trade-offs between accuracy and scalability. This paper addresses both challenges by proposing a framework that integrates interpretable machine learning with walk-forward robustness guarantees. The walk-forward approach, widely used in financial model validation, partitions the data into sequential training and testing periods, ensuring that all causal estimates are produced using only contemporaneously available information [6]. This methodology yields out-of-sample stability assessments that are essential for building confidence in the discovered causal structures.

The remainder of this paper is organized as follows. Section 2 surveys related work at the intersection of volatility modeling, causal inference, and interpretable machine learning. Section 3 describes the methodological framework for causal discovery, emphasizing the role of attention mechanisms and conditional independence testing in producing interpretable spillover graphs. Section 4 details the walk-forward robustness architecture, including algorithmic guarantees and computational considerations. Section 5 presents an empirical illustration using a multi-asset global dataset and discusses system-level trade-offs. Section 6 addresses governance, fairness, and policy implications. Section 7 outlines future research directions and sustainability challenges, and Section 8 concludes.

2. Background and Related Work

The study of volatility spillovers has a long tradition in financial econometrics. Early contributions by Granger [7] laid the foundation for testing causal relationships in time series using predictive power. Subsequent work by Engle, Ito, and Lin [8] introduced the notion of meteor showers and heat waves in volatility transmission, leading to the development of multivariate generalized autoregressive conditional heteroskedasticity models. More recently, Diebold and Yilmaz [2] proposed a connectedness measure based on forecast error variance decompositions that has become a standard tool in applied finance. However, these methods assume a linear, reduced-form structure and do not explicitly model the underlying causal graph. As a result, they can misattribute causal direction when feedback loops or common causes are present.

Causal discovery offers a more principled approach. The constraint-based PC algorithm of Spirtes, Glymour, and Scheines [9] and the score-based greedy equivalence search have been applied to economic and financial systems with promising results. More recent work leverages continuous optimization for learning directed acyclic graphs, as in the NOTEARS framework by Zheng et al. [10], which reformulates the combinatorial search for DAGs as a continuous optimization problem. These methods have been extended to handle time series data through the inclusion of lagged variables and temporal constraints [11]. The work of Liu [11], in particular, demonstrates how interpretable machine learning can be combined with realistic walk-forward constraints for volatility forecasting, highlighting the importance of out-of-sample evaluation in causal discovery.

Interpretability is a key requirement for any model used in high-stakes financial decision-making. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) provide post-hoc explanations for black-box models [12]. However, these explanations may not be faithful to the true causal structure. An alternative is to build inherently interpretable models, such as attention-based neural networks, where the attention weights can be visualized as a measure of feature importance [13]. In the context of causal discovery, attention mechanisms can be used to learn a soft adjacency matrix that approximates the causal graph, offering both scalability and interpretability. The combination of attention-based learning with conditional independence testing yields a hybrid approach that preserves the strengths of both paradigms.

3. Methodology: Causal Discovery and Interpretable Machine Learning

Our framework for causal discovery of volatility spillovers consists of three main components: a data preprocessing pipeline that transforms raw price data into volatility measures; a causal structure learning module that combines conditional independence testing with attention-based neural network training; and an interpretability layer that produces human-readable causal graphs and importance scores. The preprocessing stage involves computing realized volatility from high-frequency or daily returns, typically using squared returns or more robust estimators such as the Yang–Zhang estimator [14]. To address non-stationarity, we apply a moving-window standardization and detrending procedure that preserves the time-varying nature of volatility while mitigating the impact of level shifts.

The core causal discovery module operates on a set of preprocessed volatility series representing different markets or asset classes. We adopt a constraint-based approach, where the causal graph is constructed by performing a sequence of conditional independence tests. The test used in our implementation is based on partial correlations, which are appropriate for continuous variables under the assumption of linearity and Gaussianity. To relax these assumptions, we also incorporate a non-parametric test based on conditional mutual information [15]. The output of the testing procedure is a graph skeleton that is then oriented using rules that resolve collider structures and acyclic constraints.

Because the number of conditional independence tests grows exponentially with the number of variables, we introduce an attention-based neural network that learns a soft ordering on the variables, effectively prioritizing the most likely causal directions. The attention mechanism computes a matrix of relevance scores between each pair of variables, which serves as a prior for the constraint-based search [13]. This hybrid approach dramatically reduces the computational burden while retaining interpretability: the attention weights themselves can be inspected to understand which market pairs are most strongly connected. The final causal graph is a directed acyclic graph where edges represent volatility transmission channels, and edge weights indicate the strength of the causal relationship.

The model is trained using a loss function that combines a graph structure penalty with a prediction objective. Specifically, the network is tasked with predicting the volatility of a target market given the past volatility of all markets, and the attention weights are regularized to encourage sparsity and acyclicity. This joint optimization ensures that the learned causal structure is both consistent with the observed data and useful for forecasting. The interpretability layer then generates summary statistics, such as the total connectedness index and directional spillover indices, that can be visualized as network diagrams or heatmaps. These outputs are directly comparable to traditional connectedness measures but reflect causal rather than predictive relationships.

4. Walk-Forward Robustness Guarantees: Architecture and Implementation

A critical limitation of many causal discovery applications in finance is the reliance on a single historical estimation window. Financial markets exhibit structural breaks, regime changes, and evolving volatility dynamics that render any static graph obsolete. To address this, we embed the causal discovery pipeline within a walk-forward validation framework. The walk-forward approach divides the full time series into a sequence of overlapping or non-overlapping windows. For each window, the causal discovery is performed using only data up to the current time point, and the resulting graph is evaluated on the subsequent out-of-sample period [6]. This procedure mimics real-time deployment and provides an empirical distribution of causal structures over time.

Our implementation uses a fixed-length training window that is rolled forward by one step at each iteration. At each step, we estimate the causal graph using the methods described in Section 3, compute a set of robustness metrics – such as edge persistence, graph edit distance from the previous window, and predictive accuracy of volatility forecasts – and record the results. The collection of graphs across windows allows us to assess the stability of each causal relationship. An edge that appears consistently across many windows is considered robust, while an edge that appears only sporadically is deemed unreliable. We define a robustness guarantee threshold: an edge is claimed as a true causal relationship only if it appears in at least a predetermined percentage of windows and if the out-of-sample forecast error remains within acceptable bounds.

This architecture provides several guarantees. First, it eliminates look-ahead bias by construction, since each estimation uses only data available at the time. Second, it quantifies the uncertainty surrounding each causal edge through the empirical distribution of its presence across windows. Third, it enables early detection of structural shifts: when a previously robust edge disappears, it may signal a regime change that warrants further investigation. The computational cost of this repeated estimation is substantial, but we mitigate it through parallelization and by using the attention-based prior to reduce the search space. For a system with twenty markets and a rolling window of five years of daily data, the entire walk-forward procedure completes in a few hours on a standard server.

5. Empirical Analysis and System-Level Trade-Offs

To illustrate the framework, we consider a global set of equity index volatility series from major markets including the United States, United Kingdom, Japan, Germany, France, China, and Brazil, covering the period from January 2010 to December 2024. Daily realized volatilities are computed from intraday prices where available, and from daily closing prices otherwise. The walk-forward procedure uses a training window of 1260 days (approximately five years) and a forecast horizon of one day. The robustness threshold is set at 70% persistence over windows.

The results reveal several noteworthy patterns. The volatility spillover structure is highly asymmetric, with the US market acting as a net transmitter of volatility to all other markets in most windows, consistent with the financial dominance of the US economy [16]. However, during periods of European debt crisis (2011–2012) and the COVID-19 pandemic (2020), the directionality became more diffuse, with European markets exhibiting increased transmission to the US [17]. The causal graph also shows that connections between emerging markets, such as Brazil and China, are weaker and less stable than those between developed markets, suggesting that portfolio diversification benefits may be time-varying. The attention weights from the neural network align closely with the edges discovered by the constraint-based search, validating the use of the hybrid approach.

The system-level trade-offs are multifaceted. Increasing the number of markets included in the analysis improves the completeness of the spillover map but also increases computational cost and reduces the reliability of conditional independence tests due to the curse of dimensionality [9]. There is a fundamental trade-off between robustness and responsiveness: longer training windows yield more stable causal estimates but are slower to adapt to structural breaks. Our walk-forward framework makes this trade-off explicit, allowing the user to choose a window length that balances stability and timeliness based on the specific application. Another trade-off involves interpretability versus model complexity. The constraint-based method produces fully interpretable graphs but may miss complex nonlinear dependencies; the attention-based neural network can capture nonlinearities but its weights, while informative, are not always faithful to true causal mechanisms [12]. Our hybrid approach mitigates this by using the neural network only to prioritize the search space, while the final graph is determined by statistical tests.

6. Governance, Fairness, and Policy Implications

The deployment of causal discovery for volatility spillovers has significant governance implications. Financial regulators and central banks increasingly rely on systemic risk dashboards that track interconnectedness among institutions and markets [18]. If these dashboards are based on correlation rather than causation, they may produce misleading signals. For instance, a high correlation between two markets could arise from a common third factor, leading to a false alarm of contagion. Causal discovery addresses this by identifying the actual channels of transmission. However, the reliability of these causal estimates depends heavily on the robustness guarantees provided by the walk-forward framework. Regulators should require that any causal model used for policy decisions include out-of-sample validation and stability metrics similar to those described here.

Fairness considerations arise in cross-border spillover analysis. When a causal model attributes volatility transmission from a developing market to a developed market, it may have policy repercussions such as capital flow restrictions or international blame. The model must be transparent about its assumptions and limitations. Our interpretability layer provides a direct audit trail for each edge, showing the conditional independence tests and the window-specific evidence supporting the causal claim. This transparency is essential for accountability. Moreover, the walk-forward robustness guarantee prevents the model from making claims based on ephemeral patterns. In practice, regulators could mandate that only edges meeting a minimum persistence threshold be used for policy actions.

Policy implications extend to the design of financial infrastructure. The ability to accurately identify causal spillover channels can inform the allocation of supervisory resources, the design of stress test scenarios, and the calibration of macroprudential instruments such as countercyclical capital buffers. In a multi-jurisdictional setting, harmonizing the definitions and methodologies for volatility spillovers is critical to avoid regulatory arbitrage. Our framework offers a standardized, reproducible approach that can be adopted by international bodies such as the International Monetary Fund or the Financial Stability Board [19].

7. Future Directions and Sustainability

Looking forward, several developments could enhance the practical utility of this framework. The integration of high-frequency data, including tick-level order flow and limit order book dynamics, would allow for finer-grained causal analysis at intraday horizons. However, this would also introduce new challenges related to data volume and the need for real-time processing. The current walk-forward architecture could be adapted to an online learning setting, where the causal graph is updated incrementally as new data arrive, rather than being re-estimated from scratch at each window [20]. Such an adaptive system would be more responsive to regime changes and more computationally sustainable.

Another important direction is the incorporation of additional data modalities, such as textual information from news articles or social media sentiment, into the causal discovery process. Volatility spillovers are often driven by narratives and investor attention, which are not captured by price data alone. Natural language processing models, when combined with causal discovery, could provide a richer picture of shock propagation mechanisms [21]. The interpretability of such models remains an open challenge.

Sustainability in the sense of long-term model maintenance is also crucial. Financial markets evolve, and a causal model calibrated on a historical period may eventually become outdated. The walk-forward framework provides a natural mechanism for monitoring model drift: when the out-of-sample prediction error increases significantly, it signals that the underlying causal structure has changed, prompting a partial or full retraining. Establishing automated drift detection and retraining protocols would reduce the need for human intervention and lower operational risk.

Finally, the broader research community would benefit from standardized benchmarks for causal discovery in financial time series. Currently, there is no widely accepted dataset or evaluation metric for volatility spillover causality. Our framework, with its clear walk-forward validation and robustness thresholds, could serve as a basis for such a benchmark.

This would facilitate comparisons across different algorithmic approaches and accelerate the adoption of causal methods in financial regulation and industry.

8. Conclusion

This paper has presented a comprehensive framework for causal discovery of volatility spillovers in global markets that integrates interpretable machine learning with walk-forward robustness guarantees. By combining constraint-based causal structure learning with attention-based neural network priors, we achieve a scalable and interpretable method for estimating directed acyclic graphs of volatility transmission. The walk-forward validation architecture ensures that all causal estimates are free of look-ahead bias and that their temporal stability can be rigorously assessed. Empirical results using a multi-market global dataset illustrate the asymmetric and time-varying nature of spillovers, and highlight the structural trade-offs between model complexity, computational cost, and interpretability. The governance, fairness, and policy implications underscore the importance of causal transparency in systemic risk monitoring. As financial markets continue to evolve, the need for robust, interpretable, and causally grounded analytical tools will only grow. Our framework provides a principled path toward meeting that need.

References

1. Engle, R. F., Ito, T., & Lin, W. L. (1990). Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. Econometrica, 58(3), 525–542.

2. Diebold, F. X., & Yilmaz, K. (2009). Better to give than to receive: Predictive directional measurement of volatility spillovers. International Journal of Forecasting, 28(1), 57–66.

3. Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied Soft Computing, 90, 106181.

4. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215.

5. Pearl, J. (2009). Causality: Models, reasoning, and inference (2nd ed.). Cambridge University Press.

6. Prado, M. L. (2018). Advances in financial machine learning. Wiley.

7. Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–438.

8. Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. Econometrica, 55(2), 251–276.

9. Spirtes, P., Glymour, C., & Scheines, R. (2000). Causation, prediction, and search (2nd ed.). MIT Press.

10. Zheng, X., Aragam, B., Ravikumar, P., & Xing, E. P. (2018). DAGs with NO TEARS: Continuous optimization for structure learning. In Advances in Neural Information Processing Systems (Vol. 31).

11. Liu, T. (2026). Interpretable Machine Learning for Volatility Forecasting Under Realistic Walk-Forward Constraints.

12. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (Vol. 30).

13. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (Vol. 30).

14. Yang, D., & Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. Journal of Business, 73(3), 477–492.

15. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11), eaau4996.

16. Forbes, K. J., & Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market comovements. The Journal of Finance, 57(5), 2223–2261.

17. Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. The Journal of Finance, 56(2), 649–676.

18. Bisias, D., Flood, M., Lo, A. W., & Valavanis, S. (2012). A survey of systemic risk analytics. Annual Review of Financial Economics, 4, 255–296.

19. Financial Stability Board. (2017). Supervisory and regulatory recommendations on structural reforms: Final report. FSB.

20. Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504–507.

21. Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1–8.