Asset Graph-Based Network Metrics for Volatility Forecasting

Aleksander Janczewski*(13401017), Charel Felten*(13364022), James Kingsbury*(13437542)
*University of Amsterdam, Amsterdam, Netherlands

***Abstract*—In this investigation we explore the realised volatility forecasting power of various graphs metrics applied to financial correlation networks. We examine monthly changes in correlation between 26 global stock market indices over a period of nearly 20 years. These correlations are transformed into networks, from which we compute the minimum spanning tree (MST), planar maximally filtered graph (PMFG), mean eigenvector centrality (MEC) and transitivity (TRANS). The former two networks are investigated with the aim of reproducing the work from Magner et al. [18]. The latter two are both derived from the asset graph (AG) of the fully connected network. To our knowledge, we are the first to apply these two particular metrics in volatility forecasting. Our findings suggest that the MEC and TRANS metrics achieve lower squared errors on out-of-sample tests than VMSTL and VPMFGL.**

I. Introduction

FINANCIAL-MARKET risk, commonly measured in terms of index-return volatility, plays a fundamental role in investment decisions, risk management and has wider implications for macro-economic fundamentals.

From 2004 to early 2007, the financial markets had been reasonably stable. Both the SP500 volatility and the VIX index were below their long-run averages [19]. During the 2008 crisis, the SP500 suffered a loss of over 39%, one of its worst years on record [34]. The VIX index more than tripled, earning it the moniker: “the fear index”. Much of the fallout in 2008 bore strong resemblance to the after effects of the 1987 crisis. In both cases taxpayer funds had to be used to bail out major institutions, significant financial reform had to be undertaken and millions of individuals suffered hardship as a result. Furthermore, Choudhry demonstrates that there is a long-run causal relationship between index volatility and subsequent expenditure on both goods and services, [7]. In this way, volatility in the market indices ”spill out” into the real economy.

The relationship between market uncertainty and public confidence leads central banks to monitor market volatility as an indicator of the wider health of the system. Poon and Granger state that the Federal Reserve explicitly accounts for asset volatility when setting monetary policy, similarly the Bank of England Monetary Policy Committee is known to make reference to option-implied volatilities [31]. Effective strategies for volatility forecasting are useful for policy makers, investors and risk managers alike. The ability to take pre-emptive action before major perturbations allows market participants and regulators to adjust their behaviour to minimise post-crisis consequences.

Network analysis provides us with a convenient tool for managing the complexity inherent in the relationships between financial indices. In this analysis we abstract away from the detailed dynamics and instead focus on whether a relationship exists between indices, and if so, the magnitude of the relationship. The networks methodology is particularly well suited to volatility forecasting given that exogenous economic shocks tend to spread between countries, and network models are known to be well suited to modelling contagious dynamics [10]. In this paper we will use various network types and metrics based on the global financial correlation network to make predictions about future realised volatility for each of the individual indices.

II. Background

In this section we attempt to provide a brief overview of literature relevant to our research. We will first investigate some relevant work on the topic of volatility forecasting, then cover some literature related to financial correlation matrices and network metrics. Combining the two topics, we close the background by introducing “The Volatility Forecasting Power of Financial Network Analysis” by Magner et al. [18]. This paper represents an early attempt to exploit the correlation network method for volatility forecasting.

We first wish to motivate our research into volatility forecasting. Solnik et al. find that correlations between international markets increase in times of high market volatility [39]. One of the mentioned consequences is the loss in effectiveness of international risk diversification, as both domestic and international assets are more likely to move in tandem in times of increased volatility.

In one of the more recent volatility forecasting studies, Wang presents insights on the importance of VIX and its larger component in volatility forecasting [43]. Wang further finds that using VIX and VIX_L in the regressive models provides better predictive power in forecasting volatility than autoregression of RV only, with the emphasis on the superiority of large VIX index over original VIX [43]. Other recent approaches that can be found in literature include use of financial network metrics in predicting the volatility of stocks.

Turning to the usage of network metrics to investigate financial networks, we have Mantenga, who was one of the first to construct hierarchical networks based on the correlations between individual stock prices, [20]. The author also motivates the use of the minimum spanning tree from an economic point of view. From the time series of stock prices, and through the construction of a tree, Mantenga is able to locate grouping of stocks which he relates back to defined subindustry sectors. It transpired that these clustering results were comparable to a more formal clustering method based on Potts super-paramagnetic transitions, [28].

Onnela et al. extended Mantenga's work, first with research on asset trees [28] and later with a series of publications on asset graphs [29] [27]. Onnela's early work demonstrates how the tree evolves over time, and notably, how it shrinks significantly during a crisis. The authors also observe that the stocks which form an optimised Markowitz portfolio tend to lie on the edges of the tree. Two potential characterisation of the tree are posited, the normalised tree length and the mean occupation layer. The authors then investigate the effect of crises on these characterisations. Finding that the normalised tree length and mean occupation layer both fluctuate in time and shrink significantly during crises. In a later publication, Onnela espouses the advantages of asset graphs over asset trees [29]. When comparing the normalised tree length and the normalised graph length, the authors find that the asset graph tends to exaggerate market anomalies and consequently the asset tree length seems to track market performance more closely.

Bonanno et al. also attempt to extend the tree approach to account for time dependence in the study of index data [4]. The time evolution of the tree structure was studied and periods of significant structural change were identified. The authors experimented with different clustering methods and alternative time horizons. The authors are interested in the consequences of geographically distributed indices. In particular, the fact that stock exchanges in different time zones may not have totally overlapping trading times. The "quasi-synchronous" nature of these indices means that an orthodox synchronous analysis of global indices over short time intervals is not possible. The more general effects of non-synchronous time series analysis are documented in the econometrics literature, [17]. For example, non-synchronicity can obscure the inference of our correlation results. These effects partially explain why we choose to examine monthly windows, rather than performing a more granular analysis. At this interval, the non-synchronous hourly mismatch of data will be virtually non-existent.

More recently, Liu and Tse [16] propose to investigate complex networks created from the cross-correlations of various stocks markets over time. Their research concentrates on the relation between network synchronisation (average correlation matrix edge weight), node strength (average weight of all edges connected to a node) and observed market volatility. They investigate different time window sizes, ranging from 20 to 60 days. They find that as world stock markets fluctuate, there is high tendency for increased synchronisation between them. Additionally they find negative correlation between average price and network synchronisation. Lastly, they note that these phenomena are more pronounced in developed markets while frontier and emerging markets remain less affected.

Another investigation focused on local Korean and global stock index correlation and network topologies is conducted by Nobi et al. [25]. The authors use intra-day stock index data to construct correlation matrices and subsequently threshold networks based on periods before, during, and after the 2008 GFC. Various metrics are used to analyze the constructed complex networks, including e.g. the clustering coefficient, minimum spanning tree and the topological changes between networks are explored with the Jaccard similarity. The authors find that local indices tend to correlate very strongly in times of crisis. Moreover, they find that the Jaccard similarities between global threshold networks are higher than between local ones, and that this metric has potential as an indicator for systemic risk in markets due its effectiveness in detecting changes in market state.

Particularly interested in the effects of the 2008 GFC, Li and Pi [15] constructed MSTs from a network of global indices and investigated clustering coefficients and average thresholds. They find regional clusters among these networks as well as small world properties before and during crises - that is, a high clustering coefficient and small characteristic path length, indicative of a high number of subgraphs with global reach [21].

In related terms, Jang et al. [44] relate foreign exchange network topologies to market returns and volatility. Building a minimum spanning tree network based on moving window correlations between FX returns, they investigate network properties such as normalised tree length, node strength and closeness centrality. Of interest to our research is their conclusion that these market returns correlate positively to the centrality measures, while their volatility correlates negatively to them.

The topics of volatility forecasting and financial network analysis are both associated with reasonably large volumes of research, the former even more so than the latter. However, our work is primarily concerned with the intersection of these two topics. Finding original research in this niche proved to be quite difficult. The paper by Magner et al. [18] represents a novel attempt to apply the financial network analysis techniques pioneered by Onnela and Mantenga to the problem of stock index volatility forecasting. The authors make use of two common characterisations of financial correlation networks: the planar maximally filtered graph length and the minimum spanning tree length. Hypothesising that these quantifications of the global network contain at least some information which is useful for predicting the realised volatility of individual indices, the authors then apply relatively rudimentary regression techniques to assess the in-sample and out-of-sample performance of the proposed metrics. It is partially confirmed that for European and Asian-Pacific market indices, the log-variation in the minimum spanning tree length can act a useful predictor. The log-variation of planar maximally filtered graph length proves to be a substantially less powerful predictor, contrary to the authors' original hypothesis.

Our research will be arranged in the following manner. First, we provide a comprehensive review of the methods used to construct the financial correlation networks, compute graph metrics and ultimately, assess the power of the predictors. Second, we attempt to reproduce the in-sample and out-of-sample regression results presented by Magner et al., further discuss the challenges encountered, and finally analyse how the results we obtain differ from those of the authors. Finally, we assess the performance of our proposed metrics, using a similar testing methodology as that of Magner et al.

III. Methods

A. Constructing the network

In order to reproduce the results from Magner's publication "The Volatility Forecasting Power of financial Network Analysis", we utilize daily closing prices for a total of 26 international indices, listed in Table I. In order to reproduce Magner et al.'s results we use data provided to us by the author, obtained from Bloomberg. However, the exact constituents of the indices used are not fully disclosed in the publication. Furthermore, for the purpose of our independent research focusing on the mean eigenvector centrality and transitivity metrics, we make use of our own data sourced from Bloomberg.

For both the reproductive efforts and our independent research, we first compute logarithmic returns of the closing prices of indices on consecutive days, as presented in Equation 1. Computation of logarithmic returns provides very good approximation of the percentage change given the temporal proximity of the closing prices [20].

$$r_i = \ln P(\tau) - \ln P(\tau - 1) \tag{1}$$

where r_i stands for logarithmic return, $P(\tau)$ indicates the closing price on the trading day τ. Note we only consider consecutive trading days, the data imputation has not been performed as not only does it induce artificial correlations between indices, but also was not employed by Magner et al. For each month we calculate Pearson correlations between each index, with the formula presented in Equation 2:

$$\rho^t_{i,j} = \frac{\langle \mathbf{r}^t_i \mathbf{r}^t_j \rangle - \langle \mathbf{r}^t_i \rangle \langle \mathbf{r}^t_j \rangle}{\sqrt{\left[\langle \mathbf{r}^{t\,2}_i \rangle - \langle \mathbf{r}^t_i \rangle^2\right]\left[\langle \mathbf{r}^{t\,2}_j \rangle - \langle \mathbf{r}^t_j \rangle^2\right]}} \tag{2}$$

where ρ^t_{ij} stands for Pearson correlation between indices i and j on month t, and $\langle ... \rangle$ denotes a temporal average over the consecutive trading days in the return vectors.

Moving forward, the $N \times N$ correlation matrix $\boldsymbol{C}$ is then transformed into a $N \times N$ distance matrix $\boldsymbol{D}$, using Equation 3.

$$d_{i,j} = \sqrt{(2(1 - \rho_{i,j}))} \tag{3}$$

where, $d_{i,j}$ indicates the distance between indices i and j. The choice of this metric is advocated by Mantegna, as it fulfills the three axioms of a metric distance [20]. This transformation maps the previously computed Pearson's correlations from range $[-1, 1]$ to distances ranging from $[0, 2]$, where the highest possible distance of 2 indicates perfect negative correlation and the lowest possible distance of 0 indicates perfect positive correlation. The distances between indices i and j are further used to generate networks where indices are represented by nodes and edges by the distance matrix. The resulting network is fully connected with N nodes and $N(N-1)/2$ edges.

We use logarithmic returns to calculate realized variance RV (Equation 4) for each index i on each month t. This is done

TABLE I: Global indices used in the investigation.

Index	Abbrev.	Region	Country
S&P 500	SPX	North America	USA
NASDAQ	CCMP	North America	USA
S&P/TSX	SPTSX	North America	Canada
S&P/BMV IPC	MEXBOL	South America	Mexico
BOVESPA	IBOV	South America	Brazil
S&P CLX IPSA	IPSA	South America	Chile
S&P MERVAL	MERVAL	South America	Argentina
S&P BVL	IGBVL	South America	Peru
FTSE 100	FTSE	Europe	United Kingdom
CAC 40	CAC	Europe	France
DAX	DAX	Europe	Germany
IBEX 35	IBEX	Europe	Spain
FTSE MIB 40	FTSEMIB	Europe	Italy
AEX	AEX	Europe	Netherlands
OMX 30	OMX	Europe	Sweden
RTSI	RTS	Europe	Russia
SMI	SMI	Europe	Switzerland
NIKKEI 225	NKY	Asia	Japan
HANG-SENG	HSI	Asia	Hong-Kong
KOSPI	KOSPI	Asia	Korea
TAIEX	TWSE	Asia	Taiwan
JKSE	JCI	Asia	Jakarta
FBMKLCI	FBMKLCI	Asia	Malaysia
STI	STI	Asia	Singapore
S&P/ASX 200	ASX	Oceania	Australia
S&P/NZX	NZSE	Oceania	New Zealand

by simply summing up the squares of logarithmic returns for all the days τ in a month t of the index of interest i.

$$\mathrm{RV}_{i,t} = \sum_{\tau=1}^{T} r^2_{i,\tau,t} \tag{4}$$

where T indicates all trading days τ for index i in month t.

We further define realized volatility RVOL (Equation 5) and realized log-variance RLV (Equation 6), following conventions introduced by Barndorff-Nielsen and Shepherd [2] in the following manners:

$$\mathrm{RVOL}_{i,t} = \sqrt{\sum_{\tau=1}^{T} r^2_{i,\tau,t}} \tag{5}$$

$$\mathrm{RLV}_{i,t} = \ln\left(\sum_{\tau=1}^{T} r^2_{i,\tau,t}\right) \tag{6}$$

where, T indicates all trading days τ for index i in month t.

B. Graph Metrics

1) Minimum Spanning Tree: The minimum spanning tree (MST) is the smallest (in terms of combined edge length) possible tree which spans the whole network. It joins the N indices using $N - 1$ edges. By definition a spanning tree is a connected acyclic graph. The number of nodes in the tree will be time independent given that we have index data for all indices over the entire test time period. However, the value of the edge weights will be time dependent, given that elements in $\boldsymbol{D}$ will vary between windows. The MST can be constructed using either Prim's algorithm [32], or Kruskal's algorithm [13]. When discussed, Prim and Kruskal are often cited as both the sources of the MST problem and the first sources of efficient

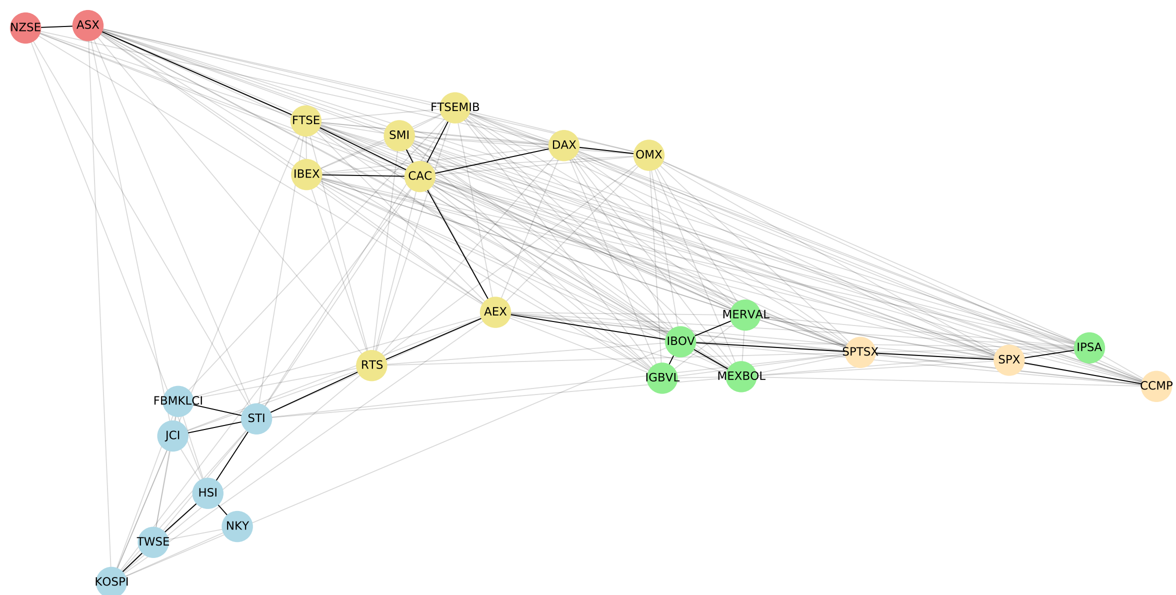

Fig. 1: Example of an MST (black edges) and AG (gray edges) with proportion $\theta = 0.6$ of the investigated index network. The length of the displayed edges is not of true scale. Particularly interesting is that while the North American indices are not directly connected to the European ones via the MST, the AG retains a lot of these connections.

solutions. However, both authors cite the 1926 paper of Czech mathematician Otakar Borůvka [6] who has since received wider credit for his formulation of the algorithm [11].

$$\text{MSTL}(t) = \left(\frac{1}{N-1}\right) \sum_{d_{ij}^t \in T^t} d_{ij}^t, \tag{7}$$

Equation 7 describes the sum of edge weights, where d_{ij}^t is defined as the distance between node i and node j, T^t is the MST, N is the number of nodes and t is the particular time period this MST is being constructed for. Additionally, the length is normalised by the number of nodes.

The idea is to encode the most important information (lowest distances) from our fully connected network into a tree with less edges, such that the most important index correlations are reflected over time through a varying minimum spanning tree length (MSTL). In general, it is possible for the MST to have multiplicity, however, given that each of our edge weights are likely to be unique and time dependent, the probability of this occurring is low. Following the method of Magner et al., we take the log variation of the MSTL, yielding the metric shown in Equation 8.

$$\text{VMSTL}(t) = \ln \text{MSTL}(t) - \ln \text{MSTL}(t-1) \tag{8}$$

An example of an MST constructed from our index correlation network is shown in Figure 1, with the black lines representing the edges of the minimum spanning tree. Of note is the natural clustering behaviour which arises, especially the intra-continental clustering among the European and Asian indices.

2) Planar Maximally Filtered Graph: The planar maximally filtered graph (PMFG), is a subgraph of representative links which is created by filtering $\boldsymbol{D}$ [41]. We can construct a maximally filtered graph by iteratively linking the nodes with the strongest connections (shortest distances), under the constraint that the newly generated graph is of a given genus k. The “filtering” level of the maximally filtered graph can be adjusted by varying the genus of the resultant graph, however we will only investigate the planar case where $k = 0$. Tuminello et al. also prove that the resulting PMFG will contain the MST and retain all of the hierarchical information originally encoded in the MST. The PMFG includes a total of $3N - 6$ edges compared to the $N - 1$ of the MST. In contrast to the MST, the PMFG is able to retain some triangular loops and 4-element cliques from the complete graph. Magner et al. hypothesise that this extra information allows the PMFG to explain to a greater degree the synchronicity between the indices [18]. Again following the method of Magner et al. we take the log variation of the PMFGL, yielding the metric shown in Equation 9.

$$\text{VPMFGL}(t) = \ln \text{PMFGL}(t) - \ln \text{PMFGL}(t-1) \tag{9}$$

3) Asset Graph: To construct an asset graph (AG) we first take the elements of $\boldsymbol{D}$ and place them in ascending rank order, yielding a sequence $d_1^t, d_2^t, ..., d_{N(N-1)/2}^t$, assuming single index notation. We create an empty subgraph and iteratively add edges in ascending order of weight. We continue adding edges until we reach a particular proportion of completeness, θ, which represents the proportion of edges from the original complete graph that we wish to retain in our new asset graph.

We will refer to the matrix representing this asset graph as $\boldsymbol{A}$. Asset graphs need not be connected and neither is there any acyclicity condition present. Clustering behaviour in the original network will be more effectively retained in the AG compared to the MST, given that cliques and cycles can be included. We conjecture that some of this retained clustering information may be useful for predicting realised volatility.

We can compute other network metrics based on the asset graph such as transitivity and mean eigenvector centrality, explained in Sections III-B4 and III-B5, respectively. The motivation behind applying these metrics on the asset graph instead of the fully connected network is that, topology-wise, a fully connected network is rather uninteresting. The asset graph on the other hand allows to convey some of the edge length information into the topology by filtering the highest correlations. Both an asset graph and MST are shown in Figure 1; it can be seen that the asset graph retains significantly more information than the minimum spanning tree.

The most problematic part of constructing asset graphs is deciding on θ. Greater values of θ mean that more of the loosely connected nodes retain connections. Lower values of θ lead to only the nodes with the shortest distances (and thus higher correlations) retaining their edges. The trick is to retain some kind of balance when deciding on the threshold. In our investigation we will use an AG with with proportion $\theta = 0.6$. Although we, much like other authors before us [29], must admit that this is somewhat of an arbitrary choice.

4) Transitivity: The transitivity (TRANS) of a graph is a measure of the degree to which nodes tend to cluster together.

$$\text{Transitivity} = \frac{3 \ \times \ \text{number of triangles in the graph}}{\text{number of connected triples of vertices}} \tag{10}$$

As seen in Equation 10, it is the number of closed triplets divided by the total number of open and closed triplets in the graph. The factor 3 is present in the numerator because each complete triangle of 3 nodes provides 3 connected triples, one centered at each of the 3 nodes. More specifically, we can consider the transitivity as the fraction of connected triples of nodes which ultimately form triangles [3]. In practice, a network with a high transitivity will tend to have many tightly connected groups with a high density of edges. This measure was originally defined by Newmann, Watts and Strogatz in 2002 [24]. The authors claimed that this measure was equivalent to the popular "global clustering coefficient"; subsequent articles also failed to make a distinction between the two metrics. While the metrics are similar as they both measure the relative frequency of triangles, they give subtly different results [36]. A complete graph of 4 nodes with one edge removed is the smallest counter example which demonstrates the difference between the two metrics.

In this work, we will apply the transitivity measure to a network which has first been filtered by an asset graph. Given this, a higher coefficient implies that nodes with particularly low distances and thus high correlations are clustered together. This means that a sufficiently large exogenous shock on the world economy can spread quickly within these clusters and thus they will move in a synchronised fashion leading to increased realised volatility. Given this proposed relationship, we hypothesise that TRANS can serve as a predictor for the future RLV.

5) Mean Eigenvector Centrality: Eigenvector centrality measures the influence of individual nodes in a network. Relative scores are given to all nodes based on the idea that connections to high-scoring nodes contribute more to the score of the node in question than connections to low-scoring nodes. A high centrality for a particular node means that this node is also connected to many nodes who themselves have high scores, [23].

We define as x_i the centrality of node i, then setup a relationship in which x_i is proportional to the sum of the centralities of the nodes adjacent to i, shown in equation 11.

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} \boldsymbol{A}_{ij} x_j \tag{11}$$

We combine the centralities of all nodes $(x_1, x_2, x_3, ...)$ into an eigenvector $\boldsymbol{x}$, which allows to rewrite the relationship in matrix form.

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \tag{12}$$

From equation 12, we can see $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ with associated eigenvalue λ. To determine the centrality values we must obtain the eigenvector associated with the dominant eigenvalue. The i^{th} value in the eigenvector then gives the relative centrality score of node i. The centrality of each node will be a function of both the quantity and quality of the connections to adjacent nodes. Having lots of neighbours is somewhat important but having high quality neighbours is even more important. That is to say, often a node with a few high quality connections will outrank a node with many low-quality connections.

We compute the centrality scores of each of the nodes in the asset graph filtered network. We continue by taking the mean of the centrality over all nodes to form the mean eigenvector centrality (MEC). We hypothesise that this new metric is indicative of the extent to which these highly correlated indices are influential upon each other, given that the elements in our asset graph already represent a subset of the most correlated indices in our complete network. Similarly to the case of transitivity, an asset graph with a high MEC implies any exogenous shock on the world economy can spread quickly within the network, encouraging index movements in a synchronised fashion and driving increases in RLV.

C. Regression Methods

We perform OLS regression with the Statsmodels package in Python, [37]. Following the method of Magner et al., we make use of a heteroskedasticity and autocorrelation consistent estimators, [22]. This is primarily performed as a precaution, ensuring that small amounts of potential auto-correlation in the time-series do not lead to spurious regression results.

We conduct two broad kinds of regression tests to determine the efficacy of our graph metrics. The first, in-sample testing, involves fitting an OLS regression on a particular set of data,

TABLE II: The benchmarks used for out-of-sample testing.

Abbreviation	Benchmark
AR(3)	$\text{RLV}_{i,t} = c + \gamma_{i,1} \cdot \text{RLV}_{i,t-1} + \gamma_{i,2} \cdot \text{RLV}_{i,t-2} + \gamma_{i,3} \cdot \text{RLV}_{i,t-3} + e_i$
AR(6)	$\text{RLV}_{i,t} = c + \gamma_{i,1} \cdot \text{RLV}_{i,t-1} + \gamma_{i,2} \cdot \text{RLV}_{i,t-2} + \gamma_{i,3} \cdot \text{RLV}_{i,t-3} + \gamma_{i,4} \cdot \text{RLV}_{i,t-4} + \gamma_{i,5} \cdot \text{RLV}_{i,t-5} + \gamma_{i,6} \cdot \text{RLV}_{i,t-6} + e_i$
AR(3)VIX	$\text{RLV}_{i,t} = c + \gamma_{i,1} \cdot \text{RLV}_{i,t-1} + \gamma_{i,2} \cdot \text{RLV}_{i,t-2} + \gamma_{i,3} \cdot \text{RLV}_{i,t-3} + \delta_i \cdot \text{VVIX}_{t-1} + e_i$
AR(6)VIX	$\text{RLV}_{i,t} = c + \gamma_{i,1} \cdot \text{RLV}_{i,t-1} + \gamma_{i,2} \cdot \text{RLV}_{i,t-2} + \gamma_{i,3} \cdot \text{RLV}_{i,t-3} + \gamma_{i,4} \cdot \text{RLV}_{i,t-4}$ $+\gamma_{i,5} \cdot \text{RLV}_{i,t-5} + \gamma_{i,6} \cdot \text{RLV}_{i,t-6} + \delta_i \cdot \text{VVIX}_{t-1} + e_i$

and then assessing the fit of the regression on that data. The second, out-of-sample testing, involves splitting the data into a training set and a test set. We first estimate the regression coefficients $\hat{\beta}_j$ on the training set and then make predictions on unseen data. While the former method provides some initial indication of the usefulness of a predictor, ultimately the latter assessment forms the basis for evidence of real predictive power.

For the in-sample regression, the dependant variable is the RLV (Equation 6) of an index i in month t, and the independent variables are our metric shifted back to month $t-1$ and autoregressive terms (AR$-n$) which are also the RLV of the same index but shifted back to month $t-n$. The transformation of RV terms into RLV terms is motivated by Santos and Ziegelmann, who provide evidence that employing models based on logarithmic RV terms allows to obtain significant gains in forecasting accuracy [35].

For the out-of-sample regressions, we have the same dependant variable in each of our models, however here we compare 4 different models against benchmarks which are identical except that they don't contain the tested metric. The 4 benchmarks are AR(3), AR(3)VIX, AR(6) and AR(6)VIX, which contain the above mentioned autoregressive terms and the VIX (mean close price of month $t-1$) respectively. These particular benchmarks were chosen to ease comparison with the results from Magner et al. [18]. Table II describes the benchmark models used when assessing our out-of-sample performance.

1) Assessment: When conducting an in-sample test we focus our attention on two particular assessment criteria. First, we examine the R^2 of our fitted regressions. This provides an indication of the amount of variance in the dependent variable (RLV in our case), which is explained by the predictors. If the R^2 of a model increases after we add our new metric, it suggests that our new metric is useful in explaining RLV. The second criteria examined is the Granger causality of our metric time-series upon the time-series of the RLV. If one time-series "granger causes" another, then the first series should contain information which is useful for predicting the second time-series over and above that which is already contained in the second time series. We perform a hypothesis test on the metric coefficient where $H_0 : \hat{\beta}_i = 0$ and $H_1 : \hat{\beta}_i > 0$. If H_0 is rejected then we can conclude with some level of confidence that our metric contains information which can be used to predict future values of RLV.

Our out-of-sample regime involves ENC-NEW, a test statistic developed by Clark and McCracken to compare the predictive power of a model with that of a related benchmark model [8]. Under the null hypothesis, the squared error loss of the benchmark model is the same as that of the model including our additional predictor. Under the alternative hypothesis, the squared error loss of the larger model is less than that of the benchmark model.

$$H_0 : E[L_{t+1}(\hat{\beta}_1)] = E[L_{t+1}(\hat{\beta}_2)], \tag{13}$$

$$H_1 : E[L_{t+1}(\hat{\beta}_1)] > E[L_{t+1}(\hat{\beta}_2)], \tag{14}$$

To formalise this test, we can define the hypotheses as in Equation 13 and Equation 14, where L_{t+1} represents the squared error loss of the prediction and $\hat{\beta}_1$ and $\hat{\beta}_2$ represent the estimated parameter vectors of model 1 and model 2, respectively. In our case, model 1 refers to a benchmark model and model 2 refers to one of our proposed models. We refer to models in this tested regime as "nested", given that our new model contain all of the predictors that are present in the benchmarks, plus one additional predictor. For example, to test the efficacy of VMSTL as a predictor, we could compare a regression model with VMSTL and three autoregressive terms with a benchmark model which is composed solely of three autoregressive terms.

$$\text{ENC-NEW} = P \cdot \frac{\sum_t (\hat{u}_{1,t+1}^2 - \hat{u}_{1,t+1}\hat{u}_{2,t+1})}{\sum_t \hat{u}_{2,t+1}^2}, \tag{15}$$

In practice, testing the hypothesis outlined above involves computing the test statistic as in Equation 15. $\hat{u}_{1,t+1}$ and $\hat{u}_{2,t+1}$ denotes the one-step-ahead forecast errors for model 1 and model 2, respectively. After computing the test statistic, we next obtain critical values and then determine the outcome of our hypothesis test using [26].

2) Stationarity: Magner et al. take the log difference of both the PMFGL and the MSTL to "allow us to work with a stationary time-series". While this transformation seems prudent, considering the effect that non-stationary time-series can have on regression, we believe it is not necessary in every case and moreover, marginally improved results are obtained in some situations with non-differenced time-series. For this reason, we perform a couple of statistical tests to verify the stationarity of the proposed metrics.

To obtain useful OLS results, it is vital that non-stationary time series are not included. If they are, then estimated coefficients in the regression fit which are not really significant may appear to be artificially significant: spurious regression. To ensure that the time series of our metrics are stationary we employ two tests, both of which rely on testing if the series in question has a unit root. A series has a unit root of 1 is a root of the series characteristic equation. The first test we employ is the Augmented Dickey-Fuller test (ADF) [9]. Using the ADF we can test the null hypothesis that our time series has a unit root (non-stationary). The alternative hypothesis is that the series is stationary. The ADF test statistic will always be negative, the larger the absolute value of the statistic, the stronger the evidence in favour of the alternative hypothesis. The second test we employ is the Phillips–Perron test (PP) [30], similar in nature to the ADF test, the null hypothesis again being that the series has a unit root.

3) Window Types: In the out-of-sample assessment, we make use of both recursive and rolling windows. To construct rolling windows, we first define the portion of our data for which to make predictions. Then, for each of these points, the regression coefficients are fitted on a window of previous time-series points. For each subsequent prediction, both the start and the end of the training window shift one position forward. The size of the training window remains fixed for each of the out-of-sample predictions.

In contrast, recursive windows involve increasing the size of our training window for each additional out-of-sample prediction. The start position of the training window remains fixed, but for each subsequent prediction, the end of the training window expands to encompass the point that was predicted in the previous time period.

It should be noted that the critical value for our ENC-NEW test statistic depends on the number of one-step-ahead forecasts and the number of examples in the training window. We define $\pi = \frac{P}{R}$ where P is the number of forecasts and R is the number of training examples. We use $\pi = 0.4$ for both rolling and recursive schemes. However, in our recursive window scheme, the size of the training window increases with each additional forecast. The size of the first training window is used, given that this leads to most conservative (largest) critical value.

IV. Reproduction Results

A. In-sample

To reproduce the results from Magner et al. as accurately as possible we contacted the authors. They kindly sent us the original returns data used in their study, ostensibly sourced from Bloomberg. The following VMSTL and VPMFGL results were computed from this data, any discrepancies are likely due to differences in data processing, the actual number of constituents in the indices used and/or due to the ambiguity surrounding the time periods ultimately used by the authors. Our reproduction uses a window of data from January 2001 to February 2002, the rationale for this particular selection will be explained in Section V.

1) Minimum Spanning Tree: Table III represents our efforts in reproducing Magner et al.'s in-sample regression results for the VMSTL. In principle, we would expect the estimated $\hat{\beta}$ for VMSTL(-1) to have a negative sign, for the regression fit on each of the indices. As explained earlier, an increase in MSTL implies decreased correlations between indices, less synchronous movement and thus lower realised volatility. All of the estimated coefficients significant at the 10% level have a negative sign, which partially confirms our economic interpretation of this metric. Broadly, we obtain similar VMSTL significance results to Magner et al.

We obtain 1 significant coefficient at the 1% level, 8 at the 5% level and 11 at the 10% level. This is in comparison to the 12 significant at the 10% level obtained by Magner et al.; we believe this discrepancy is caused by one of the justifications outlined in Section V. Our intercept terms also have a similar magnitude, although these are mostly devoid of economic interpretation.

We expect each of the AR terms to have a positive coefficient. That is, higher previous values of realised volatility imply higher future values of RLV, and vice versa. For each of the indices, the AR(-1) coefficient term is significant at the 1% level. The AR(-2) and AR(-3) terms are also reasonably significant; this is intuitive given that more recent AR terms should be better predictors of future values than older terms.

Magner et al. highlights the intra-continental patterns which emerge in their regression tests. They find substantial geographic significance clustering in both the EU and Asia-Pacific. We find analogous geographic clustering behaviour. We observe that a large portion of the European indices (FTSE, CAC, IBEX, MIB and OMX) have significant coefficients. We also observe a cluster of significant indices in the Asia-Pacific region: NKY, HSI, KOSPI, TWSE, ASX and NZSE. Each of these significant coefficients imply the existence of Granger Causality between the lagged VMSTL and RLV for that particular index. This causality implies a relationship between the correlations formed in the global index network and the volatility of the European and Asian-Pacific indices.

2) Planar Maximally Filtered Graph: Table IV shows our in-sample VPMFGL results for the 26 indices. We obtain 1 significant coefficient at the 1% level, 5 at the 5% level and 8 at the 10% in comparison to the 7 found by Magner et al. Similarly to VMSTL, an increase in VPMFGL implies a lower correlation among indices, larger probability of synchronised movements and thus larger volatility. All of our significant coefficients have a negative sign, partially confirming our hypothesis about their economic interpretation. However, the decreased significance of VPMFGL compared to VSMTL implies that the additional correlation information contained in the VPMFGL is not particularly useful for RLV prediction, if anything, it adds additional noise. This finding is consistent with the findings from Magner et al. who argued that "the MSTL is more efficient by not considering correlations of lesser magnitude" [18].

All of the fitted coefficients for the autoregressive terms are again positive, implying that larger realised volatility terms in previous periods imply larger realised volatility in the current period. All of the AR(-1) coefficients are significant, with reduced significance for longer lagged terms.

The geographic significance clustering which emerges is similar to that of the VMSTL, except on a reduced scale. Most of the significant indices are still clustered in Europe(CAC, MIB, OMX) or in Asia-Pacific (NKY, HSI, KOSPI, TSE, NZSE). Again, we conjecture that the additional loops and

Variable	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
C	**-1.448*****	**-1.236*****	**-1.458*****	**-1.780*****	**-1.862*****	**-2.955*****	**-1.563*****	**-1.644*****	**-1.487*****	**-1.373*****	**-1.304*****	**-1.332*****	**-1.160*****
	(0.390)	(0.295)	(0.415)	(0.366)	(0.377)	(0.475)	(0.305)	(0.357)	(0.312)	(0.272)	(0.270)	(0.272)	(0.233)
VMSTL(-1)	-0.276	-0.311	-0.313	0.111	0.257	-0.295	0.313	0.118	**-0.642***	**-0.901****	-0.591	**-0.602***	**-0.713****
	(0.420)	(0.362)	(0.385)	(0.326)	(0.300)	(0.359)	(0.338)	(0.411)	(0.377)	(0.395)	(0.374)	(0.344)	(0.320)
AR(-1)	**0.492*****	**0.453*****	**0.471*****	**0.581*****	**0.434*****	**0.473*****	**0.513*****	**0.491*****	**0.446*****	**0.446*****	**0.531*****	**0.512*****	**0.487*****
	(0.097)	(0.084)	(0.084)	(0.072)	(0.086)	(0.063)	(0.073)	(0.063)	(0.085)	(0.093)	(0.083)	(0.063)	(0.076)
AR(-2)	**0.164***	**0.174****	**0.268*****	0.011	**0.157****	0.041	0.013	**0.172****	**0.207*****	**0.274*****	**0.172***	**0.184****	**0.273*****
	(0.094)	(0.084)	(0.081)	(0.085)	(0.075)	(0.085)	(0.093)	(0.084)	(0.072)	(0.091)	(0.088)	(0.072)	(0.087)
AR(-3)	**0.122***	**0.170****	0.046	**0.124***	0.063	0.042	**0.164****	0.072	**0.116***	0.050	0.076	0.078	0.040
	(0.067)	(0.067)	(0.061)	(0.066)	(0.052)	(0.074)	(0.074)	(0.087)	(0.061)	(0.056)	(0.059)	(0.052)	(0.056)
R-squared	0.495	0.533	0.522	0.421	0.306	0.262	0.350	0.427	0.486	0.520	0.536	0.515	0.563

Variable	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
C	**-1.235*****	**-1.219*****	**-1.804*****	**-2.018*****	**-2.019*****	**-1.158*****	**-1.017*****	**-1.158*****	**-2.093*****	**-1.913*****	**-1.319*****	**-1.527*****	**-2.544*****
	(0.265)	(0.287)	(0.341)	(0.349)	(0.370)	(0.308)	(0.235)	(0.276)	(0.390)	(0.405)	(0.299)	(0.323)	(0.621)
VMSTL(-1)	-0.685	**-0.751****	-0.038	-0.726	**-0.877****	**-0.658****	**-0.791****	**-1.041*****	-0.217	-0.524	-0.491	**-0.684***	**-0.586****
	(0.428)	(0.356)	(0.353)	(0.482)	(0.432)	(0.322)	(0.315)	(0.305)	(0.340)	(0.399)	(0.383)	(0.379)	(0.288)
AR(-1)	**0.483*****	**0.417*****	**0.511*****	**0.401*****	**0.405*****	**0.509*****	**0.435*****	**0.314*****	**0.457*****	**0.364*****	**0.492*****	**0.433*****	**0.334*****
	(0.098)	(0.077)	(0.089)	(0.101)	(0.081)	(0.066)	(0.072)	(0.078)	(0.072)	(0.073)	(0.062)	(0.086)	(0.060)
AR(-2)	**0.257*****	**0.295*****	**0.130***	**0.202****	**0.252*****	**0.192***	**0.228*****	**0.383*****	0.100	**0.264*****	**0.193****	**0.220*****	**0.200*****
	(0.085)	(0.081)	(0.072)	(0.081)	(0.080)	(0.104)	(0.065)	(0.081)	(0.076)	(0.086)	(0.074)	(0.074)	(0.058)
AR(-3)	0.059	**0.086***	0.015	0.084	-0.005	0.106	**0.172*****	**0.118****	**0.101***	0.107	**0.114****	**0.113***	**0.121***
	(0.050)	(0.051)	(0.075)	(0.060)	(0.051)	(0.072)	(0.053)	(0.057)	(0.061)	(0.070)	(0.048)	(0.061)	(0.071)
R-squared	0.552	0.550	0.356	0.369	0.361	0.565	0.597	0.547	0.321	0.404	0.533	0.479	0.310

Window	Benchmark	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
ROL	AR(3)	-0.049	-0.025	-0.136	-0.099	0.073	-0.362	-0.177	0.129	0.089	**1.501****	**0.913***	**0.949***	**1.602****
	AR(3)VIX	-0.382	-0.295	-0.075	-0.339	-0.247	-0.189	-0.346	-0.254	-0.257	0.390	0.193	-0.167	-0.193
	AR(6)	-0.176	-0.136	-0.223	0.157	-0.040	-0.278	0.005	-0.003	0.008	**1.327****	**0.801***	0.726	**1.342****
	AR(6)VIX	-0.313	-0.238	0.086	-0.205	-0.217	-0.320	-0.435	-0.148	-0.329	0.468	0.034	-0.056	-0.090
REC	AR(3)	-0.049	0.023	-0.091	-0.439	-0.357	-0.359	-0.134	-0.084	0.181	**1.251****	**0.836***	**0.991***	**1.450****
	AR(3)VIX	-0.230	-0.148	-0.157	-0.516	-0.294	0.036	-0.289	-0.216	-0.339	0.112	0.163	-0.270	-0.173
	AR(6)	-0.159	-0.070	-0.188	-0.096	-0.071	-0.268	0.001	-0.149	-0.061	**1.019***	**0.755***	**0.742***	**1.283****
	AR(6)VIX	-0.193	-0.035	-0.126	-0.392	-0.151	-0.086	-0.264	-0.210	-0.407	0.182	0.086	-0.149	-0.054

Window	Benchmark	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
ROL	AR(3)	**0.820***	**0.969***	-0.352	**1.109***	**1.300****	**0.794***	**1.359****	**1.070***	-0.650	-0.746	-1.018	0.241	0.011
	AR(3)VIX	0.298	-0.082	-1.128	-0.523	-0.045	0.188	**1.129***	-0.039	-0.472	-0.347	-0.761	-0.127	-0.199
	AR(6)	0.496	0.678	-0.443	**1.013***	**1.200****	0.445	**1.301****	**0.859***	-0.572	-0.679	-0.986	0.162	-0.142
	AR(6)VIX	0.258	-0.055	-1.061	-0.709	0.126	0.050	**1.150***	-0.155	-0.613	-0.306	-0.823	-0.184	-0.226
REC	AR(3)	0.615	**1.199****	-0.961	**1.079****	**1.242****	**0.864***	**1.316****	**1.015***	-0.595	-0.822	-0.995	0.322	0.252
	AR(3)VIX	0.089	-0.010	-1.213	-0.172	-0.032	0.387	**0.789***	0.172	-0.447	-0.481	-0.810	-0.228	-0.173
	AR(6)	0.337	**0.928***	-1.032	**0.886***	**1.192****	0.406	**1.212****	**0.790***	-0.547	-0.511	-0.781	0.069	0.032
	AR(6)VIX	0.046	-0.024	-1.009	-0.330	0.108	0.148	**0.837***	0.118	-0.594	-0.283	-0.660	-0.294	-0.175

TABLE III: In-sample (upper tables) and out-of-sample (lower tables) results for VMSTL. The regression is performed on Magner et al.'s data for a period starting from 2001-01 until 2019-02.
The first two tables show the coefficients as well as their significance level, and their standard error for all metrics and indices. The bottom two tables show the ENC-NEW statistic and significance level for all 4 models and for both recursive window regression (used by Magner et al.), as well as rolling window regression.
***, ** and * indicate significance at the 1%, 5% and 10%-level, respectively. For the out-of-sample tests, the critical values of the ENC-NEW statistic with $\pi = 0.4$ at the 1%, 5% and 10%-level are 2.278, 1.161 and 0.764, respectively, for the rolling window regression and 2.098, 1.079 and 0.685 for recursive window regression, as proposed by Clark and McCracken [26].

clicks contained in the VPMFGL compared to VMSTL are not adding extra value and if anything, are obscuring the results.

B. *Out-of-sample Results*

While in-sample results are indicative of the usefulness of a predictor, ultimately the predictive power of a graph metric is only borne out when one-step-ahead forecasts are made. As mentioned in the previous section, for out-of-sample regressions we consider two types of training windows: the rolling and the recursive window. Moreover, for each of the types of window, we test against 4 different benchmarks previously defined in Table II. However, when considering reproduction we are mainly concerned with the recursive window case. Magner et al. do claim to have considered rolling windows, however no specific results are presented.

1) Minimum Spanning Tree: Let us first consider the out-of-sample results for the VMSTL. As can be seen in Table III, for our recursive scheme with the AR(3) benchmark we obtain 10 significant ENC-NEW test statistics at the 10% level in comparison with the 15 significant out-of-sample values at the 10% level achieved by Magner et al. For the recursive scheme with the AR(6) benchmark, we obtain 9 significant values at the 10% level, compared to the 14 achieved by the authors.

The two benchmarks including the VIX prove extremely difficult to beat, with both only having 1 significant value at

Variable	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
C	**-1.450*****	**-1.237*****	**-1.457*****	**-1.781*****	**-1.862*****	**-2.952*****	**-1.570*****	**-1.647*****	**-1.489*****	**-1.373*****	**-1.305*****	**-1.333*****	**-1.164*****
	(0.390)	(0.296)	(0.415)	(0.366)	(0.377)	(0.476)	(0.306)	(0.358)	(0.312)	(0.272)	(0.270)	(0.272)	(0.233)
VPMFGL(-1)	-0.115	-0.241	-0.267	0.051	0.258	-0.329	0.225	0.071	-0.587	**-0.870****	-0.572	-0.566	**-0.668***
	(0.438)	(0.373)	(0.387)	(0.338)	(0.312)	(0.364)	(0.340)	(0.419)	(0.406)	(0.425)	(0.417)	(0.373)	(0.340)
AR(-1)	**0.515*****	**0.464*****	**0.478*****	**0.573*****	**0.433*****	**0.472*****	**0.502*****	**0.490*****	**0.457*****	**0.456*****	**0.537*****	**0.520*****	**0.497*****
	(0.097)	(0.084)	(0.083)	(0.072)	(0.086)	(0.062)	(0.073)	(0.062)	(0.087)	(0.094)	(0.087)	(0.064)	(0.076)
AR(-2)	0.142	**0.164***	**0.262*****	0.018	**0.158****	0.041	0.022	**0.174****	**0.197*****	**0.265*****	**0.167***	**0.177****	**0.265*****
	(0.095)	(0.085)	(0.082)	(0.084)	(0.075)	(0.084)	(0.094)	(0.084)	(0.073)	(0.090)	(0.091)	(0.072)	(0.086)
AR(-3)	**0.121***	**0.169****	0.045	**0.124***	0.063	0.044	**0.164****	0.071	**0.115***	0.049	0.076	0.078	0.038
	(0.067)	(0.067)	(0.061)	(0.066)	(0.052)	(0.074)	(0.074)	(0.087)	(0.061)	(0.055)	(0.059)	(0.052)	(0.056)
R-squared	0.495	0.532	0.522	0.420	0.306	0.262	0.349	0.427	0.485	0.518	0.535	0.514	0.562

Variable	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
C	**-1.237*****	**-1.218*****	**-1.803*****	**-2.021*****	**-2.021*****	**-1.156*****	**-1.014*****	**-1.155*****	**-2.092*****	**-1.911*****	**-1.318*****	**-1.529*****	**-2.547*****
	(0.265)	(0.288)	(0.341)	(0.349)	(0.372)	(0.308)	(0.235)	(0.277)	(0.390)	(0.405)	(0.298)	(0.323)	(0.623)
VPMFGL(-1)	-0.601	**-0.749****	-0.062	-0.605	**-0.819***	**-0.696****	**-0.801****	**-1.017*****	-0.278	-0.526	-0.512	-0.623	**-0.574***
	(0.466)	(0.368)	(0.368)	(0.502)	(0.433)	(0.334)	(0.328)	(0.329)	(0.349)	(0.417)	(0.384)	(0.410)	(0.299)
AR(-1)	**0.498*****	**0.420*****	**0.509*****	**0.418*****	**0.415*****	**0.510*****	**0.441*****	**0.321*****	**0.455*****	**0.366*****	**0.492*****	**0.445*****	**0.336*****
	(0.101)	(0.077)	(0.089)	(0.102)	(0.081)	(0.066)	(0.072)	(0.079)	(0.072)	(0.073)	(0.060)	(0.087)	(0.059)
AR(-2)	**0.242*****	**0.292*****	**0.131***	**0.185****	**0.239*****	**0.192***	**0.223*****	**0.377*****	0.102	**0.262*****	**0.193*****	**0.209*****	**0.199*****
	(0.087)	(0.078)	(0.072)	(0.081)	(0.079)	(0.103)	(0.067)	(0.083)	(0.077)	(0.087)	(0.074)	(0.076)	(0.058)
AR(-3)	0.059	**0.086***	0.015	0.083	-0.002	0.105	**0.171*****	**0.116****	**0.101***	0.107	**0.114****	**0.112***	**0.120***
	(0.050)	(0.050)	(0.075)	(0.059)	(0.051)	(0.072)	(0.053)	(0.056)	(0.061)	(0.070)	(0.048)	(0.060)	(0.071)
R-squared	0.551	0.550	0.356	0.366	0.358	0.565	0.596	0.545	0.322	0.403	0.533	0.478	0.309

Window	Benchmark	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
ROL	AR(3)	-0.169	-0.117	-0.267	-0.168	0.035	-0.284	-0.532	0.077	-0.160	**1.030***	0.651	0.638	**1.129***
	AR(3)VIX	-0.416	-0.287	-0.098	-0.367	-0.317	-0.012	-0.568	-0.365	-0.196	0.442	0.242	-0.103	-0.161
	AR(6)	-0.260	-0.220	-0.356	-0.037	-0.104	-0.225	-0.319	-0.021	-0.246	**0.864***	0.535	0.399	**0.837***
	AR(6)VIX	-0.317	-0.204	0.066	-0.275	-0.327	-0.199	-0.684	-0.239	-0.307	0.561	0.109	-0.014	-0.075
REC	AR(3)	-0.157	-0.093	-0.236	-0.516	-0.345	-0.250	-0.584	-0.145	-0.061	**0.868***	0.653	**0.713***	**1.028***
	AR(3)VIX	-0.227	-0.162	-0.173	-0.501	-0.311	0.236	-0.537	-0.187	-0.312	0.144	0.177	-0.233	-0.138
	AR(6)	-0.216	-0.171	-0.352	-0.273	-0.083	-0.184	-0.405	-0.195	-0.267	0.671	0.580	0.456	**0.846***
	AR(6)VIX	-0.186	-0.041	-0.122	-0.430	-0.213	0.078	-0.485	-0.180	-0.405	0.231	0.129	-0.138	-0.023

Window	Benchmark	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
ROL	AR(3)	0.370	**0.806***	-0.555	0.557	**1.178****	**0.959***	**1.358****	**0.899***	-0.534	-0.678	-1.139	-0.023	-0.052
	AR(3)VIX	0.267	-0.020	-1.214	-0.569	0.013	0.542	**1.280****	0.008	-0.219	-0.390	-0.892	-0.079	-0.145
	AR(6)	0.048	0.504	-0.627	0.563	**0.995***	0.656	**1.165****	0.746	-0.454	-0.680	-1.143	-0.112	-0.208
	AR(6)VIX	0.233	0.000	-1.157	-0.710	0.093	0.373	**1.140***	-0.117	-0.383	-0.337	-0.862	-0.172	-0.113
REC	AR(3)	0.355	**1.016***	-1.064	0.678	**1.132****	**0.954***	**1.320****	**0.890***	-0.499	-0.694	-1.109	0.072	0.242
	AR(3)VIX	0.052	0.036	-1.223	-0.139	0.051	**0.706***	**0.979***	0.144	-0.212	-0.459	-0.873	-0.218	-0.043
	AR(6)	0.119	**0.762***	-1.171	0.526	**0.970***	0.545	**1.102****	**0.741***	-0.451	-0.487	-0.953	-0.145	0.017
	AR(6)VIX	0.026	0.023	-1.084	-0.234	0.108	0.425	**0.875***	0.067	-0.350	-0.302	-0.696	-0.307	-0.019

TABLE IV: In-sample (upper tables) and out-of-sample (lower tables) results for VPMFGL. The regression is performed on Magner et al.'s data for a period starting from 2001-01 until 2019-02.

the 10% level. This contrasts with the results presented by Magner et al., who finds 6 and 4 significant test statistics at the 10% level for the AR(3)VIX and AR(6)VIX, respectively.

Our results for one-step-ahead forecasting using the VM-STL are clearly less promising than those originally presented by Magner et al. Additionally, the fact that the authors present notably more significant out-of-sample values than in-sample significant coefficients raises some doubts about the integrity of their out-of-sample methods.

2) Planar Maximally Filtered Graph: Next, we consider the out-of-sample results presented for the VPMFGL, again for the recursive windows. The results are presented Table IV. For the AR(3) model, we obtain 8 significant ENC-NEW test statistics at 10% confidence, compared to the 13 obtained by Magner et al. For the AR(6) benchmark we obtain 5 significant at the 10% level, compared to 14 obtained by the authors. It seems again slightly counter-intuitive that the author manages to obtain more significant out-of-sample results on the harder AR(6) benchmark than on the AR(3) benchmark. When considering VPMFGL on the AR(3)VIX and AR(6)VIX, we find only 2 and 1 significant values at the 10% level with the authors finding 3 and 4, respectively. Clearly, both our results and those of the authors suggest that the VPMFGL is no match for our benchmarks models which include the VIX.

V. Reproduction Discussion

Reproduction of scientific work may appear to be a simple task, given that such work does not require any novel findings. It came to us as a surprise that reproductive efforts can not only allow one to identify all the flaws in the publication, but also provide basis for one's own research. Given that our findings regarding Magner et al.'s work proved fundamental for our own research, it is pertinent that we first discuss all the challenges that we faced while reproducing the work, identify the flaws that we found in the publication, and further provide overall assessment of the method. By doing so we intend to make it clear how we addressed these challenges and findings in our own work. The key issues faced were the lack of clarity regarding various aspects of work and misleading information regarding the actual regressions.

In particular, one of the key problems was the demystification of the transformations performed on monthly realized variance terms before the regressions were performed. Magner et al. use the terms realized variance (RV) and realized volatility (RVOL) interchangeably. The table captions with regression results state "The forecasting realized volatility [...]" suggesting that regressions were performed on RVOL terms (Equation 5). Having carefully followed Magner et al.'s steps, we were unable to reproduce the regression results, neither using RVOL nor RV. Given that not only the signs of coefficients were different, but also the magnitude of coefficients were orders of magnitude smaller, it became unlikely that either RVOL or RV terms were actually used. Our investigation concluded that RV computed using Equation 4 must have been further transformed into realized log-variance (RLV) with Equation 6, before the regression was performed. Although seemingly trivial once solved, this and similar issues constituted a major challenge in reproducing the results of the authors, particularly with respect to the total absence of any mention of RLV throughout the whole paper.

The source of confusion behind these terms could be due to a misconception across scientific fields. Barndorff-Nielsen and Shepherd highlight the fact that RV, defined as a sum of squared returns, is often referred to as RVOL in econometrics [2]. Other examples of this particular misconception can be found in literature. Audrino et al. clearly define RVOL as the square root of the sum of squares of logarithmic returns [1], whereas Vortelinos defines RVOL as simply the "sum of the observable intraday squared returns" [42].

Using RLV to perform our regressions, the in-sample coefficients for the autoregressive terms (AR) aligned closely with those presented by Magner et al. However, the coefficients and their significance for VMSTL and VPMFGL metrics were still quite far off. While investigating this problem, inconsistencies regarding the period of data used in the investigation were found. The authors state "We used daily data provided by Bloomberg from July 2001 to September 2019, totaling 216 months for a total of 26 market indices [...]", while in the "The Data" section they mention "We used daily data provided by Bloomberg from July 2001 to September 2019, totaling 223 months". Not only are there actually 219 months in the period between July 2001 and September 2019, but they also state that they use 26 indices, whereas a figure presented in the paper of the constructed MST actually contains 27 indices, with the added index being the Columbian IGBC. These findings let us to further investigate which indices were actually used and over which time periods the regressions were performed.

Upon investigation of the data provided to us by Prof. Magner, two things became clear. First, that the time period began in January 2001 and second, that the CSV file contained the IGBC index. The latter finding raised the question why it was not used in the author's final regression tests. Investigating the data, we found missing entries for IGBC from January 2001 until July 2001, hence in the scenario where one wanted to work with a different period starting from earlier than July 2001, IGBC would need to be excluded. This finding prompted us to attempt to drop IGBC from our data and experiment with different periods. After much experimentation, it turned out that shifting the period of investigation to the period starting from January 2001 until February 2019 not only allowed us to obtain VMSTL and VPMFGL coefficients very closely aligned with those of the authors, but also resulted in similar numbers of significant VMSTL and VPMFGL coefficients. These findings indicate that the actual period of time used for regression was changed, a claim further supported by the ambiguity surrounding the supposed period of investigation mentioned in the paper.

When attempting to reproduce Magner et al.'s results, we also made a deep dive into the fundamentals upon which the work was based. It came to our attention that there appears to be multiple weaknesses in the method, one of them being the computation of Pearson correlations between indices. More specifically, the fact that Pearson correlations between indices are calculated for each month based on, at most, 20 data points. Kukreti et al. suggest that correlation matrices calculated over such short time periods are usually noisy, highly singular and hence very sensitive to minor changes in correlations [14]. It is usually recommended to apply noise suppression techniques to such correlation matrices, for example Random Matrix Theory [14]. A similar notion is presented also by Guhr and Kälber, where they corroborate the fact that noise is very common in correlation matrices computed from short time series [12]. Furthermore, Bonanno et al. clarify that when the number of indices is larger than the number of time points in the covariance matrix is only positive semidefinite [5]. On the other hand, Tang provides a different point of view where he claims that that minimum spanning trees and planar maximally filtered graphs are sufficient noise filtering methods themselves [40].

Nonetheless, this issue remains a source of concern, as not only do we consider consecutive trading days (which amounts to roughly 20 days per month, while the correlations are computed between 26 indices), but also global markets have different business days, while some operate from Monday to Friday, others from Sunday till Thursday, not to mention the possibility of differing holiday schedules. While the literature surveyed previously suggests various methods of dealing with the issue, they have not been employed by Magner et al.

The next issue found was that the MST lengths provided in Figure 1 of the author's paper are incorrect. Using Magner's data and his indicated time periods, we attempted to reproduce the MSTL values for the pre-crisis, crisis and post-crisis periods. We obtained different lengths, approximately 70% larger than those reported. We hypothesise that the authors may have made an error in their computation. We experimented with different distance transformations and found that by excluding the factor 2 from the distance transformation (Equation 3), we then obtain the exact same lengths as the original authors.

Finally, another source of discrepancies that we observe are very significant differences in our and Magner et al.'s data,

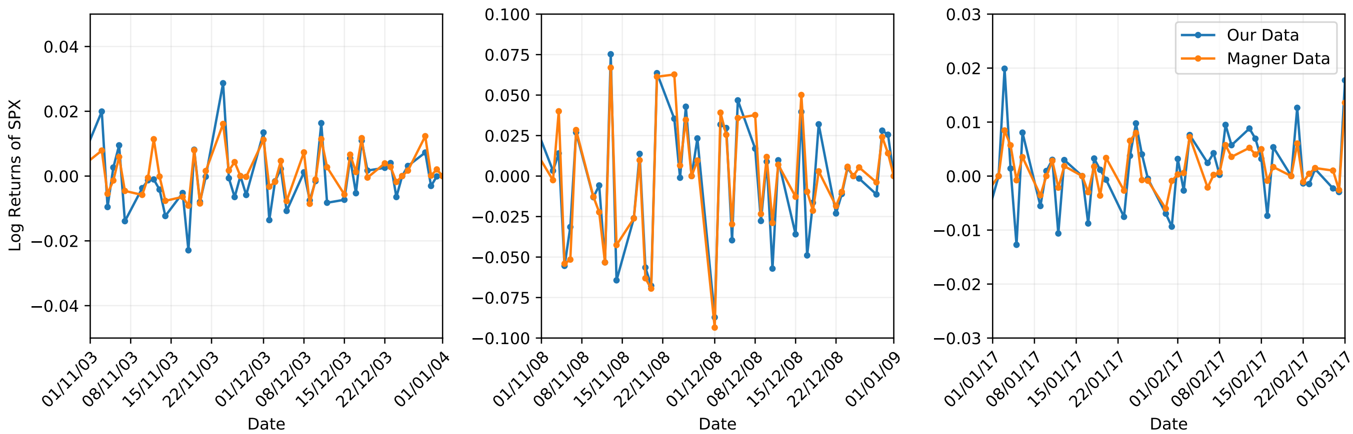

Fig. 2: The log-returns of SPX from our data and the data from Magner et al.

although both are sourced from Bloomberg. Magner et al. does not mention pre-processing data, or any imputations, thus making it unclear why we can observe such large differences in data for various indices. An example of discrepancies between returns of the SPX index is presented in Figure 2. The discrepancies could also be due to the indices only being named implicitly without further specification on how many constituting companies were included.

VI. Our Results

A. Retesting VMSTL and VPMFGL

Having outlined our efforts in reproducing the results for the VMSTL and VPMFGL and challenges faced in doing so, we will continue with the results from applying VMSTL and VPMFGL to our own data as well as results of the TRANS and MEC metrics. For these 4 metrics, we will consider the same time period, 2001-07 until 2019-09, which corresponds to the time period Magner et al. claim to have used in their publication. We will perform the same in-sample and out-of-sample regression tests that we performed in Section IV. The motivation for re-testing VMSTL and VPMFGL on our data is so that we have an equivalent basis of comparison between the VMSTL, VPMFGL, MEC and TRANS.

1) Minimum Spanning Tree: For the VMSTL, we find 2 out of 26 estimated coefficient significant, both of which are significant only at the 10% level. Additionally, we also find all 26 coefficients for the constant and AR(-1) to be significant and 21 out of 26 significant AR(-2) coefficients. The out-of-sample tests reveal that none of ENC-NEW values of the AR(3)VIX and AR(6)VIX benchmarks are significant; for the rolling window we find 2 significant coefficients for both AR(3) and AR(6) benchmarks and for the recursive window only 1 for AR(3) and 2 for AR(6). The only index both significant in-sample and out-of-sample is NZSE and no intra-continental clusters can be observed.

2) Planar Maximally Filtered Graph: For VPMFGL, the in-sample and out-of-sample regression results are presented in Table VIII. We find 2 out of 26 estimated coefficients to be significant; the same two indices are for VMSTL, with the VPMFGL coefficient being significant at the 10% level for NKY and at the 5% level for NZSE. For out-of-sample results, we find 2 and 3 estimated coefficient significant for rolling window AR(3) and AR(6) benchmarks and again 2 and 3 for recursive window AR(3) and AR(6) benchmarks. NZSE is significant for all non-VIX benchmarks.

The large number of regressions performed and low numbers of significant results, even at the 10% level, indicate that both the VMSTL and VPMFGL metrics are not particularly reliable nor robust predictors of realised volatility. It is important to remember that the confidence level of our test determines the rate of type I errors we are willing to accept. Thus, we have up to 10% probability of incorrectly rejecting a correct null hypothesis. The sheer number of regressions performed implies that false positives are inevitable. The insignificance of the results becomes particularly pertinent when compared to the two new metrics investigated, TRANS and MEC.

B. Stationarity testing

1) Mean Eigenvector Centrality: Following the explanation of our stationarity testing procedure in Section III-C2, we employ an Augmented-Dickey-Fuller test to determine if our metric time series has a unit root. Testing the monthly stationarity of the MEC of our asset graph, we obtain an ADF test statistic of -6.84, and a p-value of approximately 0. We can reject the null hypothesis of the process having a unit root. Thus, we can conclude with reasonable confidence that this process is stationary. Next, the Philips-Perron test is performed on the same series, we obtain a test statistic of -15.074 and a p-value of $8.63 \cdot 10^{-26}$. Thus we can reject H_0, we again conclude the time series is stationary. Additionally, we performed more informal tests to ensure the explicit stationarity of the mean and variance by “chunking” the process.

2) Transitivity: Next, we perform the ADF test on the transitivity of the asset graph to confirm stationarity. We obtain a test statistic of -6.618 and a p-value of approximately zero. We can reject H_0 and conclude that this process is stationary. We also perform the Philips-Perron test on this metric, obtaining a test statistic of -13.89, and a p-value of

Variable	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
C	**-10.944*** **	**-10.370*** **	**-7.896*** **	**-7.339*** **	**-6.819*** **	**-6.214** **	**-10.026*** **	**-6.524** **	**-9.812*** **	**-9.343*** **	**-8.553*** **	**-7.955*** **	**-8.269*** **
	(2.481)	(2.379)	(2.142)	(2.475)	(2.140)	(2.467)	(2.717)	(2.708)	(2.477)	(2.476)	(2.212)	(2.445)	(2.799)
MEC(-1)	**53.083*** **	**50.869*** **	**36.898*** **	**29.650** **	**28.940** **	19.899	**47.800*** **	**26.939* **	**46.811*** **	**44.528*** **	**40.721*** **	**36.914*** **	**39.291** **
	(13.890)	(13.255)	(12.162)	(13.721)	(11.614)	(13.187)	(14.960)	(14.628)	(13.697)	(13.921)	(12.331)	(13.560)	(15.447)
AR(-1)	**0.499*** **	**0.473*** **	**0.533*** **	**0.487*** **	**0.405*** **	**0.472*** **	**0.402*** **	**0.473*** **	**0.589*** **	**0.585*** **	**0.628*** **	**0.610*** **	**0.589*** **
	(0.078)	(0.071)	(0.088)	(0.077)	(0.086)	(0.074)	(0.057)	(0.068)	(0.070)	(0.076)	(0.072)	(0.060)	(0.068)
AR(-2)	**0.213*** **	**0.200*** **	**0.238*** **	0.077	**0.246*** **	0.047	**0.165** **	**0.191** **	0.105	**0.138** **	0.094	**0.107* **	**0.176** **
	(0.056)	(0.052)	(0.075)	(0.064)	(0.068)	(0.075)	(0.070)	(0.081)	(0.073)	(0.068)	(0.073)	(0.062)	(0.075)
AR(-3)	**0.104* **	**0.161*** **	0.057	**0.106* **	0.047	0.057	**0.186*** **	0.088	**0.121* **	0.090	**0.105* **	**0.091* **	0.061
	(0.056)	(0.058)	(0.063)	(0.056)	(0.049)	(0.071)	(0.058)	(0.082)	(0.062)	(0.059)	(0.060)	(0.051)	(0.057)
R-squared	0.491	0.520	0.543	0.325	0.341	0.258	0.362	0.450	0.506	0.518	0.549	0.537	0.581

Variable	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
C	**-10.284*** **	**-10.601*** **	**-5.850** **	**-9.207*** **	-1.386	**-4.716** **	**-5.499*** **	**-5.111** **	**-4.888* **	-2.173	-3.630	**-10.080*** **	**-4.601*** **
	(2.463)	(2.307)	(2.573)	(2.480)	(2.385)	(1.903)	(1.924)	(2.138)	(2.538)	(2.342)	(2.410)	(2.530)	(1.693)
MEC(-1)	**50.687*** **	**52.967*** **	22.639	**40.300*** **	-3.362	**19.085* **	**24.382** **	**21.367* **	15.774	0.262	12.104	**48.092*** **	13.220
	(13.658)	(12.716)	(14.052)	(13.884)	(12.917)	(10.480)	(10.445)	(11.970)	(13.958)	(12.923)	(13.226)	(13.972)	(9.316)
AR(-1)	**0.586*** **	**0.593*** **	**0.488*** **	**0.521*** **	**0.410*** **	**0.518*** **	**0.538*** **	**0.428*** **	**0.468*** **	**0.388*** **	**0.520*** **	**0.596*** **	**0.407*** **
	(0.079)	(0.069)	(0.078)	(0.080)	(0.067)	(0.068)	(0.098)	(0.073)	(0.072)	(0.066)	(0.070)	(0.068)	(0.065)
AR(-2)	**0.179*** **	**0.141** **	**0.148** **	**0.159** **	**0.198** **	**0.175** **	**0.136* **	**0.212*** **	0.100	**0.226*** **	**0.182** **	0.091	**0.150** **
	(0.067)	(0.062)	(0.065)	(0.066)	(0.083)	(0.083)	(0.078)	(0.071)	(0.073)	(0.074)	(0.072)	(0.072)	(0.063)
AR(-3)	0.085	**0.124** **	0.047	0.050	0.045	**0.103* **	**0.151** **	**0.163*** **	**0.111* **	0.096	0.076	**0.124** **	0.110
	(0.053)	(0.052)	(0.070)	(0.055)	(0.057)	(0.061)	(0.068)	(0.062)	(0.061)	(0.070)	(0.053)	(0.062)	(0.071)
R-squared	0.569	0.581	0.354	0.399	0.318	0.512	0.558	0.477	0.332	0.380	0.490	0.499	0.307

Window	Benchmark	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
ROL	AR(3)	**1.921** **	**1.912** **	0.614	-0.907	0.211	-0.502	2.500	-0.970	**1.779** **	**2.158** **	**1.727** **	0.623	**0.875* **
	AR(3)VIX	0.308	0.454	**1.041* **	-0.947	0.429	-1.853	**2.825*** **	-0.113	0.619	**1.046* **	0.191	**0.815* **	**1.317** **
	AR(6)	2.848	**3.174* **	0.719	-0.326	0.420	0.001	**2.309*** **	-0.360	**1.548** **	**2.148** **	2.456	0.735	**1.931** **
	AR(6)VIX	0.295	0.674	**0.946* **	-0.796	0.575	-1.832	**3.069*** **	0.562	0.573	**0.973* **	0.061	**0.798* **	**1.571** **
REC	AR(3)	**2.029** **	**2.227*** **	0.599	-0.750	0.239	-0.485	**2.544*** **	-1.028	**1.826** **	**2.170*** **	**1.989** **	**0.832* **	**1.022* **
	AR(3)VIX	0.378	**0.861* **	**1.461** **	-0.656	0.437	-1.963	**2.916*** **	-0.754	0.616	**0.961* **	0.292	**0.832* **	**1.418** **
	AR(6)	**3.195*** **	**3.595*** **	**0.854* **	-0.183	0.538	-0.032	**2.264*** **	-0.733	**1.680** **	**2.280*** **	**2.870*** **	**1.060* **	**2.141*** **
	AR(6)VIX	0.539	**1.043* **	**1.854** **	-0.538	0.677	-1.840	**2.786*** **	-0.369	**0.775* **	**1.170** **	0.225	**0.870* **	**1.625** **

Window	Benchmark	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
ROL	AR(3)	**2.485** **	2.843	0.478	**1.217** **	-0.262	**0.936* **	**2.625** **	0.707	0.344	-0.217	-0.724	**1.896** **	-0.009
	AR(3)VIX	0.256	**0.951* **	2.861	-0.087	-0.572	**1.195** **	**0.778* **	0.183	-0.201	-0.168	-1.478	**0.788* **	-0.006
	AR(6)	2.611	2.853	0.497	**0.991* **	-0.229	**1.074* **	**2.279* **	**0.823* **	0.275	-0.263	-0.552	**1.591** **	-0.117
	AR(6)VIX	0.045	**0.792* **	**1.990** **	0.037	-0.689	**1.171** **	**1.636** **	0.059	-0.203	-0.296	-1.446	0.724	-0.218
REC	AR(3)	**2.705*** **	**2.803*** **	0.232	**1.164** **	-0.202	**0.925* **	**1.876** **	0.573	0.427	-0.252	-0.553	**1.894** **	-0.068
	AR(3)VIX	0.523	0.636	**2.468*** **	0.034	-0.478	**1.222** **	**1.063* **	0.326	-0.114	-0.186	-1.433	**0.795* **	0.244
	AR(6)	**2.798*** **	**2.519*** **	0.211	**1.133** **	-0.183	**1.013* **	**1.778** **	**0.726* **	0.344	-0.273	-0.371	**1.707** **	-0.109
	AR(6)VIX	0.532	**0.828* **	**1.888** **	0.216	-0.695	**1.248** **	**1.538** **	0.260	-0.033	-0.299	-1.284	**0.960* **	0.044

TABLE V: In-sample (upper tables) and out-of-sample (lower tables) results for MEC. The regression is performed on our own data for a period starting from 2001-07 until 2019-09.

$5.88 \cdot 10^{-26}$. We can reject the null hypothesis of a unit-root and conclude with reasonable confidence that the time-series is stationary. Again, we also perform more informal tests to verify the stationarity of the processes mean and variance.

C. *In-sample Results*

1) Mean Eigenvector Centrality: The results of in-sample regressions are presented in Table V. Our investigation reveals 19 out of 26 significant estimated coefficients for mean eigenvector centrality terms. From the significant coefficients we find 19 that are significant at the 10% level, 16 that are significant at the 5%, and 12 which are significant at the 1% level. Additionally, it is evident that all of the AR(-1) coefficients are significant at the 1% level. Furthermore, although we still observe a total of 20 estimated coefficients for AR(-2) terms that are significant, they do expose substantially reduced significance, which is inline with our expectations. Moving forward, it is evident that if the estimated coefficients for the constant are not significant, the estimated coefficient for the MEC is not significant as well.

The high proportion of significant in-sample estimated coefficients for MEC implies that the lagged MEC of the global correlation network "granger causes" RLV for the majority of the global indices. Given that the MEC is indicative of the influence of the most connected nodes upon each other, it seems that this measure of influence is useful for making out-of-sample volatility predictions.

Moving forward, we also observe the inter-country patterns which emerge in our regressions tests. There is substantial geographic significance clustering in the Americas: the SPX, CCMP, TSX, MEXBOL, IBOVE and MERVAL. This behaviour is not something that was observed in the original paper by Magner et al., with the majority of their North/South American indices appearing to be insignificant in-sample for both VMSTL and VPMFGL.

It is apparent that MEC is also a great metric for European

Variable	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
C	**1.852***	**1.804****	1.021	-0.259	0.059	**-1.592***	1.254	-0.307	1.400	1.175	1.204	1.039	1.157
	(0.973)	(0.898)	(0.800)	(0.910)	(0.700)	(0.902)	(0.911)	(0.900)	(0.945)	(1.001)	(0.883)	(0.873)	(0.945)
TRANS(-1)	**-3.671*****	**-3.443*****	**-2.574*****	**-1.968****	**-1.914****	-1.169	**-3.131*****	-1.540	**-3.149*****	**-2.831*****	**-2.750*****	**-2.650*****	**-2.670****
	(0.967)	(0.928)	(0.791)	(0.928)	(0.796)	(0.945)	(1.027)	(1.032)	(0.999)	(1.039)	(0.943)	(0.941)	(1.060)
AR(-1)	**0.521*****	**0.491*****	**0.549*****	**0.502*****	**0.417*****	**0.477*****	**0.411*****	**0.479*****	**0.610*****	**0.603*****	**0.649*****	**0.634*****	**0.611*****
	(0.082)	(0.073)	(0.090)	(0.078)	(0.086)	(0.074)	(0.057)	(0.069)	(0.072)	(0.080)	(0.075)	(0.062)	(0.068)
AR(-2)	**0.200*****	**0.189*****	**0.227*****	0.068	**0.236*****	0.042	**0.152****	**0.188****	0.090	**0.130***	0.081	0.094	**0.163****
	(0.056)	(0.052)	(0.075)	(0.065)	(0.068)	(0.076)	(0.069)	(0.081)	(0.073)	(0.068)	(0.071)	(0.063)	(0.075)
AR(-3)	**0.114****	**0.169*****	0.066	**0.112****	0.053	0.063	**0.191*****	0.091	**0.130****	0.094	**0.111***	**0.097***	0.067
	(0.057)	(0.058)	(0.063)	(0.056)	(0.050)	(0.071)	(0.059)	(0.082)	(0.062)	(0.059)	(0.059)	(0.050)	(0.057)
R-squared	0.489	0.517	0.543	0.323	0.339	0.256	0.360	0.448	0.503	0.513	0.546	0.536	0.580

Variable	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
C	**1.936****	**2.131****	-0.739	0.141	**-2.604*****	-0.384	0.144	-0.098	**-1.626***	**-2.280****	-0.937	1.471	**-1.978****
	(0.957)	(0.861)	(0.849)	(0.981)	(0.855)	(0.755)	(0.606)	(0.856)	(0.969)	(0.983)	(0.867)	(0.945)	(0.809)
TRANS(-1)	**-3.517*****	**-3.610*****	-1.184	**-2.383****	0.698	-1.016	**-1.428****	-1.329	-0.476	0.177	-0.582	**-3.266*****	-0.300
	(1.028)	(0.937)	(0.946)	(1.003)	(0.916)	(0.786)	(0.697)	(0.878)	(1.018)	(0.975)	(0.953)	(1.003)	(0.655)
AR(-1)	**0.613*****	**0.622*****	**0.493*****	**0.530*****	**0.401*****	**0.523*****	**0.546*****	**0.432*****	**0.470*****	**0.385*****	**0.521*****	**0.619*****	**0.403*****
	(0.082)	(0.072)	(0.079)	(0.082)	(0.065)	(0.068)	(0.098)	(0.073)	(0.072)	(0.067)	(0.070)	(0.070)	(0.066)
AR(-2)	**0.162****	**0.126****	**0.142****	**0.153****	**0.198****	**0.170****	**0.128***	**0.211*****	0.094	**0.225*****	**0.178****	0.075	**0.153****
	(0.068)	(0.063)	(0.065)	(0.066)	(0.082)	(0.083)	(0.078)	(0.071)	(0.073)	(0.074)	(0.071)	(0.072)	(0.061)
AR(-3)	**0.092***	**0.131****	0.049	0.054	0.045	**0.105***	**0.154****	**0.166*****	**0.111***	0.097	0.079	**0.134****	0.104
	(0.053)	(0.053)	(0.070)	(0.055)	(0.057)	(0.061)	(0.069)	(0.062)	(0.062)	(0.070)	(0.053)	(0.062)	(0.072)
R-squared	0.567	0.577	0.351	0.393	0.319	0.510	0.555	0.476	0.330	0.380	0.489	0.495	0.303

Window	Benchmark	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
ROL	AR(3)	**1.903****	**1.798****	0.654	-0.709	0.096	-0.516	2.594	-0.758	**1.539****	**1.427****	**0.954***	0.381	0.528
	AR(3)VIX	0.654	**0.994***	**1.377****	-0.905	0.327	-1.891	**3.038*****	-0.526	0.404	0.657	-0.142	**0.913***	**1.722****
	AR(6)	**3.032***	**3.322****	0.746	-0.102	0.362	-0.063	**2.527*****	-0.197	**1.311****	**1.494****	**1.623****	0.607	**1.721****
	AR(6)VIX	0.760	**1.391****	**1.183****	-0.785	0.469	-1.925	**3.384*****	0.164	0.325	0.582	-0.301	**0.888***	**2.151****
REC	AR(3)	**2.052****	**2.193*****	**0.704***	-0.531	0.212	-0.491	**2.499*****	-0.776	**1.684****	**1.598****	**1.439****	**0.708***	**0.796***
	AR(3)VIX	**0.819***	**1.662****	**2.074****	-0.610	0.477	-2.023	**3.107*****	-1.061	0.565	**0.820***	0.088	**1.010***	**1.920****
	AR(6)	**3.371*****	**3.804*****	**0.939***	0.124	0.519	-0.062	**2.376*****	-0.513	**1.546****	**1.760****	**2.244*****	**1.094****	**2.082****
	AR(6)VIX	**1.148****	**2.068****	**2.408*****	-0.534	0.632	-1.993	**3.033*****	-0.633	0.677	**1.027***	-0.012	**1.134****	**2.373*****

Window	Benchmark	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
ROL	AR(3)	**1.945****	**2.150****	0.168	**0.803***	-0.077	0.374	**1.806****	0.502	-0.248	-0.013	-0.869	**1.705****	-0.216
	AR(3)VIX	-0.185	0.229	2.891	-0.328	-0.617	**1.131***	**0.805***	0.124	-0.187	-0.083	-1.600	0.555	-0.245
	AR(6)	**2.122****	**2.080****	0.238	0.651	0.011	0.654	**1.579****	0.670	-0.517	-0.120	-0.695	**1.407****	-0.274
	AR(6)VIX	-0.399	0.064	**1.951****	-0.254	-0.586	**1.092***	**1.372****	-0.030	-0.140	-0.257	-1.583	0.457	-0.354
REC	AR(3)	**2.382*****	**2.346*****	0.056	**0.821***	0.057	0.506	**1.414****	0.477	-0.182	-0.123	-0.553	**1.799****	-0.156
	AR(3)VIX	0.390	0.101	**2.579*****	-0.023	-0.517	**1.290****	**1.588****	0.496	-0.147	-0.202	-1.422	**0.744***	-0.014
	AR(6)	**2.553*****	**1.999****	0.032	**0.823***	0.068	**0.721***	**1.293****	0.663	-0.427	-0.170	-0.412	**1.627****	-0.186
	AR(6)VIX	0.413	0.215	**1.811****	0.098	-0.616	**1.346****	**1.786****	0.399	-0.082	-0.346	-1.350	**0.858***	-0.166

TABLE VI: In-sample (upper tables) and out-of-sample (lower tables) results for TRANS. The regression is performed on our own data for a period starting from 2001-07 until 2019-09.

indices, as nearly all of them appear to have significant estimated coefficients for MEC – UKX, CAC, DAX, IBEX, AEX, OMX and MIB. The coefficient estimated for IBEX appears to be the least significant, what is further revealed in lack of significance in the out-of-sample results only for that index. We hypothesise that this additional geographic clustering in Europe and the Americas is due to both the VMSTL and VPMFGL significantly reducing the number of edges in our complete graph, causing information about intra-continential patterns to be lost. Specifically, the asset graph we use to compute transitivity has significantly more edges than both the VPMFGL and the VMSTL. The MEC then appears to effectively quantify this information from the asset graph, exposing intra-continental patterns.

2) Transitivity: The in-sample results are also very promising for the transitivity (TRANS) metric. The estimated coefficients for TRANS are significant for 16 out of 26 indices. The significance is present for the exact same indices as for MEC, except from HSI, TWSE and IGBVL, which are significant only at the 10% level for MEC. While IGBVL is not significant in the out-of-sample results for either of the metrics, the HSI is significant for most of the out-of-sample for both MEC and TRANS metrics. It is evident that the significance of both of our metrics very closely aligns with each other. Overall we find that 16 that are significant at the 10% and 5% levels and 11 that are significant at the 1% level. Additionally, all of the AR(-1) coefficients are significant. The substantial number of significant estimated coefficients implies Granger causality between lagged transitivity of our asset graph and the RLV of that particular country. Specifically, a greater proportion of closed triplets in our asset graph implies a larger RLV in the next period.

D. Out-of-sample Results

1) Mean Eigenvector Centrality: Considering recursive time windows, for the AR(3) model the MEC significantly improves the prediction for 14 indices, 5 of which are significant at the 1% level. For the AR(6) model we observe 16 significantly decreased squared errors for models with the MEC metric as compared to the benchmark model, 8 of which are significant at 1% level. Both models achieve higher forecasting power in general than the VMSTL or VPMFGL.

The significance of the MEC metric decreases for nested models benchmarked against the VIX, which is not surprising given that VIX is a powerful predictor. For the AR(3)VIX model we detect 10 indices for which the MEC metric contributes significant predictive power to the model, while the same is observed for 12 indices for the AR(6)VIX model. Use of the MEC instead of the VMSTL as a forecasting tool leads to greater robustness to benchmarks including the VIX and ultimately reduced squared error losses.

For each of the benchmark models, the rolling window forecasts provide worse results than the recursive window forecasts, with fewer of the proposed models performing significantly better than the benchmark models. Broadly, this is to be expected, regression coefficients fitted on a greater number of training samples should more effective predictors of future outcomes.

2) Transitivity: Using recursive windows, TRANS applied to the asset graph obtains 14 and 15 values significant at the 10% level for the AR(3) and AR(6) models, respectively. The results also remain relatively robust to the VIX benchmarks, we achieve 11 significant values at 10% on both the AR(3)VIX and AR(6)VIX benchmark. These results are notably better than those obtained in our VMSTL reproduction. Analogous to MEC, out-of-sample results for rolling windows lead to reduced significance as compared to the recursive window scheme.

VII. Conclusion

After extensive systematic experimentation, we were able to recreate the VMSTL results from Magner et al. Even given the scale of ambiguity in the method presented, we are confident in the reproduction of the VMSTL given how well the in-sample coefficients align with the originals. We were not able to reproduce the VPMFGL estimated coefficients as precisely as those of the VMSTL. However, we can say with a degree of certainty that the inclusion of VMSTL and, to a lesser extent, VPMFGL does marginally improve one-step-ahead volatility forecasts, at least against the benchmarks presented. Also of note is the continental heterogeneity which emerges when using these metrics.

The aforementioned lack of robustness, both to different time periods and to heterogeneous global indices motivated us to search for new graph metrics. We observe that the asset graph retains a lot of potentially useful additional edges compared to the PMFG and MST. The question then becomes, how do we choose to quantify this new, more complex topology? Clearly, we need a measure more nuanced than the sum of edge weights. First, transitivity is used to determine the extent to which nodes tend to cluster together, as measured by the fraction of connected triples of nodes which ultimately form triangles. We find that a greater proportion of closed triples in our asset graph implies a larger RLV in the next period. Next, we use the MEC to characterise the extent to which the nodes in the asset graph are influential upon each other. We find that this predictor performs comparably to the transitivity metric out-of-sample. Certainly, both of the proposed metrics are substantially more effective than the VMSTL and VPMFGL.

We believe that this work demonstrates that there are significant opportunities to improve volatility forecasting through the use of network metrics. Although additional work will need to be done to verify the robustness of these metrics against more difficult benchmarks, we believe this work represents a first step towards developing more nuanced graph techniques which can capture the complex inter-index correlation patterns which partially drive global volatility trends.

A. Future Work

As briefly mentioned in the conclusion section, one of the notable further work directions would be the usage of more recent and better benchmarks in order to assess the performance of our metrics. For example, in a 2015 publication on realized volatility forecasting, Vortelinos suggests that GARCH and HAR models are currently considered the best models for realized volatility forecasting [42]. Hence, it would be interesting to compare the investigated MEC and TRANS models to these more advanced benchmarks.

Furthermore, until now our investigation has focused on one-step forecasting, the results of which appeared to be promising for both MEC and TRANS. Therefore, it seems to be necessary to follow this research with an investigation into whether these metrics also provide predictive power to the multi-step ahead forecasting framework, such as the approach presented by Santos and Ziegelmann [35].

Lastly, it could be interesting to connect our research area with information theory. A first idea for future work in this domain would be to use the mutual information (MI) between two variables as defined by Shannon [38]; in our case the variables could be the monthly returns of two indices. The nature of this metric could make it an interesting alternative to the Pearson correlation. Future work could thus be to compute the MI matrix instead of the correlation matrix and then continue with our existing metrics. We have taken some first steps towards implementing this, however, thus far the results do not appear to be as promising as those achieved with our graph metrics. We imagine this to be an interesting area of future research.

Related to this idea, another area for future work could be to explore the power of the information dissipation length (IDL) metric as defined by Quax et al. [33]. They define IDL as the "halftime of the mutual information between [interest rate swaps (IRS) with] maturity 1 and i for increasing i" [33]. They found this metric to have potential as an early-warning signal for crises in financial markets. It could be interesting to explore this metric in our context; here it could be feasible to compute the IDL as the halftime of a decaying exponential fitted to the sorted MI between indices, or the IDL of a particular index with itself over time. Given that this metric has already been shown to be useful as an early-warning signal, it could be very interesting to explore its power in forecasting realised volatility.

REFERENCES

[1] Francesco Audrino, Lorenzo Camponovo, and Constantin Roth. "Testing the lag structure of assets' realized volatility dynamics". In: *Available at SSRN 2549063* (2015).

[2] Ole E Barndorff-Nielsen and Neil Shephard. "Estimating quadratic variation using realized variance". In: *Journal of Applied econometrics* 17.5 (2002), pp. 457–477.

[3] Stefano Boccaletti et al. "Complex networks: Structure and dynamics". In: *Physics reports* 424.4-5 (2006), pp. 175–308.

[4] Giovanni Bonanno, Nicolas Vandewalle, and Rosario N. Mantegna. "Taxonomy of stock market indices". In: *Phys. Rev. E* 62 (6 Dec. 2000), R7615–R7618. DOI: 10.1103/PhysRevE.62.R7615. URL: https://link.aps.org/doi/10.1103/PhysRevE.62.R7615.

[5] Giovanni Bonanno et al. "Topology of correlation-based minimal spanning trees in real and model markets". In: *Physical Review E* 68.4 (2003), p. 046130.

[6] Otakar Borvka. "Přıspěvek k řešenı otázky ekonomické stavby elektrovodnıch sıtı". In: *Elektrotechnick obzor* 15 (1926), pp. 153–154.

[7] Taufiq Choudhry. "Stock market volatility and the US consumer expenditure". In: *Journal of Macroeconomics* 25.3 (2003), pp. 367–385.

[8] Todd E Clark and Michael W McCracken. "Tests of equal forecast accuracy and encompassing for nested models". In: *Journal of econometrics* 105.1 (2001), pp. 85–110.

[9] David A Dickey and Wayne A Fuller. "Distribution of the estimators for autoregressive time series with a unit root". In: *Journal of the American statistical association* 74.366a (1979), pp. 427–431.

[10] Prasanna Gai and Sujit Kapadia. "Contagion in financial networks". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466.2120 (2010), pp. 2401–2423.

[11] Ronald L Graham and Pavol Hell. "On the history of the minimum spanning tree problem". In: *Annals of the History of Computing* 7.1 (1985), pp. 43–57.

[12] Thomas Guhr and Bernd Kälber. "A new method to estimate the noise in financial correlation matrices". In: *Journal of Physics A: Mathematical and General* 36.12 (2003), p. 3009.

[13] Joseph B Kruskal. "On the shortest spanning subtree of a graph and the traveling salesman problem". In: *Proceedings of the American Mathematical society* 7.1 (1956), pp. 48–50.

[14] Vishwas Kukreti et al. "A perspective on correlation-based financial networks and entropy measures". In: *Frontiers in Physics* 8 (2020), p. 323.

[15] Bentian Li and Dechang Pi. "Analysis of global stock index data during crisis period via complex network approach". In: *PloS one* 13.7 (2018), e0200600.

[16] Xiao Fan Liu and Chi K Tse. "A complex network perspective of world stock markets: synchronization and volatility". In: *International Journal of Bifurcation and Chaos* 22.06 (2012), p. 1250142.

[17] Andrew W Lo and A Craig MacKinlay. "An econometric analysis of nonsynchronous trading". In: *Journal of Econometrics* 45.1-2 (1990), pp. 181–211.

[18] Nicolás S Magner et al. "The Volatility Forecasting Power of Financial Network Analysis". In: *Complexity* 2020 (2020).

[19] Kiran Manda et al. "Stock market volatility during the 2008 financial crisis". In: *GLUCKSMAN FELLOWSHIP PROGRAM STUDENT RESEARCH REPORTS: 2009-2010* 87 (2010).

[20] Rosario N Mantegna. "Hierarchical structure in financial markets". In: *The European Physical Journal B-Condensed Matter and Complex Systems* 11.1 (1999), pp. 193–197.

[21] Hendrik Mehlhorn and Falk Schreiber. "Small-World Property". In: *Encyclopedia of Systems Biology*. Ed. by Werner Dubitzky et al. New York, NY: Springer New York, 2013, pp. 1957–1959. ISBN: 978-1-4419-9863-7. DOI: 10.1007/978-1-4419-9863-7_2. URL: https://doi.org/10.1007/978-1-4419-9863-7_2.

[22] Whitney K Newey and Kenneth D West. *A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix*. Tech. rep. National Bureau of Economic Research, 1986.

[23] Mark EJ Newman. "The mathematics of networks". In: *The new palgrave encyclopedia of economics* 2.2008 (2008), pp. 1–12.

[24] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. "Random graph models of social networks". In: *Proceedings of the national academy of sciences* 99.suppl 1 (2002), pp. 2566–2572.

[25] Ashadun Nobi et al. "Correlation and network topologies in global and local stock indices". In: *Physics Letters A* 378.34 (2014), pp. 2482–2489.

[26] *Not-for-Publication Appendix to "Tests of Equal Forecast Accuracy and Encompassing for Nested Models"*. May 2000. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.161.4290&rep=rep1&type=pdf.

[27] J-P Onnela, Kimmo Kaski, and Janos Kertész. "Clustering and information in correlation based financial networks". In: *The European Physical Journal B* 38.2 (2004), pp. 353–362.

[28] J-P Onnela et al. "Dynamic asset trees and portfolio analysis". In: *The European Physical Journal B-Condensed Matter and Complex Systems* 30.3 (2002), pp. 285–288.

[29] Jukka-Pekka Onnela et al. "Asset trees and asset graphs in financial markets". In: *Physica Scripta* 2003.T106 (2003), p. 48.

[30] Peter CB Phillips and Pierre Perron. "Testing for a unit root in time series regression". In: *Biometrika* 75.2 (1988), pp. 335–346.

[31] Ser-Huang Poon and Clive WJ Granger. "Forecasting volatility in financial markets: A review". In: *Journal of economic literature* 41.2 (2003), pp. 478–539.
[32] Robert Clay Prim. "Shortest connection networks and some generalizations". In: *The Bell System Technical Journal* 36.6 (1957), pp. 1389–1401.
[33] Rick Quax, Drona Kandhai, and Peter MA Sloot. "Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series". In: *Scientific reports* 3.1 (2013), pp. 1–7.
[34] *SP 500 (GSPC) Historical Data*. May 2021. URL: https://finance.yahoo.com/quote/5EGSPC/history?period1=1180051200&period2=1241222400&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true.
[35] Douglas G Santos and Flavio A Ziegelmann. "Volatility forecasting via MIDAS, HAR and their combination: An empirical comparative study for IBOVESPA". In: *Journal of Forecasting* 33.4 (2014), pp. 284–299.
[36] Thomas Schank and Dorothea Wagner. "Approximating clustering coefficient and transitivity." In: *Journal of Graph Algorithms and Applications* 9.2 (2005), pp. 265–275.
[37] Skipper Seabold and Josef Perktold. "statsmodels: Econometric and statistical modeling with python". In: *9th Python in Science Conference*. 2010.
[38] Claude E Shannon. "A mathematical theory of communication". In: *The Bell system technical journal* 27.3 (1948), pp. 379–423.
[39] Bruno Solnik, Cyril Boucrelle, and Yann Le Fur. "International market correlation and volatility". In: *Financial analysts journal* 52.5 (1996), pp. 17–34.
[40] Yong Tang et al. "Complexities in financial network topological dynamics: Modeling of emerging and developed stock markets". In: *Complexity* 2018 (2018).
[41] Michele Tumminello et al. "A tool for filtering information in complex systems". In: *Proceedings of the National Academy of Sciences* 102.30 (2005), pp. 10421–10426.
[42] Dimitrios I Vortelinos. "Forecasting realized volatility: HAR against Principal Components Combining, neural networks and GARCH". In: *Research in international business and finance* 39 (2017), pp. 824–839.
[43] Hui Wang. "VIX and volatility forecasting: A new insight". In: *Physica A: Statistical Mechanics and its Applications* 533 (2019), p. 121951.
[44] Xin Yang et al. "Dynamic properties of foreign exchange complex network". In: *Mathematics* 7.9 (2019), p. 832.

VIII. APPENDIX

Tables VII and VIII represent the regression results when applying VMSTL and VPMFGL to our data and chosen time period.

Variable	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
C	**-1.466*****	**-1.262*****	**-1.299*****	**-2.031*****	**-1.607*****	**-2.601*****	**-1.337*****	**-1.542*****	**-1.484*****	**-1.390*****	**-1.291*****	**-1.307*****	**-1.159*****
	(0.407)	(0.344)	(0.327)	(0.396)	(0.282)	(0.454)	(0.314)	(0.353)	(0.318)	(0.282)	(0.276)	(0.267)	(0.232)
VMSTL(-1)	0.233	0.125	0.043	0.364	0.440	0.196	0.578	0.225	-0.353	-0.609	-0.380	-0.252	-0.294
	(0.351)	(0.318)	(0.349)	(0.354)	(0.302)	(0.331)	(0.356)	(0.386)	(0.368)	(0.381)	(0.418)	(0.349)	(0.356)
AR(-1)	**0.477*****	**0.436*****	**0.513*****	**0.518*****	**0.440*****	**0.487*****	**0.433*****	**0.489*****	**0.493*****	**0.454*****	**0.533*****	**0.547*****	**0.532*****
	(0.077)	(0.074)	(0.103)	(0.075)	(0.087)	(0.075)	(0.063)	(0.066)	(0.073)	(0.087)	(0.083)	(0.063)	(0.076)
AR(-2)	**0.188*****	**0.195*****	**0.228****	0.027	**0.185****	0.021	0.116	**0.179****	**0.153***	**0.244*****	**0.156***	**0.157****	**0.218****
	(0.070)	(0.062)	(0.092)	(0.063)	(0.072)	(0.082)	(0.073)	(0.084)	(0.087)	(0.083)	(0.091)	(0.068)	(0.087)
AR(-3)	0.100	**0.155****	0.053	**0.098***	0.047	0.059	**0.178*****	0.085	**0.119***	0.072	0.095	0.077	0.052
	(0.061)	(0.062)	(0.063)	(0.056)	(0.049)	(0.073)	(0.058)	(0.083)	(0.062)	(0.059)	(0.061)	(0.051)	(0.056)
R-squared	0.461	0.490	0.527	0.313	0.332	0.252	0.344	0.444	0.487	0.503	0.534	0.523	0.566

Variable	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
C	**-1.228*****	**-1.144*****	**-1.752*****	**-2.065*****	**-1.932*****	**-1.296*****	**-1.090*****	**-1.294*****	**-2.034*****	**-2.122*****	**-1.449*****	**-1.536*****	**-2.243*****
	(0.274)	(0.297)	(0.322)	(0.344)	(0.383)	(0.327)	(0.247)	(0.286)	(0.383)	(0.440)	(0.285)	(0.331)	(0.539)
VMSTL(-1)	-0.384	-0.433	0.030	-0.509	**-0.686***	-0.481	-0.474	-0.460	-0.055	-0.383	-0.221	-0.343	**-0.561***
	(0.443)	(0.356)	(0.363)	(0.379)	(0.401)	(0.310)	(0.344)	(0.317)	(0.348)	(0.345)	(0.413)	(0.374)	(0.286)
AR(-1)	**0.482*****	**0.482*****	**0.485*****	**0.415*****	**0.350*****	**0.442*****	**0.464*****	**0.355*****	**0.465*****	**0.355*****	**0.482*****	**0.501*****	**0.337*****
	(0.096)	(0.077)	(0.085)	(0.092)	(0.067)	(0.063)	(0.105)	(0.078)	(0.071)	(0.066)	(0.081)	(0.073)	(0.062)
AR(-2)	**0.248*****	**0.210*****	**0.138***	**0.231*****	**0.265*****	**0.233*****	**0.197****	**0.263*****	0.095	**0.254*****	**0.212****	0.138	**0.221*****
	(0.088)	(0.078)	(0.076)	(0.078)	(0.080)	(0.089)	(0.090)	(0.081)	(0.078)	(0.080)	(0.090)	(0.089)	(0.061)
AR(-3)	0.073	**0.111***	0.045	0.040	0.051	0.103	**0.148****	**0.165*****	**0.109***	0.101	0.075	**0.120***	0.097
	(0.055)	(0.056)	(0.070)	(0.058)	(0.056)	(0.063)	(0.072)	(0.063)	(0.061)	(0.071)	(0.054)	(0.062)	(0.070)
R-squared	0.550	0.554	0.348	0.386	0.332	0.512	0.554	0.476	0.329	0.384	0.489	0.478	0.316

Window	Benchmark	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
ROL	AR(3)	0.004	-0.031	-0.150	0.349	**0.850***	0.145	**0.963***	0.467	-0.482	0.001	-0.687	-0.631	-0.093
	AR(3)VIX	0.042	0.009	0.211	-0.042	0.226	-0.164	0.320	-0.039	-0.321	-0.327	-0.901	-0.005	0.144
	AR(6)	-0.007	-0.023	-0.148	0.567	**0.784***	0.152	**1.274****	0.323	-0.426	0.193	-0.570	-0.886	-0.171
	AR(6)VIX	0.100	0.122	0.266	0.019	0.233	-0.159	0.346	-0.087	-0.344	-0.286	-0.793	-0.343	-0.259
REC	AR(3)	-0.009	-0.074	-0.267	0.120	0.561	-0.010	0.679	0.130	-0.310	0.296	-0.026	-0.223	0.092
	AR(3)VIX	-0.050	-0.112	0.208	-0.259	0.038	-0.211	0.270	-0.227	-0.329	-0.156	-0.297	0.131	0.158
	AR(6)	0.071	0.012	-0.258	0.518	0.613	0.103	**1.188****	0.099	-0.214	0.292	0.020	-0.416	-0.102
	AR(6)VIX	-0.011	-0.058	0.201	-0.181	-0.095	-0.189	0.475	-0.208	-0.296	-0.158	-0.253	-0.028	-0.008

Window	Benchmark	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
ROL	AR(3)	-0.430	-0.359	-0.194	-0.504	0.194	-0.153	0.619	-0.217	-0.943	-0.521	-0.421	-0.475	0.525
	AR(3)VIX	-0.288	-0.190	-0.223	-0.884	-0.724	-0.227	0.425	0.621	-0.550	-0.050	-0.013	-0.279	0.153
	AR(6)	-0.346	-0.174	-0.254	-0.534	0.270	-0.132	0.465	-0.368	-1.144	-0.659	-0.765	-0.392	0.641
	AR(6)VIX	-0.236	-0.149	-0.171	-0.777	-0.510	-0.294	0.595	0.529	-0.447	-0.099	-0.040	-0.312	0.239
REC	AR(3)	-0.092	0.018	-0.492	-0.336	0.573	0.160	0.089	-0.449	-1.051	-0.604	-0.339	-0.310	**0.727***
	AR(3)VIX	-0.148	-0.237	-0.479	-0.868	-0.523	-0.141	-0.368	-0.135	-0.883	-0.199	-0.134	-0.293	0.269
	AR(6)	-0.078	0.182	-0.595	-0.387	0.520	0.255	-0.158	-0.571	-1.005	-0.595	-0.470	-0.212	**0.775***
	AR(6)VIX	-0.143	-0.188	-0.586	-0.887	-0.335	-0.151	-0.398	-0.102	-0.842	-0.167	-0.047	-0.253	0.382

TABLE VII: In-sample (upper tables) and out-of-sample (lower tables) results for VMSTL. The regression is performed on our own data for a period starting from 2001-07 until 2019-09.

Variable	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
C	**-1.466*** **	**-1.262*** **	**-1.299*** **	**-2.032*** **	**-1.606*** **	**-2.606*** **	**-1.341*** **	**-1.544*** **	**-1.485*** **	**-1.391*** **	**-1.292*** **	**-1.309*** **	**-1.161*** **
	(0.408)	(0.344)	(0.327)	(0.396)	(0.282)	(0.456)	(0.314)	(0.353)	(0.318)	(0.282)	(0.276)	(0.268)	(0.232)
VPMFGL(-1)	0.223	0.138	0.044	0.337	0.474	0.107	0.529	0.169	-0.329	-0.638	-0.396	-0.169	-0.257
	(0.371)	(0.332)	(0.359)	(0.382)	(0.320)	(0.348)	(0.377)	(0.394)	(0.385)	(0.390)	(0.444)	(0.355)	(0.359)
AR(-1)	**0.474*** **	**0.438*** **	**0.513*** **	**0.512*** **	**0.441*** **	**0.476*** **	**0.427*** **	**0.487*** **	**0.498*** **	**0.455*** **	**0.534*** **	**0.559*** **	**0.539*** **
	(0.081)	(0.076)	(0.104)	(0.078)	(0.086)	(0.075)	(0.064)	(0.066)	(0.075)	(0.085)	(0.085)	(0.063)	(0.074)
AR(-2)	**0.191*** **	**0.194*** **	**0.228** **	0.032	**0.185*** **	0.030	0.120	**0.182** **	**0.148***	**0.243*** **	**0.156***	**0.144** **	**0.212** **
	(0.072)	(0.063)	(0.092)	(0.063)	(0.071)	(0.082)	(0.073)	(0.084)	(0.087)	(0.081)	(0.093)	(0.067)	(0.084)
AR(-3)	0.100	**0.155** **	0.053	**0.099***	0.047	0.060	**0.179*** **	0.084	**0.119***	0.071	0.094	0.077	0.051
	(0.061)	(0.062)	(0.063)	(0.056)	(0.050)	(0.073)	(0.058)	(0.083)	(0.062)	(0.059)	(0.061)	(0.051)	(0.057)
R-squared	0.461	0.490	0.527	0.313	0.333	0.251	0.343	0.444	0.486	0.503	0.534	0.522	0.566

Variable	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
C	**-1.229*** **	**-1.143*** **	**-1.753*** **	**-2.067*** **	**-1.922*** **	**-1.295*** **	**-1.089*** **	**-1.293*** **	**-2.034*** **	**-2.119*** **	**-1.447*** **	**-1.536*** **	**-2.240*** **
	(0.274)	(0.297)	(0.322)	(0.343)	(0.381)	(0.327)	(0.247)	(0.286)	(0.383)	(0.441)	(0.285)	(0.331)	(0.539)
VPMFGL(-1)	-0.338	-0.444	0.041	-0.515	**-0.760***	-0.528	-0.528	-0.463	-0.099	-0.432	-0.290	-0.316	**-0.619** **
	(0.466)	(0.367)	(0.384)	(0.396)	(0.418)	(0.320)	(0.357)	(0.341)	(0.361)	(0.367)	(0.422)	(0.391)	(0.293)
AR(-1)	**0.491*** **	**0.484*** **	**0.486*** **	**0.417*** **	**0.348*** **	**0.441*** **	**0.460*** **	**0.356*** **	**0.463*** **	**0.353*** **	**0.475*** **	**0.506*** **	**0.333*** **
	(0.098)	(0.078)	(0.085)	(0.093)	(0.068)	(0.064)	(0.105)	(0.080)	(0.071)	(0.067)	(0.082)	(0.075)	(0.062)
AR(-2)	**0.238*** **	**0.208*** **	**0.137***	**0.229*** **	**0.268*** **	**0.236*** **	**0.202** **	**0.263*** **	0.098	**0.256*** **	**0.219** **	0.133	**0.225*** **
	(0.089)	(0.077)	(0.075)	(0.078)	(0.081)	(0.089)	(0.090)	(0.083)	(0.078)	(0.081)	(0.089)	(0.089)	(0.061)
AR(-3)	0.073	**0.111** **	0.045	0.039	0.052	0.102	**0.147** **	**0.164*** **	**0.108***	0.102	0.075	**0.120***	0.097
	(0.055)	(0.056)	(0.071)	(0.058)	(0.056)	(0.063)	(0.072)	(0.062)	(0.061)	(0.071)	(0.054)	(0.062)	(0.070)
R-squared	0.549	0.553	0.348	0.385	0.334	0.513	0.555	0.475	0.330	0.384	0.489	0.478	0.317

Window	Benchmark	SPX	CCMP	SPTSX	MEXBOL	IBOV	IPSA	MERVAL	IGBVL	FTSE	CAC	DAX	IBEX	FTSEMIB
ROL	AR(3)	-0.081	-0.057	-0.238	0.280	**0.966***	0.064	0.558	0.299	-0.530	-0.160	-0.793	-0.795	-0.208
	AR(3)VIX	-0.053	-0.010	0.104	-0.163	0.260	-0.204	-0.109	-0.194	-0.319	-0.393	-0.953	-0.014	0.025
	AR(6)	-0.126	-0.063	-0.241	0.376	**0.847***	-0.002	**0.804***	0.166	-0.459	0.138	-0.589	-1.012	-0.244
	AR(6)VIX	0.019	0.097	0.158	-0.142	0.224	-0.227	-0.092	-0.204	-0.336	-0.272	-0.810	-0.394	-0.307
REC	AR(3)	-0.019	-0.052	-0.349	0.021	0.628	-0.082	0.335	-0.016	-0.316	0.177	-0.089	-0.349	0.014
	AR(3)VIX	-0.051	-0.079	0.158	-0.335	0.084	-0.198	-0.159	-0.251	-0.296	-0.197	-0.313	0.161	0.113
	AR(6)	0.015	0.019	-0.332	0.290	0.629	-0.043	**0.856***	-0.055	-0.205	0.279	0.033	-0.481	-0.090
	AR(6)VIX	-0.040	-0.042	0.173	-0.292	-0.032	-0.194	0.109	-0.210	-0.257	-0.163	-0.245	-0.011	-0.013

Window	Benchmark	AEX	OMX	RTS	SMI	NKY	HSI	KOSPI	TWSE	JCI	FBMKLCI	STI	ASX	NZSE
ROL	AR(3)	-0.548	-0.385	-0.313	-0.655	0.459	0.014	0.494	-0.314	-1.053	-0.437	-0.449	-0.510	**0.799***
	AR(3)VIX	-0.351	-0.202	-0.256	-0.949	-0.625	-0.156	0.638	0.431	-0.663	-0.160	-0.187	-0.282	0.426
	AR(6)	-0.448	-0.215	-0.377	-0.660	0.588	0.083	0.390	-0.389	-1.243	-0.631	-0.731	-0.421	**0.887***
	AR(6)VIX	-0.268	-0.121	-0.239	-0.832	-0.368	-0.212	**0.771***	0.395	-0.601	-0.178	-0.199	-0.306	0.487
REC	AR(3)	-0.163	0.044	-0.475	-0.469	**0.874***	0.305	0.095	-0.469	-1.145	-0.524	-0.339	-0.303	**0.880***
	AR(3)VIX	-0.151	-0.188	-0.461	-0.845	-0.379	-0.054	-0.174	-0.180	-0.918	-0.307	-0.271	-0.260	0.450
	AR(6)	-0.102	0.211	-0.627	-0.488	**0.868***	0.471	-0.085	-0.527	-1.106	-0.580	-0.442	-0.197	**0.934***
	AR(6)VIX	-0.140	-0.134	-0.585	-0.870	-0.131	-0.062	-0.110	-0.107	-0.907	-0.271	-0.214	-0.217	0.508

TABLE VIII: In-sample (upper tables) and out-of-sample (lower tables) results for VPMFGL. The regression is performed on our own data for a period starting from 2001-07 until 2019-09.