Lappeenranta University of Technology

LUT School of Business

Strategic Finance and Business Analytics (MSF)

Aleksi Lehtonen

VARIANCE REDUCTION PROPERTIES OF DYNAMIC HEDGING MODELS – EMPIRICAL EVIDENCE FROM BRICS COUNTRIES

Supervisor Professor Eero Pätäri

Examiner Post-doctoral researcher Elena Fedorova

ABSTRACT

Author	Aleksi Lehtonen
Title	Variance reduction properties of dynamic hedging models – empirical evidence from BRICS countries
Faculty	School of Business and Management
Major	Strategic Finance and Business Analytics
Year	2015
Master's Thesis	Lappeenranta University of Technology
	70 pages, 13 figures and 7 tables
Supervisor	Professor Eero Pätäri
The 2nd examiner	Post-doctoral researcher Elena Fedorova
Key words	Dynamic hedging, volatility, CCC GARCH, DCC GARCH, BRICS

This Master's Thesis analyses the effectiveness of different hedging models on BRICS (Brazil, Russia, India, China, and South Africa) countries. Hedging performance is examined by comparing two different dynamic hedging models to conventional OLS regression based model. The dynamic hedging models being employed are Constant Conditional Correlation (CCC) GARCH(1,1) and Dynamic Conditional Correlation (DCC) GARCH(1,1) with Student's *t*-distribution. In order to capture the period of both Great Moderation and the latest financial crisis, the sample period extends from 2003 to 2014. To determine whether dynamic models outperform the conventional one, the reduction of portfolio variance for in-sample data with contemporaneous hedge ratios is first determined and then the holding period of the portfolios is extended to one and two days. In addition, the accuracy of hedge ratio forecasts is examined on the basis of out-of-sample variance reduction. The results are mixed and suggest that dynamic hedging models may not provide enough benefits to justify harder estimation and daily portfolio adjustment. In this sense, the results are consistent with the existing literature.

Acknowledgments

Writing this Master's Thesis has been a great challenge but at the same time rewarding process. I would not have manage the task by myself so I want to thank some amazing and influential people who have helped me and/or endured my ramblings as well as contributed greatly during the process.

First, I would like to thank Professor Eero Pätäri for excellent comments and instructions. Special thanks goes to Nina Oker-Blom from Finnish Financial Supervisory Authority who helped me with the data acquisition. I also would like to thank all my co-workers for their interest and support.

I would also like express my gratitude to my parents for their endless patience and all the support they have given to me. I owe you everything! Before I finish, a bow to Jaakko and Joonas who had to consume way too much coffee during the endless hours discussing about studies, finance and life in general. Cheers!

In Espoo, December 2015

Aleksi Lehtonen

Table of contents

1 INTRODUCTION 5
1.1 Background 5
1.2 Object and research problems 6
1.3 Structure of the Master's Thesis 7
2 THEORETICAL FRAMEWORK 9
2.1 Optimal hedge ratio and futures 9
2.2 Conventional model 14
2.3 GARCH models 15
2.3.1 Univariate GARCH models 15
2.3.2 Multivariate GARCH models 21
2.3.3 Copula-based GARCH models 26
2.4 Previous studies 30
3 DATA AND METHODOLOGY 33
3.1 Data properties 34
3.2 Methodology 36
3.2.1 Conventional model and risk minimising hedge ratio 36
3.2.2 GARCH models and time-varying hedge ratio 37
4 EMPIRICAL RESULTS 40
4.1 Estimation of the parameters 40
4.2 Hedge ratios on different markets 44
4.3 Effectiveness of in-sample hedging 54
4.4 Effectiveness of out-of-sample hedging 57
5 CONCLUSIONS 60
REFERENCES 63

1 INTRODUCTION

1.1 Background

Information technology together with financial innovations (internet and ETFs to give simple examples) have lowered transaction cost and opened more stock markets to retail investors and enhanced yield possibilities but also introduced greater risks (e.g., political, macro, liquidity and market risks). To be able to reduce these risks, investors can use hedging techniques for minimising the variance of the portfolio and thus create more optimal portfolios. As the want to minimise portfolio variances arise from the fundamental idea of the portfolio theory, derivatives are valuable tool for the creation of optimal portfolios. Also, as asset management has become a global business, the ability to hedge international assets is essential for both portfolio managers and retail investors.

Studies of the different hedging models go a long way and it has been shown that dynamic models dominate over static ones (see, e.g., Baillie & Myers 1991; Kroner & Sultan 1991, 1993; Brooks & Chong 2001 and Choudhry 2003, 2004). Hence, the study and development of these models is vital for the progress of greater financial stabilisation and risk management.[1] Previous studies have concentrated mostly on GARCH based models (see, e.g., studies above as well as Bollerslev 1990; Engle 2002 and Engle & Sheppard 2001; Hsu et al. 2008) because GARCH models are able to incorporate stylised facts of asset returns, such as leptokurtosis, skewness and non-Gaussian distributions shown in Jarque-Bera tests. However, most of the studies have concentrated either on currency or commodities hedging or in case of stock markets on developed markets (such as UK, US, Germany and Japan). This is most likely due the lack of data from emerging stock markets and their insignificance to investors in

[1] See Steinherr (2000) for detailed discussion how derivatives can enhance financial stability.

developed countries due the earlier access restrictions and high transaction costs.

However, it is also important to test hedging theories with emerging markets to see whether the models work in not perfectly liquid and efficient markets.[2] Especially, investors in early years of 21st century wanted better returns than stock or bond markets in developed countries yielded (lost decade due the IT bubble in 2000 as well as the latest financial crisis and its aftermath) and their focus was more transformed to emerging and especially to BRICS (Brazil, Russia, India, China and South Africa) countries. Though, with a higher potential returns the risks are also higher and most investors would benefit from the ways to decrease the portfolio variance. With different hedging models (especially with futures) investors can hedge these risks and benefit from BRICS countries with the lower variance of returns.

1.2 Object and research problems

This Master's Thesis studies different dynamic hedging models and their hedging abilities. Especially of interest is whether dynamic models can outperform conventional one when tested with BRICS market data. From these insights rises this Master's Thesis' main research problem and its sub-problem. Research problem of the Master's Thesis can be stated as follows

> *Could portfolio variance be minimised with dynamic hedging models compared to conventional hedging model on BRICS countries in period 2003-2014?*

Sub-problem regarding dynamic models can be stated as follows

> *Can dynamic conditional correlation GARCH model outperform constant conditional correlation GARCH model and thus*

[2] See e.g. Kim & Singal 2000; Omran & Farrar 2006 and Lim & Brooks 2011 for discussion about efficient market hypothesis (EMH) in emerging markets

capture persistent volatility between futures and underlying indices on BRICS countries in period 2003-2014?

To make this Master's Thesis doable some delimitations are needed. Firstly, from various dynamic hedging models two is selected for testing. First tested model is a constant conditional correlation (CCC) GARCH model, first introduced by Bollerslev (1990). However, instead of Bollerslev (1990) model we use econometric model from Kroner & Sultan (1993). Second tested model is a dynamic conditional correlation (DCC) GARCH from Engle & Sheppard (2001) and Engle (2002). These models are compared to conventional model introduced by Hsu, Tseng & Wang (2008) and Kroner & Sultan (1993). Using these models it is possible to compare hedging abilities of different methods.

Second delimitations concerns the used data. For the empirical part, data is obtained from Bloomberg Terminal. This Thesis uses data from BRICS countries which are selected as the hedging with derivatives is less studied compared to developed markets (for studies of developed markets, in addition ones mentioned above see e.g., Myers 1991; Garcia, Roh & Leuthold 1995; Park & Switzer 1995; Tong 1996; Butterworth & Holmes 2000; Brailsford, Corrigan & Heaney 2001; Lien & Tse 2002; Floros & Vougas 2004; Ku, Chen & Chen 2007; Sultan & Hassan 2008; Park & Jei 2010 and a survey of emerging derivatives markets by Lien & Zhang 2008). Time-series for testing is selected to be from 2003-2014 as the time span covers both Great Moderation[3] and financial crisis of 2007-2008.

1.3 Structure of the Master's Thesis

Rest of the Master's Thesis is organised as follows; Section 2 presents the theoretical framework of the dynamic hedging as well as presents previous studies in the field of dynamic hedging. Section 3 discusses the data and

[3] See, e.g. Bernanke 2004; Kim & Nelson 2004 and Stock & Watson 2002.

its properties as well as presents used models and methodologies for applying models and their evaluation. Section 4 presents the empirical results. Lastly, section 5 concludes and gives suggestions for further research.

2 THEORETICAL FRAMEWORK

Optimal hedging strategy, as stated above, is utmost important in the field of risk and portfolio management. This sections presents the development of hedging models from static ones to the new dynamic GARCH based models. Nevertheless, first it is needed to derive optimal hedge ratio and show why it is also risk-minimising hedge ratio. Additionally, basic fundamentals of the futures are presented before moving to hedging models. Finally, some previous studies in the field of dynamic hedging is introduced.

2.1 Optimal hedge ratio and futures

To be able to understand hedging models, one must first understand why hedge ratio is a risk-minimising. For deriving optimal hedge ratio we follow presentation from Kroner & Sultan (1993, 536-538). Optimal hedge ratio can be shown by assuming that investor has a fixed long position of one unit in the spot markets and short position of *–b* units in the futures markets. Thus, the random return of this portfolio, *x*, is

$$x = s - bf, \tag{1}$$

where s and f are changes in the spot and futures prices, respectively. If we also assume that investor faces mean-variance expected utility function,

$$EU(x) = E(x) - \gamma\sigma^2(x), \tag{2}$$

where γ is the degree of the risk aversion $(\gamma > 0)$, then investor by maximising her expected utility,

$$\max_b EU(x) = \max_b \{E(s) + bE(f) - [\sigma_s^2 + b^2\sigma_f^2 - 2b\sigma_{sf}]\}, \tag{3}$$

for *b* can find the optimal number of futures contracts in the portfolio,

$$b^* = \frac{E(f) + 2\gamma\sigma_{sf}}{2\gamma\sigma_f^2}. \quad (4)$$

If the futures rate follows a martingale (i.e. discrete-time stochastic process) Equation 4 reduces to[4]

$$b^* = \frac{\sigma_{sf}}{\sigma_f^2}. \quad (5)$$

This, however, is a hedge ratio for a single period of time and thus not usable for intertemporal hedging. But if the joint distribution of spot and futures is constant over time, it can be extended to multi-period hedging strategy by assuming a time-separable utility function. With constant joint distribution the sequence of hedge ratios $\{b_1, \dots, b_t\}$ gives $b_i = b_j \; \forall \; i, j$ (i.e. hedge ratio is constant) and the hedge ratio is often calculated as the least square estimator from a time-series regression of ΔS_t and ΔF_t (i.e., conventional hedging model presented below).

However, with arrival of new information, the joint distribution of these assets could be time-varying and then static ratio may not be adequate for an intertemporal hedging strategy (Kroner & Sultan 1993, 537 and Hsu et al. 2008, 1098). It is possible to modify the above presentation to consider a new information by letting s_t and f_t to be changes in the price of the spot and futures between time t' and t, respectively, and define $-b_{t'}$ to be the short position in futures at time t'. Then, it is possible to write a time-varying random return of the portfolio, x_t, as

$$x_t = s_t - b_{t'} f_t, \qquad t' < t, \quad (6)$$

[4] For more information about martingales, please see e.g., Ville 1939; Williams 1991; June 2009 issue of Electronic Journal for History of Probability and Statistics or Introduction to Econophysics (2007) with explanation why stock market returns appears to be random with information theory and algorithmic complex theory.

which is the payoff at time *t* to purchasing one unit of the spot and going short in $b_{t'}$ units of the futures at time t' in the past. Again, it is possible to write the expected utility function for the investor,

$$E_t U(x_{t+1}) = E_t(x_{t+1}) - \gamma \sigma_t^2(x_{t+1}), \tag{7}$$

where, risk is now measured with conditional variances and expectations and variances are calculated conditional on information available at time *t*. Following the same procedure than previously the utility-maximising hedge ratio at time *t* is

$$b_t^* = \frac{E_t(f_{t+1}) + 2\gamma \sigma_t(s_{t+1}, f_{t+1})}{2\gamma \sigma_t^2(f_{t+1})}, \tag{8}$$

and again assuming that futures prices follows martingale process, this simplifies to

$$b_t^* = \frac{\sigma_t(s_{t+1}, f_{t+1})}{\sigma_t^2(f_{t+1})}. \tag{9}$$

It is easy to see that the Equation 9 above is similar to the conventional hedge ratio except that the time-varying conditional moments replaces the static unconditional moments and the risk-minimising hedge ratio will change when a new information is obtained. Also, this conditional model will reduce to the conventional if the joint distribution of the assets remain constant over time.

Thus far, utility-maximizing hedge ratio has been proposed and models to estimate it will be presented below but before that some notions about futures should be presented. To understand the dynamic hedging with futures one needs to understand what future contracts are and how they are priced. The basic concepts of the futures and their value is derived to fully

appreciate different hedging strategies with them. As forward contracts are not *marked-to-market* daily their general results are derived first and then we consider how the futures deviate from the forwards.

We follow presentation of Munk (2013, 299-302), though there is also other ways to present the futures. If an investor wants to buy an asset at time *T*, one can lock in the effective purchasing price beforehand with a long position of forward on a given asset. Vice versa, if investor wants to sell an asset at time *T* she can take a short position of the forward and lock in the effective selling price. Using a risk-neutral probability measure $\mathbb{Q}$, the value of such a future payoff at time *t* can be written as[5]

$$\begin{aligned} V_t &= E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1}(P_T - K)\right] \\ &= E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1}P_T\right] - KE_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1}\right] \\ &= E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1}P_T\right] - KB_t^T, \end{aligned} \tag{10}$$

where *T* is maturity date, *K* is delivery price, *P* is underlying variable and $P_T - K$ provides a dividend at time *T*. Additionally, $B_t^T = E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1}\right]$ is the price of the zero-coupon bond maturing at time *T* with a unit payment. $R_{t,T}^f$ is the gross return of roll over in short risk-free investment between time *t* and *T*.

At time *t*, the delivery price *K* is usually set so that the value of a forward contract is zero. This value of *K* can be called the forward price at time *t* for the delivery date *T* and is denoted F_t^T. Also, terminal forward price needs to be $F_T^T = P_T$ (the only reasonable price for immediately delivery). Thus, setting $V_t = 0$ and the forward price is given by

[5] A risk-neutral probability measure indicates that all assets have an expected return equal to the risk-free return and the discount rate used for the expected payoff on a future is the risk-free rate. Thus, the risk-neutral probability measure simplifies derivatives pricing. In addition, it can be shown that derivatives prices are correct, when using risk-neutral probabilities, in both risk-neutral world and non-risk-neutral world $\mathbb{P}$. For more information about the risk-neutrality see e.g., Munk 2013 or Hull 2012.

$$F_t^T = \frac{E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1} P_T\right]}{B_t^T}. \tag{11}$$

Additional to the Equation 11 above, the forward price can be characterised in terms of the forward measure for maturity *T*. Then, the forward price process for contracts with delivery date T is a $\mathbb{Q}^T$-martingale. It is possible to show that F_t^t is a $\mathbb{Q}^T$-martingale[6] and thus F_t^t can be expressed

$$F_t^T = \frac{E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1} P_T\right]}{B_t^T} = E_t^{\mathbb{Q}^T}[P_T]. \tag{12}$$

Above expression assumes payments in the period [*t,T*]. However, if the underlying variable is the price of traded asset without payments in given period, it can be simplified as

$$E_t^{\mathbb{Q}}\left[\left(R_{t,T}^f\right)^{-1} P_T\right] = P_t, \tag{13}$$

and the forward price can be written as $F_t^T = {P_t}/{B_t^T}$.

After deriving general forward price it is possible to consider the futures markets and specifically how the futures price deviates from the forward prices. First, let Φ_t^T be the futures price at time *t*. As with forwards, the futures price at the delivery time is by definition equal to the price of underlying asset, $\Phi_T^T = P_T$. Between the last settlement to the market and the maturity, futures are indistinguishable from forwards and thus their values need to be identical. Before the last settlement date it is possible to write that in the discrete-time framework the dividend from the futures at time *t+1* is $\Phi_{t+1}^T - \Phi_t^T$ and in the continuous-time setting $d\Phi_t^T$. With these no-

[6] See, e.g., Munk 2013 for more information

tions it is possible to write the following theorem presented by Munk (2013, 301)[7]

Theorem 1. *The futures price Φ_t^T is a martingale under the risk-neutral probability measure $\mathbb{Q}$. In particular,*

$$\Phi_t^T = E_t^{\mathbb{Q}}[P_T]. \tag{14}$$

After deriving utility-maximizing hedge ratio and presenting futures valuation, next two sections present the theoretical foundations of estimating optimal hedge ratio and how the conditional variance models have developed.

2.2 Conventional model

The different ways of obtaining the optimal hedge ratio have been studied for decades and earlier studies pointed to a static hedge ratio which was estimated by ordinary least-square techniques (see, e.g., Ederington 1979; Figlewski 1984; Lee, Bubnys & Lin 1987 and Benet 1992). However, with development of GARCH –model by Bollerslev (1986) studies moved to the time-varying hedging strategies that can take account heteroskedasticity. These models were showed to be superior compared to the static ones (see, e.g., Baillie & Myers 1991; Kroner & Sultan 1991, 1993; Brooks & Chong 2001 and Choudhry 2003, 2004 among others). While several different models to estimate the optimal time-varying hedge ratio has been developed, two has mostly been used. First model is presented by Bollerslev (1990) and most of the proceeding models are preceding to this one. One of these models is presented by Kroner & Sultan (1993) which shows how the static ratio can be modified to the time-varying. Another highly used model was introduced by Engle & Kroner (1995). Third model was introduced by Engle & Sheppard (2001) and Engle (2001).

[7] See Munk (2013, 301) for the proof.

Before GARCH models hedging was conducted with unconditional and non-time-varying models. Ederington (1979) presented a model to estimate the hedge ratio and it was used for several studies and applied to portfolio management. This ratio can be expressed with basic OLS regression as follows

$$C_{st} = \alpha + \beta_1 C_{f1t} + \beta_2 C_{f2t} + \cdots + \beta_n C_{fnt} + u_t, \quad (15)$$

where, C_{st} is the percentage change of the spot price during period *t* and C_{fnt} the price change of n^{th} futures contract during period *t*. Also, β_n is corresponding optimal hedge ratio and these are parameter estimates obtained from OLS regression. As can be seen, these hedge ratios are not time-varying and thus not take account heteroskedasticity. The Equation 9 showed that the optimal hedge ratio can be found by dividing the time-varying covariance with the futures' time-varying variance. Hence, the development of GARCH models, which are developed to study the time-varying covariances and variances as well as take account of heteroskedasticity are presented in the next section.

2.3 GARCH models

2.3.1 Univariate GARCH models

Engle (1982) proposed a process where the conditional past variances are not constant but unconditional variances are still constant. In these kind of processes, the recent past gives information about the forecast of the current variance. As the traditional econometric models were operating under assumption of the constant variance, the new development changed greatly the perception of the variance. These new autoregressive conditional heteroskedasticity (ARCH) models capture volatility clustering in asset returns as first noted by Mandelbrot (1963). In addition, this opened new lines of study for financial economists and especially for those interested in volatility of the stock markets and its relevance to forecasting the

asset returns. For an ARCH process, the conditional mean equation for univariate time series y_t and its error term can be specified as follows (Engle 1982, 987-989)

$$y_t = E(y_t|\Psi_{t-1}) + \varepsilon_t \qquad (16)$$
$$\varepsilon_t = z_t\sigma_t \quad z_t \sim N(0,1),$$

where Ψ_{t-1} is the information set at the time *t-1*, ε_t is the error term, z_t is a sequence of normally distributed random variables with zero mean and unit variance or white noise. Engle (1982, 988) proposed how to model a time-dependent standard deviation σ_t as follows

$$\varepsilon_t = z_t\sigma_t \quad z_t \sim N(0,1), \qquad (17)$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2,$$

where σ_t^2 is the estimate of conditional variance, α_0 is constant and greater than 0, α_i is ARCH term and ≥ 0 and $i > 0$. Because the coefficients must satisfy conditions for non-negativity, the above mentioned constrains have to be met. The model above is known as ARCH(q) because its lag length (q) is not restricted and the conditional variance depends q lags of its squared errors. However, even though it was first applicable framework to model the volatility it has a few undesirable properties. The most damaging property is a linear declining lag structure in the conditional variance equation which needs to be added to ensure a positivity of the coefficients as the totally free lag distribution usually leads to violation of the non-negativity constrains (Bollerslev 1986, 307-308).

Bollerslev (1986) introduced generalised ARCH (GARCH) model to overcome these restrictions and his model has been widely used ever since. The generalisation includes the previous own lags incorporated into the

conditional variance and a general case of GARCH(p,q) is specified as follows

$$\varepsilon_t = z_t\sigma_t \quad z_t \sim N(0,1), \tag{18}$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where α_0 is constant term, α_i is the ARCH term and β_j is the GARCH term. The non-negativity constrains are, in this case, $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$. It is easy to see that if *p* reduces to zero the process is ARCH(q) and if *p=q=0* it is simply a white noise. In addition, the condition of $\sum_{i=1}^{\max(p,q)}(\alpha_i + \beta_i) < 1$ is required for the model to be stationary.[8] (Bollerslev 1986, 309-310) Necessity of the former constrain is intuitive when remembered that the variance models need to be mean-reverting and in case of $\sum_{i=1}^{\max(p,q)}(\alpha_i + \beta_i) > 1$ (non-stationary GARCH) shocks would be persistent and convergence to the mean would not happen.

However, general lag structure *(p,q)* of the GARCH model makes the estimation laborious and hence Bollerslev (1986, 311) proposed simpler GARCH(1,1) process

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_j \sigma_{t-1}^2, \tag{19}$$
$$\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0.$$

With GARCH(1,1) $\alpha_1 + \beta_1 < 1$ ensures wide-sense stationarity which is much easier to accomplish. GARCH(1,1) also adequately captures the volatility clustering and thus higher order models are rarely needed (Brooks 2008, 391-394).

[8] See proof from Bollerslev 1986, Appendix A.1.

Several extensions for a standard GARCH model have been developed to overcome some of its inadequacies. Two such examples are its inability to take account of asymmetries and the artificial constrains that must be placed to ensure the non-negativity (Brooks 2008, 404). Next, we present two commonly-used models that overcome these problems and one not very commonly-used model which nests a variety of more popular symmetric and asymmetric GARCH models.

Firstly, to take account of asymmetries (studies indicating asymmetries of the volatility see e.g., Black 1976 and Engle & Ng 1993) GJR GARCH model by Glosten, Jagannathan & Runkle (1993) is presented. The model allows for asymmetric effects, more specifically leverage effects, which are captured by a dummy variable

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i + \gamma_i N_{t-1}) \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{20}$$

where N_{t-1} is a dummy variable following

$$N_{t-1} = \begin{cases} 1 \; if \; \varepsilon_{t-i} < 0 \\ 0 \; if \; \varepsilon_{t-i} > 0. \end{cases} \tag{21}$$

From the Equations 20 and 21 can be seen that a larger impact is caused by negative (negative ε_{t-i}^2 value) shocks ($(\alpha_i + \gamma_i)\varepsilon_{t-i}^2$) than positive (positive ε_{t-i}^2 value) shocks ($\alpha_i \varepsilon_{t-i}^2$). However, the conditions for non-negativity of the constant, ARCH term and GARCH term are similar than a standard GARCH model. Additionally, a constrain of $\alpha_i + \gamma_i \geq 0$ must be satisfied. The model is acceptable with $\gamma_i < 0$ but the leverage effect occurs provided that $\gamma_i > 0$. Hence, additional constrain compared to standard GARCH model is needed.

The next GARCH model does not require constrains for non-negativity and it also captures the variance asymmetries (Nelson 1991 and Engle & Ng 1993). This model, proposed by Nelson (1991) is nowadays called EGARCH and tries to avoid three limitations of basic GARCH model, i.e. the symmetry of the change of the variance, non-negativity constrains and interpretation of the persistence of shocks to conditional variance by taking a natural logarithm of the conditional variance and making $\ln(\sigma_t^2)$ a linear function of time and lagged z_t's (Nelson 1991, 350). More formally

$$ln(\sigma_t^2) = \alpha_t + \sum_{k=1}^{\infty} \beta_k q(z_{t-k}), \quad \beta_1 \equiv 1, \tag{22}$$

where α_t and β_k are real, nonstochastic, scalar sequences. Since $\ln(\sigma_t^2)$ can be negative, by definition, non-negativity constrains to parameters is not needed. To accommodate the asymmetric reaction to shocks the value of $q(z_t)$ must be a function of the magnitude and the sign of z_t. Nelson (1991, 351) made a function to be linear combination of z_t and $|z_t|$ which gives to conditional variance well-behaved moments

$$\mathrm{q}(z_t) \equiv \theta z_t + \gamma[|z_t| - E|z_t|]. \tag{23}$$

It can be seen that by construction $\{\mathrm{q}(z_t)\}_{t=-\infty,\infty}$ is a zero-mean, i.i.d. random sequence. The two components of $q(z_t)$ are θz_t and $\gamma[|z_t| - E|z_t|]$ each with zero mean and as Nelson (1991, 351) shows if z_t is symmetric, the two components are orthogonal. Additionally, $q(z_t)$ allows the conditional variance process to respond asymmetrically to rises and falls of asset prices. This is because over the range $0 < z_t < \infty$, $q(z_t)$ the process is linear in z_t with slope $\theta + \gamma$, and over the range $-\infty < z_t \leq 0$, $q(z_t)$ the process is linear in z_t with slope $\theta - \gamma$, indicating the sign effect. The term $\gamma[|z_t| - E|z_t|]$ represents the magnitude effect which can be confirmed by assuming that $\gamma > 0$ and $\theta = 0$. In this case, the innovation in the conditional variance is positive (negative) when the magnitude of z_t is larger (smaller) than its expected value. Now suppose that $\gamma = 0$ and $\theta < 0$. The

innovation in the conditional variance is now positive (negative) when returns innovations are negative (positive). Lastly, Nelson (1991, 351) presents conditions when $\ln(\sigma_t^2)$ is strictly stationary and ergodicity. $\mathrm{Ln}(\sigma_t^2)$ is a linear process and thus its stationarity and ergodicity are easily checked.[9]

Final univariate GARCH model presented is fGARCH by Hentschel (1995). It is an omnibus model which nests several other popular symmetric and asymmetric models. For derivation of the fGARCH it is first needed to construct an absolute value version of basic GARCH(1,1) model (Hentschel 1995, 75)

$$\sigma_t = \alpha_0 + \alpha_1 \sigma_{t-1} |\varepsilon_t| + \beta_1 \sigma_{t-1}. \tag{24}$$

As is apparent, this formulation is identical to GARCH(1,1) process[10], except that the conditional variance terms has been replaced with their square roots, and the squared error has been replaced with its absolute value. However, this representation is still symmetric and Hentschel (1995, 75) proposed a way to model asymmetry

$$\sigma_t = \alpha_0 + \alpha_1 \sigma_{t-1} f(\varepsilon_t) + \beta_1 \sigma_{t-1}, \tag{25}$$

where

$$f(\varepsilon_t) = |\varepsilon_t - b| - c(\varepsilon_t - b). \tag{26}$$

This asymmetric absolute value GARCH model is interesting on its own right but to be able to nest other model for an omnibus model Hentschel (1995, 77) proposed fGARCH where, by rewriting intercept in the Equation (25) and introducing the parameters λ and ν, for $\lambda = \nu = 1$, the Equation (25) can be written as

[9] For more information, see Nelson (1991), Theorem 2.1 and Appendix II for its proof.
[10] It should be noted that Hentschel used a slightly different notation where $\sigma_{t-1}|\varepsilon_t|$ in Equation 24 is a decomposition of error term in mean equation. We use more widespread notation in Equation 18 (i.e. $\varepsilon_t = z_t \sigma_t$)

$$\frac{\sigma_t^{\lambda} - 1}{\lambda} = \alpha_0' + \alpha_1 \sigma_{t-1}^{\lambda} f^{v}(\varepsilon_t) + \beta_1 \frac{\sigma_{t-1}^{\lambda} - 1}{\lambda}. \quad (27)$$

Equation (27) is a law of motion for Box & Cox (1964) transformation of the conditional standard deviation, where parameter λ determines the shape of the transformation.[11] After transformation σ_t is convex when $\lambda > 1$, while for $\lambda < 1$ it is concave. The parameter *v* transforms the absolute value function $f(\cdot)$. Hentschel (1995, 79-83) shows that fGARCH is able to nest most of the other popular models and show these to be merely a restricted versions of fGARCH. Some of the nested models are above presented GJR-GARCH and EGARCH as well as TGARCH by Zakoian (1994), absolute value GARCH from Taylor/Schwert (1986 and 1989), NAGARCH (Engle & Ng, 1993), NARCH and NGARCH by Higgins & Bera (1992). However, QGARCH by Sentana (1995) is excluded but Engle & Ng (1993) showed that there is little empirical difference between QGARCH and NAGARCH which can be nested. After representing some univariate models we expand these to multivariate models. As we have shown, univariate models can be used to derive the variance part of hedging models but multivariate versions are needed to estimate covariances between the spot and futures markets.

2.3.2 Multivariate GARCH models

In case of the multivariate GARCH (MGARCH) models a field can be divided roughly to half (Bauwens, Laurent & Rombouts 2006, 79). The most obvious application for MGARCH models is the study of the relations between volatilities and co-volatilities of several markets, whereas another application is time-varying hedge ratios. A related issue to both applications is whether the correlations between asset returns change over time. Firstly, we present general definition of the MGARCH and then review dif-

[11] With Box & Cox (1964) transformation dependent variables, usually assumed to be independently normally distributed with constant variances, can be estimated with less restrictive assumptions.

ferent specifications for the conditional variance matrix. These specification can be divided to three nonmutually exclusive approaches: direct generalisations of the univariate GARCH, linear combinations of the univariate GARCH models and nonlinear combinations of the univariate GARCH models (Bauwens et al. 2006, 81). From the first approach we present VEC and BEKK models by Bollerslev, Engle & Wooldridge (1988) and Engle & Kroner (1995), respectively. As the linear combinations of univariate GARCH models (and particularly the orthogonal models) are particular F-GARCH models and thus nested in to the BEKK model these are not presented here.[12] From the third approach the constant correlation model by Bollerslev (1990) and the dynamic correlation model by Engle (2002) are presented. Additionally, subsection 2.3.3 introduces a new copula-based GARCH models and the theory behind them.

Following denotation from Bauwens et al. (2006, 80-81) for a mean equation consider a vector stochastic process $\{y_t\}$ of dimension N x 1 generated by the past information until time *t-1* (denoted I_{t-1}). Parameters of the finite vector is denoted θ and then it is possible to write

$$y_t = \mu_t(\theta) + \varepsilon_t, \tag{28}$$

where $\mu_t(\theta)$ is the conditional mean vector and

$$\varepsilon_t = H_t^{\frac{1}{2}}(\theta) z_t, \tag{29}$$

where $H_t^{\frac{1}{2}}(\theta)$ is a N x N positive definite matrix. For the N x 1 random vector z_t, as in univariate models, the following first two moments are assumed

$$\begin{aligned} E(z_t) &= 0 \\ Var(z_t) &= I_N, \end{aligned} \tag{30}$$

[12] For more information, see e.g., Bauwens et al. 2006 or van der Weide 2002.

where I_N is the identity matrix of order *N*. To be able to explain $H_t^{\frac{1}{2}}$ (θ is leaved out for convenience), the conditional variance matrix of y_t is calculated

$$\begin{aligned} Var(y_t|I_{t-1}) &= Var_{t-1}(y_t) = Var_{t-1}(\varepsilon_t) \\ &= H_t^{\frac{1}{2}} Var_{t-1}(z_t)(H_t^{\frac{1}{2}})' \\ &= H_t \end{aligned} \tag{31}$$

Hence, $H_t^{\frac{1}{2}}$ can be stated as any N x N matrix that may be obtained from H_t with Cholesky decomposition when H_t is the conditional variance matrix of y_t's. Next, we review different specifications of H_t. Again, using denotation from Bauwens et al. (2006) models are only presented '(1,1)' form to keep the notational burden as low as possible.

Bollerslev et al. (1988) proposed a general VEC model to explain H_t. Each element of H_t are linear function of the lagged squared errors and cross-products of errors and lagged values of the elements of H_t and can be specified as

$$vech(H_t) = c + Avech(\varepsilon_{t-1}\varepsilon_{t-1}') + Bvech(H_{t-1}), \tag{32}$$

where $vech(\cdot)$ denotes the column stacking operator of the lower portion of N x N matrix as a *N(N+1)/2x1* vector, *c* is a *(N+1)N/2x1* parameter vector. *A* and *B* are square parameter matrices of order *(N+1)N/2* (Bollerslev et al. 1988, 119 and Bauwens et al. 2006, 82).

However, the number of parameters is *N(N+1)(N(N+1)+1)/2* (e.g., for *N=3*, the number of parameters equals 78) and thus model is practical only in the bivariate case (Bauwens et al. 2006, 82). Because of this, Bollerslev et

al. (1988, 120) proposed a version of the model where each element h_{ijt} depends only on its own lag and on the previous value of $\varepsilon_{it}\varepsilon_{jt}$ or in which the *A* and *B* are assumed to be diagonal. Diagonal VEC (DVEC) can be expressed with Hadamard products (denoted by $\odot$) and written as (Bollerslev et al. 1988, 120 and Bauwens et al. 2006, 82)

$$h_t = C^\circ + A^\circ \odot (\varepsilon_{t-1}\varepsilon'_{t-1}) + B^\circ \odot h_{t-1} \tag{33}$$

or more practically oriented way as follows

$$h_{ijt} = \gamma_{ij} + \alpha_{ij}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \beta_{ij}h_{ij,t-1}, \;\; i,j = 1,\dots,N \tag{34}$$
$$\varepsilon_t|\psi_{t-1} \sim N(0, H_t).$$

VEC and DVEC suffer from the same difficulties than univariate GARCH models, and strong restrictions for parameters are required to ensure positivity of h_t. To overcome this, Engle & Kroner (1995) proposed a new parameterisation for h_t that easily imposes its positivity. Proposed BEKK(1,1,K) is introduced next.

The BEKK(1,1,K) is defined as

$$H_t = C^{*\prime}C^* + \sum_{k=1}^{K} A_k^{*\prime}\,\varepsilon_{t-1}\varepsilon'_{t-1}A_k^* + \sum_{k=1}^{K} G_k^{*\prime}\,H_{t-1}G_k^*, \tag{35}$$

where C^*, A_k^* and G_k^* are N x N matrices, but C^* is upper triangular. However, the greatest shortcoming of the model is that the parameters do not represent directly the impact of the different lagged terms on the elements of H_t. The BEKK model is a special case of the VEC model. Both models are usually exercised only for bivariate situations as the models are difficult to estimate due the high number of unknown parameters, even after imposing several restrictions. To circumvent this problem factor and orthogonal models are developed and they impose a common dynamic

structure on all the elements of H_t, thus resulting less parameterised models.[13] (Bauwens et al. 2006, 83-84)

Next, some non-linear combinations of the univariate GARCH models is presented. Even though the theoretical results on stationarity, ergodicity and moments may not be as straightforward as with the previous models but they are less greedy in parameters than VEC and BEKK models (Bauwens et al. 2006, 88). This ensures more easy estimation. First model presented is the constant correlation model by Bollerslev (1990).

The CCC model is defined as

$$H_t = D_t R D_t = \left(\rho_{ij}\sqrt{h_{ii,t}h_{jj,t}}\right), \tag{36}$$

where

$$D_t = diag\left(h_{11,t}^{\frac{1}{2}} \dots h_{NN,t}^{\frac{1}{2}}\right). \tag{37}$$

$h_{ii,t}$ is univariate GARCH model, and

$$R = (\rho_{ij}) \tag{38}$$

is a symmetric positive definite matrix with $\rho_{ii} = 1,\ \forall i$ (Bollerslev 1990, 499 and Bauwens et al. 2006, 88). As earlier, H_t needs to be a positive definite and is if and only if all the N conditional variances are positive and *R* is positively defined (Bollerslev 1990, 499 and Bauwens et al 2006, 89). Conditional correlations are constant and this may seem a bit unrealistic for the empirical applications. Thus, several generalisations of the CCC

[13] For more information concerning factor and orthogonal models, see among other Engle et al. 1990; Bollerslev & Engle 1993; Vrontos et al. 2003 and Bauwens et al. 2006.

model have been proposed where the conditional correlation matrix is a time-varying. A version from Engle (2002) is presented because we test dynamic hedging models.

Engle (2002, 341) specifies the DCC model as

$$H_t = D_t R_t D_t, \tag{39}$$

where specification of D_t's is same as in Equation (37) and

$$R_t = diag\{Q_t\}^{-1} Q_t diag\{Q_t\}^{-1}, \tag{40}$$

where

$$Q_t = S(\mu' - A - B) + A\varepsilon_{t-1}\varepsilon'_{t-1} + BQ_{t-1}. \tag{41}$$

Q_t is N x N symmetric positive definite matrix, *S* is the N x N unconditional variance matrix of ε_t and *A* and *B* are non-negative scalar parameters which need to satisfy *A+B<1* (Engle 2002, 342 and Bauwens et al. 2006, 90). Finally, the next subsection introduces a new class of copula-based GARCH models.

2.3.3 Copula-based GARCH models

To be able to understand copula-based GARCH models, one needs first an introduction to copulas and their theory. First, a basic theory behind the unconditional copula is presented and then, following Patton (2006), extension to conditional copulas are offered. Then we show how one can estimate marginal and dependence processes and their log-likelihoods.

Following Sklar (1959) and Patton (2006) some notations are first defined. Let the two variables of interest to be X and Y and the conditioning varia-

ble be W, which may be a vector. The joint distribution of (X,Y,W) is defined as F_{XYW}, the conditional distribution of (X,Y) given W as $F_{XY|W}$ and the conditional marginal distributions $F_{X|W}$ and $F_{Y|W}$. Also, one can recall that $F_{X|W}(x|w)=F_{XY|W}(x,\infty|w)$ and $F_{Y|W}(y|w)=F_{XY|W}(y,\infty|w)$. This review assumes, again following notation from Sklar (1959), that the all necessary derivatives of distribution function F_{XYW} exists and that $F_{XY|W}$, $F_{X|W}$ and $F_{Y|W}$ are continuous. The cumulative distribution function (*or cdf*) of random variable will be denoted with an uppercase letter and the corresponding density function (*or pdf*) with the lowercase letter.

Copula theory enables to decompose a joint distribution into its marginal distributions and its dependence function, or copula. Following the notation above this can be shown as

$$F_{XY}(x,y) = C\big(F_X(x), F_Y(y)\big), or \tag{42}$$

$$f_{xy}(x,y) = f_x(x) \cdot f_y(y) \cdot c\big(F_X(x), F_Y(y)\big) \tag{43}$$

where the Equation 42 decomposes a bivariate *cdf*, and the Equation 43 a bivariate *pdf*.[14] However, the copula theory presented by Sklar (1959) describes unconditional copulas which does not take account time-varying features of the dependencies. Patton (2006) presents an extension to Sklar's theorem by letting parameters to be autoregressive or conditional on the set of past information. Patton's extension is anecdotal to Engle's ARCH development with ability to take account past information. First, some definition are presented (Patton 2006, 533).

Definition 1. *The conditional copula of* $(X,Y)|W=w$*, where* $X|W=w \sim F_{X|W}(\cdot\,|w)$ *and* $Y|W=w \sim F_{Y|W}(\cdot\,|w)$*, is the conditional joint distribution function of* $U \equiv F_{X|W}(X|w)$ *and* $V \equiv F_{Y|W}(Y|w)$ *given* $W=w$*.*

[14] For more throughout review of unconditional copulas, see Joe (1997) and Nelsen (1999).

The variables U and V are known as the conditional “probability integral transforms” of X and Y given W (Patton 2006, 533). These random variables have the Unif(0,1) distribution regardless of the original distribution as Fisher (1932) and Rosenblatt (1952) have shown. Patton (2002) has also shown that a conditional copula has the properties of an unconditional copula, for each $w \in W$.

With above definitions it is possible to presented extension to the Sklar theorem (Patton 2006, 533).

Theorem 2. *Let $F_{X|W}(\cdot|w)$ be the conditional distribution of $X|W = w$, $F_{Y|W}(\cdot|w)$ be the conditional distribution of $Y|W = w$, $F_{XY|W}(\cdot\,|w)$ be the joint conditional distribution of $(X,Y)|W = w$ and W be the support of W. Assume that $F_{X|W}(\cdot|w)$ and $F_{Y|W}(\cdot|w)$ are continuous in x and y for all $w \in W$. Then there exists a unique conditional copula $C(\cdot|w)$ such that*

$$F_{XY|W}(x,y|w) = C(F_{X|W}(x|w), F_{Y|W}(y|w)|w), \quad \forall (x,y) \in \mathbb{R}x\mathbb{R} \text{ and each } w \in W. \tag{44}$$

Conversely, if we let $F_{X|W}(\cdot|w)$ be the conditional distribution of $X|W = w$, $F_{Y|W}(\cdot|w)$ be the conditional distribution of $Y|W = w$, and $\{C(\cdot|w)\}$ be a family of conditional couplas that is measurable in w, then the function $F_{XY|W}(\cdot\,|w)$ defined by the Equation 44 is a conditional bivariate distribution function with conditional marginal distributions $F_{X|W}(\cdot|w)$ and $F_{Y|W}(\cdot|w)$.

However, to be able to extend Sklar’s theorem, the conditioning variable(s), W, must be same for both marginal distributions and the copula. It is necessary to use the same conditioning variable for FX|W, FY|W and C or the function FXY|W will not satisfy the conditions for it to be a joint conditional distribution function.

The density function equivalent of the Equation 44 can be estimated with maximum likelihood method given that $F_{X|W}$ and $F_{Y|W}$ are differentiable and $F_{XY|W}$ and C are twice differentiable. First a conditional bivariate distribution function is transformed to density function

$$f_{XY|W}(x,y|w) \equiv \frac{\partial^2 F_{XY|W}(x,y|w)}{\partial x \partial y}$$
$$= \frac{\partial F_{X|W}(x|w)}{\partial x} \times \frac{\partial F_{Y|W}(y|w)}{\partial y} \times \frac{\partial^2 C\left(F_{X|W}(x|w), F_{Y|W}(y|w)|w\right)}{\partial u \partial v} \quad (45)$$
$$f_{XY|W}(x,y|w) \equiv f_{X|W}(x|w) \times f_{Y|W}(y|w) \times c(u,v|w),$$
$$\forall (x,y,w) \in \mathbb{R} \times \mathbb{R} \times \mathcal{W}$$

and then estimated with maximum likelihood method

$$\mathcal{L}_{XY} = \mathcal{L}_X + \mathcal{L}_Y + \mathcal{L}_C \quad (46)$$

where $u \equiv F_{X|W}(x|w)$, and $v \equiv F_{Y|W}(y|w)$, $\mathcal{L}_{XY} \equiv \log f_{XY|W}(x,y|w)$, $\mathcal{L}_X \equiv \log f_{X|W}(x|w)$, $\mathcal{L}_Y \equiv \log f_{Y|W}(y|w)$, and $\mathcal{L}_C \equiv \log c(u,v|w)$.

As with other GARCH models, selected distribution function is essential for estimation of correlations between assets.

After presenting the development of GARCH models that can be used to estimate the optimal hedge ratios, some previous studies concerning the dynamic hedging is presented.

2.4 Previous studies

Even though the academic interest to volatility models and volatility forecasting has been great in recent years and decades, an interest to forecast the hedge ratios has been less scarce. However, in recent years research has seen a new dawn and many authors have applied the models presented in the previous section to the hedge ratio estimation (e.g. Cecchetti, Cumby & Figlewski 1988; Baillie & Myers 1991; Kroner & Sultan 1991, 1993; Myers 1991; Garcia, Roh & Leuthold 1995; Park & Switzer 1995; Tong 1996; Butterworth & Holmes 2000; Brailsford, Corrigan & Heaney 2001; Brooks & Chong 2001; Lien & Tse 2002; Choudry 2003, 2004; Floros & Vougas 2004; Ku, Chen & Chen 2007; Hsu et al. 2008; Sultan & Hassan 2008 and Park & Jei 2010). However, there are no studies conducted with the BRICS countries. In addition, the articles that studies the dynamic hedging are divided between the different type of markets (i.e. equity, commodity and exchange rate markets). Thus, comparing the effectiveness of the hedging performance is not straightforward as different markets has their own characteristics and the markets have evolved from the earlier studies. We focus on the studies conducted with the stock-index futures as these have most resemblance to our research.

The stock-index futures and the hedging with them is the least studied area of the dynamic hedging. This could be because the international diversification has been deemed to provide an adequate natural hedge due the low correlation between the national stock markets and the hedging of the exchange rates is more a concern of the foreign direct investments (FDIs). However, the economic and financial integration has increased the correlations among the different markets and thus decreased the benefits from the international diversification (see e.g. Baele & Inghelbrecht 2009, Bekaert & Harvey 2000, Brooks & Del Negro 2004, Driessen & Laeven 2007, Errunza, Hogan & Hung 2009, Goetzmann et al. 2005, Kizys & Pierdzioch 2009 and Longin & Solnik 1995). In addition, the emerging markets are started to open for the foreign investors the portfolio invest-

ments need also be hedged. Thus, hedging the international portfolios is nowadays more important than it was previously.

Most of the authors have found that dynamic hedging models outperform conventional model when studied with the stock-index futures (see e.g., Park & Switzer 1995; Butterworth & Holmes 2000; Choudry 2003; Floros & Vougas 2004 and Hsu et al. 2008). However, most of the studies are conducted with the market data from the developed countries and Choudry (2003) was the only one who incorporated a developing country (South Africa) to his study. In addition, Park & Switzer (1995) and Butterworth & Holmes (2000) incorporated weekly adjustment and thus can be considered to be a different methodology compared to daily adjustment. Also, Butterworth & Holmes (2000) employed the Johnson (1960) minimum variance strategy which is not a time-varying hedge ratio. Floros & Vougas (2004) estimated the hedge ratios of the different hedging models and no specific variance reduction was studied. Hsu et al. (2008) found mixed results as the effectiveness of in-sample hedging did not provide significant variance reduction compared to the conventional model. However, out-of-sample results showed significant variance reduction with both CCC GARCH and DCC GARCH. Thus, it can be stated that the result from the previous studies with the stock-index futures are mixed and they are very sensitive for the selection of the markets, studied time periods and the portfolio adjustment frequency.

Similar results can be found from the studies conducted with the different markets (namely, commodity and exchange rate futures). Most of the authors (see e.g., Cecchetti, Cumby & Figlewski 1988; Baillie & Myers 1991; Kroner & Sultan 1991, 1993; Myers 1991; Garcia, Roh & Leuthold 1995; Tong 1996; Brooks & Chong 2001 and Ku, Chen & Chen 2007) found evidence that the dynamic models outperform conventional one. However,

results are again mixed, showing similar pattern than with the stock-index futures.

The previous literature can be summarised to be on the side of the dynamic hedging models (see e.g. Cecchetti, Cumby & Figlewski 1988; Baillie & Myers 1991; Kroner & Sultan 1991, 1993; Myers 1991; Garcia, Roh & Leuthold 1995; Park & Switzer 1995; Tong 1996; Butterworth & Holmes 2000; Brailsford, Corrigan & Heaney 2001; Brooks & Chong 2001; Choudry 2003, 2004; Floros & Vougas 2004 and Ku, Chen & Chen 2007). However, some of the more recent studies have found support that the GARCH type dynamic hedging models would not significantly outperform conventional model (see e.g. Lien & Tse 2002; Hsu et al. 2008; Sultan & Hassan 2008 and Park & Jei 2010). Lien & Tse (2002) conducted a survey of the recent developments in the futures hedging and authors' conclusion was that the results are mixed when considering performance of the GARCH hedge ratios. Their review showed that in-sample hedging performance varies between studies and out-of-sample performance mostly favours conventional model. Interesting, the only study (Brooks & Chong 2001) we are aware of that tested Exponentially Weighted Moving Average (EWMA) for hedging purposes outperformed GARCH type models. Moreover, some authors (Brooks & Chong 2001 and Tong 1996) argued that the advantages of the dynamic hedging are more significant with the cross hedge with currency futures.

It also seems that most of the authors want to interpret the variance reductions (which are usually minor) in favour of the dynamic hedging strategies and usually ignore the transaction costs from the regular portfolio adjustment. Only Park & Switzer (1995) acknowledged the shortfall of the frequent rebalancing.

3 DATA AND METHODOLOGY

Research methodology is a quantitative in nature and empirical analyses are conducted with Matlab using MFE Toolbox by Kevin Sheppard. For the empirical data, BRICS countries indices as well as their futures are used. Hedging models are, as stated above, dynamical models and their hedging performance is studied. Methodology for hedging performance follows the one proposed by Kroner & Sultan (1993) and Hsu et al. (2008).

Firstly, diagnostic analysis of the assets is presented. Secondly, parameter estimation of the models is conducted with the maximum likelihood method. After the parameter estimation in-sample comparison of the hedging performance is done. For comparison, measurement variance of the returns of the portfolios is calculated using equation below

$$var(s_t - \delta_t^* f_t). \tag{47}$$

To check robustness of the results, holding period of the portfolios will be extended to one and two days. Another measurement for in-sample comparison that will be used is the percentage reduction of the variance over conventional method. In addition to the in-sample comparison, out-of-sample comparison will be conducted. The method adopted to determine the series of out-of-sample hedge ratios involves rolling over a particular number of samples (80% of the whole range). The variation in portfolio returns is evaluated using the spot and futures returns that are observed at the forecasted date. This rolling method is used until all of the data set is exhausted. Then, the same comparison measures are calculated as with the in-sample comparison. Using these methods it is possible to conclude which, if any, of the models outperformed static method in the in-sample hedging and which in the out-of-sample hedging.

3.1 Data properties

This section examines diagnostic analysis of the assets. All data is obtained from Bloomberg Terminal, running from 25 March 2003 to 16 April 2014 and Table 1 presents the asset specific time series. As stated in Section 1, we study how the dynamic hedging models perform with market data from BRICS countries. We have used stock indices and index futures from CSI 300 for China, Ibovespa for Brazil, JSE Top 40 for South Africa, MICEX for Russia and Sensex for India. All indices contain country's most highly traded stocks and they all are capitalisation weighted. The asset returns are the logarithmic changes of the daily closing prices.

Table 1. Asset time-series and number of observations

Asset	Time-series	No of observations
CSI 300	16 April 2010 – 16 April 2014	967 obs.
Ibovespa	1 April 2003 – 16 April 2014	2 724 obs.
JSE Top 40	1 April 2003 – 16 April 2014	2 761 obs.
MICEX	3 August 2005 – 3 March 2014	2 112 obs.
Sensex	1 April 2003 – 16 April 2014	2 645 obs.

Table 2 depicts summary statistics of the data and stylised facts, such as skewness, leptokurtosis and highly significant Jarque-Bera statistics, imply that unconditional distributions of the spot and future returns are asymmetric, fat-tailed and non-Gaussian. All series (excl. Sensex) show a negative skewness indicating frequent small gains and a few extreme losses. Thus, providing an additional reason to hedge returns. In addition, the kurtosis present in all series shows that all markets are fat-tailed and thus extreme gains/losses are more probable than with Gaussian distributed series. The Ljung-Box test (Q(10)) shows serial correlation in all returns series except CSI 300. Same trend is present at $Q^2(10)$ and ARCH(5) tests where squared returns are autocorrelated. Exceptions are ARCH(5) and $Q^2(10)$ tests for CSI 300.

Table 2. Summary statistics

Note. J-B is the Jarque-Bera test for normality; Q(10) is Ljung-Box statistics for up to 10th-order serial correlation in the returns; $Q^2(10)$ is the Ljung-Box statistics for the serial correlations in the squared returns; and ARCH(5) is the LM test for up to the fifth-order ARCH effects. * indicates significance at the 1% level.

	Asset									
	CSI 300		Ibovespa		JSE Top 40		MICEX		Sensex	
Statistics	Stock	Futures	Stock	Futures	Stock	Futures	Stock	Futures	Stock	Futures
Mean	-0.0004	-0.0004	0.0005	-0.0002	0.0006	0.0006	0.0003	0.0002	0.0007	0.0008
Std. Dev.	0.0140	0.0137	0.0179	0.0144	0.0138	0.0140	0.0238	0.0266	0.0158	0.0163
Skewness	-0.1629	-0.0435	-0.0656	-0.2158	-0.1381	-0.0614	-0.4876	-0.3576	0.0015	0.0337
Kurtosis	4.8528	6.2044	8.3729	6.6696	6.4883	5.8228	24.0566	14.3827	10.4899	10.1937
J-B	142*	414*	3,322*	1,533*	1,408*	918*	39,083*	11,441*	6,180*	5,702*
Q(10)	13.20	14.45	56.73*	55.78*	25.33*	31.24*	24.49*	32.79*	33.95*	24.57*
$Q^2(10)$	16.39	14.45	4,057*	3,104*	2,204*	1,917*	626*	1,345*	742*	795*
ARCH(5)	6.96	3.15	584*	495*	564*	503*	378*	454*	250*	262*

Table 3. Unit Root and Cointegration Tests

Note. The ADF tests are applied to test the null hypothesis of a unit root for the spot and futures prices and the returns; Trace is the Johansen trace test, with null hypothesis being that there is no cointegration; λ is the estimated cointegrating parameter. * indicates significance at the 1% level.

	Asset									
	CSI 300		Ibovespa		JSE Top 40		MICEX		Sensex	
	Stock	Futures	Stock	Futures	Stock	Futures	Stock	Futures	Stock	Futures
ADF (price)	-1.23	-1.32	0.15	0.13	1.91	1.88	-0.11	-0.35	1.10	1.05
ADF (return)	-31.28*	-30.28*	-52.22*	-52.63*	-51.94*	-51.89*	-46.68*	-44.40*	-47.39*	-48.46*
Trace	650*	-	1,888*	-	1,873*	-	1,376*	-	1,633*	-
λ	0.4901	-	0.5002	-	0.4928	-	0.4791	-	0.4609	-

Table 3 depicts the results from the unit root and cointegration tests. The augmented Dickey-Fuller (ADF) tests show that the spot and futures prices have a unit root but stationarity is obtained with the first-difference. The Johansen trace statistics show the cointegration between the spot and futures prices and thus the error-correction terms should be considered in section 3.2 when the model specifications are done. As the cointegrating parameters (λ) deviate significantly from one, actual values are used and error-correction term used in section 3.2 is $(S_{t-1} - \lambda F_{t-1})$.

3.2 Methodology

3.2.1 Conventional model and risk minimising hedge ratio

This Thesis examines three models more closely and uses these for empirical testing, namely conventional hedge, CCC GARCH and DCC GARCH (Kroner & Sultan 1993; Engle & Sheppard 2001 and Engle 2002). To be able to present the above models it is needed to start with conventional model from Kroner & Sultan (1993). In the case of conventional modelling, the hedged portfolio has been found by regressing the returns of the spot asset on the returns of the hedging instruments (Kroner & Sultan 1993, 535). Formally

$$s_t = \alpha + \beta f_t + \varepsilon_{t.} \tag{48}$$

If the joint distribution of the spot and futures remains the same over time, risk-minimising hedge ratio[15] δ^* (corresponding β in Equation 48) will be (Kroner & Sultan 1993, 537 and Hsu et al. 2008, 1098)

$$\delta^* = \frac{cov(s_t, f_t)}{var(f_t)} \; or \; \delta^* = \frac{\sigma_{s,f}}{\sigma_f^2}. \tag{49}$$

[15] To follow the conventional terminology we use delta (δ) to nominate a hedge ratio.

3.2.2 GARCH models and time-varying hedge ratio

However, if the distribution of the spot and futures prices is time-varying an alternative model is needed (Kroner & Sultan 1993, 537 and Hsu et al. 2008, 1098). If the conditional information set at time *t*-1 is used, optimal time-varying hedge ratio is as follows

$$\delta_{t-1}^{*} = \frac{cov_{t-1}(s_t, f_t)}{var_{t-1}(f_t)}, \tag{50}$$

or with econometric terms

$$\hat{\delta}_t^{*} = \frac{\hat{h}_{sf,t}}{\hat{h}_{f,t}^{2}}. \tag{51}$$

In the case of the time-varying hedge ratio, conditional variance and covariances need to be specified as shown above in the Equation 51. Kroner & Sultan (1993, 538) have presented bivariate error-correction model for this purpose. To be able to use the bivariate GARCH error-correction model, the first two conditional moments of s_t and f_t need to be parameterised (Kroner & Sultan 1993, 538). The first moment is modelled with a bivariate error-correction model from Engle & Granger (1987) and the second moment is modelled with a bivariate GARCH model. The econometric model of first moment is as follows

$$s_t = \alpha_{0s} + \alpha_{1s}(S_{t-1} - \lambda F_{t-1}) + \varepsilon_{st} \tag{51}$$

$$f_t = \alpha_{0f} + \alpha_{1f}(S_{t-1} - \lambda F_{t-1}) + \varepsilon_{ft},$$

$$\begin{bmatrix} \varepsilon_{st} \\ \varepsilon_{ft} \end{bmatrix} | \, \Psi_{t-1} \sim Dist(0, H_t). \tag{52}$$

Where S_{t-1} and F_{t-1} are the spot and futures prices, respectively, $S_{t-1} - \lambda F_{t-1}$ is the error-correction term, ψ_{t-1} is the information set at time *t-1*, and the disturbance term $\varepsilon_t = (\varepsilon_{st}, \varepsilon_{ft})'$ follows a certain bivariate distribu-

tion with a zero mean and a conditional covariance matrix H_t. From Equations 51 and 52 can be seen that the disturbance term will be time-varying and hence indicating time-varying hedge ratios and the error-correction term characterises the long-run relationship between the spot and futures prices (Hsu et al. 2008, 1099).

For the optimal hedge ratio, we need to obtain conditional variances which can be estimated with a univariate GARCH model

$$\begin{aligned} h_{s,t}^2 &= c_s + \alpha_s \varepsilon_{s,t-1}^2 + b_s h_{s,t-1}^2 \\ h_{f,t}^2 &= c_f + \alpha_f \varepsilon_{f,t-1}^2 + b_f h_{f,t-1}^2. \end{aligned} \tag{53}$$

Conditional covariances are obtained with the multivariate GARCH models and the first model we use for estimation is the CCC GARCH by Bollerslev (1990)

$$H_t = \begin{bmatrix} h_{ss,t} & h_{sf,t} \\ h_{sf,t} & h_{ff,t} \end{bmatrix} = \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix}, \tag{54}$$

where ρ is constant correlation coefficient.

The assumption of Equation 54 that conditional correlation is constant could, however, be too restrictive to fit in with reality (Engle 2002, 341 and Hsu et al. 2008, 1099). To release this restriction and improve the flexibility of the hedging models Engle and Sheppard (2001) and Engle (2002) proposed the DCC GARCH model. It can be stated to be generalisation of CCC GARCH model from Bollerslev (1990). To get an intuition for the DCC GARCH firstly above Equation 54 need to be expanded as follows

$$H_t = \begin{bmatrix} h_{ss,t} & h_{sf,t} \\ h_{sf,t} & h_{ff,t} \end{bmatrix} = \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} = D_t R D_t. \tag{55}$$

Where D_t is the N x N stochastic diagonal matrix ($D_t = diag\{\sqrt{h_{i,t}}\}$) and R is a correlation matrix containing the conditional correlations (Engle 2002, 341). The DCC GARCH model allows the correlation matrix to be time-varying. This can be written as follows (Engle 2002, 341)

$$H_t = D_t R_t D_t. \tag{56}$$

Where $R_t = diag\{Q_t\}^{-1} Q_t diag\{Q_t\}^{-1}$, $Q_t = S(u' - A - B) + A\varepsilon_{t-1}\varepsilon'_{t-1} + BQ_{t-1}$ and D_t is the N x N stochastic diagonal matrix as with CCC GARCH model.

For the models to be usable, the parameter estimations are needed which then are used to test the variance reduction. The next section presents estimated parameters for the models and shows the effectiveness of in-sample and out-of-sample hedging.

4 EMPIRICAL RESULTS

This section provides empirical results derived from the estimated parameters. Firstly, parameter values for each testable model is presented. Then, we present hedge ratios on different markets and evaluate whether there is difference between pre- and post-financial crisis (i.e. before and after 31 December 2007). Next, we present the comparison of the effectiveness of in-sample hedging between considered models. Lastly, the effectiveness of out-of-sample hedging is presented to show how the models fare in a real like situation.

4.1 Estimation of the parameters

This section presents the estimation results of the different hedging models when all parameters have been estimated with the maximum likelihood method. For the CCC GARCH model Table 4 depicts the estimated parameters for both mean and variance functions. As can be seen from the table, constant correlations (ρ) between the spot and futures returns are all positive and close to one (over 0.95), with exception of MICEX (approximately 0.87). In addition, all constant correlations show extremely high statistical significance. Hence, correlations between the spot and futures returns are assumed to be high throughout the estimation period. It can also be noted that the value of the log-likelihood function for CSI 300 is considerably lower than for the other markets. It is most likely due the shorter estimation period as for CSI 300 only 967 observation was available compared to approximately 2,500 observations for the other markets.

Table 4. Constant Correlation GARCH estimations

Note. The table presents the maximum likelihood estimates of the constant correlation GARCH model with Student's *t* distribution. Figures in the parentheses are *p*-values. The model is described as follows

$$s_t = \alpha_{0s} + \alpha_{1s}(S_{t-1} - \delta F_{t-1}) + \varepsilon_{st}$$
$$f_t = \alpha_{0f} + \alpha_{1f}(S_{t-1} - \delta F_{t-1}) + \varepsilon_{ft},$$
$$\begin{bmatrix}\varepsilon_{st}\\ \varepsilon_{ft}\end{bmatrix} | \Psi_{t-1} \sim Student's\ t(0, H_t),$$
$$H_t = \begin{bmatrix} h_{ss,t} & h_{sf,t} \\ h_{sf,t} & h_{ff,t}\end{bmatrix} = \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t}\end{bmatrix}\begin{bmatrix} 1 & \rho \\ \rho & 1\end{bmatrix}\begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t}\end{bmatrix},$$
$$h_{s,t}^2 = c_s + \alpha_s \varepsilon_{s,t-1}^2 + b_s h_{s,t-1}^2$$
$$h_{f,t}^2 = c_f + \alpha_f \varepsilon_{f,t-1}^2 + b_f h_{f,t-1}^2.$$

	CSI 300		Ibovespa		JSE Top 40		MICEX		Sensex	
Parameters	i=s	i=f	i=s	i=f	i=s	i=f	i=s	i=f	i=s	i=f
a_{0i}	-0.0058	-0.0053	0.0012	0.0012	0.0006	0.0006	-0.0054	-0.0065	0.0015	0.0016
	(0.0986)	(0.1249)	(0.2522)	(0.2335)	(0.4003)	(0.4171)	(0.0485)	(0.0343)	(0.0780)	(0.0781)
a_{1i}	3.36E-06	3.02E-06	-2.25E-08	-2.49E-08	3.56E-09	4.24E-09	7.67E-06	8.94E-06	-1.18E-07	-1.24E-07
	(0.1226)	(0.1565)	(0.5219)	(0.4825)	(0.9461)	(0.9368)	(0.0342)	(0.0269)	(0.3324)	(0.3210)
c_i	9.64E-06	4.98E-06	6.95E-06	7.02E-06	2.32E-06	2.30E-06	1.12E-05	1.29E-05	6.24E-06	6.65E-06
	(0.0618)	(0.0513)	(0.0001)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
a_i	0.0311	0.0197	0.0582	0.0532	0.0461	0.0457	0.1448	0.1236	0.0725	0.0731
	(0.0239)	(0.0164)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
b_i	0.9290	0.9557	0.9494	0.9527	0.9514	0.9524	0.8747	0.8903	0.9525	0.9515
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
p	0.9561		0.9839		0.9822		0.9119		0.9842	
	(0.0000)		(0.0000)		(0.0000)		(0.0000)		(0.0000)	
Log-likelihood	6,840		19,031		20,935		12,578		19,411	

Table 4 also shows that all $a_i + b_i$ estimates are also close to one (over 0.97), implicating that the shocks in the stock and futures markets have a high persistence in volatility. Thus, indicating that the long-run average variance is not good proxy for calculating hedge ratio in the short time horizon.[16]

Table 5 depicts the estimated parameters for the DCC GARCH model. However, the marginal processes are identical to those in Table 4 and thus only estimations for the variance functions are shown. Estimations of A + B are not unequivocally close to one unlike with the CCC GARCH model. For example in case of CSI 300 (0.317) and JSE Top 40 (0.759). This indicates that correlations between futures and underlying assets are not always highly persistent and thus the DCC GARCH would not be better for capturing the principal phenomena. Hence, the results should not be unanimous and it can be expected that the CCC GARCH can outperform the DCC model in some cases.

[16] Thus, Table 4 provides evidence that the Equation 49 does not describe an optimal hedging ratio as the denominator ($var(f_t)$) does not consider a persistence in volatility.

Table 5. Dynamic Correlation GARCH model estimations

Note. The table presents the maximum likelihood estimates of the dynamic conditional correlation GARCH model. The marginal processes are identical to those in the CCC GARCH model presented in Table 4, hence only the estimates of the correlation parameters are reported. Figures in parentheses are *p*-values. The correlation process is specified as follows

$$H_t = D_t R_t D_t.$$

$$R_t = diag\{Q_t\}^{-1} Q_t diag\{Q_t\}^{-1}$$

$$Q_t = S(u' - A - B) + A\varepsilon_{t-1}\varepsilon_{t-1}' + BQ_{t-1}$$

D_t is the N x N stochastic diagonal matrix.

	CSI 300		Ibovespa		JSE Top 40		MICEX		Sensex	
Parameters	i=s	i=f	i=s	i=f	i=s	i=f	i=s	i=f	i=s	i=f
A	0.0386		0.0045		0.1619		0.0829		0.1081	
	(0.0000)		(0.0000)		(0.0000)		(0.0000)		(0.0000)	
B	0.2780		0.9955		0.5973		0.8995		0.8674	
	(0.0000)		(0.0000)		(0.0000)		(0.0000)		(0.0000)	
Log-likelihood	6,777		18,394		20,714		12,591		19,286	

Table 5 provides the same insight for the log-likelihood functions as Table 4 as CSI 300 shows significantly lower values than the other markets. Low persistence between the spot and futures markets in CSI 300 needs to be considered with certain scepticism and we cannot say that the DCC GARCH is a better way to model the volatility.

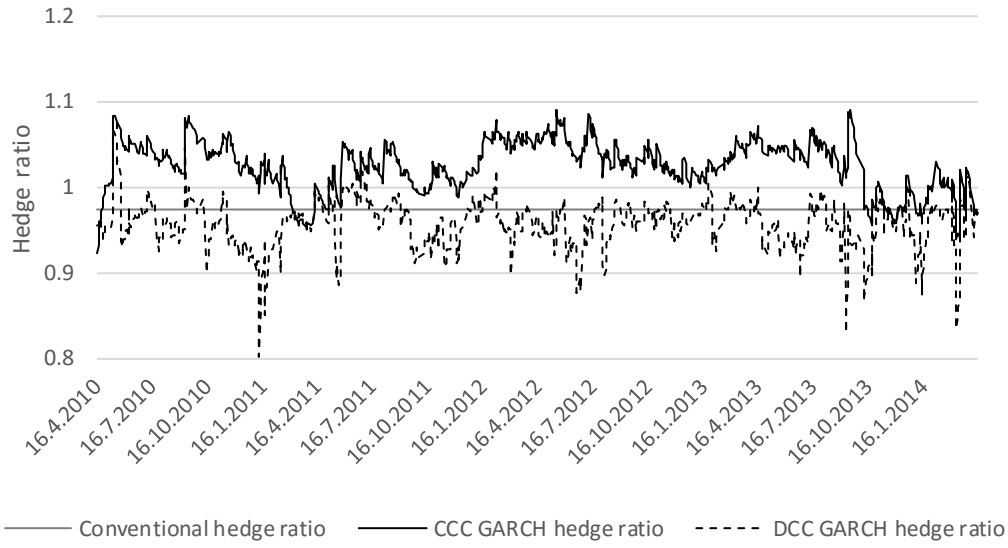
4.2 Hedge ratios on different markets

This section compares the hedge ratios of the different models on each studied market. First, the hedge ratios on full period is presented and then the hedge ratios are divided to two sub-periods (before and after 31 December 2007) to be able to compare whether there is significant changes between the periods. Each subsequent figure shows the hedge ratios of the conventional model as a grey (straight) line, the hedge ratios of the CCC GARCH model as a black line and the hedge ratios of the DCC GARCH as a dashed line.

Figure 1 depicts hedge ratios on Chinese (CSI 300) market. As the period being studied is relative short and starts from 2010, we have not divided it to the sub-periods. As can be seen from the figure, both dynamic hedging models provide very volatile hedge ratios which is to be expected as underlying assets are rather volatile in nature. Conventional hedge ratio is close to one (0.97) and the dynamic hedge ratios also move around the complete hedge. The largest daily hedge ratios are 1.09 and 1.07 for CCC GARCH and DCC GARCH. On the other hand, the lowest respective hedge ratios are 0.92 and 0.80. Hence, following either one of the dynamic models an investor needs to balance her portfolio regularly and transactions could be quite large as daily changes are significant. It can be also noted that the CCC GARCH and the DCC GARCH give a very different estimation for the optimal hedge ratio as most of the time CCC GARCH's hedge ratio is higher and DCC GARCH's hedge ratio is lower than conventional one. Thus, we can expect to see the significant divergences as

we compare the variance reductions to the conventional model on the next section.

Figure 1. Hedge ratios on CSI 300 (full period)



Next we examine the hedge ratios for the Brazilian (Ibovespa) market. Figure 2 shows hedge ratios of the different models on full examination period. For Ibovespa, the conventional hedge ratio is also close to one (0.96) and the dynamic hedge ratios are also close to one most of the days. However, as can be seen from the figure both dynamic models show large deviations from the mean for single days. For the CCC GARCH, the largest hedge ratio is 1.07 and for the DCC GARCH it is 1.24. Similarly, the lowest hedge ratios diverge between dynamic models as the lowest hedge ratios for CCC GARCH and DCC GARCH are 0.76 and 0.29.

Figure 2. Hedge ratios on Ibovespa (full period)

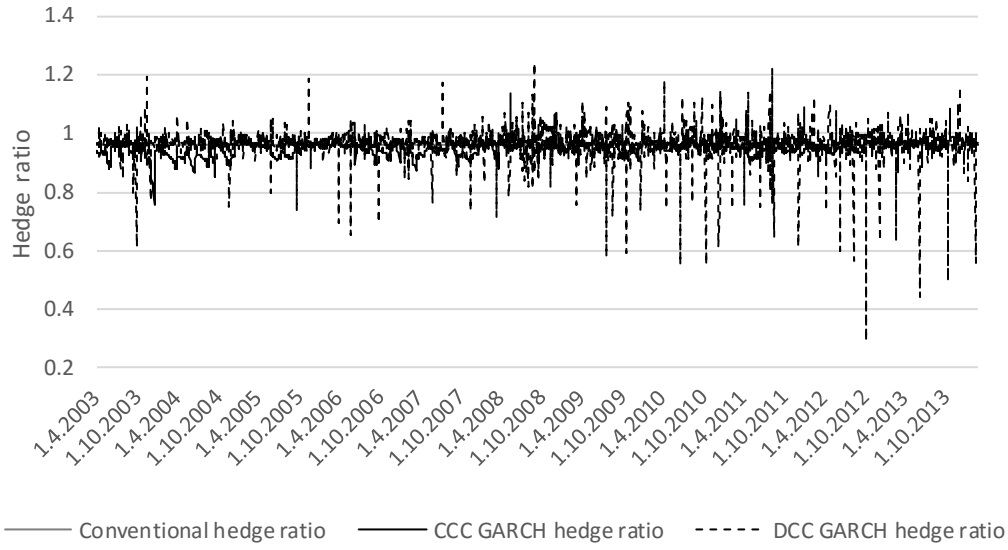
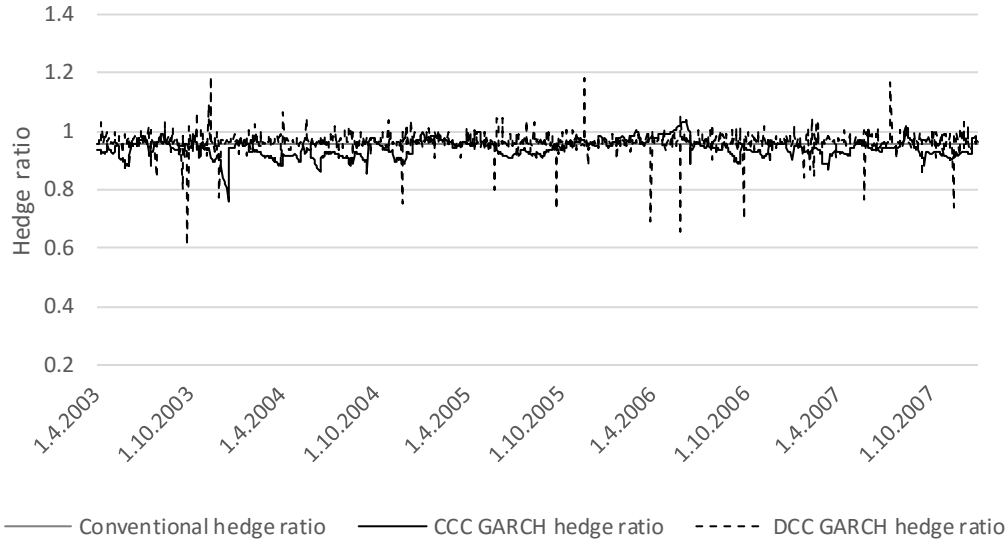


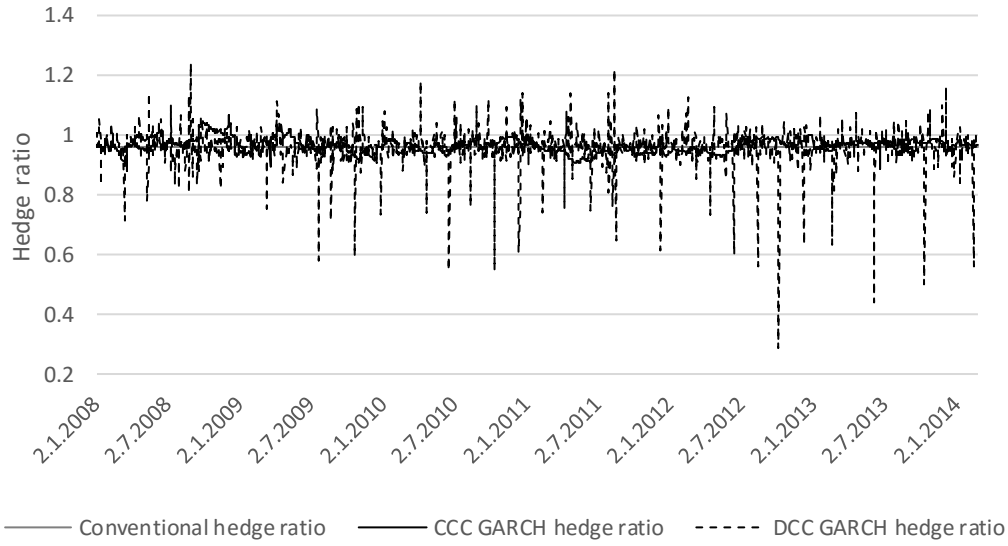
Figure 3 presents the hedge ratios on the sub-period before 31 December 2007. We can observe that the hedge ratio calculated with the CCC GARCH has remained fairly stable as sub-period maximum ratio is 1.04 and minimum 0.76, whereas the DCC GARCH based hedge ratio has been more volatile (max 1.19 and min 0.61). Comparison to the previous figure shows that especially the hedge ratio calculated with CCC GARCH has been more stable at early years of the period. In addition, the single day volatilities have been rather stable when compared to the sub-period presented in Figure 4. This is most likely due the fact that Figure 3 depicts period known as Great Moderation when market volatilities were low and most of the investors believed that financial markets has overcome risks.

Figure 3. Hedge ratios on Ibovespa (before 31 December 2007)



As can be seen from Figure 4, the Great Moderation did not last and financial crisis and its aftermath brought high volatility back to the financial markets, which can be clearly seen from the hedge ratios calculated with the DCC GARCH. During the second sub-period, the maximum hedge ratios for CCC GARCH and DCC GARCH are 1.07 and 1.24, whereas the minimum hedge ratios are 0.88 and 0.29. In addition, large deviations from the mean are much more regular than before the financial crisis when calculated with the DCC GARCH.

Figure 4. Hedge ratios on Ibovespa (31 December 2007 onwards)



Next, we move to South African (JSE Top 40) markets and Figure 5 depicts the hedge ratios from the full sample period. In JSE we see the same pattern as in previously studied markets as the conventional hedge ratio is close to one (0.95) and the dynamic hedge ratios show different kind of volatility patterns. When considering the whole period, maximum hedge ratios for the CCC GARCH and DCC GARCH are 1.07 and 1.31, while the minimum ratios are 0.76 and 0.12. All hedge ratios are close to one most of the time and again, the large deviations from the mean, especially with the DCC GARCH are documented.

Figure 5. Hedge ratios on JSE Top 40 (full period)

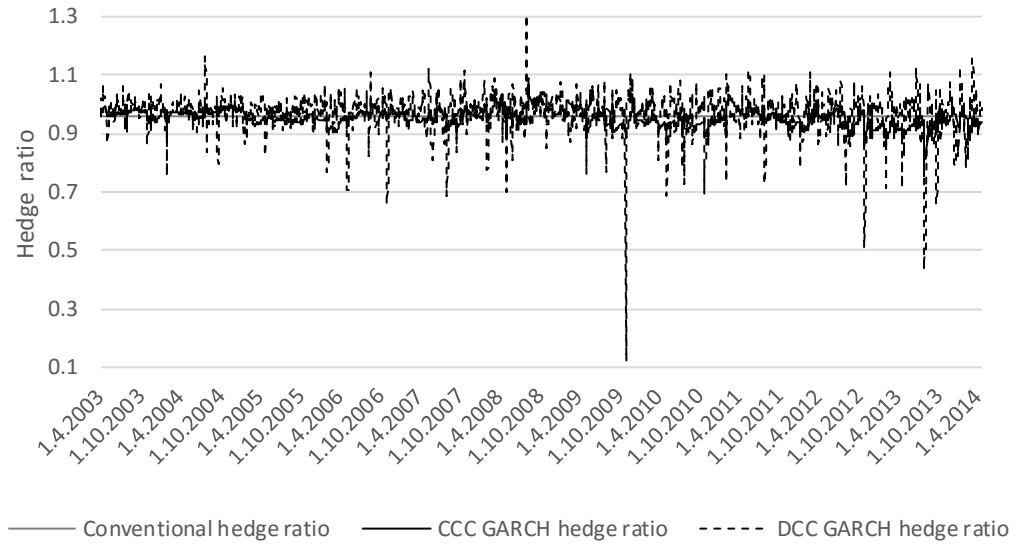


Figure 6 depicts the hedge ratios for the sub-period before 31 December 2007. Following the same pattern as Brazilian market, the CCC GARCH hedge ratios are more stable than the DCC GARCH ratios. For the CCC GARCH, the maximum hedge ratio for the sub-period was 1.01 while minimum was 0.89, whereas the maximum hedge ratio for the DCC GARCH was 1.16 and the minimum 0.66. However, Figure 6 shows that the large deviations from the mean were still rather irregular compared to the Figure 7 and the sub-period after 31 December 2007.

Figure 6. Hedge ratios on JSE Top 40 (before 31 December 2007)

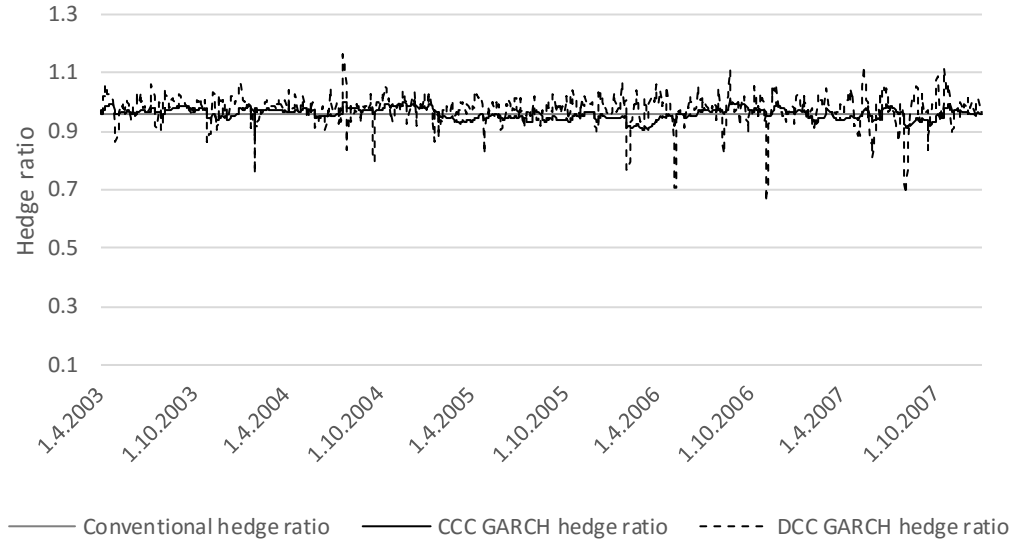
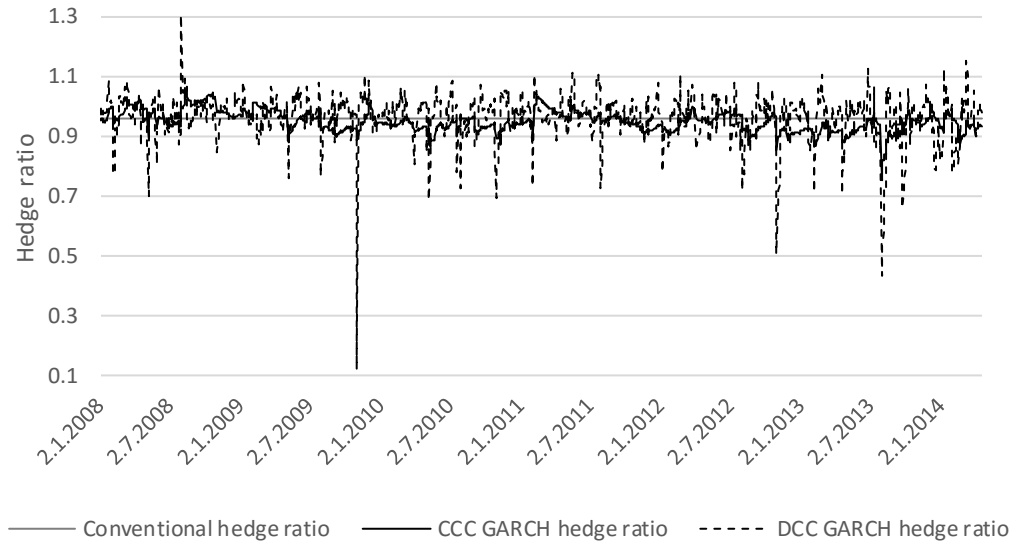


Figure 7 shows the hedge ratios for the sub-period after 31 December 2007. During and after the financial crisis the DCC GARCH hedge ratios started to oscillate more (the largest hedge ratio 1.31 and the smallest 0.12) and deviations from the conventional hedge ratio were much larger than before 31 December 2007. At the same time, the CCC GARCH hedge ratios were more stable as the largest hedge ratio was 1.07 and the smallest 0.76. Thus, the South African market provides further evidence of the existence of the Great Moderation and the financial markets' perception of the existence of the risk after the financial crisis.

Figure 7. Hedge ratios on JSE Top 40 (31 December 2007 onwards)



The Russian market (MICEX) provides the same big picture even though the noise in this particular market is greater. Figure 8 shows the hedge ratios for the full sample period. The conventional hedge ratio (0.76) is significantly lower than on the other markets and the dynamic hedge ratios are more volatile throughout the whole period. The largest hedge ratios for the CCC GARCH and DCC GARCH were 1.26 and 1.21 and Russia is the other market beside China where the maximum CCC GARCH hedge ratio is larger than the maximum DCC GARCH hedge ratio.

Figure 8. Hedge ratios on MICEX (full period)

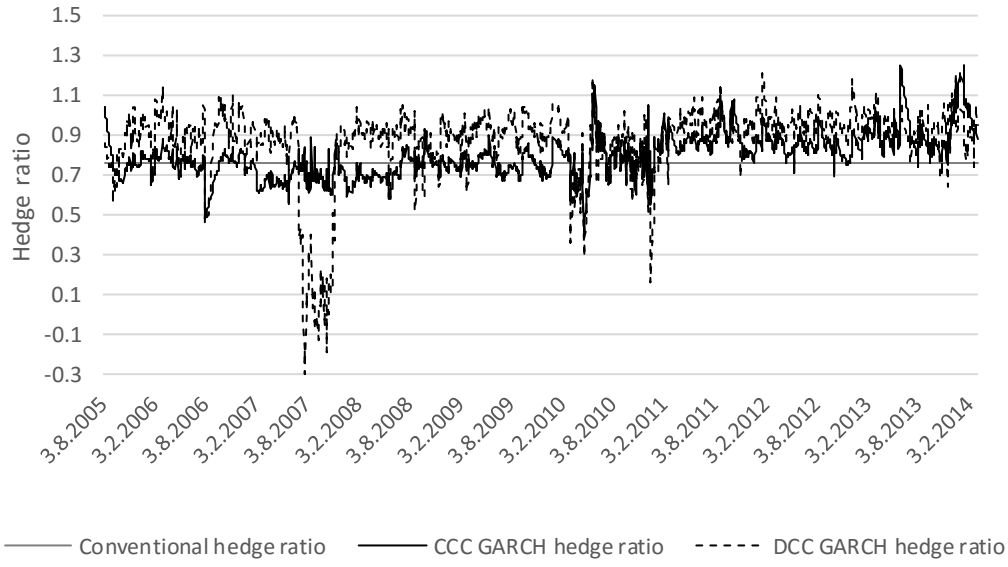


Figure 9 shows that for most of the sub-period from August 2005 to December 2007, the dynamic hedge ratios deviate greatly from the mean and the DCC GARCH hedge ratio remain higher. As the largest hedge ratios were 1.04 and 1.14 and the smallest were 0.48 and -0.30 for CCC GARH and DCC GARCH, respectively, the markets can be interpreted to be unsure of whether or not to hedge against Russian market risks. However, the significant drop in the DCC GARCH hedge ratio from early July 2007 onwards occurred, whereas the CCC GARCH hedge ratio remained close to the conventional one. No obvious reasons for the drop of the hedge ratio can be given but the significantly lower hedge ratio could indicate that the investors believed that hedging of the Russian market was not essen-

tial anymore as the drop is dated very close to the start of the sharp rise in share prices in Russia.

Figure 9. Hedge ratios on MICEX (before 31 December 2007)

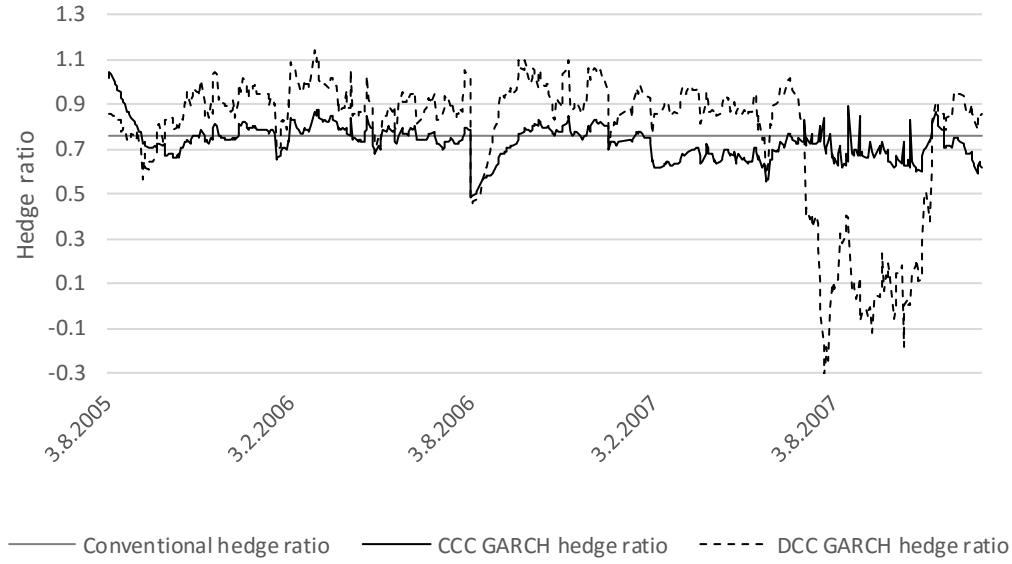


Figure 10 shows that after the large drop in the hedge ratio, the DCC GARCH rose back to approximately same level as the conventional hedge. During the latter sub-period the dynamic hedge ratios have stayed higher than the conventional hedge most of the time and thus we would expect to see the variance reduction compared to the conventional model. The hedge ratios are rather stable when compared to the other models with the exception of year 2010 when volatility was greater. Hence, the highest hedge ratios for the sub-period are 1.26 and 0.41 for the CCC GARCH and DCC GARCH, whereas the lowest ratios were 1.21 and 0.16, respectively.

Figure 10. Hedge ratios on MICEX (31 December 2007 onwards)

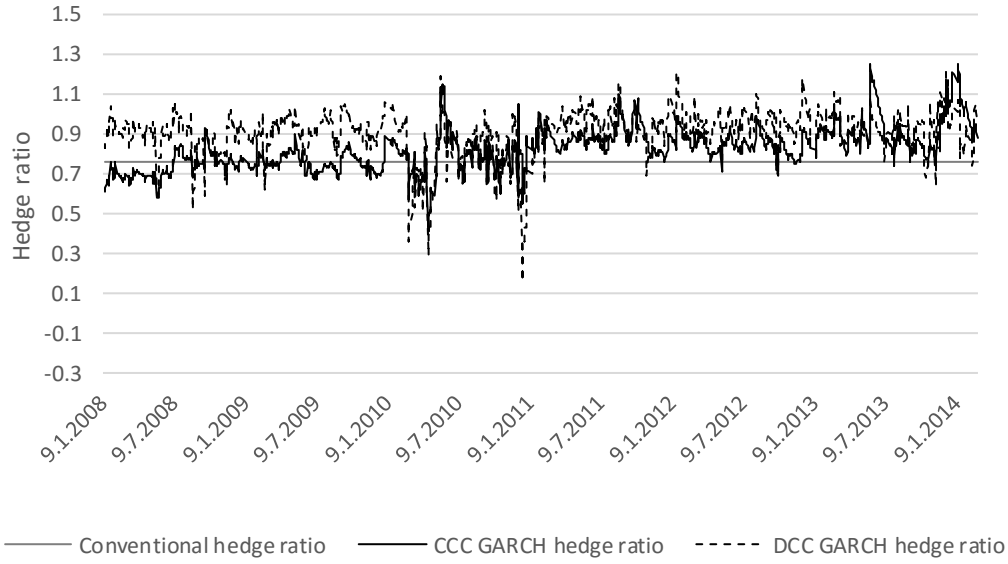


Figure 11 provides an information about the hedge ratios in India (Sensex). The Indian market gives a similar conventional hedge ratio than the other markets (excl. Russia) and was close to one (0.94). During the full sample period the highest dynamic hedge ratios were 1.10 and 1.17 for the CCC GARCH and DCC GARCH, whereas at their lowest, they were 0.82 and 0.21. Following the same pattern than most of the other markets, the CCC GARCH hedge ratios were much more stable than the DCC GARCH ratios.

Figure 11. Hedge ratios on Sensex (full period)

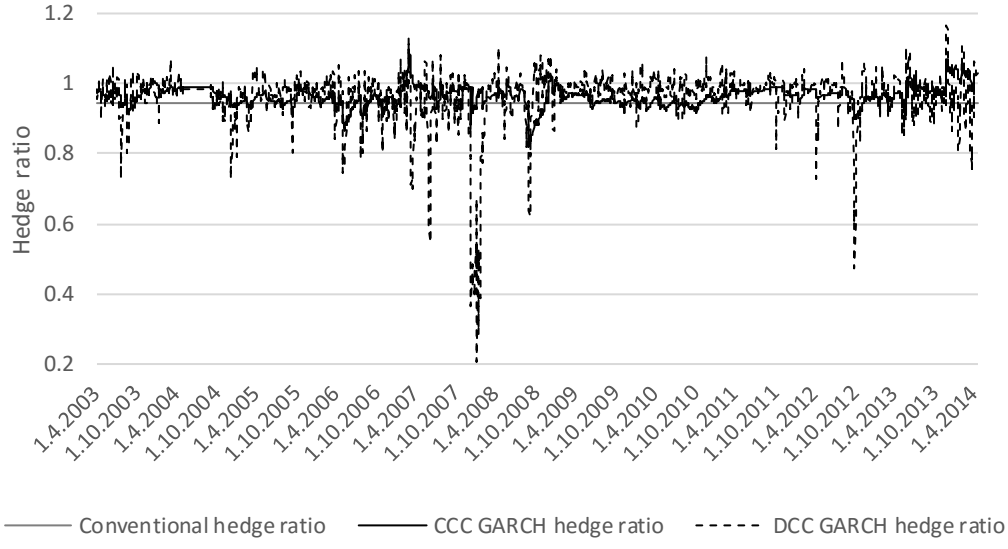
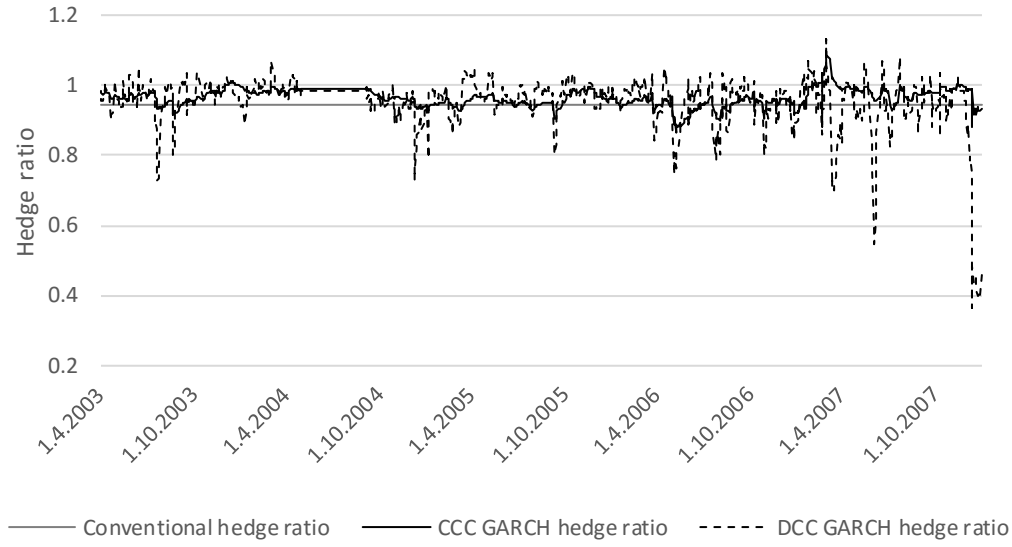


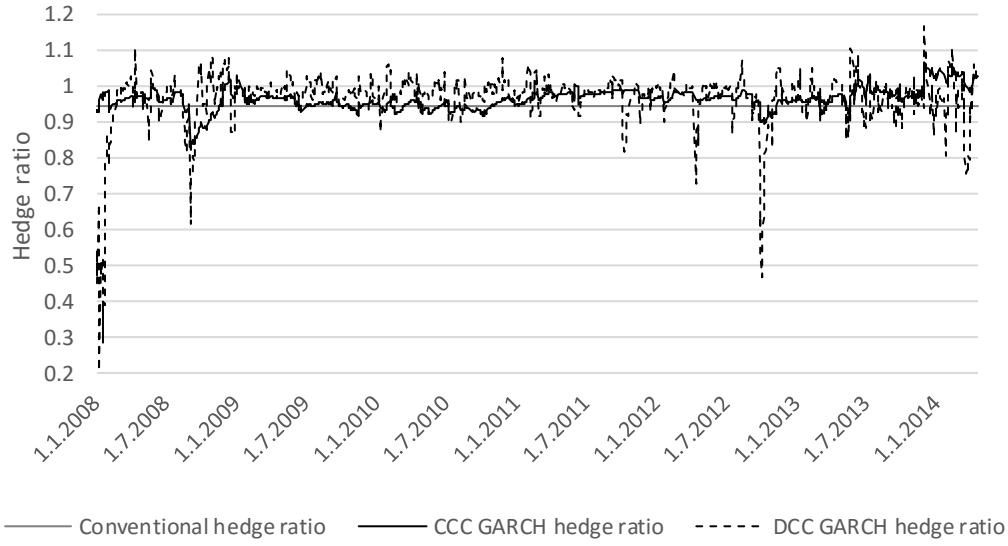
Figure 12 shows that the CCC GARCH hedge ratio was the most stable from all the markets studied during the sub-period from April 2003 to December 2007, whereas the DCC GARCH hedge ratio showed significant variations during the time period and especially at the end of 2007. The highest hedge ratios were 1.10 and 1.13 for the CCC GARCH and DCC GARCH, whereas the lowest were 0.87 and 0.36, respectively.

Figure 12. Hedge ratios on Sensex (before 31 December 2007)



The same pattern continued in the latter sub-period as the CCC GARCH hedge ratios remained rather stable and DCC GARCH hedge ratios varied more. The highest CCC GARCH and DCC GARCH hedge ratios were 1.07 and 1.17 and the lowest 0.82 and 0.21, respectively.

Figure 13. Hedge ratios on Sensex (31 December 2007 onwards)



Overall, the hedge ratio graphs provide evidence that during the Great Moderation the hedge ratios varied less than during and after the financial crisis indicating that the markets did not anticipate significant risks in the early years of the 21 century. In addition, with the exception of the Russian markets, the conventional hedge ratio was close to one as can be expected if one wants to minimise the daily variance of the portfolio. Hence, we expect that the variance reductions compared to the conventional model are small as we cannot observe large long-term deviations from the conventional level. We also expect that the dynamic hedging models work best in the Russian market as the conventional hedge ratio was lower, leaving more room to reduce the variance.

4.3 Effectiveness of in-sample hedging

This section provides the results of in-sample hedging performance of the different models. A hedge portfolio is composed of a spot asset and δ units of futures and is calculated as indicated in the Equation 47. Table 6 summarises the results of the in-sample hedging performance of the different models.

Table 6 shows that the variance reductions over the conventional model are negligible at best and in most cases negative. Panel A presents the results with the contemporaneous prices. The variance reduction varies between -10.49 and 0.74 percent. For JSE Top 40, the DCC GARCH model underperformed conventional model by 10.49% while the greatest variance reduction (0.74%) was also with the DCC GARCH model on Sensex portfolio. It can be also noted that while the DCC GARCH on South African market underperformed greatly, the CCC GARCH provides approximately the same variance as the conventional model. We cannot observe any trend showing that either of the dynamic models would provide a greater variance reduction in all markets.

The same trend continues when one-day holding period is considered (Panel B) even though slightly more markets show the variance reduction compared to the conventional model. The variance reduction is between -7.71 and 2.36 percent. Again, JSE Top 40 has the largest underperformance (7.71%) with the DCC GARCH model and again the CCC GARCH provides much better variance reduction in the same market. For one-day holding period, the greatest variance reduction (2.36%) is documented for the Russian market with the CCC GARCH model. Panel B does not provide additional evidence that the dynamic models would reduce the portfolio variance compared to the conventional model. However, compared to the contemporaneous prices, the dynamic models provide greater variance reduction, except for Ibovespa where variances of both dynamic models increased, when one-day holding period was employed.

Table 6. Comparison of the Effectiveness of In-Sample Hedging

Note. The effectiveness of the hedging models is evaluated with the contemporaneous, one-day-later, and two-day-later prices, for which the respective results are reported in Panels A, B and C. Computation of the variance in portfolio returns is carried out by Equation (47) and all reported values of variance have been multiplied by 10^4. The variance reduction over the conventional method is given as $-(\sigma_j^2 - \sigma_{conventional}^2)$. The numbers in the parentheses, in percentage terms, are given as

$$-100(\sigma_j^2 - \sigma_{conventional}^2)/\sigma_{conventional}^2$$

	Portfolio Variance					Variance reduction over the conventional model (%)				
Models	CSI 300	Ibovespa	JSE Top 40	MICEX	Sensex	CSI 300	Ibovespa	JSE Top 40	MICEX	Sensex
Panel A: $var(s_t - \delta^* f_t)$										
Conventional	0.1780	0.2476	0.0956	1.6105	0.1357					
CCC GARCH	0.1853	0.2464	0.0958	1.6174	0.1363	-0.0072 (-4.06)	0.0012 (0.48)	-0.0002 (-0.17)	-0.0068 (-0.43)	-0.0007 (-0.49)
DCC GARCH	0.1804	0.2543	0.1056	1.6048	0.1347	-0.0024 (-1.35)	-0.0067 (-2.72)	-0.0100 (-10.49)	0.0057 (0.36)	0.0010 (0.74)
Panel B: $var(s_{t+1} - \delta^* f_{t+1})$										
Conventional	0.1766	0.2476	0.0956	1.6043	0.1357					
CCC GARCH	0.1844	0.2474	0.0950	1.5665	0.1365	-0.0077 (-4.38)	0.0002 (0.09)	0.0006 (0.61)	0.0378 (2.36)	-0.0008 (-0.56)
DCC GARCH	0.1788	0.2611	0.1030	1.6084	0.1338	-0.0021 (-1.20)	-0.0134 (-5.42)	-0.0074 (-7.71)	-0.0041 (-0.26)	0.0019 (1.41)
Panel C: $var(s_{t,t+2} - \delta^* f_{t,t+2})$										
Conventional	0.1765	0.2477	0.0957	1.6048	0.1358					
CCC GARCH	0.1850	0.2472	0.0953	1.5787	0.1367	-0.0085 (-4.82)	0.0005 (0.20)	0.0003 (0.36)	0.0261 (1.62)	-0.0009 (-0.68)
DCC GARCH	0.1805	0.2596	0.1007	1.6503	0.1356	-0.0040 (-2.28)	-0.0119 (-4.78)	-0.0050 (-5.24)	-0.0455 (-2.83)	0.0002 (0.13)

When moving to two-day holding period (Panel C), the same pattern recurs where no significant variance reduction compared to the conventional model is observed. This time the variance reduction is between -5.24 and 1.62 percent. Once again, the largest underperformance was with the DCC GARCH in JSE Top 40 even though the underperformance decreased from the one-day holding period and the CCC GARCH provides much better performance and reduction to the conventional model. With the two-day holding period, the CCC GARCH for Russian market gave again the largest variance reduction (1.62%) even though a slightly lower reduction than in the one-day holding period.

The effectiveness of in-sample hedging can be summarised by stating that in most of the markets the dynamic hedging models could not provide significant variance reductions to the conventional model. Nevertheless, the dynamic hedging models provided the variance reduction at least in one of the examination periods for each market with the exception of CSI 300. Underperformance of CSI 300 can be partly due the short estimation period as the length of the period was only approximately 39% of average length of estimation periods in the other markets.

4.4 Effectiveness of out-of-sample hedging

Table 7 provides the effectiveness of out-of-sample hedging and picture is largely the same than in the Table 6. Thus, it can be stated again that the variance reductions are mostly negligible or slightly negative and varies between -5.26 and 23.91 percent. For the out-of-sample hedging the largest underperformance was on Chinese market (-5.26%) and the largest outperformance on Russian market (23.91%).

Table 7. Comparison of the Effectiveness of Out-of-Sample Hedging

Note. The method adopted to determine the series of out-of-sample hedge ratios involves rolling over a particular number of samples (80% of the whole range). The variation in the portfolio returns is evaluated using the spot and futures returns that are observed at the forecasted date. Computation of the variance in the portfolio returns is carried out by Equation (47) and all reported values of variance have been multiplied by 10^4. The variance reduction over conventional method is given as $-(\sigma_j^2 - \sigma_{conventional}^2)$. The numbers in the parentheses, in percentage terms, are given as

$$-100(\sigma_j^2 - \sigma_{conventional}^2)/\sigma_{conventional}^2$$

	Portfolio Variance					Variance reduction over the conventional model (%)				
Models	CSI 300	Ibovespa	JSE Top 40	MICEX	Sensex	CSI 300	Ibovespa	JSE Top 40	MICEX	Sensex
$var(s_{t+1} - \delta_{t+1\|t} * f_{t+1})$										
Conventional	0.1547	0.3070	0.1225	0.8421	0.1213					
CCC GARCH	0.1629	0.3073	0.1215	0.6407	0.1214	-0.0081 (-5.26)	-0.0003 (-0.10)	0.0010 (0.79)	0.2014 (23.91)	-0.0001 (-0.11)
DCC GARCH	0.1587	0.3142	0.1242	0.6524	0.1202	0.0040 (-2.56)	-0.0072 (-2.36)	-0.0017 (-1.37)	0.1897 (22.53)	0.0011 (0.93)

The significantly large variance reduction on Russian market was anticipated as stated in Section 4.2. The lowest conventional hedge ratio left room for the larger variance reduction and only few large daily deviations provided a good starting point for the hedge ratio forecasting. On the other hand, the dynamic models underperformed the conventional model for every tested period on Chinese market. We can assume that this is due the short estimation period as stated above.

When comparing our results to those in the earlier literature we can observe that regardless of the underlying assets, the hedge ratios are close to one (see e.g. Baillie & Myers 1991; Kroner & Sultan 1993; Brooks & Chong 2001; Choudry 2003; Hsu et al. 2008 and Park & Jei 2010). When considering the hedging performance, the results from the previous studies are ambiguous as some have found that the dynamic models outperform the static model (e.g. Baillie & Myers 1991; Brooks & Chong 2001 and Choudry 2003). However, most of the recent studies (e.g. Lien & Tse 2002; Hsu et al. 2008; Sultan & Hassan 2008 and Park & Jei 2010) have not found evidence that the dynamic models would significantly outperform the static one or the benefits seemed too minimal to warrant the efforts. Especially, when considering the out-of-sample results the dynamic models does not provide better forecasting ability. Interestingly, Brooks & Chong (2001) found that rather simple EWMA model outperforms more advanced GARCH models.

When making a comparison between the studies it is essential to remember that they have been conducted with variety of different assets ranging from the commodities to currencies. As all markets have their own idiosyncrasies, the results can vary. In addition, when comparing the results presented above with the earlier literature it is noteworthy that none of the previous studies have not been implemented with a similar data set, namely the stock-index futures on BRICS countries.

5 CONCLUSIONS

This Master's Thesis analysed the effectiveness of the different hedging models on BRICS (Brazil, Russia, India, China, and South Africa) countries during the 21st century. The hedging performance was studied by comparing two different dynamic hedging models to conventional OLS regression based model. The dynamic hedging models being tested were Constant Conditional Correlation (CCC) and Dynamic Conditional Correlation (DCC) GARCH(1,1) with Student's *t*-distribution. In order to capture both the Great Moderation and the latest financial crisis, the sample period from 2003 to 2014 was employed.

We obtained daily data from the stock market indices and corresponding index-futures. To be able to analyse returns from the time-series the logarithmic transformation was done for all the data. We first provided descriptive statistics of the stock indices to show that some stylised facts (such as skewness, kurtosis and non-normality of the returns) are present in all indices, implying that the unconditional distributions of the spot and futures returns are asymmetric, fat-tailed and non-Gaussian. Then we provided the unit root and cointegration tests and saw that the spot and futures prices for stock indices are cointegrated and the error-correction terms should be considered in the model specifications. After that, we estimated the models and calculated the time-varying hedge ratios. To compare possible variance reductions, we formed the portfolios by multiplying the futures returns with the hedge ratios and deducting these from the spot returns. After the construction of the portfolios, we calculated the portfolio variances. The variance reductions were calculated for the in-sample data and the effectiveness of the hedge ratio forecasts were evaluated on the basis of out-of-sample variance reductions. For in-sample hedging we used the contemporaneous hedge ratios and then extended the holding period of the portfolios to one day and two days.

For the effectiveness of in-sample hedging we found ambiguous results as most of the variance reductions were negligible or slightly negative. The results imply that the dynamic hedging models does not outperform the static one with the daily data. However, the results showed a great variability as we found both large negative and positive but singular variance reductions. For the Chinese market the dynamic models showed the negative variance reduction to the contemporaneous as well as one-day and two-day holding periods. Other markets did not show systematically under- or outperformance. Similar results were found when testing the effectiveness of out-of-sample hedging. Again, the dynamic models underperformed with the Chinese market data but the Russian market showed a very significant variance reduction over the conventional model.

When compared the results to the previous studies, we can state that they are parallel to ours even though the assets and markets examined are different. The result are divided between studies as some authors found evidence that the dynamic hedging models outperform conventional one whereas most of the newer studies showed ambiguous results about the hedging performance of the dynamic models.

Results show that the portfolio variance can be somewhat reduced from that of the conventional model in some markets but not systematically. In addition, we did not find evidence that the DCC GARCH would reduce portfolio variance compared to the CCC GARCH. Thus, we conclude that even though the conventional hedge can be costly (in sense of the higher portfolio variance) in some cases, it is still more cost efficient as the daily portfolio adjustment provides large transaction costs. The result could be different if the study had conducted with weekly or monthly adjustments.

The results of this Master's Thesis has implications for the academics and practitioners. This Thesis has contributed to the scarce literature on the

effectiveness of dynamic hedging stock-index futures. As we do not know similar studies with the BRICS market data, this Thesis provides valuable information for the academics and practitioners alike. As estimation of the dynamic hedging is more difficult than the conventional hedging, practitioners should carefully consider whether the dynamic hedging is beneficial. However, the picture could be different with less frequent portfolio adjustment or if estimation of the dynamic hedging were done with the different probability distributions.

For the future research, the less frequent portfolio adjustment would be interesting to study. In addition, the different probability distributions used for the estimation would be interesting as a recent studies[17] have shown that the stock markets most likely does not follow normal (or Gaussian) distribution but instead the stable distribution with stability parameter unequal to two (i.e. $\alpha \neq 2$). The third line of new studies would be to employ a new copula-based GARCH models. Lastly, the study of the hedging models during the non-continuous movements would provide an information concerning the risk of portfolio hedging.

[17] Mandelbrot (1963) was the first to point this out and Fama (1963) found similar evidence. For more comprehensive description, see, e.g., Sornette (2004) or Mantegna et al. (2007) and their bibliography. In addition, Taleb (2010) provides non-technical introduction. Nolan provides more technical information about stable distributions and their applications in his website (http://academic2.american.edu/~jpnolan/stable/stable.html). Especially useful are Nolan (2015a, 2015b).

REFERENCES

Ang, A. & Chen, J. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63, 443-494.

Baele, L. & Inghelbrecht, K. 2009. Time-Varying Integration and International Diversification Strategies, *Journal of Empirical Finance* 16, 368-387.

Baillie, R.T. & Myers, R.J. 1991. Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics* 6, 109-124.

Bauwens, L., Laurent, S. & Rombouts, J. 2006. Multivariate GARCH models: A survey. *Journal of Applied Econometrics* 21, 79-109.

Bekaert, G. & Harvey, C. 2000. Foreign Speculators and Emerging Equity Markets. *Journal of Finance* 55, 565-613.

Benet, B.A. 1992. Hedge period length and ex-ante futures hedging effectiveness: The case of foreign-exchange risk cross hedges. *Journal of Futures Markets* 12, 163-175.

Bernanke, B. 2004. The Great Moderation. Speech. The Federal Reserve Board, Washington, DC. February 20, 2004.

Black. F. 1976. Studies of stock market volatility changes. *1976 Proceedings of the American Statistics Association, Business and Economic Statistics Section*, 177-181.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.

Bollerslev, T. 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. *Review of Economics and Statistics* 72, 498-505.

Bollerslev, T. & Engle, R.F. 1993. Common persistence in conditional variances. *Econometrica* 61, 167-186.

Bollerslev, T., Engle, R.F. & Wooldridge, JM. 1988. A capital asset pricing model with time varying covariances. *Journal of Political Economy* 96, 116-131.

Brailsford, T., Corrigan, K. & Heaney, R. 2001. A comparison of measures of hedging effectiveness: A case study using the Australian all ordinary share price index futures contract. *Journal of Multinational Financial Management* 11, 465-483.

Brooks, C. 2008. Introductory Econometrics for Finance. 2nd Ed. New York, Cambridge University Press.

Brooks, C. & Chong, J. 2001. The cross-currency hedging performance of implied versus statistical forecasting models. *Journal of Futures Markets* 21, 1043-1069.

Brooks, R. & Del Negro, M. 2004. The Rise in Comovement Across National Stock Markets: Market Integration or IT Bubble? *Journal of Empirical Finance* 11, 659-680.

Butterworth, D. & Holmes, P. 2000. Ex ante hedging effectiveness of UK stock index futures contracts: Evidence for the FTSE 100 and FTSE Mid 250 contracts. *European Financial Management* 6, 441-457.

Cecchetti, S.G., Cumby, R.E. & Figlewski, S. 1988. Estimation of optimal futures hedge. *Review of Economics and Statistics* 70, 623-630.

Choudry, T. 2003. Short-run deviations and optimal hedge ratio: evidence from stock futures. *Journal of Multinational Financial Management* 13, 171-192.

Choudry, T. 2004. The hedging effectiveness of constant and time-varying hedge ratios using three Pacific Basin stock futures. *International Review of Economics and Finance* 13, 371-385.

Driessen. J. & Laeven, L. 2007. International Portfolio Diversification Benefits: Cross-Country Evidence from a Local Perspective. *Journal of Banking and Finance* 31, 1693-1712.

Ederington, L.H. 1979. The hedging performance of the new futures markets. *Journal of Finance* 34, 157-170.

Engle, R.F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50 (4), 987-1007.

Engle, R.F. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20, 339-350.

Engle, R.F. & Granger, C.W.J. 1987. Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55, 251-276.

Engle, R.F. & Kroner, K.F. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122-150.

Engle, R.F. & Ng, V.K. 1993. Measuring and Testing the Impact of News on Volatility. *Journal of Finance* 48 (5), 1749-1778.

Engle, R.F., Ng, V.K. & Rothschild, M. 1990. Asset pricing with a factor-ARCH covariance structure: empirical estimates for treasury bills. *Journal of Econometrics* 45, 213-238.

Engle, R.F. & Sheppard, K. 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. *Working Paper 8554. National Bureau of Economic Research.* Cambridge: Massachusetts.

Errunza, V., Hogan, K. & Hung, M-W. 1999. Can the Gains from International Diversification Be Achieved without Trading Abroad? *Journal of Finance* 54, 2075-2107.

Fama, E.F. 1963. Mandelbrot and the Stable Paretian Hypothesis. *Journal of Business*, 36 (4), 420-429.

Figlewski, S. 1984. Hedging performance and basis risk in stock index futures. *Journal of Finance* 39, 657-669.

Fisher, R.A. 1932. Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd.

Floros, C. & Vougas, D.V. 2004. Hedge ratios in Greek stock index futures market. *Applied Financial Economics* 14, 1125-1136.

Garcia, P., Roh, J-S. & Leuthold, R.M. 1995. Simultaneously determined, time-varying hedge ratios in the soybean complex. *Applied Economics* 27, 1127-1134.

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. *Journal of Finance* 48, 1779-1801.

Goetzmann, W., Li, L. & Rouwenhorst, G. 2005. Long-Term Global Market Correlations. *Journal of Business* 78, 1-38.

Hansen, B. 1994. Autoregressive conditional density estimation. *International Economic Review* 35, 705-730.

Hentschel, L. 1995. All in the family – Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* 39, 71-104.

Higgins, M. & Bera, A. 1992. A class of nonlinear ARCH models. *International Economic Review* 33, 137-158.

Hsu, CC., Tseng, CP. & Wang, YH. 2008. Dynamic Hedging with Futures: A copula-based GARCH model. *Journal of Futures Markets* 28 (11), 1095-1116.

Hull, J. 2012. Options, futures, and other derivatives. 8th ed. Boston: Pearson Education.

Joe, H. 1997. Multivariate Models and Dependence Concepts. London: Chapman & Hall.

Johnson, L. 1960. The theory of hedging and speculation in commodity futures. *Review of Economic Studies* 27, 139-151.

Kim, C. & Nelson, C.R. 1999. Has the U.S. economy become more stable? A Bayesian approach based on Markov-switching model of the business cycle. *Review of Economics and Statistics* 81 (4), 608-616.

Kim, E.H. & Singal, V. 2000. Stock market openings: Experience of emerging economies. *Journal of Business* 73 (1), 25-66.

Kizys, R. & Pierdzioch, C. 2009. Changes in the International Comovement of Stock Returns and Asymmetric Macroeconomics Shocks. *Journal of International Financial Markets, Institutions and Money* 19, 289-305.

Kroner, K.F. & Sultan, J. 1991. Exchange Rate Volatility and Time Varying Hedge Ratios. Rhee, S.G. and Chang, R.P., eds *in Pacific-Basin Capital Markets Research, Volume II*. Amsterdam: Elsevier Science Publishers, 397-412.

Kroner, K.F. & Sultan, J. 1993. Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures. *Journal of Financial and Quantitative Analysis* 28 (4), 535-551.

Ku, Y-H.H., Chen, H-C. & Chen, K-H. 2007. On the application of the dynamic conditional correlation model in estimating optimal time-varying hedge ratios. *Applied Economics Letters* 14 (7), 503-509.

Lee, C.F., Bubnys, E.L. & Lin, Y. 1987. Stock index futures hedge ratios: Tests on horizon effects and functional form. *Advances in Futures and Options Research* 2, 291-311.

Levy, H. & Sarnat, M. 1984. Portfolio and Investment Selection: Theory and Practice. New Jersey: Prentice-Hall.

Lien, D. & Zhang, M. 2008. A Survey of Emerging Derivatives Markets. *Emerging Markets Finance & Trade* 44 (2), 39-69.

Lien, D. & Tse, Y.K. 2002. Some recent developments in futures hedging. *Journal of Economic Surveys* 16 (3), 357-396.

Lim, K-P. & Brooks, R. 2011. The evolution of stock market efficiency over time: A survey of empirical literature. *Journal of Economic Surveys* 25 (1), 69-108

Longin, F. & Solnik, B. 1995. Is the Correlation in International Equity Returns Constant: 1960-1990? *Journal of International Money and Finance* 14, 3-26.

Longin, F. & Solnik, B. 2001. Extreme correlation of international equity markets. *Journal of Finance* 56, 649-676.

Mandelbrot, B. 1963. The Variation of Certain Speculative Prices. *Journal of Business* 36, 394-419.

Mantegna, R.N. & Stanley, E.H. 2007. An Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge: Cambridge University Press.

Markowitz, H. 1952. Portfolio Selection. *Journal of Finance* 7 (1), 77-91.

Marmer, H.S. 1986. Portfolio Model Hedging with Canadian Dollar Futures: A Framework for Analysis. *Journal of Futures Markets* 6, 84-92.

Munk, C. 2013. Financial Asset Pricing Theory. Oxford: Oxford University Press.

Myers, R.J. 1991. Estimating time-varying optimal hedge ratios on futures markets. *Journal of Futures Markets* 11, 39-53.

Myers, R.J. & Thompson, S.R. 1989. Generalized optimal hedge ratio estimation. *American Journal of Agricultural Economics* 71, 858-868.

Nelsen, R.B. 1999. An Introduction to Copulas. New York: Springer-Verlag.

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

Nolan, J.P. 2015a. StableBibliography. [PDF]. <http://academic2.american.edu/~jpnolan/stable/StableBibliography.pdf>, accessed 8 November 2015.

Nolan, J.P. 2015b. Stable Distributions – Models for Heavy Tailed Data. In progress, Chapter 1 online at academic2.american.edu/$\sim$jpnolan. Boston: Birkhauser.

Omran, M. & Farrar. S.V. 2006. Tests of weak form efficiency in the Middle East emerging markets. *Studies in Economics and Finance* 23 (1), 13-26.

Park, S.Y. & Jei, S.Y. 2010. Estimation and hedging effectiveness of time-varying hedge ratio: Flexible bivariate GARCH approaches. *Journal of Futures Markets* 30 (1), 71-99.

Park, T.H. & Switzer, L.N. 1995. Time-varying distributions and optimal hedge ratios for stock index futures. *Applied Financial Economics* 5, 131-137.

Patton, A.J. 2002. Applications of Copula Theory in Financial Econometrics. Unpublished Ph.D. dissertation, University of California, San Diego.

Patton, A.J. 2006. Modeling asymmetric exchange rate dependence. *International Economic Review* 47, 527-556.

Rosenblatt, M. 1952. Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics* 23, 470-472.

Schwert, W. 1989. Why does stock market volatility change over time? *Journal of Finance* 44, 1115-1153.

Sentana, E. 1995. Quadratic ARCH Models. *Review of Economic Studies* 62, 639-661.

Sornette, D. 2004. Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton and Oxford: Princeton University Press.

Steinherr, A. 2000. Derivatives: the wild beast of finance: a path to effective globalisation? Chichester: Wiley.

Stock, J.H. & Watson, M.J. 2002. Has the Business Cycle Changed and Why? NBER Macroeconomics Annual 2002, Volume 17. Massachusetts: MIT Press.

Sultan, J. & Hasan, M.S. 2008. The effectiveness of dynamic hedging: evidence from selected European stock index futures. *European Journal of Finance* 14 (6), 469-488.

Taleb, N.N. 2010. The Black Swan: The Impact of the Highly Improbable. London: Penguin Group.

Taylor, S. 1986. Modelling financial time series. New York: Wiley.

The Splendors and Miseries of Martingales. *Electronic Journal for History of Probability and Statistics* 5 (1).

Tong, W.H.S. 1996. An examination of dynamic hedging. *Journal of International Money and Finance* 15 (1), 19-35.

Ville, J. 1939. Étude critique de la notion de collectif. Monographies des Probabilités I (in French). Paris: Gauthier-Villars.

Williams, D. 1991. Probability with Martingales. Cambridge: Cambridge University Press.

Zakoian, J-M. 1994. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* 18 (5), 931-955.