ESSAYS IN FINANCIAL RISK MANAGEMENT

by

Ibrahim Ergen

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Mahmoud Amin El-Gamal, Chair
Professor of Economics and Statistics

Kenneth Medlock III,
James A. Baker, III, and Susan G. Baker
Fellow in Energy and Resource Economics and Adjunct Professor of Economics

Barbara Ostdiek,
Associate Professor of Finance and Statistics

HOUSTON, TEXAS

JANUARY 2010
ABSTRACT

ESSAYS IN FINANCIAL RISK MANAGEMENT

by

IBRAHIM ERGEN

In Chapter 1, the usefulness of Extreme Value Theory (EVT) methods, GARCH models, and skewed distributions in market risk measurement is shown by predicting and backtesting the one-day-ahead VaR for emerging stock markets and the S&P 500 index. It has been found that the conventional risk measurement methods, which rely on normal distribution assumption, grossly underestimate the downside risk.

In Chapter 2, the dependence of the extreme losses of the emerging stock market indices is analyzed. It is shown that the dependence in the tails of their loss distributions is much stronger than that implied by a correlation analysis. Economically speaking, the benefits of portfolio diversification are lost when investors need them most. The standard methodology for bivariate extremal dependence analysis is slightly generalized into a multi-asset setting. The concept of hidden extremal dependence for a multi-asset portfolio is introduced to the literature and it is shown that the existence of such hidden dependence reduces the diversification benefits.

In Chapter 3, the mechanisms that drive the international financial contagion are discussed. Trade competition and macroeconomic similarity channels are identified as significant drivers of financial contagion as measured by extremal dependence.

In Chapter 4, the determinants of short-term volatility for natural gas futures are investigated within a GARCH framework augmented with market fundamentals. New findings include the asymmetric effect of storage levels and maturity effect across seasons. More importantly, I showed that, the augmentation of GARCH models with market fundamentals improves the accuracy of out-of-sample volatility forecasts.
ACKNOWLEDGEMENTS

Many people, in one way or another, have contributed to the completion of this dissertation to whom I want to express my gratitude.

First and foremost, my deepest gratitude must go to my advisor, Dr. Mahmoud A. El-Gamal. He encouraged me to develop independent thinking and research skills and patiently provided the vision and advice necessary to be successful in the doctoral program. He was always aware of a supervisor’s professional responsibilities. During the hard times of graduate study, he became a friend and a counselor and put me back on track. I believe the completion of this dissertation would never be possible without his support and encouragement. It is impossible for me to express my appreciation for him with words.

I am also grateful to the dissertation committee members Dr. Kenneth Medlock and Dr. Barbara Ostdiek who aroused my curiosity and shared with me their deep knowledge on energy commodities. They put serious effort into reading this dissertation under significant time constraints. Their thoughtful criticisms and suggestions improved the manuscript and ultimately made this a better work.

Dr. Bryan Brown, Dr. Robin Sickles and Dr. Yoosoon Chang taught me the most contemporary statistical and econometric methods and as such they deserve to be recognized among other professors. I am also thankful to Dr. Erkan Ture, Dr. Mert Demir, Dr. Guldal Buyukdamgaci and Dr. Serol Bulkan of Marmara University for encouraging me to undertake graduate studies in the US.

Many other people provided assistance in obtaining academic resources and the database which was necessary for undertaking this study. Specifically, I would like to express my thanks to Peggy Shaw, the director of Business Information Center at Rice University and Billy Paul Coxsey from her team.

I am grateful to Ms. Altha Rodgers, the graduate secretary, David Blazek, Jennifer
Rothal, Ruben Juarez, Osman Nal, Jiaqi Hao, Sinan Ertemel, Levent Kutlu, Islam Rizvanoglu, Baris Esmerok, Emre Dogan, Burcu Cigerli and other PhD students at Rice University Economics Department for creating a family like environment which made learning and working here fun.

The Turkish-American community has been very supportive and friendly over the years. I am particularly thankful to my academic brother Hulusi Inanoglu, Kadim abi, Cafer abi, Cahit hocam, Can hocam, Nurullah hocam, Ozgur hocam, Indrit hocam, Hulusi hocam, Kaptan Feryat, Fatih Golgi, Ibrahim Bakir, Kadir Bahar, Mustafa Cesur and the Legend, Emre Kocatulum. I am also thankful to some wonderful friends in Turkey: Bilal Cinar, Bahadir Icel, Esat Hizir and all other group Bizimtekne members + Ozan KK. from Marmara University and 6FenA99 class of Yamanlar Science School.

This dissertation would never be complete were it not the support of my beloved family. My parents, Ismet and Zehra Ergen and my sisters Ayse, Betul and Rukiye were always positive, encouraging and supportive at hard times. I always felt the power of their prayers on my work. They also made a big sacrifice by agreeing on my leaving home and coming to the USA for graduate study. I am also grateful to the love of my life, my wife Aysegul, who has given a continuous support and has provided love, comfort and encouragement since the day we met.

Finally, I am very grateful to the One who blessed me with all of the above: supportive and encouraging professors, very nice friends and a loving family. All beauties in life are due to Him.
## Contents

1 Extreme Value Theory and VaR Prediction for Emerging Market Stock Indices ........................................... 1

1.1 Introduction ............................................................. 2

1.2 Data and Statistical Analysis ........................................ 6

1.3 Extreme Value Theory ................................................. 21

1.4 VaR Estimation .......................................................... 24

1.4.1 Historical Simulation ............................................... 24

1.4.2 Normal Distribution Model ....................................... 25

1.4.3 Student’s t-Distribution Model .................................. 25

1.4.4 Skewed Normal Distribution Model ......................... 26

1.4.5 Skewed t-Distribution Model .................................... 27

1.4.6 EVT Model ............................................................ 28

1.4.7 GARCH Model ....................................................... 30

1.4.8 GARCH-EVT Model ................................................. 31

1.4.9 GARCH-HS Model .................................................. 32

1.4.10 EWMA Model ....................................................... 33

1.5 Backtesting ............................................................ 34

1.6 Conclusion .............................................................. 40
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Hidden Extremal Dependence and Implications for Portfolio Risk</td>
<td>42</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>43</td>
</tr>
<tr>
<td>2.2</td>
<td>Extremal Dependence</td>
<td>47</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Hill's Estimation</td>
<td>48</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Extremal Dependence Measures</td>
<td>49</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Extremal Dependence Analysis of Emerging Markets</td>
<td>53</td>
</tr>
<tr>
<td>2.3</td>
<td>Extremal Dependence and Diversification Benefits</td>
<td>56</td>
</tr>
<tr>
<td>2.4</td>
<td>Hidden Asymptotic Dependence and Diversification Benefits</td>
<td>63</td>
</tr>
<tr>
<td>2.5</td>
<td>Conclusion</td>
<td>69</td>
</tr>
<tr>
<td>3</td>
<td>Extremal Dependence and International Financial Contagion</td>
<td>71</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>71</td>
</tr>
<tr>
<td>3.2</td>
<td>Channels of Financial Contagion</td>
<td>73</td>
</tr>
<tr>
<td>3.3</td>
<td>Testing for Trade Channel</td>
<td>76</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Existing Methodologies</td>
<td>76</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Alternative Model</td>
<td>78</td>
</tr>
<tr>
<td>3.4</td>
<td>Testing for the Macro-Similarity Channel</td>
<td>80</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusion</td>
<td>81</td>
</tr>
<tr>
<td>4</td>
<td>Fundamentals and Natural Gas Volatility Dynamics: Predictive Power of Augmented GARCH Models in Volatility and Risk Forecasting</td>
<td>82</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>83</td>
</tr>
<tr>
<td>4.2</td>
<td>Data and Statistical Analysis</td>
<td>87</td>
</tr>
<tr>
<td>4.2.1</td>
<td>A First Look at the Natural Gas Volatility</td>
<td>91</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Levene Tests for Variance Equality</td>
<td>95</td>
</tr>
<tr>
<td>4.3</td>
<td>Empirical Model and Estimation Results</td>
<td>98</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Day-of-the-Week, Seasonality and Maturity Modeling</td>
<td>99</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Summary Statistics for S&P/IFCI Returns .......................... 13
1.2 VaR Backtesting Tables ................................................. 38
1.3 Scoring Table for VaR Prediction Models ............................ 38

2.1 \(\bar{\chi}\) Estimates and Probability Values for Daily S&P/IFCI Losses ................................. 55
2.2 \(\chi\) Estimates and Standard Errors for Daily S&P/IFCI Losses .............................................. 55
2.3 Correlation Matrix for Daily S&P/IFCI Losses ........................ 56
2.4 Diversification Benefits at \(\alpha = 0.999^{th}\) Quantile .......................... 58
2.5 Average Diversification Benefits for Different Tail Dependence Structures .................................. 59
2.6 Marginal Risks Ratio at \(\alpha = 0.999^{th}\) Quantile .......................... 61
2.7 Diversification Benefits and Extremal Dependence ........................ 62
2.8 Diversification Benefits and Correlation Coefficient ........................ 63
2.9 Average Diversification Benefits for Different Dependence Structures in Three Asset Setting ............................ 67

3.1 Regression Results of Extremal Dependence on Trade Competition .............................. 79
3.2 Regression Results Including Macro-Similarity Control Variables .............................. 81

4.1 Summary Statistics for Nearby Month Futures Returns ........................................ 88
4.2 Standard Deviation of Daily Returns Broken into Groups ........................................ 92
<table>
<thead>
<tr>
<th>Table Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>Brown-Forsythe Type Levene Tests for Equality of Variances</td>
<td>97</td>
</tr>
<tr>
<td>4.4</td>
<td>GARCH Estimation Results For Nearby Month Futures Returns</td>
<td>101</td>
</tr>
<tr>
<td>4.5</td>
<td>GARCH Estimation Results with Storage and Weather Variables For Nearby Month Futures Returns</td>
<td>107</td>
</tr>
<tr>
<td>4.6</td>
<td>Accuracy Statitics for Out-of-Sample Forecasting</td>
<td>111</td>
</tr>
<tr>
<td>4.7</td>
<td>VaR Violations and Probability Values for Binomial Hypothesis</td>
<td>114</td>
</tr>
</tbody>
</table>
## List of Figures

1.1 Graphical Representation of VaR ............................................. 3
1.2 The Level of S&P/IFCI Price Index ................................. 9
1.3 S&P/IFCI Return Plots ...................................................... 11
1.4 Normal QQ-Plots for S&P/IFCI Returns ......................... 15
1.5 Correlograms for Raw Returns of S&P/IFCI Index .............. 18
1.6 Correlograms for Squared Returns of S&P/IFCI Index ......... 20
1.7 Densities for Several Skewed Normal Distributions with $\mu = 0$ and $\sigma = 1$ .............................. 26
1.8 Densities for Several Skewed t Distributions with $\mu = 0$ and $\sigma = 1$ ........ 28

2.1 Reduction in VaR from Diversification ............................... 57
2.2 No Reduction in VaR from Diversification ...................... 59
2.3 Densities of Diversification Benefits for Different Dependence Structures ...... 60
2.4 Portfolio Level Hidden Asymptotic Dependence .................... 64
2.5 Weights Used in the Three Asset Portfolio Problem ............... 66
2.6 Densities of Diversification Benefits in Three Asset Setting ...... 68

4.1 Working Gas In Underground Storage (a) and Its Deviation From Five Year Historical Mean (b) ................................. 90
4.2 Standard Deviations of Nearby Month Futures Returns Across Days, Seasons, Bidweeks and Storage Levels ....................... 94
LIST OF FIGURES

4.3 National Weather Shock Variables .................................. 104
4.4 Density Fits For Natural Gas and Taiwan Stock Indices .... 115
Chapter 1

Extreme Value Theory and VaR Prediction for Emerging Market Stock Indices

Summary. Using comprehensive state-of-the-art market risk models, I show the usefulness of extreme value theory (EVT) methods, generalized autoregressive conditional heteroscedasticity (GARCH) models, and skewed distributions in market risk measurement by predicting and backtesting the day-ahead Value at Risk (VaR) for emerging stock markets and the S&P 500. It is found that the conventional methods of risk measurement such as the exponentially weighted moving average (EWMA) model of RiskMetrics greatly underestimate the risk. The results indicate that EVT is the best way of modeling fat distribution tails. The GARCH-EVT model that accomplishes dynamic volatility and fat-tail modeling in a three-step procedure provides the best backtesting performance in VaR prediction.
1.1 Introduction

Emerging countries are characterized by a market economy that is in between developing and developed status. These countries frequently experience very high volatility and extreme movements in their stock markets. Examples of such highly volatile periods include the Mexican debt crisis in 1994, Asian crisis in 1997, Russian default in 1998, the Turkish banking crisis in 2001, and the Brazilian crisis in 2002. VaR, as a measure of portfolio risk, was developed as a result of this high volatility and extreme market movements since the 1990s. VaR estimates the amount that the loss of a portfolio may exceed with a specified small probability within a specified time interval. Therefore, it is a high quantile of the loss distribution. More formally,

\[ \text{VaR}_\alpha = \inf \{ x \in \mathbb{R} : F_X(x) \geq \alpha \} = F^{-1}(\alpha), \]  

where \( X \) stands for the loss, \( F_X \) is the distribution function of losses, and \( \alpha \) is the quantile at which VaR is calculated. \( F^{-1} \) is known as the generalized inverse of \( F_X \), or the quantile function associated with \( F_X \). If \( F_X \) is a monotonically increasing function, then \( F^{-1} \) reduces to the regular inverse function. Graphical representation of VaR can be seen in Figure 1.1 where \( \alpha \) is chosen to be 0.99 so that there is a 1 percent chance of incurring a loss that exceeds the \( \text{VaR}_{0.99} \).

VaR as a measure of extreme downside risk was first proposed by the RiskMetrics team operating under J.P. Morgan. However, the critical turning point for its popularity in the financial industry was the amendment made to Basel-I by the Basel Committee on Banking Supervision in 1996. Until this amendment, banks were required to hold capital only against their credit risk. In 1996, the Basel Committee announced that commercial banks have to hold regulatory capital against their mar-
1.1 Introduction

Probability Distribution of Portfolio Value

Figure 1.1: Graphical Representation of VaR

ket risk as well. Moreover, this additional capital has to cover the losses of the bank's portfolio 99 percent of the time. This was basically another way of defining VaR. Consequently, VaR became the industry standard for measuring the risk of a portfolio in the following years. Basel Committee also allowed the financial institutions to choose their market risk models freely subject to regulation and supervision from national regulating bodies. The performance and suitability of market risk models for the estimation of VaR is an important issue because underestimation of VaR results in less cash holdings than required, which in turn increases the insolvency risk. On the other hand, overestimation of VaR causes more cash holdings than required which creates inefficiency in the allocation of resources.

Another quantile-based risk measure to address is the expected shortfall or the Conditional Value at Risk (CVaR). Expected shortfall is the average of all possible VaR values beyond some high quantile of the loss distribution and is given by:

$$ES_{\alpha} = \frac{1}{1 - \alpha} \int_{\alpha}^{1} VaR_{u} du,$$  \hspace{1cm} (1.2)

It can be shown that $ES_{\alpha}$ is the expectation of the loss, given that a loss that exceeds
1.1 Introduction

the \( VaR_\alpha \) has already occurred. That is,

\[
ES_\alpha = E[X | X \geq VaR_\alpha],
\]

(1.3)

Therefore, in Figure 1.1, \( ES_{0.99} \) is the gravity center of the dashed tail region. As the definitions suggest, the expected shortfall provides an idea about the magnitude of an extreme loss in case it occurs, whereas VaR provides no information regarding this point. The expected shortfall also satisfies some ideal theoretical properties such as subadditivity (see Artzner et al. 1997). However, VaR somehow accomplished dominating the financial industry and became the widely accepted risk measure for a portfolio. Therefore, I will focus on VaR as the risk measure of a portfolio in this study.

There have been many studies on alternative ways of risk measurement with VaR. Historical simulation and variance-covariance method are fairly simple, and the most commonly used methods in practice. Historical simulation is a totally non-parametric way of VaR estimation. Sample quantile of the loss data is used as the VaR estimate in this model. Variance-covariance method assumes a distribution for asset losses that is closed under linear transformations. Using this property, the distribution of any linear portfolio of assets is still the same distribution, and VaR can be calculated easily using the quantile function of the chosen distribution. A normality assumption for the loss distribution is being made most of the time in the variance-covariance method.

It is a known fact that asset returns distributions had fat tails during the 1960s (Mandelbrot 1963). With the introduction of mathematical finance literature in the 1970s, this fact was ignored, and most risk management methodologies were developed based on the assumption of geometric Brownian motion (GBM) for asset prices. Upon
1.1 Introduction

The observation of catastrophic market events during the 1990s, the weaknesses of these models were uncovered. Awareness of rare, but devastating market events, led more researchers to point out the need to go back and study the implications of fat tails. EVT is a strong tool that addresses the modeling of fat tails. EVT provides a way of calculating the probabilities of rare events that have never been observed in the sample. Among many others, Gencay and Selcuk (2004) and McNeil and Frey (1999, 2000) exhibited the usefulness of EVT in VaR estimation.

GARCH models are a natural candidate for VaR estimation. GARCH models assume "conditional normality" instead of "normality"; that is, losses are normally distributed, but the volatility of the losses are changing over time. These models are very strong in modeling the dynamic nature of market volatility. Also, in this case, the unconditional distribution of returns is not normal but a mixture of normal distributions, which implies fatter tails than the normal distribution. However, the empirical literature suggests that GARCH models are not able to capture the tail behavior of financial data because most of the time the residuals from a GARCH fit are still fat tailed.

In order to model dynamic volatility and fat tails at the same time, McNeil and Frey (2000) suggests a three-step procedure for VaR estimation known as the GARCH-EVT model. The GARCH model is fitted to the data in the first step with quasi-maximum likelihood estimation. In the second step the EVT methods are applied to the residuals extracted from this fit and a VaR estimate is calculated for these residuals. Lastly, in the third step, a VaR estimate for the original data is calculated. Barone-Adesi, Bourgoin, and Giannapolis (1998) suggest a similar methodology for VaR estimation, but they use historical simulation on the residuals of the GARCH model in step two.

Using comprehensive state-of-the-art market risk models, I show the usefulness
of EVT methods and GARCH models as well as skewed distributions in market risk modeling. This is accomplished by predicting and backtesting the one-day-ahead VaR for a comprehensive set of emerging stock market indices and the S&P 500 index for several quantiles. For emerging market indices, a similar study was undertaken by Gencay and Selcuk (2004). This study differs from theirs in several aspects. First, the gaussian GARCH model, the GARCH-EVT model and the GARCH-HS model are included. Second, the EWMA volatility model is included in this study as a benchmark, because of its importance as the suggested model by the RiskMetrics Group. Third, skewed distributions with variance-covariance approach are included. Lastly, the data for all countries are coming from the same time period in this study. So, the results can be compared across countries as well as across models.

The remaining sections of this chapter are organized as follows. In section 1.2, the statistical analysis of the data is presented, credibility of the geometric Brownian motion assumption is questioned by means of normality tests, and the motivation for using conditional risk management by including GARCH models is presented. Section 1.3 provides the background knowledge for EVT methods. All market risk models for VaR prediction is reviewed in section 1.4 and the predictions are compared by backtesting in section 1.5. In section 1.6, conclusions are discussed.

### 1.2 Data and Statistical Analysis

Standard & Poors and International Financial Corporation (S&P/IFC) daily equity price index data\(^1\) are obtained from DataStream for 13 emerging markets: Turkey, Malaysia, Indonesia, Philippines, Korea, China, Taiwan, Thailand, Mexico, Argentina, Brazil, Chile, and Peru. The S&P/IFCI is the investable index, and is a weighted

\(^1\)Stock index data for all countries are dollar denominated prices.
average of only those equities that can be traded by international investors. The dataset runs from June 30, 1995 to November 1, 2007. This time span covers the period of the Asian financial crisis in 1997, the Russian default in 1998, the Turkish banking crisis in 2001, and the Brazilian crisis of 2002.

The level of stock indices are plotted in Figure 1.2 for all countries. From these plots, one can easily detect the effects of the Asian currency crisis in the summer of 1997 on the financial markets of these countries. The crisis started with the depreciation of the Thai baht and quickly spread to the whole region, causing from 35 to 90 percent depreciation in the equity indices of Indonesia, Malaysia, Philippines, Thailand, Taiwan, and lastly South Korea. In 2001 and 2002, Turkey and Brazil experienced a serious downturn in their financial markets together with a huge depreciation in their currencies. By the beginning of 2003, a global boom was observed in almost all emerging countries until the summer of 2007, which marks the beginning of observable effects of the subprime mortgage crisis.

The daily percentage logarithmic returns are calculated as

\[ r_t = 100 \log \left( \frac{S_t}{S_{t-1}} \right) , \tag{1.4} \]

for all countries, where \( S_t \) is the level of the equity price index at the end of the day \( t \). The logarithmic return approach is useful because the geometric Brownian motion assumption for the price process can be tested through a normality test on logarithmic returns because returns are log-normally distributed under this assumption. Working with logarithmic returns is a standard approach in the financial literature. For ease of exposition, the logarithmic returns and logarithmic losses are referred to simply as "returns" and "losses" throughout this chapter.

Returns calculated with (1.4) are plotted in Figure 1.3. The level of extreme daily
1.2 Data and Statistical Analysis

(a) Turkey

(b) Indonesia

(c) Malaysia

(d) Philippines

(e) South Korea

(f) China

(g) Taiwan

(h) Thailand
1.2 Data and Statistical Analysis

Figure 1.2: The Level of S&P/IFCI Price Index
1.2 Data and Statistical Analysis

(a) Turkey

(b) Indonesia

(c) Malaysia

(d) Philippines

(e) South Korea

(f) China

(g) Taiwan

(h) Thailand
Figure 1.3: S&P/IFCI Return Plots
1.2 Data and Statistical Analysis

losses reaches up to 15 percent in Brazil and Mexico, 20 percent in Turkey, Malaysia, and South Korea, and even 30 percent in Argentina and 40 percent in Indonesia. This behavior in emerging markets renders them good candidates for studying the behavior of extremes values. Relatively stable emerging markets in this sense are Taiwan and Chile, which have returns between 6% and -6%. This kind of behavior is generally observed in developed markets. Returns of the S&P 500 index in Figure 1.3(n), for example, have a similar range as Chile. Volatility clustering is a common observation for all countries, but the level of volatility is obviously much higher for emerging markets than for the S&P 500.

Table 1.1 summarizes the basic statistics of the S&P/IFCI return data for all countries. Some countries display positive skewness, while others display negative skewness. The kurtosis reported is the excess kurtosis over three, the kurtosis for a normal distribution. Apparently, the loss distributions for all countries exhibit higher kurtosis than implied by a normal distribution. Taiwan and Chile have the smallest excess kurtosis, but even those are not seemingly consistent with a normality assumption.

A quantile-quantile plot (QQ plot) is a very useful graphical tool to analyze the tail behavior of a distribution and evaluate Gaussianity. A QQ plot shows the relationship between empirical quantiles of the data and theoretical quantiles of some reference distribution. The attractiveness of the QQ plots come from the statistical result that quantiles of two distributions from the same parametric family are linear transformations of each other. Therefore, linearity of the QQ plot reveals that the data are coming from the family of the reference distribution but possibly with different parameters. Further, an S shape reveals that the data have fatter tails than the reference distribution while an inverse S shape reveals that the data have thinner tails than the reference distribution assuming the theoretical quantiles of the reference
Table 1.1: Summary Statistics for S&P/IFCI Returns

distribution are on the horizontal axis.

Normal QQ plots of all emerging countries, as well as the S&P 500 are plotted in Figure 1.4. A normal QQ plot is a QQ plot for which the reference distribution is chosen to be a normal distribution. In these QQ plots, the mean and the variance of the theoretical normal distribution are chosen as the sample mean and variance of the return data. All of the QQ-Plots display an S shape. Therefore, the return data for all countries and the S&P 500 have fatter tails than implied by the normal distribution. The ones that are relatively closer to be normal are those from Taiwan, Chile, and the S&P 500. This is consistent with the summary statistics in Table 1.1.
1.2 Data and Statistical Analysis

(a) Turkey

(b) Indonesia

(c) Malaysia

(d) Philippines

(e) South Korea

(f) China

(g) Taiwan

(h) Thailand
1.2 Data and Statistical Analysis

Figure 1.4: Normal QQ-Plots for S&P/IFCI Returns

(i) Brazil
(j) Chile
(k) Mexico
(l) Argentina
(m) Peru
(n) S&P500
Besides these graphical tools, formal numerical tests are administered to evaluate Gaussianity. The Jarque Bera test statistic (JBTS) is given by

$$JBTS = \frac{1}{6}n(s + \frac{1}{4}(k - 3)^2)$$

where $s$ is the square of sample skewness and $k$ is the sample kurtosis. The asymptotic distribution for JBTS is a chi-square with a two degree of freedom under the null hypothesis of Gaussianity. JBTS and associated probability values for all emerging markets and the S&P 500 returns are presented in Table 1.1. The null hypothesis of Gaussianity is rejected for all countries at all reasonable significance levels. Note that the smallest three JBTS are obtained from Chile, Taiwan, and the S&P 500 index. These results are consistent with the previous judgment that these countries are closer to being normal.

Next, the serial dependence of returns is analyzed to determine the plausibility of an independence assumption and to investigate the structure of dependence if it exists. The most fundamental measures of dependence in a series are autocovariance and autocorrelation. A correlogram displays the sample autocorrelation of a data up to some number of lags. Correlograms up to lag 25 for the raw returns and the squared returns are presented in Figures 1.5 and 1.6, respectively. The dashed horizontal lines are at the 95 percent confidence bands for the autocorrelation of iid Gaussian noise. These correlograms reveal that the autocorrelation of raw returns is mostly concentrated in the first lag. On the other hand, the squared returns exhibit a high degree of autocorrelation even up to lags 20-25. High returns are generally followed by other high returns, and low returns are generally followed by other low returns, but not necessarily with the same sign. This general behavior in financial time series data is known as “volatility clustering” and is captured well by the GARCH models.
1.2 Data and Statistical Analysis

(a) Turkey

(b) Indonesia

(c) Malaysia

(d) Philippines

(e) Korea

(f) China

(g) Taiwan

(h) Thailand
1.2 Data and Statistical Analysis

Figure 1.5: Correlograms for Raw Returns of S&P/IFCI Index

(i) Brazil  
(j) Chile  
(k) Mexico  
(l) Argentina  
(m) Peru  
(n) S&P500
1.2 Data and Statistical Analysis

(a) Turkey

(b) Indonesia

(c) Malaysia

(d) Philippines

(e) Korea

(f) China

(g) Taiwan

(h) Thailand
Figure 1.6: Correlograms for Squared Returns of S&P/IFCI Index
There are also formal numerical tests designed to evaluate the serial dependence. These are known as the portmanteau tests, and the Ljung-Box test is the most widely used of these. The Ljung-Box Q statistic is given by

\[
LBQ = n(n + 2) \sum_{i=1}^{h} \frac{\hat{\rho}_i^2}{n - i},
\]

where \( \rho_i \) is the autocorrelation in lag \( i \). Under the null hypothesis of independent and identically distributed returns, the Ljung-Box Q statistic has an asymptotic chi-squared distribution with a two degree of freedom. The Ljung-Box test is administered at the first lag \( h = 1 \), for both the raw returns and squared returns. The results are presented in Table 1.1. The independence hypothesis is rejected for all countries. The hypothesis of autocorrelation in the first lag can be rejected only for Turkey and the S&P 500. However, squared returns exhibit significant autocorrelation for these countries as well, which contradicts with independence.

Both graphical and formal numerical methods have shown that neither normality nor independent identical distribution assumption for returns are plausible.

### 1.3 Extreme Value Theory

EVT is a strong method to study the tail behaviors of loss distributions. It allows the modeling of fat-tailed distributions. There are two kinds of approaches to model extreme values. Modeling the maxima and modeling the observations that exceed a high threshold. For maxima modeling, see Embrechts, Kluppelberg, and Mikosch (1997). In this study, I concentrate on the threshold exceedances methodology, which is a method used to estimate the distribution of exceedances above a high threshold.\(^2\)

\(^2\)This is because Ljung-Box test at the first lag was enough for the rejection of independence.
1.3 Extreme Value Theory

Assuming \( X \) as the univariate variable of interest, \( F_X \) as its distribution function, \( u \) as the high threshold, and \( x_o \) as the right endpoint of the support of \( X \), the distribution of exceedances is defined as

\[
F_u(y) = \Pr(X - u \leq y | X > u) \quad \text{for} \quad 0 < y < x_o - u.
\]

If there is no finite right endpoint for the support of \( X \), then \( x_o = \infty \), which is the case for most distributions of financial data. The distribution of exceedances \( F_u(y) \) represents the probability of the “exceedance over the threshold” being less than \( y \) given the fact that an exceedance over the threshold has already occurred.

The following limit result that relates the distribution of exceedances \( F_u(y) \) with the generalized Pareto distribution (GPD) is the key point in univariate EVT.

\[
\lim_{u \to x_o} \sup_{0 < y \leq x_o} |F_u(y) - G_{\xi, \beta}(y)| = 0, \quad (1.5)
\]

where \( G_{\xi, \beta}(y) \) is the distribution function of GPD given by:

\[
G_{\xi, \beta}(y) = \begin{cases} 
1 - (1 + \frac{\xi y}{\beta})^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\
1 - \exp\left(-\frac{y}{\beta}\right) & \text{if } \xi = 0 
\end{cases}, \quad (1.6)
\]

where \( \beta > 0 \) and \( (1 + \frac{\xi y}{\beta}) > 0 \).

Here, \( \beta \) is the scale parameter and \( \xi \) is the shape parameter of GPD. The limit result (1.5) basically reveals that the distribution of exceedances uniformly converges to GPD as the threshold converges to the right end point of the support of \( X \). By exploiting this result, it can be assumed that not only the limit distribution, but also \( F_u \) itself, is GPD for a high enough threshold. In the threshold exceedances
1.3 Extreme Value Theory

methodology, a suitable threshold is chosen, and the observations that exceed this high threshold are filtered from the data. Let \( N_u \) observations exceed \( u \), and they are labeled as \( X_1, X_2, \ldots, X_{N_u} \). Then, the exceedances over the threshold are calculated as \( Y_j = X_j - u \), and GPD is fit to \( Y \) by maximizing the following likelihood function

\[
\log L(\xi, \beta, Y) = -N_u \log(\beta) - \sum_{j=1}^{N_u} \log \left( 1 + \frac{\xi Y_j}{\beta} \right),
\]

subject to the parameter constraints \( \beta > 0 \) and \( 1 + \frac{\xi Y_j}{\beta} > 0 \). This log-likelihood function can easily be verified using the distribution function of the GPD given by (1.6). The methodology to estimate the VaR by using the MLE estimates of the GPD fit is explained in section 1.4.6.

The choice of an appropriate threshold is an important issue in this procedure. If a low threshold is chosen, then GPD approximation is biased. This is because GPD holds only in the limit as \( u \) approaches infinity. The assumption that the exceedances over the chosen threshold are GPD distributed is an approximation. For this assumption to be reliable, the threshold should be set as high as possible. On the other hand, a very high threshold results in very few observations exceeding the threshold for the GPD parameter estimation. This results in high standard errors for estimated parameters. Therefore, in threshold choice, there is a bias-variance trade-off that is a standard statistical problem. There are several graphical methods proposed, such as the Hill plots, to determine the appropriate threshold. However, since I predict the VaR for each period with a rolling window approach, it is impractical to use these methods. Following the common approach in the literature, the 0.95th quantile of the data is used as the threshold in GPD estimations.
1.4 VaR Estimation

The S&P/IFCI dataset runs from June 30, 1995 to November 1, 2007 and has a total of \( T = 3,219 \) daily price observations for each of the 13 emerging markets and the S&P 500 index. The return data are obtained by taking the differences in logarithmic prices as in (1.4), and the loss data are obtained by negating the returns.

The empirical methodology used for out-of-sample forecasting is known as a sliding window scheme, and the length of the sliding window is chosen as 500 observations.\(^3\) To make a prediction for day \( t \) where \( t \in \{501, 502, ..., T\} \), only the loss observations \( \{\ell_{t-1}, \ell_{t-2}, ..., \ell_{t-500}\} \) are used. That is, observations 1 through 500 are used to estimate the VaR for day 501, observations 2 through 501 are used to estimate the VaR for day 502. Since \( T = 3,219 \), there are 2,719 VaR predictions. VaR is calculated for 0.95th, 0.975th, 0.99th and 0.995th quantiles. In what follows, the methods used for VaR estimation are presented briefly.

1.4.1 Historical Simulation

Historical simulation is a purely non-parametric method that uses \( \alpha_{th} \) quantile of the empirical distribution of loss data as \( \text{VaR}_\alpha \). It is estimated by ordering the losses, \( \{\ell_{t-1}, \ell_{t-2}, ..., \ell_{t-500}\} \) in descending order as \( \{\ell(1), \ell(2), ..., \ell(500)\} \) so that \( \ell(1) \) is the largest and \( \ell(500) \) is the smallest loss and choosing

\[
\text{VaR}_\alpha = \ell(500(1 - \alpha)) .
\]

If \( 500(1 - \alpha) \) is not an integer, basic linear interpolation is used.

\(^3\)In GARCH literature, at least 500 observations are suggested in order to have stable parameter estimates. I choose to follow this minimum requirement in order to have a longer backtesting period.
1.4 VaR Estimation

1.4.2 Normal Distribution Model

In this model, losses are assumed to be normally distributed; that is, \( \ell_t \sim iid \, N(\mu, \sigma^2) \). Using the loss data in the sliding window \( \{ \ell_{t-1}, \ell_{t-2}, ..., \ell_{t-500} \} \), MLE estimators of \( \mu \) and \( \sigma \) are calculated and \( VaR_\alpha \) is estimated by

\[
VaR_\alpha = \hat{\mu} + \hat{\sigma} \, \Phi^{-1}(\alpha),
\]

where \( \Phi() \) is the standard normal distribution function.

1.4.3 Student’s t-Distribution Model

In this model, the losses are assumed to be student’s \( t \) distributed; that is, \( \ell_t \sim iid \, t(\mu, \sigma^2, \nu) \). Student’s \( t \) distribution can account for the fat tails, and therefore is included in the study. The density function for a \( t \)-distributed random variable is given by

\[
f(x) = \frac{\Gamma(\nu + 1)}{\sqrt{\nu \pi} \, \Gamma(\frac{\nu}{2})} \left( 1 + \frac{(x-\mu)^2}{\nu} \right)^{-\frac{\nu+1}{2}},
\]

where \( \Gamma \) is the well-known gamma function. The parameters \( \{\mu, \sigma, \nu\} \) are estimated by MLE methods, and \( VaR_\alpha \) is calculated by

\[
VaR_\alpha = \hat{\mu} + \hat{\sigma} \, \hat{t}_\nu^{-1}(\alpha),
\]

where \( \hat{t}_\nu^{-1}() \) is the quantile function of standard \( t \)-distribution with \( \hat{\nu} \) degrees of freedom. The above relationship of quantiles is actually valid for all distributions in the location scale family.
1.4 VaR Estimation

1.4.4 Skewed Normal Distribution Model

In this model, losses are assumed to be coming from a skewed normal distribution; that is, \( \ell_t \sim iid \ SN(\mu, \sigma^2, \theta) \). Skewed normal model is included in the study in order to account for the skewness in the emerging market return distributions. The density function for the skewed normal distribution, which was developed by Azzalini(1985), is given by:

\[
f(x) = 2 \phi \left( \frac{x - \mu}{\sigma} \right) \Phi \left( \theta \frac{x - \mu}{\sigma} \right),
\]

where \( \phi() \) and \( \Phi() \) denote the standard normal pdf and cdf, respectively. Note that when \( \theta = 0 \), this reduces to symmetric normal pdf, and as \( \theta \) increases the skewness also increases. So, \( \theta \) is the parameter that controls the skewness of the distribution.

The skewed normal distribution preserves some properties of the normal distribution, such as the linear combinations of skewed normal variables are also skewed normals. Density functions for several skewed normal distributed variables can be seen in Figure 1.7 (a,b,c).

\[\text{(a) } \theta = -2 \quad \text{(b) } \theta = 0 \quad \text{(c) } \theta = 2\]

Figure 1.7: Densities for Several Skewed Normal Distributions with \( \mu = 0 \) and \( \sigma = 1 \)
1.4 VaR Estimation

The empirical procedure is to estimate \( \{\mu, \sigma, \theta\} \) with MLE methods and calculate the VaR with numerical methods as

\[
VaR_\alpha = F^{-1}_{\mu, \sigma, \theta}(\alpha),
\]

where \( F^{-1}_{\mu, \sigma, \theta}(\cdot) \) is the quantile function of skewed normal distribution associated with the estimated parameters.

1.4.5 Skewed t-Distribution Model

In this model, the losses are assumed to be coming from a skewed \( t \)-distribution; that is, \( \ell_t \sim iid \, St(\mu, \sigma^2, \gamma, \nu) \). Azzalini and Capitanio (2003) extended the approach of introducing a skewness parameter to any elliptical distribution, in particular to student’s \( t \)-distribution. Both the skewness and the fat tails in emerging market returns are accounted for with this model. The density function for the skewed \( t \)-distribution is given by:

\[
f(x) = c \frac{K_{(v+1)/2} \sqrt{2^2(\nu + \frac{(x-\mu)^2}{\sigma^2})} \exp\left(\frac{(x-\mu)^2}{\sigma^2} \gamma\right)}{\left(\sqrt{2}\sigma(\nu + \frac{(x-\mu)^2}{\sigma^2})\right)^{\frac{v+1}{2}} \left(1 + \frac{(x-\mu)^2}{\sigma^2 \nu}\right)^{\frac{v+1}{2}}},
\]

where the normalizing constant \( c = \frac{2^{(v+1)/2}}{\Gamma(\frac{1}{2})\sqrt{\pi \nu}} \), \( K \) is a modified Bessel function of the third kind, and \( \gamma \) is the parameter that controls the skewness. As \( \gamma \) converges to 0, the above density function in the limit converges to the density function of a symmetric student’s \( t \)-distribution. Density functions for several skewed \( t \)-distributed variables with different skewness parameters are presented in Figure 1.8 (a,b,c). For convenience, densities of skewed normal counterparts are also plotted with dashed lines on the same figures.
1.4 VaR Estimation

The empirical procedure is to estimate the parameters \( \{\mu, \sigma, \gamma, \nu\} \) by MLE methods and calculate the VaR by numerical methods as

\[
VaR_\alpha = F_{\hat{\mu}, \hat{\sigma}, \hat{\gamma}}^{-1}(\alpha),
\]

where \( F_{\hat{\mu}, \hat{\sigma}, \hat{\gamma}}^{-1}(\cdot) \) is the quantile function associated with the estimated parameters.

1.4.6 EVT Model

In section 1.3, the procedure to fit GPD to the tails of loss distribution is explained. In this section, the results of the GPD fit are used to estimate the VaR. Note that for a loss observation greater than the threshold, \( x > u \), we have

\[
\Pr(\ell > x) = \Pr(\ell > u) \Pr(\ell > x | \ell > u) \\
= \Pr(\ell > u) \Pr(\ell - u > x - u | \ell > u) \\
= (1 - \Pr(\ell \leq u)) (1 - \Pr(\ell - u \leq x - u | \ell \geq u)) \\
= (1 - F_\ell(u)) (1 - F_u(x - u)).
\]
1.4 VaR Estimation

By exploiting the limit result (1.5), $F_u$ can be approximated as if it were an exact GPD distribution function. Thus, the following expression can be obtained by substituting the GPD distribution function with the MLE parameter estimates of $\xi$ and $\beta$ in place of $F_u$,

$$\Pr(\ell > x) = (1 - F_\ell(u)) \left( 1 + \frac{\xi (x - u)}{\beta} \right)^{-1/\xi}.$$

Further, if $x = VaR_\alpha$ is substituted in this expression, then $Pr(\ell > x)$ becomes $1 - \alpha$ by the definition of VaR. Then,

$$1 - \alpha = (1 - F_\ell(u)) \left( 1 + \frac{\xi VaR_\alpha - u}{\beta} \right)^{-1/\xi}.$$

Arranging the terms to solve for $VaR_\alpha$, the following result can be obtained:

$$VaR_\alpha = u + \frac{\hat{\beta}}{\xi} \left( \left( \frac{1 - \alpha}{1 - F_\ell(u)} \right)^{-\xi} - 1 \right).$$

In EVT methodology, the 0.95th quantile of the loss data is used as the threshold. This is because a sliding window methodology is employed to produce 2,719 one-day-ahead predictions, and it is impractical to use Hill plots for threshold selection in every step. The choice of the 0.95th quantile as the threshold is very common in EVT literature. Therefore, the threshold $u$ is chosen to be the 25th largest loss in the sliding window. In EVT methodology, a GPD fit is used for only the tail region beyond the threshold. However, the empirical distribution is used up to the threshold. Therefore, $F_\ell(u) = \Pr(x \leq u)$ can be estimated non-parametrically as 0.95. As a result, $VaR_\alpha$ is obtained as:

$$VaR_\alpha(\ell_t) = \ell_t(25) + \frac{\hat{\beta}}{\xi} \left( \left( 20(1 - \alpha) \right)^{-\xi} - 1 \right),$$ (1.7)
where $\ell(25)$ denotes the 25th largest loss in the sliding window. In addition to the modeling of fat tails, the EVT method implicitly accounts for skewness because it uses the empirical distribution up to the threshold.

### 1.4.7 GARCH Model

All of the methods explained above assume the observations are independent and identically distributed. Therefore, these methods cannot capture the volatility clustering that is a common behavior of financial time series. To incorporate dynamic volatility in the emerging market returns, I use GARCH models. The following model is estimated, which allows an AR(1) structure in conditional mean and GARCH(1,1) structure in conditional variance equation with shocks $z_t \sim iid N(0,1)$,

\[
\begin{align*}
    r_t &= \mu_t + \epsilon_t \\
    \epsilon_t &= \sigma_t z_t \\
    \mu_t &= c_1 + c_2 r_{t-1} \\
    \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2,
\end{align*}
\]

where $r_t$ is the return on day $t$. The $Var_t$ is calculated in three steps. First, the above Gaussian-GARCH model is fit to the data with the MLE method. Specifically, parameter estimates for $\hat{\theta} = \{\hat{c}_1, \hat{c}_2, \hat{\omega}, \hat{\alpha}, \hat{\beta}\}$ are obtained by maximizing the log-likelihood function

\[
logL(\theta; r_{t-1}, \ldots, r_{t-500}) \propto \sum_{i=1}^{i=500} \left( \log \sigma_{t-i}^2(\theta) - \frac{\epsilon_{t-i}^2}{\sigma_{t-i}^2(\theta)} \right)
\]
1.4 VaR Estimation

Then, $\hat{\mu}_t$ and $\hat{\sigma}_t$ are obtained by substituting the MLE parameter estimates into the conditional mean and conditional variance equations as:

$$\hat{\mu}_t = \hat{c}_1 + \hat{c}_2 r_{t-1}; \quad \hat{\sigma}_t^2 = \hat{w} + \hat{\alpha} \varepsilon_{t-1}^2 + \hat{\beta} \sigma_{t-1}^2.$$  

Lastly, the VaR estimate is calculated as:

$$VaR_{\alpha}(\ell_t) = -\hat{\mu}_t + \hat{\sigma}_t \Phi^{-1}(\alpha),$$

where $\Phi()$ is the normal cdf. Here, mean has a negative sign because the GARCH model is fit to the returns not to the losses. Therefore, the mean and standard deviation of the loss distribution is $-\mu_t$ and $\sigma_t$, respectively.

1.4.8 GARCH-EVT Model

The GARCH model above assumes Gaussian innovations. So, it is incapable of modeling fat tails, although it can model volatility changes over time. The GARCH-EVT model also assumes an AR(1) in conditional mean and GARCH(1,1) in conditional variance, but the shocks $z_t$ are not assumed to be Gaussian any more. Instead, the only assumptions are that $z_t$ is iid, $E[z_t] = 0$ and $E[z_t^2] = 1$. The GARCH-EVT model first proposed by McNeil and Frey (2000) follows a three-step methodology to estimate the VaR. First, the model parameters $\hat{\theta} = \{\hat{c}_1, \hat{c}_2, \hat{w}, \hat{\alpha}, \hat{\beta}\}$ are estimated by quasi Maximum Likelihood Estimation (quasi-MLE). The quasi-MLE refers to fitting a Gaussian-GARCH model to the data as for the regular GARCH model. Fitting Gaussian-GARCH to the data may seem unreasonable because the Gaussianity assumption for the residuals is dropped. However, Gourieroux (1997) showed that Gaussian-GARCH estimation, when the actual residual distribution is not Gaussian
(quasi-MLE) delivers consistent and asymptotically normal parameter estimates under some regularity conditions. The output of the first step is the conditional mean and conditional variance predictions, $\hat{\mu}_t, \hat{\sigma}_t$. In the second step, standardized residuals are extracted from the fit by

$$z_{t-i} = \frac{\hat{e}_{t-i}}{\hat{\sigma}_{t-i}} \forall i \in \{1, 2, ..., 500\}.$$

The implied residuals from the GARCH fit are still fat tailed, and this is modeled by applying EVT methodology to the standardized residuals as described in section 1.4.6 to obtain

$$VaR_\alpha(Z_t) = z(25) + \frac{\hat{\beta}}{\hat{\xi}} \left( (20(1 - \alpha))\hat{e} - 1 \right),$$

where $z(25)$ is the 25th largest standardized residual and $\hat{\xi}, \hat{\beta}$ are MLE estimates of GPD fit to standardized residuals. Lastly, in step three, the VaR of the original losses can be calculated by

$$VaR_\alpha(L_t) = -\hat{\mu}_t + \hat{\sigma}_t VaR_\alpha(Z_t).$$

Using the EVT methods on residuals together with the GARCH model, both the dynamic volatility and fat tails are addressed and better VaR predictions can be obtained.

### 1.4.9 GARCH-HS Model

The methodology in GARCH-HS is similar to the GARCH-EVT in essence. Instead of fitting a GPD to the tails of implied residuals, this method applies historical simulation, as described in section 1.4.1 to the implied residuals. First, parameters of the GARCH(1,1) model are estimated by quasi-MLE, and the standardized residuals are
extracted. Then the VaR of the standardized residuals are estimated with historical simulation. Lastly, VaR is estimated by

\[
VaR_\alpha(\ell_t) = -\mu_t + \hat{\sigma}_t VaR_\alpha(Z_t).
\]

### 1.4.10 EWMA Model

The EWMA is an ad-hoc and model-free technique for volatility forecasting. The volatility of next period is given by a weighted average of volatility from the last period and squared return from the last period. With recursive substitution, it can be shown that EWMA volatility is a weighted average of past squared returns where the weight given to older observations decreases exponentially.

\[
\begin{align*}
    r_t &= \sigma_t z_t \\
    \sigma_t^2 &= \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2 \quad (1.8) \\
    z_t &\sim N(0, 1).
\end{align*}
\]

The current Risk-Metrics VaR estimation method proposes using an EWMA model with \( \lambda = 0.94 \). This parameterization is used in order to have it as a benchmark because most of the financial industry uses this model. To start the recursion in volatility equation, I used the standard deviation of the 500 observations in the sliding window. Once the volatility forecast is obtained from EWMA model, the VaR is predicted as

\[
VaR_\alpha = \hat{\sigma}_t \Phi^{-1}(\alpha).
\]
1.5 Backtesting

Backtesting is a way of model validation. The most recent 500 loss observations \( \{\ell_{t-1}, \ell_{t-2}, \ldots, \ell_{t-500}\} \) are used to predict the \( VaR_t \). Then, the loss \( \ell_t \) is observed at time \( t \). After \( \ell_t \) becomes known, there is an opportunity to measure the performance of the \( VaR \) prediction. Monitoring the performance of \( VaR \) prediction methods continuously over time is known as backtesting.

Backtesting of \( VaR \) is administered through the violations of the \( VaR \). Whenever \( \ell_t > VaR_t \), it is said that a violation of the \( VaR \) has occurred. By definition, \( VaR_\alpha \) is a number that the loss can exceed with \( 1 - \alpha \) probability. Therefore, the ratio of the total number of such violations to the total number of observations in the backtesting period should be close to \( 1 - \alpha \). The dataset has \( T=3,219 \) loss observations for each country, and 2,719 of them can be used in the backtesting procedure since the sliding window length is 500. Therefore, the expected number of violations over \( VaR_{0.995} \) is 14, over \( VaR_{0.99} \) is 27, over \( VaR_{0.95} \) is 68, and over \( VaR_{0.95} \) is 136. These are the targets for the number of violations of the \( VaR \) for corresponding level of quantiles. The performances of the methods can be compared by checking how close their \( VaR \) violations to these targets.

More formally, the process of \( VaR \) violations can be thought as a Bernoulli trial. For each trial the probability of a violation is, \( \Pr(\ell_t > VaR_{t,\alpha}) = 1 - \alpha \). Moreover, each of these trials should be independent of each other by the assumptions of models. Therefore, the total number of violations should be binomially distributed; that is,

\[
\sum_{i=1}^{i=2719} I\{\ell_t > VaR_{t,\alpha}\} \sim B(2719, 1 - \alpha),
\]

where \( I\{\} \) is the indicator function. This backtesting methodology is adopted from
Detailed backtesting results for all countries and all models are given in Table 1.2. The expected number of violations for all quantiles is given in the first rows of the tables. Then, the number of VaR violations for all methods is given in the respective rows together with the probability values from a one-sided likelihood ratio test of the null hypothesis of binomial distribution for the number of exceedances. For 14 countries and 4 separate quantiles, the hypothesis testing procedure is repeated. Therefore, each model is tested for a total of 56 cases. A probability value of 0.05 is used as the rejection threshold in these tests. A scoring table for the risk models is presented in Table 1.3. The number of cases that resulted in a failure to reject the null hypothesis of binomial distribution is reported for all quantiles separately as well as the totals. Therefore, the higher the score, the better the model.

Overall, the GARCH-EVT model is the best model. Binomial hypothesis cannot be rejected in 52 of the 56 cases. Additionally, the performance of this model is fairly uniform at all quantiles because it has the ability to model both the fat tails and the dynamic volatility. Since EVT is a semiparametric method and focuses only on the left (loss) tail of the distribution, it also implicitly accounts for the skewness.

One important observation is that any model with a normality assumption performs very badly at all quantiles except the 0.95th. These are the normal model, skewed normal model, Gaussian GARCH model, and the EWMA model of Risk Metrics. Any kind of normality assumption results in significant underestimation of the risk. At the 0.95th quantile, almost all models show similar performances. The investment banking industry is calculating the VaR at 0.95th quantile, and this may result in drawing the wrong conclusions in model validation. The performance of the risk measurement techniques should be monitored for higher quantiles as well as in validating the models.
### 1.5 Backtesting

<table>
<thead>
<tr>
<th>(a) Turkey</th>
<th>(b) Indonesia</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.95 )</td>
<td>( \alpha = 0.95 )</td>
</tr>
<tr>
<td>Target</td>
<td>Target</td>
</tr>
<tr>
<td>136</td>
<td>136</td>
</tr>
<tr>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>127 (0.43)</td>
<td>125 (0.33)</td>
</tr>
<tr>
<td>Student's t</td>
<td>Student's t</td>
</tr>
<tr>
<td>146 (0.30)</td>
<td>158 (0.08)</td>
</tr>
<tr>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>135 (0.95)</td>
<td>149 (0.28)</td>
</tr>
<tr>
<td>EVT</td>
<td>EVT</td>
</tr>
<tr>
<td>135 (0.95)</td>
<td>149 (0.26)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>Skewed Normal</td>
</tr>
<tr>
<td>124 (0.29)</td>
<td>119 (0.13)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>Skewed-t</td>
</tr>
<tr>
<td>147 (0.34)</td>
<td>152 (0.17)</td>
</tr>
<tr>
<td>GARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>125 (0.33)</td>
<td>129 (0.54)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>GARCH-HS</td>
</tr>
<tr>
<td>140 (0.32)</td>
<td>131 (0.58)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>GARCH-EVT</td>
</tr>
<tr>
<td>140 (0.32)</td>
<td>131 (0.66)</td>
</tr>
<tr>
<td>EWMA</td>
<td>EWMA</td>
</tr>
<tr>
<td>133 (0.79)</td>
<td>148 (0.25)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(c) Malaysia</th>
<th>(d) Philippines</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.95 )</td>
<td>( \alpha = 0.95 )</td>
</tr>
<tr>
<td>Target</td>
<td>Target</td>
</tr>
<tr>
<td>136</td>
<td>136</td>
</tr>
<tr>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>130 (0.6)</td>
<td>137 (0.93)</td>
</tr>
<tr>
<td>Student's t</td>
<td>Student's t</td>
</tr>
<tr>
<td>157 (0.07)</td>
<td>158 (0.08)</td>
</tr>
<tr>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>156 (0.08)</td>
<td>162 (0.23)</td>
</tr>
<tr>
<td>EVT</td>
<td>EVT</td>
</tr>
<tr>
<td>156 (0.08)</td>
<td>162 (0.23)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>Skewed Normal</td>
</tr>
<tr>
<td>133 (0.79)</td>
<td>155 (0.1)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>Skewed-t</td>
</tr>
<tr>
<td>156 (0.06)</td>
<td>159 (0.05)</td>
</tr>
<tr>
<td>GARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>110 (0.02)</td>
<td>123 (0.25)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>GARCH-HS</td>
</tr>
<tr>
<td>127 (0.43)</td>
<td>139 (0.79)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>GARCH-EVT</td>
</tr>
<tr>
<td>127 (0.43)</td>
<td>131 (0.66)</td>
</tr>
<tr>
<td>EWMA</td>
<td>EWMA</td>
</tr>
<tr>
<td>139 (0.79)</td>
<td>148 (0.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(e) South Korea</th>
<th>(f) China</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.95 )</td>
<td>( \alpha = 0.95 )</td>
</tr>
<tr>
<td>Target</td>
<td>Target</td>
</tr>
<tr>
<td>136</td>
<td>136</td>
</tr>
<tr>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>139 (0.79)</td>
<td>139 (0.79)</td>
</tr>
<tr>
<td>Student's t</td>
<td>Student's t</td>
</tr>
<tr>
<td>162 (0.03)</td>
<td>158 (0.06)</td>
</tr>
<tr>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>135 (0.93)</td>
<td>154 (0.12)</td>
</tr>
<tr>
<td>EVT</td>
<td>EVT</td>
</tr>
<tr>
<td>135 (0.93)</td>
<td>154 (0.12)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>Skewed Normal</td>
</tr>
<tr>
<td>124 (0.25)</td>
<td>126 (0.38)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>Skewed-t</td>
</tr>
<tr>
<td>143 (0.56)</td>
<td>159 (0.05)</td>
</tr>
<tr>
<td>GARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>150 (0.23)</td>
<td>155 (0.1)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>GARCH-HS</td>
</tr>
<tr>
<td>133 (0.79)</td>
<td>153 (0.14)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>GARCH-EVT</td>
</tr>
<tr>
<td>133 (0.79)</td>
<td>155 (0.1)</td>
</tr>
<tr>
<td>EWMA</td>
<td>EWMA</td>
</tr>
<tr>
<td>171 (0)</td>
<td>155 (0.1)</td>
</tr>
</tbody>
</table>
### 1.5 Backtesting

<table>
<thead>
<tr>
<th>(g) Taiwan</th>
<th>(h) Thailand</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.95 )</td>
<td>( \alpha = 0.95 )</td>
</tr>
<tr>
<td>Target</td>
<td>136</td>
</tr>
<tr>
<td>Normal</td>
<td>129 (0.54)</td>
</tr>
<tr>
<td>Student's t</td>
<td>148 (0.33)</td>
</tr>
<tr>
<td>HS</td>
<td>145 (0.23)</td>
</tr>
<tr>
<td>EVT</td>
<td>148 (0.33)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>134 (0.86)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>154 (0.12)</td>
</tr>
<tr>
<td>GARCH</td>
<td>155 (0.01)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>159 (0.05)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>159 (0.05)</td>
</tr>
<tr>
<td>EWMA</td>
<td>158 (0.05)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(i) Brazil</th>
<th>(j) Chile</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.95 )</td>
<td>( \alpha = 0.95 )</td>
</tr>
<tr>
<td>Target</td>
<td>136</td>
</tr>
<tr>
<td>Normal</td>
<td>165 (0.01)</td>
</tr>
<tr>
<td>Student's t</td>
<td>193 (0)</td>
</tr>
<tr>
<td>HS</td>
<td>168 (0.01)</td>
</tr>
<tr>
<td>EVT</td>
<td>168 (0.01)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>147 (0.34)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>167 (0.01)</td>
</tr>
<tr>
<td>GARCH</td>
<td>154 (0.12)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>145 (0.43)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>145 (0.43)</td>
</tr>
<tr>
<td>EWMA</td>
<td>150 (0.22)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(k) Mexico</th>
<th>(l) Argentina</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.95 )</td>
<td>( \alpha = 0.95 )</td>
</tr>
<tr>
<td>Target</td>
<td>136</td>
</tr>
<tr>
<td>Normal</td>
<td>133 (0.79)</td>
</tr>
<tr>
<td>Student's t</td>
<td>163 (0.02)</td>
</tr>
<tr>
<td>HS</td>
<td>155 (0.1)</td>
</tr>
<tr>
<td>EVT</td>
<td>155 (0.1)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>127 (0.43)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>148 (0.23)</td>
</tr>
<tr>
<td>GARCH</td>
<td>148 (0.3)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>138 (0.86)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>138 (0.86)</td>
</tr>
<tr>
<td>EWMA</td>
<td>134 (0.86)</td>
</tr>
</tbody>
</table>

| \( \alpha = 0.975 \) | \( \alpha = 0.975 \) |
| \( \alpha = 0.99 \) | \( \alpha = 0.99 \) |
| \( \alpha = 0.995 \) | \( \alpha = 0.995 \) |

| \( \alpha = 0.95 \) | \( \alpha = 0.95 \) |
| \( \alpha = 0.975 \) | \( \alpha = 0.975 \) |
| \( \alpha = 0.99 \) | \( \alpha = 0.99 \) |
| \( \alpha = 0.995 \) | \( \alpha = 0.995 \) |
1.5 Backtesting

### (m) Peru

<table>
<thead>
<tr>
<th>Model</th>
<th>α = 0.95</th>
<th>α = 0.975</th>
<th>α = 0.99</th>
<th>α = 0.995</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>136</td>
<td>88</td>
<td>27</td>
<td>14</td>
</tr>
<tr>
<td>Normal</td>
<td>134 (0.86)</td>
<td>91 (0.01)</td>
<td>62 (0)</td>
<td>46 (0)</td>
</tr>
<tr>
<td>Student's t</td>
<td>163 (0.02)</td>
<td>90 (0.01)</td>
<td>36 (0.11)</td>
<td>18 (0.25)</td>
</tr>
<tr>
<td>HS</td>
<td>157 (0.07)</td>
<td>86 (0.03)</td>
<td>38 (0.05)</td>
<td>24 (0.01)</td>
</tr>
<tr>
<td>EVT</td>
<td>157 (0.07)</td>
<td>84 (0.06)</td>
<td>36 (0.11)</td>
<td>22 (0.04)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>125 (0.33)</td>
<td>88 (0.02)</td>
<td>59 (0)</td>
<td>45 (0)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>163 (0.02)</td>
<td>85 (0.04)</td>
<td>35 (0.18)</td>
<td>21 (0.06)</td>
</tr>
<tr>
<td>GARCH</td>
<td>133 (0.79)</td>
<td>85 (0.04)</td>
<td>53 (0)</td>
<td>38 (0)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>148 (0.32)</td>
<td>82 (0.1)</td>
<td>36 (0.11)</td>
<td>23 (0.03)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>148 (0.32)</td>
<td>77 (0.28)</td>
<td>31 (0.47)</td>
<td>22 (0.04)</td>
</tr>
<tr>
<td>EWMA</td>
<td>124 (0.29)</td>
<td>81 (0.12)</td>
<td>43 (0)</td>
<td>34 (0)</td>
</tr>
</tbody>
</table>

### (n) S&P500

<table>
<thead>
<tr>
<th>Model</th>
<th>α = 0.95</th>
<th>α = 0.975</th>
<th>α = 0.99</th>
<th>α = 0.995</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>136</td>
<td>68</td>
<td>27</td>
<td>14</td>
</tr>
<tr>
<td>Normal</td>
<td>135 (0.93)</td>
<td>89 (0.01)</td>
<td>50 (0)</td>
<td>31 (0)</td>
</tr>
<tr>
<td>Student’s t</td>
<td>152 (0.17)</td>
<td>83 (0.07)</td>
<td>32 (0.37)</td>
<td>19 (0.17)</td>
</tr>
<tr>
<td>HS</td>
<td>143 (0.54)</td>
<td>88 (0.02)</td>
<td>40 (0.02)</td>
<td>25 (0.01)</td>
</tr>
<tr>
<td>EVT</td>
<td>143 (0.54)</td>
<td>83 (0.07)</td>
<td>34 (0.21)</td>
<td>21 (0.08)</td>
</tr>
<tr>
<td>Skewed Normal</td>
<td>133 (0.78)</td>
<td>85 (0.04)</td>
<td>41 (0.01)</td>
<td>33 (0)</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>148 (0.3)</td>
<td>81 (0.12)</td>
<td>32 (0.37)</td>
<td>17 (0.37)</td>
</tr>
<tr>
<td>GARCH</td>
<td>152 (0.17)</td>
<td>90 (0.01)</td>
<td>50 (0)</td>
<td>34 (0)</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>152 (0.17)</td>
<td>84 (0.06)</td>
<td>40 (0.02)</td>
<td>28 (0)</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>152 (0.17)</td>
<td>80 (0.15)</td>
<td>35 (0.15)</td>
<td>18 (0.25)</td>
</tr>
<tr>
<td>EWMA</td>
<td>150 (0.22)</td>
<td>93 (0)</td>
<td>54 (0)</td>
<td>33 (0)</td>
</tr>
</tbody>
</table>

### Table 1.2: VaR Backtesting Tables

<table>
<thead>
<tr>
<th>Model</th>
<th>α = 0.95</th>
<th>α = 0.975</th>
<th>α = 0.99</th>
<th>α = 0.995</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>t-Model</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>HS</td>
<td>12</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>EVT</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td>Skewed-Norm</td>
<td>14</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Skewed-t</td>
<td>11</td>
<td>7</td>
<td>10</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>GARCH</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>GARCH-HS</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>46</td>
</tr>
<tr>
<td>GARCH-EVT</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>12</td>
<td>53</td>
</tr>
<tr>
<td>EWMA</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>

### Table 1.3: Scoring Table for VaR Prediction Models

Besides these general results, I briefly discuss the relative improvements obtained from fat tail, dynamic volatility, and skewness modeling in what follows.

Fat Tail Modeling:

Except for the 0.95th quantile, the EVT model outperforms the normal model. The normality assumption results in a significant underestimation of the risk. EVT accomplishes modeling of heavy tails by assigning more probability weight to the occurrence of unexpected extreme events. Also, the GARCH-EVT model gives better results.
compared to the Gaussian GARCH model. The Gaussian GARCH model underestimates the risk because the residuals from the GARCH model still have fatter tails than implied by a normal distribution, which requires further modeling. Additionally, the student’s $t$-distribution results in less rejection than the normal and skewed $t$-distribution results in less rejection than the skewed normal model in all quantiles except for the 0.95th quantile. All these results exhibit the importance of fat-tail modeling in VaR estimation. Comparing the EVT model with the $t$-distribution models, the EVT model accomplishes the fat-tail modeling much better. One common observation is that fat-tail modeling gets more and more important as the quantile of interest increases. While there is not much difference in the performances of different models in the 0.95th quantile, there is a huge gap at higher quantiles.

Dynamic Volatility Modeling:
At the 0.95th quantile, performances of the GARCH model and the normal model are almost the same. However, at the 0.975th quantile, the GARCH model outperforms the normal model. Additionally, the GARCH-EVT model outperforms the EVT model at all quantiles. Similarly, the GARCH-HS model results in a great improvement over the HS model. All these results imply that dynamic volatility is another important aspect in VaR estimation.

Skewness Modeling:
Both skewed normal distribution and skewed $t$-distribution outperform their symmetric counterparts in VaR estimation performance resulting in less rejection of the binomial hypothesis. This implies that skewness modeling can also improve the VaR estimation performance.
Emerging markets are characterized with high volatility and extreme movements in their financial markets. VaR as a measure of market risk has been developed in response to these sudden and extreme market events observed during the 1990s. For an investor who wants to be protected against catastrophic losses, VaR provides the level of loss to be exceeded with a specified small probability. Estimation of VaR requires the choice of a market risk model among numerous alternatives, and this brings an additional risk known as model risk.

In this study, the reliability of normality and independentness assumptions are investigated by means of graphical tools as well as formal numerical tests for the emerging stock markets and the S&P 500 returns. QQ plots reveal that the return data have much fatter tails than the normal distribution, and correlograms reveal that there is strong volatility clustering in the data. Consequently, normality and independentness hypotheses were strongly rejected.

Then, the relative importance of modeling these facts for risk management is examined. A comprehensive set of market risk models were studied to estimate VaR for emerging market stock indices and the S&P 500. Backtesting of one-day-ahead VaR predictions in several quantiles revealed that the most important aspects of the emerging market stock indices that needs careful modeling are the fat tails and volatility clustering. Skewness modeling can also help in improving VaR estimation performance. EVT proved to be a strong way of modeling the fat tails, much better than the t-distributions. Additionally, the GARCH models are strong in modeling the stochastic nature of market volatility. As a result, the GARCH-EVT method, which accomplishes modeling of dynamic volatility and fat tails simultaneously and takes account of skewness implicitly, exhibited superior performance in backtesting.
1.6 Conclusion

The more conventional risk measurement techniques, especially those employing normality assumptions, whether it be the normal model, the skewed normal model, the Gaussian GARCH model or the EWMA model, result in a significant underestimation of the VaR. This is because the tails of the normal distribution decays much faster than the actual data. Consequently, normality assumptions result in underestimation of the probability of extreme losses. This is probably why after each episode of sharp movement in market prices we used to hear from market practitioners that it was a one in a million day event. In fact they are not. The reality is that the models assuming normality for simplicity are not reliable at all. Another interesting finding is that backtesting performance at the 0.95th quantile is misleading in model validation. The performances of the models become distinguishable only at higher quantiles. Especially, the importance of fat-tail modeling increases at higher quantiles.

In this study, I used quasi-MLE methods in parameter estimation of the GARCH-EVT and the GARCH-HS models. Also, skewed models are used only in a static way assuming independent identically distributed losses. Future research may try to maximize the exact likelihood function of a the GARCH model with \( t \)-distributed innovations or even skewed \( t \)-distributed innovations. I argue that the GARCH-EVT model implicitly takes care of the skewness because it is a semiparametric method focusing only on the loss tail. However, it would be interesting to see results from a model that explicitly models skewness within a dynamic volatility setting.
Chapter 2

Hidden Extremal Dependence and Implications for Portfolio Risk

Summary. Using extremal dependence measures that focus on the joint tail region of the bivariate loss distributions, it is found that most emerging equity markets are independent in their limiting joint extremes. However, the dependence in finite levels of extremes is still much stronger than that implied by a correlation analysis. This implies that the benefits of portfolio diversification are lost when an investor needs them most. These results are consistent with the previous studies in the literature.

In a two-asset portfolio setting, it is documented that the diversification benefits, as measured by the percentage reduction in portfolio VaR, are negatively associated with the strength of extremal dependence between the components of the portfolio and to the ratio of marginal VaRs. The superiority of extremal dependence measures over the correlation coefficient in several aspects is emphasized in many studies. However, their superiority in explaining the diversification benefits better than the correlation coefficient is shown in this study for the first time.

Poon et al.'s (2003a,b) methodology for bivariate extremal dependence analysis
2.1 Introduction

is slightly generalized into a multi-asset setting following El-Gamal and Jaffe (2008). This is achieved by first generating two subportfolios that are mutually exclusive and collectively exhaustive and then analyzing the dependence structure between these two subportfolios. El-Gamal and Jaffe (2008) showed that it is possible to find subportfolios that exhibit stronger extremal dependence than all possible bivariate asset pairs within the portfolio. I will refer to this as “portfolio level hidden extremal dependence.” In this study, I show for the first time that the existence of hidden extremal dependence greatly reduces the diversification benefits. This new finding makes the dependence structure discovered by El-Gamal and Jaffe (2008) very relevant for portfolio management decisions.

2.1 Introduction

In Chapter 1, it was argued that the high volatility and the extreme movements in financial markets during the 1990s inspired the use of quantile-based risk measures such as the VaR and the expected shortfall. At the same time, the contagious nature of these extreme movements inspired the development of new dependence measures in monitoring the extreme comovement of asset prices. For a long period of time, Pearson’s correlation has been the most widely used measure of dependence. However, recent research uncovered its weakness in capturing the dependence during turmoil periods in financial markets. Consequently, much effort has been focused on developing new dependence measures focusing on the extremes. In fact, if an investor is comfortable with the assumption of multivariate normally distributed asset returns, Pearson’s correlation is the perfect choice for a dependence measure. However, it is well documented that asset returns exhibit skewness and a high degree of excess kurtosis, which implies fat tails. Additionally, correlation can only measure the linear
dependence between the variables, whereas nonlinear dependence is omitted. Moreover, Pearson's correlation averages the deviations from the mean giving equal weight to observations in the body and in the tails. Consequently, the body of the distribution dominates the calculated average (Embrechts et al. 1999). If the dependence in the joint tail region is stronger than the dependence in the body of the distribution, a correlation analysis may result in a significant underestimation of the probability of co-crashes. This is equivalent to overestimating the benefits of portfolio diversification.

In this study, I use extreme value dependence measures developed by Ledford and Tawn (1996, 1997, 1998) and Coles et al. (1999). These authors' studies made a clear and important distinction between the asymptotically independent and asymptotically dependent classes of extremal dependence structures. This distinction is related to the behavior of the variables as they approach their respective extremes. For asymptotically independent variables, the conditional probability of one variable being very extreme, in the sense that it exceeds a high quantile of its marginal distribution, given that the other is already very extreme, goes to zero as the quantile of interest approaches to 1. Formally, let

$$
\chi = \lim_{\alpha \to 1} \Pr(Z_1 > Z_{1,\alpha} | Z_2 > Z_{2,\alpha}),
$$

(2.1)

where $Z_{i,\alpha}$ \{i=1,2\} stands for the $\alpha^{th}$ quantile of the marginal distribution of variable $Z_i$. If $\chi = 0$, then $Z_1$ and $Z_2$ are said to be asymptotically independent. This independence needs careful interpretation. The independence here is in the joint tail region in a limiting sense. At the finite levels of extremes, these variables may still exhibit high levels of dependence. However, as we move further in the joint tail, they become independent from each other in the limit. Consequently, very extreme
2.1 Introduction

observations in these variables cannot happen simultaneously. On the other hand, for asymptotically dependent variables, the above limiting conditional probability does not converge to zero; that is, \( \chi = c > 0 \). Therefore, very extreme values in each variable might be observed simultaneously for asymptotically dependent variables.

Since \( \chi = 0 \) for all asymptotically independent pairs of variables, Coles et al. (1999) proposed a novel dependence measure \( \bar{\chi} \) in order to measure the extremal dependence within the family of asymptotically independent variables. For such variables, \( \bar{\chi} \) measures the speed of conversion to zero of the conditional probability in (2.1). Poon et al. (2003a,b) proposed a new methodology for the estimation and inference of these two dependence measures. Following their empirical methodology, I identify the asymptotic dependence structures of all possible pairs of emerging markets in the dataset. It is found that most emerging market pairs are asymptotically independent, but the extremal dependence in the finite levels of extremes is still much stronger than that suggested by a multivariate normal distribution and a correlation analysis. This result is consistent with the previous research. It implies that a multivariate normal distribution assumption for asset returns together with a correlation analysis overestimates the benefits of portfolio diversification. Actually, those benefits are lost when the investors need them most.

Then, the relationship between the diversification benefits and the extremal dependence structure between the components of a portfolio is studied in a two-asset portfolio setting. I quantify the diversification benefits as the maximum achievable percentage reduction in portfolio VaR and measure it for all possible two asset portfolios. It is found that asymptotically independent pairs of assets provide much higher diversification benefits than asymptotically dependent pairs. Another important determinant of diversification benefits is the ratio of marginal VaRs from the components of the portfolio. As this ratio gets higher, diversification benefits are reduced.
2.1 Introduction

The superiority of extremal dependence measures over the correlation coefficient in several aspects is emphasized in many studies. However, the ability of extremal dependence measures to explain diversification benefits is never studied. I estimated two separate regression models using diversification benefits as the dependent variable. The first model uses $\chi$, and the other uses the correlation coefficient as an explanatory variable in addition to the ratio of marginal VaRs. The results indicate that extremal dependence measure $\chi$ explains the variation in the diversification benefits better than the correlation coefficient.

Somewhat differently from most extremal dependence studies, I extend the bivariate analysis into higher dimensions by slightly generalizing Poon et al.’s (2003a,b) methodology. For simplicity, only three-asset portfolios are considered in this study. A linear portfolio of three assets is defined as

$$Z(w) = w_1 Z_1 + w_2 Z_2 + w_3 Z_3,$$

with $0 \leq w_i \leq 1$ and $w_1 + w_2 + w_3 = 1$. I investigate the extremal dependence between the subportfolio generated by the first and second assets, given by

$$\frac{w_1}{w_1 + w_2} Z_1 + \frac{w_2}{w_1 + w_2} Z_2,$$

and the third asset $Z_3$. It is found that in some cases the extremal dependence between the two-asset subportfolio and the third asset is stronger than the extremal dependence between any pairs of assets in the portfolio. In particular, it is possible for the two-asset subportfolio to be asymptotically dependent with the third asset even though all bivariate pairs of assets exhibit asymptotic tail independence. This point is first noted by El-Gamal and Jaffe (2008). However, there have been no stud-
ies looking at the portfolio implications of it. This type of dependence is referred as "portfolio level hidden asymptotic dependence," and its implications for portfolio risk and diversification benefits are investigated in this study for the first time. This is achieved by identifying the tail dependence structures of all possible three-asset portfolios and comparing the diversification benefits across two groups: those exhibiting hidden extremal dependence and those that do not. It is found that the existence of hidden asymptotic dependence greatly reduces diversification benefits, as measured by the opportunities of VaR reduction.

The rest of the chapter is organized as follows. Section 2.2 provides the background knowledge for extremal dependence analysis. In section 2.3 the relationship between the VaR of a portfolio and the extremal dependence structure between the components of the portfolio is investigated. Section 2.4 slightly generalizes the bivariate dependence analysis of Poon et. al. (2003a,b) into a multi-asset setting. Portfolio level hidden extremal dependence is introduced and the implications for portfolio management is analyzed. In section 2.5, conclusions is discussed.

2.2 Extremal Dependence

As background, it is necessary to provide a very brief overview of Hill's (1975) estimation method, which is an alternative way of estimating the shape parameter, $\xi$, in the limiting GPD for threshold exceedances. The idea is very similar to threshold exceedances that was explained in Chapter 1.
2.2 Extremal Dependence

2.2.1 Hill’s Estimation

The fat-tailed distributions with $\xi > 0$ exhibit a power decay rate in their tails given by:

$$F(x) = \Pr(X > x) = x^{-1/\xi} \ell(x) \quad \text{for} \quad x \geq u,$$

where $u$ is a high threshold and $\ell(x)$ is a slowly varying function. Further, $\xi$ is the same shape parameter in the limiting GPD for threshold exceedances given in (1.5) and (1.6). Note that

$$\Pr(X > x|X > u) = \frac{\ell(x) x^{-1/\xi}}{\ell(u) u^{-1/\xi}} = \left(\frac{x}{u}\right)^{-1/\xi}.$$

The second equality follows by assuming that $\ell(x) = c$ for any $x \geq u$ and for some constant $c$. This approximation can be made for a high enough threshold $u$ because $\ell(x)$ is a slowly varying function. Using those $N_u$ observations that exceed the threshold $u$, the log-likelihood function can be written as

$$\text{Log} \mathbf{L}(\xi, X) = \sum_{j=1}^{j=N_u} \left( - (\log \xi + \log u) - \left(1 + \frac{1}{\xi}\right) \log \left(\frac{x_j}{u}\right) \right).$$

Hill’s estimator is the closed form solution to this problem given by:

$$\hat{\xi} = \frac{1}{N_u} \sum_{j=1}^{j=N_u} \log \left(\frac{x_j}{u}\right).$$

Substituting $x = u$ in (2.3), and estimating $\Pr(X > u)$ nonparametrically with $N_u/N$, an estimator for the constant $c$ can be obtained as well.

$$\hat{c} = \frac{N_u}{N} u^{1/\xi}.$$

\footnote{A slowly varying function is identified with $\lim_{x \to \infty} \frac{\ell(tx)}{\ell(x)} = 1 \forall t$.}
2.2.2 Extremal Dependence Measures

The tail dependence measure $\chi$ is defined as

$$\chi = \lim_{\alpha \to 1} \Pr(Z_1 > Z_{1,\alpha} | Z_2 > Z_{2,\alpha}), \quad (2.6)$$

where $Z_{i,\alpha} \{i=1,2\}$ stands for the $\alpha^{th}$ quantile of the marginal distribution of the variable $Z_i$. If $\chi = 0$, then $Z_1$ and $Z_2$ are said to be asymptotically independent. If $\chi = c > 0$, then they are said to be asymptotically dependent. A bivariate joint distribution can be completely specified by the distributions of its marginal components and the dependence structure between these marginals. When studying dependence structures, it is very handy to get rid of the effects of marginals by transforming them to a common distribution. This can be achieved by using probability integral transforms, and it does not have any effect on the dependence measures since these measures are all quantile-based quantities. Following the common convention in the literature, unit Frechet margins are used in this study. Note that if the marginal distributions of two variables $X_1$ and $X_2$ are unit frechet,\(^2\) then the definition of $\chi$ above reduces to:\(^3\)

$$\chi = \lim_{r \to \infty} \Pr(X_1 > r | X_2 > r). \quad (2.7)$$

Ledford and Tawn (1996) showed that any bivariate joint distribution for with unit Frechet margins has to satisfy:

$$\lim_{r \to \infty} \Pr(X_1 > r, X_2 > r) = \ell(r) r^{-1/\eta}, \quad (2.8)$$

where $\ell(r)$ is a slowly varying function and $\eta \in (0, 1]$ is the coefficient of tail depen-

\(^2\)Pr($X_1 \leq r$) = exp($-1/r$) and $\lim_{r \to \infty} \Pr(X > r) = 1/r$.

\(^3\)Apply the change of variable $\alpha = 1 - \frac{1}{r}$ and let $r \to \infty$. 
2.2 Extremal Dependence

dence. Note that the coefficient of tail dependence is closely related to \( \chi \) and helps to identify asymptotic dependence and independence because by (2.7) and (2.8)

\[
\chi = \lim_{r \to \infty} \frac{\Pr(X_1 > r, X_2 > r)}{\Pr(X_2 > r)},
\]

\[
= \lim_{r \to \infty} \frac{\ell(r)r^{-1/\eta}}{1/r},
\]

\[
= \lim_{r \to \infty} \ell(r)r^{1-\frac{1}{\eta}}.
\]

The only way for \( \chi \) to be nonzero is to have \( \eta = 1 \) and \( \ell(r) \to c > 0 \). Otherwise, if \( \eta < 1 \), the variables will be asymptotically independent.\(^4\)

Since \( \chi = 0 \) for all asymptotically independent pair of variables, Coles et al. (1999) proposed a novel dependence measure \( \tilde{\chi} \) in order to measure the dependence within the family of asymptotically independent variables. They also showed its relation to \( \chi \) and the coefficient of tail dependence \( \eta \). Coles et al. defined \( \tilde{\chi} \) as

\[
\tilde{\chi} = \lim_{\alpha \to 1} \frac{2 \log \Pr(Z_2 > Z_{2,\alpha})}{\log \Pr(Z_1 > Z_{1,\alpha}, Z_2 > Z_{2,\alpha})} - 1.
\]

\(^4\)r\(^\alpha\) is a regularly varying function, whereas \( \ell(r) \) is a slowly varying function. If \( r\(^\alpha\) \) converges to 0, the whole expression will be converging to 0 as well, even if \( \ell(r) \) explodes. Regularly varying functions dominate the slowly varying functions.
the dependence measure $\bar{x}$ reduces to

$$\bar{x} = \lim_{r \to \infty} \frac{2 \log \Pr(X_2 > r)}{\log \Pr(X_1 > r, X_2 > r)} - 1$$

$$= \lim_{r \to \infty} \frac{2 \log \frac{1}{r}}{\log(\ell(r)r^{-1/\eta})} - 1$$

$$= \lim_{r \to \infty} \frac{-2 \log r}{\frac{1}{\eta} \log r + \log \ell(r)} - 1$$

$$= 2\eta - 1. \quad (2.11)$$

As a result, the only way to have asymptotic dependence is $\bar{x} = 1$. Therefore, we also have:

$$\bar{x} = 1 \text{ iff } \chi = c > 0$$

$$\bar{x} < 1 \text{ iff } \chi = 0. \quad (2.12)$$

The dependence measures $\chi$ and $\bar{x}$ together provide a complete characterization of the joint tail behavior. If the variables are asymptotically dependent, $\chi$ provides the strength of asymptotic dependence. If they are asymptotically independent, then $\chi = 0$ and the strength of dependence within this class is measured by $\bar{x}$. For a multivariate normal distribution, it is shown that $\bar{x} = \rho$, and this provides a way of comparing the strength of dependence in extremes with that implied by a multivariate normal distribution.

To test for asymptotic dependence, Poon et al. (2003a,b) suggest a three-step procedure that uses Hill’s estimator. First, the bivariate variables $(Z_1, Z_2)$ are transformed into $(X_1, X_2)$ to have unit Frechet margins by using the probability integral
2.2 Extremal Dependence

Transform:

\[ X_1 = \frac{-1}{\log F_{Z_1}(z_1)} \quad \text{and} \quad X_2 = \frac{-1}{\log F_{Z_2}(z_2)}, \]

(2.13)

where \( F_{Z_1} \) and \( F_{Z_2} \) are the marginal distribution functions of \( Z_1 \) and \( Z_2 \), respectively.\(^5\)

This transformation does not have any impact on the extremal dependence measures since they are quantile-based measures, and the transformation does not change the order of the data. In the second step, \( T = \min(X_1, X_2) \) is defined. Note that

\[ \Pr(T > r) = \Pr(X_1 > r, X_2 > r) = \ell(r) r^{-\eta}, \]

(2.14)

where the second equality follows from (2.8). Therefore, \( \eta \) is the shape parameter for the univariate variable \( T \) (see 2.3). Lastly, the coefficient of tail dependence \( \eta \) can be estimated by Hill’s estimator as explained in section 2.2.1. Then we have

\[ \hat{\chi} = 2\hat{\eta} - 1 = 2 \left( \frac{1}{N_u} \sum_{j=1}^{N_u} \log \left( \frac{T_j}{u} \right) \right) - 1, \]

where \( u \) is the high threshold for variable \( T \). Tests for the null hypothesis of \( \hat{\chi} = 1 \) (asymptotic dependence) can be administered using the MLE properties of Hill’s estimator. It can be shown that the asymptotic variance is

\[ \text{Variance}(\hat{\chi}) = \frac{(\hat{\chi} + 1)^2}{N_u} \]

If the asymptotic dependence hypothesis cannot be rejected, only then can \( \chi \) be

\(^5\)1/log(p) is the quantile function associated with unit Frechet distribution
2.2 Extremal Dependence

estimated as (see 2.5 and 2.9)

\[ \hat{\xi} = \lim_{r \to \infty} \ell(r) = \hat{c} = \frac{N_u}{N} u, \]

and its variance is given by

\[ \text{Variance}(\hat{\xi}) = \frac{u^2 N_u(N - N_u)}{N^3} \]

2.2.3 Extremal Dependence Analysis of Emerging Markets

In this section, Poon et al.’s (2003a,b) methodology is applied to determine the extremal dependence structures in the loss tails of all bivariate pairs of emerging markets in the dataset. While transforming marginals by

\[ X_1 = \frac{-1}{\log F_{Z_i}(z_i)} \text{ and } X_2 = \frac{-1}{\log F_{Z_2}(z_2)}, \]

the distribution functions of marginals \( F_{Z_i}(z_i) \) and \( F_{Z_2}(z_2) \) are needed. A GPD distribution is fit to the tails of marginals using the 0.95th quantile of the data as the threshold. Then, transformation to unit Frechet marginals is accomplished by using the empirical distribution function up to the threshold and the GPD fit beyond the threshold. More formally,

\[ F_{Z_i}(z_i) = \begin{cases} \tilde{F}_{Z_i}(z_i) & \text{if } z_i < u_{z_i} \\ 1 - (1 - \tilde{F}_{Z_i}(u_{z_i})) \left(1 + \frac{\xi_{z_i}(z_i - u_{z_i})}{\beta_{z_i}}\right)^{-1/\xi_{z_i}} & \text{if } z_i \geq u_{z_i}, \end{cases} \]

where \( \tilde{F}_{Z_i}(z_i) \) is the empirical distribution function of \( Z_i \) for \( i = \{1, 2\} \). Moreover, \( \xi \) and \( \beta \) are the MLE parameter estimates from the GPD fit to \( Z_i \), \( u_{z_i} \) is the 0.95th quantile of \( Z_i \) and so \( \tilde{F}_{Z_i}(u_{z_i}) = 0.95 \).
2.2 Extremal Dependence

Estimates for $\bar{\chi}$ are obtained by Hill's estimation on the univariate variable $T = \min(X_1, X_2)$, and the 0.95th quantile of $T$ is used as the threshold at this step as well. Parameter estimates and the associated probability values for the null hypothesis of $\bar{\chi} = 1$ are presented in Table 2.1. Most pairs of international stock markets are asymptotically independent (78 out of 91) from each other. There are only 13 pairs for which the null hypothesis of asymptotic dependence cannot be rejected at the 5 percent significance level, and these are given in bold. The estimates of $\chi$ for these 13 pairs together with their standard errors are given in Table 2.2.

The pairs that are asymptotically dependent are generally in geographic proximity to each other. In 12 of the 13 asymptotically dependent pairs, either both countries are Asian or both countries are Latin American. This is possibly because the linkages that can transfer the shocks are stronger among the countries that are in geographic proximity. These linkages might include trade competition and macroeconomic similarities, which will be discussed in Chapter 3.

Additionally, for almost all pairs (90 out of 91), the estimate of $\bar{\chi}$ is greater than Pearson’s correlation. The correlation matrix is given in Table 2.3. Considering the fact that $\bar{\chi} = \rho$ for a multivariate normal distribution, this comparison implies that the dependence in finite levels of extremes even within the asymptotically independent pairs is much stronger than that implied by a multivariate normal distribution. Therefore, a multivariate normality assumption and a simple correlation analysis, which has been the traditional methodology in finance for the last half century, is grossly underestimating the chances of joint crashes in multiple markets.
### Table 2.1: \( \chi \) Estimates and Probability Values for Daily S&P/IFCI Losses

<table>
<thead>
<tr>
<th></th>
<th>Tur</th>
<th>Ind</th>
<th>Mal</th>
<th>Phil</th>
<th>Kor</th>
<th>Chn</th>
<th>Twn</th>
<th>Thai</th>
<th>Bra</th>
<th>Chl</th>
<th>Mex</th>
<th>Arg</th>
<th>Per</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ind</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mal</td>
<td>0.34</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phil</td>
<td>0.18</td>
<td>0.72</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kor</td>
<td>0.34</td>
<td>0.6</td>
<td>0.67</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chn</td>
<td>0.34</td>
<td>0.55</td>
<td>0.54</td>
<td>0.49</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twn</td>
<td>0.43</td>
<td>0.3</td>
<td>0.33</td>
<td>0.4</td>
<td>0.48</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td>0.26</td>
<td>0.8</td>
<td>0.68</td>
<td>0.82</td>
<td>0.78</td>
<td>0.49</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bra</td>
<td>0.61</td>
<td>0.4</td>
<td>0.35</td>
<td>0.47</td>
<td>0.41</td>
<td>0.44</td>
<td>0.39</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td>0.64</td>
<td>0.39</td>
<td>0.47</td>
<td>0.46</td>
<td>0.44</td>
<td>0.44</td>
<td>0.38</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mex</td>
<td>0.64</td>
<td>0.4</td>
<td>0.44</td>
<td>0.43</td>
<td>0.36</td>
<td>0.47</td>
<td>0.27</td>
<td>0.34</td>
<td>0.92</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>0.49</td>
<td>0.34</td>
<td>0.24</td>
<td>0.27</td>
<td>0.41</td>
<td>0.25</td>
<td>0.24</td>
<td>0.82</td>
<td>0.69</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per</td>
<td>0.67</td>
<td>0.31</td>
<td>0.37</td>
<td>0.36</td>
<td>0.22</td>
<td>0.3</td>
<td>0.3</td>
<td>0.72</td>
<td>0.75</td>
<td>0.8</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>0.5</td>
<td>0.23</td>
<td>0.18</td>
<td>0.22</td>
<td>0.25</td>
<td>0.16</td>
<td>0.22</td>
<td>0.05</td>
<td>0.56</td>
<td>0.53</td>
<td>0.82</td>
<td>0.49</td>
<td>0.36</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2.2: \( \chi \) Estimates and Standard Errors for Daily S&P/IFCI Losses

<table>
<thead>
<tr>
<th></th>
<th>Tur</th>
<th>Ind</th>
<th>Mal</th>
<th>Phil</th>
<th>Kor</th>
<th>Chn</th>
<th>Twn</th>
<th>Thai</th>
<th>Bra</th>
<th>Chl</th>
<th>Mex</th>
<th>Arg</th>
<th>Per</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mal</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phil</td>
<td></td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bra</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mex</td>
<td></td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td></td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per</td>
<td></td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: \( \chi \) Estimates and Probability Values for Daily S&P/IFCI Losses

Table 2.2: \( \chi \) Estimates and Standard Errors for Daily S&P/IFCI Losses
2.3 Extremal Dependence and Diversification Benefits

In this section, diversification benefits in terms of risk reduction and their relation to the extremal dependence between the components of a portfolio is studied in a two asset portfolio setting. Diversification benefits are measured by the magnitude of maximum achievable percentage reduction in the VaR, and this quantity is calculated for all possible two-asset portfolios. Twenty-one portfolios from each pair of assets are created. The losses of portfolios are obtained by

$$Z(w) = w_1 Z_1 + w_2 Z_2,$$  \hspace{1cm} (2.15)

where $Z_i$ is the percentage losses of individual assets, and $0 \leq w_i \leq 1; w_1 + w_2 = 1,$ and $w_i \in G = \{0, 0.05, 0.1, ..., 1\}$. The VaRs for each of these portfolios are estimated.
2.3 Extremal Dependence and Diversification Benefits

with the EVT methodology introduced in Chapter 1. In Figure 2.1, this is illustrated for the Malaysia-Philippines pair. On the left end of the plot, all wealth is invested in the Philippines, and the $VaR_{0.999}$ is calculated as 8.359. On the right end of the plot, all wealth is invested in Malaysia, and the $VaR_{0.999}$ is calculated as 11.031. All other points are the VaR values calculated at the $0.999^{th}$ quantile for portfolios of these two countries. Weight given to Malaysia is incremented by 0.05 each time resulting in 21 points on the plot.

![Figure 2.1: Reduction in VaR from Diversification](image)

Instead of holding just the less risky asset (Philippines in this case), the VaR can be reduced by diversifying into the riskier asset (Malaysia in this case). “Maximum Percentage Reduction In Risk” is defined as:

$$MPRIR = 100 \left(1 - \frac{\min_{w_1 \in B} VaR_\alpha(Z(w))}{\min_{w_1 \in G} VaR_\alpha(Z(w))}\right), \quad (2.16)$$
where $B = \{0,1\}$ is the boundary set in which the investor is constrained to invest in only one country and $G = \{0,0.05,0.1,\ldots,1\}$ includes all possible portfolios. For the Philippines-Malaysia pair, $MPRIR_{0.999} = 100(1 - 7.898/8.359) = 5.51$, and so $Var_{0.999}$ can be reduced at most by 5.51 percent.

Since the VaR is an extreme loss risk of a portfolio, intuition suggests that those pairs with stronger extremal dependence should provide smaller diversification benefits resulting in smaller $MPRIR$. This conjecture turns out to be correct. Diversification benefits for each pair measured by $MPRIR_{i,j}$ at $\alpha = 0.999th$ quantile are presented in Table 2.4.

<table>
<thead>
<tr>
<th></th>
<th>Tur</th>
<th>Ind</th>
<th>Mal</th>
<th>Phil</th>
<th>Kor</th>
<th>Chn</th>
<th>Twn</th>
<th>Thai</th>
<th>Bra</th>
<th>Chl</th>
<th>Mex</th>
<th>Arg</th>
<th>Per</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ind</td>
<td>26.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mal</td>
<td>19.8</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phil</td>
<td>17.3</td>
<td>0.61</td>
<td>5.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kor</td>
<td>28.6</td>
<td>15.3</td>
<td>9.24</td>
<td>4.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chn</td>
<td>15.7</td>
<td>4.32</td>
<td>22.55</td>
<td>10.92</td>
<td>18.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twn</td>
<td>4.17</td>
<td>8.51</td>
<td>17.32</td>
<td>28.5</td>
<td>8.29</td>
<td>8.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td>15.44</td>
<td>7.07</td>
<td>15.93</td>
<td>10.27</td>
<td>13.64</td>
<td>23.63</td>
<td>10.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bra</td>
<td>9.86</td>
<td>13.09</td>
<td>25.92</td>
<td>15.65</td>
<td>25.01</td>
<td>29.38</td>
<td>17.74</td>
<td>27.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td>0</td>
<td>0</td>
<td>4.3</td>
<td>7.08</td>
<td>2.77</td>
<td>4.63</td>
<td>17.15</td>
<td>6.34</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mex</td>
<td>3.65</td>
<td>5.72</td>
<td>16.45</td>
<td>27.74</td>
<td>13.53</td>
<td>23.76</td>
<td>29.69</td>
<td>23.26</td>
<td>1.68</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>17.22</td>
<td>9.97</td>
<td>18.34</td>
<td>16.74</td>
<td>24.24</td>
<td>23.87</td>
<td>19.4</td>
<td>27.78</td>
<td>0</td>
<td>0</td>
<td>3.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per</td>
<td>4.38</td>
<td>8.03</td>
<td>19.05</td>
<td>26.94</td>
<td>13.54</td>
<td>20.63</td>
<td>31.94</td>
<td>22.33</td>
<td>0.27</td>
<td>1.56</td>
<td>13.72</td>
<td>1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>0.09</td>
<td>4.59</td>
<td>12.77</td>
<td>14.54</td>
<td>5.33</td>
<td>15.65</td>
<td>14.62</td>
<td>12.41</td>
<td>0.59</td>
<td>16.86</td>
<td>0</td>
<td>0.06</td>
<td>12.92</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4: Diversification Benefits at $\alpha = 0.999th$ Quantile

There are eight pairs in this table that have $MPRIR_{i,j} = 0$. For these pairs, diversification cannot reduce the risk at all, since the minimum VaR is achieved at a boundary point. In Figure 2.2, the case of Mexico-Chile pair is illustrated as an example. Five of these eight pairs with no diversification benefits (Ind-Mal, Bra-Chl,
Bra-Arg, Chl-Mex, Mex-USA) are among the asymptotically dependent pairs as well (see Table 2.2).

![Figure 2.2: No Reduction in VaR from Diversification](image)

The impact of asymptotic dependence structure on the diversification benefits is obvious from these extreme examples of no diversification cases. Besides that, the averages of MPRIRs for different dependence structures can provide a better understanding of this issue. These averages are presented for several quantiles of interest in Table 2.5.

<table>
<thead>
<tr>
<th>Tail Dependence Structure</th>
<th>Asymptotically Dependent</th>
<th>Asymptotically Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>α = 0.999</td>
<td>4.347</td>
<td>14.128</td>
</tr>
<tr>
<td>α = 0.99</td>
<td>5.747</td>
<td>11.539</td>
</tr>
<tr>
<td>α = 0.975</td>
<td>6.362</td>
<td>10.89</td>
</tr>
</tbody>
</table>

Table 2.5: Average Diversification Benefits for Different Tail Dependence Structures
2.3 Extremal Dependence and Diversification Benefits

At the 0.999th quantile, the average MPRIR for 13 asymptotically dependent pairs is 4.347 percent. On the other hand, the average MPRIR for remaining 88 asymptotically independent pairs is 14.128 percent. Therefore, the opportunity of risk reduction for asymptotically independent pairs is much higher than for asymptotically dependent pairs. The same conclusion can be driven for all levels of quantiles. The more interesting observation is that the average diversification benefit is decreasing for the asymptotically dependent pairs, whereas it is increasing for the asymptotically independent pairs as the level of quantile increases. Therefore, the effect of an asymptotic dependence structure becomes more significant for stronger market crashes.

In Table 2.5, only the average diversification benefits for different tail dependence structures were reported, which provides useful but limited information. Instead, in Figure 2.3 the kernel density plots for diversification benefits are presented, which provides the full information in terms of the distribution of diversification benefits across different dependence structures. These density plots further support the argument that asymptotically independent pairs result in higher diversification benefits compared to the asymptotically dependent pairs.

![Figure 2.3: Densities of Diversification Benefits for Different Dependence Structures](image)
2.3 Extremal Dependence and Diversification Benefits

The other critical factor affecting the diversification benefits is the “Marginal Risk Ratio” defined as the ratio of marginal $\text{VaR}$ measures for the two assets in the portfolio; that is, $MRR_{i,j} = \frac{V_{\alpha}R_i}{V_{\alpha}R_j}$. Bradley and Taqqu (2004) studied the effects of $MRR$ on the optimal weight of risky assets in a two-asset portfolio allocation problem. Inspired by their results, I investigated the effects of $MRR$ on diversification benefits. In general, the results indicate that, as one of the assets in the portfolio gets comparatively riskier than the other asset, the opportunity for $\text{VaR}$ reduction gets smaller. In particular, two of the eight pairs in the dataset with $MPRIR_{i,j} = 0$ have very high marginal risk ratios. These are Chile-Turkey with an $MRR = 3.33$ and Chile-Indonesia with an $MRR = 4.4$. These values are among the top four $MRR$s in 91 pairs. Marginal risk ratios at $\alpha = 0.999th$ quantile for all pairs can be seen in Table 2.6.

<table>
<thead>
<tr>
<th></th>
<th>Tur</th>
<th>Ind</th>
<th>Mal</th>
<th>Phil</th>
<th>Kor</th>
<th>Chn</th>
<th>Twn</th>
<th>Thai</th>
<th>Bra</th>
<th>Chi</th>
<th>Mex</th>
<th>Arg</th>
<th>Per</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ind</td>
<td>1.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mal</td>
<td>1.55</td>
<td>2.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phil</td>
<td>2.05</td>
<td>2.7</td>
<td>1.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kor</td>
<td>1.18</td>
<td>1.55</td>
<td>1.32</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chn</td>
<td>1.57</td>
<td>2.08</td>
<td>1.01</td>
<td>1.3</td>
<td>1.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twn</td>
<td>2.27</td>
<td>3</td>
<td>1.46</td>
<td>1.11</td>
<td>1.93</td>
<td>1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td>1.62</td>
<td>2.01</td>
<td>1.02</td>
<td>1.34</td>
<td>1.29</td>
<td>1.03</td>
<td>1.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bra</td>
<td>1.47</td>
<td>1.94</td>
<td>1.06</td>
<td>1.39</td>
<td>1.25</td>
<td>1.07</td>
<td>1.55</td>
<td>1.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi</td>
<td>3.33</td>
<td>4.4</td>
<td>2.15</td>
<td>1.63</td>
<td>2.83</td>
<td>2.12</td>
<td>1.47</td>
<td>2.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mex</td>
<td>2.06</td>
<td>2.72</td>
<td>1.33</td>
<td>1.01</td>
<td>1.78</td>
<td>1.31</td>
<td>1.1</td>
<td>1.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>1.34</td>
<td>1.77</td>
<td>1.16</td>
<td>1.53</td>
<td>1.14</td>
<td>1.18</td>
<td>1.7</td>
<td>1.14</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per</td>
<td>2.24</td>
<td>2.96</td>
<td>1.44</td>
<td>1.09</td>
<td>1.9</td>
<td>1.42</td>
<td>1.01</td>
<td>1.47</td>
<td>1.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>3.48</td>
<td>4.6</td>
<td>2.25</td>
<td>1.7</td>
<td>2.96</td>
<td>2.21</td>
<td>1.54</td>
<td>2.29</td>
<td>2.37</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.6: Marginal Risks Ratio at $\alpha = 0.999th$ Quantile
This idea can be better summarized on Figure 2.1. As the investor goes from left to right diversifying more into the riskier Malaysia, diversification benefits pull the curve down. This pull gets stronger as the extremal dependence gets weaker. At the same time, riskier Malaysia pulls the curve up, and this pull gets stronger as the MRR gets higher. The resulting diversification benefit will be a result of the combination of these two forces.

To investigate the relative importance of these two effects, I designed a regression model using $X_{ij}$ as a proxy for the strength of asymptotic dependence and included $MRR_{ij}$ as the other explanatory variable. The dependent variable in this regression is the diversification benefits measured by $MPRIR_{ij}$. The results of this regression model are provided in Table 2.7. Both the strength of extremal dependence and the ratio of marginal risks are significant determinants of diversification benefits at all conventional levels with expected negative signs.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>t stat.</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>41.2846</td>
<td>22.88</td>
<td>0.00***</td>
</tr>
<tr>
<td>$X_{ij}$</td>
<td>-27.6227</td>
<td>-11.52</td>
<td>0.00***</td>
</tr>
<tr>
<td>$MRR_{ij}$</td>
<td>-8.8084</td>
<td>-12.65</td>
<td>0.00***</td>
</tr>
<tr>
<td>Multiple $R^2$</td>
<td></td>
<td>0.7542</td>
<td></td>
</tr>
</tbody>
</table>

$MPRIR_{ij} = \alpha + \beta_1 X_{ij} + \beta_2 MRR_{ij} + \epsilon_{ij}$

Table 2.7: Diversification Benefits and Extremal Dependence

To compare the explanatory power of different dependence measures, I replaced $X_{ij}$ with Pearson’s correlation coefficient $\rho_{ij}$ as an explanatory variable and estimated the model again. The results of this regression model are presented in Table 2.8. Pearson’s correlation is also significant at all conventional confidence levels but with a smaller t-statistic than for the $\bar{X}$ in the first regression. Also, the $R^2$ of the regres-
2.4 Hidden Asymptotic Dependence and Diversification Benefits

The extremal dependence measures are developed for bivariate variables. When an investor holds a portfolio of more than two assets, he or she may be interested in monitoring the dependence between a group of assets as a subportfolio and remaining part of the portfolio. This is important, because sometimes a crash in just one market may not increase the concerns for the prospects of a crash in other markets, but a joint crash in multiple markets may lead to panic, making the propagation of crashes to other markets easier. In this section, this point is addressed by applying the extremal dependence analysis to subportfolios within a portfolio. For simplicity, only three-asset portfolios are considered in this study, but this empirical approach can be extended to any dimension. More specifically, the extremal dependence between the

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>t stat.</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>37.6103</td>
<td>21.89</td>
<td>0.00***</td>
</tr>
<tr>
<td>$\rho_{ij}$</td>
<td>-43.7829</td>
<td>-10.45</td>
<td>0.00***</td>
</tr>
<tr>
<td>$MRR_{ij}$</td>
<td>-8.9607</td>
<td>-12.14</td>
<td>0.00***</td>
</tr>
<tr>
<td>Multiple $R^2$</td>
<td>0.7251</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$MPRIR_{ij} = \alpha + \beta_1 \rho_{ij} + \beta_2 MRR_{ij} + \epsilon_{ij}$

Table 2.8: Diversification Benefits and Correlation Coefficient

The extremal dependence measure $\bar{\chi}$ explains the variation in VaR-diversification benefits better than Pearson's correlation coefficient.
subportfolio generated by two of the assets

\[ \frac{w_1}{w_1 + w_2} Z_1 + \frac{w_2}{w_1 + w_2} Z_2, \]

and the other asset \( Z_3 \) is investigated. An interesting example, which makes the main point of this section, is shown in Figure 2.4.

Figure 2.4: Portfolio Level Hidden Asymptotic Dependence

A triple of assets is chosen: Turkey-Mexico-Indonesia. The portfolio is split into two subportfolios. The first subportfolio is the Turkey-Indonesia combination, and the second subportfolio is Mexico alone. The horizontal axis shows the weight of Turkey in the first subportfolio; that is, \( \frac{w_1}{w_1 + w_2} \). The solid line shows the estimate of \( \bar{x} \) between the two subportfolios, and the dashed lines are the 95 percent confidence intervals for \( \bar{x} \). On the very left end of the curve, \( \bar{x} \) for the Mexico-Indonesia pair is equal to 0.4, and the confidence intervals reveal that the asymptotic dependence (\( \bar{x} = 1 \) hy-
2.4 Hidden Asymptotic Dependence and Diversification Benefits

Hypothesis is rejected. On the very right end of the curve, \( \bar{\chi} \) for the Mexico-Turkey pair is equal to 0.64, and the null hypothesis of asymptotic dependence is rejected again. However, there are some portfolio weights for which Mexico is asymptotically dependent with the subportfolio generated by Turkey and Indonesia (0.6 ≤ \( \frac{w_1}{w_1 + w_2} \) ≤ 0.8). This is referred to as the “portfolio level hidden asymptotic dependence” because the conventional bivariate dependence analysis is unable to identify such a dependence structure. Hidden dependence implies that the prospects of a crash in Mexico conditional on a crash of a portfolio composed of Turkey and Indonesia is stronger than the prospects of a crash in Mexico conditional on a crash only in Turkey or only in Indonesia.

To my knowledge, the only paper that employs the above-defined methodology is El-Gamal and Jaffe(2008). The contribution of this study is the application of this idea to a large set of emerging market data, thereby investigating the implications for portfolio management. The percentage loss of a three-asset linear portfolio can be defined as

\[
Z(w) = w_1Z_1 + w_2Z_2 + w_3Z_3,
\]

(2.17)

where \( Z_i \) is the percentage loss of individual assets, and \( w \in G = \{ w : w_1 + w_2 + w_3 = 1, \ 0 \leq w_i \leq 1, \ w_i = 0.05k, \ k \in Z \} \). The portfolios are created again using a grid of an increment size equal to 0.05. This grid generates 231 possible portfolios from each choice of three assets, as illustrated in Figure 2.5. Possible weights that can be assigned to first and second assets are shown on the \( x \) and \( y \) axes respectively. Weight given to the third asset is determined by \( 1 - w_1 - w_2 \). The corners of the triangle correspond to single-asset portfolios, and the sides correspond to two-asset portfolios.

The definition of \( MPRI R \) that was introduced in the previous section is extended
2.4 Hidden Asymptotic Dependence and Diversification Benefits

Figure 2.5: Weights Used in the Three Asset Portfolio Problem

into three asset setting as:

\[ MPRIR = 100 \left( 1 - \frac{\min_{w \in B} VaR_{\alpha}(Z(w))}{\min_{w \in G} VaR_{\alpha}(Z(w))} \right), \tag{2.18} \]

where \( B = \{ w : w_1 + w_2 + w_3 = 1; \ 0 \leq w_i \leq 1; \ w_i = 0.05k; \ k \in \mathbb{Z}; \ \exists i \ w_i = 0 \} \)

is the boundary set with which the investor is constrained to invest in at most two countries. The set \( G \) includes all possible portfolios as before. In Figure 2.5, the set \( B \) is shown with the bullet circles, whereas the set \( G \) includes all points in the plot. Therefore, in a three-asset setting, MPRIR measures the additional benefit of diversifying into the third asset. The VaR for all portfolios is calculated by following the EVT methodology explained in Chapter-1.
2.4 Hidden Asymptotic Dependence and Diversification Benefits

Triples of assets can be chosen in 364 different ways from a pool of 14 assets. All of these choices are investigated and classified according to their tail dependence structure. First, the conventional bivariate dependence analysis is administered. If any two assets in the portfolio are asymptotically dependent with each other, such portfolios are eliminated from consideration to control for the effects of pairwise direct asymptotic dependence. In the remaining 229 triples, it is found that all possible pairs are asymptotically independent from each other. From these remaining 229 triples, portfolios are created using weights illustrated in Figure 2.5 and existence of portfolio-level hidden dependence is checked. It is found that 62 triples exhibit hidden asymptotic dependence, whereas 167 of them do not. The average diversification benefits, as measured by MPRIR at several quantiles are presented for both groups in Table 2.9. Comparing the reported numbers with those reported in Table 2.5 reveals that the average reduction in the VaR by additional diversification into a third asset is much lower than the average reduction in the VaR obtained from diversification into the second asset. This can be regarded as the diminishing marginal benefit of diversification.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Hidden Asymptotic Dependent</th>
<th>Asymptotically Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.999</td>
<td>2.167</td>
<td>4.04</td>
</tr>
<tr>
<td>0.99</td>
<td>1.052</td>
<td>2.949</td>
</tr>
<tr>
<td>0.975</td>
<td>0.981</td>
<td>2.69</td>
</tr>
</tbody>
</table>

Table 2.9: Average Diversification Benefits for Different Dependence Structures in Three Asset Setting

The average MPRIR at the 0.999th quantile for the 62 triples that exhibit hidden asymptotic dependence is 2.167 percent. On the other hand, the average MPRIR at the 0.999th quantile for the 167 triples that does not have hidden asymptotic
dependence is 4.04 percent. Therefore, the existence of hidden asymptotic dependence reduces the diversification benefits, and this same conclusion can be drawn for all quantiles. However, as the level of the quantile is increased, the behavior of MPRIR is different from what it was for pairwise direct asymptotic dependence given in Table 2.5. In pairwise direct asymptotic dependence, MPRIR was decreasing as the quantile of interest increased. Here, it is just the opposite. This implies that the consequences of a pairwise direct asymptotic dependence is more severe than the hidden asymptotic dependence.

In Table 2.9, only the average diversification benefits for different tail dependence structures were reported. In Figure 2.6 the kernel density plots for diversification benefits are presented that provide the full information on the distribution of diversification benefits across different tail dependence structures. These density plots further support the argument that the portfolios with hidden asymptotic dependence result in lower diversification benefits.

Figure 2.6: Densities of Diversification Benefits in Three Asset Setting
2.5 Conclusion

Market turmoil spread quickly through the global markets in recent episodes of financial crises. The analysis of comovement in asset prices, especially that focusing on joint losses provides valuable information for the riskiness of financial positions. In this study, the extremal dependence measures driven from extreme value theory are estimated for all pairs of emerging market stock indices in the dataset. Thirteen pairs out of 91 exhibit asymptotic tail dependence. These are generally pairs of countries that are in geographic proximity to each other. Most pairs exhibit asymptotic independence, but this point needs careful interpretation. The independence here is in the joint tail region in a limiting sense. At the finite levels of extremes, almost all pairs exhibit stronger dependence than that implied by a multivariate normal distribution and correlation analysis. The implications are particularly important in understanding the market events of recent decades. Markets are dominated by large investors who use standard Geometric Brownian Motion models in constructing portfolios. These models not only underestimate the risk of a crash for individual assets, as shown in Chapter 1, but also the risk of joint crashes in multiple markets.

The diversification benefits in a two-asset portfolio allocation problem are quantified with the MPRIR for all pairs of assets. The MPRIR is defined to be the percentage reduction in the VaR as a result of choosing the minimum VaR portfolio instead of allocating all resources to the less risky asset. Two main forces affecting diversification benefits are identified. These are the extremal dependence structure between the components of the portfolio and the ratio of the marginal risks of these components. The benefits of diversification are much higher for asymptotically independent pairs compared to the asymptotically dependent pairs. Additionally, the benefits of diversification increase as the ratio of marginal risks decreases.
To investigate the dependence within a portfolio of more than two assets, the bivariate dependence methodology proposed by Poon et al. (2003a,b) is slightly generalized. It is shown that in a multi-asset portfolio, the dependence between the subportfolios may be stronger than the bivariate dependencies between the individual assets. In particular, it is possible to find asymptotically dependent subportfolios even when all pairs of assets in the portfolio are asymptotically independent of each other. This type of dependence structure is referred to as the “portfolio level hidden asymptotic dependence” because it cannot be uncovered with the regular practice of bivariate dependence analysis. By generalizing the MPRIR measure into a multi-asset portfolio setting, it has been shown that those three-asset portfolios exhibiting hidden asymptotic dependence provide less diversification benefits. This study was the first in the literature that investigated the implications of hidden asymptotic dependence in portfolio allocation.
Chapter 3

Extremal Dependence and International Financial Contagion

Summary. Financial contagion is defined as the spread of financial market turmoil beyond the borders where it originated. In this study, the channels through which crisis may propagate are discussed by a review of the theoretical literature. Then, trade and macro-similarity channels of financial contagion are tested with a novel regression model. The trade competition is confirmed as a significant determinant of financial contagion. The similarities of macroeconomic indicators are also jointly significant and increase the strength of financial contagion.

3.1 Introduction

In most episodes of financial crises after the 1990s, financial distress behaves somewhat like a contagious disease among emerging markets. In fact, the term “contagion” acquired a new meaning among economists after these episodes of crises. In economics, contagion is defined as the spread of financial market turmoil beyond the borders
3.1 Introduction

where it originated.

There are many alternative channels through which international financial contagion may occur resulting in the extreme comovement of asset prices around the globe. In this paper, a brief overview of these channels is discussed, and I proceed with testing the trade and macro-similarity channel of financial contagion.

There have been two common empirical approaches in testing the trade channel. One approach is choosing a country as the first victim of a financial crisis and assuming that the crisis spreads only in one direction from the ground-0 country to others. Glick and Rose (1999) is the leading empirical study in this line of research. In their approach, a cross-country probit model is estimated. A binary crisis indicator variable for each country $i$ is used as the dependent variable, and it is regressed on a proxy variable measuring the level of trade competition between the ground-0 country and country $i$.

The second approach assumes that contagion may occur in any direction. Hence, there is no ground-0 country. The dependent variable is still a binary crisis indicator for each country $i$. However, the explanatory variable is the summation of trade competition variables across all other countries in crisis. Eichengreen et al. (1999) applies this approach with a Probit-MLE as the method of estimation.

While the first approach suffers from the assumption of a ground-0 country so that contagion may not occur in any direction, the second approach suffers from the fact that the trade variables cannot be used as the regressors, but they are pooled across all countries in crisis.

In this paper, I propose an alternative regression model for testing the trade channel of financial contagion. The point estimate of $\bar{\chi}$, which is introduced in Chapter 2, is used to measure the strength of contagion between two countries. It will be regressed on the trade competition variable. This approach does not assume a ground-0
country so that contagion is possible in any direction. Also, trade competition variables are not pooled across several countries. Instead, they are used directly as the explanatory variables. Another advantage of this new regression model is that the dependent variable is continuous instead of binary.

After testing for the trade channel, macro-similarity indices are included in the regression as control variables. Although this model is novel, the conclusions are consistent with the previous literature. Trade competition in export markets is identified as a possible explanation of the strength of financial contagion between any two countries. Similarities of macroeconomic indicators, such as GDP growth rate are not significant determinants of financial contagion. However, macroeconomic similarities are jointly a significant determinant of financial contagion.

The rest of the paper is organized as follows. In section 3.2, theoretical literature on the channels of financial contagion is reviewed. In section 3.3, the two common empirical methodologies in testing the trade channel of contagion are reviewed in detail. The alternative model of this study is proposed, and the estimation results are presented. In section 3.4, macro-similarity channel of financial contagion is tested. In section 3.5, conclusions are discussed.

3.2 Channels of Financial Contagion

Trade linkages are well emphasized in the literature. In emerging markets, currency devaluation and financial market turmoil are generally observed simultaneously. As a result of currency devaluation, trade partners of a crisis country experience deterioration in their trade balances. This is because they lose their competitiveness in both bilateral trade with the crisis country and trade with third countries. This, in turn, places pressure on the currencies of other countries that have high levels of bilateral
3.2 Channels of Financial Contagion

trade with the crisis country or that have high levels of trade competition in third markets with the crisis country.

Another channel for contagion is the similarities of macroeconomic indicators. After a price shock in the financial markets of a country, investors may get concerned about other countries, which have similar macroeconomic indicators. This concern may cause selling behavior for assets in these countries, which in turn results in comovement of prices and extremal dependence.

Financial links are another fundamental source of contagion. Countries, which are financially integrated to the global economy and whose financial assets are being traded in global markets, might be more vulnerable to contagion because it is a common behavior that investors sell in multiple markets when they are faced with a shock in one of them. It is sometimes caused by the rules of thumb in portfolio allocation as pointed out by Schinasi and Smith (2001). In their model, when there is a shock to an asset in the portfolio, it has two separate effects on the demand for other risky assets. The substitution effect makes them more attractive, whereas the income effect makes the entire portfolio less attractive. The selling-off in other markets may also be a result of the existence of liquidity-constrained investors. Leveraged investors may face margin calls after a shock in one market, and they try to sell assets in other markets whose prices have not collapsed yet in anticipation of raising cash and meeting margin calls, as in Calvo (1999). Common bank lending is another financial link that can produce similar effects. When credits in one emerging market default, international banks might cut the credits to other emerging markets in order to reduce their exposure to emerging markets in general. They may refuse to roll over the short-term credits and call back the callable ones.

On the other hand, contagion might not be related to any macroeconomic or financial fundamental linkages. For example, it can be a result of market imperfections
such as information asymmetry. Calvo (1999) models the high costs of obtaining accurate information on emerging markets. The investors who can afford the cost of information are likely to be those having leveraged portfolios. When informed investors start selling, uninformed investors cannot distinguish between the possible reasons for this selling. Informed investors may be selling due to the arrival of new information, or they may be selling simply because they are faced with margin calls. Under some conditions, uninformed investors believe that there is a worsening of fundamentals that they cannot observe but informed investors can. Consequently, they just imitate the selling behavior of informed investors even if there is nothing wrong with the fundamentals of the economy. Kodres and Pritsker (1999) model a portfolio rebalancing problem that has three types of agents: informed, uninformed, and liquidity (noise) traders. After showing how portfolio rebalancing causes contagion, they also show that the magnitude of contagion increases as information asymmetries increase.

Another channel of contagion not related to economic fundamentals is multiple equilibria. In the second-generation models of currency crises first proposed by Obstfeld (1994), speculators' beliefs that the currency will be depreciated turn out to be self-fulfilling. Shifts in self-fulfilling expectations can cause the economy to jump from good equilibrium to bad equilibrium in a very similar way to the bank run model of Diamond and Dybvig (1983). Jeanne (1997) and Masson (1998) identified three regions for an economy's fundamentals. If the fundamentals are very strong, there is only a good equilibrium; if they are too weak, the only equilibrium is the bad one. If the fundamentals are in the intermediate region, both equilibria are possible. A successful speculative attack on one country's currency may play the role of a sunspot for some other country if the latter's fundamentals are in the intermediate region. This may result in a jump from good equilibrium to bad equilibrium in the latter.
3.3 Testing for Trade Channel

There have been many empirical studies testing for different channels of contagion. In this paper, a new method of testing for the trade channel of financial contagion is proposed.

3.3.1 Existing Methodologies

There have been two common empirical approaches in testing for the trade channel. One approach is choosing a country as the first victim of a financial crisis and assuming that the crisis spreads only in one direction, from the ground-zero country to others. Glick and Rose's study (1999) is one of the leading studies using this approach. To test for the trade channel of contagion with this approach, one may simply run a regression of the form:

\[ \text{Crisis}_i = \beta T_{0,i} + \epsilon_i, \]  
(3.1)

where \( \text{Crisis}_i \) is a binary indicator of crisis for country \( i \) determined with anecdotal evidence from financial newspapers and \( T_{0,i} \) is a measure of trade competition between country \( i \) and ground-zero country given by

\[ T_{0,i} = \sum_{k \neq i} \left\{ \frac{x_{0k} + x_{ik}}{x_0 + x_i} \cdot \left[ 1 - \frac{x_{ik} - x_{0k}}{x_{ik} + x_{0k}} \right] \right\}, \]  
(3.2)

where \( x_{ik} \) denotes bilateral exports from country \( i \) to country \( k(k \neq i) \) and \( x_i \) denotes aggregate bilateral exports from country \( i \). This index is a weighted average of the similarities of the countries \( i \) and 0 in their export patterns to all other countries around the globe. The summation of first the terms in the above expression across all \( k \) is equal to 1. These terms can be thought of as weights that represent the importance of country \( k \) in the aggregate trade of countries 0 and \( i \). The second
3.3 Testing for Trade Channel

term measures the similarities of country 0 and i in their exports to country k. This similarity gets larger as the amount of exports of countries i and 0 to country k get closer to each other.

The idea behind this measure can be understood better by thinking of country k as the United States, since it is the most important export market for emerging countries. As the proportion of exports to the United States from the two countries to their aggregate exports gets higher, the weight given to the United States increases. Also, as the exports of these two countries to the United States get closer to each other, the similarity term gets closer to its maximum value, which is 1. As such, this measure captures the effect of a consumption shock in a big export market as a reason for financial contagion among its exporters.

Glick and Rose (1999) estimate the above model with a Probit-MLE. A positive and significant $\beta$ implies that higher trade competition in export markets with the initial victim of a crisis increases the probability of successive crises in other countries.

The second approach in testing the trade channel of contagion assumes that contagion may occur in any direction. Eichengreen et al.'s study (1996) is the leading study using this approach. With this approach, one may simply run the regression

$$Crisis_i = \beta \sum_{j \neq i} T_{ij}(Crisis_j) + \epsilon_i,$$

where

$$T_{ij}(Crisis_j) = \begin{cases} T_{ij} & \text{if } Crisis_j = 1 \text{ for any } j \neq i \\ 0 & \text{otherwise}, \end{cases}$$

Here, $Crisis_i$ is a crisis indicator for country i, and $T_{ij}$ represents the level of trade competition between countries i and j. Eichengreen et al. (1996) used MERM weights
3.3 Testing for Trade Channel

published by the IMF for the trade competition variable and extreme observations (2 standard deviations away from the mean) of an exchange market pressure index (EMPI) to identify the binary crisis indicator.\(^1\)

3.3.2 Alternative Model

Note that in the first approach where a ground-0 country is chosen, the binary crisis variable is regressed on the trade competition variable. However, in the second approach, it is regressed on a weighted average of “crisis elsewhere” variables where the weights are the trade competition variables. Therefore, the trade competition variables of all crisis countries are pooled in a single variable. While the first approach suffers from the assumption of a ground-0 country so that contagion may not occur in any direction, the second approach suffers from the fact that the trade variables cannot be used as the regressors but only as the weighting coefficients in constructing the regressors.

The alternative procedure proposed in this paper aims to solve this problem. By constructing pairs of countries and measuring the strength of contagion between countries \(i\) and \(j\) with the point estimate of \(\bar{x}_{ij}\), it is possible to use the trade competition variable \(T_{ij}\) as the regressor and also allow the contagion to occur in any direction. Another advantage of this approach is that it uses a continuous variable that directly quantifies the strength of contagion in the left-hand side instead of using a binary crisis indicator. The regression proposed has the following form:

\[
\bar{x}_{ij} = \alpha + \beta T_{ij} + \epsilon_{ij}. \tag{3.3}
\]

Here, the trade competition variable is constructed by generalizing the measure

\(^{1}\text{EMPI is constructed as a weighted average of the percentage change in exchange rates, central bank reserves, and interest rates.}\)
suggested by Glick and Rose (1999). It can be calculated for any pair of countries (instead of fixing a ground-0 country) as

\[ T_{ij} = \sum_{k \neq i,j} \left\{ \frac{x_{ik} + x_{jk}}{x_i + x_j} \cdot \left[ 1 - \frac{x_{jk} - x_{ik}}{x_{jk} + x_{ik}} \right] \right\}, \]  

(3.4)

where \( x_{ik} \) denotes bilateral exports from country \( i \) to country \( k(k \neq i) \) and \( x_i \) denotes aggregate bilateral exports from country \( i \). The direction of trade statistics database of the IMF is used to construct the trade competition variable \( T_{ij} \). Trade data for 162 countries were downloaded from this database because in calculating the trade competition variable, (3.4) uses export data from the thirteen countries in the dataset to all countries in the world. The data for Taiwan are not available, and so trade competition variables are calculated for 78 pairs constructed from the thirteen remaining countries. The results of the OLS regression given in (3.3) are presented in Table 3.1.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t.stat.</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.358</td>
<td>0.045</td>
<td>7.935</td>
<td>1.47e-11</td>
</tr>
<tr>
<td>Trade</td>
<td>0.431</td>
<td>0.125</td>
<td>3.446</td>
<td>0.00093</td>
</tr>
<tr>
<td>Multiple R-Squared</td>
<td>0.1351</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint-F-statistic</td>
<td>11.87</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Regression Results of Extremal Dependence on Trade Competition

The coefficient for trade competition is significant at all conventional significance levels. By using probit models, the previous empirical literature established that higher trade competition with a crisis country significantly increases the probability of having a successive crisis. Our approach using a continuous dependent variable establishes that higher trade competition between any pair of countries increases the strength of extremal dependence between their financial market losses. This is con-
sistent with the results of the extremal dependence analysis in Chapter 2. The 13 asymptotically dependent pairs were all in geographic proximity to each other. The countries in geographic proximity are trade partners of each other and also compete in export markets with each other. As a result, they exhibit stronger extremal dependence with each other.

3.4 Testing for the Macro-Similarity Channel

To test for the macro-similarity channel of financial contagion, macro-similarity indices are created and added to the regression model as control variables. The similarity indices are created for (1) GDP growth rate, (2) domestic credit growth rate, (3) current account/GDP ratio, (4) CPI growth rate. Following Gregorio and Valdes (2001), the similarity indices are calculated as

$$M_{ij,k} = \exp(-|x_{i,k} - x_{j,k}|)$$

(3.5)

where $x_{i,k}$ is the $k^{th}$ standardized macro variable for country $i$ and where $k \in \{1, 2, 3, 4\}$. Standardization is done by subtracting the country mean and dividing by the country standard deviation. After calculating these macro-similarity indices, the following model is estimated.

$$\hat{\chi}_{ij} = \alpha + \beta T_{ij} + \sum_{k=1}^{4} \theta_k M_{ij,k} + \epsilon_{ij}. $$

(3.6)

The results of this estimation are presented in Table 3.2. The trade competition variable is still significant at 95 percent confidence level. All the coefficients for macro-similarity variables have correct positive signs. Only the credit growth similarity is significant at the 90 percent confidence level. None of the macro-similarity
variables are significant at 95 percent confidence level. However, the joint F-statistic for all macro-similarity variables is significant at all conventional confidence levels; hence, macro-similarities are significant all together. This implies colinearity between different macro-similarity variables, which is not surprising.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t stat.</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.17329</td>
<td>0.18063</td>
<td>-0.959</td>
<td>0.3406</td>
</tr>
<tr>
<td>Trade</td>
<td>0.30340</td>
<td>0.13752</td>
<td>2.206</td>
<td>0.0306</td>
</tr>
<tr>
<td>Current Account/GPD</td>
<td>0.28315</td>
<td>0.19922</td>
<td>1.421</td>
<td>0.1595</td>
</tr>
<tr>
<td>Credit Growth</td>
<td>0.51863</td>
<td>0.27164</td>
<td>1.909</td>
<td>0.0602</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>0.02263</td>
<td>0.25750</td>
<td>0.088</td>
<td>0.9302</td>
</tr>
<tr>
<td>CPI Growth</td>
<td>0.32895</td>
<td>0.22052</td>
<td>1.492</td>
<td>0.1401</td>
</tr>
<tr>
<td>Multiple R-Squared</td>
<td></td>
<td></td>
<td></td>
<td>0.2813</td>
</tr>
<tr>
<td>Joint F-statistic</td>
<td></td>
<td></td>
<td></td>
<td>5.116</td>
</tr>
<tr>
<td>Joint F- (Macro-similarities)</td>
<td></td>
<td></td>
<td></td>
<td>3.73</td>
</tr>
</tbody>
</table>

Table 3.2: Regression Results Including Macro-Similarity Control Variables

3.5 Conclusion

To explore the mechanisms driving the financial contagion among emerging markets, a new regression model is developed with the extremal dependence measure on the left-hand side and trade competition and macro-similarity indices on the right-hand side. Trade competition is a significant determinant of extremal dependence, whereas macro-similarity variables are not individually significant. However, macro-similarity variables exhibit colinearity and they are jointly significant.
Chapter 4

Fundamentals and Natural Gas
Volatility Dynamics: Predictive Power of Augmented GARCH Models in Volatility and Risk Forecasting

Summary. I investigate the determinants of the short-term volatility for natural gas nearby month futures within a GARCH framework augmented with market fundamentals. Consistent with the previous literature, I found that volatility is much higher on storage report announcement days and Mondays. Additionally, volatility is higher on winter days, defined as December, January, and February. Samuelson’s (1965) hypothesis is also investigated, and a significant time to maturity effect is detected. The effect of the interaction between the storage levels and the seasonality on the short-term volatility is discussed as well. During the winter, lower storage levels than
the seasonal norms result in higher volatility. However, during non-winter months the effect is just the opposite: higher storage levels than the seasonal norms result in higher volatility. Another factor contributing to the high short-term volatility is the weather shocks in excess of seasonal norms. There is a gap in the literature concerning the out-of-sample forecasting accuracy of GARCH models augmented with market fundamentals. This gap is filled with this study. I found that GARCH models augmented with market fundamentals reduce the mean absolute error and mean squared error of simple GARCH models in out-of-sample volatility forecasting.

4.1 Introduction

Natural gas futures contracts began trading on the New York Mercantile Exchange (NYMEX) in April 1990. One contract is written on 10,000 MMBTU of natural gas to be delivered to Henry Hub.\footnote{1 MMBTU = 1 Million British Thermal Units (BTU). A BTU is a unit used to describe the heat value (energy content) of fuels. A BTU is defined as the amount of heat required to raise the temperature of one pound of liquid water by one degree from 60°F to 61°F at a constant pressure of one atmosphere. (www.wikipedia.org)}\footnote{2 Henry Hub is a point on the natural gas pipeline system in Erath, Louisiana. It interconnects with nine interstate and four intrastate pipelines. Spot and future prices set at Henry Hub are denominated in $/MMBTU and are generally regarded as the primary price set for the North American natural gas market. (www.wikipedia.org)} Contracts for delivery in each month and for up to six years out are traded at any point in time. Trading in a given contract ends three business days before the first calendar day of the delivery month.

Trading in natural gas futures has skyrocketed in recent years. Open interest in Nymex natural gas futures grew at a rate of 15.2 percent per year during the last decade (Wei, Linn and Zhu, 2007). Daily volume is in the order of 60,000 to 100,000 contracts for the nearby month futures and 20,000 to 60,000 contracts for the second
nearby futures.\textsuperscript{3} \textsuperscript{4}

The volatility of natural gas prices has received increasing attention in recent years. The extreme fluctuations in both spot and futures prices caused researchers and market practitioners to focus on the sources of this high volatility. Whether news about natural gas market fundamentals, or excessive speculation and irrational investor behavior is responsible for the high volatility is an ongoing debate. In this chapter, empirical evidence is provided that natural gas futures price volatility is driven by market fundamentals within a GARCH type of dynamic volatility framework.

The response of prices to shifts in supply and demand depends on price elasticity of the commodity. In general, natural gas markets are highly inelastic in both the supply and demand side; hence, the price is very responsive to short-term changes in both, which results in high volatility. Two key fundamental pieces of information affecting the natural gas markets are the level of working gas in storage facilities and weather changes. These are generally regarded as proxies for supply and demand.

For supply conditions, the storage report is perceived as the most important piece of information by natural gas market participants. The report is currently prepared by the Energy Information Administration (EIA) of the Department of Energy (DOE). Anecdotal evidence on the effect of the natural gas storage report is abundant in the financial press. The following appeared in Communications, Energy and Paperworkers Union of Canada (CEP) News on Thursday, October 30, 2008:

\textit{Underground natural gas storage in the U.S. increased 46 billion cubic feet (Bcf) in the week ending Oct. 24, according to the Energy Information Administration (EIA)’s weekly report on Thursday. Expectations had been}

\textsuperscript{3}Nearby Future Contract is the earliest maturing contract. This corresponds to the next month delivery contract for natural gas futures.

\textsuperscript{4}Second Nearby Contract is the second earliest maturing contract in natural gas markets; it corresponds to the contract for delivery on the month after the next month.
for a 41 Bcf increase. Following the report, natural gas prices bottomed out, and are now at new session lows of 6.528 from pre-report levels of 6.811.

The report provides the level of total underground working gas and the historical average of this quantity for the equivalent time periods of last five years. Although the effects of storage surprises on short-term volatility have been emphasized in many studies (Gregoire and Boucher 2008; Mu 2007; Linn and Zhu (2004)), I did not identify any research considering the effects of the interaction between the storage surprises and seasonality. This chapter addresses this issue. In the winter, natural gas demand spikes, and the supply is unable to react quickly since the production of natural gas is uniform across seasons. When this happens, low storage levels, as compared to historical averages may be regarded as tight supply situations and put pressure on gas prices, which results in high volatility. Conversely, during the other seasons, higher storage levels than historical averages may increase concerns regarding the capacity of storage facilities and result in highly volatile natural gas prices. In this chapter, supporting evidence is presented for this hypothesis, which implies asymmetric effect of storage surprises in different seasons.

For demand conditions, the key information for short-term volatility dynamics is a change in the weather. Upon the observation of unexpected cold weather during the winter months, the extra demand for heating pushes the prices up causing higher volatility since the supply cannot adjust to such changes in the short run. In recent years, power generation plants have used more natural-gas-fueled technology. Natural gas usage in electricity generation rose from 12 percent to 17 percent between 1990 and 2006 (Hartley et al., 2007.a). As a result, hotter than expected temperatures

---

5 Working gas in storage is the volume of gas in the reservoir that is in addition to the cushion or base gas. Base gas is the volume of gas needed as a permanent inventory to maintain adequate reservoir pressures. (www.doe.eia.gov)
during the summer months increase the demand for cooling and might have a similar effect on volatility. Empirical evidence is presented that weather anomalies, measured as the degree days in excess of seasonal norms, result in increased short-term volatility.

In addition to the supply-and-demand driven volatility, the previous literature found significantly higher volatility on Mondays (Mu 2007; Murry and Zhu 2004). The study of Fleming et al. (2004) explains this effect based on the continuous weather information flow during the long nontrading weekend period.

Additionally, time to maturity may be another determinant of futures market volatility. Samuelson (1965) was the first to claim that the volatility of futures prices increases as the contract maturity gets closer. Using a very large futures dataset on 6,805 contracts, Daal et al. (2006) found that the maturity effect was much stronger for agricultural and energy commodity futures than it was for financial futures. Using extreme value method, to measure the daily volatility of natural gas futures from daily low and daily high prices, Gregorio and Boucher (2008) found that the maturity effect was significant even after controlling for storage surprises. On the other hand, Mu (2007) tested for the maturity effect by fitting separate GARCH models to the nearby futures contracts and the second nearby contracts and comparing the fitted daily volatilities. In this chapter, I present empirical evidence of the maturity effect by directly including a time to maturity variable in the conditional variance equation of the GARCH model, a substantially different approach from past research.

In this study, GARCH models are used as an econometric tool in order to account for the dynamic nature of short-term market volatility. Some of the more recent studies on natural gas volatility augment the GARCH models with market fundamentals in order to focus on the determinants of volatility (Ates and Wang 2008; Mu 2007; Pyndick 2004; Murry and Zhu 2004). However, the literature does not address the out-of-sample forecasting accuracy of GARCH models augmented with market
4.2 Data and Statistical Analysis

Natural gas futures price data from February 2001 to May 2008 were obtained from NYMEX contracts. Contract-by-contract price data are available from DataStream. The return series for nearby month futures are constructed in two steps. First, returns for individual contracts \( i \) are calculated by

\[ r_{t,i} = \ln(F_{t,i}/F_{t-1,i}) \]  

(4.1)

where \( F_{t,i} \) is the price of the futures contract \( i \) at time \( t \). Then, the nearby month contract return for time \( t \) is obtained as,

\[ r_{t,nb} = r_{t,j} \]  

(4.2)
where \( j \) is the earliest maturing contract. In other words, day \( t \) is in month \( j - 1 \). By first obtaining the returns and then rolling over the contracts, constructing price series from different contracts is avoided, which may distort the data. Therefore, all nearby futures returns \( r_{nb,t} \) are tradable and realizable. At the end of this procedure, the nearby month futures return data are obtained that run from January 4, 2001 to April 23, 2008, a total of \( T = 1,823 \) daily observations.

Summary statistics for nearby month futures returns are presented in Table 4.1. The numbers in parenthesis are the probability values for the associated tests. The returns are right skewed for this sample period and exhibit excess kurtosis. Consequently, the Jarque-Bera test rejects the null hypothesis of normal distribution. These statistics imply that natural gas futures returns are not normally distributed. However, a comparison with the summary statistics of emerging market stock returns given in Table 1.1 reveals that natural gas futures returns are much closer to being normal despite their higher standard deviation.

<table>
<thead>
<tr>
<th>Mean</th>
<th>-0.09538</th>
<th>LBQ(5)</th>
<th>5.31 (0.379)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>-0.09935</td>
<td>LBQ(10)</td>
<td>7.526 (0.675)</td>
</tr>
<tr>
<td>Variance</td>
<td>12.7719</td>
<td>( LBQ^2(5) )</td>
<td>61.81 (5.13e-12)</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.5738</td>
<td>( LBQ^2(10) )</td>
<td>74.195 (6.82e-12)</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.3779</td>
<td>ADF</td>
<td>-12.34 (0.00)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.9046</td>
<td>ARCH-LM(5)</td>
<td>54.44 (0.00)</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>1877.15 (0.00)</td>
<td>ARCH-LM(10)</td>
<td>59.83 (0.00)</td>
</tr>
</tbody>
</table>

Table 4.1: Summary Statistics for Nearby Month Futures Returns

The Ljung-Box test at lags 5 and 10 does not reject the null hypothesis of no autocorrelation in raw returns. However, for squared returns, the null hypothesis of no autocorrelation is rejected strongly. This suggests that there is strong volatility persistence in the data. Therefore, the LM-ARCH test (Engle, 1982) is administered and the null hypothesis of no ARCH effects is strongly rejected. An augmented Dickey
4.2 Data and Statistical Analysis

Fuller test (Dickey and Fuller, 1979) is administered with a constant and 12 lags in the unit root regression without a time trend. The null hypothesis of non-stationarity is rejected, so non-stationarity is not a problem in the analysis.

The gas storage report was being announced by the American Gas Association (AGA) until May 2002 on Wednesdays at 2:00 pm. Since then, the report is released every Thursday at 10:30 am by the Energy Information Administration (EIA). The report provides information on storage levels the Friday before, net weekly changes in storage levels, and the storage levels one-year before. In addition, the five-year historical average for the equivalent time period and the difference between the current level and the five-year average are reported. The deviation from the historical average is the key variable in this study.

The storage data are publicly available from the EIA website. From all the above-mentioned variables, the downloadable data only includes the storage levels. I followed a two-step procedure to construct the deviations from historical levels. First, the weekly data are interpolated to obtain a daily storage level data. This is needed so that storage data for the same day in each of the previous five years are available. In the second step, the storage deviation variable $SD_t$ is constructed as

$$SD_t = S_{t,s} - \frac{1}{5} \sum_{i=1}^{5} S_{t,s-i}$$

where $S_{t,s}$ is the level of storage on day $t$ in year $s$. This two-step procedure is the same methodology followed by the EIA while preparing the storage report.

A time series plot of the working gas in underground storage is presented in Figure 4.1 panel a.

---

6The data are downloadable from http://tonto.eia.doe.gov/dnav/ng/ng_stor_wkly_sl_w.htm
7Simple linear interpolation is used here.
8A complete documentation of the EIA methodology can be found at the link: http://www.eia.doe.gov/oil_gas/natural_gas/ngs/methodology.html.
4.2 Data and Statistical Analysis

Figure 4.1: Working Gas In Underground Storage (a) and Its Deviation From Five Year Historical Mean (b)
4.2 Data and Statistical Analysis

The demand is highly seasonal, and the supply is uniform across seasons in the natural gas market. As such, storage levels also exhibit very strong seasonality because it balances the difference between the supply and demand providing a buffer to the market. In Figure 4.1 panel b, the time series plot of the storage deviation from the five-year historical average $SD_t$ is provided. Note that during the winters of 2001 and 2003 the storage levels were significantly lower than their five-year historical averages, and these periods also coincide with high volatility in natural gas markets.

I obtained the daily realized temperature data running from January 1960 to April 2008 from the trading floor of a very active natural gas trading firm. The dataset includes daily minimum and daily maximum temperatures for seven locations: Atlanta Hartsfield Airport, Chicago Midway Airport, Chicago O’Hare Airport, Dallas Forth Worth Airport, New York Central Park, New York JFK Airport, and New York LaGuardia Airport. The discussion of the weather modeling is left to Section 4.3.2.

4.2.1 A First Look at the Natural Gas Volatility

In this section, I analyze the nearby month futures volatility with respect to natural gas market fundamentals without imposing any econometric structure. The previous literature emphasizes higher volatility on Mondays and storage report announcement days. In some studies, winter was found not to be associated with higher volatility after controlling for other factors. I believe this result is driven by using a broad definition for winter: November to March. This coincides with the period when withdrawal from storage is higher than the injection to storage; hence, known as the withdrawal season. Here, I restrict the definition of winter to include only December, January, and February.
### 4.2 Data and Statistical Analysis

In Table 4.2, the standard deviations of nearby month futures returns are presented for several subgroups based on certain characteristics. There are several important patterns in this table.

<table>
<thead>
<tr>
<th>Panel-A: By Mondays SDDAYs and Winter</th>
<th>Winter</th>
<th>Non-Winter</th>
<th>All Seasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>6.30</td>
<td>4.24</td>
<td>4.79</td>
</tr>
<tr>
<td>SDDAY</td>
<td>4.37</td>
<td>3.66</td>
<td>3.85</td>
</tr>
<tr>
<td>Other Days</td>
<td>3.43</td>
<td>2.79</td>
<td>2.96</td>
</tr>
<tr>
<td>All Days</td>
<td>4.27</td>
<td>3.31</td>
<td>3.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: By Winter and Bidweek</th>
<th>Bidweek</th>
<th>Non-Bidweek</th>
<th>All Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>5.57</td>
<td>4.00</td>
<td>4.27</td>
</tr>
<tr>
<td>Non-Winter</td>
<td>3.34</td>
<td>3.31</td>
<td>3.31</td>
</tr>
<tr>
<td>All Seasons</td>
<td>4.03</td>
<td>3.49</td>
<td>3.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: By SD and Winter</th>
<th>$SD &gt; 0$</th>
<th>$SD &lt; 0$</th>
<th>All Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>3.73</td>
<td>5.66</td>
<td>4.27</td>
</tr>
<tr>
<td>Non-Winter</td>
<td>3.42</td>
<td>2.84</td>
<td>3.31</td>
</tr>
<tr>
<td>All Seasons</td>
<td>3.51</td>
<td>3.80</td>
<td>3.57</td>
</tr>
</tbody>
</table>

Table 4.2: Standard Deviation of Daily Returns Broken into Groups

In panel A, standard deviations are calculated for Mondays, storage report announcement days (SDDAYs), and all other days along with winter and non-winter days. The standard deviation is highest for Monday returns, 4.79. Although lower than Mondays, the standard deviation on storage report announcement days, 3.85, is higher than the standard deviation for all other days, 2.96. The winter effect is also evident. The return volatility for winter days, 4.27, is higher than the return volatility for non-winter days, 3.31. The reason for higher volatility on storage report announcement days is obvious. The storage report is regarded as the most important piece of information by market practitioners, and it is priced as soon as it becomes
available. Using intraday data, Linn and Zhu (2004) showed that the new information is absorbed into the prices within minutes. The volatility of the 10:30-10:35 am interval is much higher than any other five-minute interval during the trading day. However, when Thursdays were excluded from their sample, this effect is completely gone. Fleming et al. (2004) explain the Monday effect with the continuous flow of weather information. They find that the variance ratios of trading to nontrading periods are significantly lower for weather sensitive markets compared to equity markets. They attribute the difference to the continuing flow of weather information over the nontrading period, whereas the information flow for equity markets is reduced in the nontrading period. Additionally, the ratios get even lower over the weekend compared to weekdays, which supports their hypothesis further.

The volatilities for subgroups generated by the Cartesian product of days and seasons also make complete sense. Standard deviations for different days have the same order both in and outside of winter, and winter volatility is consistently higher than non-winter volatility in any subgroup of days. A bar plot of the statistics presented in Table 4.2 panel A is presented in Figure 4.2(a).

Panel B of Table 4.2 analyzes the standard deviations of the returns in the same way, but now the sample is split as the bidweek and non-bidweek days across winter and non-winter days. In the natural gas market, the largest volume of spot trading occurs in the last five business days of every month known as “bidweek.” This is when producers are trying to sell their core production and consumers are trying to buy for their core natural gas needs for the upcoming month (see www.naturalgas.org). The average prices set during bidweek are commonly the prices used in physical contracts over the next month. Since the trading in futures contracts terminates on the third business day before the first business day of the next month and bidweek is the last five business days of the month, the last three business days of trading for the nearby
month contract coincides with the bidweek. In analyzing the maturity effect first proposed by Samuelson (1965), I use bidweek as a natural cutoff point. In panel B of Table 4.2, the standard deviation of the returns on the bidweek days is 4.03, whereas the standard deviation of returns not on the bidweek days is 3.49. The maturity effect is particularly strong during winter. The standard deviation of bidweek days in winter is 5.57, but the standard deviation of non-bidweek days in winter is 4. On the other hand, there is only a marginal difference between the standard deviations outside of winter, 3.34 for bidweek days and 3.31 for non-bidweek days. A bar plot of the statistics presented in Table 4.2 panel B is provided in Figure 4.2(b).

In panel C, SD is the storage deviation variable constructed by (4.3). The short-term volatility studies in the literature found that lower than expected storage levels results in increased volatility because this signals a tight supply situation to the natural gas market (Mu 2007). However, to my knowledge, there are no studies analyzing the effects of the interaction of storage levels with seasonality. In panel
4.2 Data and Statistical Analysis

C, the finding of the previous literature is first confirmed and then challenged. The standard deviation of the periods in which \( SD < 0 \); that is, storage level is lower than five-year historical average, is 3.8, whereas the standard deviation of the periods in which \( SD > 0 \) is 3.51. This confirms the previous literature. However, a more careful examination of the table reveals that this relationship is valid only during the winter months when supply tightness is really a big problem. During the winter, those periods with \( SD < 0 \) have a standard deviation of 5.66, whereas those periods with \( SD > 0 \) have a standard deviation of 3.73. During the non-winter months, the effect is just the opposite: Returns of those periods with \( SD > 0 \) have a standard deviation of 3.42, whereas the returns of those periods with \( SD < 0 \) have a standard deviation of 2.84. This should be because of the concerns regarding the storage capacities. Very high storage levels during non-winter months when demand is minimal increase the concerns about whether there will be enough storage space to store the production for winter demand. This puts a pressure on the price of storage space, which naturally spills over to natural gas prices, causing excessive volatility. Lee Van Atta (2008) cites the excess volatility as one of the most important reasons leading to excessive storage construction over the past few years. This view is consistent with the finding here. A bar plot of the statistics presented in Table 4.2 panel C is presented in Figure 4.2(c).

4.2.2 Levene Tests for Variance Equality

Besides the graphical evidence presented in Figure 4.2, I formally test for the equality of variances among some subsamples of the data. The robust Brown-Forsythe (1974) type Levene (1960) test statistics and associated probability values are presented in Table 4.3. In each row of the table the null hypothesis of equal variances across
the \( J \) groups in the second column is tested. In the first row, the null hypothesis of equal variances for winter and non-winter returns is rejected. In the second row, equality of variances for Mondays, storage report announcement days, and all other days is rejected as well. In the third row, six groups are constructed as the Cartesian product of the winter groups in the first row and days in second row. The equality of variances across these six groups is rejected again. These results are consistent with those presented in Table 4.2 and Figure 4.2. The result in row four—equal variances for bidweek days and non-bidweek days cannot be rejected—is somewhat surprising. In row five, four groups are produced from the Cartesian product of winter groups in row one and bidweek groups in row four. The equality of variances is rejected in this case. However, this may be because unequal variances of winter and non-winter dominating the analysis. To get rid of that effect, I test the equality of variances for those bidweek and non-bidweek days only in winter. The results in row six still cannot reject the equality of variances, although the probability value gets much smaller compared to that in row four. Therefore, there is not strong evidence for unequal variances for bidweek and non-bidweek days. In row seven, the equality of variances for periods with positive storage deviation and negative storage deviation is tested, and the equal variance hypothesis can not be rejected. Constructing four groups based on the Cartesian product of winter and the sign of storage deviation, the equal variance hypothesis is rejected. In the last row, to control for the winter effect, I constructed two groups with positive and negative storage deviations only from winter returns. Now, the null hypothesis of equal variances is rejected. This is consistent with my hypothesis that storage deviation has asymmetric effects during winter and outside of winter.
### Table 4.3: Brown-Forsythe Type Levene Tests for Equality of Variances

<table>
<thead>
<tr>
<th>Groups</th>
<th>J</th>
<th>BFL Test Statistic</th>
<th>Distribution under $H_0$</th>
<th>p.Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) W, NW</td>
<td>2</td>
<td>17.596</td>
<td>$F_{1,1821}$</td>
<td>2.863e-05</td>
</tr>
<tr>
<td>(2) MD, STD, OD</td>
<td>3</td>
<td>39.896</td>
<td>$F_{2,1820}$</td>
<td>0.00</td>
</tr>
<tr>
<td>(3) W&amp;MD, W&amp;STD, W&amp;OD, NW&amp;MD, NW&amp;STD, NW&amp;OD</td>
<td>6</td>
<td>20.656</td>
<td>$F_{5,1817}$</td>
<td>0.00</td>
</tr>
<tr>
<td>(4) BW, NBW</td>
<td>2</td>
<td>0.2512</td>
<td>$F_{1,1821}$</td>
<td>0.6163</td>
</tr>
<tr>
<td>(5) W&amp;BW, W&amp;NBW, NW&amp;BW, NW&amp;NBW</td>
<td>4</td>
<td>6.2663</td>
<td>$F_{3,1819}$</td>
<td>0.00031</td>
</tr>
<tr>
<td>(6) W&amp;BW, W&amp;NBW</td>
<td>2</td>
<td>1.2409</td>
<td>$F_{1,425}$</td>
<td>0.2659</td>
</tr>
<tr>
<td>(7) SD &gt; 0, SD &lt; 0</td>
<td>2</td>
<td>0.8743</td>
<td>$F_{1,1821}$</td>
<td>0.3499</td>
</tr>
<tr>
<td>(8) W&amp;SD &gt; 0, W&amp;SD &lt; 0, NW&amp;SD &gt; 0, NW&amp;SD &lt; 0</td>
<td>4</td>
<td>12.978</td>
<td>$F_{3,1819}$</td>
<td>2.18e-8</td>
</tr>
<tr>
<td>(9) W&amp;SD &gt; 0, W&amp;SD &lt; 0</td>
<td>2</td>
<td>7.4212</td>
<td>$F_{1,425}$</td>
<td>0.0067</td>
</tr>
</tbody>
</table>

Notes:
1. The abbreviations in the group names are as follows: W for winter, NW for non-winter, MD for Monday, STD for storage report announcement day, OD for other days, BW for bidweek, NBW for non-bidweek.
2. The Brown-Forsythe-Levene test statistic is computed as

$$ F = \frac{\frac{\sum_{j=1}^{J} n_j (\bar{D}_j - \bar{D})^2}{\sum_{j=1}^{J} \sqrt{\frac{\sum_{t=1}^{n_j}(D_{tj} - \bar{D}_j)^2}{n_j}}}}{N - J} $$

where $D_{tj} = |r_{tj} - \bar{M}_j|$ and $r_{tj}$ is the return for day $t$ in group $j$; $\bar{M}_j$ is the sample median of the $n_j$ returns in group $j$; $\bar{D}_j = \frac{\sum_{t=1}^{n_j}(D_{tj})}{n_j}$ is the mean absolute deviation from the median $\bar{M}_j$ in group $j$; and $\bar{D} = \frac{\sum_{j=1}^{J} \sum_{t=1}^{n_j}(D_{tj}/N)}{\sum_{j=1}^{J} n_j}$ is the grand mean where $N = \sum_{j=1}^{J} n_j$. The test statistic is distributed as $F_{J-1, N-J}$ under the null hypothesis of equality of variances across the $J$ groups.
3. The original Levene test uses the mean instead of the median. The optimal choice depends on the underlying distribution. However, Brown-Forsythe type test based on the median is recommended since it provides good robustness against non-normal data while retaining good statistical power.
4.3 Empirical Model and Estimation Results

The Ljung-Box test statistics for squared returns in Table 4.1 suggest that there is strong volatility persistence for natural gas nearby month futures returns. Consequently, the LM-ARCH tests confirmed the existence of ARCH effects. In order to take this persistence into account, a GARCH volatility model is adopted as the econometric tool in this section. The focus of this study is completely on the estimation and out-of-sample prediction of daily volatility. Therefore, no structure is specified for the mean equation of the GARCH model. Instead, zero expected return is assumed. Since the day-ahead return is very difficult to forecast, this approach is common for volatility forecasting studies. Consequently, the specification of the empirical model is as follows:

\[ r_t = \sigma_t z_t \]

\[ \sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma X_t \]  \hspace{1cm} (4.4)

where \( r_t \) is the nearby month futures return on day \( t \) given by (4.1) and (4.2), \( \sigma_t \) is the conditional volatility, \( z_t \) is the shocks to the data generating process with \( E[z_t] = 0, \ E[z_t^2] = 1 \). Lastly, \( X_t \) is a vector of exogenous variables capturing the dynamics of the natural gas market volatility. The parameters of the model are obtained by maximizing the following log-likelihood function:

\[ \log L(\omega, \alpha, \beta, \gamma) \propto \sum_{t=1}^{t=n} \left( \log \sigma_t^2(\omega, \alpha, \beta, \gamma) - \frac{r_t^2}{\sigma_t^2(\omega, \alpha, \beta, \gamma)} \right). \]  \hspace{1cm} (4.5)

This likelihood function assumes that the shocks \( z_t \) are normally distributed.
4.3.1 Day-of-the-Week, Seasonality and Maturity Modeling

Inspired by the statistical analysis presented in panel A of Table 4.2, the following variables are included in the model.

\(SDDAY_t\): A dummy variable for the storage report announcement days

\(MON_t\): A dummy variable for Mondays

\(WIN_t\): A dummy variable for winter days, with the winter defined as December, January, and February

Additionally, panel B of Table 4.2 presents preliminary evidence regarding the maturity effect on futures volatility. However, the Brown-Forsythe type Levene test does not confirm the unequality of the variances for the bidweek and non-bidweek days. Therefore, using a dummy variable for bidweek is not justified. Instead, I construct more general variables to capture the maturity effect:

\(TTM\): The number of business days until the maturity of nearby month futures contract

\(TTMWIN\): Time to maturity variable on winter days. The variable is constructed as: \(TTMWIN_t = TTM_t \times WIN_t\).

The latter variable is constructed because it was found that the maturity effect is particularly strong during the winter months. The results of the estimations including these first set of exogenous variables are presented in Table 4.4. The first estimation is for a simple \(GARCH(1,1)\) model. Starting with the second estimation, one more variable is included in the model at each time. The estimation results are consistent with the previous data analysis. Volatility is significantly higher on storage report announcement days, Mondays, and winter days. In estimation-5, the time to maturity (TTM) variable is significant and has the correct negative sign. So, the volatility of
futures returns increases as the maturity gets closer. However, it lost its significance in estimation-6 after TTMWIN is also included in the estimation. This is consistent with the previous idea that the maturity effect is present only during winter months. More formally, the coefficient of time to maturity variable during non-winter months is $\gamma_4$. On the other hand, during the winter months, it is $\gamma_4 + \gamma_5$. Therefore, testing for the conditional hypothesis that the maturity effect is present only during the winter requires a statistical test of the null hypothesis $H_0: (\gamma_4 = 0 \text{ and } \gamma_4 + \gamma_5 < 0)$. The $t$ statistic for $\gamma_4$ is -0.15. Also, the $t$ statistic for $\gamma_4 + \gamma_5$ is calculated using the variance-covariance matrix of estimated parameters and reported in Table 4.4 as -8.416. This confirms the null hypothesis that the maturity effect is present only in winter. In the final estimation, the non-significant TTM variable is dropped from the estimation. All remaining variables are significant at all conventional levels.

A high level of persistence in natural gas futures volatility, as measured by the sum $\alpha + \beta$, is evident in these estimations. The volatility literature suggests that the persistence in volatility might be the result of driving exogenous variables that are persistent themselves. Therefore, such variables should reduce the level of volatility persistence once they are included in the conditional variance equation of GARCH models. This kind of behavior is observed in the parameter estimates with exogenous variables in the volatility equation. While $\alpha + \beta = 0.985$ in the simple GARCH estimation, it reduced to 0.938 in the final estimation. The half-life of a volatility shock is defined as the time it takes for half of the shock to vanish and is given by:

$$\text{Half-Life} = \frac{\log(0.5)}{\log(\alpha + \beta)} \quad (4.6)$$

The half-life estimates are also presented in Table 4.4. The half-life decreases to 10.83 days in the augmented model from the 45.86 days in the simple GARCH model.
### Table 4.4: GARCH Estimation Results For Nearby Month Futures Returns

<table>
<thead>
<tr>
<th>Model</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\gamma_3$</th>
<th>$\gamma_4$</th>
<th>$\gamma_5$</th>
<th>H-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>0.255</td>
<td>0.077</td>
<td>0.908</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.53***)</td>
<td>(8.62***)</td>
<td>(75.94***)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45.86</td>
</tr>
<tr>
<td>(2)</td>
<td>0.500</td>
<td>0.073</td>
<td>0.902</td>
<td>4.211</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.30**)</td>
<td>(7.02***)</td>
<td>(62.99***)</td>
<td>(3.99***)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.38</td>
</tr>
<tr>
<td>(3)</td>
<td>-0.997</td>
<td>0.080</td>
<td>0.873</td>
<td>4.32</td>
<td>3.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.997)</td>
<td>(9.39***)</td>
<td>(66.08***)</td>
<td>(4.56***)</td>
<td>(6.32***)</td>
<td></td>
<td></td>
<td></td>
<td>14.4</td>
</tr>
<tr>
<td>(4)</td>
<td>-1.02</td>
<td>0.078</td>
<td>0.859</td>
<td>4.09</td>
<td>4.47</td>
<td>0.341</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-5.53***)</td>
<td>(6.00***)</td>
<td>(44.65***)</td>
<td>(4.73***)</td>
<td>(7.73***)</td>
<td>(4.35***)</td>
<td></td>
<td></td>
<td>10.65</td>
</tr>
<tr>
<td>(5)</td>
<td>-0.503</td>
<td>0.070</td>
<td>0.874</td>
<td>4.31</td>
<td>4.01</td>
<td>0.162</td>
<td>-0.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.66*)</td>
<td>(6.61***)</td>
<td>(44.09***)</td>
<td>(4.96***)</td>
<td>(6.81***)</td>
<td>(1.645*)</td>
<td>(-2.69***)</td>
<td></td>
<td>12.02</td>
</tr>
<tr>
<td>(6)</td>
<td>-1.06</td>
<td>0.073</td>
<td>0.863</td>
<td>4.69</td>
<td>4.44</td>
<td>2.40</td>
<td>-0.0029</td>
<td>-0.217</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.11***)</td>
<td>(5.68***)</td>
<td>(38.13***)</td>
<td>(5.38***)</td>
<td>(6.98***)</td>
<td>(7.18***)</td>
<td>(-0.15)</td>
<td>(-6.73***)</td>
<td>10.48</td>
</tr>
<tr>
<td>(7)</td>
<td>-1.10</td>
<td>0.073</td>
<td>0.865</td>
<td>4.67</td>
<td>4.41</td>
<td>2.38</td>
<td>-0.215</td>
<td>-0.215</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-5.50***)</td>
<td>(5.68***)</td>
<td>(38.84***)</td>
<td>(4.35***)</td>
<td>(7.30***)</td>
<td>(9.38***)</td>
<td>(-8.29***)</td>
<td>(-8.29***)</td>
<td>10.83</td>
</tr>
</tbody>
</table>

\[ t.stat(\hat{\gamma}_4 + \hat{\gamma}_5) = -8.416 \text{ (In Estimation-6)} \]

Significance Codes: * 10%, ** 5%, *** 1%

\[
\begin{align*}
  r_t & = \sigma_t z_t \\
  \sigma_t^2 & = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma_1 SDDAY_t + \gamma_2 MON_t + \gamma_3 WIN_t + \gamma_4 TTM_t + \gamma_5 TTMWIN_t
\end{align*}
\]

Table 4.4: GARCH Estimation Results For Nearby Month Futures Returns
4.3 Empirical Model and Estimation Results

4.3.2 Storage and Weather Modeling

In order to incorporate the asymmetric effects of storage levels during different seasons, two additional variables are constructed: $SD_t$ and $SDWIN_t$. The first variable $SD_t$ is the same variable constructed in Section 4.2. It is the deviation of storage level from its five-year historical average. Once $SD_t$ is calculated for a storage report announcement day, the same number is used on the following days until the next storage report announcement. This is different from the previous literature that included the storage surprise variable only for the announcement days (Mu 2007; Gregoire and Boucher 2008). This is because I regard this variable not only as a proxy for storage surprise but as a proxy for supply tightness during winter and as a proxy for tightness in storage space supply in other seasons. The second variable $SDWIN_t$ is a proxy for supply tightness during winter. It is constructed as $SDWIN_t = SD_t \times WIN_t$. This variable enables us to model the asymmetric effect of storage levels for different seasons. The expectation is a positive coefficient for $SD_t$ and a higher negative coefficient for $SDWIN_t$ to confirm the hypothesis that low storage levels increase the volatility in winter, whereas high storage levels increase the volatility at other times.

The weather modeling is accomplished by the well-known degree day variables. These are quantitative indices used to reflect the demand for energy. Experience shows that there is no need for heating or cooling if the outside temperature is $65^\circ F$. Consequently Heating Degree Days and Cooling Degree Days variables are defined as:

$$HDD_t = \text{Max}(0, 65 - T_{ave,t})$$
$$CDD_t = \text{Max}(0, T_{ave,t} - 65),$$

(4.7)
where $T_{ave,t}$ is the average of the maximum and minimum observed temperature on day $t$. There are two common ways of modeling weather shocks, either with ex-post forecast errors or with temperature anomalies, defined as the deviation of degree days variables from their seasonal norms (Mu 2007). Here, I follow the second approach because the forecast data are not available. The following weather shock variables are constructed: $HDD.Shock_t$: This is defined as the deviation of Heating Degree Days from the seasonal norms over the forecasting horizon. Following Mu (2007), the forecasting horizon is chosen to be seven days since the weather forecasts from the public media are typically broadcast for seven days ahead.\(^9\)

\[
HDD.Shock_t = \sum_{i=t+1}^{i=t+7} (HDD_i - HDD.Norm_t),
\]  

(4.8)

where $HDD.Norm_t$ is the historical 30-year average of HDD on day $t$. The historical 30-year average is the definition of the National Weather Service (NWS) for the seasonal norm. Since I do not have the actual forecast data, the realized HDD is used in creating this variable.

$CDD.Shock_t$: This is defined as the deviation of CDD from the seasonal norms and calculated in the same way as:

\[
CDD.Shock_t = \sum_{i=t+1}^{i=t+7} (CDD_i - CDD.Norm_t),
\]  

(4.9)

where $CDD.Norm_t$ is the 30-year historical average of CDD on day $t$.

Both weather variables are constructed for Chicago, New York, Atlanta, and Dallas. Then, the natural gas consumption weighted average of these locations is calculated.\(^10\) The national weather shock variables are plotted in Figure 4.3.

\(^9\)Results are robust to the choice of a forecasting horizon as 8, 9, or 10 days.

\(^10\)I used the same weights used in Mu (2007): 0.42 for Chicago, 0.28 for New York, 0.17 for
4.3 Empirical Model and Estimation Results

Figure 4.3: National Weather Shock Variables

Atlanta, and 0.13 for Dallas.
4.3 Empirical Model and Estimation Results

Heating degree day shocks are closer to zero in summer months, whereas cooling degree day shocks are closer to zero in winter months since both the thirty-year historical averages and the actual realizations get closer to zero.

Estimation results including the storage and weather variables are presented in Table 4.5. In estimation-8, only the two storage variables are added to the last estimation in Table 4.4. Both variables have the correct sign and are significant at all conventional levels. The positive sign for the coefficient of $SD_t$ suggests that high storage levels increase the short-term volatility during the non-winter period. Also, the negative coefficient for $SDWIN_t$ is greater than the positive coefficient of the $SD_t$, which suggests that it is the low storage levels resulting in high volatility during winter months. This asymmetric effect is tested more formally later in estimation-11. In estimation-9, only the two weather variables are added to the last estimation in Table 4.4. They are both significant at the 10 percent confidence level with a correct positive sign. Higher degree days than the seasonal norms increase short-term volatility. One unexpected result is that the significance of $CDD.Shock$ is stronger than the significance of $HDD.Shock$. This might be due to the other variables accounting for higher volatility in winter. If the winter dummy variable and its interaction term with the maturity were excluded from the model as in estimation-10, the significance of $HDD.Shock$ becomes stronger than the significance of $CDD.Shock$. In this estimation, HDD is significant at the 1 percent level and CDD is significant at the 5 percent level. Lastly, in estimation-11, I include the storage and weather variables together. The inference for the storage variables remains the same. The asymmetric effect of storage deviation variable $SD_t$ across seasons can be formally tested with the null hypothesis of $H_0 : (\gamma_5 > 0 \ and \ \gamma_5 + \gamma_6 < 0)$. The $t$ statistic for $\gamma_5$ is 3.34. Also, the $t$ statistic for $\gamma_5 + \gamma_6$ is calculated using the variance-covariance matrix of estimated parameters and reported in Table 4.5 as -3.094. This confirms the asymmetric effect of
<table>
<thead>
<tr>
<th>Model</th>
<th>(\omega)</th>
<th>(\alpha)</th>
<th>(\beta)</th>
<th>(\gamma_1)</th>
<th>(\gamma_2)</th>
<th>(\gamma_3)</th>
<th>(\gamma_4)</th>
<th>(\gamma_5)</th>
<th>(\gamma_6)</th>
<th>(\gamma_7)</th>
<th>(\gamma_8)</th>
<th>H-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8)</td>
<td>-1.29</td>
<td>0.07</td>
<td>0.839</td>
<td>5.58</td>
<td>5.43</td>
<td>3.54</td>
<td>-0.28</td>
<td>0.00062</td>
<td>-0.0017</td>
<td>0.00062</td>
<td>-0.0017</td>
<td>7.24</td>
</tr>
<tr>
<td></td>
<td>(-5.46)</td>
<td>(5.09)</td>
<td>(31.85)</td>
<td>(6.55)</td>
<td>(7.94)</td>
<td>(9.24)</td>
<td>(-7.49)</td>
<td>(3.72)</td>
<td>(-5.00)</td>
<td>(3.72)</td>
<td>(-5.00)</td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>-1.05</td>
<td>0.076</td>
<td>0.850</td>
<td>5.03</td>
<td>4.36</td>
<td>2.74</td>
<td>-0.24</td>
<td>0.027</td>
<td>0.062</td>
<td></td>
<td></td>
<td>9.01</td>
</tr>
<tr>
<td></td>
<td>(-5.47)</td>
<td>(5.67)</td>
<td>(35.08)</td>
<td>(5.87)</td>
<td>(6.29)</td>
<td>(7.54)</td>
<td>(-6.67)</td>
<td>(1.69)</td>
<td>(1.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10)</td>
<td>-1.05</td>
<td>0.073</td>
<td>0.888</td>
<td>4.64</td>
<td>3.04</td>
<td>-0.29</td>
<td>0.00064</td>
<td>0.027</td>
<td>0.055</td>
<td></td>
<td></td>
<td>9.01</td>
</tr>
<tr>
<td></td>
<td>(-5.62)</td>
<td>(6.47)</td>
<td>(49.87)</td>
<td>(5.14)</td>
<td>(4.89)</td>
<td>(-6.98)</td>
<td>(3.34)</td>
<td>(3.75)</td>
<td>(2.04)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11)</td>
<td>-1.28</td>
<td>0.068</td>
<td>0.833</td>
<td>5.73</td>
<td>5.54</td>
<td>3.77</td>
<td>-0.29</td>
<td>0.00064</td>
<td>-0.0017</td>
<td>0.017</td>
<td>0.015</td>
<td>6.65</td>
</tr>
<tr>
<td></td>
<td>(-5.13)</td>
<td>(4.98)</td>
<td>(30.34)</td>
<td>(6.64)</td>
<td>(7.35)</td>
<td>(8.66)</td>
<td>(-6.98)</td>
<td>(3.34)</td>
<td>(-4.19)</td>
<td>(0.842)</td>
<td>(0.407)</td>
<td></td>
</tr>
</tbody>
</table>

\(t.stat(\gamma_5 + \gamma_6) = -3.094\) (In estimation-11)

Significance Codes: * 10%, ** 5%, *** 1%

\[ r_t = \sigma_t z_t \]
\[ \sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma_1 SDDAY_t + \gamma_2 MON_t + \gamma_3 WIN_t + \gamma_4 TTMWIN_t + \gamma_5 SD_t + \gamma_6 SDWIN_t + \gamma_7 HDD.Shock_t + \gamma_8 CDD.Shock_t \]

Table 4.5: GARCH Estimation Results with Storage and Weather Variables For Nearby Month Futures Returns
the storage variable across seasons. As for weather variables, they still have the correct positive signs, but they lost their significance after the addition of the storage variables. Note that the storage variables were included in the model as a proxy for supply, and the weather shocks were included as a proxy for demand in the natural gas market. However, storage levels could be thought of as the result of the combination of supply and demand forces, thereby providing an explanation for the reduction in the significance of weather variables. The winter dummy is another factor that reduces the significance of weather variables due to nonorthogonality as discussed before.

4.4 Out-of-Sample Forecast Accuracy

Recent papers employing GARCH models to investigate the effects of natural gas market fundamentals on the volatility dynamics of natural gas futures report results only for in-sample estimations. These estimation results provide valuable information for understanding the volatility dynamics of natural gas futures. However, the out-of-sample predictive power of these augmented GARCH models has not been tested. In this section, day-ahead volatility predictions are made for natural gas nearby month futures using simple GARCH models, as well as their augmented counterparts, and the accuracy of these forecasts are compared.

The empirical methodology followed here is known as a sliding window scheme. To make a prediction for day $t$ where $t \in \{501, 502, ..., T\}$, only the returns $\{r_{t-1}, r_{t-2}, ..., r_{t-500}\}$ are used. So, the length of the sliding window is chosen as 500 observations. That is, returns 1 through 500 are used to predict the volatility for day 501; returns 2 through 501 are used to predict the volatility for day 502, and so on. Since $T = 1,823$, there are $T - 500 = 1,323$ volatility predictions.
4.4 Out-of-Sample Forecast Accuracy

One problem in out-of-sample forecasting is that the variables $HDD.Shock_t$ and $CDD.Shock_t$ are using information from the future. On day $t$, the volatility for day $t + 1$ is being forecasted, but at that time these variables are not available yet in the information set. To solve this problem, first weather shock forecasts are obtained by fitting an ARIMA(1,2,1) model to the last 500 calendar days of temperature data for all four cities. The ARIMA(1,2,1) is chosen based on the Schwarz information criterion (SIC). The natural gas consumption weighted average of weather shock variables across the four cities is calculated as the final weather shock forecasts. Then, the weather shock forecasts are used in augmented GARCH models to forecast the volatility. A simpler approach is to use the appropriate lags of weather shock variables. This can be regarded as forecasting the next seven days’ weather shocks as being equal to the last seven days’ weather shocks. The presented results are from the ARIMA forecasting approach, but using a simpler lag approach provides the same results.

4.4.1 Other Simple Models for Forecasting

Random Walk Model:
With a random walk assumption, the volatility forecast for day $t$ is the realized volatility on day $t - 1$. It is used as the benchmark model.

$$\hat{\sigma}_t = |\sigma_{t-1}|.$$  \hspace{1cm} (4.10)

Moving Average Model:
Moving average models are widely used by natural gas market practitioners. In this paper, 20-day and 60-day moving averages are used that correspond to one-month
4.4 Out-of-Sample Forecast Accuracy

and three-month trading days.

\[ \hat{\sigma}_t = \sqrt{\frac{1}{m} \sum_{i=1}^{m} r_{t-1}^2} \]  \hspace{1cm} (4.11)

Implied Volatility:

Annualized implied volatilities for the closest-to-the-money call options are obtained from the trading floor of a very active natural gas trading firm. Dividing annualized implied volatility by \(\sqrt{250}\), the daily implied volatilities are obtained. Forecasting can be done as follows:

\[ \hat{\sigma}_t = \sigma_{t-1}^{imp} / \sqrt{250} \]  \hspace{1cm} (4.12)

4.4.2 Forecast Accuracy Results

After making the forecast and observing the realized volatility the next day, the volatility forecast error can be defined as

\[ FE = \sigma_t - \hat{\sigma}_t \]

\[ FE = |r_t| - \hat{\sigma}_t \]

where \(\hat{\sigma}_t\) is the forecasted volatility, and \(\sigma_t\) is the realized volatility for day \(t\). The latter equality follows because, in the absence of intraday data, the most common approach in the literature is to use the absolute value of return as the realized volatility. Three statistical measures are used for measuring forecast accuracy. These are mean
absolute error (MAE), mean squared error (MSE), and Theil’s U statistic.

\[
MAE = \frac{1}{n} \sum_{t=1}^{n} |r_t - \hat{\sigma}_t|, \quad (4.13)
\]

\[
MSE = \frac{1}{n} \sum_{t=1}^{n} (|r_t| - \hat{\sigma}_t)^2, \quad (4.14)
\]

\[
\text{Theil's } U = \frac{\sqrt{\sum_{t=1}^{n} (|r_t| - \hat{\sigma}_t)}}{\sqrt{\sum_{t=1}^{n} (|r_t| - |r_{t-1}|)}}, \quad (4.15)
\]

Theil’s U statistic can be thought as a relative accuracy measure. It is the ratio of the root mean squared error of the chosen model to the root mean squared error of the random walk model. Out-of-sample forecasting accuracy measures for the simple and augmented GARCH models as well as other simple forecasting methods are presented in Table 4.6.

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>Rank (MAE)</th>
<th>MSE</th>
<th>Rank (MSE)</th>
<th>Theil’s U</th>
<th>Rank (Theil’s U)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garch(1,1)</td>
<td>1.984</td>
<td>6</td>
<td>6.284</td>
<td>5</td>
<td>0.7908</td>
<td>5</td>
</tr>
<tr>
<td>Model (7)</td>
<td>2.005</td>
<td>8</td>
<td>6.366</td>
<td>6</td>
<td>0.7959</td>
<td>6</td>
</tr>
<tr>
<td>Model (8)</td>
<td>1.931</td>
<td>3</td>
<td>6.016</td>
<td>2</td>
<td>0.7737</td>
<td>2</td>
</tr>
<tr>
<td>Model (9)</td>
<td>1.970</td>
<td>5</td>
<td>6.271</td>
<td>4</td>
<td>0.7899</td>
<td>4</td>
</tr>
<tr>
<td>Model (11)</td>
<td>1.8913</td>
<td>2</td>
<td>5.974</td>
<td>1</td>
<td>0.7710</td>
<td>1</td>
</tr>
<tr>
<td>MA(20) Model</td>
<td>1.8730</td>
<td>1</td>
<td>6.195</td>
<td>3</td>
<td>0.7852</td>
<td>3</td>
</tr>
<tr>
<td>MA(60) Model</td>
<td>1.963</td>
<td>4</td>
<td>6.459</td>
<td>8</td>
<td>0.8017</td>
<td>8</td>
</tr>
<tr>
<td>Implied Vol</td>
<td>2.004</td>
<td>7</td>
<td>6.411</td>
<td>7</td>
<td>0.7987</td>
<td>7</td>
</tr>
<tr>
<td>Random Walk</td>
<td>2.239</td>
<td>9</td>
<td>10.049</td>
<td>9</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 4.6: Accuracy Statistics for Out-of-Sample Forecasting

Model 7 is based on the estimation-7 in Table 4.4. The forecasting accuracy of this model is slightly worse than the simple GARCH model giving marginally higher MAE, MSE, and Theil’s U statistics. Model 8 and 9 are based on estimations 8
and 9 in Table 4.5. Both models perform better than the simple GARCH model. Moreover, storage variables in model 8 improve the performance much more than the weather variables in model 9. Lastly, model 11 includes all variables and among all the GARCH models, it provides the lowest MAE, MSE, and Theil’s U statistics, making a great improvement on the simple GARCH model.

So, the evidence is not conclusive that any augmentation of the GARCH models would increase the accuracy in out-of-sample forecasting as in the case of model 7. However, with carefully chosen variables to account for natural gas market fundamentals, it is possible to increase the forecast accuracy as in models 8, 9, and 11.

The simple forecasting schemes generally perform very poorly with the exception of the 20-day moving average. MA(20) method ranks first with a very small margin in terms of MAE and third in terms of MSE and Theil’s U. The MSE measure penalizes models with the square of their errors, and the MA(20) does not perform as good as with the linear penalty function used in MAE calculation. Therefore, this simple model is producing very good forecasts in general, but once it is wrong it is way of the target. On the other hand, model 11 that includes all fundamental variables within a GARCH framework ranks second in terms of MAE and first in terms of MSE and Theil’s U, thereby consistently providing good forecast accuracy.

4.5 Out-of-Sample Risk Estimation

Applications of augmented GARCH models are not limited to volatility forecasting. In fact, the volatility forecasts obtained from these models can be further used to estimate the extreme tail risk of a natural gas position. In this section, VaR forecasting and backtesting methodology, which was applied to emerging market stock indices in Chapter 1, is repeated for the Nymex natural gas futures. However, here I focus on
the effects of model augmentation with market fundamentals and applying EVT on the VaR prediction performance. Therefore, the models from Table 4.6 are employed with and without EVT, in this section but other models from the first chapter are omitted.

After the estimation of the simple and augmented GARCH models with MLE methods, VaR is estimated in two separate ways. One way is to assume that the residuals follow a normal distribution and calculate VaR as

\[ \text{VaR}_\alpha = \hat{\sigma}_t \Phi^{-1}(\alpha), \]

where \( \hat{\sigma}_t \) is the volatility forecast from the model and \( \Phi \) is the normal distribution function. The other methodology is relying on the EVT-GARCH model of McNeil and Frey (2000), which was discussed in detail in Chapter 1. In this case, the GPD distribution is fit to the residuals extracted from the GARCH model in the first step. Then, VaR is calculated as

\[ \text{VaR}_\alpha = \hat{\sigma}_t \left( u + \frac{\hat{\beta}_z}{\hat{\xi}_z} \left( \frac{1 - \alpha}{1 - F(u)} \right) \hat{\xi}_z - 1 \right), \]

where \( z \) are the residuals, \( u \) is the threshold set at 0.95th quantile of the residuals, \( \hat{\xi} \) is the shape, and \( \hat{\beta} \) is the scale parameter estimates from the GPD fit.

By definition, \( \text{VaR}_\alpha \) can be exceeded by the loss \( 1 - \alpha \) of the time. With a total sample size of 1,823 and a rolling sample size of 500 business days, 1,323 VaR predictions are obtained. Therefore, the expected (target) number of violations of the VaR for 0.95th, 0.975th, 0.99th, and 0.995th quantiles are 66, 33, 13, and 7, respectively. These expected numbers of violations and the actual number of violations from each of the models are reported in Table 4.7. Since the VaR violation process can be
thought of as a Bernoulli process, the number of total violations needs to be binomially distributed. The probability values associated with the one-sided likelihood ratio test of the null hypothesis of binomial distribution for the number of violations are also reported on the same table in parenthesis.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\alpha = 0.95$</th>
<th>$\alpha = 0.975$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 0.$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>66</td>
<td>33</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Garch(1,1)</td>
<td>50 (0.03)</td>
<td>29 (0.46)</td>
<td>13 (0.95)</td>
<td>9 (0.38)</td>
</tr>
<tr>
<td>EVT-Garch</td>
<td>38 (0)</td>
<td>23 (0.06)</td>
<td>6 (0.03)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>Model-7</td>
<td>54 (0.11)</td>
<td>28 (0.36)</td>
<td>15 (0.63)</td>
<td>7 (0.88)</td>
</tr>
<tr>
<td>EVT-Model7</td>
<td>41 (0)</td>
<td>19 (0.01)</td>
<td>6 (0.03)</td>
<td>3 (0.11)</td>
</tr>
<tr>
<td>Model8</td>
<td>64 (0.79)</td>
<td>34 (0.87)</td>
<td>15 (0.63)</td>
<td>7 (0.88)</td>
</tr>
<tr>
<td>EVT-Model8</td>
<td>49 (0.02)</td>
<td>25 (0.14)</td>
<td>7 (0.06)</td>
<td>6 (0.81)</td>
</tr>
<tr>
<td>Model9</td>
<td>59 (0.36)</td>
<td>29 (0.46)</td>
<td>12 (0.73)</td>
<td>9 (0.38)</td>
</tr>
<tr>
<td>EVT-Model9</td>
<td>52 (0.06)</td>
<td>18 (0)</td>
<td>8 (0.12)</td>
<td>6 (0.81)</td>
</tr>
<tr>
<td>Model11</td>
<td>65 (0.88)</td>
<td>40 (0.24)</td>
<td>21 (0.05)</td>
<td>13 (0.03)</td>
</tr>
<tr>
<td>EVT-Model11</td>
<td>55 (0.15)</td>
<td>30 (0.58)</td>
<td>13 (0.95)</td>
<td>7 (0.88)</td>
</tr>
</tbody>
</table>

Table 4.7: VaR Violations and Probability Values for Binomial Hypothesis

In terms of the usefulness of EVT, the results for natural gas futures are somewhat different from those of emerging market stock indices reported in Chapter 1. The EVT methodology for fat-tail modeling does not improve the backtesting performance of the Gaussian GARCH model. More surprisingly, there is no room for improvement because the number of VaR violations from Gaussian GARCH model is very close to the expected number of violations. This is because the tail fatness of natural gas futures is not as pronounced as in the case of emerging market stocks, although its returns are more volatile. Higher volatility, larger range of data, or a bigger domain for a distribution does not automatically imply fat tails. In EVT, fat tails are defined as having a power decay rate in the tails of distribution, and extremes are defined in terms of quantiles. A return distribution with tails wide open from -15 to 15 percent may have an exponential decay rate in its tails and closely follow a normal
distribution. Natural gas return distribution is like this, and 8 percent loss in a day is not an extreme event for this market. Conversely, a narrow return distribution extending from -6 to 6 percent may have a power decay rate in its tails, and its shape may be very deviated from that of a normal distribution. As an example, Taiwan's stock market return distribution is like this, and 8 percent loss in one day is an extreme event for this market. This point can be visually investigated in Figure 4.4.

Figure 4.4: Density Fits For Natural Gas and Taiwan Stock Indices

In panel A, the Kernel and normal density plots for natural gas futures returns is presented. Same graphs are shown for the Taiwan stock index returns in panel B. Taiwan is chosen because in Chapter 1, it was shown that Taiwan is one of the
4.5 Out-of-Sample Risk Estimation

countries that exhibit stock return distributions closer to being normal compared to other emerging stock markets. Figure 4.4 suggests that natural gas futures return distribution is even closer to being normal, although it exhibits higher volatility with wide open tails. Therefore, natural gas is characterized with very high price volatility, but it does not really exhibit fat tails.

In section 4.3, it is shown that the persistence in volatility is driven by the persistent market fundamental variables. Once these variables are added to the model, the volatility persistence measured by the half-life significantly decreases. The motivation for this section originally was whether the fundamental market variables could be the drivers of the tail fatness as well. However, as just explained, the tail fatness present in natural gas futures returns is so moderate that it can be accounted for by the Gaussian GARCH model. As a result, a comparison of the simple GARCH model with the augmented GARCH models provides similar conclusions. The Gaussian GARCH model is performing very close to the targets, so there is not much room for improvement. Therefore, the augmented GARCH models do not improve the performance of the simple GARCH model. However, they do not decrease the performance either. Models 7, 8 and 9 all exhibit similar performance to the simple GARCH model. The null hypothesis of binomial distribution is accepted at all four quantiles for these models with similar and sometimes even higher probability values. However, model 11 results in significant underestimation of risk in higher quantiles. This underestimation can be fixed by EVT modeling, and model 11 with EVT performs even better than the simple GARCH model. In general, there is no need for augmentation of the GARCH model for risk estimation because there is no additional improvement. This finding suggests that a simple GARCH model can be used for risk measurement purposes of linear pay-off instruments such as natural gas futures. On the other hand, the findings of previous section suggests that augmented
4.6 Conclusion

GARCH models provides better performance in volatility forecasting and therefore can be used in pricing non-linear pay-off instruments such as options.

4.6 Conclusion

Recently, a new literature emerged on modeling short-term volatility dynamics of natural gas futures. This research focuses on the augmentation of GARCH models and its variants with the natural gas market fundamentals in order to understand the sources of high volatility in natural gas prices. In this paper, several new findings contributing to this literature have been presented, and more importantly forecasting the accuracy of these models is analyzed for the first time.

First of all, the effect of storage levels on short-term volatility is asymmetric across the seasons. During the winter months, lower storage levels than the five-year historical average were found to be increasing the short-term volatility. In contrast, it is the high levels of storage causing excess volatility in other seasons. This can be attributed to the changing concerns of market players at different seasons. In the winter, low storage levels are perceived as a tight supply situation causing excess volatility. At other times, the market is mainly concerned about the storage space supply. Therefore, high levels of storage cause excess volatility.

Secondly, the maturity effect for natural gas nearby month futures is found to be a significant determinant of volatility only in the winter months. This result is confirmed by both data analysis and econometric estimation of the GARCH models, including the maturity variable and its interaction with the seasonality in the volatility equation. Since winter is the season when demand is highly inelastic, traders might be overreacting to new information arrival closer to the maturity and this causes excess volatility.
4.6 Conclusion

In addition to these new findings, this study confirms some of the previous results in the literature. Higher volatility is observed on Mondays possibly due to the accumulation of weather information over the non-trading weekend. Storage report announcement days also exhibit higher volatility than other days since new arriving information is priced very fast in this case. Volatility on winter days, defined as December, January, and February, is found to be higher than other seasons. Lastly, weather shocks in excess of the seasonal norms increase the short-term volatility. However, when storage variables and the winter dummy is included in the model, they take the significance of weather variables. This might be because the storage variables are taking care of both supply and demand dynamics and/or non-ortgonality with the winter dummy.

As for forecasting accuracy, the augmented GARCH models with carefully chosen fundamental variables have the potential to decrease MAE, MSE, and Theil’s U statistics. The model using storage and weather variables in addition to other variables capturing the Monday, storage report announcement, winter, and maturity effects provides the best forecasting accuracy, thereby greatly improving on the simple GARCH model forecasts. This is a very important finding because better volatility forecasts can be used in option pricing and hedging natural gas exposures. Fleming et al. (2001) suggests focusing on the economic significance of time varying predictable volatility instead of evaluating the statistical performance of volatility models. Future research can be conducted in such applications of augmented GARCH models for natural gas volatility.

Lastly, we found that a simple GARCH model provides very good backtesting result in risk estimation, and there is no room for improvement by augmenting the model with market fundamentals. This is because natural gas futures return distribution does not exhibit very fat tails. Overall, the results suggest that the volatility
4.6 Conclusion

forecasting performance can be increased by augmentation of the GARCH models, whereas a simple GARCH model can be preferable for the risk measurement of linear portfolios considering the simplicity advantage.
Bibliography


