Which is Worse: Heavy Tails or Nonlinear Dynamics?

Joshua Traut[*], Wolfgang Schadner[†]

January 13, 2023

Abstract

Heavy tails and time-varying autocorrelation (also called nonlinear dynamics) are both stylized facts of financial returns that destabilize markets. The former are extreme events by definition and the latter is known to cause volatility clustering. This work disentangles the two sources and examines which one does the greater damage, whether the threat can be reduced via diversification, and how an inclusion of nonlinear dynamics can enhance the quality of risk modeling. The analysis is carried out for index time series representing seven different asset classes and for individual stock portfolio time series. The isolation of the stylized facts is achieved under recent developments in surrogate analysis (IAAFT, IAAWT). While tail risk historically received more attention, especially in financial regulation, we observe that the change in autocorrelation is the greater driver of maximum drawdowns and aggregate losses across all time series. We further find that diversification does not yield any protection from those risks. These insights have widespread implications for appropriate policy-making, efficient risk hedging, and investment management.

Keywords: Financial Stability, Nonlinear Dynamics, Tail Risk, Autocorrelation, Volatility Clustering, Heavy Tails, Risk Management

[*]University of St. Gallen, Swiss Institute of Banking and Finance, Unterer Graben 21, 9000 St. Gallen, Switzerland | Email: joshua.traut@unisg.ch | https://www.joshuatraut.com

[†]University of St. Gallen, Swiss Institute of Banking and Finance, Unterer Graben 21, 9000 St. Gallen, Switzerland | wolfgang.schadner@unisg.ch | https://www.w-schadner.com

1 Introduction

Asset returns are known to be neither normally distributed nor of perfect random order. In contrast, they appear to exhibit a heavy-tailed distribution and their randomness (autocorrelation) changes over time, a phenomenon that is also known as nonlinear dynamics. We will use the terms 'time-varying autocorrelation' and 'nonlinear dynamics' interchangeably. Empirical evidence for heavy tails and nonlinear dynamics in asset returns is extensive, which is why these two statistical properties are mutually accepted for financial markets and called stylized facts by now (see Cont, 2001; Jiang et al., 2019; Lux & Alfarano, 2016, for a profound review). Both facts are known to adversely affect market stability. Heavy tails are extreme and thus potentially destabilizing events by definition. Time-varying autocorrelation in returns (nonlinear dynamics) affects the development of returns. For example, equity returns tend to be positively autocorrelated (trending) during bullish times, and negatively autocorrelated (overreacting) during bearish times (Schadner, 2022a). These time-varying dynamics are known to cause volatility clustering, i.e. longer subsequent periods of high volatility. Both stylized facts—heavy tails and nonlinear dynamics—are argued to be nested within investor behavior and therefore hard to rationalize (Lux & Alfarano, 2016). Given these two destabilizing forces, the important question as to which one is worse rises.

Answering this question is important as it directly affects risk management, investment management, and academic research. On the one hand, sound knowledge about the destabilizing effects of heavy tails and nonlinear dynamics can be insightful for proper and cost-effective hedging strategies. On the other hand, exposure to such risks may be rewarded with a risk premium and could be an interesting source of returns for investment managers. From an academic perspective, the influence of both factors on financial stability may answer questions on how to realistically model asset returns.

Besides the areas listed above, we observe that the answer to which one is worse may be most crucial for policy-making. Currently, the regulation appears to focus mostly on the reduction of heavy tails while the effects of nonlinear dynamics remain mostly untouched. The regulation of banks is centralized through the Basel Framework that is developed by the Basel Committee on Banking Supervision (BCBS). It sets minimum regulatory requirements for its 45 members which comprise central banks and bank supervisors from 28 jurisdictions. Therefore, a systematic neglection of an important risk driver may have catastrophic consequences for global financial stability. Given this potentially strong need for action, our paper is tilted toward the policy applications of the research question. Its goal is to provide a basis for whether the focus on heavy tails is appropriate to ensure financial stability or whether nonlinear dynamics of financial markets should be granted more attention. However, the insights are also applicable to the areas listed above.

Within this work, we disentangle the two sources of financial instability via surrogate analysis. The method allows us to transform a return series into surrogates that first, follow a normal distribution while preserving the nonlinear dynamics, second, keep the initial distribution while removing the nonlinear dynamics, and third, follow a normal distribution without nonlinear dynamics. The resulting surrogate series can be interpreted as the original return series for which one or both stylized facts were removed while leaving all other characteristics untouched. Upon these disentangled series we compute maximum drawdown and worst year-over-year losses to assess the adverse impact of each of the two stylized facts. The analysis is carried out for daily index log-returns representing the seven asset classes of stocks, bonds, private equity, foreign exchange, commodities, real estate, and cryptocurrency. Further, we analyze daily stock log-return series of S&P 500 constituents of which we build portfolios with different numbers of portfolio holdings. From this portfolio analysis, we examine both, how diversification impacts the presence of heavy tails and nonlinear dynamics and how those two stylized facts translate into drawdown measures. Lastly, to demonstrate the potential flaws of the current regulation we perform a simplified VaR modeling example that incorporates nonlinear dynamics.

The main contribution of our paper is threefold. To our knowledge, we are the first to evaluate risk metrics on return series from which heavy tails and nonlinear dynamics were isolated while preserving all other statistical properties. Thus, our first contribution is that we quantify which of the two stylized facts more adversely affects market stability. Our second contribution is that we expand this analysis to an equity portfolio context. This allows us to evaluate the benefits of diversification, i.e. whether it reduces the presence and severity of the two risk drivers in portfolios. Lastly, we not only point out the weaknesses of the current regulation but also provide a simple workaround on how potentially flawed regulatory risk metrics can be fixed. In general, this paper attempts to build a bridge between the latest developments in surrogate analysis and real-world applications. We observe that especially in the literature on nonlinear dynamics, asset returns are taken as a go-to dataset in empirical analyses, but most works solely focus on the measurement of statistical characteristics. Our goal is to close the gap between this rather technical literature and broader finance applications. In doing so, we try to avoid overly technical jargon and excessive descriptions of model specifications. Instead, we focus more on the interpretation of our results and attempt to give sound understanding to an audience that may not have been exposed to research on nonlinear dynamics before.

Our findings show that all financial time series that are used in the analysis are exposed to excess heavy tails and nonlinear dynamics. Diversification seems to increase heavy tails while levels of nonlinearity remain unaffected. We observe that removing heavy tails from the original return series reduces maximum drawdowns by about 5% to 19% across asset classes. In contrast, removing nonlinear dynamics has an even greater potential for market stability with a respective reduction in maximum drawdowns of about 20% to 40%. We further find that the removal of both stylized facts only leads to marginal differences compared to the cases when only nonlinearity is removed. The results for worst year-on-year returns yield a very similar picture. In a portfolio context, we discover that both drivers of drawdowns can not be tamed by using diversification as the weapon of choice. We find that the better diversified the portfolio, the stronger the effect of removing heavy tails and/or nonlinear dynamics on maximum drawdowns and worst year-on-year returns. Lastly, our simple model example demonstrates how the neglection of nonlinear dynamics can introduce estimation errors and shows that the inclusion of locally estimated scaling exponents can enhance the quality of risk models.

We conclude that while policy makers and potentially some investment managers largely focus on heavy tails, our empirical insights reveal that hedging/removing nonlinear dynamics is of even greater importance for financial stability and risk reduction. It is, however, worth noting that the topic has gained attention among policy makers who acknowledge that the nonlinear dynamics of financial markets have an impact on financial stability and are often ignored by the assumption of independent and identically distributed (iid) returns in regulations (Anderson & Noss, 2013). However, with our novel analysis, the severity of nonlinear dynamics is quantifiable for the first time and our findings show that a simple acknowledgment is not enough. Rather, the effects of nonlinear dynamics should be considered for future iterations of the regulation as soon as possible.

The structure of the remaining article is as follows: Section 2 gives an overview of the relevant literature for our work. Section 3 provides an introduction to nonlinear dynamics that is designed for readers who are unfamiliar with that matter. Section 4 gives a brief overview of the relevant Basel regulatory framework. Section 5 introduces the methodological concepts for the segregation of the two stylized facts from the return series. Section 6 contains the empirical analysis of the paper. It first assesses the effects of heavy tails and nonlinear dynamics on financial stability. Then it provides a simple example of how nonlinear dynamics can be used in a regulatory modeling context to enhance the quality of risk models. Section 7 concludes and lays out the implications of our work. The Appendix holds supplementary material such as technical details of our model specification.

2 Related Literature

Our paper relates to multiple strains of literature. It expands on studies of nonlinearity and higher order moments in financial returns. Also, it is related to the literature of risk-modeling and return forecasting.

Empirical evidence about the existence of nonlinearity in asset returns was first gathered for US stocks (Ding et al., 1993; Lo, 1991). Later studies confirm the findings for the US and expand the analysis to other financial markets. Those include developed stock markets (Alvarez-Ramirez et al., 2008; Al-Yahyaee et al., 2018; Bogachev et al., 2007; Caraiani, 2012; Di Matteo, 2007; Di Matteo et al., 2003; Grech, 2016; Shahzad et al., 2017; Turiel & Pérez-Vicente, 2003, 2005), emerging stock markets (Cajueiro & Tabak, 2004; Caraiani, 2012; Di Matteo, 2007; Di Matteo et al., 2003; Du & Ning, 2008; Kumar & Deo, 2009; H. Y. Wang & Wang, 2018), including high-frequency analysis of both (Gu & Huang, 2019; Kwapień et al., 2005), fixed-income markets (Aloui et al., 2018; H. Y. Wang & Wang, 2018), currency

markets (Al-Yahyaee et al., 2018; Bogachev et al., 2007; Di Matteo, 2007; Di Matteo et al., 2005; Drozdz et al., 2010; Schmitt et al., 2000), cryptocurrency markets (Al-Yahyaee et al., 2018; Ghazani & Khosravi, 2020; Stosic et al., 2019), commodity markets (Al-Yahyaee et al., 2018; Bogachev et al., 2007; Ghazani & Khosravi, 2020; Jiang et al., 2014; Shao, 2020), and credit default swap markets (Aloui et al., 2018). Also, nonlinearity was studied in simulated return series (Bogachev et al., 2008; Bogachev et al., 2007; Grech, 2016).
A theoretical foundation to model the nonlinear dynamics in financial markets was proposed by Peters (1994) based on his criticism of the efficient market hypothesis (Peters, 1991). At its core, his fractal market hypothesis.focuses on heterogeneity of agents with respect to their investment horizons and information processing. With equally represented agents, supply and demand clears smoothly as the same information may cause different actions among investors. Periods of crisis occur if one group of investors dominates the market and their actions cannot be cleared by others. Works that build on the fractal market hypothesis include Blackledge et al. (2019), Dar et al. (2017), Kristoufek (2013), Lamphiere et al. (2021), Li et al. (2014), Rachev et al. (1999), and Weron and Weron (2000). A current more in depth overview of the fractal market hypothesis and its applications in modeling can be found in Blackledge and Lamphiere (2022).

There is a vast literature that studies higher moments in financial returns. However, the literature gets thinner the higher the moment of interest. Regarding the first moment, return volatility, Ang et al. (2006), Ang et al. (2009) find that idiosyncratic volatility has an effect on future returns of stocks. Blitz et al. (2013) and Blitz and van Vliet (2007) make similar discoveries for total return volatility. These findings were reinforced by other studies who find similar results in different financial markets (Baker & Haugen, 2012; Cao & Han, 2013; Chung et al., 2019; Detzel et al., 2019; Joshipura & Joshipura, 2016, 2019).
Empirical works on the third moment, skewness, of asset returns include those of Amaya et al. (2015), Bali and Murray (2013), Barberis and Huang (2008), Boyer et al. (2010), Brunnermeier et al. (2007), Conrad et al. (2013), T. C. Green and Hwang (2012), Harvey and Siddique (2000), and Mitton and Vorkink (2007). Others focus on the co-skewness of assets with the market (Dittmar, 2002; Harvey & Siddique, 2000; Schneider et al., 2020). Theoretical foundations for why skewness and co-skewness levels influence asset prices are

laid out by (Barberis & Huang, 2008; Brunnermeier et al., 2007; Harvey & Siddique, 2000; Kraus & Litzenberger, 1976, 1983; Rubinstein, 1973). It is worth noting that Jin et al. (2022) report that the investment horizon plays a significant role for the pricing of co-skewness. They find that the co-skewness of a stock scales differently between investment horizon and is only priced for a small range of short horizons.
Literature on the fourth moment, kurtosis, of asset returns appears to be relatively scarce in comparison to that on the second and third moment. Studies that examine kurtosis and co-kurtosis in asset returns include Ang et al. (2006), Chang et al. (2013), Conrad et al. (2013), and Dittmar (2002). Theoretical foundations for the pricing of kurtosis are laid out by Dittmar (2002).

Nonlinearity and heavy tails can not be observed in isolation as it is found that both of these stylized facts are interrelated. In general, it is found that nonlinear dynamics can either be caused by heavy tailed distributions or temporal correlation (Kantelhardt et al., 2002). In literature no consensus as to which of the two contribute most to the nonlinear dynamics in financial data is found. While some find that the distribution has a higher contribution (Barunik et al., 2012; Buonocore et al., 2016; Matia et al., 2003; Zhou, 2009) others have provided evidence of the opposite (Benbachir & Alaoui, 2011; H. Chen & Wu, 2011; E. Green et al., 2014; Kwapień et al., 2005; Suárez-García & Gómez-Ullate, 2014; Y. Wang et al., 2011) or that both sources are significantly present (Kumar & Deo, 2009; Norouzzadeh & Rahmani, 2006).

In the history of literature on risk modeling there have been several milestones. Until Mandelbrot (1963) proposed that financial returns exhibit heavy tails and should be modeled following a Pareto-Lévy distribution, returns were thought to follow geometric Brownian motion (Bachelier, 1900; Osborne, 1959). Following works neglected Mandelbrot's theory and found that while returns are indeed heavy-tailed, they they are not so heavy as to be described by Lévy laws (Fama, 1976; Hsu et al., 1974; Officer, 1972). The seminal autoregressive conditional heteroskedasticity (ARCH) model by Engle (1982) became widely acknowledged and was able to capture volatility clustering of returns. As a side effect, the distributions created by this model exhibit heavier tails than the normal. A generalization of the ARCH model (GARCH) was proposed by Bollerslev (1986). Multiple variations of the ARCH model were developed to better capture further characteristics of financial returns. For example, the exponential ARCH (EGARCH) model was developed to account for the asymmetric autocorrelation between positive and negative price variations (Nelson, 1991). To better capture long-term memory of variance integrated models such as the IGARCH model (Engle & Bollerslev, 1986) and its fractional successor FIGARCH (Baillie et al., 1996) were proposed. The ARCH model (or some derivation thereof) is one of the most acknowledged models for modeling financial returns to date. However, these models are argued to spuriously handle the fractal nature of financial returns (Calvet & Fisher, 2002; He & Wang, 2017; Lv & Shan, 2013; Schmitt et al., 1999) and various studies have shown that fractal models have a higher goodness of fit than different variations of ARCH models (H. Chen & Wu, 2011; W. Chen et al., 2014; Frezza, 2014; Günay, 2016; Lux, 2001; Segnon & Trede, 2018; Wei & Wang, 2008).

3 Introduction to Nonlinear Dynamics

While most readers with a background in economics are familiar with heavy-tailed distributions, the concept of nonlinear dynamics may require a little more explanation. A time series is said to exhibit nonlinear dynamics whenever the autocorrelation between observations varies in time. For example, imagine you observe the return series of an asset. You find that ten years ago the return of a given day was always positively correlated to the return of the previous day. Also, you find that this pattern is not observable for the past year. This exemplary asset is exposed to nonlinear dynamics because its autocorrelation structure was not constant over time. In general, the serial dependence in time series measures the series' self-similarity and nonlinear dynamics measure how those similarity patterns change over time.

The concept of nonlinear dynamics has a wide range of application that goes far beyond finance. For illustrative purposes, take, for example, a tree. It can be seen as a complex network of branches that become continually smaller. You will notice that if you zoom in, the shape of successive branches resembles that of the previous generation so the structure of the whole tree exhibits self-similarity. However, it may be that at different levels of resolution the degree of self-similarity varies (i.e. that larger branches may look less alike than their smaller counterparts and some may even look random). The tree exhibits nonlinear dynamics as it is a network that cannot be described by linear equations but yet it is not fully random.

Let us take the analogy of examining a tree from different levels of resolution to financial markets. In a similar fashion we can study the similarity of financial return series at different time scales. An example is to examine whether long-term economic cycles exhibit the same degree of self-similarity as individual trading days or months. The measurement of self-similar structures is not a simple task to do. However, self-similarity implies that the series under investigation exhibits some level of persistence. A way to measure that persistence is to quantify the distance traveled between two points. For illustrative purposes, think about the Pacific coast line of the US. Same as with the tree, it exhibits a degree of self-similarity as certain sections of the coast line may look very similar to others and they are not formed completely random. If you were to measure the length of that coast line between two points with an imaginary mile-long ruler, you would end up with a certain length. If you now take a meter-long ruler, the resulting length would grow as you are able to measure small twists and turns more accurately. Consequently, the length you measure depends on your ruler length (or scale). In fact, it grows as a power law function of your ruler length.

Applying this example to financial markets, the outcome of the persistence estimation strongly depends on the time scaling of returns (the ruler). We can be certain that if we calculate the sum of absolute returns over any time interval of a financial time series using daily returns the result would be smaller as if we used 5-minute returns instead. Another way of measuring the distance between two points in a financial series is to compute the volatility of returns over that time interval defined as

$$V_t(s) = \sqrt{\mathbb{E}\left[(R_t(s) - \mathbb{E}[R_t(s)])^2\right]} \tag{1}$$

for a log-return series $R_t(s)$ over a time window of s from time t to $t+s$. $\mathbb{E}$ denotes expectations to enhance readability. Following the above discovery we can expect $V(s)$ to exhibit the following power-law relation

$$V(s) = V_0 s^H \tag{2}$$

where V_0 denotes the base level of the volatility. Here, we are only interested in the scaling relation with respect to the time scale s so we can reduce the equation as

$$V(s) \sim s^H \qquad \Longrightarrow \qquad \log(V(s)) \sim H \log(s) \tag{3}$$

In the above equations H denotes the scaling exponent that is also known as the Hurst exponent (Hurst, 1956). For price-like series it has the property of $H \in (0,1)$ and is a measure of the returns' autocorrelation or persistence.[1] If returns are iid they are uncorrelated and $H = 0.5$. If $H > 0.5$ returns are persistent whereas if $H < 0.5$ they are anti-persistent. In finance the iid assumption is widely employed which is why the volatility of returns is commonly scaled via the square-root of time (i.e., $s^{0.5}$). Nonlinearity refers to the circumstance that H is not constant but rather varies in s. This is also called 'multifractality' or 'multiscaling'. For financial return series this means that returns in small time scales may exhibit higher (or lower) autocorrelation than those in larger time scales. The degree of nonlinear dependence measures the fluctuation of H in s (i.e., the range of H across time scales). In general, the degree of autocorrelation has a strong influence on volatility and thus drawdowns and tail risk. If there is a high degree of autocorrelation, returns tend to trend and deviate strongly from their mean, leading to a high volatility. Whereas low to negative autocorrelation will cause returns to cancel out each other and stay closer around their mean, leading to lower volatility.

In the presence of high nonlinear dynamics, scaling volatility via a constant factor can lead to estimation errors as each time scale requires its unique scaling exponent. An example for these errors can be found in Table 1. The table holds the volatilities for log-returns of the S&P 500 index at different time scales. Note, that the underlying return series are daily returns and higher return intervals were generated using accumulation. Thus, the sample

[1] There is vast empirical evidence that the power-law scaling properties of Equation 3 hold for financial returns (see Jiang et al., 2019; Lux & Alfarano, 2016, for a profound review).

size for the volatility estimates is constant for each time interval but samples overlap. The bold numbers are the calculated volatilities using the return series of the same time scale. The remaining entries in each row hold volatility estimates from a different time scale that were scaled up or down following the common convention of scaling via the square-root of time, $s^{0.5}$. The values in parentheses underneath the scaled estimates show the respective estimation error. The results show that using the common scaling convention can lead to large errors. For example, estimating yearly return volatility using the scaled daily return volatility yields an overestimation of 19%.

Table 1: Volatility Scaling Errors of S&P 500

This table reports volatilities of the S&P 500 log returns at different return intervals between 1 Januar 1996 and 31 December 2021. The returns are scaled to the respective intervals via accumulation. Each row holds the volatility of returns for each period. The bold numbers are the actual volatilities as calculated from the respective return interval. The remaining entries in each row hold the volatility estimates of estimates from a different interval that are scaled down/up following the common convention of multiplying the volatility estimate with the square-root of the respective time interal, $s^{0.5}$, to which the volatility should be scaled. The numbers in parantheses underneath each row report the estimation error of the scaled volatility estimates.

	Daily	Weekly	Monthly	Yearly	10 Years
Daily	**1.21%**	1.08%	1.06%	1.02%	0.92%
		(-10.7%)	*(-12.9%)*	*(-16.0%)*	*(-24.5%)*
Weekly	2.71%	**2.42%**	2.73%	2.23%	2.01%
	(12.0%)		*(12.7%)*	*(-7.7%)*	*(-17.0%)*
Monthly	6.64%	4.19%	**4.72%**	4.65%	4.18%
	(40.6%)	*(-11.2%)*		*(-1.6%)*	*(-11.5%)*
Yearly	19.17%	17.45%	16.36%	**16.10%**	14.48%
	(19.0%)	*(8.4%)*	*(1.6%)*		*(-10.1%)*
10 Years	60.62%	55.19%	51.73%	50.92%	**45.80%**
	(32.4%)	*(20.5%)*	*(13.0%)*	*(11.2%)*	

4 Regulatory Framework

Within this paper we focus on the MAR standard of the Basel Framework that describes how to calculate capital requirements for market risk. In general, the Basel Framework sets our minimum regulatory requirements for its 45 members that comprise central banks and bank supervisors from 28 jurisdictions. These requirements are then adapted by its members into local laws. At the end of 2022 there was a change of the Basel regulation. Because the duration for a regional implementation of the regulations may be different between members, we lay out the most current and prior version of the regulation below briefly and highlight how heavy tails and nonlinear dynamics are included or neglected.

In the recent Basel 2.5 framework that was effective until December 31, 2022, in determining their market risk for capital requirements banks could choose between a standardized and an internal model approach (BCBS, 2019, MAR10.11). The primary risk measure for the assessment of the capital requirements was the 99^{th} percentile value-at-risk (VaR). In calculating VaR, banks had to use 10-day returns. They were allowed to use "VaR numbers calculated according to shorter holding periods scaled up to ten days by, for example, the square root of time" (BCBS, 2019, MAR30.14 (3)). Heavy tails were accounted for in this regulation as the VaR measure with its focus on an outer percentile of the return distribution is a tail-measure by construction. However, the effect of nonlinear dynamics was neglected twice. First, only one time horizon (ten days) is to be used in the assessment of risk. Second, the use of a uniform scaling factor (square root of time) does not acknowledge for the unique scaling behavior of different financial assets.[2]

In the subsequent Basel 3 framework banks can again opt for an internal risk model approach (BCBS, 2020, MAR11.7). Within this approach, risk is assessed using the 97.5^{th} percentile expected shortfall (ES) (BCBS, 2020, MAR33.3). Banks are required to calculate the ES at a base horizon of ten days that is not to be scaled from shorter time horizons. They further have to specify a set of risk factors for which the regulation sets a respective time horizon. Depending on the risk factor exposure the bank has to scale a modified version of the 10-day ES with the square root of a time delta that is determined by the prescribed time horizon. The capital requirements are then determined by the sum of the base ES and the sum of the scaled ES that are determined by the risk factor exposure (BCBS, 2020, MAR33.4). Again, heavy tails are captured by the ES measure by construction. This comes at no surprise as one of the reasons for the regulatory revision was to better capture tail risk (BCBS, 2019). Nonlinear dynamics are again not rightfully accounted for as the base ES measure is scaled

[2] Readers who are unfamiliar with the concept of nonlinear dynamics may address section 3 to follow along this evaluation of the regulation.

with the square root of time for periods beyond ten days.

5 Methodology

Surrogate analysis is a well-established tool for hypothesis testing across scientific disciplines (see e.g., Keylock, 2019). It allows to transform time series to match a specific feature—like the distribution or the autocorrelation—while preserving the other characteristics of the original series. The advantage over a comparable simulation approach is that surrogates do not require knowledge about the underlying stochastic process, hence are less assumption sensitive. As with simulations, robust conclusion can be only drawn from a larger number of surrogates. Our methodology is closely related to that of Barunik et al. (2012), Buonocore et al. (2016), E. Green et al. (2014), and Zhou (2009) and many others who compare financial data to surrogate time series and draw conclusions on certain characteristics of the original data.

5.1 Keep Heavy Tail, remove Nonlinearity: IAAFT

The simplest example to create a surrogate series that removes all dependence among observations is a random shuffle. While the distribution itself stays untouched, its order in time is randomized which in turn destroys all linear and nonlinear autocorrelation. However, asset returns do not follow a perfect random walk empirically and typically have a modest degree of linear dependence (see e.g., Li et al., 2016). The level of linear dependence is one of the factors that determine the scaling of volatility and consequently impacts time-aggregated risk metrics. Thus, keeping the same linear dependence structure of the original series while removing all nonlinear dynamics is crucial to make a fair evaluation of the latter in terms of risk. A simple random shuffle of returns is thus inappropriate in our context.

To keep the original level of linear dependence, we use the iterated amplitude adjusted Fourier transform (IAAFT) of Schreiber and Schmitz (1996) to create our surrogate data. This algorithm allows us to destroy the nonlinear dependence while preserving the original return distribution as well as their original linear dependence. In a nutshell, the algorithm reorders the original returns in time while preserving their linear autocorrelation structure.

In more detail, the algorithm can be split into three steps:

1. Perform a Fourier transform of the original data. This transformation splits the original series into its power spectrum and phase. The power spectrum describes the periodicity of the original series and reveals repetitive patterns as well as their strength. It is the linear autocorrelation structure of the original data. The phase gives information about how the different patterns of the original series are ordered in time.
2. Perform a random shuffle of the original data
3. Iterate the following two steps until a convergence criterion is fulfilled or any changes are so small that they do not cause a re-ordering of the values of the previous iteration:
   (a) Perform a Fourier transform of the shuffled (newly generated) series and retrieve its power spectrum and phase. Replace the power spectrum with the power spectrum of the original series that was retrieved in step 1. Perform an inverse Fourier transform to generate a new dataset. Given the initial random sort, this means that this new dataset has the same spectrum (linear autocorrelation structure) as the original data but with phases (different order in time) of the shuffled (newly generated) series.

   (b) Replace the values of the series created above with those of the original series in a rank-order matching process. This means that the largest observation of the new series is replaced with the largest observation of the original series and so forth until all observations of the new series are replaced. This preserves the original distribution of the series. However, the replacement process has an influence on the power spectrum of the resulting series and thus deteriorates the quality of the matching of the linear autocorrelation structure. Because of this, steps (a) and (b) are repeated until the difference between the linear autocorrelation structure of the original series and the surrogate series is minimized.

The resulting surrogate series of this algorithm precisely matches the distribution function of the original series as it simply reorders the original observations in time. The matching of the linear autocorrelation structure is fulfilled until a given error tolerance. Because the algorithm performs a random shuffle before entering the iterative matching process, multiple different surrogates can be created from the same original time series.[3]

5.2 Remove Heavy Tail, keep Nonlinearity: IAAWT

Preserving the time-varying autocorrelation structure is a bit more difficult to handle (see Keylock, 2019). Recall, that nonlinear dependence can be understood as the linear autocorrelation structure to vary in time. To preserve this time-dependent structure while transforming the original returns into a surrogate series that has equal statistical properties we use Keylock's (2017) innovation of the iterated amplitude adjusted wavelet transform (IAAWT). This algorithm resembles the nonlinear dependence structure and not solely the overall degree of nonlinearity (like Paluš, 2008). A natural consequence of keeping the nonlinear dependence structure is for example, that the surrogate's volatility clusters around the same time of when crises occurred. In short, the algorithm works relatively similar to the IAAFT—it reorders the observations of the original time series while preserving its nonlinear autocorrelation structure.

In more detail the algorithm can be split into the following steps:

1. Perform a wavelet transform of the original data. Same as the Fourier transform, the transformation allows to split the original series into its power spectrum and phase. However, the wavelet transform provides more detailed insights into the autocorrelation structure of the original series. While the Fourier transform only retrieves the overall linear autocorrelation structure , the wavelet transform also documents where and how that structure changed in time. In other words, the wavelet transform captures the nonlinear dynamics of the signal. Same as with the Fourier transform, that information is captured in the power spectrum component of the transformation.

2. Perform a random shuffle to the original data.

3. Iterate the following two steps until a convergence criterion is fulfilled or any changes are so small that they do not cause a re-ordering of the values of the previous iteration:

[3]For readers who seek for further explanatory material of this algorithm: A good visualization of the algorithm can be found in Venema et al. (2006). An explanation why the power spectrum of the Fourier transform equals the linear autocorrelation structure of the original series can be found in Vaseghi (2008).

(a) Perform a wavelet transform to the randomly shuffled (newly generated) data series and obtain the new phases, combine these with the original power spectrum that was retrieved in step 1. Perform an inverse wavelet transform to generate a new dataset. Again, the resulting series has the same nonlinear autocorrelation structure as the original series but that structure is differently ordered in time.

(b) Replace the values of the new series with those of the original series in the same rank-order matching process as in the IAAFT algorithm. The highest observation of the newly created series is replaced with the highest observation of the original series and so forth until all observations of the new series are replaced. Again, the replacement process has an influence on the power spectrum of the resulting series so the above two steps have to be repeated until the difference in nonlinear structure between the original series and the surrogate series is minimized.

Again, the resulting surrogate series of the algorithm precisely matches the distribution function of the original series as it simply reorders the original observations in time. The matching of the nonlinear autocorrelation structure is matched until a given error tolerance. Same as with the IAAFT, because the algorithm performs a random shuffle before entering the iterative matching process, multiple different surrogates can be created from the same original time series.[4]

5.3 Illustrative Example of the Surrogate Algorithms

As we have explained above, the IAAFT and IAAWT algorithm can be used to create surrogates that have the same linear/nonlinear autocorrelation structure as the original time series while preserving the original distribution. For parts of our application however, we want to change the distribution of the surrogate data while keeping all other statistical properties of the original data. We achieve this simply by using a different dataset in the rank-order matching process within step 3(b) of the above two algorithms. For our purposes we fit a normal distribution of same mean and variance to the original data and thus create a normalized version of the original dataset with the same number of observations. Now, instead

[4]For readers who seek for further explanatory material of this algorithm: A more technical layout of the algorithm can be found in the paper that developed this algorithm (Keylock, 2017).

of replacing the observations in the surrogate time series with the observations of our original data and step 3(b), we take the observations of our normalized data series. This change does not affect the matching of the linear/nonlinear autocorrelation structure because it is explicitly controlled for as the convergence criterion in both algorithms after the match-making process. Within this paper, whenever we perform normalization of the distribution function we will mark the respective algorithm with the subscript n. Hence, if we perform a normalization within the IAAFT (IAAWT) algorithm, we will show it in the notation as IAAFT_n (IAAWT_n). Note, that this change of distribution can also be performed using any other distribution function without affecting the desired correlation structure matching of both algorithms.

An illustration of how we use the IAAFT and IAAWT algorithms to change the statistical properties of return series can be found in Figure 1. The figure holds the original log-return series of the S%P 500 index between the beginning of 1926 and the end of 2021 in the upper time series. It can be observed that the time series exhibits bursty behavior that is especially prominent during periods of financial distress. This volatility clustering shows that the time series is most likely prone to nonlinear dynamics. Further, it can be seen that while most returns fall closely around 0% there are a couple of very extreme observations, which make the return distribution heavy-tailed.

Below the original time series, three exemplary surrogate time series are illustrated. The series underneath the original returns was created using the IAAWT_n algorithm. It can be seen that the algorithm normalized the distribution as all heavy tails from the original returns are removed, making the whole distribution much more compact (Note the change of scale on the y-axis). The nonlinear dynamics of the original return series remain untouched as the bursts in times of financial distress can also easily be identified in the surrogate series. The series in the third row of the figure was created using the IAAFT algorithm. It can be seen that it exhibits the same number and amplitude of extreme observations as the original return series. This comes at no surprise because the individual observations within this surrogate series are the same as within the original returns. When looking at the course of this surrogate series it appears much smoother than the original and does not exhibit such bursty, nonlinear behavior. The bottom series was created using the IAAFT_n algorithm. Similar to the data series in the second row, because of the normalization its extreme observations are not as severe as that of the original return series. Additionally, and very similar to the series in row 3, its course appears to be much smoother and no volatility clusters are apparent.

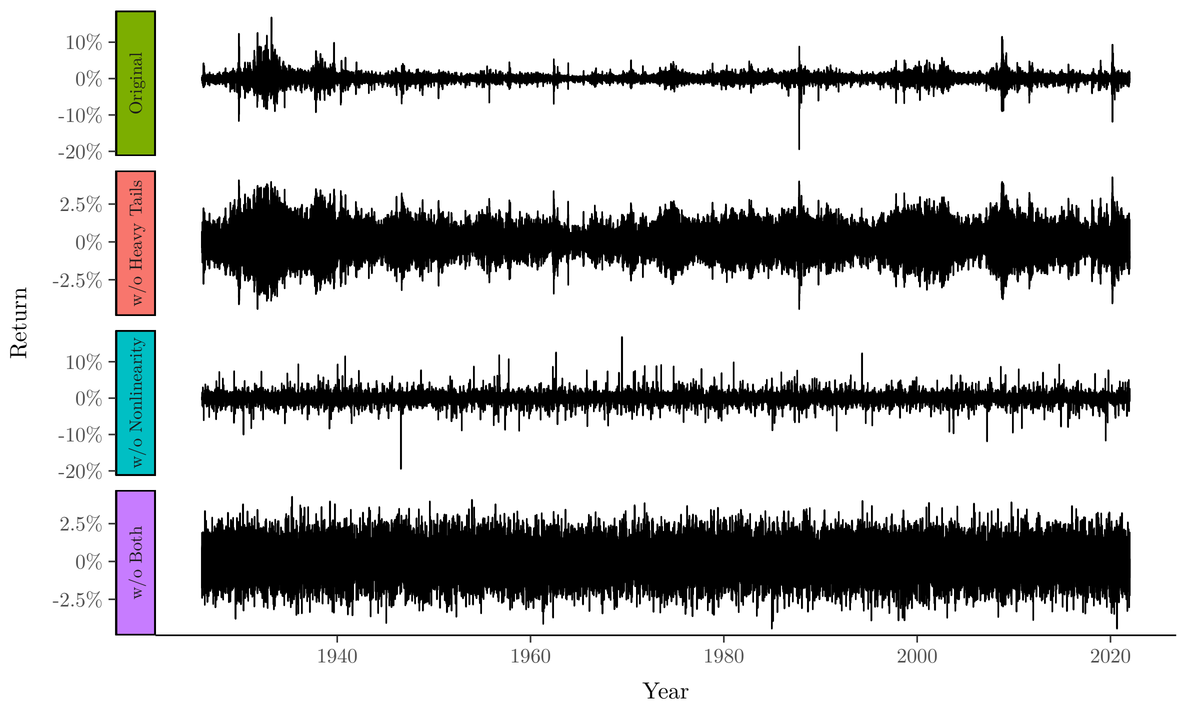

Figure 1: Exemplary comparison of original and surrogate time series for the S%P 500 Index
The figure shows the returns for the original S%P 500 Index and for two surrogate series that do not exhibit heavy tails or nonlinear dynamics. The time series range from the beginning of 1926 until the end of 2021. The surrogate series without heavy tails was created using the IAAWT$_n$ method. The surrogate series without nonlinearity was created using the IAAFT method. The surrogate series without both stylized facts was created using the IAAFT$_n$ method.

6 Empirical Analysis

The empirical analysis of our work can be separated into four parts. First, we provide an overview of the data that is used in the following chapters. Second, we disentangle the sources of financial instability for seven index return series to assess how different financial markets are affected by heavy tails and nonlinear dynamics. Third, we analyze whether diversification can protect investors from the destabilizing effects of the two stylized facts. Lastly, we show how the neglection of nonlinear scaling dynamics can cause errors in risk modeling and how those flaws can be resolved in an exemplary example that is inspired by current regulatory requirements.

6.1 Data

Our analysis focuses on two datasets. First, we analyze seven index time series of daily log-returns, each representing a different asset class. The series have a common ending date at 12/31/2021, the starting date was chosen individually with respect to data availability. The data is described in more detail in Table 2. Second, to gain insights of the performance of portfolios, we analyze daily log-returns of constituents of the S&P 500 index that we pull from Datastream. We solely use constituents for which uninterrupted return data is available between 01/01/1992 and 31/12/2021. Also, we exclude stocks for which returns did not change for a trading week to exclude highly illiquid stocks from the analysis. The total number of stocks in this dataset equals 144.[5] The choice for the time horizon of this dataset was made based on two counteracting objectives. First, the sample size was to be maximized to increase statistical significance of results. Second, the time horizons of each sample constituent was to be kept as long as possible because nonlinearity analyses yield more robust results for longer time series (Shao et al., 2012). The dataset at hand finds a compromise between having a fairly large sample with 144 stocks that covers a relatively long time horizon of above 30 years.

From these 144 stocks we form portfolios at different levels of diversification. We chose the number of portfolio holdings following theoretical and empirical findings in the literature. Studies examining the effect of diversification subject to the number of portfolio holdings report that diversification benefits decrease in the number of holdings and that they plateau with the number of total holdings being around 40 (Evans & Archer, 1968; Mao, 1970; Statman, 1987; Upson et al., 1975). Thus, the number of portfolio holdings we use are 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, and 45. For each of these holding numbers we create 40 equal-weighted portfolios of which the constituents are randomly chosen from the 144 available stocks. This choice to create 40 portfolios is based on the assumption that the central limit theorem holds at that sample size and results are not strongly biased due to outliers. Overall, our resulting total number of portfolios for the analysis is thus 800.

[5] A list of the respective company names can be found in Appendix A.

To get an impression on the presence of the two stylized facts in our dataset, we compute excess kurtosis and nonlinearity of the data. While the computation of excess kurtosis is straight forward, quantifying the degree of nonlinear dependence is a little more involved. For its assessment we compute spreads in Hölder exponents, $\Delta\alpha$, using the basic MF-DFA method of Kantelhardt et al. (2002).[6] These spreads are corrected for spurious inflation by deducting the mean Hölder spread from IAAFT surrogates. An example for such spurious sources is that false nonlinear dependence is introduced as a result of using data sets with finite length and/or heavy tails (Grech & Pamuła, 2012; Grech & Pamuła, 2013; López & Contreras, 2013; Mukli et al., 2015; Pamuła & Grech, 2014; Rak & Grech, 2018). Both false drivers of nonlinear dynamics can be estimated by transforming the original time series into IAAFT surrogate series. As described in subsection 5.1 the transformed IAAFT surrogate series keep the same distribution as the original but do not exhibit nonlinear dynamics. Therefore, nonlinearity measured on the surrogate time series should only be composed of the two spurious sources. The true nonlinearity can be retrieved via subtracting the Hölder spreads of the IAAFT surrogates, $\Delta\alpha_{surr}$, from the Hölder spreads of the original time series, $\Delta\alpha$:

$$\Delta\alpha_{nl} = \Delta\alpha - \Delta\alpha_{surr} \tag{4}$$

The excess spreads, $\Delta\alpha_{nl}$, thus represent the magnitude of pure nonlinear dependence in the original time series. We follow Schreiber and Schmitz (2000) who propose to use at least 19 (39) surrogate series for one-sided (two-sided) statistically significant results. Thus, we create 40 IAAFT surrogates and use their mean Hölder spread as $\Delta\alpha_{surr}$ in the above equation. We find a strong convergence of $\Delta\alpha_{surr}$ for IAAFT surrogate samples larger than 20. From this we conclude that the use of a larger surrogate sample would not alter the results of our nonlinearity estimation. In other words, we find that the results of the nonlinearity measurement are robust. This method to retrieve a pure measure of nonlinear dynamics was

[6] MF-DFA is one of the standard tools that is most widely used for characterizing nonlinear dynamics (see Grech, 2016; Jiang et al., 2019). Its implementation is discussed in thorough detail by Ihlen (2012), thus not repeated here. A recap on the MF-DFA algorithm and the choice of our model specifications are laid out in Appendix B.

also carried out by Pamuła and Grech (2014), Rak and Grech (2018), Schadner (2022b), and Zhou (2012) and many more.

Results for the assessment of the two stylized facts are reported in Table 2 for the index data and in Figure 2 for the stock portfolio data. The excess kurtosis for both data sets indicates that all series are characterized by a heavy-tailed distribution. From the excess Hölder spreads we observe that all return series are characterized by nonlinear dynamics. With regards to the stock portfolios two more patterns can be observed. First, heavy-tailedness seems to increase with the number of portfolio holdings while the degree of nonlinear dynamics appears to be independent of the number of holding. Second, the distribution of both measures decreases with the number of holdings. This is likely attributable to the fact that the more holdings a portfolio has, the more idiosyncratic risk is diversified away and the more similar the portfolios get.

Table 2: Description and summary statistics of the index data

This table holds data description and summary statistics for daily log-returns of the seven indices that were used in the analysis. All series end with 12/31/2021. The second-last column reports excess kurtosis, which indicates heavy-tailedness. The last column holds nonlinearity measured by the excess spreads in Hölder exponents using the MF-DFA method of Kantelhardt et al. (2002).

Asset Class	Instrument	Source	Start	Kurt.	Nonlin.
Stocks	S&P 500 Index	CRSP	1/2/1926	16.9	.183
Private Equity	LPX Composite Listed P.E. Index	Bloomberg	4/2/2004	23.8	.117
Bonds	S&P 500 Bond Index	Refinitiv	1/3/1995	6.52	.441
Foreign Exchange	GBP/USD rate	Refinitiv	8/19/1971	6.73	.811
Cryptocurrency	FTSE Bitcoin Index	Refinitiv	10/13/2015	3.83	.159
Commodity	Bloomberg Commodity Index	Refinitiv	1/3/1991	4.37	.145
Real Estate	MSCI U.S. Liquid Real Estate Index	MSCI	6/1/2001	18.9	.302

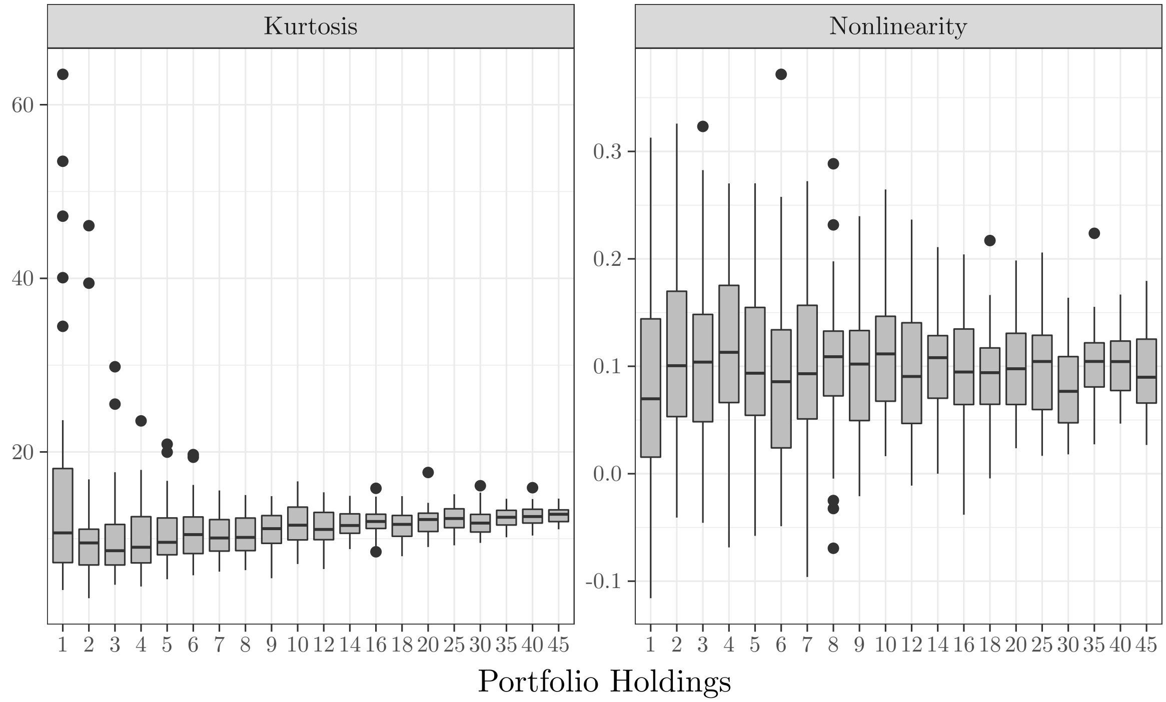

Figure 2: Summary statistics of the stock portfolio data

The figure holds excess kurtosis in the left subfigure and nonlinearity measured by the excess spreads in Hölder exponents using the MF-DFA method of Kantelhardt et al. (2002) in the right subfigure. The portfolios are composed using daily log-returns of randomly chosen constituents of the S&P 500 index for which return data is available between 01/01/1992 and 31/12/2021, resulting into an investible pool of 144 stocks. For each number of holdings 40 equal-weighted portfolios were formed.

6.2 Disentangling the Sources for Indices

For each of the seven return series we create 200 many IAAFT, IAAWT$_n$, and IAAFT$_n$ surrogates of same mean and variance as the original return series. The IAAFT surrogates can be interpreted as if the original returns had no nonlinear dynamics, the IAAWT$_n$ surrogates as if returns would be normally distributed, and the IAAFT$_n$ as a benchmark where both stylized facts are removed from the original return series. Next, upon the surrogate series we compute the risk measures of maximum drawdowns and worst year-over-year return to compare them to each other as well as to the original series. The decision to focus on these two risk measures was made because they are path dependent and can thus be influenced by nonlinear dynamics and heavy tails in the distribution. Other risk measures such as the empirical value at risk or the semi-standard deviation solely focus on the distribution of returns and are unaffected by nonlinear dynamics. In other words, because the distribution of the surrogates is either equal to that of the original dataset or set to be normal, risk measures that solely focus on the distribution will yield inconclusive results. Results for the original series will be identical to those of the IAAFT series and results for the IAAWT_n series will be identical to those of the IAAFT_n series because either of the two pairs share the same distribution.[7] Also, we prefer drawdown measures because they give a good proxy of how prone a return series is to shocks. Thus these measures can be interpreted as proxies for financial stability. The results for maximum drawdowns and worst year-over-year returns are reported in Table 3. Figure 3 visualizes the disentangled maximum drawdowns. A visualization for disentangled worst year-on-year returns is not presented because it looks very similar to that of the maximum drawdowns.

The results show that each of the seven indices has a lower downside risk when nonlinear dynamics rather than heavy tails are removed. We can see that in absolute terms the reduction is greater for asset classes that are more risky, i.e. stocks vs. bonds. However, when looking at the relative change, we find that those are relatively similar across asset classes, at least for maximum drawdowns. In general, we observe that the removal of heavy tails reduces maximum drawdowns by roughly 17%. The removal of nonlinear dynamics reduces maximum drawdowns by around 30%. For worst year-over-year returns the percentages are a little more spread out. Furthermore, we find that when both stylized facts are removed, the improvement in downside risk compared to a return series without nonlinear dynamics is marginal. Hence, reducing nonlinear dynamics in returns has a greater potential for financial stability.

[7]Note, that this holds true only if non-path-dependent risk measures are calculated with the same time scale as the time series under investigation. If those measures were calculated under a different time scale, for example, yearly (rolling) value at risk calculations for daily data, the results would be different for all time series as a result of varying degrees of scaling properties among the time series.

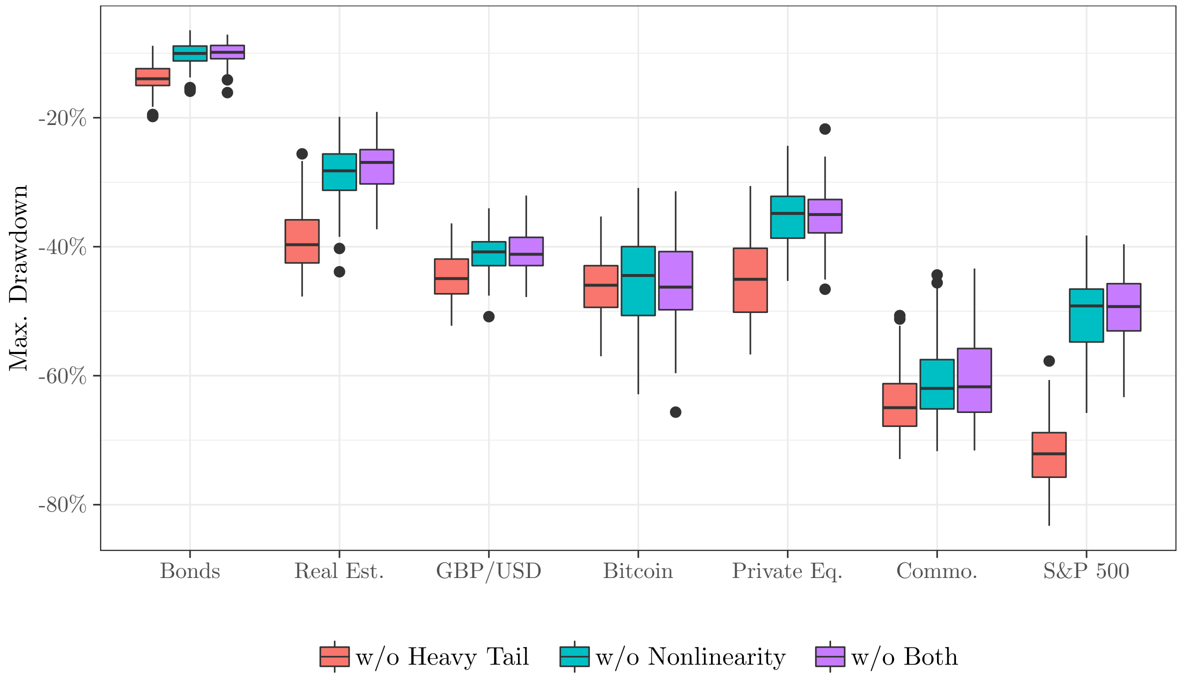

Figure 3: Maximum drawdowns of the different asset classes after disentangling stylized facts
The figure holds maximum drawdowns for three surrogates series of seven indices. Each surrogate series comprises 200 surrogates. The surrogates without heavy tails were created using the IAAWT_n method. The surrogates without nonlinearity were created using the IAAFT method. The surrogates without both stylized facts were created using the IAAFT_n method.

6.3 Disentangling the Sources for Stock Portfolios

In practice the go to weapon against risk is diversification. Hence, it makes sense to evaluate whether the nonlinearity or heavy-tailedness of returns can be influenced with diversification and how these changes translate into risk measures. In Figure 2 we already showed that an increase in portfolio holdings leads to an increase in excess kurtosis while the degree of nonlinear dynamics remains fairly constant across portfolio holdings. Hence, we can already see that diversification does not reduce the presence of the two stylized facts. To determine whether diversification yields protection regarding how the two stylized facts translate into downside measures, we repeat the analysis that we carried out on an index level below.

We create five IAAFT, IAAWT_n, and IAAFT_n surrogate time series for each portfolio, leading to a total of 200 surrogate series for each surrogate methodology and portfolio holding. Again,

Table 3: Maximum drawdowns of original series compared to mean values of surrogates

The table compares maximum drawdowns and worst YoY returns for surrogates series without heavy tails, without nonlinearity, and without both stylized facts of seven indices. Each surrogate series comprises 200 surrogates. The surrogates without heavy tails were created using the IAAWT method. The surrogates without nonlinearity were created using the IAAFT method. The surrogates without both stylized facts were created using the IAAFT method under the constraint that the return distribution of the surrogates is normal. The numbers in parentheses indicate relative reduction compared to original series.

	Bonds	Real Est.	F.X.	Crypto	P.E.	Comm.	Stocks
	max. Drawdown						
Original	-14.8%	-45.8%	-55.0%	-56.3%	-54.4%	-77.2%	-84.8%
w/o Heavy Tail	-14.0%	-39.0%	-44.8%	-46.1%	-45.0%	-64.2%	-72.3%
	(-5%)	(-15%)	(-19%)	(-18%)	(-17%)	(-17%)	(-15%)
w/o Nonlinearity	-10.2%	-28.7%	-41.0%	-45.3%	-35.1%	-61.0%	-50.7%
	(-31%)	(-37%)	(-25%)	(-20%)	(-36%)	(-21%)	(-40%)
w/o Both	-10.1%	-27.6%	-40.9%	-45.6%	-35.1%	-60.6%	-49.8%
	(-32%)	(-40%)	(-26%)	(-19%)	(-36%)	(-22%)	(-41%)
	worst YoY return						
Original	-12.4%	-35.0%	-28.1%	-50.1%	-43.0%	-52.1%	-64.5%
w/o Heavy Tail	-11.2%	-33.2%	-28.6%	-38.3%	-33.8%	-44.0%	-53.2%
	(-10%)	(-5%)	(2%)	(-24%)	(-22%)	(-16%)	(-18%)
w/o Nonlinearity	-8.2%	-23.1%	-25.8%	-35.0%	-26.1%	-37.8%	-39.2%
	(-34%)	(-34%)	(-8%)	(-30%)	(-39%)	(-28%)	(-39%)
w/o Both	-8.2%	-21.9%	-26.1%	-36.2%	-25.7%	-38.1%	-38.1%
	(-34%)	(-37%)	(-7%)	(-28%)	(-40%)	(-27%)	(-41%)

upon the surrogate series we compute the risk measures of maximum drawdowns and worst year-over-year returns to compare them to each other as well as to the original series. Figure 4 holds the results for the maximum drawdown analysis. It also includes the drawdowns of the original portfolios for a better comparison. However, it has to be noted that the data for the original portfolios differs form those of the surrogates as it comprises 40 observations per holding whereas the surrogate time series hold 200 observations per holding. Again, we only show results for maximum drawdowns because those for worst year-on-year returns look very similar.

Five important insights can be gained from the figure. First, diversification works. We observe that the average drawdown for better diversified portfolios is smaller than for concentrated portfolios. Second, removing the stylized facts from the return series leads to a reduction in maximum drawdowns across all levels of diversification. Third, nonlinear dynamics seem to be a bigger driver for drawdown risk as drawdowns of the surrogate series without nonlinearity are always below those of the surrogate series without heavy tails. Fourth, removing both, nonlinear dynamics and heavy tails from the return distribution only leads to a marginal increase in drawdowns compared to a time series where only nonlinear dynamics were removed. Fifth, the removal of the two stylized facts has a larger influence on the reduction of drawdowns at higher degrees of diversification. For the surrogates without heavy tails these results are consistent with the increase of excess kurtosis in portfolio holdings that is reported in Figure 2. If we believe that excess kurtosis has a positive effect on drawdowns it is intuitive that drawdown reduction of the surrogate series without heavy tails increases in portfolio holdings as these portfolios are more exposed to a driver drawdowns. However, with a constant level of nonlinearity across portfolio holdings there is no such obvious explanation for the positive relationship between drawdown reduction and portfolio holdings for the surrogate series without nonlinearity.

To further uncover the different drawdown reduction properties of the surrogate series Figure 5 compares the relative reduction in maximum drawdowns between the surrogate series without heavy tails and surrogate series without nonlinear dynamics and the original portfolio data. It can be observed that the higher the number of portfolio holdings the higher the relative reduction in drawdowns for both surrogate time series. When comparing the ratio between relative drawdown reduction of both surrogate series we find that there seems to be a constant relationship. Except for portfolios with less than three holdings, the relative reduction in drawdowns that is achieved with a surrogate series without heavy tails is roughly 50% of the reduction that can be achieved with a surrogate series without nonlinearity.

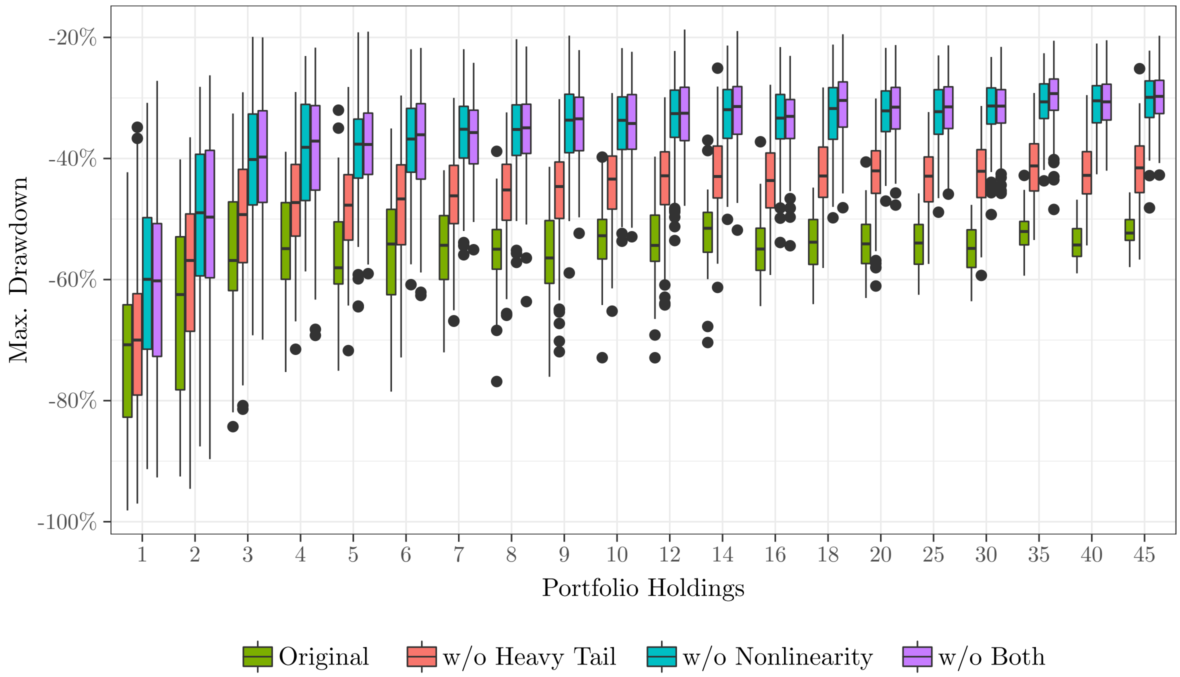

Figure 4: Maximum drawdowns of stock portfolios with different holdings after disentangling stylized facts

The figure holds maximum drawdowns for three surrogates series of stock portfolios with different holdings. The portfolios are composed using daily log-returns of randomly chosen constituents of the S&P 500 index for which return data is available between 01/01/1992 and 31/12/2021, resulting into an investible pool of 144 stocks. For each number of holdings 40 equal-weighted portfolios were formed, for which the drawdowns are illustrated by the green candle. For each portfolio we create five surrogates leading to a total number of 200 time series per portfolio holding. The surrogates without heavy tails were created using the IAAWT_n method. The surrogates without nonlinearity were created using the IAAFT method. The surrogates without both stylized facts were created using the IAAFT_n method.

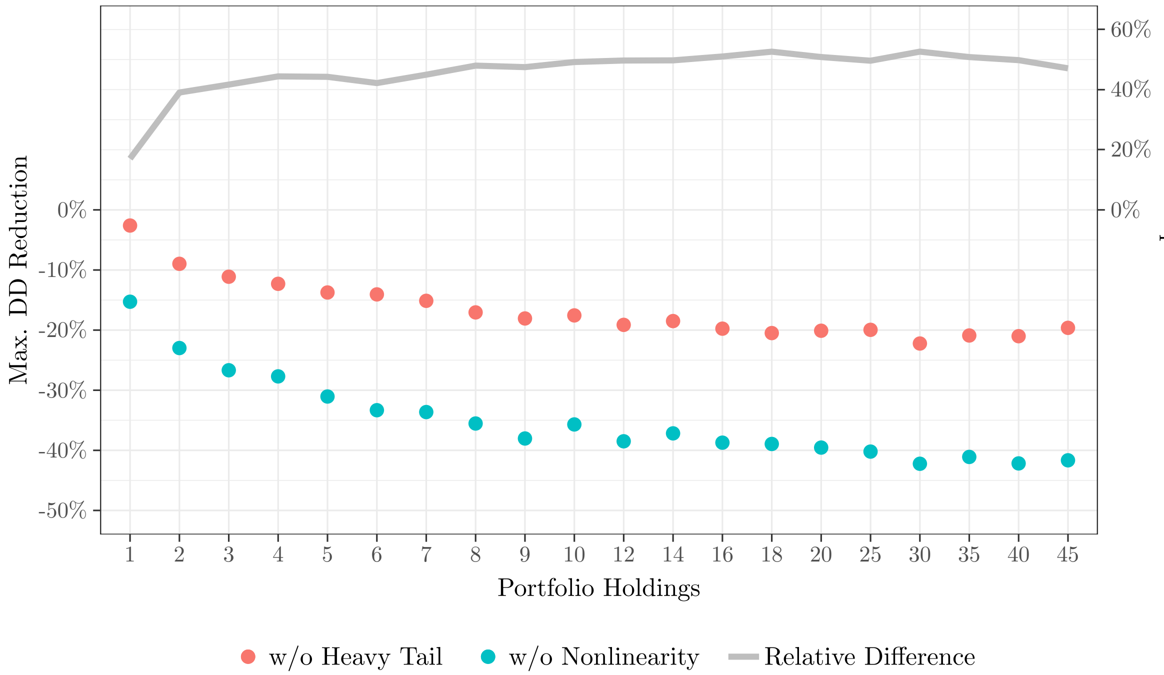

Figure 5: Relative maximum drawdown reduction of stock portfolios with different holdings after disentangling stylized facts

The figure holds mean maximum drawdown reductions for two surrogates series of stock portfolios with different holdings. The portfolios are composed using daily log-returns of randomly chosen constituents of the S&P 500 index for which return data is available between 01/01/1992 and 31/12/2021, resulting into an investible pool of 144 stocks. For each number of holdings 40 equal-weighted portfolios were formed. For each portfolio we create five surrogates leading to a total number of 200 time series per portfolio holding. The surrogates without heavy tails were created using the IAAWT$_n$ method. The surrogates without nonlinearity were created using the IAAFT method. The dots show the relative reduction of maximum drawdowns of the surrogate time series relative to the drawdown of the original time series on the left axis. The grey line shows the ratio between the two time series on the right axis. It thus measures the ratio between the relative reductions of spreads.

6.4 Example for Improved VaR Scaling

To demonstrate how the gained insights from the above analyses can be used in practice, a simple example on backtesting a risk model is outlined below. This exemplary test helps to answer two questions. First, whether improper scaling of variance leads to worse risk estimates and second, whether that flaw can be corrected by the use of correct scaling exponents.

In the setup of the example we are guided by the regulation. Within the old and the new Basel Framework banks are required to backtest their risk models based on the VaR risk measure over the most recent 12 months of data (BCBS 2019, MAR32.4, BCBS 2020, MAR32.4).

The framework is a very straightforward procedure for comparing the risk measures with the actual trading outcomes. To assess the quality of the risk model, banks have to calculate the number of times that the trading outcomes were not covered by the risk measures. This means that a 99% daily VaR should cover 99% of the daily trading outcomes (BCBS 2019, MAR99.34).

Following the regulation we use the VaR as the risk metrics of interest for our model backtest. For this test we use our daily S%P 500 Index time series for the most recent 40 years of data. To keep things simple, our expected VaR equals the VaR that we calculate from the observed mean and variance from the most recent 12 months of data. Then we compare how many times that estimate was exceeded thereafter. We update our estimate on ten day non-overlapping rolling windows leading to a total sample of 1,059 observations. To test the overall quality of the VaR model we repeat the above steps for confidence levels between 90% and 99.9% at steps of 0.5%.

To find an answer to the above two questions, we test three VaR risk models. First, we calculate a daily VaR model that is evaluated against the following one day return. This model serves as a baseline benchmark as the risk model and the evaluation period have the same time scale and no adjustments are necessary. Second, we use the results of the base line model and scale them with $10^{0.5}$ to retrieve a 10-day VaR estimate (Recap, that the base horizon for the capital requirement regulation is 10 days). This model is then compared to the following 10-day return. A comparison between the quality of this second model and our base line model should help answering the question whether ignoring the scaling properties of returns alters VaR estimates. Third, similar to our second risk model, we scale the results of our base model with the local scaling exponent of the most recent estimate within our in sample period to retrieve 10-day VaR estimates. The scale exponent estimates are based on the algorithm of (Ihlen & Vereijken, 2014)[8] using the most recent five years of available data. The decision to use a relatively long history of returns for the scaling exponent estimation is based on the fact that such estimates can yield spurious results for short time series (Shao et al., 2012). Thus we follow the recommendation to set the sample size greater than 1,000 observations (Ihlen, 2012). Also, we estimate the local 10-day scaling exponent using an

[8]More technical details about the algorithm can be found in Ihlen (2012) and Ihlen and Vereijken (2013).

interpolated value for local scale estimations of 9, 10, and 11 days because including more scales is recommended to increase robustness (Ihlen & Vereijken, 2014). We arrive at our 10-day VaR estimate by taking our base 1-day VaR and multiplying it with 10^{H_l} where H_l is our local scaling exponent estimation.[9]

The results of our analysis are summarized in Figure 6. It shows the percentiles of the VaR models on the y-axis and the respective fractions of returns that fall out of the confidence interval on the x-axis. The black line refers to the null hypothesis, i.e. that 1% of returns fall below the 99% VaR estimate. The 95% level confidence bands around that null hypothesis are calculated assuming normally distributed shocks. The results for our base 1-day VaR estimate are represented with black X's. We find that though simple, the VaR estimate performs relatively well as all of its estimates fall within the confidence bands of the null hypothesis. When comparing the results of our estimate using $10^{0.5}$ as the uniform scaling coefficient (red dots) we find that the performance of the VaR model becomes worse. At lower confidence levels between 90% and 95% the model is too conservative whereas at higher levels between 97.5% and 100% the model underestimates the frequency of extreme events. Lastly, when comparing the 10-day VaR model that was constructed using a local scaling exponent we find that the performance of the model is much more in line with the base 1-day model. The flaws of under- and overestimation that were introduced with the uniform scaling exponent appear to be corrected when using correct time scaling. This observation is supported by the data as the average distance from the null hypothesis is 0.88% for the uniformly scaled VaR model and 0.47% for the model using correct scaling. Even more so, the model using local time scales shows the best performance as it most closely scattered around the null hypothesis, meaning that the predictions made by the model are the closest to out of sample outcomes.

[9] We acknowledge that multiplying a VaR forecast with a scaling factor comes at the additional assumption that the mean of the return series is zero. We keep this assumption as it is implicitly embedded in the regulation. Recap on section 1 for details.

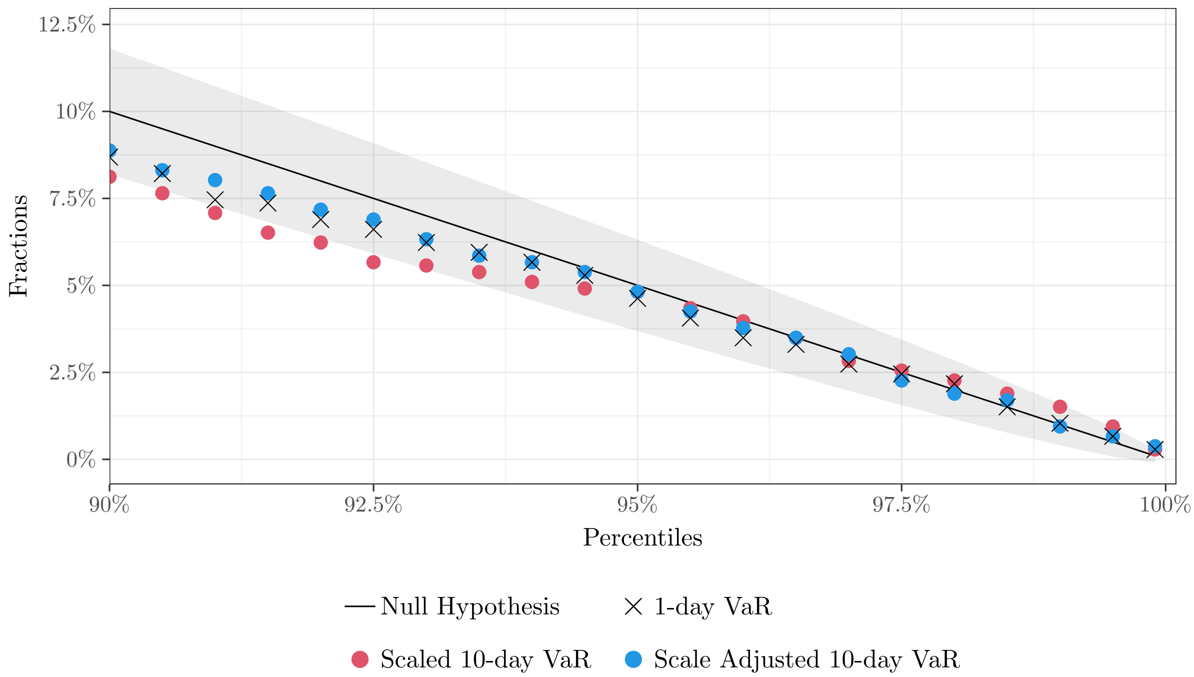

Figure 6: Comparison between scale-adjusted and unadjusted 10-day historical VaR forecast
The figure compares the accuracy of three VaR risk models. The black line refers to the null hypothesis, i.e. that 1% of returns fall below the 99% VaR estimate. The grey area around the hypothesis is the 95% level confidence interval that was calculated assuming normally distributed shocks. The black crosses refer to a 1-day VaR model that was estimated using the historical distribution of the most recent 12 months of data and was evaluated against the following 1-day return. The red dots refer to a scaled version of the 1-day model. This model was scaled with $10^{0.5}$ to retrieve a 10-day VaR estimate and was evaluated against the following 10-day return. The blue dots also refer to a scaled version of the 1-day model. This model was scaled with 10^{H_l} where H_l is the local scaling exponent to arrive at a 10-day VaR estimate and was evaluated against the following 10-day return. The local scaling exponent H_l was estimated based on the algorithm of (Ihlen & Vereijken, 2014) using the past five years of data with an interpolation for local scale estimations of 9,10, and 11 days.

To demonstrate that the results originate from the wrong scaling of variance, we repeat the example with a more sophisticated model setup. Again, we estimate three VaR models. First, a base 1-day VaR model and then two scaled 10-day VaR models of which one is scaled with the square root of time and the other is scaled with the estimated local scaling component. Now we estimate the base 1-day model using an AR(1) model for the mean and a GARCH(1,1) for the volatility. Again, we fit the base model to the most recent 12 months of data and make a 1-day forecast out of sample. All other specifications as well as the scaling for our two 10-day VaR forecasts are the same as before. The results are reported in Figure 7.

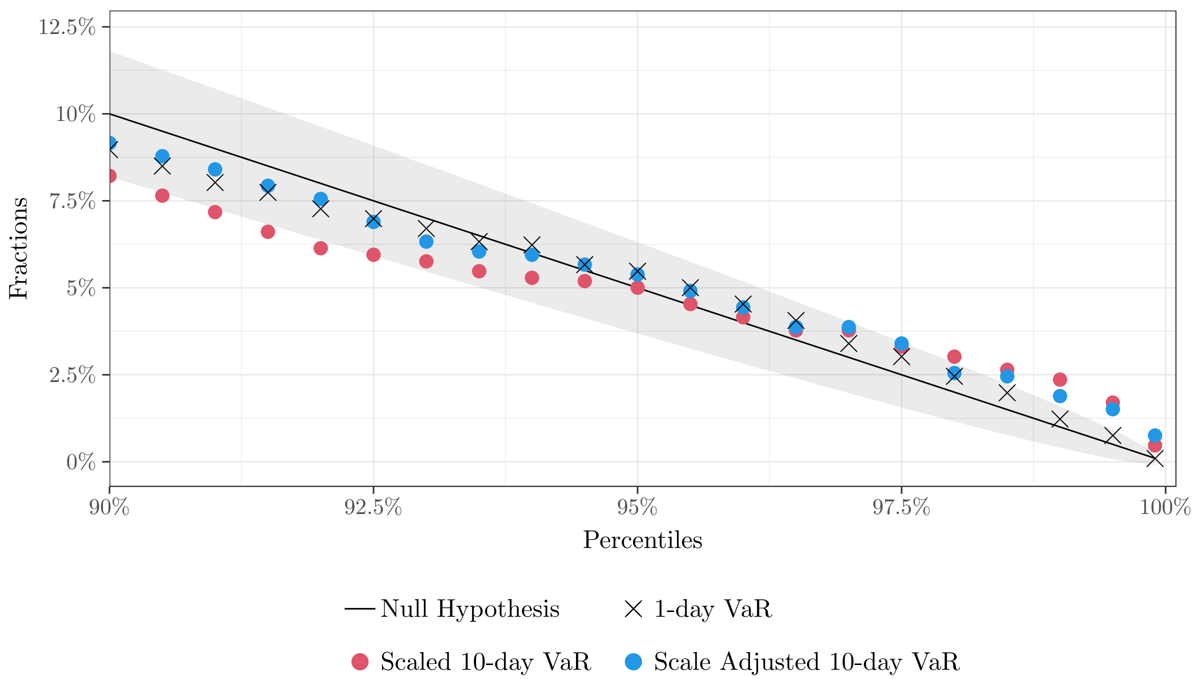

Figure 7: Comparison between scale-adjusted and unadjusted 10-day AR(1) GARCH(1,1) VaR forecast

The figure compares the accuracy of three VaR risk models. The black line refers to the null hypothesis, i.e. that 1% of returns fall below the 99% VaR estimate. The grey area around the hypothesis is the 95% level confidence interval that was calculated assuming normally distributed shocks. The black crosses refer to a 1-day VaR model that was estimated using an AR(1) model for the mean and a GARCH(1,1) model for the variance. We fit both models to the most recent 12 months of data and make a VaR forecast assuming normally distributed shocks that was evaluated against the following 1-day return. The red dots refer to a scaled version of the 1-day model. This model was scaled with $10^{0.5}$ to retrieve a 10-day VaR estimate and was evaluated against the following 10-day return. The blue dots also refer to a scaled version of the 1-day model. This model was scaled with 10^{H_l} where H_l is the local scaling exponent to arrive at a 10-day VaR estimate and was evaluated against the following 10-day return. The local scaling exponent H_l was estimated based on the algorithm of (Ihlen & Vereijken, 2014) using the past five years of data with an interpolation for local scale estimations of 9,10, and 11 days.

Looking at the graph we receive a very similar picture as before. Again, the base 1-day VaR forecast is relatively accurate with all forecasts falling into the 95% confidence interval. Scaling this base forecast to ten days without using the appropriate scaling behavior worsens the performance of the forecast. Respecting the local scaling behavior improves the performance of the scaled VaR estimate such that it is very similar to the base forecast. Again, looking at the average distance from the null hypothesis yields support for what we see from the visual inspection of the figure. The uniformly scaled VaR model misses the null by 1.06% on average whereas the model using the correct scaling has an average error of only 0.62%. Overall, we again observe that wrong scaling worsens the accuracy of our VaR forecast and incorporating the respective scaling behavior fixes the problem.

7 Conclusion

We analyze the effect of the two stylized facts, heavy tails and nonlinear dynamics, on financial stability. We find that financial markets across various asset classes are clearly more destabilized from nonlinear dynamics than from heavy-tailed distributions per se. We also observe that the effect gets more pronounced with an increasing degree of portfolio diversification. However, the relative reduction of drawdowns between a return series without heavy tails and a series without nonlinear dynamics appears to be rather constant. In a simple example, we show not only how the neglection of nonlinear dynamics can cause errors in risk models but also how the integration of scaling dynamics can enhance the performance of such models.

Overall, we conclude that nonlinear dynamics in asset returns are clearly more destabilizing for markets than heavy-tailed distributions. Our findings call for a rethinking of risk and have strong implications for financial regulation within which nonlinear dynamics of asset returns are mostly neglected. To enhance the stability of financial markets regulators should either include measures of nonlinearity in capital requirement regulations or take other measures to reduce nonlinear dependencies in asset returns in the first place. Though this paper focuses largely on financial regulation its implications reach beyond regulators and also touch on the areas of risk management, investment management, and financial theory. For example, the finding that nonlinear dynamics appear to cause more harm to investors raises the question of how to properly hedge against this risk or whether a risk premium can be earned. From a more theoretical perspective, it has to be questioned whether the vast literature of linear asset pricing models is worthwhile if the nonlinear dynamics in asset returns appear to be a large driver of downside risk. We leave the answers to these questions to future analysis.

Appendices

A Company Names

The company names according to Datastream of the constituents in our investment pool for the equity portfolios are:

ABBOTT LABORATORIES, ADVANCED MICRO DEVICES, AFLAC, AIR PRDS CHEMS, ALTRIA GROUP, AMERICAN EXPRESS, AMERICAN INTL GP, ANALOG DEVICES, AON CLASS A, APA, APPLE, APPLIED MATS, AUTOMATIC DATA PROC, AVERY DENNISON, BALL, BANK OF AMERICA, BANK OF NEW YORK MELLON, BATH AND BODY WORKS, BECTON DICKINSON, BOEING, BRISTOL MYERS SQUIBB, BROWN FORMAN B, CAMPBELL SOUP, CATERPILLAR, CHEVRON, CHURCH DWIGHT CO, CINCINNATI FINL, CLOROX, COCA COLA, COLGATE PALM, COMCAST A, COMERICA, CONAGRA BRANDS, CONSOLIDATED EDISON, CORNING, CSX, CUMMINS, CVS HEALTH, DANAHER, DEERE, DOVER, DXC TECHNOLOGY, EATON, EDISON INTL, ELI LILLY, EMERSON ELECTRIC, ENTERGY, EQUIFAX, EXELON, EXXON MOBIL, FEDEX, FIFTH THIRD BANCORP, FIRSTENERGY, FMC, FORD MOTOR, FRANKLIN RESOURCES, GENERAL ELECTRIC, GENERAL MILLS, GLOBE LIFE, HALLIBURTON, HASBRO, HERSHEY, HESS, HOME DEPOT, HOWMET AEROSPACE, HP, HUMANA, ILLINOIS TOOL WORKS, INTEL, INTERNATIONAL BUS MCHS, INTERNATIONAL PAPER, INTERPUBLIC GROUP, INTL FLAVORS FRAG, JACOBS SOLUTIONS, JOHNSON JOHNSON, JP MORGAN CHASE CO, KELLOGG, KEYCORP, KIMBERLY CLARK, KLA, L3HARRIS TECHNOLOGIES, LENNAR A, LINCOLN NATIONAL, LOEWS, LOWE S COMPANIES, LUMEN TECHNOLOGIES, M T BANK, MARSH MCLENNAN, MASCO, MCCORMICK COMPANY NV, MCDONALDS, MEDTRONIC, MERCK COMPANY, MOTOROLA SOLUTIONS, NEWELL BRANDS XSC, NEWMONT, NIKE B, NORTHROP GRUMMAN, NUCOR, OMNICOM GROUP, PACCAR, PARKER HANNIFIN, PEPSICO, PFIZER, PNC FINL SVS GP, PPG INDUSTRIES, PROCTER GAMBLE, PROGRESSIVE OHIO, PULTEGROUP, RAYTHEON TECHNOLOGIES, REGIONS FINL NEW, ROCKWELL AUTOMATION, S P GLOBAL, SCHLUMBERGER, SEALED AIR, SHERWIN WILLIAMS, SOUTHWEST AIRLINES, STANLEY BLACK DECKER, STATE STREET, STRYKER, SYSCO, TARGET, TELEFLEX, TERADYNE XSC, TEXAS INSTRUMENTS, TEXTRON, THERMO FISHER SCIENTIFIC, TJX, TRANE TECHNOLOGIES, TRAVELERS COS, TYSON FOODS A, UNION PACIFIC, V F, VIATRIS, W R BERKLEY, WALGREENS BOOTS ALLIANCE, WALMART, WALT DISNEY, WELLS FARGO CO, WEYERHAEUSER, WHIRLPOOL, WILLIAMS, WW GRAINGER, X3M

B MF-DFA Model Specification

The MF-DFA algorithm is one of the most widely employed workflow to analyze nonlinear dynamics of time series. It slices the time series into sub series of equal length. This process is repeated for varying lengths, which are also referred to as the model's scales, s. For each iteration, the trend within each sub series is removed and fluctuations are measured using the error function $F(q, s)$ defined as

$$F(q,s) = \left\{ \frac{1}{2N} \sum_{k=1}^{2N} [\hat{F}^2(s,k)]^{q/2} \right\}^{1/q} \tag{B.1}$$

for $q \neq 0$, and

$$F_0(s) = exp\left\{ \frac{1}{4N} \sum_{k=1}^{2N} ln[\hat{F}^2(s,k)] \right\} \tag{B.2}$$

for $q = 0$, where $\hat{F}^2(s,k)$ is the variance of the signal $x_i (i = 1, \ldots, N_s)$ around its local trend, defined as

$$\hat{F}^2(s,k) = \frac{1}{s} \sum_{k=1}^{s} \left\{ x_{(k-1)s+j} - P_{k,j} \right\}^2 \tag{B.3}$$

where N is the finite length of the time series and $P_{k,j}$ is the polynomial trend subtracted for jth data in kth sub series ($k = 1, ..., N$). The error function $F(q, s)$ can be interpreted as an extended version of a root-mean-squared error function whose exponents change in q. The overall goal of the analysis is to determine whether the fluctuations around the local trend depend on the time scale s of the sub series for different values of q. The strength of nonlinear dynamics can be quantified using the power law

$$F(q,s) \sim s^{h(q)} \tag{B.4}$$

The strength of nonlinear dynamics is then measured as the maximum Δh of the generalized Hurst exponent profile defined as

$$\Delta h = \max_q h(q) - \min_q h(q) \tag{B.5}$$

This means that if $h(q)$ varies with the change of q, the time series is exposed to nonlinear dynamics. For series that only exhibit linear autocorrelation $h(q)$ is a constant and thus $\Delta h = 0$. Another measure of nonlinear dynamics is the maximum Hölder spread $\Delta\alpha$ which measures the width of the so called Hölder or singularity spectrum $f(\alpha)$. This spectrum can be retrieved with the help of a Legendre transform (Feder, 1988; Peitgen et al., 2004) of the spectrum of generalized Hurst exponents as

$$\alpha = h(q) + qh'(q), \qquad f(\alpha) = q[\alpha - h(q)] + 1 \tag{B.6}$$

The Legendre transform is a way to display the spectrum of generalized Hurst exponents differently. It does not have an influence on the results of the analysis. The width of the singularity spectrum, $\Delta\alpha$ is the most widely adopted measure for the strength of nonlinear dynamics (Jiang et al., 2019). We also use this indicator to measure nonlinear dynamics to make our work comparable to that of others.

In the model setup, determining an optimal detrending polynomial $P_{k,j}$, scales s, and range for q is crucial for retrieving meaningful results from the analysis. However, finding optimal parameters is still a widely debated task in literature (Grech & Pamuła, 2013; Kantelhardt et al., 2002; Oswiecimka et al., 2013; Pamuła & Grech, 2014; Shao et al., 2012).

To tackle this task we follow the recommendations made by Ihlen (2012). To avoid overfitting we choose a to let $P_{k,j}$ be of first order. Further, we use 20 loq-equally spaced scales starting at $s_{min} = 32$ to $s_{max} = N/10$, where N denotes the series' finite length. Log-equal spacing is recommended because of the power law relation between $F(q, s)$ and s (Ihlen, 2012). We set $s_{min} = 32$ because literature finds that spurious nonlinear dynamics are introduced for financial time series with less than 30 observations (Buonocore et al., 2016). This finding is further bolstered by literature focusing on measurement of nonlinear dynamics. One spurious source are short samples where fluctuations not related to nonlinear dynamics may dominate over fluctuations that originate from nonlinear dynamics due to small statistics of short data (Grech & Pamuła, 2012; Grech & Pamuła, 2013; López & Contreras, 2013; Mukli et al., 2015; Pamuła & Grech, 2014; Rak & Grech, 2018). We set $s_{max} = N/10$ because in the fitting procedure the number of subseries would be very small, making estimates statistically unreliable. These restrictions are more conservative than those by Kantelhardt et al. (2002) who proposed the MF-DFA method. They put an upper limit s_{max} to their scales at $N/4$ and a lower limit of $s_{min} > 10$. It is worth noting however, that the choice of the scale does not appear to affect results of a DFA analysis as much as it does for other analysis methods of nonlinear dynamics (Shao et al., 2012). We set our moment order $q \in [-5, 5]$. This decision is guided by the fact that the potential range of scaling exponents becomes broader as the corresponding PDF deviates further from a Gaussian distribution. Thus, in choosing the range for our q we relate exponents of power law tails in financial time series that are found to range between two and five (e.g., Ghashghaie et al., 1996; Gopikrishnan et al., 1998; Longin, 1996; Pagan, 1996; Rak & Grech, 2018). We chose to limit $|q|$ within that ballpark to avoid the introduction of spurious nonlinear dynamics which is known to accelerate for larger ranges of q and is especially relevant for distributions exhibiting heavy tails (Ihlen, 2012; Pamuła & Grech, 2014; Rak & Grech, 2018). The decision to employ an equal width of the range of q around 0 is to keep estimates unbiased. This is recommended by Ihlen (2012) to make the analysis balanced because estimations in the MF-DFA procedure will be dominated by segments with small (large) fluctuations if $q < 0$ ($q > 0$) due to the way the error function Equation B.1 is set up (Kantelhardt et al., 2002).

References

Aloui, C., Shahzad, S. J. H., & Jammazi, R. (2018). Dynamic efficiency of European credit sectors: A rolling-window multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *506*, 337–348.

Alvarez-Ramirez, J., Alvarez, J., Rodriguez, E., & Fernandez-Anaya, G. (2008). Time-varying Hurst exponent for US stock markets. *Physica A: Statistical Mechanics and its Applications*, *387*(24), 6159–6169.

Al-Yahyaee, K. H., Mensi, W., & Yoon, S. M. (2018). Efficiency, multifractality, and the long-memory property of the Bitcoin market: A comparative analysis with stock, currency, and gold markets. *Finance Research Letters*, *27*(March), 228–234.

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*(1), 135–167.

Anderson, N., & Noss, J. (2013). The Fractal Market Hypothesis and its implications for the stability of financial markets. *Bank of England Financial Stability Paper*, (23).

Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. *The Review of Financial Studies*, *19*(4), 1191–1239.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, *91*(1), 1–23.

Bachelier, L. (1900). Théorie de la spéculation Annales. *Annales scientifiques de l'É.N.S.*, *17*, 21–86.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*(1), 3–30.

Baker, N. L., & Haugen, R. A. (2012). Low Risk Stocks Outperform within All Observable Markets of the World. *SSRN Electronic Journal*.

Bali, T. G., & Murray, S. (2013). Does risk-neutral skewness predict the cross section of equity option portfolio returns? *Journal of Financial and Quantitative Analysis*, *48*(4), 1145–1171.

Barberis, N., & Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, *98*(5), 2066–2100.

Barunik, J., Aste, T., Di Matteo, T., & Liu, R. (2012). Understanding the source of multifractality in financial markets. *Physica A: Statistical Mechanics and its Applications*, *391*(17), 4234–4251.

BCBS. (2019). Basel III - Explanatory note on the minimum capital requirements for market risk, (January), 1–21.

Benbachir, S., & Alaoui, M. E. (2011). A Multifractal Detrended Fluctuation Analysis of the Moroccan Stock Exchange. *International Research Journal of Finance and Economics*, *78*(6).

Blackledge, J., Kearney, D., Lamphiere, M., Rani, R., & Walsh, P. (2019). Econophysics and fractional calculus: Einstein's evolution equation, the fractal market hypothesis, trend analysis and future price prediction. *Mathematics*, *7*(11).

Blackledge, J., & Lamphiere, M. (2022). A Review of the Fractal Market Hypothesis for Trading and Market Price Prediction. *Mathematics*, *10*(1), 1–46.

Blitz, D., Pang, J., & van Vliet, P. (2013). The volatility effect in emerging markets. *Emerging Markets Review*, *16*(1), 31–45.

Blitz, D., & van Vliet, P. (2007). The volatility effect. *The Journal of Portfolio Management*, *34*(1), 102–113.

Bogachev, M. I., Eichner, J. F., & Bunde, A. (2008). The effects of multifractality on the statistics of return intervals. *European Physical Journal: Special Topics*, *161*(1), 181–193.

Bogachev, M. I., Eichner, J. F., & Bunde, A. (2007). Effect of nonlinear correlations on the statistics of return intervals in multifractal data sets. *Physical Review Letters*, *99*(24), 1–4.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *The Review of Financial Studies*, *23*(1), 170–202.

Brunnermeier, M. K., Gollier, C., & Parker, J. A. (2007). Optimal beliefs, asset prices, and the preference for skewed returns. *American Economic Review*, *97*(2), 159–165.

Buonocore, R. J., Aste, T., & Di Matteo, T. (2016). Measuring multiscaling in financial time-series. *Chaos, Solitons and Fractals*, *88*, 38–47.

Cajueiro, D. O., & Tabak, B. M. (2004). The Hurst exponent over time: Testing the assertion that emerging markets are becoming more efficient. *Physica A: Statistical Mechanics and its Applications*, *336*, 521–537.

Calvet, L., & Fisher, A. (2002). Multifractality in Asset Returns: Theory and Evidence. *The Review of Economics and Statistics*, *84*(3), 381–406.

Cao, J., & Han, B. (2013). Cross section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics*, *108*(1), 231–249.

Caraiani, P. (2012). Evidence of multifractality from emerging European stock markets. *PLoS ONE*, *7*(7), 1–9.

Chang, B. Y., Christoffersen, P., & Jacobs, K. (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, *107*(1), 46–68.

Chen, H., & Wu, C. (2011). Forecasting volatility in Shanghai and Shenzhen markets based on multifractal analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(16), 2926–2935.

Chen, W., Wei, Y., Lang, Q., Lin, Y., & Liu, M. (2014). Financial market volatility and contagion effect: A copula-multifractal volatility approach. *Physica A: Statistical Mechanics and its Applications*, *398*, 289–300.

Chung, K. H., Wang, J., & Wu, C. (2019). Volatility and the cross-section of corporate bond returns. *Journal of Financial Economics*, *133*(2), 397–417.

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex Ante Skewness and Expected Stock Returns. *The Journal of Finance*, *68*(1), 85–124.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*(2), 223–236.

Dar, A. B., Bhanja, N., & Tiwari, A. K. (2017). Do global financial crises validate assertions of fractal market hypothesis? *International Economics and Economic Policy*, *14*(1), 153–165.

Detzel, A., Duarte, J., Kamara, A., Siegel, S., & Sun, C. (2019). The Cross-Section of Volatility and Expected Returns: Then Then and Now. *SSRN Electronic Journal*.

Di Matteo, T. (2007). Multi-scaling in finance. *Quantitative Finance*, *7*(1), 21–36.

Di Matteo, T., Aste, T., & Dacorogna, M. M. (2003). Scaling behaviors in differently developed markets. *Physica A: Statistical Mechanics and its Applications*, *324*(1-2), 183–188.

Di Matteo, T., Aste, T., & Dacorogna, M. M. (2005). Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *Journal of Banking and Finance*, *29*(4), 827–851.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*(1), 83–106.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance*, *57*(1), 369–403.

Drozdz, S., Kwapień, J., Oświęcimka, P., & Rak, R. (2010). The foreign exchange market: Return distributions, multifractality, anomalous multifractality and the Epps effect. *New Journal of Physics*, *12*.

Du, G., & Ning, X. (2008). Multifractal properties of Chinese stock market in Shanghai. *Physica A: Statistical Mechanics and its Applications*, *387*(1), 261–269.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedacity with Estimates of variance of United Kingdom Inflation.

Engle, R. F., & Bolerslev, T. (1986). Modeling the Persistence of Conditional Variances. *Econometric Reviews*, *5*(1), 1–50.

Evans, J. L., & Archer, S. H. (1968). Diversification and the Reduction of Dispersion: An Empirical Analysis. *The Journal of Finance*, *23*(5), 761–767.

Fama, E. F. (1976). *Foundations of Finance*. Basic Books.

Feder, J. (1988). *Fractals*. Plenum Press.

Frezza, M. (2014). Goodness of fit assessment for a fractal model of stock markets.

Ghashghaie, S., Breymann, W., Peinke, J., Talkner, P., & Dodge, Y. (1996). Turbulent cascades in foreign exchange markets. *Nature*, *381*(6585), 767–770.

Ghazani, M. M., & Khosravi, R. (2020). Multifractal detrended cross-correlation analysis on benchmark cryptocurrencies and crude oil prices. *Physica A: Statistical Mechanics and its Applications*, *560*, 125172.

Gopikrishnan, P., Meyer, M., Amaral, L. A., & Stanley, H. E. (1998). Inverse cubic law for the distribution of stock price variations. *European Physical Journal B*, *3*(2), 139–140.

Grech, D., & Pamuła, G. (2012). Multifractal background noise of monofractal signals. *Acta Physica Polonica A*, *121*(2 B), 34–39.

Grech, D. (2016). Alternative measure of multifractal content and its application in finance. *Chaos, Solitons and Fractals*, *88*, 183–195.

Grech, D., & Pamuła, G. (2013). On the multifractal effects generated by monofractal signals. *Physica A: Statistical Mechanics and its Applications*, *392*(23), 5845–5864.

Green, E., Hanan, W., & Heffernan, D. (2014). The origins of multifractality in financial time series and the effect of extreme events. *European Physical Journal B*, *87*(6), 1–9.

Green, T. C., & Hwang, B. H. (2012). Initial public offerings as lotteries: Skewness preference and first-day returns. *Management Science*, *58*(2), 432–444.

Gu, D., & Huang, J. (2019). Multifractal detrended fluctuation analysis on high-frequency SZSE in Chinese stock market. *Physica A: Statistical Mechanics and its Applications*, *521*, 225–235.

Günay, S. (2016). Performance of the multifractal model of asset returns (MMAR): Evidence from emerging stock markets. *International Journal of Financial Studies*, *4*(2).

Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, *55*(3), 1263–1295.

He, S., & Wang, Y. (2017). Revisiting the multifractality in stock returns and its modeling implications. *Physica A: Statistical Mechanics and its Applications*, *467*, 11–20.

Hsu, D. A., Miller, R. B., & Wichern, D. W. (1974). On the stable paretian behavior of stock-market prices. *Journal of the American Statistical Association*, *69*(345), 108–113.

Hurst, H. E. (1956). The Problem of Long-Term Storage in Reservoirs. *Hydrological Sciences Journal*, *1*(3), 13–27.

Ihlen, E. A. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *Frontiers in Physiology*, *3*(141).

Ihlen, E. A., & Vereijken, B. (2013). Identifying multiplicative interactions between temporal scales of human movement variability. *Annals of Biomedical Engineering*, *41*(8), 1635–1645.

Ihlen, E. A., & Vereijken, B. (2014). Detection of co-regulation of local structure and magnitude of stride time variability using a new local detrended fluctuation analysis. *Gait and Posture*, *39*(1), 466–471.

Jiang, Z. Q., Xie, W. J., & Zhou, W. X. (2014). Testing the weak-form efficiency of the WTI crude oil futures market. *Physica A: Statistical Mechanics and its Applications*, *405*, 235–244.

Jiang, Z. Q., Xie, W. J., Zhou, W. X., & Sornette, D. (2019). Multifractal analysis of financial markets: A review. *Reports on Progress in Physics*, *82*(12).

Jin, C., Conlon, T., & Cotter, J. (2022). Co-Skewness across Return Horizons. *Journal of Financial Econometrics*, *22-14*.

Joshipura, N., & Joshipura, M. (2016). The Volatility Effect: Evidence from Indian Markets. *Applied Finance Letters*, *5*(1), 12–17.

Joshipura, N., & Joshipura, M. (2019). The Volatility Effect: Recent Evidence from Indian Markets. *Theoretical Economics Letters*, *09*(06), 2152–2164.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, *316*, 87–114.

Keylock, C. J. (2019). Hypothesis Testing for Nonlinear Phenomena in the Geosciences Using Synthetic, Surrogate Data. *Earth and Space Science*, *6*(1), 41–58.

Keylock, C. J. (2017). A multifractal surrogate data generation algorithm that preserves pointwise Holder regularity structure , with initial applications to turbulence. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *95*(032123).

Kraus, A., & Litzenberger, R. H. . (1976). Skewness Preference and the Valuation of Risk Assets. *The Journal of Finance*, *31*(4), 1085–1100.

Kraus, A., & Litzenberger, R. H. . (1983). On the Distributional Conditions for a Consumption-oriented Three Moment CAPM. *The Journal of Finance*, *38*(5), 1381–1391.

Kristoufek, L. (2013). Fractal markets hypothesis and the global financial crisis: Wavelet power evidence. *Scientific Reports*, *3*.

Kumar, S., & Deo, N. (2009). Multifractal properties of the Indian financial market. *Physica A: Statistical Mechanics and its Applications*, *388*(8), 1593–1602.

Kwapień, J., Oświęcimka, P., & Drozdz, S. (2005). Components of multifractality in high-frequency stock returns. *Physica A: Statistical Mechanics and its Applications*, *350*(2-4), 466–474.

Lamphiere, M., Blackledge, J., & Kearney, D. (2021). Carbon futures trading and short-term price prediction: An analysis using the fractal market hypothesis and evolutionary computing. *Mathematics*, *9*(9).

Li, D., Nishimura, Y., & Men, M. (2014). Fractal markets : Liquidity and investors on different time horizons. *Physica A: Statistical Mechanics and its Applications*, *407*, 144–151.

Li, D., Nishimura, Y., & Men, M. (2016). Why the long-term auto-correlation has not been eliminated by arbitragers: Evidences from NYMEX. *Energy Economics*, *59*, 167–178.

Lo, A. W. (1991). Long-Term Memory in Stock Market Prices. *Econometrica*, *59*(5), 1279–1313.

Longin, F. M. (1996). The Asymptotic Distribution of Extreme Stock Market Returns. *The Journal of Business*, *69*(3), 383–408.

López, J. L., & Contreras, J. G. (2013). Performance of multifractal detrended fluctuation analysis on short time series. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *87*(2), 1–8.

Lux, T. (2001). Turbulence in financial markets: The surprising explanatory power of simple cascade models. *Quantitative Finance*, *1*(6), 632–640.

Lux, T., & Alfarano, S. (2016). Financial power laws: Empirical evidence, models, and mechanisms. *Chaos, Solitons and Fractals*, *88*, 3–18.

Lv, X., & Shan, X. (2013). Modeling natural gas market volatility using GARCH with different distributions. *Physica A*, *392*(22), 5685–5699.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, *36*(4), 394.

Mao, J. C. T. (1970). Essentials of Portfolio Diversification Strategy. *The Journal of Finance*, *25*(5), 1109–1121.

Matia, K., Ashkenazy, Y., & Stanley, H. E. (2003). Multifractal properties of price fluctuations of stocks and commodities. *Europhysics Letters*, *61*(3), 422–428.

Mitton, T., & Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies*, *20*(4), 1255–1288.

Mukli, P., Nagy, Z., & Eke, A. (2015). Multifractal formalism by enforcing the universal behavior of scaling functions. *Physica A: Statistical Mechanics and its Applications*, *417*, 150–167.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *59*(2), 347–370.

Norouzzadeh, P., & Rahmani, B. (2006). A multifractal detrended fluctuation description of Iranian rial-US dollar exchange rate. *Physica A: Statistical Mechanics and its Applications*, *367*, 328–336.

Officer, R. R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association*, *67*(340), 807–812.

Osborne, M. F. M. (1959). Brownian Motion in the Stock Market. *Operations Research1*, *7*(2), 145–173.

Oswiecimka, P., Drod, S., Kwapie, J., & Górski, A. Z. (2013). Effect of detrending on multifractal characteristics. *Acta Physica Polonica A*, *123*(3), 597–603.

Pagan, A. (1996). The econometrics of financial markets. *Journal of Empirical Finance*, *48*(1-3), 15–102.

Paluš, M. (2008). Bootstrapping multifractals: Surrogate data from random cascades on wavelet dyadic trees. *Physical Review Letters*, *101*(13), 1–4.

Pamuła, G., & Grech, D. (2014). Influence of the maximal fluctuation moment order q on multifractal records normalized by finite-size effects. *EPL*, *105*(50004).

Peitgen, H.-O., Jürgens, H., & Saupe, D. (2004). *Chaos and Fractals: New Frontiers of Science*. Springer.

Peters, E. E. (1991). *Chaos and Order in the Capital Markets: A New View of Cycles, Prices, and Market Volatility*. John Wiley & Sons.

Peters, E. E. (1994). *Fractal market analysis: Applying chaos theory to investment and economics*. John Wiley & Sons.

Rachev, S. T., Weron, A., & Weron, R. (1999). CED model for asset returns and fractal market hypothesis. *Mathematical and Computer Modelling*, *29*(10-12), 23–36.

Rak, R., & Grech, D. (2018). Quantitative approach to multifractality induced by correlations and broad distribution of data. *Physica A: Statistical Mechanics and its Applications*, *508*, 48–66.

Rubinstein, M. (1973). The Fundamental Theorem of Parameter-Preference Security Valuation. *Journal of Financial and Quantitative Analysis*, *8*(1), 61–69.

Schadner, W. (2022a). Expected Return Auto-Correlation: Believes, Efficiency and Meltdowns. *SSRN Electronic Journal*.

Schadner, W. (2022b). U.S. Politics from a multifractal perspective. *Chaos, Solitons and Fractals*, *155*, 111677.

Schmitt, F., Schertzer, D., & Lovejoy, S. (2000). Multifractal Fluctuations in Finance. *International Journal of Theoretical and Applied Finance*, *3*(3), 361–364.

Schmitt, F., Schertzer, D., & Lovejoy, S. (1999). Multifractal analysis of foreign exchange data. *Applied Stochastic Models and Data Analysis*, *15*(1), 29–53.

Schneider, P. G., Wagner, C., & Zechner, J. (2020). Low-Risk Anomalies? *The Journal of Finance*, *75*(5), 2673–2718.

Schreiber, T., & Schmitz, A. (1996). Improved surrogate data for nonlinearity tests. *Physical Review Letters*, *77*(4), 635–638.

Schreiber, T., & Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, *142*(3-4), 346–382.

Segnon, M., & Trede, M. (2018). Forecasting market risk of portfolios: copula-Markov switching multifractal approach. *European Journal of Finance*, *24*(14), 1123–1143.

Shahzad, S. J. H., Nor, S. M., Mensi, W., & Kumar, R. R. (2017). Examining the efficiency and interdependence of US credit and stock markets through MF-DFA and MF-DXA approaches. *Physica A: Statistical Mechanics and its Applications*, *471*, 351–363.

Shao, Y. H. (2020). Does Crude Oil Market Efficiency Improve After the Lift of the U.S. Export Ban? Evidence From Time-Varying Hurst Exponent. *Frontiers in Physics*, *8*(October), 1–8.

Shao, Y. H., Gu, G. F., Jiang, Z. Q., Zhou, W. X., & Sornette, D. (2012). Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series. *Scientific Reports*, *2*, 835.

Statman, M. (1987). How Many Stocks Make a Diversified Portfolio? *The Journal of Financial and Quantitative Analysis*, *22*(3), 353–363.

Stosic, D., Stosic, D., Ludermir, T. B., & Stosic, T. (2019). Multifractal behavior of price and volume changes in the cryptocurrency market. *Physica A: Statistical Mechanics and its Applications*, *520*, 54–61.

Suárez-García, P., & Gómez-Ullate, D. (2014). Multifractality and long memory of a financial index. *Physica A: Statistical Mechanics and its Applications*, *394*, 226–234.

Turiel, A., & Pérez-Vicente, C. J. (2003). Multifractal geometry in stock market time series. *Physica A: Statistical Mechanics and its Applications*, *322*, 629–649.

Turiel, A., & Pérez-Vicente, C. J. (2005). Role of multifractal sources in the analysis of stock market time series. *Physica A: Statistical Mechanics and its Applications*, *355*(2-4), 475–496.

Upson, R. B., Jessup, P. F., & Matsumoto, K. (1975). Portfolio Diversification Strategies. *Financial Analysts Journal*, *31*(3), 86–88.

Vaseghi, S. V. (2008). Power Spectrum Analysis. In S. V. Vaseghi (Ed.), *Advanced digital signal processing and noise reduction* (4th ed., pp. 271–294). John Wiley & Sons.

Venema, V., Ament, F., & Simmer, C. (2006). Nonlinear Processes in Geophysics A Stochastic Iterative Amplitude Adjusted Fourier Transform algorithm with improved accuracy. *Nonlinear Processes in Geophysics*, *13*, 321–328.

Wang, H. Y., & Wang, T. T. (2018). Multifractal analysis of the Chinese stock, bond and fund markets. *Physica A: Statistical Mechanics and its Applications*, *512*, 280–292.

Wang, Y., Wu, C., & Pan, Z. (2011). Multifractal detrending moving average analysis on the US Dollar exchange rates. *Physica A: Statistical Mechanics and its Applications*, *390*(20), 3512–3523.

Wei, Y., & Wang, P. (2008). Forecasting volatility of SSEC in Chinese stock market using multifractal analysis. *Physica A: Statistical Mechanics and its Applications*, *387*(7), 1585–1592.

Weron, A., & Weron, R. (2000). Fractal market hypothesis and two power-laws. *Chaos, Solitons and Fractals*, *11*, 289–296.

Zhou, W. X. (2009). The components of empirical multifractality in financial returns. *Europhysics Letters*, *88*(28004).

Zhou, W. X. (2012). Finite-size effect and the components of multifractality in financial volatility. *Chaos, Solitons and Fractals*, *45*(2), 147–155.