ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis

INTERNATIONAL BACHELOR ECONOMICS AND BUSINESS ECONOMICS

Unraveling the Nexus between the Dollar Exchange Rate and Sectoral Risk: a VaR-based Comparative Study of USD Risk Spillovers

AUTHOR

Samir SABITLI

557678ss

SUPERVISOR

Ruben DE BLIEK, PH.D.

SECOND ASSESSOR

Haikun ZHU, PH.D.

Abstract

This paper investigates the spillover of risk from the USD exchange rate to 11 different sectors of the U.S. economy using the Value-at-Risk (VaR) framework. We build 1% and 5% VaR models under the variance-covariance method by constructing various GARCH-type models incorporating the skewed Generalised Error Distribution. VaR-based risk spillovers are subsequently identified by means of the Granger Causality in Risk test by Hong et al. (2009). Initially, we find that the use of SGED error specifications improves forecasted VaR performance over Student's t. Results from the Granger tests indicate that generally, there does not exist a significant consistent spillover from USD Broad to various sectors. However, a deeper dive reveals that the most significant spillovers decay rapidly, supporting the notion of low cross-market dependence. While the overarching debate may remain for the foreseeable future, this paper has demonstrated the merit that VCM-VaR-SGED-Granger-Causality methodologies present over other approaches.

Keywords: Risk Spillover, Granger Causality Test in Risk, USD Exchange Rate

JEL Codes: C58, C01

Date Final Version: July 12, 2023

Acknowledgement

Special thanks to Gang-Jin Wang, Ph.D., Professor at Hunan University, for providing the necessary Matlab code to compute the one and two-way kernel-based Granger Causality in Risk Test Statistic and p-values.

Contents

Symbols

VaR(α)	Value-at-Risk at a specified confidence level
α	Confidence Level
ε_t	Time-varying Error/Shock/Innovation term
z_α	Quantile of specified distribution at α confidence level
σ_t	Conditional time-varying Volatility
Q_i	Granger Causality in Risk test statistic where $i = 1$ and $i = 2$ refer to one-way and two-way directionality, respectively
χ^2_{VP}	Goodness of Fit test statistic based on Vlaar and Palm (1993)
PQ_A	Automatic Portmanteau test for serial correlation based on Escanciano and Lobato (2009)
$\mathrm{LR}_{\mathrm{CC}}$	Conditional coverage Likelihood-Ratio test statistic
$\mathrm{LR}_{\mathrm{UC}}$	Unconditional coverage Likelihood-Ratio test statistic
$\hat{P}$	GARCH Persistence Parameter
$\cdot_{m,t}$	Returns Series (or equivalently, Market) m at time t attached to specified Variable
$\mathbb{E}[r_t \vert I_{t-1}]$	Mean of Returns series Conditional on Information Set I_{t-1}
$\mathbb{E}\left[r_{m,t} < \mathrm{VaR}_{m,t}\right]$	VaR Coverage

1 Introduction

Following the collapse of the Bretton Woods exchange rate regime in the 1970s, the world largely began relaxing foreign capital controls and gradually transitioning into a system of floating exchange rates. These major developments in the global financial order combined with a period of rapidly-growing entanglement of markets contributed to significant market volatility at the time, as exemplified by the Sterling Crisis of 1976 and the 1985 Plaza Accord. These periods of extreme market downturn underscored the necessity to improve our knowledge of the complex dynamics behind the interaction of exchange rates and market forces (Jacque, 1981), and to build the tools necessary to combat risk. Recognising this, in 1995 the financial world formalised the Basel Regulations, which established the metric of 'Value at Risk' (VaR) as a standard method of assessing general risk for high-risk financial institutions (van den Goorbergh & Vlaar, 1999). Since then, a rich field of academia spanning multiple disciplines has embraced VaR as a powerful tool of risk management.

Classical economics can serve as the basis of our knowledge when discussing the potential relationship between markets and exchange rates. Taking common equity markets as a proxy for market performance, the 'flow-oriented' model asserts that a positive relationship exists between stock prices and exchange rates, whereby an increase in rates will improve the current account, making domestic firms' exports more competitive, thereby improving firms' bottom line (Dornbusch & Fischer, 1980a). This is reflected in a higher stock price. However, per Branson (1983) and Frankel (1976), the 'stock-oriented' model considers causality to be in reverse, arguing that exchange rates are influenced by the supply and demand of financial assets – the increase in stock price will generally lead to greater demand for domestic assets, resulting in investors selling off foreign assets to acquire more domestic currency to pursue more domestic investments. Through the dynamics of local money supply and interest rates, this would appreciate the domestic currency. However, other well-known theories the likes of Frenkel's (1976) asset market model disagree with both of the aforementioned, arguing that there cannot exist any relationship between the two variables.

Empirically, we observe a multitude of different methodologies, samples, and ultimately, conclusions. Owing to the prevailing econometric practices of the time, the earliest studies on the topic tended to favour simpler methodologies. The research from Aggarwal (1981) is characterised for its usage of basic multivariate regressions to arrive at the finding of a significantly positive association between exchange rates and stock prices. Using a similar methodology and dataset, Giovannini and Jorion (1987) reach a similar results. However, many other prominent comparable studies, such as that of Soenen and Hennigar (1988) and Bahmani-Oskooee and Sohrabian (1992) reach the exact opposite conclusion, arguing for a negative relationship. Many of these earlier studies, however, suffered from econometric inconsistencies as publicised by Bahmani-Oskooe (Bahmani-Oskooee & Saha, 2015). Through improvements to their econometric methodologies, later studies such as those of Tsagkanos & Siriopoulos (2013) and Ratanapakorn & Sharma (2007), were able to shed more light on the relationship, now decomposing it into long and short run effects. However, despite improvements to methodologies and a growing body of literature to build upon, according to the extensive literature review of Xie et al. (2020), the conclusions reached on the relationship at hand remain discordant.

Alongside a lack of consensus on the direction and sign of the causality from exchange rates to equity markets, there appear to be numerous other sources of uncertainty. Firstly, much of existing literature opts to analyse volatility or price spillovers, arguably as a psuedo-proxy for risk. However, recent developments in our understanding of volatility and risk, thanks in part due to the research of Engle (2004), have shown them to be separate and distinct concepts, and one too often erroneously considered equal. For this reason, we find it scientifically valuable to extend existing literature by pursuing specifically the spillover of risk, defined in a form most universally accepted for the purposes of regulatory risk management. Furthermore, it appears that it is predominantly indices of aggregated markets that are most investigated. However, Oral and Akkaya (2015) suggest that VaR, if modelled correctly, can be an effective tool in combating risk across many non-financial industries, as exemplified by its requirement by a diverse range of institutions. Considering the practical necessity of a well-specified VaR model by a range of stakeholders across various sectors, and the possibility of a systemic exchange rate spillover effect in US markets, we aim to further knowledge in this field by studying the spillover of exchange rate risk across several sectors of the U.S. economy. This, ultimately, leads us to the core research question: **How does the volatility and risk of the US Dollar exchange rate affect various US market sectors and how can it be modelled effectively?**

In constructing this investigation, we borrow from the methodology formulated by Fan et al. (2008). We first access daily returns on indices on the largest sectors of the U.S markets by The Center for Research in Security Prices (CRSP) spanning between March 31, 2011 and December 31, 2022. We then gather daily 'returns' on the US 'Broad Index' , an indicator from the Federal Reserve, to assess the value of the US. Dollar relative to the currencies of the U.S's trade partners, weighed by trade volume for the same time period. Following the standard methodology behind VaR construction and specific recommendations from Fan et al. (2008), for the purposes of forecasting VaR, we first construct several GARCH-type volatility models incorporating mean modelling. Per the well-known stylised facts of financial returns highlighted by Cont (2001) and the empirical studies in section 2.2, we attempt to construct statistically-adequate VaR methodologies by incorporating the standard GARCH, GJR-GARCH, APARCH, TGARCH, and iGARCH variance models, which explicitly take advantage of these features. Additionally, due to the well-known leptokurtotic nature of financial returns, we require a returns distribution which accommodate the idiosyncrasies of each of our returns series. As recommended by Lee and Su (2012), the skewed Generalised Error Distribution is able to conform to the specific skewness and kurtosis of our data, enabling us to more accurately forecast VaR. Following that, we perform backtests to determine whether our VaR model has correctly estimated the downside risk by the method formulated by Kupiec (1995) and Christoffersen (1998). Finally, the spillover of the VaR of USD into market sectors is tested by means of the Granger causality test for risk formulated by Hong et al. (2009). We utilise both the one-way and two-way-instantaneous tests to also assess the directionality of Granger-causality.

We structure this investigation as follows. The subsequent section attempts to shed some light on some prevailing theories on the relationship between exchange rates and equity markets and VaR modelling as it is done conventionally in literature. Further, the section concludes with a review of literature on the spillover of exchange rates into equity markets, taking into account the preceding discussions. Following the Theoretical Framework, the Data and Methodology expand upon the description of the specific dataset and procedures utilised to arrive at our conclusion on the core research question. With the general methodology clarified, we continue to the outcome of the variance and VaR modelling in accordance with the previously-established procedure. Finally, the Discussion and Conclusion interjects economic insights into the outcome attained in the preceding section and summarises the investigation with respect to the existing body of literature in this field.

2 Theoretical Framework

Our investigation serves to bridge past research on two prominent fields: the spillover of exchange rates and risk modelling. As such, we proceed by introducing each. In section 2.1, we discuss the various accepted theories on the interactions of exchange rates with financial markets, and in section 2.2, we outline the main approach to downside risk modelling in literature. Finally, with respect to the previous discussions, in section 2.3 we conduct a review of literature on the spillover of currency risk on equity markets.

2.1 Theoretical Foundation for Exchange Rate Interactions

In aiming to understand the interactions of the international USD exchange rate with various sectors of the U.S. economy, we introduce two widely-accredited, yet fundamentally opposing families of theories claiming to govern these relationships. These are the so-called "flow-oriented" and "stock-oriented" models of exchange rates, formulated by Dornbusch & Fisher (1980b) and Branson (1983) & Frankel (1992), respectively.

According to the "Flow-Oriented" model, changes in exchange rates directly affect the international competitiveness of companies, influencing their revenues, costs, and output which ultimately affect the present value of future cash flows (Moeljadi & Fauziah, 2015). The latter, as is commonly known in corporate finance, translates directly to stock prices (Sloan, 1996). More specifically, the direction and sign of causality are claimed to be from the exchange rate to stock prices and positive. For instance, currency depreciation would make a country's firms' exports more competitive on the international stage, increasing revenues and thereby, driving up the value of the firms. On the other hand, firms affected more by import levels would observe increases in production costs and lower net earnings, lowering firm value (Moeljadi & Fauziah, 2015). These effects, however, need not only apply to firms engaging in international trade; according to Adler & Dumas (1984), changes to a country's macroeconomic conditions caused by exchange rate fluctuations may directly affect firms' input-output price levels, despite no direct exposure to trade. Moreover, exchange rates have implications regarding the competitive standing of domestic firms over foreign competitors that operate under multiple exchange rates.

The "Stock-Oriented" family of models, on the other hand, take a different approach as they emphasise the role of the supply and demand of financial assets to derive the direction of causality and its sign.

Specifically, the 'Portfolio Balance Model" of Branson (1983) & Frankel (1992) claims a negative causality from stock prices to exchange rates through both direct and indirect channels. Per Xie et al. (2020), assuming that investors are risk-averse and they move investments to countries with greater stock returns, we observe currency appreciation where there is greater stock returns, and depreciation otherwise because higher returns lead to greater demand by investors for the local currency and vice versa. Hence, high-return countries are likely to experience exchange rate depreciation [1].

However, Frenkel's (1976) asset market models oppose both conclusions, taking a stance against any relationship. Broadly, the theory argues that exchange rates should be treated similarly to the price of assets, namely, that we consider its present value a function of the expected rate of return. Consequently, assuming that their expected rates of return are not affected by common factors, there should not be a relationship between the two (Kollias et al., 2012). However, where common factors may exist, for instance, with interest rates, there may be an association.

2.2 Quantification of Downside Financial Risk

Since the 1990s, Value-at-Risk, or VaR, has risen to prominence as the premiere method of quantifying downside financial risk (Hartz et al., 2006). It can be defined as the maximum expected loss associated with a financial position at a certain probability of the realised loss exceeding this amount (Odening & Hinrichs, 2002). This probability is referred to as the confidence level and is typically set at 95 or 99 percent. As outlined by Abad et al. (2013), in literature VaR is mainly computed and forecasted through the variance-covariance (VCM), historical, and Monte Carlo methods, however, VCM complements our methodology best as it enables integration with time-varying volatility models and permits the explicit selection of a customised distribution function. VaR computed under VCM can be depicted as

$$\mathrm{VaR}(\alpha) = \mu + z_\alpha \sigma, \tag{1}$$

where μ is the mean of a returns series, z_α is the p-quantile of a specified distribution's density function corresponding to the pre-specified confidence level[2] α, and σ is the conditional variance of a returns series. The expression implies that the main considerations when implementing VCM are the specification of a distribution from which to derive quantiles and a method of modelling the conditional variance of a series.

According to Yu and So (2006), the adequacy and validity of a VaR model under VCM depends strongly on the quality of the variance parameter σ. Taking advantage of well-known salient features of returns data, we are able to construct time-varying variance models. One of the most well-known features is the presence of volatility clustering, namely that large changes in volatility tend to be followed by large changes and small changes by other small changes (Jacobsen

[1] It is important to keep in mind that the exchange rate here refers to the price of one unit of foreign currency for one unit of local currency. Hence, a depreciation indeed corresponds to an appreciation of the local currency

[2] Throughout this report, the standard confidence levels of 95 and 90 percent may be used interchangeably with α levels of 0.01 and 0.05. The 95% and 99% confidence level corresponds to an α and quantile probability of 0.05 and 0.01, respectively.

& Dannenburg, 2003). This stylised fact gives rise to some level of predictability to returns, enabling variance to be split into predictable and unpredictable components. The predictable component, also known as conditional volatility, is the basis of the most known variance models, such as the family of Generalised Autoregressive Heteroskedasticity (GARCH) models first formulated by Bollerslev (1986). Taking squared returns as a proxy for volatility, the decomposition of returns into conditional variance is shown below.

$$\begin{aligned} r_t = \mathbb{E}[r_t|I_{t-1}] + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t \\ \mathbb{E}\left[r_t^2\right] = \mathbb{E}\left[\left(\mathbb{E}\left[r_t|I_{t-1}\right] + \sigma_t z_t\right)^2\right] = \sigma_t^2. \end{aligned} \tag{2}$$

Expression 2 shows that returns can be shown to be comprised of the sum of $\mathbb{E}[r_t|I_{t-1}]$, the expectation of the series conditional on the past information set I_{t-1}, and a time-varying error term ε_t. The error itself is broken down into the predictable conditional variance σ_t and a random error process z_t, which is independently and identically distributed (i.i.d) and has a zero-mean. If we assume that the conditional mean is zero, we see that the expectation of the squared series is simply conditional volatility.

To ensure that other salient features of financial returns are sufficiently exploited for an optimal variance fit, we test three different classes of GARCH-type models widely documented in literature. First, there may be sufficient evidence to consider so-called "symmetric effect" models, where shocks of ε_t do not have a different impact on conditional volatility relative to their sign. In their landmark study, Hansen and Lunde (2005) assessed 330 different ARCH-type models using daily IBM stock and the Deutsche Mark to USD exchange rate data on their ability to forecast one-day-ahead conditional variance. They determined forecasting performance out-of-sample using six different loss functions and tested for the statistical significance of those results using the Superior Predictive Ability (SPA) test of Hansen (2001) and the Reality Check for Data Snooping (RC) by White (2000). Ultimately, they arrived at three different conclusions: first, regarding exchange rate data, there was no evidence that any other model outperformed the basic GARCH(1,1) specification. However, in terms of IBM data, GARCH(1,1) is decisively beaten by other models. Interestingly, it was noted that forecasts for DM-USD improved when its models were fitted with a Student's t error distribution assumption as opposed to standard normal, but the converse was seen with IBM. Moreover, the authors observed that the models that were able to best fit IBM data were those accommodating an asymmetric response to positive and negative values of each shock ε_t, or in other words, a "leverage effect".

Extending the methodology of Hansen and Lunde (2005), Angelidis and Benos (2007) investigated the ability of several asymmetric response ARCH-type models to forecast VaR computed under the parametric (VCM), semi-parametric, and non-parametric methods. They evaluate peformance for confidence levels of 97.5% and 99% for the Dow-Jones Euro Stoxx large and small cap indices. To ensure the robustness of their results, forecasts are conducted over two time periods. Then, the authors perform two-stage diagnostics, where they first VaR adequacy is first statistically evaluated using Kupiec's (1995) backtesting methodology. Then, similar to Hansen et al. (2005), the authors utilise the SPA test to assess improvements in forecasts across models. Specifically, they introduce the Exponential GARCH (EGARCH) model of Nelson (1991) and threshold GARCH (TGARCH) model, arguing that unlike GARCH, they can better model stock returns as they accomodate leverage effects. However, Angelidis et al. (2007) assert that conditional variance need not necessarily be a linear function of lagged squared returns (Brooks & Persand, 2003) and as such, incorporate the asymmetric-power ARCH (APARCH) into their analysis. Generally, the authors found that for long positions on the large-cap indices, the VaR forecasts for both confidence intervals were adequate[3] for all ARCH-type models. In opposition to several other similar studies such as those by Guermat and Harris (2002) alongside Billion and Pelizzon (2000), VaR forecasts of large-cap indices seemed to improve when ARCH models were estimated under symmetric and skewed Student's t distributions.

Notwithstanding, there is a growing body of literature endorsing a third class of model due to its ability to take advantage of a different stylised fact of financial returns; according to Cont (2001), returns exhibit a "long memory", or in other words, the very slow decay of shocks to the conditional volatility. So and Yu (2006) show that these long-term dependencies can significantly influence the forecasting of market volatility. Berggren (2017) researched the impact of several long-memory GARCH models in the accuracy of forecasting VaR and Expected Shortfall [4] benchmarked against standard GARCH using the main Swedish market index, OMXS30. Unlike the procedure of Hansen et al. (2005), Berggren opted not to utilise Hansen's (2001) SPA test as it fails for Expected Shortfall. Rather, his procedure is similar to the first stage of Angelidis's (2007) diagnostics whereby Kupiec's (1995) backtest is used to determine VaR adequacy at 2.5% confidence level. Extending their methodology, the author also incorporates Christoffersen's (1998) test for unconditional coverage, where an additional requirement of VaR exceedance independence is imposed. His results indicated that the long-memory Integrated GARCH (IGARCH) and GJR-GARCH models proved most effective, as they passed both backtests with optimal coverages. Reinforcing the conclusion of Hansen et al. (2005), the author cites the research of Ane (2006), arguing that more complex models such as APARCH and EGARCH may not necessarily improve fit, but rather add redundancy. However, they acknowledge that some long-memory properties of APARCH may capture similar to that of IGARCH, improving convergence. Similar to Angelidis et al. (2007), Berggren observed that fitting Student's t and skewed t distributions to these models generally improved fit, suggesting that there may be gains to taking advantage of return distributions' asymmetries.

Thus far, we have discussed the selection of a variance model for VCM, however, as outlined in expression 1, it is just as crucial to select a distribution from which to derive VaR quantiles and to ensure that it corresponds to our variance models' error distribution. Typically, variance models assume that the error process z_t follows a standard normal distribution. However, many researchers recognise that this may fail to capture the well-known fact that returns often exhibit some degree of leptokurtosis (Lee & Su, 2012), meaning that their distributions are skewed with heavy tails (Cont, 2001). Being a symmetric distribution with thin tails, using a standard normal

[3]When we mention VaR adequacy, we refer to the convergence of forecasted VaR being at approximately the alpha level and the Kupiec test result returning insignificant. This is further explained in the methodology section of this paper

[4]Conceptually like VaR, Expected Shortfall is another tool utilised to compute downside risk. Broadly speaking, the main difference with VaR is its variable weighting of a range of quantiles. VaR places a 100% weight on the defined α quantile of a series. Expected Shortfall is out of the scope of this research and as such, it will not be evaluated.

distribution specification for z_α is likely to lead to systemic under or over estimation of VaR, causing our models to be inadequate, even if variance itself is modelled correctly. This is the primary motivation for why, according to Lee and Su (2012), researchers instead began opting for non-normal distributions such as the Student's t and the Generalised Error Distribution (GED), which can accommodate excess kurtosis. This is exemplified by the research of So and Yu (2006), who, researching the performance of asymmetric, symmetric, and long-memory GARCH models on a range of global market indices and exchange rates, found that almost across the board, models incorporating Student's t yielded better VaR estimates at a confidence level of 1% over the standard normal. Other researchers such as Berggren (2017) and Angelidis et al. (2007) covered previously reach similar conclusions. However, Lee and Su (2012) argue that VaR estimates can be further improved by adopting a distribution that not only deals with excess kurtosis, but also excess skewness. Hence, to overcome the issue of systemic under or estimation of VaR, we adopt the skewed Generalised Error Distribution (SGED) to derive returns quantiles under VCM. Through the the maximum likelihood technique, this distribution enables the estimation of parameters for skewness, kurtosis, and centering, enabling it to take on a wide range of different shapes to adapt to to the idiosyncrasies of our returns series. Further, as leptokurotic returns are likely to indicate asymmetric model error distributions, we apply SGED to the variance models discussed above (Siwen Zhou, 2018). Throughout the following sections, we outline a methodology that enables us to compare the performance of the SGED specification over Student's t to assess claims made in previous literature regarding its superiority in modelling VaR.

Although the preceding discussion on the choice of variance model and VaR distribution is by no means conclusive, it gives us several hints at the specifications we should consider. The research of Hansen et al. (2005) gives credence to the idea that the simple symmetric standard GARCH model may prove effective with exchange rate data such as our Broad Index. However, investigating VaR directly, Angelidis et al. (2007) and Berggren (2017) assert that a range of asymmetric models such as EGARCH, TGARCH, GJR-GARCH, and APARCH may capture leverage effects that standard GARCH cannot. Further, Berggren (2017) suggests that IGARCH can be a simple, yet effective method of capturing long-memory effects that other models may largely fail to do. We extend existing VaR methodologies through our specification of an SGED distribution as literature typically favours simpler specifications. We discuss the specific procedures to arrive at our VaR in section 4.

2.3 Past Empirical Studies

The widely-accredited exchange rate theories discussed in section 2.1 indicate that, on a theoretical level, there may or may not exist some sort of relationship between equity markets and exchange rates. Due to the sheer differences in samples, methodologies, and shifting attitudes towards econometric research, this same uncertainty regarding the sign, direction, or even the existence of causality flowing from either variable has been a distinct feature of empirical studies spanning back to the 1980s (Xie et al., 2020).

2.3.1 Early Methodologies

According to the extensive literature review of Bahmani-Oskooee and Saha (2015), the first study investigating the dynamics of stock prices and exchange rates was that of Aggarwal (1981). With monthly data on aggregate indices of stock prices and the effective USD exchange rate between 1974-1978 at his disposal, by means of a series of simple multivariate regression models, he demonstrated a positive association between the two variables. The result, per Bahmani-Oskooee et al. (2015), seems to suggest that firms generally did not benefit from the depreciation of the dollar, rather, they appear to have lost some degree of competitiveness, hurting their performance, and subsequently, their stock price. While the conclusion reached seems to directly dispute the "Flow-oriented" models of Dornbusch & Fisher (1980a), the more granular study conducted by Soenen and Hennigar (1988) reaches precisely the opposite result. This time, the authors researched the influence of the USD exchange rate on seven sectors of the U.S. economy, being automobile, computer, machinery, paper, textile, steel, and chemical (Bahmani-Oskooee & Saha, 2015). Through a similar methodology, they find that there existed a negative relationship between these sectors and the exchange rate. In other words, the depreciation of the dollar saw in response a larger volumes of exports, increasing firms' profits.

2.3.2 Risk and Volatility Spillover via Cointegrating Relationships

While these earlier studies were important in establishing the viability of research in the field and putting forth the foundations for modern methodologies, the results are arguably unreliable. Bahmani-Oskooee (2015) argue that these studies failed to take into account the integrating properties and the cointegration relationship between the variables, rendering their econometric models potentially spurious.

In an attempt to mitigate the issue, literature instead began favouring Engle and Granger's (1987) cointegration methodology and the Granger causality framework. Some research adopting the methodology on developed countries seem to demonstrate that the price spillover across equity markets and exchange rates may be positive (Xie et al., 2020). Investigating the Portfolio Balance Theory, Tsagkanos and Siriopoulos (2013) researched the existence of a long-term causal relationship from stock returns to exchange rate returns within the EU and US across two time periods: the 2008 financial crisis (2008-2012) and 2003-2007, during a more stable market. To do this, the researchers adopted a special econometric method called the Structural Nonparametric Cointegration Regression (SNCR) in order to more clearly decompose the relationship into the short and long run. To determine the price level spillover across markets, the authors construct first-differenced Vector Autoregressive (VAR) models with which they use Granger Causality tests. Largely in line with the results of Tsai (2012) and Kollias et al. (2012), during the crisis period they find that stock prices influence exchange rate fluctuations in both regions. Contrary to Tsai and Kollias et al., they find that the relationship is long-run in the EU whereas it is short-run in the US. Investigating the stable period, they conclude that the aforementioned causality is observed in the short-run for both the EU and US.

Using an augmented VAR and Granger Causality framework, Ratanapakorn and Sharma (2007) have a broader focus and investigate the existence of short and long run relationships between the S&P500 and six macroecnomic variables across 1975-1999. With the presence of a cointegrating relationship between their models, the authors use the Vector Error Correction Model (VCEM) in tandem with Granger causality to separate a potential relationship in the short and long run and detect a statistically-significant spillover. Generally, their results corroborate those of Tsagkanos and Siriopoulos (2013), as they observe a positive bidirectional long and short run relationship from the US equity market to exchange rates. However, no Granger causality is observed from the exchange rate to the stock market in the short run.

Other researchers instead adopt variance models to examine the volatility spillover, rather than price spillover, across the two markets in question. Applying the Multivariate GARCH (MGARCH) model to the Euro-Dollar exchange rate and two US market indices, Dow-Jones and S&P500, Tastan (2006) examines the presence of significant time-varying conditional covariance and correlation coefficients. Through his analysis, the author demonstrates that the nature of the correlation between the two variables is nuanced and difficult to capture through traditional models. With this methodology, the use of MGARCH models makes the determination of the directionality of volatility spillover prohibitively difficult and forbids a direct comparison with the results of research like Ratanapakorn and Sharma (2007) and Tsagkanos et al. (2013) as the spillover is no longer decomposed into the short and long run. Nevertheless, Tastan concludes that while unconditional correlation between the market indices and exchange rate is low and negative, the conditional correlation is significant, both economically and statistically. In other words, excluding the effects of short-term shocks, the average longer-term correlation can be said to be very small, whereas correlation taking into account all available fluctuations over the specified time period shows that the relationship may be more significant. The implication of these results is that each individual shock is important in the joint correlation across the variables.

3 Data

To investigate risk spillovers across the U.S. Dollar exchange rate and sectors of the U.S. economy, data on relevant indices was collected. Daily closing price levels for indices by the Center for Research in Security Prices (CRSP) on equity markets, split across 10 different sectors were obtained from Wharton Research Data Services (WRDS). The CRSP sector indices, similar to the commonly used series of S&P sector indices, provide a sufficient tradeoff between the variety of industries covered and the volume of data available. In addition to the 10 sectors shown in Table 2, we also include a total market index to assess the market as a whole. However, the sample size differs as data begins from 31/12/2011 for the total market index, 20/09/2012 for the sectors, and end on 31/12/2022. To capture movements in the U.S. Dollar exchange rate, the Trade-Weighted U.S. Dollar Index, also known as the Broad Index, was utilised; it jointly captures the exchange rates of the U.S. Dollar against the currencies of the largest U.S. trading partners, weighed by trade volume. While some studies such as that of Roberedo et al. (2016) and Tastan (2006) choose to study individual set of currencies, particularly for research pertaining to non-developed regions, using a trade-weighted index such as the Broad Index remains a popular choice among many other researchers including Aggarwal (1981) and Bahmani-Oskooee and Sohrabian (1992) who aim to investigate exchange rate effects in the context of developed markets. Arguably, one of the factors empowering a trade-weighted index over individual exchange rates in this case is the similarity of markets it encompasses. To maintain the relevance of the index, the weights are updated annually (Mico et al., 2005). However, it should be noted that the index levels are nominal rather than in real terms due to the fact that our sector indices are not inflation-adjusted; conceptually, the units of analysis must match. Daily Broad index levels from 02/01/2006 until 14/04/2023 were obtained directly from a personal API key from FRED, the economic research branch of the Federal Reserve, using the `R` package `fredr`.

Many of our upcoming analyses will incorporate time series models. These models require our time series variables to be stationary, however, it is well-known that index price levels are typically non-stationary as they often have trends and fluctuate around non-constant means. A simple method of obtaining a usable, yet stationary time series is to compute log-returns from our variables as shown below (S. J. Taylor, 2007).

$$r_t = \ln \frac{P_t}{P_{t-1}} \cdot 100,$$

where r_t is the percentage daily return, P_t is the price level, and P_{t-1} is the 1-period lagged price level at time t, respectively. To then test the stationarity of the resulting time series, an Augmented Dickey-Fuller test incorporating drift and trend + drift terms were conducted. The results, as shown in Table A1, indicate that in all cases, our data is highly stationary. In addition to this requirement, most models require that the time series is complete without any breaks in the dates. However, both the Broad and CRSP Indices are discontinuous time series, as there is no published data on non-working days. While some researchers opt for data imputation techniques to fill in the gaps (see Fan et al. (2008)), we believe that it is best to maintain the integrity of the original dataset in our case since the sheer number of new artificial data points created has the potential to induce erroneous results. Fortunately, the `R` package `xts` provides an easy solution to the issue. Using the `xts` time format, we can simply force `R` to ignore the breaks in the time series and treat it as a continuous unit. Consequently, we can proceed to input these series in our time series models, albeit with the caveat that when going back in steps of one day across the time series, it does not always correspond to a jump of 1 day between dates. For the purposes of forecasting, however, this does not present an issue.

Merging all of the returns series together, we arrive at an effective sample size of 2585. Our models require data not only for calibration and parameter estimation, but also to evaluate forecasting performance. Consequently, we reserve 80% of the sample, 1936 observations, for model estimation, and the remaining 20%, 485 observations, for forecast evaluation.

Table 2: Descriptive Statistics for Cleaned Returns Data

	First Observation	Final Observation	Obs.	Mean	Standard Deviation	Min	Max	Skewness	Kurtosis
USD Broad	03/01/2006	14/04/2023	4,509	0.011	0.313	-2.088	1.874	0.179	6.575
Total Market	01/04/2011	31/12/2022	2,958	0.037	1.131	-13.134	9.028	-0.933	19.398
Real Estate	21/09/2012	31/12/2022	2,585	0.013	1.271	-19.798	8.392	-1.837	34.014
Energy	21/09/2012	31/12/2022	2,585	0.005	1.878	-23.413	14.641	-1.008	20.032
Basic Materials	21/09/2012	31/12/2022	2,585	0.024	1.381	-11.093	11.424	-0.474	11.277
Industrials	21/09/2012	31/12/2022	2,585	0.039	1.231	-13.175	11.031	-0.739	16.966
Consumer Staples	21/09/2012	31/12/2022	2,585	0.031	0.963	-10.581	7.647	-1.109	20.846
Health Care	21/09/2012	31/12/2022	2,585	0.047	1.114	-11.231	7.517	-0.518	11.746
Telecommunications	21/09/2012	31/12/2022	2,585	0.007	1.116	-9.433	7.997	-0.362	8.776
Utilities	21/09/2012	31/12/2022	2,585	0.028	1.159	-12.236	12.389	-0.347	22.955
Financials	21/09/2012	31/12/2022	2,585	0.037	1.312	-15.791	11.291	-0.848	21.898
Technology	21/09/2012	31/12/2022	2,585	0.053	1.455	-14.575	10.541	-0.533	11.921

Note. The table shows descriptive statistics for the log-returns of the specified indices. The first price level observation is lost when computing log-returns as lagged values are required. A daily date series accompanies each returns series to comprise the time-series required for future analysis. These times series, however, are incomplete with gaps as weekends are non-trading days. Data for the CRSP sector and total market indices obtained from Wharton Research Data Services are only available until 31/12/2022, whereas the Broad Index data obtained from a proprietary API key which is updated continuously.

Table 2 shows various descriptive statistics for the daily CRSP and Broad returns series. First, while relatively different, in absolute terms, all series means fluctuate approximately around zero. While the standard deviations vary, they are mostly between 1.3-1.5% and the Energy sector appears to the most volatile at 1.88%. On the other hand, the Broad index seems to be the least volatile. Differences between Broad and the rest of the indices regarding volatility, extremes, and mean may suggest that Broad is affected by shocks different in characteristics than those of the CRSP indices. Observing skewness and kurtosis, we see that all CRSP indices are considerably negatively skewed, indicating fatter tails on the left side of their distributions. Broad, on the other hand, is positively skewed. Thus, if one were to randomly draw from these distributions, they would be considerably more likely to observe an extreme negative return from CRSP indices than the Broad index. With high kurtosis, all returns series also have very fat tails, overall. Considering that the normal distribution has a skewness of zero and kurtosis of 3, it is reasonable to conclude that none of the returns exhibit a normal distribution. The results of a series of Jarque-Bera tests for normality reinforce the aforementioned notion.

In addition to basic descriptive statistics, we analyse preliminary indicators of a relationship across our variables. Observing basic Pearson correlations shown in Figure 1, we see that overall, the Broad Index is weakly negatively correlated with all sector indices, suggesting that increases in the USD index correspond to weak drops in sector index levels. While only a mere association, this idea does seem to correspond well with the flow-oriented family of theories outlined, whereby the performance of equity markets is directly linked to the export-competitiveness of the currency. In this case, a higher USD Broad index level makes the USD relatively more expensive, lowering export competitiveness, presumably hurting some sectors. The sector indices, however, appear to have strong positive correlations amongst themselves with some approaching a perfect $r = 1$. Of these, Total Market appears to have the greatest consistently strong links to other indices. The weakest, on the other hand, is the Energy market. These indicators do appear to be realistic as intuitively, individual sectors are usually strongly interlinked with the overall state of the economy. There may, however, exist some exceptions to this given certain idiosyncrasies of specific industries. Energy, may in fact be such a case.

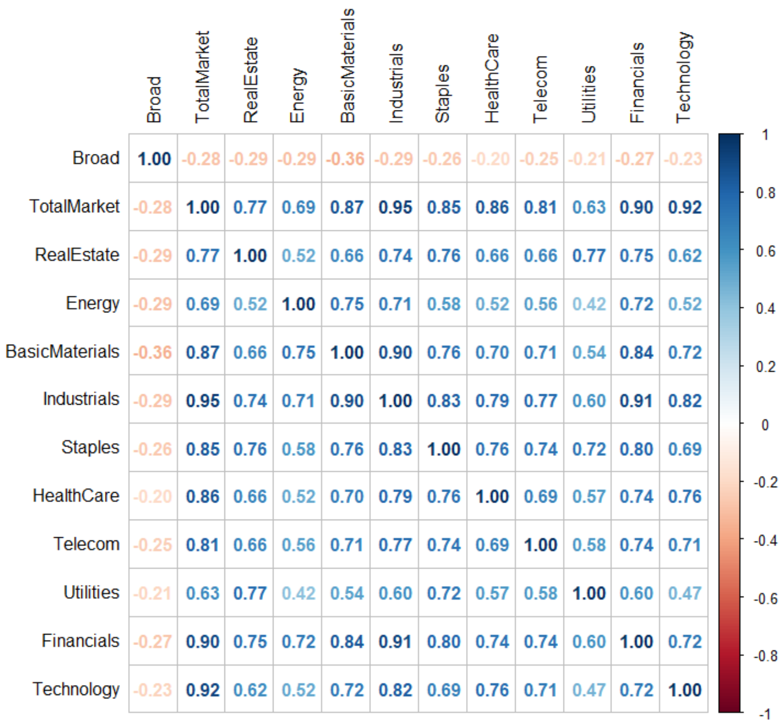

Figure 1: Basic Correlation Matrix across Indices

Note. The diagram depicts the correlation coefficients of each Index against each other. All coefficients are tested against the null hypothesis of no significant correlation with insignificant statistics marked with a cross.

These basic correlations, while helpful in obtaining a general picture, fail to account for possible time-dependencies. To account for this, we also consider cross-correlation coefficients, also known as pairwise-autocorrelations, as shown in Figure 2. With USD Broad serving as the baseline, we see that correlations are negative and largely only significant at around ±0-5 lags. These results suggest that these indices are unable to maintain a significant relationship across a longer period of time and whatever link does exist is unlikely to be positive for any lag length. Ultimately then, we observe weak associations across USD Broad and Sectors, regardless of whether it's viewed across all lag lengths aggregated or segregated.

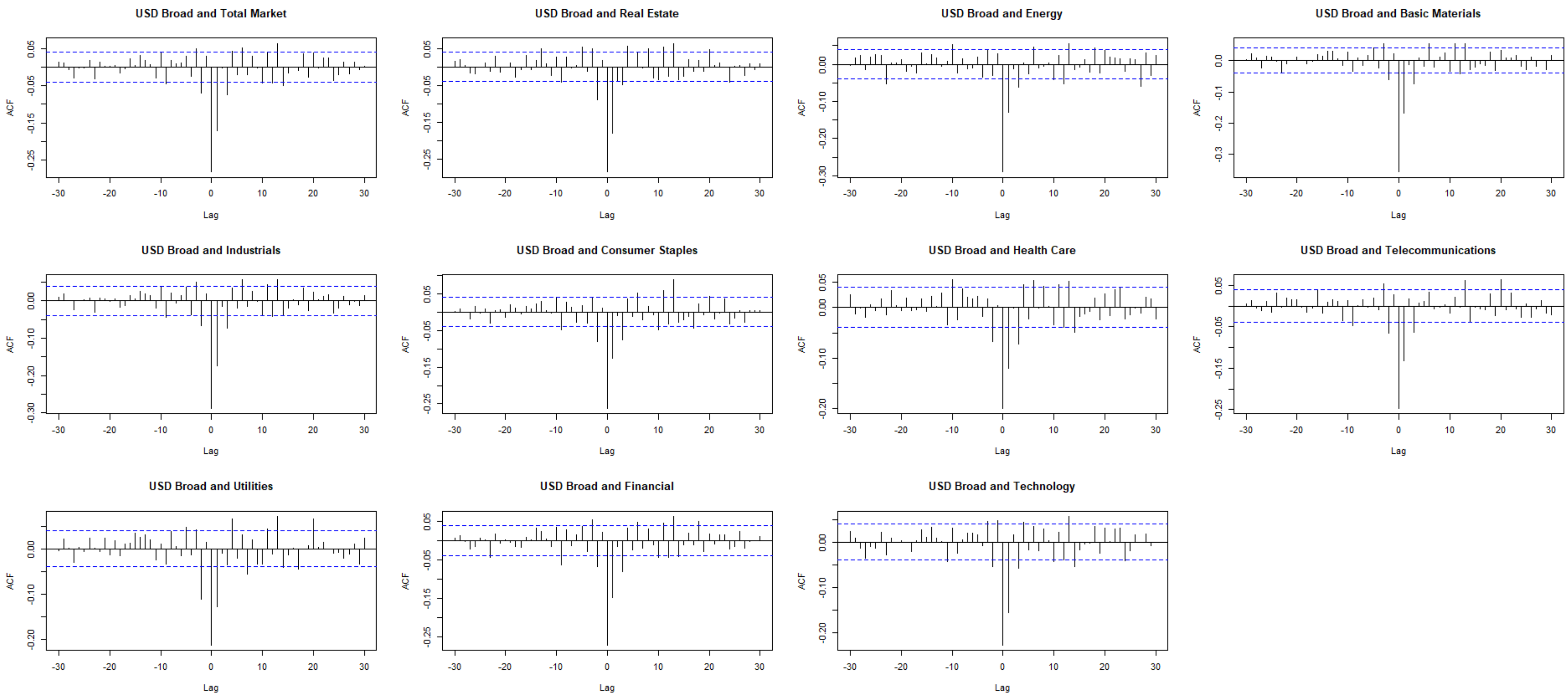

Figure 2: Cross Correlation Matrix across Indices

Note. The above diagram depicts the cross-correlation coefficients for all sector indices against the USD Broad Index. This particular type of correlation enables the decomposition of Pearson correlation into time-varying parts, hence, the lag dimension. Blue lines indicate the threshold from which the autocorrelation coefficient is deemed statistically significant.

4 Methodology

The latter part of the central research question entails investigating the construction of VaR models as a means to understand its spillover from exchange rates to sectors of the U.S. economy. Taking into preceding discussions and prevailing ideas in literature, the following sections outline a set of robust approaches.

4.1 Box-Jenkins Methodology

Fundamentally, VaR can be modelled and forecasted using time-series models that successfully capture the salient features of our financial returns. We mainly utilise the Box-Jenkins methodology to construct necessary models as it provides a systematic, yet comprehensive guideline to construct theoretically-sound models (Lu & AbouRizk, 2009). Generally, the approach utilises three different strategies: identification, estimation, and diagnostics. Firstly, a model, or data-generation process (DGP) is found that fits the data appropriately. Then, the aforementioned model is fitted to the data, often with techniques such as Maximum-Likelihood or Ordinary-Least-Squares. Finally, to ensure statistical adequacy, the estimated model is tested for violations against its core assumptions. The specific considerations within each step are discussed with every type of model necessary in the subsections below.

4.2 Mean and Variance Models

4.2.1 Mean Modelling

To arrive at statistically valid variance models, we must first ensure that the conditional mean $\mathbb{E}[r_t|I_{t-1}]$ in expression 2 is adequately modelled. Namely, we require that it is stationary with its residuals resembling a white noise process, where they are i.i.d with zero mean (Sohail Chand et al., 2012). Generally, the presence of autocorrelation in returns series necessitates fitting conditional mean models to eliminate excessive serial correlation. We can test for autocorrelation through the use of Portmanteau-type tests. For our purposes, an "Automatic Portmanteau" (AQ) test by Escancianoa et al. (2009) is used due to its robustness against conditional heteroskedasticity (i.e., volatility-clustering effects), the automatic selection of test lags through an information criterion algorithm, and its greater statistical power. Given an insignificant test result, then the returns series in question will be used in its unadulterated form in proceeding variance models. Otherwise, an autoregressive-moving-average (ARMA) model will be fitted through a lag-selection algorithm. Specifically, using the `R` `rugarch` package's `auto.afirma()` functionality, we determine which combination of AR(P) and MA(Q) lags in the ARMA(P,Q) model minimize the Akaike Information Criterion (AIC), which can be said to show the model that optimally forecasts from an unknown data generation process while preventing overfitting by including too many lags (Bisgaard & Kulahci, 2009). To ensure that the remaining autocorrelation has been eliminated, the AQ test will be utilised again on the models' residuals.

4.2.2 Variance Modelling

4.2.2.1 ARCH-type Models Utilised

To formally describe the volatility clustering phenomenon outlined previously and to model the unpredictable component $\varepsilon_t = \sigma_t z_t$, Engle (1982) first formulated the Autoregressive Heteroskedasticity (ARCH) model. However, as a response to instances when ARCH lags became too large, Bollerslev (1986) developed the more parsimonious GARCH model. In its "standard" form, it can be shown to be

$$r_t = \mathbb{E}[r_t|I_{t-1}] + \varepsilon_t, \quad \varepsilon|I_{t-1} \sim N(0, \sigma_t),$$
$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{3}$$

where ω is the constant, ε_t^2 is the mean equation residual, and $N(0, \sigma_t)$ denotes that the residuals follow a normal distribution with zero mean and standard deviation σ_t. For the model to be stationary, we require that the persistence $\hat{P} = \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j < 1$. Typically, the model is denoted by GARCH(p,q) where they correspond to the ARCH and GARCH parameter(s) α_j and β_j, respectively. The above expressions imply that to estimate a GARCH(p,q) model, both the conditional mean and variance equations must be jointly computed. To do so, the Maximum Likelihood method is commonly utilised.

While standard GARCH may suffice in capturing basic variance dynamics, it fails to account for other salient features of returns data. Its main omission is arguably the fact that there exist leverage effects in financial returns, as first observed by Black (1976), where positive and negative shocks of ε_{t-j} have differing magnitudes of effects on volatility. In such cases, a class of asymmetric GARCH-type models should be considered (Lee & Su, 2012). Following the discussion in section 2.2, we shall introduce the EGARCH, GJR-GARCH, APARCH and TGARCH models. Though these models have a common goal of integrating asymmetries, they differ in their approaches. Finally, we outline iGARCH, a special case.

First, the Exponential GARCH, or EGARCH, model was formulated by Nelson (1991) and expresses the conditional variance in the logarithmic form. The variance equation can be shown as

$$\ln \sigma_t^2 = \omega + \sum_{j=1}^{q} \left(\alpha_j z_{t-j} + \gamma_j \left(|z_{t-j}| - \mathbb{E}|z + t - j| \right) \right) + \sum_{j=1}^{p} \beta_j \ln \sigma_{t-j}^2, \tag{4}$$

where the γ_j parameters dictate the magnitude of the asymmetric shocks. Where $\gamma_j = 0$, there is no asymmetric effect, however, $\gamma_j < 0$ indicates that negative shocks have greater effect than positive ones and $\gamma_j > 0$ indicates the converse. Thanks to the model's logarithmic specification, the only requirement for stationarity is that $\hat{P} = \beta_j < 1$. The α_j parameter can be interpreted similarly to GARCH.

The so-called "GJR-GARCH" model composed by Glosten et al. (1993) instead utilises an indicator function I to determine the asymmetric sign effect. Their variance model is expressed as

$$\sigma_t^2 = \omega + \sum_{j=1}^{q}\left(\alpha_j \varepsilon_{t-j}^2 + \gamma_j I_{t-j}\varepsilon_{t-j}^2\right) + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2, \tag{5}$$

where γ_j is the so-called 'leverage' term indicating the effect when I holds value 1 when $\varepsilon \leq 0$ and zero otherwise. According to Ghalanos (2022), the persistence of the model must now be $\hat{P} = \sum_{j=1}^{q}\alpha_j + \sum_{j=1}^{p}\beta_j + \sum_{j=1}^{q}\gamma_j\kappa$ as the persistence is driven by the asymmetry of the conditional distribution. Here, κ can be seen as the expectation of $z_t < 0$.

The asymmetric power ARCH model, or APARCH, was formulated by Ding et al. (1993) and permits the variance and residual terms to hold a power different than 1. The notable innovation with the model arises from Taylor's (1986) observation that autocorrelations of absolute returns tend to be larger than those of squared returns. Incorporating power asymmetries and the aforementioned, we can express the model as

$$\sigma_t^\delta = \omega + \sum_{j=1}^{q}\alpha_j\left(|\varepsilon_{t-j}| - \gamma_j\varepsilon_{t-j}\right)^\delta + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^\delta, \tag{6}$$

where $\delta > 0$ and γ_j is the coefficient measuring the sign effect. Here, persistence is a function of the asymmetry of the conditional distribution: $\hat{P} = \sum_{j=1}^{p}\beta_j + \sum_{j=1}^{q}\alpha_j\kappa_j$. Where $\delta = 1$ in the previous model, we arrive at the Threshold GARCH (TGARCH) model formulated by Zakoian (1994). TGARCH is quite similar to GJR-GARCH in its fundamental formulation, however, instead of squared residuals, it utilises absolute residuals.

Finally, we introduce the Integrated GARCH (iGARCH) model by Engle and Bollerslev (1986). The model is an augmented version of the standard GARCH that assumes a unit persistence, i.e., $\hat{P} = \sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{p}\beta_i = 1$. This assumption is crucial for financial returns, which according to the empirical study carried out by Chou (1988) on U.S. stocks, tends to be extremely high. As a result, iGARCH may prevail in instances where volatility clustering is extremely high. Additionally, by enforcing $\hat{P} = 1$, β_i is not estimated, but computed; it follows that only α_i is estimated.

4.2.2.2 Variance Modelling Identification and Diagnostics

However, before estimating any variance model, we must ensure that our returns contain volatility effects that GARCH-type models can capture. This is done by testing for so-called ARCH effects through the Li-Mak test. Per Lundbergh & Terasvirta (1999), a rejection of the null hypothesis of no autocorrelation in the returns' squared residuals would necessitate variance modelling.

After this identification phase, we fit the GARCH-type models discussed to our returns series. There is an incredibly influential body of literature advocating for the philosophy of parsimony in time series modelling. Arguably the most well known investigation into ARCH-type models, the investigation of Hansen et al. (2005), presents compelling evidence that the most parsimonious GARCH/ARCH specifications may in fact be the most optimal fit for financial returns. Indeed, their empirical findings are supported by theory; addressing criticisms regarding the accuracy of GARCH-derived forecasts, Anderson and Bollerslev (1998) defend parsimony, arguing that complex specifications require specific empirical backing and imply that they lead to the well-known trap of overfitting. To avoid these issues and to reduce computation time, we therefore only test the most parsimonious specifications of the models discussed, i.e., ARCH and GARCH lag orders of one.

It is important to note that some returns series may also be fitted to ARMA(P,Q) models; in this case, the ARMA specification is used in the conditional mean component. Further, as discussed in section 2.2.2, financial returns tend to have leptokurtic distributions and are typically skewed to the left (Lee & Su, 2012). It follows that their GARCH residuals may not abide by the commonly-assumed standard normal distribution. Thus, to account for these deviations, we introduce the skewed GED distribution to fit to all GARCH residual series. The probability density function of the skewed GED can be shown to be

$$PDF_{SGED} = \frac{k^{1-\frac{1}{k}}\Gamma\left(\frac{1}{k}\right)^{-1}\exp\left(-\frac{1}{k}\frac{|u|^k}{\varphi^k(1+\text{sign}(u)\lambda^k)}\right)}{2\varphi},$$
$$\Gamma(a) = \int_0^\infty \frac{x^{a-1}}{e^x}\,\mathrm{d}x \quad \& \quad k(a) = \frac{\sin a\pi}{a\pi}, \tag{7}$$

where φ is a scaling constant corresponding to the standard deviation of the series, and λ is the skewness parameter, and k is the kurtosis parameter. After fitting every GARCH-type model discussed to every returns series, we must ultimately select the prevailing models. To do so, we borrow from the ARMA lag selection procedure and use AIC as our criterion for attaining a balance between models' ability to forecast and parsimony. The model that minimises AIC will then continue onto the two-stage diagnostics phase, where its statistical validity will be tested. Here, we first check whether the model has captured all remaining ARCH effects by performing the AQ test on the model residuals - the presence of a white noise process would indicate an adequate fit (Escanciano & Lobato, 2009). Further, we compare the distribution of the residuals series to the corresponding custom skewed GED distribution to ensure that the error is also correctly specified. As a benchmark for the adequacy of the fit, we also compare it to the theoretical Student's t distribution. To do so, we employ the χ^2 goodness of fit test by Palm (1993). When most bins (hereafter denoted as O_i where $i = \{1,2,3,4\}$) return insignificant, or equivalently, that the null hypothesis of no difference between the distributions is not rejected, we consider the fit adequate.

4.2.3 VaR Evaluation, Forecast, and Diagnostics

To evaluate the performance of the variance models constructed and compare it against models derived under Student's t, we estimate our winning variance models using the first 80% of our available sample and then use the remaining 20% for out-of-sample forecasting. In doing so, conduct 1-day-ahead rolling forecasts, where each moving estimation window is 25 days, after which the model parameters are re-estimated. Then, we compute daily VaR for all forecasted series under VCM as is shown below.

$$\text{VaR}_{m,t}(\alpha) = \mathbb{E}[r_t|I_{t-1}]_{m,t} + z_{m,\alpha}\sigma_{m,t}, \tag{8}$$

where $\mathbb{E}[r_t|I_{t-1}]_{m,t}$ is the conditional expectation of returns series m at day t, $z_{m,\alpha}$ is the α quantile of the custom skewed GED or Student's t distribution specified in the previous subsection belonging to series m, and $\sigma_{m,t}$ is the conditional variance specified for series m. The specific quantiles are shown in the Results section. Note that since we evaluate VaR at $\alpha = \{0.01, 0.05\}$, we end up with two sets of VaRs for each series.

Similar to the GARCH and ARMA diagnostics, we must be able to ensure that our VaRs are adequate. One way of doing so involves computing the number of times that the realised returns were observed to be lower than the VaR series, also referred to as being in violation of VaR, proportional to the total number of observations. Hereafter, we refer to this proportion as the coverage and denote it by $\mathbb{E}\left[r_{m,t} < \mathrm{VaR}_{m,t}\right]$. Generally, the closer the coverage is to the corresponding VaR α, the better the model (Su, 2015). Beyond this informal assessment, we formally test for the gap between the coverage and the expected proportion of violations using Kupiec's (1995) Unconditional Coverage backtest (hereafter denoted by LR_{UC}). Using the asymptotic χ^2 distribution, it assesses the null hypothesis that $\mathbb{E}\left[r_{m,t} < \mathrm{VaR}_{m,t}\right] = \alpha$. Thus, failing to reject the null hypothesis is an indicator of an adequate VaR model. However, a shortcoming of the LR_{UC} is that it does not inform us of whether violations are correlated over time (Glyn A, 2013). Fortunately, the Conditional Coverage backtest by Christoffersen (1998) (denoted by LR_{CC}) addresses this issue by evaluating the independence of violations over time using the same distribution as Kupiec's test. A failure to reject the null hypothesis further indicates the adequacy of our VaR model by demonstrating that violations are independent of each other and observed with the expected frequency. Ultimately, we compare the backtests and convergences across both distributions specified to determine whether SGED is first, adequate in modelling VaR, and if so, whether it is an improvement over a Student's t specification.

4.3 Granger Causality Tests for One and Two-Way Risk Spillover

Finally, to determine whether there is a spillover of risk across markets, we utilise a special test formulated by Hong (2009). Using the Granger Causality framework, the test takes as inputs the violations of each VaR series (in binary form) and indicates whether information regarding risk in one market improves your ability to forecast violations in another (Fan et al., 2008). First, we must define a function that indicates VaR violations:

$$Z_{m,t} = \xi\left(r_{m,t} < \mathrm{VaR}_{m,t}\right), \tag{9}$$

where ξ is an indicator function holding value 1 when, at a given market m and time t, VaR is greater than realised return. In order to derive more insights regarding the direction of potential Granger-causality across exchange rates and sector indices, we use both one-way and two-way tests. With the one-way test, we assess causality from the Broad Index to the sector indices and vice versa. The null and alternate hypotheses for the one-way test can be shown to be

$$H_0^1 : \mathbb{E}\left[Z_{1,t}|I_{1,t-1}\right] = \mathbb{E}\left[Z_{1,t}|I_{t-1}\right] \quad \text{vs} \quad H_A^1 : \mathbb{E}\left[Z_{1,t}|I_{1,t-1}\right] \neq \mathbb{E}\left[Z_{1,t}|I_{t-1}\right]. \tag{10}$$

In other words, if we are testing a risk spillover from $m = 2$ to $m = 1$ and there is insufficient evidence to reject H_0^1, then we can assert that a VaR violation in $m = 2$ is not necessarily useful in predicting future VaR violations in $m = 1$. The converse reasoning is valid when the test p-value returns significant, i.e., $p \geq 0.05$. A similar logic is used to arrive at the null and alternate hypotheses of the two-way variant of the test.

The statistical foundations of the aforementioned tests lie in the cross-correlation function (CCF) between the indicator functions of VaR series in question. Based on this, Hong (2009) developed the one and two-way test statistics as shown below

$$Q_1(M) \equiv \frac{T\sum_{j=1}^{T-1}\left(k^2\left(\frac{j}{M}\right)\hat{\rho}(j) - C_{1T}(M)\right)}{\left(D_{1T}(M)\right)^{\frac{1}{2}}} \quad \& \quad Q_2(M) \equiv \frac{T\sum_{|j|=1}^{T-1}\left(k^2\left(\frac{j}{M}\right)\hat{\rho}^2(j) - C_{2T}(M)\right)}{\left(D_{2T}(M)\right)^{\frac{1}{2}}}, \tag{11}$$

where the centering and standardisation constants are:

$$\begin{gathered} C_{1T}(M) \equiv \sum_{j=1}^{T-1}\left(1-\frac{j}{T}\right)k^2\left(\frac{j}{M}\right) \quad \& \quad C_{2T}(M) \equiv \sum_{|j|=1}^{T-1}\left(1-\frac{|j|}{T}\right)k^2\left(\frac{j}{M}\right), \\ D_{1T}(M) \equiv 2\sum_{j=1}^{T-1}\left(1-\frac{j}{T}\right)\left(1-\frac{j+1}{T}\right)k^4\left(\frac{j}{M}\right) \quad \& \\ D_{2T}(M) = 2\left[1+\hat{\rho}^4(0)\right]\sum_{|j|=1}^{T-1}\left(1-\frac{|j|}{T}\right)\left(1-\frac{|j|+1}{T}\right)k^4\left(\frac{j}{M}\right). \end{gathered} \tag{12}$$

Here, Q_1 and Q_2 denote the one and two-way test statistics, respectively. Additionally, we consider $k(\cdot)$ to be a kernel function with the purpose of assigning weights to different lags, denoted by M. While there a variety of different kernels suggested by Hong (2009), we implement the recommendations of Wang et al. (2019) and utilise the Daniel kernel, which can be shown to be $k_d(z) = \sin(\pi z)/\pi z$. This specific kernel function assigns higher discounts to larger lags M, which is intuitive as time-series and financial analysis typically shows that distant historical information (shocks) are less informative than those of today. Regardless, Hong (2009) demonstrates that the choice of $k(\cdot)$ does not have a significant impact on the efficiency of the lag-weighting process as long as the weighing is not uniformly applied.

5 Results and Discussion

5.1 Mean and Variance Modelling Results

In line with the Box-Jenkins methodology, we evaluate the identification, estimation, and diagnostics of our fitted models. First, the results of the mean modelling phase can be seen in Table A2. We observe that the Automatic Portmanteau statistics for all returns series except for Total Market, Healthcare, and Technology are insignificant, signifying an absence of autocorrelation. This result seems to be in line with the Efficient Market Hypothesis of Fama (1991), which suggests that the pricing of assets should take into account all available information and that patterns would be immediately exploited such that they would no longer remain.

Subsequently, autocorrelation in raw financial returns should not, in theory, exist. However, as demonstrated by Fama (1965) himself, certain inefficiencies in markets can indeed result in short-term autocorrelation that enables some predictability in financial returns. This is what we likely observe with the significant autocorrelations of Total Market, Healthcare, and Technology. Consequently, we fit an ARMA(P,Q) model to eliminate the remaining autocorrelation. The ARMA lag selection criterion identifies ARMA(0,2), ARMA(2,2), and ARMA(1,0) as the models that best minimise AIC for each respective returns series. Repeating the diagnostics, we observe near-zero PQ statistics, confirming that autocorrelation has likely been eliminated. Having addressed the modelling of the raw returns, we confirm the necessity of variance models by formally testing for volatility clustering. Table A2 shows that all indices, regardless of whether they have a conditional mean model or not, exhibit extremely high ARCH-LM statistics, indicating the presence of strong so-called "ARCH" effects. Hence, variance modelling is a valid procedure in this case.

Fitting all discussed GARCH-type models with the appropriate mean equation, we observe that AIC is minimised when the Broad Index is fitted to EGARCH(1,1), Total Market to APARCH(1,1)-ARMA(0,2), Real Estate and Telecom to GJR-GARCH(1,1), Energy, Basic Materials, Industrials, Consumer Staples and Financial to TGARCH(1,1), Healthcare to APARCH(1,1)-ARAMA(2,2) and Technology to TGARCH(1,1)-ARMA(1,0). In each of these cases, the performed diagnostics confirmed that the AIC-chosen model eliminated all remaining ARCH and residual autocorrelation, reflecting that the models are statistically sound. However, the procedure proved more difficult for the Utilities index, because the top 4 models chosen by AIC all continued to yield significant ARCH effects, meaning that they failed to sufficiently capture variance effects. However, it was only the final model, iGARCH(1,1) that exhibited an insignificant statistic, succeeding in the diagnostics. For this reason, it instead was chosen. Even with iGARCH, though, the ARCH-LM statistic is considerably higher than other successful models, indicating a particular complexity in its variance structure, at least for our assortment of GARCH-type models. In general, however, the models' AICs and subsequent diagnostics were extremely similar, resulting in most prevailing by very close margins. In other words, it appears for the choice of model is not very important when determined by AIC. A possible clue for this phenomenon could arise the sheer similarity of the conditional variances across our array of models. Figure 3 shows that with the exception of the Broad Index, the conditional variance generated by the winning models is very similar when observing trends and peaks. These features may indicate that these indices have some level of interdependence or significant cross-correlations, resulting in the assortment of models capturing similar variance dynamics across the returns series. This assertion is reinforced by the investigation of Bordoloi (2009) into U.S and world equity markets, that finds that in general although the largest dependencies occur in response to sectors' own shocks, there is also significant interdependence on other sectors' shocks. In particular, they observe this interdependence to be greatest at times of extreme volatility, which is observed in our case since our final set of estimation periods took place during the heightened volatility period brought on by the SARS-CoV-2 pandemic.

Pursuant to the discussion in Section 2 about the nature of volatility models best suited to model financial returns, our results are inconclusive on account of several factors. Firstly, the general assertions of Hansen and Lunde (2005) that simpler GARCH-type models can better model exchange rate data and more complex models perform better on equity indices do not appear to completely hold. For instance, for USD Broad, AIC was markedly higher for GARCH(1,1) than EGARCH(1,1), suggesting that it is a considerably worse fit than the more complex model. However, in line with Hansen et al., winning sector index models were more complex asymmetric models the likes of TGARCH(1,1). These disagreements concerning the right level of model complexity may be partially due to the selection criteria utilised. Hansen (2001) largely bases their results on the Superior Predictive Ability (SPA) test, which works by determining whether the mean squared prediction error (MSPE) of out-of-sample forecasts is statistically different across a range of models. AIC, on the other hand, focuses predominantly on in-sample models and enforcing parsimony, penalising models deemed too complex. Arguably, these differences are such that applying AIC can better aid with initial model selection and is superior in systematic application across a large number of models. However, the deep-dive of Cai et al. (2010) into SPA suggests that based purely on the selection of a forecasting model, it may be superior in performance. Additionally, in-line with evidence from Hansen et al. (2005), the most parsimonious specification, i.e., ARCH and GARCH lags of 1, always yielded adequate diagnostics results. This particular result should not stand to surprise since it is widely believed to be a core tenet of time-series modelling (Andersen & Bollerslev, 1998). This principle extends not only to univariate GARCH modelling, but also to that of multivariate, such as with Tastan's (2006) M-GARCH study. According to Mcaleer et al. (2008), parsimonious multivariate models can help circumvent the so-called "curse of dimensionality"; as higher-order models tend to incorporate more than two dimensions, the volume of data necessary to reliably conduct parametric analysis can grow exponentially, potentially rendering it computationally infeasible.

Fitting each winning model, we utilise the Maximum Likelihood Estimation technique and apply an assumption of SGED to the errors of the models. The results of the estimations are shown in Table 3. First, all variance equation parameters are significant at 5%, suggesting conditional variance does indeed depend on lagged values. Moreover, all persistence levels other than iGARCH are all below 1, indicating that the models exhibit stationarity, as required by their core assumptions. However, all models' persistence levels are above 0.94, suggesting the presence of high levels of volatility clustering and long memory across the board. In other words, shocks decay very slowly and are unable to stabilise to an equilibrium swiftly. Persistence in this manner cannot be extended to Utilities's iGARCH specification; when attempting to observe a long-term unconditional variance, the process explodes and tends to infinity. As such we only interpret the shorter-term volatility effect. The process yields an $\alpha = 0.0970$, implying that the Utilities sector's volatility does not have a large reaction to new shocks. Interestingly, we observe differing leverage effects across the models. The USD Broad EGARCH(1,1) $\gamma > 0$, meaning that when shocks are positive, there is an additional effect of 0.1137. In contrast, the $\gamma > 0$ parameters of the GJR-GARCH, APARCH and η_1 of TGARCH all demonstrate an asymmetric effect opposite that of USD Broad's EGARCH – namely that negative shocks affect conditional volatility more than positive shocks. This asymmetric negative effect we observe is in line with what is typically seen in financial research, as leverage effects are a recognised stylised fact of financial returns

(Cont, 2001). Further, reinforcing the discussion on other stylised facts of returns, we observe that the results of tests for excess kurtosis and skewness, indicated by λ and ξ respectively, are all statistically significant at 1%. Thus, we have no reason to suspect that our returns series deviate from these norms and possess evidence for non-normality in their distributions. Analysing the trends in each conditional variance series depicted in Figure 3, the similarities between the sector indices are apparent; in addition to the overall levels of variance being similar, the various indices exhibit a similar pattern of shocks and peaks. Notably, it appears that all series have a sudden jump in volatility at around February 2020, which coincides with the beginning of the SARS-CoV-2 pandemic that resulted in extreme distress in financial markets. Unlike the sector indices, however, the Broad Index exhibits much lower overall conditional variance levels, especially when other sector indices have large peaks.

Table 3: Results of fitted Models

	USD Broad	Total Market	Real Estate	Energy	Basic Materials	Industrials	Consumer Staples	Health Care	Telecom	Utilities	Financial	Technology
	E-GARCH(1,1)	**ARMA(0,2) APARCH(1,1)**	**GJR-GARCH(1,1)**	**TGARCH(1,1)**	**TGARCH(1,1)**	**TGARCH(1,1)**	**TGARCH(1,1)**	**ARMA(2,2) APARCH(1,1)**	**GJR-GARCH(1,1)**	**iGARCH(1,1)**	**TGARCH(1,1)**	**ARMA(1,0) TGARCH(1,1)**
μ	0.0103	0.0175***	0.0164	-0.0198	0.0030	0.0264	0.0081	0.0445***	-0.0059	0.0401**	0.0236	0.0416***
	[0.0065]	[0.0061]	[0.0180]	[0.0237]	[0.0207]	[0.0268]	[0.0136]	[0.0041]	[0.0211]	[0.0193]	[0.0172]	[0.0088]
AR(1)	-	-	-	-	-	-	-	-1.5027***	-	-	-	-0.0695***
								[0.0158]				[0.0241]
AR(2)	-	-	-	-	-	-	-	-0.6772***	-	-	-	-
								[0.0299]				
MA(1)	-	-0.0554***	-	-	-	-	-	1.4687***	-	-	-	-
		[0.0098]						[0.0189]				
MA(2)	-	0.0268**	-	-	-	-	-	0.6312***	-	-	-	-
		[0.0083]						[0.0285]				
ω	-0.0251**	0.0457***	0.0344***	0.0198***	0.0263***	0.0338***	0.0388***	0.0496***	0.0555***	0.0108***	0.0485***	0.0467***
	[0.0121]	[0.0068]	[0.0092]	[0.0060]	[0.0065]	[0.0085]	[0.0071]	[0.0086]	[0.0138]	[0.0036]	[0.0084]	[0.0087]
α_1	0.0292**	0.1463***	0.0504**	0.0831***	0.0755***	0.0995***	0.1382***	0.1079***	0.0468**	0.0970***	0.1528***	0.1108***
	[0.0119]	[0.0149]	[0.0210]	[0.0145]	[0.0135]	[0.0218]	[0.0178]	[0.0133]	[0.0203]	[0.0145]	[0.0193]	[0.0192]
β_1	0.9901***	0.8500***	0.8588***	0.9233***	0.9211***	0.8929***	0.8475***	0.8759***	0.8572***	0.9030†	0.8374***	0.8782***
	[0.0046]	[0.0165]	[0.0253]	[0.0134]	[0.0142]	[0.0234]	[0.0189]	[0.0167]	[0.0248]	[NA]	[0.0194]	[0.0268]
γ_1	0.1137***	0.8200***	0.1071***	-	-	-	-	0.8418***	0.0818***	-	-	-
	[0.0191]	[0.0786]	[0.0270]					[0.1024]	[0.0287]			
δ	-	0.8006***	-	-	-	-	-	0.6720***	-	-	-	-
		[0.1246]						[0.1858]				
η_1	-	-	-	0.5293***	0.9194***	0.8187***	0.6977***	-	-	-	0.7486***	0.6801***
				[0.1149]	[0.1652]	[0.1771]	[0.1049]				[0.0951]	[0.0178]
λ	1.0937***	0.8030***	0.8322***	0.9271***	0.8220***	0.8003***	0.8133***	0.8481***	0.8381***	0.8331***	0.8441***	0.7915***
ξ	1.3599***	1.3923***	1.5130***	1.4721***	1.5245***	1.3811***	1.5287***	1.4622***	1.5106***	1.3964***	1.4298***	1.2762***
$\hat{P}$	0.9901	0.9522	0.9589	0.9869	0.9795	0.9687	0.9543	0.9494	0.9494	1.0000	0.9541	0.9615
Log-Likelihood	-242	-2194	-2553	-3290	-2868	-2513	-2087	-2522	-2649	-2475	-2499	-2775
AIC	0.2574	2.2772	2.6442	3.4064	2.9701	2.6034	2.1633	2.6180	2.7441	2.5625	2.5884	2.8752
PQ_A	1.8922	1.9544	1.1592	1.047	1.746	2.687	2.313	1.6046	2.5077	1.7040	3.3350*	2.329
Weighted ARCH-LM (7)	1.6263	2.3557	3.3878	3.5030	0.2200	3.6503	4.7170	5.518	0.8096	6.48711	2.3632	1.1625

Note. † By construction, iGARCH requires that the ARCH and GARCH parameters equal one by computing $1-\sum_{i=1}^{q}\alpha_i-\sum_{i>1}^{p}\beta_i$. As such, the β_i is simply calculated, not estimated, meaning that it has no associated standard error or p-value. ‡ The ARCH-LM test utilised is that of a weighted Portmanteau type, as formulated by Fisher and Gallagher (2012). λ and ξ represent the results of tests for excess kurtosis and skewness, respectively. $\hat{P}$ denotes the persistence of each respective GARCH model. Significance levels are denoted as follows: $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

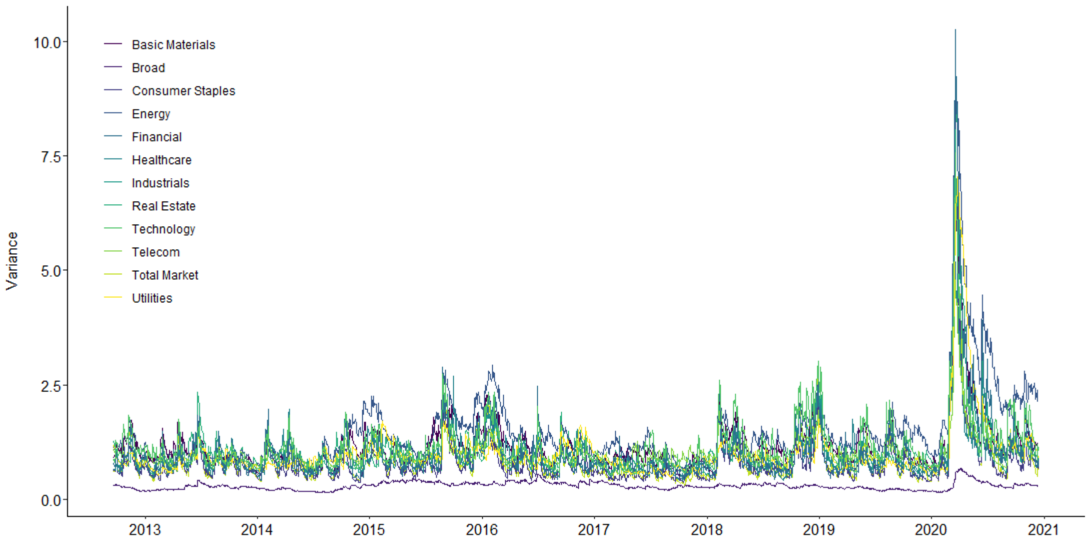

Figure 3: Conditional Variance for each Winning GARCH-type Model

Note. The illustration depicts the trends in the conditional variance of the models fitted to the various indices discussed. The time series observed is that of the estimation window, i.e., the first 80% of our observations amounting to 1936 days. The remaining 20% is used to carry out diagnostics on out-of-sample forecasts.

5.2 VaR Estimation

5.2.1 Derivation of Theoretical Returns Quantiles

Having estimated and verified the validity of our variance models, we may begin the procedure of forecasting and computing VaR. In doing so, we must first evaluate the relevant 95% and 99% quantiles under SGED. Table 4 shows the actual quantiles of each index compared with those of their custom SGED and Student's t counterparts. Analysing actual quantiles, it is apparent that the overwhelming majority of sector indices have 1% and 5% tails that are much fatter than those of Student's t, meaning that it would greatly underestimate true downside risk, making subsequent VaR models inadequate. SGED quantiles, on the other hand, appear to better correspond to actual returns as they are generally lower than t for 1%. However, at 5%, Student's t appears to have some edge over SGED, as it is closer for Total Market, Energy, Consumer Staples, and Utilities. Thus, in some instances, SGED seems to overestimate risk at 5%. An outlier in the discussion is the Broad Index, which has comparatively thin and even a positive 5% tail. In this case, Student's t 1% and SGED 5% fit better, although both appear to greatly overestimate true risk.

Further, to determine the veracity of our assertion that SGED specifies the error specification in GARCH models better than Student's t, we assess the statistical similarities across each distribution and the distribution of each winning GARCH model's residuals by means of a special χ^2 Goodness of Fit test by Vlaar and Palm (1993). The results show that for SGED, bins from O_1 to O_4 largely return insignificant test statistics, indicating that we cannot reject the test's null hypothesis that the two aforementioned distributions are identical. On the contrary, the results from adopting Student's t are almost all highly significant, demonstrating that using that error specification would lead to less statistically-sound variance models. The only exception to this appears to be USD Broad, whose null hypothesis is not rejected across the board. A possible reason for this could be that the Broad index has a significantly thinner distribution in the tails than the rest of the indices, making t a comparatively better fit. Notwithstanding, the corresponding SGED test statistics, thereby, fit, are not too dissimilar.

Table 4: Quantile and Distribution Results

	Original Distribution		Skewed GED Error Specificiation						Student's t Error Specification					
	Quantiles		Quantiles		χ^2 Goodness of Fit (GARCH Residuals)				Quantile		χ^2 Goodness of Fit (GARCH Residuals)			
	0.01	0.05	0.01	0.05	O_1	O_2	O_3	O_4	0.01	0.05	O_1	O_2	O_3	O_4
USD Broad	-0.770	0.471	-2.404	-1.573	25.59	43.41**	45.49	59.14	-2.552	-1.585	18.11	33.15	42.18	46.08
Total Market	-3.258	-1.569	-2.828	-1.775	29.70*	41.64*	56.07**	53.98	-2.511	-1.524	50.71***	56.12***	68.05***	69.42**
Real Estate	-3.235	-1.757	-2.716	-1.750	30.36**	27.33	44.37	56.25	-2.478	-1.557	52.76***	62.16***	75.07***	79.29***
Energy	-4.617	-2.565	-2.638	-1.724	29.06*	47.16**	48.71	64.67*	-2.508	-1.600	41.91***	58.69***	62.31***	78.36***
Basic Materials	-3.559	-2.054	-2.737	-1.769	20.94	30.98	42.97	56.87	-2.463	-1.576	40.14***	56.55***	60.24**	74.64***
Industrials	-3.535	-1.754	-2.830	-1.769	14.48	30.08	33.01	50.47	-2.492	-1.517	35.59**	59.87***	64.91***	89.88***
Consumer Staples	-2.856	-1.347	-2.741	-1.769	27.60*	38.48	48.17	66.79**	-2.470	-1.579	44.00***	53.27***	68.59***	84.97***
Health	-3.227	-1.664	-2.694	-1.714	19.06	27.67	35.45	46.13	-2.458	-1.532	39.89***	46.48**	57.43**	73.56**
Telecommunications	-3.156	-1.721	-2.731	-1.768	13.96	26.18	31.48	54.03	-2.484	-1.583	33.44**	47.41**	64.91***	70.46**
Utilities	-3.564	-1.534	-2.759	-1.732	17.95	43.01**	47.31	69.32**	-2.506	-1.507	33.69**	49.70***	57.10**	77.95***
Financial	-3.610	-1.769	-2.739	-1.739	24.62	36.75	44.66	58.16	-2.483	-1.560	34.39**	53.70***	59.45**	72.52**
Technology	-4.078	-2.059	-2.900	-1.765	26.05	37.18	38.67	51.65	-2.503	-1.449	61.56***	77.35***	81.23***	96.80***

Note. The values 0.01 and 0.05 under the Quantile columns represent probability levels. The values O_1 to O_4 under χ^2_{VP} signify the various bins tested. Significance levels are denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5.2.2 Forecast and Estimation of VaR under VCM

In accordance with the central question, we determine the adequacy of our VaR models by conducting various diagnostics on one-step ahead forecasts for both SGED and t-specified error distributions. Table 5 shows the convergence for the forecast period for $\alpha = 0.01$ and $\alpha = 0.05$ alongside each respective series mean. According to Su (2015), a preliminary indicator of a good VaR model fit is when the convergence is close or identical to the desired alpha level. A glance at the figures suggests that all SGED-based models seem to be relatively close to the alpha values, with some exceptions. On the contrary, t-based estimates appear comparatively worse, and in some cases, heavily misspecified. However, we cannot definitively conclude on the quality of our models based simply on these convergences. Rather, we do so using statistics from Christoffersen's (1998) and Kupiec's (1995) tests for convergence. Indeed, in line our assertions, all SGED models' statistics are highly insignificant, suggesting that VaR violations do not exceed theoretical levels expected and that each violation is likely independent of each other. In other words, all VaR forecasted under ARMA-GARCH-SGED can be considered to be adequately modelled. This conclusion does not hold for SGED's counterpart; for several indices such as Total Market, Real Estate, Consumer Staples, and Technology several test results are significant at either 5% or 10%, suggesting that exceedances are above expected levels. Where LR_{CC} and LR_{UC} are insignificant, we still observe that p-values are considerably lower than corresponding SGED estimates, suggesting that these are comparatively worse VaR models.

Though all SGED models pass this diagnostics stage, some estimates are more precise than others. While the Conditional Convergence p-values indicate that all 1% estimates are well specified, at 5%, Health, Telecommunications, and Real Estate p-values are relatively close to being significant at 10% confidence level. Christoffersen's test results agree with Kupiec at 1%, but generally penalise 5% results markedly more. As we see it, there are two distinct reasons for these results. The first revolves around the backtesting methods themselves. Per Glyn (2013), as VaR violations are naturally rare, both test have relatively limited power with a sample size like this, impacting their reliability. Additionally, the accuracy of Christoffersen's test is uniquely weakened by the fact that it does not account for instances where there are no consecutive VaR violations. Nevertheless, both tests combined with convergences themselves are known to provide an adequate glance at the quality of VaR models. Besides possible shortcomings in these backtesting methods, the uniquely volatile nature of financial markets throughout 2020-2022 may also be responsible. In their detailed comparison study of various different VaR computation techniques, Abad and Benito (2013) opted to utilise various world equity indices, including the S&P500 and Dow Jones. Crafting GARCH specifications under Box-Jenkins and evaluating VaR through both Kupiec's and Christoffersen's tests, they reveal that the precision of GARCH forecasts, and subsequently, VaR, suffers considerably when estimated during a period of excess volatility. To improve the veracity of their results, they test for differences across various loss functions to understand whether VaR estimation quality depends on the level of volatility in the sample utilised. Finally, the authors find that their results are robust to different samples. Thus, having began the forecast right at the beginning of the SARS-CoV-2 pandemic, it is plausible that challenges arising from modelling during such an unusually high period of volatility harmed the quality of our variance forecasts, ultimately increasing the deviation of convergence from the respective α levels.

Table 5: VaR Modelling Results

	Skewed GED Error and VaR Distribution						Student's t Error and VaR Distribution					
	$\mathbb{E}[r_{m,t} < \text{VaR}_{m,t}]$		LR_{CC}		LR_{UC}		$\mathbb{E}[r_{m,t} < \text{VaR}_{m,t}]$		LR_{CC}		LR_{UC}	
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$
USD Broad	0.0083	0.0475	0.2233	2.365	0.1565	0.0636	0.0041	0.0413	2.1784	2.5418	2.1617	0.8134
			(0.8944)	(0.3066)	(0.6924)	(0.8008)			(0.3365)	(0.2857)	(0.1415)	(0.3671)
Total Market	0.0103	0.0579	0.1099	0.6917	0.0053	0.5992	0.0145	0.0847	1.0616	10.3426***	0.8557	10.2534***
			(0.9465)	(0.7076)	(0.9420)	(0.4389)			(0.5881)	(0.0057)	(0.3549)	(0.0014)
Real Estate	0.0145	0.0517	1.0620	2.7580	0.8557	0.0276	0.0207	0.0620	4.6720*	5.3406*	4.2492**	1.3642
			(0.5881)	(0.2518)	(0.3549)	(0.8682)			(0.0967)	(0.0692)	(0.0393)	(0.2482)
Energy	0.0145	0.0455	1.0620	1.0840	0.8557	0.2169	0.0145	0.0620	1.0616	2.0444	0.8557	1.3642
			(0.5881)	(0.5816)	(0.3549)	(0.6414)			(0.5881)	(0.3598)	(0.3549)	(0.2482)
Basic Materials	0.0083	0.0434	0.2233	0.4742	0.1565	0.4654	0.0165	0.0682	4.2490	3.3005	1.7413	3.0397*
			(0.8944)	(0.7889)	(0.6924)	(0.4951)			(0.1195)	(0.1920)	(0.1870)	(0.0812)
Industrials	0.0124	0.0475	0.4119	2.365	0.2609	0.0636	0.0145	0.0744	1.0616	5.3437*	0.8557	5.3014**
			(0.8139)	(0.3066)	(0.6095)	(0.8008)			(0.58814)	(0.0691)	(0.3549)	(0.0213)
Consumer Staples	0.0145	0.0496	1.0620	0.5144	0.8557	0.0017	0.0227	0.0744	6.3339**	5.3437*	5.8211**	5.3014**
			(0.5881)	(0.7732)	(0.3549)	(0.9667)			(0.0421)	(0.0691)	(0.0158)	(0.0213)
Health	0.0083	0.0393	0.2233	2.8220	0.1565	1.2660	0.0124	0.0579	0.4119	0.6917	0.2609	0.5992
			(0.8944)	(0.2439)	(0.6924)	(0.2605)			(0.8139)	(0.7076)	(0.6050)	(0.4389)
Telecommunications	0.0103	0.0641	0.1099	2.5270	0.0053	1.8540	0.0145	0.0682	1.0616	4.0278	0.8557	3.0397*
			(0.9465)	(0.2827)	(0.9420)	(0.1733)			(0.5881)	(0.1335)	(0.3549)	(0.0812)
Utilities	0.0124	0.0413	0.4119	0.8500	0.2609	0.8134	0.0124	0.0620	0.4119	1.3753	0.2609	1.3642
			(0.8139)	(0.6538)	(0.6095)	(0.3671)			(0.8139)	(0.5028)	(0.6095)	(0.2482)
Financial	0.0103	0.0496	0.1099	0.0380	0.0053	0.0017	0.0186	0.0682	3.2236	4.3447	2.8812*	3.0397*
			(0.9465)	(0.9812)	(0.9420)	(0.9667)			(0.1995)	(0.1139)	(0.0896)	(0.0812)
Technology	0.0083	0.0517	0.2233	0.1075	0.1565	0.0276	0.0145	0.0826	1.0616	9.1874***	0.8557	9.1514***
			(0.8944)	(0.9477)	(0.6924)	(0.8682)			(0.5881)	(0.0101)	(0.3549)	(0.0025)

Note. The α levels of 0.01 and 0.05 represent the confidence levels of the VaR series. The LR_{CC} and LR_{UC} correspond to the conditional and unconditional convergence Likelihood-Ratio statistics, respectively. A VaR violation is counted when $r_t < \text{VaR}_t$, as recorded in the indicator function $\xi(\cdot)$. Columns under $\mathbb{E}[r_{m,t} < \text{VaR}_{m,t}]$ indicate the proportion of violations; values near 0.01 and 0.05 hint at a good VaR model. The values in brackets represent p-values. Significance levels are denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

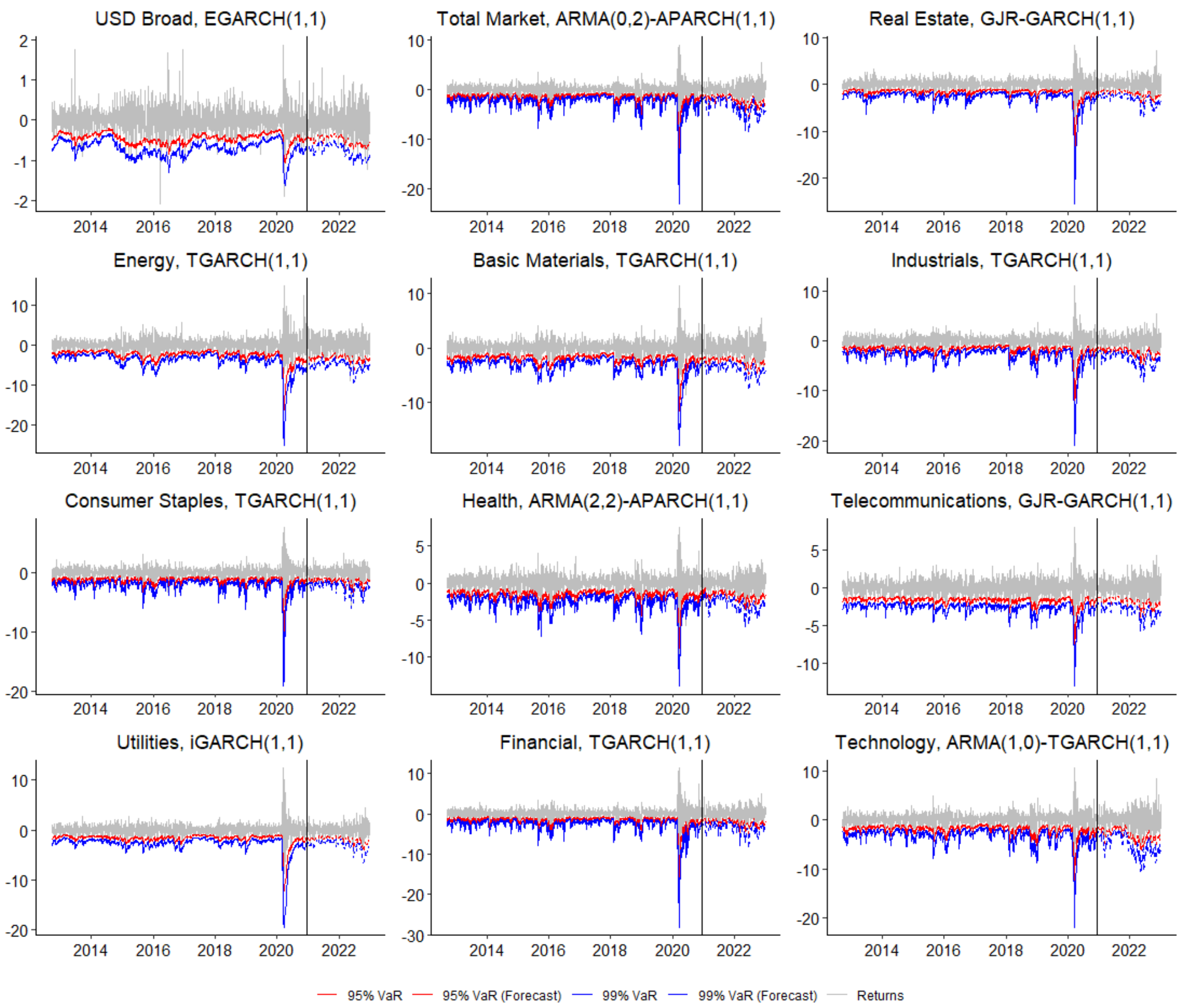

Figure 4: VaR Models and Forecasts for each Return series at $\alpha = 0.01$ and $\alpha = 0.05$

Note. The assortment of VaR models combines both the in-sample and out-of-sample periods outlined in the Data section. The transition from these periods is shown by the vertical line at observation 1936, which corresponds approximately to 2021.

5.3 Risk Spillover Results

To arrive at a conclusion on the central question, we now test for Granger causality in risk at $\alpha = 0.01$ and $\alpha = 0.05$ across our SGED-derived VaR models. Much of relevant literature studying cointegrating or multivariate relationships decomposes the spillover of price or volatility into short-term and long-term components. While the nature of our univariate models cannot explicitly capture these characteristics, we, similarly to Fan et al. (2008), attempt to reveal whether potential spillovers are lag-sensitive by conducting the test statistic for lag orders of $M = 1$ to $M = 40$, representing possible risk transmissions across 1 to 40 day increments.

These results are summarised in Table 6, however, Granger Causality in risk can be better understood in graphical form, as shown in Figures 5 and 6. As the test follows an asymptotic Normal Distribution, for a given M lag length, we can say that if the Q_i statistic is above the threshold of 1.645, there is risk transmission from one market into another significant at a 5% confidence level. In statistical terms, that conclusion entails the idea that if there is a large drop in index returns in one market, then it is useful for its prediction in another. Further, a Q greater in magnitude can be interpreted as representing a stronger cumulative spillover effect relative to a weaker one. Taking a closer look, we see that for several indices such as Total Market, Consumer Staples, Utilities, and Industrials, Technology, Q_1 and Q_2 statistics for both $\alpha = 0.01$ and $\alpha = 0.05$ have a tendency to start near, or at the 5% significance threshold and as the lag order increases, very rapidly decay in magnitude. On a base level, this seems to suggest that for most indices, there doesn't seem to be any significant one or two-way risk transmission across any reasonable lag length, and that of those that exhibit high spillovers, they very quickly fade after just a couple days after occurring in the respective origin market. Interestingly, this outcome appears, at least in some part, reflect and reinforce the pattern of cross-correlation shown in Figure 2, where we saw that autocorrelations across indices lose all statistical significance after just a couple lags. This relation is not a coincidence – while far from a substitute, the Granger Causality concept is partially derived from cross-correlations (Hong et al., 2009), making them invariably linked. The one outlier in these general remarks is the spillover to Energy, whose Q_1 and Q_2 statistics at $\alpha = 0.01$ rapidly increase to above the 99% confidence interval threshold and decay relatively more slowly, converging to around the 95% confidence interval threshold. Moreover, for many of the indices, Q_1 and Q_2 at $\alpha = 0.05$ frequently appears to be higher than $\alpha = 0.01$, suggesting that those risk spillovers are greater in strength. However, compared to the wider acdemic field, it can be considered rather surprising. Empirical studies point to codependencies in risk spillovers being higher when the downside is larger in absolute terms. Hong et al. (Hong et al., 2009) applies his test and VaR computed through CaViAR, method of directly modelling the tails of distributions, to the Yen/Dollar and Euro/Dollar series and found that across $M = 5$ to $M = 40$, the $\alpha = 0.01$ VaR series yielded greater Q statistics than $\alpha = 0.05$. While forgoing the Granger Causality framework, Bordoloi (2009) also studies financial contagion and ultimately finds that interdependence across markets, robust to a variety of different indices, is strongest "at times of extreme volatility", implying that a lower quantile such as 1% would carry with it a greater weight in this co-dependence.

Evaluating spillovers from Sector indices to USD Broad, we see a greater frequency of spillover effects crossing the 5% significance threshold, especially at $\alpha = 0.01$. Moreover, many such effects are longer in duration, falling to non-significance far less rapidly. For instance, at around 10-lagged days, spillovers from Total Market, Real Estate, Industrials, and Financial shoot to above either the 95% or 99% confidence intervals within the first 0-8 days after a VaR violation and usually do not decay significantly past the the 95% threshold. In this aspect, spillovers from Telecommnications are an exception as its effects fall well past the 95% threshold within the first couple days after the initial VaR exceedance.

Table 6: Summary of Granger Causality Test in Risk Results

H_0	Q_i							
	$M = 5$		$M = 10$		$M = 20$		$M = 40$	
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$
Broad $\not\Rightarrow$ Total Market	-1.0573	0.2923	-1.5771	0.2190	-1.2335	0.3872	-0.9537	0.3218
Broad $\not\Rightarrow$ Real Estate	-1.1896	0.4379	-1.7055	1.0375	-2.2655	1.4566*	-2.9749	1.7951**
Broad $\not\Rightarrow$ Energy	9.1691***	-0.5461	5.7108***	-0.0457	4.0692***	0.4478	3.1314***	0.0989
Broad $\not\Rightarrow$ Basic Materials	-0.9268	0.0657	-1.4643	-0.0304	-0.2198	0.3754	1.1275	0.2807
Broad $\not\Rightarrow$ Industrials	-1.0511	0.1285	-1.5730	-0.1070	-1.2872	0.9695	-1.0914	1.6262*
Broad $\not\Rightarrow$ Cons. Staples	-1.0600	-0.7435	-1.5642	-0.3109	-1.3184	0.1804	-1.1256	0.3195
Broad $\not\Rightarrow$ Health	-0.9949	0.0689	-1.5312	-0.0887	-0.6538	-0.1652	0.3499	-0.5109
Broad $\not\Rightarrow$ Telecom	-1.0348	-0.2939	-1.5573	-0.5382	-1.2486	-0.7341	-0.8391	-0.9372
Broad $\not\Rightarrow$ Utilities	-1.0290	-0.4544	-1.5496	0.1972	-1.3336	-0.1935	-1.1073	-1.1436
Broad $\not\Rightarrow$ Financial	-0.9252	0.1904	-1.3383	-0.2591	-0.5115	0.1663	1.0103	0.4989
Broad $\not\Rightarrow$ Technology	-1.2021	0.5982	-1.7272	0.4689	-2.0244	0.1091	-2.6494	-0.2998
Total Market $\not\Rightarrow$ Broad	1.2367	-0.9008	2.7819***	-0.7180	3.0696***	-0.6544	2.7051***	-1.0442
Real Estate $\not\Rightarrow$ Broad	1.2403	-0.7607	2.7901***	-0.8116	3.0927***	-0.9083	2.6976***	-0.7471
Energy $\not\Rightarrow$ Broad	-1.1873	-0.5061	-1.6840	-0.4099	-2.0203	-0.3832	-2.3730	-0.2111
Basic Materials $\not\Rightarrow$ Broad	-1.1887	0.5618	-1.6773	-0.0384	-2.1556	-0.6225	-2.5407	-0.6305
Industrials $\not\Rightarrow$ Broad	0.9888	-0.0381	2.2154**	-0.0929	1.7420**	-0.3650	1.6265*	-0.0780
Cons, Staples $\not\Rightarrow$ Broad	-1.0236	-0.0456	-1.5475	-0.0994	-1.5207	-0.5614	-0.6055	-0.6798
Health $\not\Rightarrow$ Broad	-1.0052	-0.1685	-1.5464	-0.2255	-0.8741	-0.2006	-0.1945	-0.0802
Telecom $\not\Rightarrow$ Broad	-1.1633	-0.8066	-1.6775	-0.7043	-2.3986	-0.7029	-2.4396	0.5001
Utilities $\not\Rightarrow$ Broad	0.8690	-0.1985	2.2437**	-0.5817	1.3155*	-0.7937	-0.3560	-1.1108
Financial $\not\Rightarrow$ Broad	-1.1405	4.7378***	-1.6426	4.6625***	-2.2348	3.1212***	-1.9573	2.7023***
Techology $\not\Rightarrow$ Broad	-1.1737	-1.0517	-1.7039	-1.0924	-0.4089	-1.4168	0.5939	-1.5538
Total Market $\not\Leftrightarrow$ Broad	-0.2579	-0.1026	0.5269	-0.1428	1.0646	-0.0522	1.0835	-0.4085
Real Estate $\not\Leftrightarrow$ Broad	-0.3340	-0.0247	0.4485	0.2648	0.3789	0.4562	-0.3229	0.7862
Energy $\not\Leftrightarrow$ Broad	4.3498***	-1.0104	2.3623***	-0.5721	1.2071	-0.1517	0.3936	-0.2172
Basic Materials $\not\Leftrightarrow$ Broad	-1.6267	0.0104	-2.3132	-0.3004	-1.8037	-0.3494	-1.1148	-0.3720
Industrials $\not\Leftrightarrow$ Broad	-0.3966	-0.3152	0.1635	-0.3914	0.1280	0.2253	0.2417	0.9410
Cons. Staples $\not\Leftrightarrow$ Broad	-1.5914	-0.2777	-2.2821	-0.1336	-2.1105	-0.1639	-1.3292	-0.1821
Health $\not\Leftrightarrow$ Broad	-1.5544	-0.0042	-2.2686	-0.1660	-1.2257	-0.2210	-0.0259	-0.3902
Telecom $\not\Leftrightarrow$ Broad	-1.6632	-1.0111	-2.3652	-1.0649	-2.6618	-1.1581	-2.4036	-0.4325
Utilities $\not\Leftrightarrow$ Broad	0.0312	0.0000	-0.2957	-0.5014	-0.4095	-0.7679	-0.2206	-0.8875
Financial $\not\Leftrightarrow$ Broad	-1.5850	3.1700***	-2.1998	3.0506***	-2.0495	2.3636***	-0.7872	2.3115**
Techology $\not\Leftrightarrow$ Broad	-1.7724	-0.5484	-2.4960	-0.6048	-1.8389	-1.0298	-1.5571	-1.3857

Note. The symbols $\not\Rightarrow$ and $\not\Leftrightarrow$ represent the null hypotheses of the Granger Causality in Risk test statistics Q_1 and Q_2, respectively. Where Q statistics are significant, this null can be rejected in favour for Granger-causality. The α levels of 0.01 and 0.05 correspond to the 99% and 95% confidence levels of the VaR indicator series being tested. Significance levels are denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

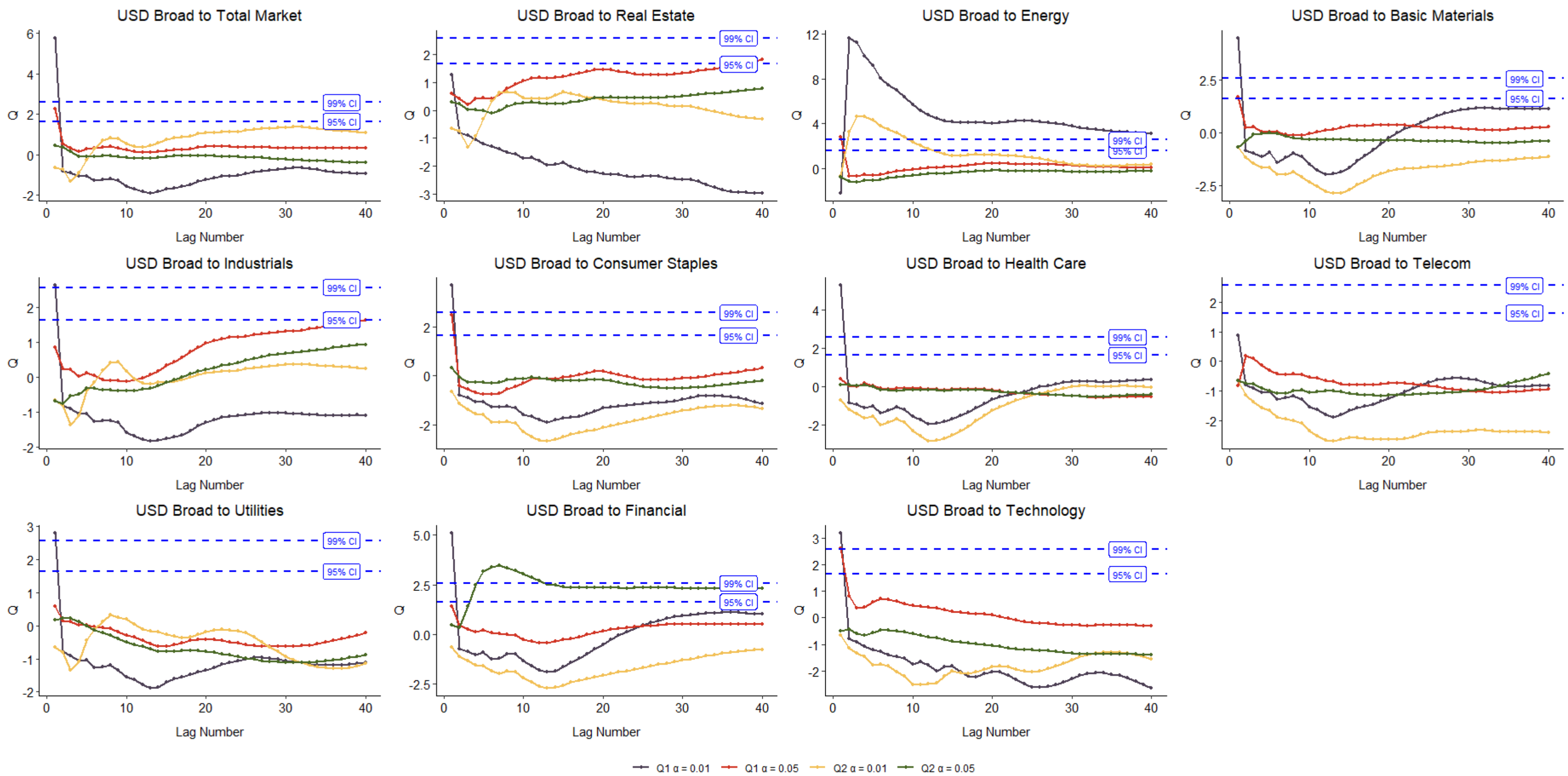

Figure 5: One and Two-Way Granger Causality Statistics from USD Broad to Sector Indices

Note. The series of graphs depict the one and two-way spillovers from the USD Broad to Sector indices using the Granger Causality in Risk statistic. The test is assessed over lag lengths from 1 till 40. In each plot are Q_1 and Q_2 statistics split into $\alpha = 0.01$ and $\alpha = 0.05$. The dashed horizontal blue lines are thresholds above which the risk spillover can be considered significant at their respective confidence level for that particular lag length M. The highest line corresponds to the critical value of $Q = 2.576$, a confidence level of 99%, and the lower one $Q = 1.645$, a confidence level of 95%.

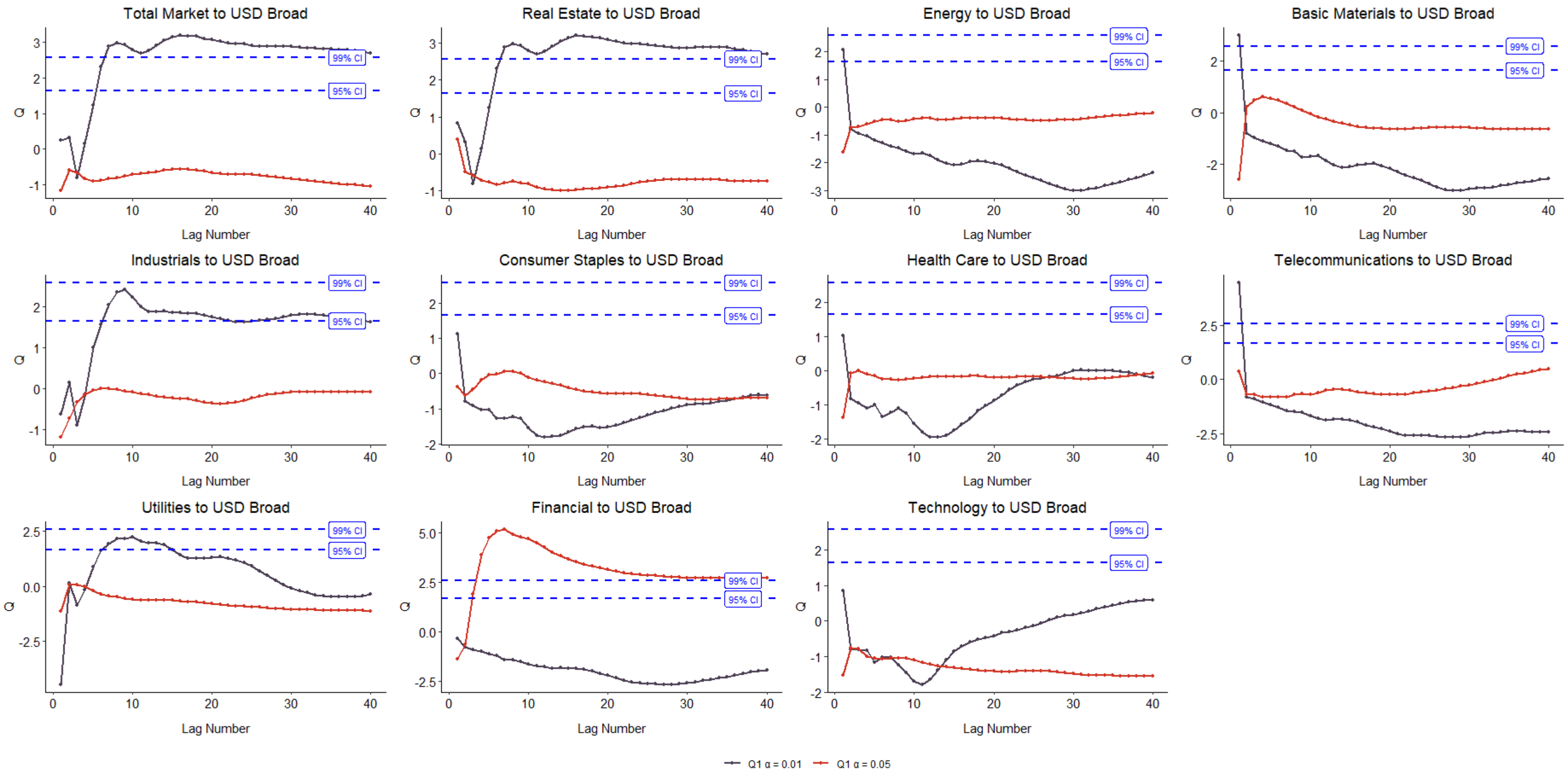

Figure 6: One Way Granger Causalities from Sector Indices to USD Broad

Note. The series of graphs depict the one way risk spillovers from each sector index to the USD Broad index assessed over lag lengths from 1 till 40. In each plot are Q_1 statistics split into $\alpha = 0.01$ and $\alpha = 0.05$. The dashed horizontal blue lines are thresholds above which the risk spillover can be considered significant at their respective confidence level for that particular lag length M. The highest line corresponds to the critical value of $Q = 2.576$, a confidence level of 99%, and the lower one $Q = 1.645$, a confidence level of 95%.

While the specific nature of these spillovers vary by the direction of Granger causality and the α level, we can derive two general conclusions. First, we were unable to establish the existence of consistent and systematic significant spillovers from either Broad to USD Broad or vice versa. Second, we find that there is a considerable diminishing effect, whereby spillovers high in strength tend to dissipate quickly. The former stands in agreement with the results of Tsagkanos and Siriopoulos (2013) who find no significant one-way long-run Granger Causality between the EUR/USD exchange rate and the US stock market. However, their Granger Causality tests do indicate short-run spillovers, especially when measured during their sample's "crisis" period, which encompasses the 2008 Financial meltdown. One key difference between our studies concerns the sample utilised. The use of two different periods capturing different variance dynamics aids in the robustness of Tsagkanos and Siriopoulos's findings, however, in these circumstances there is a case to be made for the superiority of the Granger Causality in Risk test over the standard Granger Causality test typically utilised. According to Hong et al. (2009), these test have one particular trait that makes them relatively unreliable indicators of spillovers, especially in the context of risk management. Due to the fact that they weigh all lags uniformly, they essentially imply that all past financial information is equally relevant in predicting spillovers in the current day. Considering the stylised financial fact that today's markets are dictated by current events than by remote distant ones (Hong et al., 2009), we arrive at a fallacy. Furthermore, studies of price spillovers do not necessarily translate well into those of risk or volatility due to the fundamental differences between these types of data that necessitate different methodologies. One such factor is the existence of nonstationarity and non-linear dependencies among price series, which often require wholly different approaches such as cointegration to isolate potentially-spurious elements from subsequent analyses. Returns series, which in our case is used to derive volatility[5], is typically considered stationary and as such, can be "treated" directly. Consequently, it may be more proper to compare our results with studies of volatility. Tastan (2006) utilises a series of Multivariate GARCH models to study the conditional and unconditional correlations between the DOW Jones, S&P500, and the USD/EUR exchange rate. He finds that while conditional correlations are significant both statistically and in magnitude, unconditional ones are not. The implication, then, is that there is evidence that the USD/EUR exchange rate moves strongly together with the US market indices in the relative short-term, but in the absence of a rigorously defined conditional variance model, such a relationship cannot be seen over a larger time horizon. As Tastan highlights, the interpretation of these models is rather difficult due to the non-trivial variance dynamics at play. It is also in this instance that the Granger test formulated by Hong et al. (2009) is advantageous; due to the fact that the test takes the result of VaR indicator functions as the input, there is a high degree of freedom in the methodology chosen to model said VaR. Moreover, lag-dependent hypothesis testing makes the interpretation of the Q_i statistics across the time dimension simple. Conceptually, this enables researchers to bypass the limitations enacted by the nonstationarity of price series, as faced by Tsagkanos and Siriopoulos (2013), while also overcoming difficulties in spillover interpretation, as faced by Tastan (2006).

Notwithstanding, it should be noted that the reliability of our results may be partially

[5]As outlined in the Theoretical Framework, we proxy volatility in GARCH models by taking the square of returns.

hindered due to the nature of our sample. As highlighted in the discussion regarding VaR forecasts, Abad et al. (2013) argue that the estimation of variance models using periods of relatively low volatility in order to forecast a period of relatively higher volatility has the potential to deteriorate the quality of said forecasts. Using a VaR-Granger-Causality-in-risk framework, Du et al. (2015) study the spillover of risk from equity to commodity markets, separating their sample into pre and post 2008 financial crisis. Ultimately, they find that the crisis leads to a structural change in the nature of risk transmission across these markets – compared to the pre-crisis period, the strength of the instantaneous spillover is greatly increased, indicating that a possible consequence of the crisis was a far greater temporary interdependence across markets. Extending these findings to our investigation, it is possible that a consequence of estimating our models using a 2011-2020 period to forecast 2020-2022 is that the changed behaviour of market participants following the SARS-CoV-2 pandemic induces a structural change in variance that our GARCH models cannot capture, leading to a bias in the spillover results. As such, for future research we recommend investigating possible structural breaks in the data such that samples used for estimation are consistent without extreme deviations in conditional variance levels. Further, the investigations of Hong et al. (2009), Chernozhukov & Umantsev (2001), Danielsson & de Vries (2000), and Engle & Manganelli (2004) provide compelling evidence that the tails of financial data may be governed by mechanisms different than those of the rest of the distribution. They argue that rather than modelling whole distributions as we have attempted with SGED, it may be more theoretically sound to focus only on the tails. To this end, Engle & Manganelli proposed the Conditional Autoregressive Value at Risk (CAViaR) model, which directly outputs VaR. Hong et al. (2009) in particular argues that such a model may be more accurate and generalised than a GARCH framework when facing unusually-shaped or time-varying error distributions. Hence, as a further refinement to our methodology, we propose a deeper dive into CAViaR, as it has the potential to eliminate the types of difficulties that we faced when, for instance, attempting to find a theoretical distribution that functioned well as both our VaR quantile distribution and GARCH error specification.

6 Conclusion

As this paper has shown, understanding and evaluating the connectedness of multiple markets is no trivial task. Prior literature has tackled this challenge through a variety of creative methods, yet, only a handful have garnered widespread adoption. Using the methodology outlined by Fan et al. (2008), and more specifically, the new innovative Granger Causality in Risk test formulated by Hong et al. (2009), this study aimed to uncover potential risk spillovers between sectors of the US economy and the trade-weighted USD exchange rate. To achieve this, we first fitted a series of ARMA-GARCH-type models to our indices in accordance with the Box-Jenkins methodology, observing the quality, optimality, and diagnostics related to their fit. It was found that conditional mean models were not frequently needed, indicating a lack of autocorrelation in our raw returns series. Overall, the models' AICs suggest that there are not significant differences in quality of fit across our range of models. Additionally, while sector indices exhibited similar conditional variance levels, USD Broad was comparatively very low. This was expected and signalled prior by the low kurtosis and skewness of its raw distribution. Extending Fan et al's (2008) procedures, we also check for the validity of our SGED error specification by conducting χ^2 goodness of fit tests. In line with expectations, they confirm that SGED is a comparatively better error specification than the commonly-used Student's t. With the validity of our models confirmed, we forecast and compute VaR under VCM for both $\alpha = 0.01$ and $\alpha = 0.05$, checking for the statistical adequacy of its fit. Kupiec's (1995) and Christoffersen's (1998) respective tests for Unconditional and Conditional Convergences confirm that all forecasted VaRs are valid and in-line with the theoretically-expected rates of violation. Finally, we compute the Granger Causality in Risk test statistics for Q_1 and Q_2 to determine the one-way and instantaneous spillovers. We take advantage of the ability to determine spillovers across any lag length by choosing a range from $M = 1$ to $M = 40$, covering all plausible effects.

Ultimately, we find minimal evidence for a systematic spillover, both one-way and two-way, of risk from either USD Broad to sector indices or vice versa. While there were several instances of spillovers significant at the 99% confidence level, the common tendency was for such spillovers to dissipate quickly over the next 5-10 days. Further, there was not a strong indication that spillovers were inherently different across $\alpha = 0.01$ or $\alpha = 0.05$ – this particular result is contrary to what is expected, because empirical research such as that of Abad & Bordoloi (2013) and Hong et al. (2009) argue that interdependence across financial markets should be stronger when losses greater in magnitude, i.e., at $\alpha = 0.01$. Compared against other research, we find our results corroborate those of Tsagkanos and Siriopoulos (2013), who determined that price spillovers are not significant in the long run. However, observing the short-run, we do not see major differences in the dynamics of these spillovers, while they do. Moreover, Tastan (2006) models volatility spillovers across the US/EUR exchange rate in US markets. Generally speaking, his results are the opposite of Tsagkanos and Siriopoulos, and subsequently, quite dissimilar to ours. Studying their methodologies, we find several advantages to spillover measurements based on the Granger Causality test formulated by Hong et al. (2009) over traditional methods. As the test does not require a regression model as an input, there is greater flexibility in using alternate VaR computation or estimation procedures. Moreover, the ease of the test's interpretation helps reduce uncertainties that may arise with comprehending complex multi-dimensional models.

These empirical findings have important implications for the wider field of risk management and academia. The advantages of the methodology presented benefit researchers and risk managers alike as its flexibility shows its merit as a general go-to way of quickly uncovering latent spillovers. As indicated by researchers by the likes of Engle and Manganelli (1982), for the purposes of understanding extreme risk, focusing on the tails of returns distributions is arguably more important than attempting to models entire distributions, and as such, we wish to bring more attention to methods incorporating VaR directly. While the topic will likely remain elusive with its numerous caveats, attempts at unraveling its mysteries bring us a step closer to mastering the concept of risk itself.

References

Abad, P., & Benito, S. (2013). A detailed comparison of value at risk estimates. *Mathematics and Computers in Simulation*, *94*, 258–276. https://doi.org/10.1016/j.matcom.2012.05.011

Adler, M., & Dumas, B. (1984). Exposure to currency risk: Definition and measurement [Publisher: [Financial Management Association International, Wiley]]. *Financial Management*, *13*(2), 41–50. https://doi.org/10.2307/3665446

Aggarwal, R. (1981). Exchange rates and stock prices: A study of the u.s. capital markets under floating exchange rates. *Akron Business and Economic Review*, *12*.

Alexios Ghalanos. (2022, October 26). Introduction to the rugarch package. https://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts [Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University]]. *International Economic Review*, *39*(4), 885–905. https://doi.org/10.2307/2527343

Ané, T. (2006). An analysis of the flexibility of asymmetric power GARCH models. *Computational Statistics & Data Analysis*, *51*(2), 1293–1311. https://doi.org/10.1016/j.csda.2005.11.002

Angelidis, T., Benos, A., & Degiannakis, S. (2007). A robust VaR model under different time periods and weighting schemes. *Review of Quantitative Finance and Accounting*, *28*(2), 187–201. https://doi.org/10.1007/s11156-006-0010-y

Bahmani-Oskooee, M., & Saha, S. (2015). On the relation between stock prices and exchange rates: A review article. *Journal of Economic Studies*, *42*(4), 707–732. https://doi.org/10.1108/JES-03-2015-0043

Bahmani-Oskooee, M., & Sohrabian, A. (1992). Stock prices and the effective exchange rate of the dollar [Publisher: Routledge]. *Applied Economics*, *24*(4), 459–464. https://doi.org/10.1080/00036849200000020

Billio, M., & Pelizzon, L. (2000). Value-at-risk: A multivariate switching regime approach. *Journal of Empirical Finance*, *7*(5), 531–554. https://doi.org/10.1016/S0927-5398(00)00022-0

Bisgaard, S., & Kulahci, M. (2009). Quality quandaries*: Time series model selection and parsimony [Publisher: Taylor & Francis]. *Quality Engineering*, *21*(3), 341–353. https://doi.org/10.1080/08982110903025197

Black, F., & Scholes, M. (1976). TAXES AND THE PRICING OF OPTIONS [Publisher: John Wiley & Sons, Ltd]. *The Journal of Finance*, *31*(2), 319–332. https://doi.org/10.1111/j.1540-6261.1976.tb01889.x

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Branson, W. (1983, June). *A model of exchange-rate determination with policy reaction: Evidence from monthly data* (w1135). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w1135

Brooks, C., & Persand, G. (2003). Volatility forecasting for risk management [Publisher: John Wiley & Sons, Ltd]. *Journal of Forecasting*, *22*(1), 1–22. https://doi.org/10.1002/for.841

Chernozhukov, V., & Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics*, *26*(1), 271–292. https://doi.org/10.1007/s001810000062

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using garch [Publisher: Wiley]. *Journal of Applied Econometrics*, *3*(4), 279–294. Retrieved June 6, 2023, from http://www.jstor.org/stable/2096644

Christoffersen, P. F. (1998). Evaluating interval forecasts [Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University]]. *International Economic Review*, *39*(4), 841–862. https://doi.org/10.2307/2527341

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*(2), 223–236. https://doi.org/10.1080/713665670

Danielsson, J., & De Vries, C. G. (2000). Value-at-risk and extreme returns [Publisher: [GENES, ADRES]]. *Annales d'Économie et de Statistique*, (60), 239–270. https://doi.org/10.2307/20076262

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. https://doi.org/10.1016/0927-5398(93)90006-D

Dornbusch, R., & Fischer, S. (1980a). Exchange rates and the current account [Publisher: American Economic Association]. *The American Economic Review*, *70*(5), 960–971. Retrieved March 30, 2023, from http://www.jstor.org/stable/1805775

Dornbusch, R., & Fischer, S. (1980b). Exchange rates and the current account [Publisher: American Economic Association]. *The American Economic Review*, *70*(5), 960–971. Retrieved March 30, 2023, from http://www.jstor.org/stable/1805775

Du, L., & He, Y. (2015). Extreme risk spillovers between crude oil and stock markets. *Energy Economics*, *51*, 455–465. https://doi.org/10.1016/j.eneco.2015.08.007

Dutta Bordoloi, S. (2009, October 28). *Interdependence of US industry sectors UsingVector autoregression* (Doctoral dissertation). Worcester Polytechnic Institute. https://digitalcommons.wpi.edu/etd-theses/1073brought%20to%20you%20by%20COREView%20metadata,%20citation%20and%20similar%20papers%20at%20core.ac.ukprovided%20by%20DigitalCommons@WPI

Engle, R. (2004). Risk and volatility: Econometric models and financial practice [Publisher: American Economic Association]. *The American Economic Review*, *94*(3), 405–420. Retrieved June 21, 2023, from http://www.jstor.org/stable/3592935

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *50*(4), 987–1007. https://doi.org/10.2307/1912773

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50. https://doi.org/10.1080/07474938608800095

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles [Publisher: [American Statistical Association, Taylor & Francis, Ltd.]]. *Journal of Business & Economic Statistics*, *22*(4), 367–381. Retrieved July 10, 2023, from http://www.jstor.org/stable/1392044

Erik Berggren. (2017, January). *Are GARCH models necessary for expected shortfall?* (Doctoral dissertation). Lund University. https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8902182&fileOId=8902880

Escanciano, J. C., & Lobato, I. N. (2009). An automatic portmanteau test for serial correlation. *Recent Advances in Time Series Analysis: A Volume Honouring Peter M. Robinson*, *151*(2), 140–149. https://doi.org/10.1016/j.jeconom.2009.03.001

Eugene F. Fama. (1991). Efficient capital markets: II [Publisher: [American Finance Association, Wiley]]. *The Journal of Finance*, *46*(5), 1575–1617. https://doi.org/10.2307/2328565

Fama, E. F. (1965). The behavior of stock-market prices [Publisher: University of Chicago Press]. *The Journal of Business*, *38*(1), 34–105. Retrieved July 1, 2023, from http://www.jstor.org/stable/2350752

Fan, Y., Zhang, Y.-J., Tsai, H.-T., & Wei, Y.-M. (2008). Estimating 'value at risk' of crude oil price and its spillover effect using the GED-GARCH approach. *Energy Economics*, *30*(6), 3156–3171. https://doi.org/10.1016/j.eneco.2008.04.002

Frankel, J. A. (1992, January 1). Monetary and portfolio-balance models of exchange rate determination. In J. M. Letiche (Ed.), *International economic policies and their theoretical foundations (second edition)* (pp. 793–832). Academic Press. https://doi.org/10.1016/B978-0-12-444281-8.50038-6

Frenkel, J. A. (1976). A monetary approach to the exchange rate: Doctrinal aspects and empirical evidence. *The Scandinavian Journal of Economics*, *78*(2), 200. https://doi.org/10.2307/3439924

Giovannini, A., & Jorion, P. (1987). Interest rates and risk premia in the stock market and in the foreign exchange market. *Journal of International Money and Finance*, *6*(1), 107–123. https://doi.org/10.1016/0261-5606(87)90016-7

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks [Publisher: [American Finance Association, Wiley]]. *The Journal of Finance*, *48*(5), 1779–1801. https://doi.org/10.2307/2329067

Glyn A, H. (2013). *Value-at-risk: Theory and practice, second edition*. https://www.value-at-risk.net

Guermat, C., & Harris, R. (2002). Forecasting value at risk allowing for time variation in the variance and kurtosis of portfolio returns. *International Journal of Forecasting*, *18*(3), 409–419. https://EconPapers.repec.org/RePEc:eee:intfor:v:18:y:2002:i:3:p:409-419

Hansen, P. (2001). *An unbiased and powerful test for superior predictive ability* [Issue: 2001-06]. Brown University, Department of Economics. https://EconPapers.repec.org/RePEc:bro:econwp:2001-06

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, *20*(7), 873–889. https://doi.org/10.1002/jae.800

Hartz, C., Mittnik, S., & Paolella, M. (2006). Accurate value-at-risk forecasting based on the normal-GARCH model. *Nonlinear Modelling and Financial Econometrics*, *51*(4), 2295–2312. https://doi.org/10.1016/j.csda.2006.09.017

Hong, Y., Liu, Y., & Wang, S. (2009). Granger causality in risk and detection of extreme risk spillover between financial markets. *Recent Development in Financial Econometrics*, *150*(2), 271–287. https://doi.org/10.1016/j.jeconom.2008.12.013

Jacobsen, B., & Dannenburg, D. (2003). Volatility clustering in monthly stock returns. *Journal of Empirical Finance*, *10*(4), 479–503. https://doi.org/10.1016/S0927-5398(02)00071-3

Jacque, L. L. (1981). Management of foreign exchange risk: A review article [Publisher: Palgrave Macmillan Journals]. *Journal of International Business Studies*, *12*(1), 81–101. Retrieved March 26, 2023, from http://www.jstor.org/stable/154420

Kollias, C., Mylonidis, N., & Paleologou, S.-M. (2012). The nexus between exchange rates and stock markets: Evidence from the euro-dollar rate and composite european stock indices using rolling analysis. *Journal of Economics and Finance*, *36*(1), 136–147. https://doi.org/10.1007/s12197-010-9129-8

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, *3*(2), 73–84. https://doi.org/10.3905/jod.1995.407942

Lee, C.-F., & Su, J.-B. (2012). Alternative statistical distributions for estimating value-at-risk: Theory and evidence. *Review of Quantitative Finance and Accounting*, *39*(3), 309–331. https://doi.org/10.1007/s11156-011-0256-x

Lu, Y., & AbouRizk, S. (2009). Automated box–jenkins forecasting modelling. *Automation in Construction*, *18*(5), 547–558. https://doi.org/10.1016/j.autcon.2008.11.007

Mcaleer, M., & Da Veiga, B. (2008). Forecasting value-at-risk with a parsimonious portfolio spillover GARCH (PS-GARCH) model. *Journal of Forecasting*, *27*(1), 1–19. https://doi.org/10.1002/for.1049

Mico, L., Autria Mazda, & Sarita Subramanian. (2005). *Indexes of the foreign exchange value of the dollar*. https://www.federalreserve.gov/pubs/bulletin/2005/winter05_index.pdf

Moeljadi, M., & Fauziah, F. (2015). Dynamic relationship between exchange rates and stock prices in asia on period 2009-2013. *Pressacademia*, *2*(1), 124–124. https://doi.org/10.17261/Pressacademia.201519783

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *59*(2), 347–370. https://doi.org/10.2307/2938260

Odening, M., & Hinrichs, J. (2002). Using extreme value theory to estimate value-at-risk [Publisher: MCB UP Ltd]. *Agricultural Finance Review*, *63*(1), 55–73. https://doi.org/10.1108/00215000380001141

Oral, C., & CenkAkkaya, G. (2015). Cash flow at risk: A tool for financial planning. *2nd GLOBAL CONFERENCE on BUSINESS, ECONOMICS, MANAGEMENT and TOURISM*, *23*, 262–266. https://doi.org/10.1016/S2212-5671(15)00358-5

Ratanapakorn, O., & Sharma, S. C. (2007). Dynamic analysis between the US stock returns and the macroeconomic variables. *Applied Financial Economics*, *17*(5), 369–377. https://doi.org/10.1080/09603100600638944

Reboredo, J. C., Rivera-Castro, M. A., & Ugolini, A. (2016). Downside and upside risk spillovers between exchange rates and stock prices. *Journal of Banking & Finance*, *62*, 76–96. https://doi.org/10.1016/j.jbankfin.2015.10.011

Siwen Zhou. (2018). *Exploring the driving forces of the bitcoin exchange rate dynamics: An EGARCH approach* (Doctoral dissertation). https://mpra.ub.uni-muenchen.de/89445/

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? [Publisher: American Accounting Association]. *The Accounting Review*, *71*(3), 289–315. Retrieved May 13, 2023, from http://www.jstor.org/stable/248290

So, M. K., & Yu, P. L. (2006). Empirical analysis of GARCH models in value at risk estimation. *Journal of International Financial Markets, Institutions and Money*, *16*(2), 180–197. https://doi.org/10.1016/j.intfin.2005.02.001

Soenen, L., & Hennigar, E. S. (1988). An analysis of exchange rates and stock prices: The u.s. experience between 1980 and 1986. *Akron Business and Economic Review*, *19*.

Sohail Chand, Shahid Kamal, & Imran Ali. (2012). Modeling and volatility analysis of share prices using ARCH and GARCH models. *World Applied Sciences Journal*. https://doi.org/10.5829/idosi.wasj.2012.19.01.793

Stefan Lundbergh & Timo Terasvirta. (1999). Evaluating GARCH models. https://papers.tinbergen.nl/99008.pdf

Su, J.-B. (2015). Value-at-risk estimates of the stock indices in developed and emerging markets including the spillover effects of currency market. *Economic Modelling*, *46*, 204–224. https://doi.org/10.1016/j.econmod.2014.12.022

Tastan, H. (2006). Estimating time-varying conditional correlations between stock and foreign exchange markets. *Physica A: Statistical Mechanics and its Applications*, *360*(2), 445–458. https://doi.org/10.1016/j.physa.2005.06.062

Taylor, S. (1986). *Modelling financial time series*. Wiley.

Taylor, S. J. (2007, December). *Modelling financial time series* (2nd ed.). WORLD SCIENTIFIC. https://doi.org/10.1142/6578

Tsagkanos, A., & Siriopoulos, C. (2013). A long-run relationship between stock price index and exchange rate: A structural nonparametric cointegrating regression approach. *Journal of International Financial Markets, Institutions and Money*, *25*, 106–118. https://doi.org/10.1016/j.intfin.2013.01.008

Tsai, I.-C. (2012). The relationship between stock price index and exchange rate in asian markets: A quantile regression approach. *Journal of International Financial Markets, Institutions and Money*, *22*(3), 609–621. https://doi.org/10.1016/j.intfin.2012.04.005

van den Goorbergh, R., & Vlaar, P. (1999, March). *Value-at-risk analysis of stock returns: Historical simulation, variance techniques or tail index estimation?* De Nederlandsche Bank.

Vlaar, P., & Palm, F. (1993). The message in weekly exchange rates in the european monetary system: Mean reversion, conditional heteroscedasticity, and jumps. *Journal of Business & Economic Statistics*, *11*(3), 351–60. https://EconPapers.repec.org/RePEc:bes:jnlbes:v:11:y:1993:i:3:p:351-60

Wang, G.-J., Xie, C., Wen, D., & Zhao, L. (2019). When bitcoin meets economic policy uncertainty (EPU): Measuring risk spillover effect from EPU to bitcoin. *Finance Research Letters*, *31*, S1544612318305749. https://doi.org/10.1016/j.frl.2018.12.028

White, H. (2000). A reality check for data snooping [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *68*(5), 1097–1126. Retrieved June 16, 2023, from http://www.jstor.org/stable/2999444

Xie, Z., Chen, S.-W., & Wu, A.-C. (2020). The foreign exchange and stock market nexus: New international evidence. *International Review of Economics & Finance*, *67*, 240–266. https://doi.org/10.1016/j.iref.2020.01.001

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*(5), 931–955. https://EconPapers.repec.org/RePEc:eee:dyncon:v:18:y:1994:i:5:p:931-955

Zongwu Cai, Jiancheng Jiang, & Jingshuang Zhang. (2010). A new test for superior predictive ability. https://core.ac.uk/download/pdf/41362936.pdf

A ADF Test Results

Table A1: Augmented Dickey-Fuller Tests for Stationarity of Returns Series

	Augmented Dickey Fuller Test Statistic	
	With Trend and Intercept	**With Intercept**
U.S. Broad Index Returns	-33.5159***	-33.5151***
Total Market Index Returns	-34.0568***	-34.0606***
Real Estate Index Returns	-31.8518***	-31.8551***
Energy Index Returns	-32.5154***	-32.5035***
Basic Materials Index Returns	-33.0400***	-33.0452***
Industrials Index Returns	-32.8389***	-32.8428***
Consumer Staples Index Returns	-33.6658***	-33.6728***
Health Care Index Returns	-34.0703***	-34.0727***
Telecommunications Index Returns	-34.5989***	-34.6027***
Utilities Index Returns	-33.4168***	-33.4221***
Financial Index Returns	-33.3034***	-33.3098***
Technology Index Returns	-36.7039***	-36.7094***

Note. The 5% critical values for the combined trend and intercept and solely intercept are $t_{DF} = -3.41$ and $t_{DF} = -2.86$, respectively. Significance levels are denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B Detailed Mean and Variance Models

Table A2: Results of Mean and Variance Model Selection Process

	ARMA(0,0) No Mean Model		ARMA(1,0)		ARMA(0,2)		ARMA(2,2)		GARCH(1,1) ARMA(P,Q)			GJR-GARCH(1,1) ARMA(P,Q)			TGARCH(1,1) ARMA(P,Q)			APARCH(1,1) ARMA(P,Q)			EGARCH(1,1) ARMA(P,Q)			iGARCH(1,1) ARMA(P,Q)		
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)†	(3)	(1)	(2)†	(3)	(1)	(2)†	(3)	(1)	(2)†	(3)	(1)	(2)†	(3)	(1)	(2)†	(3)
USD Broad	1.89	587***	-	-	-	-	-	-	1.89	0.90	0.262	1.89	0.81	0.2618	1.89	2.30	0.2579	1.89	1.55	0.2580	**1.89**	**1.63**	**0.2574**	1.89	0.84	0.2619
Total Market	4.29**	185***	-	-	0.00	267***	-	-	0.69	0.14	2.3141	1.18	2.77	2.2911	1.78	2.33	2.2772	**1.95**	**2.36**	**2.2772**	1.30	1.89	2.2882	0.66	2.10	2.3158
Real Estate	1.16	335***	-	-	-	-	-	-	1.16	3.71	2.6513	**1.16**	**3.39**	**2.6442**	1.16	6.52	2.6460	1.16	4.37	2.6443	1.16	6.28	2.6498	1.16	3.26	2.6556
Energy	1.05	361***	-	-	-	-	-	-	1.05	2.93	3.4225	1.05	2.22	3.4123	**1.05**	**3.50**	**3.4064**	1.05	3.68	3.4074	1.05	3.27	3.4105	1.05	2.75	3.4223
Basic Materials	1.75	246***	-	-	-	-	-	-	1.75	0.13	3.0020	1.75	0.39	2.9823	**1.75**	**0.22**	**2.9701**	1.75	0.25	2.9709	1.75	0.34	2.9744	1.75	0.36	3.0049
Industrials	2.69	183***	-	-	-	-	-	-	2.69	1.46	2.6278	2.69	2.73	2.6101	**2.69**	**3.65**	**2.6034**	2.69	3.37	2.6041	2.69	2.83	2.6098	2.69	1.56	2.6309
Consumer Staples	2.31	122***	-	-	-	-	-	-	2.31	3.24	2.1907	2.31	2.42	2.1711	**2.31**	**4.72**	**2.1633**	2.31	4.65	2.1644	2.31	3.26	2.1694	2.31	3.69	2.1945
Health Care	4.41**	189***	-	-	-	-	0.28	235***	0.44	0.61	2.6411	0.60	0.42	2.6285	1.23	2.45	2.6184	**1.60**	**5.52**	**2.6180**	0.97	2.14	2.6246	0.45	0.20	2.6444
Telecom	2.51	198***	-	-	-	-	-	-	2.51	0.97	2.7473	**2.51**	**0.81**	**2.7441**	2.51	1.02	2.7467	2.51	0.78	2.7448	2.51	1.01	2.7497	2.51	1.35	2.7567
Utilities	1.70	157***	-	-	-	-	-	-	1.70	8.20**	2.5575	1.70	8.27**	2.5561	1.70	12.63***	2.5598	1.70	9.26**	2.5568	1.70	11.72***	2.5615	**1.70**	**6.49**	**2.5625**
Financial	3.34*	213***	-	-	-	-	-	-	3.34*	1.75	2.6223	3.34*	3.40	2.5976	**3.34***	**2.36**	**2.5884**	3.34*	2.37	2.5894	3.34*	1.76	2.5978	3.34*	3.91	2.6251
Technology	5.79**	230***	0.01	276***	-	-	-	-	1.53	1.03	2.8956	1.82	1.67	2.8806	**2.33**	**1.16**	**2.8752**	2.25	1.28	2.8760	2.00	0.94	2.8806	1.41	2.18	2.8985

Note. Column (1) represents the PQ_A statistic outlined in the methodology. Column (2) denotes the ARCH-LM test statistic and column (3) signifies the AIC value. The ARCH-LM test is assessed at a lag order of 12 to capture a large portion of potential clustering. The statistics in column (2)† are the Weighted-ARCH LM test since certain shortcomings of rugarch R package do not allow for the normal ARCH-LM test to be conducted with GARCH-type models. However, for the purposes of evaluating remaining ARCH effects, it should not present a threat to the results' validity. The models that have been determined to be optimal for variance modelling based on these preliminary tests are highlighted in yellow. Significance levels are denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.