



An exploration of the mathematical structure and behavioural biases of 21st century financial crises

Nick James^{a,*}, Max Menzies^{b,1}

^a School of Mathematics and Statistics, University of Melbourne, Victoria, Australia

^b Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing, China

ARTICLE INFO

Keywords:

Financial market crises
Nonlinear time series analysis
Portfolio optimisation

ABSTRACT

In this paper we contrast the dynamics of the 2022 Ukraine invasion financial crisis with notable financial crises of the 21st century — the dot-com bubble, global financial crisis and COVID-19. We study the similarity in market dynamics and associated implications for equity investors between various financial market crises and we introduce new mathematical techniques to do so. First, we study the strength of collective dynamics during different market crises, and compare suitable portfolio diversification strategies with respect to the unique number of sectors and stocks for optimal systematic risk reduction. Next, we introduce a new linear operator method to quantify distributional distance between equity returns during various crises. Our method allows us to fairly compare underlying stock and sector performance during different time periods, normalising for those collective dynamics driven by the overall market. Finally, we introduce a new combinatorial portfolio optimisation framework driven by random sampling to investigate whether particular equities and equity sectors are more effective in maximising investor risk-adjusted returns during market crises.

1. Introduction

Perhaps the most challenging part of investing is making portfolio decisions during economic and financial market crises [1]. Most investors and fund managers can make money during bull markets [2], it is bear markets where many investment managers are revealed, as the majority fail to beat index funds [3]. Even in academia, a disproportionate amount of research focuses on financial market characteristics that ultimately derive from periods of distress. This includes the identification and modelling of heavy-tailed distributions [4,5], the identification of regime switching behaviours [6,7] and structural breaks [8,9], optimal portfolio construction [10], heteroskedasticity [11], nonstationarity [12,13] and perhaps most of all, the nature of heightened volatility [14,15]. There is an old saying in financial markets that is often quoted during times of crisis, “*history may not repeat itself, but it does rhyme*” [16]. In this work, we address this question mathematically and in a data-driven study — where we introduce new mathematical methods and frameworks to explicitly test the persistence of financial market structure and investor behavioural biases during times of crisis. We are also motivated to compare and contrast the most recent financial crisis brought about by the Russian invasion of Ukraine [17] with prior crises such as the global financial crisis (GFC).

Even if the present day is not strictly speaking a financial crisis, this research is still highly relevant today, given the contemporary market period is quite unprecedented. Following a strong bull market (outside the COVID-19 drawdown and subsequent recovery), we have begun to observe patches of an economic slowdown [18]. Many market commentators, investment banks, hedge funds and

* Corresponding author.

E-mail addresses: nick.james@unimelb.edu.au (N. James), max.menzies@alumni.harvard.edu (M. Menzies).

¹ Equal first authors, playing an equal role in every aspect of the manuscript.

other high profile stakeholders in the investment business are predicting a sharp economic slowdown and recession, often termed a “hard landing” [19]. We are currently experiencing high inflation and money supply around the world [20], and this has triggered central banks to commence structured monetary tightening programmes, delivered through systematic increase in interest rates [21]. However, inflation is proving to be sticky and still remains too high according to numerous policymakers [22]. Furthermore, we have seen a clear disconnect between fixed income and equity markets, unlike prior financial crises and periods of financial distress [23]. To further confuse investors and other financial market participants, we have observed a variety of geopolitical and economic events that fail to align with a singular outlook on market conditions. For instance, the ongoing Ukraine war has had massive ramifications on energy prices [17], while the technology and financial services sectors have instituted multiple rounds of retrenchments [24]. By contrast, the continued development and hype surrounding generative artificial intelligence potentially brings great opportunity and disruption [25]. Such changes have caused great contradiction within equity markets. In this work, we explore the differences between the recent financial market period and prior crises. In doing so, we also wish to build on numerous academic work that has analysed financial performance and volatility during market crises, including the Dot-com bubble [26], GFC [27], COVID-19 [28,29], smaller regional crashes [30] and a limited amount of work regarding the Ukraine crisis [15,31]. In particular, we seek to investigate the consistency of investor behavioural biases during times of crises, that is, do we see repeat behaviours that can be exploited to generate investor returns.

Our work draws on a long history of applying statistical and physically-inspired models to capture the dynamics of real world phenomena. In financial markets, these methodologies have been applied to a wide range of asset classes such as equities [32–34], foreign exchange [35] and cryptocurrencies [36–42] and debt-related instruments [43]. Such methods from applied mathematics have been used in a variety of other disciplines including epidemiology [44–52], environmental sciences [53–61], crime [62–64], the arts [65,66], and other fields [67–69]. Readers interested in recent work related to temporal dynamics with various societal impacts on the economy should consult [70–72].

In particular, there is a significant body of work in financial markets that focuses on the cross-correlation matrix of asset price returns as the central object of study. This captures the interdependencies of assets in the market or smaller communities of equities such as GICS sectors. To study the evolution of such correlation structures, techniques such as principal components analysis has been employed, with a particular focus on studying the evolution of the leading eigenvectors and their associated eigenvalue. Various authors have shown that just a few components can describe most of the observed variability of the market [73–76]. Employing techniques such as random matrix theory, further nuances in the market can be uncovered [77–79]. A wealth of research has also covered network analysis, where the stock is viewed as a complex network and the cross-correlation matrix captures the underlying strength between individual assets, or groups of assets [80,81]. A wide variety of mathematical and statistical methods have been used to study correlation structures in global equity markets over time, and to identify temporal dependence exhibited by various communities of securities [82–85]. The econophysics community have applied such techniques to a wide variety of asset classes. Specific attributes of financial time series, such as their nonstationarity and volatility have led to numerous works exploring methods to capture nuanced effects such as their heavy-tailedness and volatility clustering [86,87]. These methods often try to overcome limiting assumptions of the correlation metric, such as its assumption of linearity.

Our final motivating topic of financial portfolio optimisation has been addressed from different perspectives across a variety of research communities. Researchers from fields such as operations research, applied mathematics, statistics, econometrics and more, have tackled the problem of forecasting and optimising financial risk and portfolio selection [88–93]. In the econophysics community, various approaches have been used to study methods for reducing market risk, “best-value” portfolio construction and evolutionary asset allocation. One benefit in tackling such a problem from an econophysics perspective, is the flexibility in candidate problem solving — where econophysicists can draw on techniques developed in a wide variety of mathematical disciplines.

This paper is structured as follows. In Section 2, we outline the specific dates of our financial crises, and comment on the change in the distribution of those equity correlations over time. In Section 3, we introduce our random sampling model for the study of evolutionary collective dynamics, and focus on contrasting the benefit in diversification across and within equity sectors in various financial crises. In Section 4, we investigate the similarity in the market performance of various sectors performance after normalising for different periods — this provides a new approach to directly comparing different economic crises appropriately. We conclude with Section 5, where we introduce a new combinatorial optimisation framework for identifying the most frequently occurring equities and equity sectors in top-performing portfolios during times of crisis.

2. Data and overview of financial crisis periods

Throughout this paper, we are interested in the study of collective dynamics over time and within distinct market periods. The core object of study is a collection of equity time series sampled from a variety of GICS sectors across two decades of daily price data.

We define dates of our economic crises as follows:

- **Dot-com bubble:** 01-03-2000 - 01-10-2002
- **Global financial crisis (GFC):** 03-01-2007 - 03-05-2010
- **COVID-19:** 11-03-2020 - 31-08-2020
- **2022 crash associated with Ukraine invasion:** 03-01-2022 - 13-05-2022

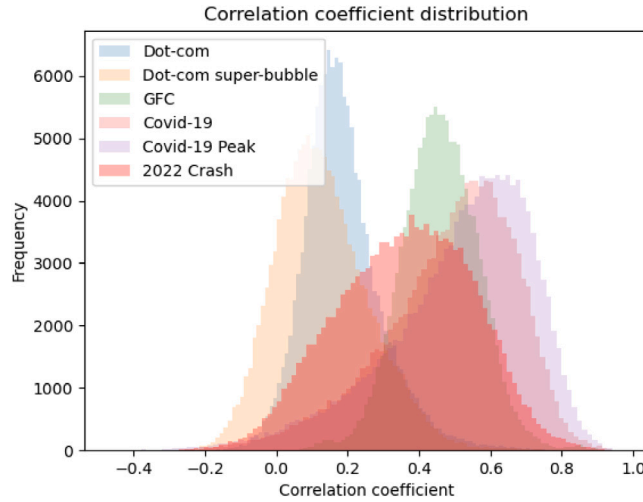


Fig. 1. The distribution of equity market correlation coefficients computed in different market crises. Until the Ukraine crisis, there was a clear tendency for equity correlations to become more strongly positive with significantly less variance. The collective severity of correlations decrease among crises in the order COVID-19, the GFC, the Ukraine crash, and the Dot-com.

Fig. 1 displays the non-trivial correlations between all equities’ log returns time series during our periods of investigation. We also include the “dot-com super-bubble” [94] (comprising the dot-com bubble up until end of March 2000) and the peak of the COVID-19 financial crisis (11-03-2020 to 29-05-2020). We remark with caution that there may be significant survivorship bias here, particularly with respect to the dot-com bubble (and associated super-bubble) as many stocks that were most significantly affected during this time went bankrupt and do not exist in our analysis. There is a persistent trend from the dot-com bubble until the peak period of the COVID-19 crash, that the average equity correlation during market crises increased quite substantially. This would imply that during crises there is potentially less opportunity for diversification benefit (with less low and negative correlation available), and that economic crises may be shorter and more severe. By contrast, the 2022 market crisis exhibits meaningfully lower average correlation than the GFC, COVID-19 and peak COVID-19 market periods. It is this precise point of difference we use as our springboard for this paper, and seek to investigate the points of difference, and how optimal investor behaviour would deviate during a period such as this, relative to other crises.

3. Crisis-dependent risk reduction within and across sectors

Let $p_i(t)$ be the multivariate time series of equities’ daily closing prices where $i = 1, \dots, N$ index the $N = 503$ equities under examination. We generate a multivariate time series of log returns as follows:

$$R_i(t) = \log \left(\frac{p_i(t)}{p_i(t-1)} \right), t = 1, \dots, T. \tag{1}$$

We standardise our equity returns via $\tilde{R}_i(t) = [R_i(t) - \langle R_i \rangle] / \sigma(R_i)$, where $\langle \cdot \rangle$ is an expectation and σ represents a standard deviation, both of which are computed over the same time interval. We then choose a smoothing window S and define a time-varying correlation matrix Ψ as follows:

$$\Psi(t) = \frac{1}{S} \tilde{R}(t) \tilde{R}^T(t), t = S, \dots, T. \tag{2}$$

There is a wide range of previous literature that focuses on the appropriate value of the parameter S [74,95]. When S is too small, changes in the correlation structure will be highly dynamic, however, the results may be excessively noisy. By contrast, when S is large, the evolutionary model may be insensitive to abrupt changes in correlation structure. In this paper, we set $S = 60$, shorter than typical values of 90 or 120, as we are investigating shorter and more intense periods of market crisis.

Subsequently, we apply time-varying principal components analysis (PCA) to these time-varying correlation matrices. The key object we examine is the explanatory variance conveyed by the correlation matrix’s first eigenvalue. At each instance in time, our evolutionary correlation matrix will generate a sequence of ordered eigenvalues $\lambda_1(t), \dots, \lambda_N(t)$. We normalise these eigenvalues by $\tilde{\lambda}_i(t) = \frac{\lambda_i(t)}{\sum_{k=1}^N \lambda_k(t)}$.

We now build upon the work first introduced in [95,96], and compare the results from an extensive sampling experiment to highlight differences in across/within equity sector diversification in different market crises. We draw inspiration from the evolutionary collective strength of the market’s correlation matrix, measured by $\tilde{\lambda}_1(t)$, to quantify the diversification benefit from various portfolio mixes. For varying values of (w, a) spanning $2 \leq w, a \leq 9$, we draw $D = 1000$ random portfolios of wa equities consisting of a sectors and w equities per sector. That is, within each sector (chosen randomly), we sample w equities per sector, and

Table 1

Average $\mu_{w,a}$ of the median normalised eigenvalue $\tilde{\lambda}_{w,a}^{0.50}(t)$ for different pairs (w,a) diversifying across a sectors and w equities per sector during the dot-com crisis. In red, we display a greedy path reducing the value of $\mu_{w,a}$ (implying an increase in the overall diversification benefit) by gradually increasing the portfolio size, starting from the smallest portfolio (2, 2).

Number of sectors	Number of equities per sector							
	2	3	4	5	6	7	8	9
2	0.435	0.375	0.354	0.344	0.332	0.331	0.308	0.305
3	0.341	0.320	0.294	0.290	0.284	0.280	0.268	0.274
4	0.311	0.289	0.271	0.265	0.259	0.252	0.252	0.251
5	0.290	0.268	0.255	0.253	0.243	0.237	0.240	0.239
6	0.271	0.256	0.248	0.240	0.236	0.231	0.228	0.228
7	0.262	0.247	0.237	0.234	0.231	0.230	0.224	0.226
8	0.248	0.240	0.232	0.228	0.223	0.225	0.222	0.223
9	0.242	0.232	0.228	0.221	0.221	0.220	0.218	0.216

we sample across a sectors. Both the individual securities and equity sectors are drawn randomly and independently with uniform probability.

To quantify diversification benefit for a portfolio that consists of wa equities, we compute the $wa \times wa$ correlation matrix Ψ for each draw and calculate the respective normalised first eigenvalue $\tilde{\lambda}_{w,a}(t)$. We register the 50th percentile (median) of the $D = 1000$ values, which we denote $\tilde{\lambda}_{w,a}^{0.50}(t)$. We analyse this quantity in numerous subsequent experiments. We commence by computing the temporal mean of the median of the normalised first eigenvalue,

$$\mu_{w,a} = \frac{1}{T - \tau + 1} \sum_{t=\tau}^T \tilde{\lambda}_{w,a}^{0.50}(t) \tag{3}$$

as a measure of the diversification benefit of a portfolio with w equities sampled uniformly (and independently) across a equity GICS sectors, averaged both over time and across our random sampling. We record the values of $\mu_{w,a}$ calculated for each of the four crises under examination in Tables 1, 2, 3 and 4. We term these tables diversification pathways as they offer insight as to how best to increase the collective diversification benefit of a portfolio, by increasing either the number of sectors a or stocks per sectors w . Within the tables, we mark in red *greedy paths* that sequentially increase either w or a to maximally decrease the value of $\mu_{w,a}$, that is increase the overall diversification benefit.

We start by examining diversification pathways during the dot-com bubble, shown in Table 1. In exploring the evolution of our greedy strategy, the $\mu_{w,a}$ quantity begins at 0.435 for the (2,2) portfolio and reaches a low-point of 0.221 for the (5,9) portfolio. The greedy path suggests that there is limited further improvement beyond a portfolio consisting of 3 or 4 equities sampled randomly from 5 different equity sectors. The significant reduction in market strength displayed by the average portfolio during the dot-com bubble is representative of decreased correlation among equities compared to either crises (Fig. 1), and larger potential to diversify.

We next turn to Table 2 where we study optimal portfolio construction during the GFC. By contrast, the (2,2) portfolio has a higher level of collective market strength with a $\mu_{w,a} = 0.581$ and only reducing to $\mu_{w,a} = 0.439$ by the time it reaches a (9,6) portfolio. This analysis suggests that during the GFC, there was less opportunity to successfully diversify a portfolio with fewer equities, demonstrated by a significantly larger optimal portfolio.

Optimal portfolio diversification during the COVID-19 market crisis, shown in Table 3, commences at the highest point of any crisis we examine. The (2,2) portfolio has a value of $\mu_{w,a} = 0.631$, and the optimal portfolio only decreases to a value of $\mu_{w,a} = 0.500$, which is achieved by the (9,5) portfolio. There is essentially no diversification benefit beyond the (6,3) portfolio, which yields a value of $\mu_{w,a} = 0.501$. The limited reduction in average market strength is consistent with the high degree of collective behaviour amongst equities during the COVID-19 market crash and reflects the near impossibility of obtaining a comfortably diversified portfolio during this crisis.

Finally, we turn to the Ukraine market crisis displayed in Table 4. The (2,2) portfolio yields a value of $\mu_{w,a} = 0.512$ and reaches a low-point of $\mu_{w,a} = 0.351$ with the (5,9) portfolio. This significant reduction is consistent with the clear spread in the distribution of correlation among equities (Fig. 1), allowing for substantial reduction in portfolio risk. Furthermore, there appears to be the existence of a reasonable “best-value” portfolio, as there is limited reduction in average portfolio risk beyond a portfolio with 3 equities samples from either 5 or 6 equity sectors.

To complement the above tables, we directly display the (red) greedy paths in Fig. 2, plotting the value of $\mu_{w,a}$ against the number of steps for each of the four crises. This figure reveals several key findings. First, we see a consistent pattern where there is limited diversification benefit beyond 5 or 6 diversification decisions. Typically this is a portfolio mix of (5,3) or (6,3), where the portfolio is of total size 15 or 18. Together with the tables above, this demonstrates that most diversification steps are made across (rather than within) equity sectors). It is interesting to look at the change in curvature of the this greedy path across various market crises, and how these change after more diversification decisions have been made. Most notably, we see the greatest sequential diversification benefit during the dot-com bubble, and the least during COVID-19 and the GFC. Of particular note is the greedy path that is observed during the COVID-19 crisis, where we observe essentially no increase in diversification potential, just noisy oscillation, beyond step 5. This is consistent with the sharp and abrupt decline in equity markets during this time period, and the indiscriminate selling by various investor types that was observed. By contrast, more protracted crises such as the global financial

Table 2

Average $\mu_{w,a}$ of the median normalised eigenvalue $\bar{\lambda}_{w,a}^{0.50}(t)$ for different pairs (w, a) diversifying across a sectors and w equities per sector during the GFC. In red, we display a greedy path reducing the value of $\mu_{w,a}$ (implying an increase in the overall diversification benefit) by gradually increasing the portfolio size, starting from the smallest portfolio (2, 2).

Number of sectors	Number of equities per sector							
	2	3	4	5	6	7	8	9
2	0.581	0.550	0.530	0.526	0.510	0.499	0.505	0.497
3	0.527	0.509	0.480	0.481	0.483	0.479	0.472	0.472
4	0.501	0.479	0.475	0.470	0.464	0.460	0.456	0.456
5	0.482	0.470	0.467	0.460	0.453	0.450	0.452	0.455
6	0.476	0.465	0.451	0.453	0.448	0.445	0.448	0.444
7	0.461	0.454	0.452	0.444	0.443	0.444	0.442	0.442
8	0.461	0.454	0.446	0.442	0.439	0.439	0.438	0.438
9	0.457	0.444	0.445	0.440	0.439	0.439	0.439	0.436

Table 3

Average $\mu_{w,a}$ of the median normalised eigenvalue $\bar{\lambda}_{w,a}^{0.50}(t)$ for different pairs (w, a) diversifying across a sectors and w equities per sector during the COVID-19 crisis. In red, we display a greedy path reducing the value of $\mu_{w,a}$ (implying an increase in the overall diversification benefit) by gradually increasing the portfolio size, starting from the smallest portfolio (2, 2).

Number of sectors	Number of equities per sector							
	2	3	4	5	6	7	8	9
2	0.631	0.591	0.588	0.577	0.56	0.556	0.553	0.562
3	0.578	0.566	0.543	0.543	0.543	0.553	0.538	0.531
4	0.555	0.541	0.535	0.520	0.522	0.520	0.517	0.524
5	0.533	0.516	0.520	0.501	0.517	0.515	0.504	0.504
6	0.522	0.501	0.518	0.512	0.510	0.503	0.506	0.507
7	0.526	0.514	0.507	0.497	0.508	0.506	0.502	0.495
8	0.518	0.508	0.507	0.502	0.501	0.503	0.497	0.493
9	0.511	0.507	0.507	0.500	0.502	0.502	0.501	0.495

Table 4

Average $\mu_{w,a}$ of the median normalised eigenvalue $\bar{\lambda}_{w,a}^{0.50}(t)$ for different pairs (w, a) diversifying across a sectors and w equities per sector during the 2022 market crisis. In red, we display a greedy path reducing the value of $\mu_{w,a}$ (implying an increase in the overall diversification benefit) by gradually increasing the portfolio size, starting from the smallest portfolio (2, 2).

Number of sectors	Number of equities per sector							
	2	3	4	5	6	7	8	9
2	0.512	0.489	0.467	0.473	0.469	0.466	0.447	0.444
3	0.457	0.414	0.425	0.420	0.411	0.408	0.405	0.399
4	0.433	0.402	0.393	0.395	0.393	0.381	0.377	0.379
5	0.408	0.370	0.376	0.373	0.363	0.363	0.372	0.364
6	0.392	0.370	0.365	0.368	0.368	0.364	0.353	0.357
7	0.391	0.369	0.364	0.359	0.365	0.360	0.355	0.359
8	0.382	0.361	0.351	0.353	0.353	0.350	0.353	0.348
9	0.378	0.361	0.358	0.351	0.351	0.346	0.352	0.346

crisis and the dot-com bubble see continued diversification benefit throughout the course of the respective greedy path. These findings mirror the distributions that are shown in Fig. 1, where tighter distributions more strongly translated to the right are consistent with market crises where it is harder to generate diversification benefits.

We conduct one further supporting experiment, where we take an alternative approach to examining the diversification benefit in sequential steps taken across and within equity sectors. In Fig. 3, we compute and display averaged diversification benefits $\mu_{w,\cdot} = \mathbb{E}_a \mu_{w,a}$ averaged across values of a and $\mu_{\cdot,a} = \mathbb{E}_w \mu_{w,a}$ averaged across values of w . Concretely, these are computed as the column and row averages, respectively, of the above tables of $\mu_{w,a}$, and we display them for each crisis under consideration. In all four market crises, this figure again confirms that beyond 4 or 5 steps, there is greater further reduction in diversifying across sectors, rather than diversifying with different equities sampled from within the same sector. In addition, it again confirms the specific market crises where diversification benefits were most significant. Diversification benefits were most substantial in the dot-com crisis, then Ukraine, the GFC and finally the COVID-19 market crash.

4. Crisis-to-crisis linear projection modelling

In this section, we start with the following motivating question, *how would equity X have performed in financial period Y?* Here, we introduce a mathematical framework to address this question, and seek to compare the distribution of equity sector returns

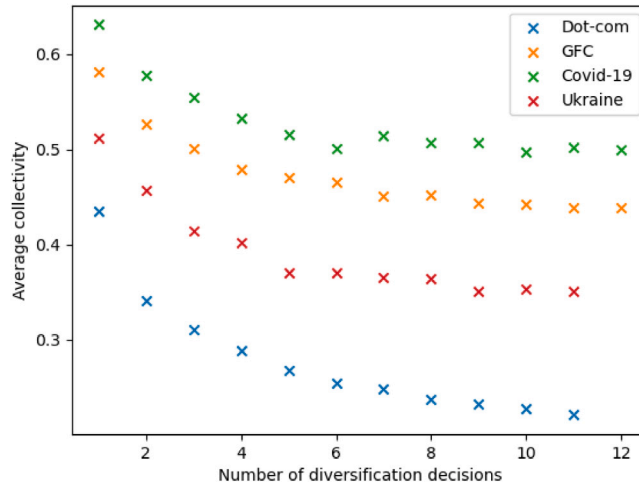


Fig. 2. The greedy paths demonstrate the sequential reduction in average collectivity, or increase in collective diversification benefit, as the portfolio increases in size via increases of w or a . It is interesting to note the deviation in total risk reduction that is achieved during the dot-com crisis (maximal reduction) when compared to the COVID-19 crisis (minimal reduction). The same order in decreasing collective strength of correlations is observed as in Fig. 1: COVID-19, the GFC, the Ukraine crash, and the Dot-com.

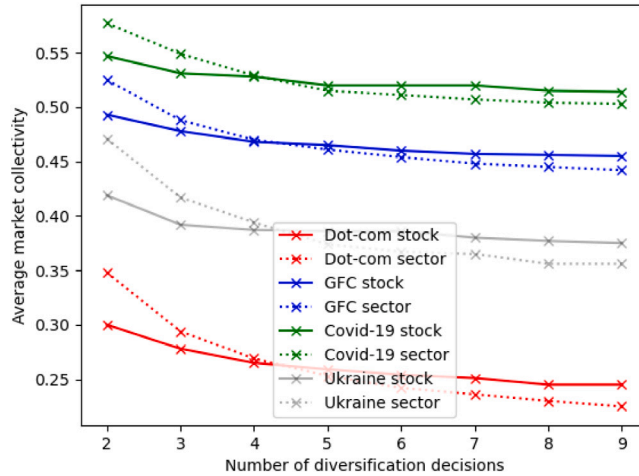


Fig. 3. In this figure we contrast the sequential benefit in diversification steps that are taken either across or within equity sectors, having averaged over all possibilities. The paths in the figure depict the change in values of both $\mu_{w,a} = \mathbb{E}_a \mu_{w,a}$ averaged across values of a (a column average of Tables 1–4) and $\mu_{,a} = \mathbb{E}_w \mu_{w,a}$ averaged across values of w (a row average of Tables 1–4). The figure shows that initially, diversification by stock may provide greater reduction in portfolio risk, however by the time 4 or 5 steps have been taken, there is a clear benefit in subsequent diversification steps taken across sectors, rather than within them. This finding is consistent across all four financial market crises studied. Again, the same order in decreasing collective strength of correlations is observed as in Fig. 1: COVID-19, the GFC, the Ukraine crash, and the Dot-com.

between financial crises and make inference as to whether sector dynamics or financial crises have greater discriminatory power. For example, do equities from the technology sector behave more similarly in the GFC and dot-com bubble than two different equity sectors (such as consumer discretionary and healthcare) during the same financial crisis. For this purpose, we wish to determine and use an appropriate operator to fairly transform distributions from one period to another. We construct an optimal linear operator as follows.

Given a distribution f , there are two natural affine-linear transformations to change it. More simply, there is translation, $f(x) \mapsto h(x) = f(x - b)$ for a constant b . Given a distribution, this moves the distribution f b units to the right to yield h . In addition, there is scaling, $f(x) \mapsto h(x) = \frac{1}{a} f(\frac{x}{a})$ for $a > 0$. The factor of $\frac{1}{a}$ ensures that the transformed h is still a distribution integrating (or summing) to 1. We can combine these two operations to define the following family of linear operators on the space of distributions:

$$T_{a,b}f(x) = \frac{1}{a} f\left(\frac{x-b}{a}\right). \tag{4}$$

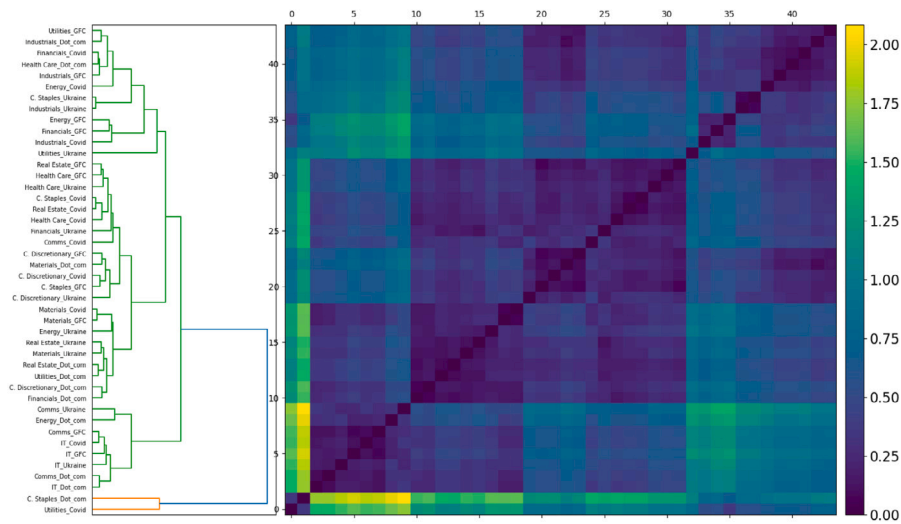


Fig. 4. Hierarchical clustering on transformed sector return distributions assesses whether the discriminatory power of individual equity sectors or distinct market crises is stronger. The left y-axis orders and groups different sectors across different time periods according to the similarity of return profiles after normalising for different underlying crisis conditions. The right y-axis gives a colour scale that summarises the size of entries in the distance matrix that forms the bulk of the figure. The cluster structure in the splitting tree of the left y-axis reveals that both individual sectors and market periods exhibit a similar influence in equity market returns. For instance, the information technology sector behaves most similarly with itself across various economic crises while a strong association between the financial, real estate and energy sectors is observed within distinct economic crises.

Next, we discuss how to choose the most suitable linear operator to compare two periods. To do this, let f be the distribution of market returns for the GFC, for example, and g be the distribution of market returns for the dot-com. There is unlikely to be a pair a, b such that the transformed distribution $T_{a,b}f$ coincides with g exactly. Instead, we select a, b to minimise the Wasserstein distance $d_W(g, T_{a,b}f)$. This will yield a choice T such that T approximately maps returns during the GFC period to returns of the dot-com period.

In Fig. 4, we take every sector S and every crisis C and compare the distributions under the Wasserstein distance. To fairly compare different crises, we use our linear operators determined through our optimisation scheme to map every sector from each crisis to the GFC as a sort of reference period. We explain this in several steps.

First, for notational clarity, we consider each pair of coefficients a, b as already determined and drop them from our notation. Thus, between any two crisis periods C_1, C_2 there is an operator $T(C_1, C_2)$. When $C_1 = C_2$, this is simply the identity operator. Now, let C_0 refer to the GFC as our reference crisis period. For each other crisis C , we may use our determined linear operator $T(C, C_0)$ to map the returns of C to the GFC C_0 . Again, for the GFC itself, $T(C, C_0)$ is simply the identity operator, doing nothing. For each sector S of every crisis C , we use the comparison operator $T(C, C_0)$ to map S and thus obtain an adjusted distribution of returns $T(C, C_0)S$. Finally, we use the Wasserstein metric to compare all adjusted distributions $T(C, C_0)S$ for each crisis C and sector S . Effectively, we have transformed all sectors' return distributions from all different crises into a common metric space where they may be fairly compared. Hierarchical clustering is shown in Fig. 4.

The dendrogram structure in this figure indicates that both individual equity sector idiosyncratic returns and crisis-related impacts exhibit similar discriminatory power. Across the dendrogram one can see multiple instances of association among homogeneous economic crises and equity sectors. Notable self-association can be observed where the same equity sector clusters together when sampled from different economic crises. For example, in a cluster of very high affinity we see strong association between the information technology (IT) and communications sector. This may be due to many equities with a latent IT classification being classified as communications companies under GICS definitions. We see a strong association for energy, real estate and financial sectors to cluster together within any candidate economic crisis. This may be indicative of these sectors' strong relationship with equity beta, and their uniformly strong and weak performance during bull and bear markets, respectively. By contrast, we also see examples of the same equity sector clustering together when sampled from distinct economic crises. For instance, the IT sector returns distribution during the COVID-19 crisis, GFC, Ukraine crisis and dot-com bubble associate together in a small cluster of very high affinity. This indicates that the strength of the IT sector's homogeneity outweighs the collective affinity of distinct economic crises.

5. Combinatorial simulation for optimal portfolio structure

The problem of optimal portfolio weights has been the subject of great debate among financial academics and practitioners for many years. There is a huge range in what various portfolio managers believe is an optimal number of stocks to hold, and the respective weight each of these stocks should comprise within a portfolio. Many quantitative portfolio managers hold hundreds or

Table 5

Sector allocation of top-performing portfolios during crises. For example, among our top 1% performing portfolios during the COVID-19 market crisis, energy stocks made up 3.8% across these portfolios, despite comprising 15.1% of the whole market. We see striking similarity during crises, and significant deviation from the index.

Sector	COVID-19	Dot-com	GFC	Ukraine	Index
Energy	3.8%	2.7%	2.4%	2.6%	15.1%
Communications	3.9%	5.1%	5.1%	6.0%	14.1%
Real Estate	4.3%	6.0%	5.4%	6.4%	13.1%
Utilities	5.5%	6.9%	6.4%	8.3%	12.7%
Materials	6.1%	7.5%	6.5%	8.6%	11.5%
Financials	9.7%	8.2%	8.4%	9.1%	6.6%
Consumer staples	10.3%	9.1%	8.9%	10.2%	6.2%
Consumer discretionary	11.2%	12.7%	12.7%	11.5%	5.8%
Industrials	14.8%	13.5%	13.9%	11.9%	5.6%
Information technology	15%	13.6%	14.5%	12.3%	5.2%
Healthcare	15.4%	14.5%	16.0%	13.1%	4.2%

thousands of equities in a market neutral capacity, seeking to minimise their exposure to the overall market. By contrast, many retail investors or high-conviction fund managers hold as few as 10–15 securities in a long-only capacity — justifying this behaviour as only putting capital behind their very best ideas.

In this section, we begin from the oft-observed assumption that it is difficult to beat equal weightings between assets in portfolio optimisation [97,98]. Our philosophy for this section is — when it is so difficult to determine optimal portfolio weights, why even try? Instead, we only vary the selection of assets, using just equal weightings to form our portfolios. We investigate the performance of different stocks (and their composite sectors) in how often they appear in optimised portfolios chosen from random sampling of different assets with equal weightings.

We proceed by performing a portfolio optimisation based on random sampling. We form 100 000 random portfolios, each determined by randomly sampling 40 distinct equities and forming an equally weighted portfolio, with each equity contributing 2.5% to the overall portfolio. We draw 100 000 random portfolios and extract the top 1% (or 1000) of these based on ranking their *Sharpe ratios*. Usually the Sharpe ratio optimisation problem takes the following form: one selects weights w_i to optimise the following quantity:

$$\frac{\sum_{i=1}^n w_i R_i - R_f}{\sqrt{\mathbf{w}^T \Sigma \mathbf{w}}}, \tag{5}$$

subject to: $0 \leq w_i \leq 1, i = 1, \dots, n,$ (6)

$$\sum_{i=1}^n w_i = 1, \tag{7}$$

where R_i are historical returns, Σ the historical covariance matrix, and R_f is the risk-free rate of investment (which we set to zero). This serves as a trade-off between expected portfolio returns and variance.

We make two changes in our setting. First, we reverse the domain of optimisation - w_i are always fixed at $w_i = \frac{1}{40}$ for an equally weighted portfolio, and we optimise only over the choice of assets in the portfolio; second, we sample randomly and deliberately record the best 1% of portfolios with respect to their Sharpe ratios rather than only the single best. We then study the composition of these top-performing portfolios, and investigate whether this optimal construction varies during different economic crises.

For each crisis C and sector S , let $p_S^{(C)}$ be the proportion of assets in sector S selected in the optimal random portfolios. These form probability vectors $\mathbf{p}^{(C)} \in \mathbb{R}^{11}$ of length 11, and constitute the columns of Table 5. Given two crises C_1, C_2 , we may define a distance between two distributions of chosen sectors as

$$d(C_1, C_2) = \|\mathbf{p}^{(C_1)} - \mathbf{p}^{(C_2)}\|_1 = \frac{1}{2} \sum_S |p_S^{(C_1)} - p_S^{(C_2)}|. \tag{8}$$

where the sum is taken over all 11 sectors S . This has the property that $d(C_1, C_2) = 0$ if and only if crises C_1 and C_2 have an identical distribution of stocks chosen across sectors. Furthermore, $d(C_1, C_2) \leq 1$ is the maximal possible distance, with equality if and only if the distributions are disjoint, with no sectors in common at all between the two crises. We also compute the distance between the four crises and the size distribution of sectors according to the count of stocks in each sector (the index column of Table 5.)

While quite simple, this distance may be interpreted as possibly the most suitable distance between two distributions of discrete sets, for it can be shown to be equivalent to the *discrete Wasserstein metric* between distributions on a discrete metric space. Specifically, (8) can be recast in a more general form

$$W^1(\mu, \nu) = \sup_F \left| \int_X F d\mu - \int_X F d\nu \right|. \tag{9}$$

In (9), X is a discrete set of size 11 equipped with the 0–1 distance, μ and ν are the discrete probability distributions associated with $\mathbf{p}^{(C_1)}$ and $\mathbf{p}^{(C_2)}$, respectively, and the supremum is taken over all 1-Lipschitz functions $F : X \rightarrow \mathbb{R}$, meaning $|F(x, y)| \leq d(x, y) = 1$

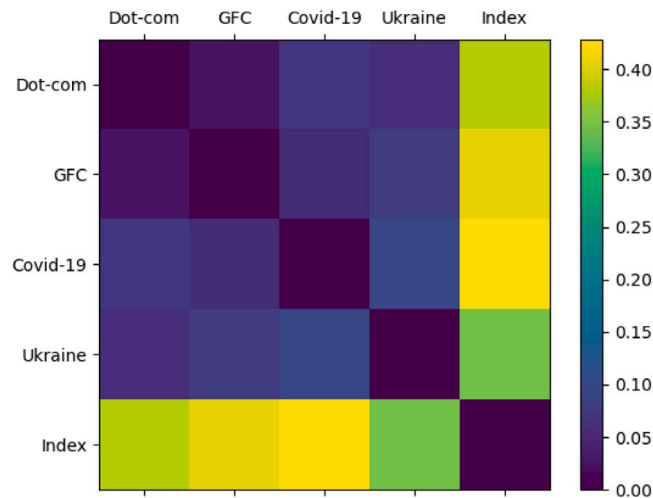


Fig. 5. In this figure, we measure the discrepancy between five distributions over equity sectors using the discrete Wasserstein distance (8). The five distributions are the determined distribution of portfolio weights during the four crises, Dot-com, GFC, COVID-19 and Ukraine war and finally the distribution of equity sectors by count. These form the rows and columns of the distance matrix, while the colour scale gives the size of the distance between distributions. The dark square indicates the four distributions of portfolio weights are highly similar (low mutual distance) among all four crises and differ substantially from the distribution by number of sectors. This is seen in Table 5 as well, where certain sectors are dramatically over- or under-represented in top-performing portfolios.

for all distinct x, y . More details are provided in [99]. We display the matrix of distances between crises C (as well as the index) in Fig. 5.

Table 5 shows the frequency of equities sampled from candidate sectors occurring in top-performing portfolios during times of crisis, and reveals numerous surprising findings. First, there is a remarkable consistency of the optimal portfolio membership distribution between market crises, as visible in the four highly similar crisis columns. That is, during the dot-com bubble, GFC, COVID-19 crash and Ukraine war, all periods which exhibited high equity correlations (to varying degrees) and persistent negative returns, the best-performing portfolios displayed incredible similarity in the distribution of sector weightings. Second, and by contrast, the index distribution is meaningfully different to expected optimal portfolio construction during times of crisis. This suggests that during economic crises investors seeking to construct portfolios optimising for risk-adjusted returns will have to deviate materially from index weightings and overweight select equity sectors. The aforementioned two findings are also clearly confirmed by Fig. 5, where under our discrete Wasserstein distance, the index distribution is clearly differentiated from the four crises and their close similarity.

The third finding is the emergence of a clear split in equity sectors that have a tendency to be overweight and underweight. The most predominant top-performing sectors include: financials, consumer staples, consumer discretionary, industrials, information technology and healthcare. Equities from these sectors tend to comprise $\sim 70\%$ – 75% of total portfolio weight. Equity sectors that occur less frequently include energy, communications, real estate, utilities and materials. These sectors tend to make up $\sim 20\%$ – 25% of total portfolio weight. Again, one can see material deviation from optimal weighting during all four crises to the weights one would expect after randomly sampling from the index. It is important to note that the index percentages are based on the number of equity securities existing within each sector and do not reflect each equity's market capitalisation. We are assuming that each security has an equal chance of being held in the portfolio, regardless of other factors such as market capitalisation (which can be highly dependent on recent share price performance).

6. Conclusion

It is worth framing our findings within a broader context of previous econophysics literature, and practical implications for equity investment management. Throughout this paper, we introduce new methodologies and apply them to financial market crises. Although the focal point of this paper is 21st century financial market crises, the methodologies we introduce are flexible and can be generalised well beyond the crises we examine. Concretely, this paper serves two purposes. First, we introduce techniques that could be applied to any time-varying financial system where collective dynamics are of interest (equities, fixed income, currencies, and so on). Second, we apply these techniques on US equity data to learn how equity investors should be managing stock portfolios during times of crisis. We specifically focus on financial market crises as there is greater homogeneity in market dynamics when comparing equity bear markets, than equity bull markets. We have undertaken a detailed and principled study of equity market performance during previous market crises. Our analysis focuses on optimal portfolio construction by way of marginal diversification benefit across and within equity sectors, the discriminatory power and association power of equity sectors and individual market crises, and the most probable portfolio structure to maximise risk-adjusted returns during crises. In each section of our paper, we

either extend or develop new mathematical techniques to answer these questions. The methodologies introduced in this paper may be extended quite naturally to any comparative modelling of two candidate financial periods.

The analysis in this paper has direct implications to investors, and their investment decision-making during times of crisis. First, in Section 3, where we examine optimal portfolio construction with respect to equities held within and across equity sectors, we demonstrate several interesting findings. Most notably, our analysis identifies a relatively consistent best-value portfolio ranging from 15–18 equities, beyond which the marginal risk reduction from holding an additional equity security is diminished. In the case of the COVID-19 crisis, there is no additional diversification benefit observed at all beyond this point. Our analysis across crises reveals that market crises with higher levels of correlation, unsurprisingly, show materially less diversification benefit as they traverse down the greedy path from (2,2) down to (5,9) or (6,9). In essence, we see surprising consistency in the emergence of a best-value portfolio for risk reduction across all examined market crises. The second key implication to investors comes in Section 5, where we introduce a new simulation framework to test for optimal portfolio construction during different market crises. Our framework demonstrates remarkable consistency in the expected optimal portfolio's equity sector weights. Fig. 5 crystallises this idea, where the similarity in these weighting schemes is revealed, and perhaps just as importantly, the material deviation from expected index weightings is highlighted. Further, we demonstrate that in the event of a market crisis, implementing successful weighting schemes of prior crises is perhaps the most appropriate strategy. This suggests that equity investors who wish to outperform the index during times of crisis should be guided by successful asset allocation mixes from prior crises. In particular, investors should take note of the fact that select equity sectors such as Healthcare, Information technology, Industrials, Consumer discretionary and Consumer staples are materially overrepresented in optimal portfolios, while sectors such as Energy, Communications, Real estate, Utilities and Materials should be underweighted.

Finally, all the analysis in the paper could be extended beyond US equities. There are many smaller regional financial crises that were not investigated in this paper. Our linear projection analysis (Section 4) demonstrates that both crisis and equity sector homogeneity apply to individual sector performance, to varying degrees. This framework could be extended to the individual security level, and applied more broadly than US equities. This technique is a general approach to answer the question, “*Find a security today that is behaving similarly to how security X behaved in the past*”. This approach could apply to equity and fixed income security selection, and could even be extended to fund manager selection — where we study distributions of hedge fund returns. Future research could investigate equities and other assets more globally to see if there are consistent patterns of sectors or asset classes that are overrepresented in strong-performing portfolios during financial crises or periods of distress more broadly.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] Himanshu, Ritika, N. Mushir, R. Suryavanshi, Impact of COVID-19 on portfolio allocation decisions of individual investors, *J. Public Aff.* 21 (4) (2021) <http://dx.doi.org/10.1002/pa.2649>.
- [2] T.S. Bernard, Amateur investors rode the bull up. Now the bear looms, 2022, <https://www.nytimes.com/2022/05/18/your-money/stock-market-crash-trading-retail.html>. The New York Times, May 18, 2022.
- [3] A. Cain, Index funds outperform active managers in a bear market, 2020, <https://www.afr.com/companies/financial-services/index-funds-outperform-active-managers-in-a-bear-market-20200421-p541wl>. The Australian Financial Review, April 30, 2020.
- [4] J. Danielsson, B.N. Jorgensen, M. Sarma, C.G. de Vries, Comparing downside risk measures for heavy tailed distributions, *Econom. Lett.* 92 (2) (2006) 202–208, <http://dx.doi.org/10.1016/j.econlet.2006.02.004>.
- [5] H.V. Long, H.B. Jebreen, I. Dassios, D. Baleanu, On the statistical GARCH model for managing the risk by employing a fat-tailed distribution in finance, *Symmetry* (ISSN: 2073-8994) 12 (10) (2020) 1698, <http://dx.doi.org/10.3390/sym12101698>.
- [6] S. Gray, Modelling the conditional distribution of interest rates as a regime-switching process, *J. Financ. Econom.* 2 (1996) 211–250.
- [7] J. Cai, A Markov model of switching-regime ARCH, *J. Bus. Econ. Stat.* 12 (1994) 309–316.
- [8] C.G. Lamoureux, W.D. Lastrapes, Persistence in variance, structural change, and the GARCH model, *J. Bus. Econ. Stat.* 8 (2) (1990) 225–234, <http://dx.doi.org/10.2307/1391985>.
- [9] R.T. Baillie, C. Morana, Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach, *J. Econom. Dynam. Control* 33 (8) (2009) 1577–1592, <http://dx.doi.org/10.1016/j.jedc.2009.02.009>.
- [10] E. Vercher, J.D. Bermúdez, J.V. Segura, Fuzzy portfolio optimization under downside risk measures, *Fuzzy Sets and Systems* 158 (7) (2007) 769–782, <http://dx.doi.org/10.1016/j.fss.2006.10.026>.
- [11] D. Nelson, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59 (1991) 347–370.
- [12] J.D. Hamilton, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57 (2) (1989) 357–384, <http://dx.doi.org/10.2307/1912559>.
- [13] G. Livan, J. ichi Inoue, E. Scalas, On the non-stationarity of financial time series: impact on optimal portfolio selection, *J. Stat. Mech. Theory Exp.* 2012 (07) (2012) P07025, <http://dx.doi.org/10.1088/1742-5468/2012/07/p07025>.
- [14] R. Cerqueti, M. Giacalone, R. Mattera, Skewed non-Gaussian GARCH models for cryptocurrencies volatility modelling, *Inform. Sci.* 527 (2020) 1–26, <http://dx.doi.org/10.1016/j.ins.2020.03.075>.
- [15] Y. Fang, Z. Shao, The Russia-Ukraine conflict and volatility risk of commodity markets, *Finance Res. Lett.* 50 (2022) 103264, <http://dx.doi.org/10.1016/j.frl.2022.103264>.

- [16] T. Salamone, For stocks, history may not repeat itself, but it does rhyme, 2022, <https://www.nasdaq.com/articles/for-stocks-history-may-not-repeat-itself-but-it-does-rhyme>. Schaeffer's Investment Research, April 11, 2022.
- [17] B.M. Jenkins, Consequences of the war in Ukraine: The economic fallout, 2023, <https://www.rand.org/blog/2023/03/consequences-of-the-war-in-ukraine-the-economic-fallout.html>. Rand Corporation, March 7, 2023.
- [18] D. Bachman, United States economic forecast, 2023, <https://www2.deloitte.com/us/en/insights/economy/us-economic-forecast/2023-q2.html>. Deloitte, June 15, 2023.
- [19] K. Gilchrist, IMF warns hard landing 'within the realm of possibilities' for U.S. economy, 2023, <https://www.cnbc.com/2023/04/12/imf-warns-of-hard-landing-risk-for-us-economy.html>. CNBC, April 12, 2023.
- [20] D. DeSilver, In the U.S. and around the world, inflation is high and getting higher, 2022, <https://www.pewresearch.org/short-reads/2022/06/15/in-the-u-s-and-around-the-world-inflation-is-high-and-getting-higher/>. Pew Research Center, June 15, 2022.
- [21] D. McHugh, Europe's central bank hikes key interest rate to record high even as recession threat grows, 2023, <https://apnews.com/article/european-central-bank-interest-rates-inflation-recession-1b7ad9f3f6e9522de05d7b3c1f93e437>. The Associated Press, September 14, 2023.
- [22] C. Smith, H. Clarfelt, Jay Powell warns inflation 'too high' in Jackson Hole speech, 2023, <https://www.ft.com/content/a1506faf-ced9-4be7-b2f7-2f4c12375786>. The Financial Times, August 26, 2023.
- [23] R. Mathieson, Navigating a trilemma, 2023, <https://www.blackrock.com/us/financial-professionals/insights/systematic-equity-market-outlook>. Blackrock, April 21, 2023.
- [24] A. Stringer, A comprehensive list of 2023 tech layoffs, 2023, <https://techcrunch.com/2023/09/07/tech-industry-layoffs-2023/>. TechCrunch, September 7, 2023.
- [25] The economic potential of generative AI: The next productivity frontier, 2023, <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>. McKinsey & Company, June 14, 2023.
- [26] F. Lillo, R.N. Mantegna, Power-law relaxation in a complex system: Omori law after a financial market crash, *Phys. Rev. E* 68 (1) (2003) <http://dx.doi.org/10.1103/physreve.68.016119>.
- [27] A.M. Petersen, F. Wang, S. Havlin, H.E. Stanley, Market dynamics immediately before and after financial shocks: Quantifying the Omori, productivity, and Bath laws, *Phys. Rev. E* 82 (3) (2010) <http://dx.doi.org/10.1103/physreve.82.036114>.
- [28] S. Priscilla, S.E. Hatane, J. Tarigan, COVID-19 catastrophes and stock market liquidity: evidence from technology industry of four biggest ASEAN capital market, *Asia-Pac. J. Bus. Adm.* (2022) <http://dx.doi.org/10.1108/apjba-10-2021-0504>.
- [29] M. Lashkaripour, ESG tail risk: The Covid-19 market crash analysis, *Finance Res. Lett.* 53 (2023) 103598, <http://dx.doi.org/10.1016/j.frl.2022.103598>.
- [30] R. Fauzi, I. Wahyudi, The effect of firm and stock characteristics on stock returns: Stock market crash analysis, *J. Financ. Data Sci.* 2 (2) (2016) 112–124, <http://dx.doi.org/10.1016/j.jfds.2016.07.001>.
- [31] E. Keleş, Stock market response to the Russia-Ukraine war: Evidence from an emerging market, *J. East-West Bus.* 29 (3) (2023) 307–322, <http://dx.doi.org/10.1080/10669868.2023.2210121>.
- [32] D. Wilcox, T. Gebbie, An analysis of cross-correlations in an emerging market, *Physica A* 375 (2) (2007) 584–598, <http://dx.doi.org/10.1016/j.physa.2006.10.030>.
- [33] N. James, M. Menzies, K. Chin, Economic state classification and portfolio optimisation with application to stagflationary environments, *Chaos Solitons Fractals* 164 (2022) 112664, <http://dx.doi.org/10.1016/j.chaos.2022.112664>.
- [34] N. James, M. Menzies, A new measure between sets of probability distributions with applications to erratic financial behavior, *J. Stat. Mech. Theory Exp.* 2021 (12) (2021) 123404, <http://dx.doi.org/10.1088/1742-5468/ac3d91>.
- [35] M. Ausloos, Statistical physics in foreign exchange currency and stock markets, *Physica A* 285 (1–2) (2000) 48–65, [http://dx.doi.org/10.1016/s0378-4371\(00\)00271-5](http://dx.doi.org/10.1016/s0378-4371(00)00271-5).
- [36] R. Gebarowski, P. Oświęcimka, M. Wątopek, S. Drożdż, Detecting correlations and triangular arbitrage opportunities in the forex by means of multifractal detrended cross-correlations analysis, *Nonlinear Dynam.* 98 (3) (2019) 2349–2364, <http://dx.doi.org/10.1007/s11071-019-05335-5>.
- [37] M. Wątopek, S. Drożdż, J. Kwapien, L. Minati, P. Oświęcimka, M. Stanuszek, Multiscale characteristics of the emerging global cryptocurrency market, *Phys. Rep.* 901 (2021) 1–82, <http://dx.doi.org/10.1016/j.physrep.2020.10.005>.
- [38] N. James, M. Menzies, Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time, *Nonlinear Dynam.* 107 (4) (2022) 4001–4017, <http://dx.doi.org/10.1007/s11071-021-07166-9>.
- [39] J. Kwapien, M. Wątopek, M. Bezbradica, M. Crane, T.T. Mai, S. Drożdż, Analysis of inter-transaction time fluctuations in the cryptocurrency market, *Chaos* 32 (8) (2022) 083142, <http://dx.doi.org/10.1063/5.0104707>.
- [40] M. Wątopek, J. Kwapien, S. Drożdż, Multifractal cross-correlations of bitcoin and ether trading characteristics in the post-COVID-19 time, *Future Internet* 14 (7) (2022) 215, <http://dx.doi.org/10.3390/fi14070215>.
- [41] M. Wątopek, J. Kwapien, S. Drożdż, Cryptocurrencies are becoming part of the world global financial market, *Entropy* 25 (2) (2023) 377, <http://dx.doi.org/10.3390/e25020377>.
- [42] S. Drożdż, J. Kwapien, M. Wątopek, What is mature and what is still emerging in the cryptocurrency market? *Entropy* 25 (5) (2023) 772, <http://dx.doi.org/10.3390/e25050772>.
- [43] J. Driessen, B. Melenberg, T. Nijman, Common factors in international bond returns, *J. Int. Money Finance* 22 (5) (2003) 629–656, [http://dx.doi.org/10.1016/s0261-5606\(03\)00046-9](http://dx.doi.org/10.1016/s0261-5606(03)00046-9).
- [44] N. James, M. Menzies, P. Radchenko, COVID-19 second wave mortality in Europe and the United States, *Chaos* 31 (2021) 031105, <http://dx.doi.org/10.1063/5.0041569>.
- [45] C. Manchein, E.L. Brugnago, R.M. da Silva, C.F.O. Mendes, M.W. Beims, Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies, *Chaos* 30 (4) (2020) 041102, <http://dx.doi.org/10.1063/5.0009454>.
- [46] H.-J. Li, W. Xu, S. Song, W.-X. Wang, M. Perc, The dynamics of epidemic spreading on signed networks, *Chaos Solitons Fractals* 151 (2021) 111294, <http://dx.doi.org/10.1016/j.chaos.2021.111294>.
- [47] B. Blasius, Power-law distribution in the number of confirmed COVID-19 cases, *Chaos* 30 (9) (2020) 093123, <http://dx.doi.org/10.1063/5.0013031>.
- [48] N. James, M. Menzies, Estimating a continuously varying offset between multivariate time series with application to COVID-19 in the United States, *Eur. Phys. J. Spec. Top.* 231 (2022) 3419–3426, <http://dx.doi.org/10.1140/epjs/s11734-022-00430-y>.
- [49] M. Perc, N.G. Miksić, M. Slavinec, A. Stožer, Forecasting COVID-19, *Front. Phys.* 8 (2020) 127, <http://dx.doi.org/10.3389/fphy.2020.00127>.
- [50] J.A.T. Machado, A.M. Lopes, Rare and extreme events: the case of COVID-19 pandemic, *Nonlinear Dynam.* 100 (2020) 2953–2972, <http://dx.doi.org/10.1007/s11071-020-05680-w>.
- [51] N. James, M. Menzies, H. Bondell, Comparing the dynamics of COVID-19 infection and mortality in the United States, India, and Brazil, *Physica D* 432 (2022) 133158, <http://dx.doi.org/10.1016/j.physd.2022.133158>.
- [52] A.S. Sunahara, A.A.B. Pessa, M. Perc, H.V. Ribeiro, Complexity of the COVID-19 pandemic in Maringá, *Sci. Rep.* 13 (1) (2023) <http://dx.doi.org/10.1038/s41598-023-39815-x>.
- [53] N. James, M. Menzies, Global and regional changes in carbon dioxide emissions: 1970–2019, *Physica A* 608 (2022) 128302, <http://dx.doi.org/10.1016/j.physa.2022.128302>.

- [54] M.K. Khan, M.I. Khan, M. Rehan, The relationship between energy consumption, economic growth and carbon dioxide emissions in Pakistan, *Financ. Innov.* 6 (1) (2020) <http://dx.doi.org/10.1186/s40854-019-0162-0>.
- [55] R.G. Derwent, D.R. Middleton, R.A. Field, M.E. Goldstone, J.N. Lester, R. Perry, Analysis and interpretation of air quality data from an urban roadside location in Central London over the period from July 1991 to July 1992, *Atmos. Environ.* 29 (8) (1995) 923–946, [http://dx.doi.org/10.1016/1352-2310\(94\)00219-b](http://dx.doi.org/10.1016/1352-2310(94)00219-b).
- [56] N. James, M. Menzies, Spatio-temporal trends in the propagation and capacity of low-carbon hydrogen projects, *Int. J. Hydrogen Energy* 47 (38) (2022) 16775–16784, <http://dx.doi.org/10.1016/j.ijhydene.2022.03.198>.
- [57] E.J. Westmoreland, N. Carslaw, D.C. Carslaw, A. Gillah, E. Bates, Analysis of air quality within a street canyon using statistical and dispersion modelling techniques, *Atmos. Environ.* 41 (39) (2007) 9195–9205, <http://dx.doi.org/10.1016/j.atmosenv.2007.07.057>.
- [58] N. James, M. Menzies, Equivalence relations and L^p distances between time series with application to the Black Summer Australian bushfires, *Physica D* 448 (2023) 133693, <http://dx.doi.org/10.1016/j.physd.2023.133693>.
- [59] S.K. Grange, D.C. Carslaw, A.C. Lewis, E. Boleti, C. Hueglin, Random forest meteorological normalisation models for swiss PM₁₀ trend analysis, *Atmos. Chem. Phys.* 18 (9) (2018) 6223–6239, <http://dx.doi.org/10.5194/acp-18-6223-2018>.
- [60] N. James, M. Menzies, Distributional trends in the generation and end-use sector of low-carbon hydrogen plants, *Hydrogen* 4 (2023) 174–189, <http://dx.doi.org/10.3390/hydrogen4010012>.
- [61] C. Libiseller, A. Grimvall, J. Waldén, H. Saari, Meteorological normalisation and non-parametric smoothing for quality assessment and trend analysis of tropospheric ozone data, *Environ. Monit. Assess.* 100 (1–3) (2005) 33–52, <http://dx.doi.org/10.1007/s10661-005-7059-2>.
- [62] N. James, M. Menzies, Dual-domain analysis of gun violence incidents in the United States, *Chaos* 32 (2022) 111101, <http://dx.doi.org/10.1063/5.0120822>.
- [63] M. Perc, K. Donnay, D. Helbing, Understanding recurrent crime as system-immanent collective behavior, *PLoS One* 8 (10) (2013) e76063, <http://dx.doi.org/10.1371/journal.pone.0076063>.
- [64] N. James, M. Menzies, J. Chok, A. Milner, C. Milner, Geometric persistence and distributional trends in worldwide terrorism, *Chaos Solitons Fractals* 169 (2023) 113277, <http://dx.doi.org/10.1016/j.chaos.2023.113277>.
- [65] H.Y.D. Sigaki, M. Perc, H.V. Ribeiro, History of art paintings through the lens of entropy and complexity, *Proc. Natl. Acad. Sci.* 115 (37) (2018) <http://dx.doi.org/10.1073/pnas.1800083115>.
- [66] M. Perc, Beauty in artistic expressions through the eyes of networks and physics, *J. R. Soc. Interface* 17 (164) (2020) 20190686, <http://dx.doi.org/10.1098/rsif.2019.0686>.
- [67] N. James, M. Menzies, H. Bondell, In search of peak human athletic potential: a mathematical investigation, *Chaos* 32 (2022) 023110, <http://dx.doi.org/10.1063/5.0073141>.
- [68] A. Clauset, M. Kogan, S. Redner, Safe leads and lead changes in competitive team sports, *Phys. Rev. E* 91 (6) (2015) 062815, <http://dx.doi.org/10.1103/physreve.91.062815>.
- [69] N. James, M. Menzies, Optimally adaptive Bayesian spectral density estimation for stationary and nonstationary processes, *Stat. Comput.* 32 (2022) 45, <http://dx.doi.org/10.1007/s11222-022-10103-4>.
- [70] H.Y.D. Sigaki, M. Perc, H.V. Ribeiro, Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market, *Sci. Rep.* 9 (1) (2019) 1440, <http://dx.doi.org/10.1038/s41598-018-37773-3>.
- [71] M. Jusup, P. Holme, K. Kanazawa, M. Takayasu, I. Romić, Z. Wang, S. Geček, T. Lipić, B. Podobnik, L. Wang, W. Luo, T. Klanjšček, J. Fan, S. Boccaletti, M. Perc, Social physics, *Phys. Rep.* 948 (2022) 1–148, <http://dx.doi.org/10.1016/j.physrep.2021.10.005>.
- [72] M. Perc, The social physics collective, *Sci. Rep.* 9 (1) (2019) <http://dx.doi.org/10.1038/s41598-019-53300-4>.
- [73] R.K. Pan, S. Sinha, Collective behavior of stock price movements in an emerging market, *Phys. Rev. E* 76 (4) (2007) <http://dx.doi.org/10.1103/physreve.76.046116>.
- [74] D.J. Fenn, M.A. Porter, S. Williams, M. McDonald, N.F. Johnson, N.S. Jones, Temporal evolution of financial-market correlations, *Phys. Rev. E* 84 (2) (2011) 026109, <http://dx.doi.org/10.1103/physreve.84.026109>.
- [75] M.C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T.H. Seligman, T. Guhr, H.E. Stanley, Identifying states of a financial market, *Sci. Rep.* 2 (1) (2012) <http://dx.doi.org/10.1038/srep00644>.
- [76] A.J. Heckens, S.M. Krause, T. Guhr, Uncovering the dynamics of correlation structures relative to the collective market motion, *J. Stat. Mech. Theory Exp.* 2020 (10) (2020) 103402, <http://dx.doi.org/10.1088/1742-5468/abb6e2>.
- [77] L. Laloux, P. Cizeau, J.-P. Bouchaud, M. Potters, Noise dressing of financial correlation matrices, *Phys. Rev. Lett.* 83 (7) (1999) 1467–1470, <http://dx.doi.org/10.1103/physrevlett.83.1467>.
- [78] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, Random matrix approach to cross correlations in financial data, *Phys. Rev. E* 65 (6) (2002) <http://dx.doi.org/10.1103/physreve.65.066126>.
- [79] P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley, Quantifying and interpreting collective behavior in financial markets, *Phys. Rev. E* 64 (3) (2001) <http://dx.doi.org/10.1103/physreve.64.035106>.
- [80] J.-P. Onnela, K. Kaski, J. Kertész, Clustering and information in correlation based financial networks, *Eur. Phys. J. B* 38 (2) (2004) 353–362, <http://dx.doi.org/10.1140/epjb/e2004-00128-7>.
- [81] D.-H. Kim, H. Jeong, Systematic analysis of group identification in stock markets, *Phys. Rev. E* 72 (4) (2005) <http://dx.doi.org/10.1103/physreve.72.046133>.
- [82] S. Drożdż, F. Grümmer, F. Ruf, J. Speth, Towards identifying the world stock market cross-correlations: DAX versus dow jones, *Physica A* 294 (1–2) (2001) 226–234, [http://dx.doi.org/10.1016/s0378-4371\(01\)00119-4](http://dx.doi.org/10.1016/s0378-4371(01)00119-4).
- [83] S. Drożdż, R. Gebarowski, L. Minati, P. Oświęcimka, M. Wątopek, Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects, *Chaos* 28 (7) (2018) 071101, <http://dx.doi.org/10.1063/1.5036517>.
- [84] S. Drożdż, L. Minati, P. Oświęcimka, M. Stanuszek, M. Wątopek, Signatures of the crypto-currency market decoupling from the Forex, *Future Internet* 11 (7) (2019) 154, <http://dx.doi.org/10.3390/fi11070154>.
- [85] S. Drożdż, L. Minati, P. Oświęcimka, M. Stanuszek, M. Wątopek, Competition of noise and collectivity in global cryptocurrency trading: Route to a self-contained market, *Chaos* 30 (2) (2020) 023122, <http://dx.doi.org/10.1063/1.5139634>.
- [86] C.-S.J. Chu, G.-J. Santoni, T. Liu, Stock market volatility and regime shifts in returns, *Inform. Sci.* 94 (1–4) (1996) 179–190, [http://dx.doi.org/10.1016/0020-0255\(96\)00117-x](http://dx.doi.org/10.1016/0020-0255(96)00117-x).
- [87] G. Chen, M. Gan, G. Chen, Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications, *Inform. Sci.* 438 (2018) 46–57, <http://dx.doi.org/10.1016/j.ins.2018.01.029>.
- [88] W.F. Sharpe, Mutual fund performance, *J. Bus.* 39 (S1) (1966) 119–138, <http://dx.doi.org/10.1086/294846>.
- [89] N. James, M. Menzies, J. Chan, Semi-metric portfolio optimization: a new algorithm reducing simultaneous asset shocks, *Econometrics* 11 (1) (2023) 8, <http://dx.doi.org/10.3390/econometrics11010008>.
- [90] C. Calvo, C. Ivorra, V. Liern, Fuzzy portfolio selection with non-financial goals: exploring the efficient frontier, *Ann. Oper. Res.* 245 (1–2) (2014) 31–46, <http://dx.doi.org/10.1007/s10479-014-1561-2>.
- [91] A. Prakash, N. James, M. Menzies, G. Francis, Structural clustering of volatility regimes for dynamic trading strategies, *Appl. Math. Finance* 28 (3) (2021) 236–274, <http://dx.doi.org/10.1080/1350486x.2021.2007146>.
- [92] V. Bhansali, Putting economics (back) into quantitative models, *J. Portfolio Manag.* 33 (3) (2007) 63–76, <http://dx.doi.org/10.3905/jpm.2007.684755>.
- [93] J. Moody, M. Saffell, Learning to trade via direct reinforcement, *IEEE Trans. Neural Netw.* 12 (4) (2001) 875–889, <http://dx.doi.org/10.1109/72.935097>.

- [94] S. Drożdż, F. Grümmer, F. Ruf, J. Speth, Log-periodic self-similarity: an emerging financial law? *Physica A* 324 (1–2) (2003) 174–182, [http://dx.doi.org/10.1016/s0378-4371\(02\)01848-4](http://dx.doi.org/10.1016/s0378-4371(02)01848-4).
- [95] N. James, M. Menzies, G.A. Gottwald, On financial market correlation structures and diversification benefits across and within equity sectors, *Physica A* 604 (2022) 127682, <http://dx.doi.org/10.1016/j.physa.2022.127682>.
- [96] N. James, M. Menzies, Collective dynamics, diversification and optimal portfolio construction for cryptocurrencies, *Entropy* 25 (6) (2023) 931, <http://dx.doi.org/10.3390/e25060931>.
- [97] V. DeMiguel, L. Garlappi, R. Uppal, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financ. Stud.* 22 (5) (2009) 1915–1953, <http://dx.doi.org/10.1093/rfs/hhm075>.
- [98] A. Farago, E. Hjalmarrsson, Small rebalanced portfolios often beat the market over long horizons, *Rev. Asset Pricing Stud.* 13 (2) (2023) 307–342, <http://dx.doi.org/10.1093/rapstu/raac020>.
- [99] N. James, M. Menzies, H. Bondell, Understanding spatial propagation using metric geometry with application to the spread of COVID-19 in the United States, *EPL (Europhys. Lett.)* 135 (4) (2021) 48004, <http://dx.doi.org/10.1209/0295-5075/ac2752>.