ISSN 2223-0025

THE LAHORE JOURNAL OF BUSINESS

Lahore School of Economics

Hervé Alexandre and Antonin de Benoist
Oil Prices and Government Bond Risk Premiums

Bushara Bano and Rajiv Kumar Jha
Organizational Role Stress Among Public and Private Sector Employees: A Comparative Study

Mahreen Mahmud
The Forecasting Ability of GARCH Models for the 2003–07 Crisis: Evidence from S&P500 Index Volatility

Aliya Bushra
Job Satisfaction and Women's Turnover Intentions in Pakistan's Public Universities

Tareena Musaddiq
Modeling and Forecasting the Volatility of Oil Futures Using the ARCH Family Models

Volume 01, No.1 May-Sept, 2012

THE LAHORE JOURNAL OF BUSINESS

Editorial Staff: Tele. No: 0092 - 42 - 36561230
Telefax: 0092 - 42 - 36561230
E-mail: ljb@lahoreschool.edu.pk
Publisher: Lahore School of Economics, Lahore, Pakistan.

Correspondence relating to subscriptions and changes of address should be sent to *The Lahore Journal of Business*, Intersection Main Boulevard Phase VI DHA and Burki Road, Burki Lahore - 53200 – Pakistan.

Instructions to authors can be found at the end of this issue. No responsibility for the views expressed by authors and reviewers in *The Lahore Journal of Business* assumed by the Editors and the Publisher.

01 2012

THE LAHORE JOURNAL OF BUSINESS

Contents Vol. 01, No.1, 2012

Notes for Authors

1. The first page of the manuscript should have the title of the paper, the names(s) of author(s), and a footnote giving the current affiliation of the author(s) and any acknowledgments.

2. The following page should include an abstract (maximum 400 words). The abstract should be followed by Journal of Economic Literature (JEL) classifications and keywords (minimum of two and maximum of five). Please note that JEL classifications are also available for marketing and management related topics. For more on JEL classifications please consult http://www.aeaweb.org/journal/jel_class_system.php.

3. The main text of the article should be single column format, double line spaced with font style of Times New Roman. All main headings should be bold with size 16, while subheadings should be bold with font size 14. The main text should be justified with size 12.

4. The numerals used for headings and subheadings should be 1, 1.1 and 1.1.1. It is advised to limit the level of headings within article.

5. Write one to nine in words, 10 and above should be in figures. Write numbers in figures (not words) for exact measurements, quantities and percentages. Use thousands, millions, billions and not crores and lakhs. In text use 'per cent'; in tables and figures '%'. In the case of decimal, use "0.8" rather than ".8". Maintain consistency in the number of decimal places after the decimal point. Thus, use either "7.8" and "10.0" or "7.89" and "10.00" throughout the article.

6. Give specific dates in the form "22 December 1999". Decades should be referred to as 'the 1990s'. Please spell out the "nineteenth century", etc.

7. Displayed formulae should be numbered consecutively throughout the manuscript as (1), (2), etc. against the right-hand margin of the page. In cases where the derivation of formulae has been abbreviated, it is of great help to the referees if the full derivation can be presented on a separate sheet (not to be published).

8. Do not embed "graphically designed" equations, but prepare these using the word processor's facility **(Example: Microsoft equation editor)**.

9. Distinguish between figures (diagrams) and tables (statistical material) and number them in separate sequences. Each table/figure should have a brief but descriptive title. All table/figures should be as self-explanatory as possible, incorporating any necessary descriptive material in a note at the base of the table.

10. Do not import the tables/figures into the text file but, instead, indicate their approximate locations directly in the electronic text. **Example: [Insert Table 1 about here], [Insert Figure 1 about here]**.

11. Footnotes should be numbered sequentially.

12. All references used in the text should be listed in alphabetical order of the authors' surnames at the end of the text. References in the text should include the name(s) of author(s) with the year of publication in parentheses, and all references should conform to the style of the Journal. Further information on questions of style may be obtained from the Editor, The Lahore Journal of Business, Lahore – Pakistan.

13. To avoid unnecessary errors you are strongly advised to use the "spellchecker" function of your word processor.

14. Any manuscript which does not conform to the above instructions may be returned for the necessary revision.

15. Before publication, page proofs will be sent to the corresponding author. Proofs should be corrected carefully; the responsibility for detecting errors lies with the author.

16. All submissions should be in electronic format to ljb@lahoreschool.edu.pk

Lahore Journal of Business – Call for Papers

Aims and Scope

Lahore Journal of Business (LJB), a peer reviewed publication of Lahore School of Economics, is aimed at providing a specialized forum for dissemination of qualitative and quantitative research in various areas of business administration. LJB invites researchers, policy makers and analysts to submit competitive theoretical and empirical papers that explore and contribute to the understanding of various areas in business domain. The Journal aims at bringing together state of art research findings, particularly from emerging markets, in various business disciplines including (but not limited to) accounting, banking, management, marketing, finance, investments, human resource management and organizational behavior.

Paper Submission and Review Process

The submission of articles will be followed by an editorial review to determine the relevance of manuscript and scope of the Journal. On editorial acceptance, there will be a double-blind review by our advisory panel that comprise of eminent international and local researchers and scholars. The recommendations/comments of the referees will be communicated to the corresponding authors. The papers are now being considered for publication. The manuscript can be submitted in electronic format (.doc and .docx only) to ljb@lahoreschool.edu.pk

Key Areas of Interest: *Accounting, Banking, Management, Marketing, Finance, Strategy, Human Resource Management, Organizational Behavior etc.*

Contact Info

Dr. Nawazish Mirza
Associate Professor – Finance
Editor, Lahore Journal of Business
Lahore School of Economics, Lahore, Pakistan
T +92 (0) 42 36560936 Ext 182
ljb@lahoreschool.edu.pk

THE LAHORE JOURNAL

OF

BUSINESS

Lahore School of Economics

INTERSECTION MAIN BOULEVARD PHASE VI DHA AND BURKI ROAD, BURKI LAHORE, PAKISTAN.
TEL: 0092-42-36561230 FAX: 0092-42-36561230

SUBSCRIPTION FOR PUBLICATIONS

1. **Kindly enter a subscription for the following publication of the Lahore School of Economics:**

Publication	Yearly subscription (within Pakistan)	Yearly subscription (outside Pakistan)	Period (No. of years)	Total Payment
1. The Lahore Journal of Business	Rs. 600.00	US $ 50	….........	….........

2. **The subscription is to be in the following name and address:**

Name: ---

Address: ---

----

3. **A crossed cheque/demand draft for the sum of Pakistan Rupees/US $________ is enclosed made out in the name of The Lahore School of Economics to cover the above subscription.**

4. **Please address your order to: Editor, Lahore Journal of Business, Centre for Research in Economics and Business, Lahore School of Economics, Intersection Main Boulevard Phase VI DHA and Burki Road, Burki Lahore, 53200, Pakistan.**

Signature: -------------------------------------

Name: ---

Date: ---

The Lahore School of Economics

The Lahore School of Economics (established 1993) is one of Pakistan's leading centres of learning for teaching and research in economics, finance and business administration. The Lahore School of Economics' objectives include: (i) The training of young Pakistanis as professional economists, finance managers, accountants, financial analysts, bankers and business executives, and (ii) Undertaking research in economics, management, finance and banking to further deepen understanding of major economic facts, issues and policies.

The Lahore School was granted a Charter in January, 1997 by an Act of the Provincial Assembly of the Punjab: The Lahore School of Economics Act 1997 (Act II of 1997). The Charter vests the powers of an independent degree granting institution to The Lahore School.

The Lahore School has both undergraduate and graduate programs in economics, business information systems and finance. Its postgraduate program leading to the MPhil and PhD degree is administered through the Lahore School's Centre for Research in Economics and Business (CREB). The student body and faculty comprise both national and expatriate Pakistanis and The Lahore School encourages expatriate Pakistanis to join as students or as faculty.

The Lahore School's publication program comprises The Lahore Journal of Economics, The Lahore Journal of Business (bi-annual publications), Lahore Journal of Policy Studies, a Seminar Paper Series and a Text Book Series. The Program encourages both in-house and external contributors.

For further information, please call (Pakistan 92-42-) 35714936 or 36560969 or visit the Web page: www.lahoreschoolofeconomics.edu.pk

Printed by Lahore School of Economics Press

The Lahore Journal of Business
1:1 (Summer 2012): pp. 1–21

Oil Prices and Government Bond Risk Premiums

Hervé Alexandre and Antonin de Benoist*

Abstract

This article analyses the impact of oil prices on bond risk premiums issued by emerging economies. No empirical study has yet focused on the effects of oil prices on government bond risk premiums. We develop a model of credit spread with data from the EMBIG index of 17 countries, from 1998 to 2008. An analysis in time series is carried out on each country and a panel analysis used to determine the global impact of oil prices on investors' risk perceptions. We suggest a new estimator for oil prices to take into account the effect of the price variance, and show that oil prices influence the risk premiums of sovereign bonds along with the price volatility that increases the accuracy of the model.

Keywords: Oil prices, sovereign debt, risk premium.

Classification: F30, G12, G15.

1. Introduction

For decades, government financing has been an increasing problem for countries worldwide. Globalization offers more ways of financing but also more uncertainty, particularly for countries with political problems or for those that do not fall within the most developed subsample.

Emerging economies such as Brazil or China have been developing steadily since the beginning of the 1990s. During this period, their issuing of government bonds has increased considerably, underlining their need for substantial investment in infrastructure and long-term projects. At the same time, they have had to face a series of financial crises, which has greatly reduced their credit capacity and increased the spread—and therefore the cost—of financing. The determinants of this spread and macroeconomic indicators alone cannot explain investors' perception of risk.

---

* The authors are based at the Université Paris-Dauphine in France and can be contacted at herve.alexandre@u-dauphine.fr. They are grateful to this article's referees for the latter's very helpful comments.

Among the factors determining the risk associated with bonds, the price of oil price is a key element to be considered. As Edwards (1985) points out, nine of the ten last recessions have been preceded by oil crises. Moreover, the volatility of oil prices has strongly increased since January 1998. The impact of variations in oil price on economic performance—growth, productivity, and inflation—has, therefore, been the subject of some empirical studies (see Blanchard & Gali, 2007).

If international liquidity has been a subject of interest that has been well discussed in the financial literature, the impact of oil prices has not excited nearly the same interest. Min's (1998) article is, to our knowledge, the only study to include oil prices as an explanatory variable, but the author does not find significant links between variations in oil price and government bond spreads.

This article analyzes the impact of oil prices on the bond risk premiums issued by emerging economies. It is part of the wider framework of studies on the risk measure associated with foreign bonds (Edwards, 1985; Min, 1998; Pan & Singleton, 2007). In spite of the multiplicity of works, however, it would seem that no empirical study has yet focused on the effects of oil prices on government bond risk premiums. In order to do so, we develop a model of credit spread[1] with data from the emerging markets bond index global (EMBIG) of 17 countries from 1998 to 2008.

Our empirical study proceeds with three estimates. The first is an analysis of each country in time series to show how its individual characteristics influence the relation between oil prices and the risk premium of government bonds. The second is a panel analysis that examines the essential role of oil prices as a risk factor that is both global and external. The third is an improved variant of the panel analysis that confirms the essential role of the volatility of oil prices in estimating the risk associated with sovereign bonds.

2. A Review of the Literature

2.1. Determinants of Risk Associated with Government Bonds

The issuing of government bonds in emerging countries increased heavily during the 1990s, improving the liquidity of this financial product.

[1] This spread corresponds to the difference between the yield of a risky bond and that of one that is virtually risk-free, such as the treasury bills issued by the US.

At present, government bonds are the first source of financing in developing countries, after foreign direct investment. At the same time the public debt market in developing countries has experienced serious upheavals following several financial crises. In these conditions, it is important to understand the factors determining the price of government bonds.

2.1.1. Macroeconomic Determinants

Different models allow for the selection of macroeconomic variables by modeling the borrowing status of a country. The principal indicators of solvency are (i) weak stock of debt, (ii) weak interest rate, and (iii) production growth. The economy's level of openness is another key factor in the international solvency of the country. Other variables of competitiveness, such as the exchange rate, can also play an important role in the country's credit risk.

One of the founding empirical studies on international credit markets is that of Edwards (1985), who is particularly interested in the measurement of government bonds in the context of the debt crises of the 1980s. The author develops a model that considers emerging countries as small borrowers within an almost perfect financial market: the spread depends on the default probability, which itself results from macroeconomic variables. The data comes from the World Bank and International Monetary Fund, and concerns 727 debt instruments of 19 of the least developed countries for the period 1976–1980. Several determinants are considered: (i) the ratio of investment growth as a percentage of gross domestic product (GDP), (ii) the average amount to be invested, (iii) the growth of GDP per unit of capital, (iv) the rate inflation, (v) international reserves, (vi) the deflation rate, and (vii) government spending added to GDP.

Edwards' (1985) empirical results show that the development of spread takes into account the economic characteristics particular to the country under consideration. For example, the debt-to-GDP ratio is positively correlated with the spread level. The sum of international reserves is significant and plays an important role in determining the amount of the risk premium. The proportion invested has a negative impact on default probability and the spread of government securities. The study shows that temporal differences within the same country are more important than differences between countries, and concludes that investors take into account a country's individual macroeconomic specificities.

Cantor and Packer (1996) analyze the determinants of the spread of government bonds in over 49 countries in 1995, taking into account their GDP growth, inflation, current account, debt, indicators of economic development, and agency ratings (Moody's, Standard & Poor's). The authors do not find any significant relationship between countries' macroeconomic indicators and the fluctuations in spread. A subsequent empirical study by Kamin and von Kleist (1999) analyzes more than 304 sovereign bonds issued in the 1990s. The authors find that the spread in Latin American countries is, on average, more than 39 percent higher than those in Asia, indicating a segmented government bonds market.

Min (1998) studies the economic determinants of the yield from government bonds (made out in American dollars) in emerging countries for the period 1991–1995. The determinants of default probability are regrouped into four groups of variables that explain the spread level: (i) liquidity and solvency, (ii) basic macroeconomics, (iii) variables of external shocks, and (iv) indicative variables. The results show that an economy's liquidity and solvency play an essential role through the GDP-to-debt ratio and ratio of international reserves to GDP, and integrates the effect of external shocks on the country's risk of default. Most of the regional specificities considered are not significant, which indicates that common factors determine the spread level—this result is confirmed by recent studies such as that by Longstaff, Pan, Pedersen, and Singleton (2007). Finally, Min shows that the volatility of spread is symmetrically influenced both by liquidity and macroeconomic fundamentals such as the rate of inflation, the GDP-to-debt ratio, and the ratio of international reserves to GDP.

Following the increase in volatility of sovereign debt at the end of the 1990s, many investors appeared on the market to take advantage of investment opportunities. Ades, Kaune, Leme, Masih, and Tenengauzer (2000) first established an arbitration model on bond spreads issued in 15 emerging economies from 1996 to 2000. In the same way as Min (1998), Ades et al. break down the macroeconomic variables into four categories: (i) solvency (real gross national product [GNP]); (ii) liquidity (debt/public investment budget, international reserves, budgetary balance/GNP, and LIBOR); (iii) external shocks (exchange rate, exports/GNP); and (iv) indicative variables (previous default). From the evaluation model, it would seem that 12 of the 15 countries considered had undervalued bonds while only one country reached its basic value.

Eichengreen and Mody (1998) analyze more than 1,300 bonds issued by 55 emerging economies between 1991 and 1997. The explanatory factors of the spread are the ten-year interest rate on American bonds, the ratio of external debt to GDP, the ratio of debt to exports, the ratio of international reserves to GDP, the level of growth to GDP, and finally the ratio of budget deficit to GDP. The authors' results confirm that an increase in the quality of a country's credit increases the probability of bond issuing, and reduces the government bond premium. The market differentiates between countries as a function of the quality of the borrower.

2.1.2. Government Bonds: A Real Class of Assets?

Is the risk associated with government bonds merely idiosyncratic or is it determined by global economic factors? To answer this question, Longstaff et al. (2007) examine the profile yield/risk of government bonds to analyze the associated credit risk determinants in the context of the capital asset pricing model. Their study concerns the monthly credit default swap (CDS) premiums made out in US dollars between October 2000 and May 2007 for each of the 26 countries studied. CDS premiums have the advantage of directly reflecting investors' risk perceptions. Moreover, the CDS government bonds market is often more liquid than the bonds market itself.

In order to analyze the determinants of the CDS sovereign bonds premium, the authors include four categories of explicative variables: (i) local variables (yield of local market, exchange rate, and sum of reserves); (ii) financial variables (stock market and American bonds market); (iii) variables of global risk premiums (yield of the Standard and Poor's (S&P) 500 index and variation of the spread between historical volatility estimated on the options of the volatility index (VIX); and (iv) an investment flow variable.

Longstaff et al.'s (2007) results reveal a very strong correlation between the CDS premiums of the different countries. Three principal global factors explain 50 percent of their variation. The spread of CDS on government bonds depends on American shares (S&P500 and NASDAQ indices), the bonds market, and a global risk premium. Pan and Singleton (2007) show that the spreads of different countries are strongly correlated with the VIX. The global risk, identical in the 26 countries, determines more than 30 percent of the total development of CDS spreads on government bonds. The macroeconomic determinants specific to each country represent only a small part of the total development of premiums. As in the work of Min (1998) or Ferrucci (2003), it seems that the risk premiums are higher in Latin America than in other regions.

2.2. *Impact of Oil Prices on Growth, Inflation, and Productivity*

2.2.1. *Oil Crises and Economic Performance: Empirical Evidence*

Economists have long examined the relationship between oil prices and macroeconomic performance. Since the 1970s and the growing dependence of oil-importing economies, many oil crises have occurred: the crash of 1986, the price increase of 1990/91 associated with the Gulf war, the growth of 2000, and the crisis of 2003 associated with the war in Iraq. In 2007/08, oil prices rose strongly as the result of a conjunction of economic, political, geological, and climatic factors. Thus, oil price variations depend to a large extent on exogenous events, such as those linked to the political situation in the Middle East, the development of cartels, or military conflicts.

Increases in oil price are deemed responsible for recessions, inflation, and the reduction in productivity of the mid-1970s. Numerous empirical studies such as that of Hamilton (1983) show that the relationship between oil prices and GDP is more than a simple statistical coincidence. Determining the mechanisms by which oil prices affect macroeconomic conditions is essential to quantify the former's impact on a country's solvency and to measure the spread of sovereign bonds. Nevertheless, some economists have questioned this idea (see Barsky & Kilian, 2004), arguing that macroeconomic variables partly determine fluctuations in oil price.

2.2.2. *How Oil Prices Influence Growth, Inflation, and Productivity*

Effect of Energy Consumption on GDP

The elasticity of production to energy prices depends on the proportion of energy being used in production. Empirically, this proportion is relatively small. For example, in 2000, the US's consumption of oil—which reached 7.2 million barrels—represented only 2.2 percent of GDP. However, it is important to point out that this percentage has risen substantially following recent rises in oil price, although it remains small relative to production.

Nevertheless, the relations of cause and effect between variations in oil price and GDP are all but simple. Bohi (1991) shows that there is no empirical evidence to support the idea that countries with higher energy costs are more severely affected by an oil crisis than those that rely less on oil as a source of energy. Empirical studies show that the cost of an oil crisis is not so much the result of a rise in oil prices as that of a fall in the consumption of other factors of productions to which it leads. Rotemberg and Woodford (1996) show that a 10-percent increase in the oil price can lead to a 2.5-percent fall in GDP over an 18-month period.

Sectoral Reallocations

An oil crisis can have a different impact on capital and employment, and cause reallocations between sectors of activity. Hamilton (1983) shows that oil crises reduce demand in other industries, which leads to a redistribution of work between sectors of activity. The costs of capital adjustment and work, following an oil crisis, have been the subject of much research (see Lee, Ni, & Ratti, 1995).

Monetary Policies and Inflation

Certain studies highlight the role of monetary policy in the relationship between oil prices and GDP. According to Barsky and Kilian (2004), the recession of 1973/74 was one of the consequences of the US Federal Reserve's monetary expansion in response to fears about inflation, which led to an increase in the oil price. Monetary policies can also cause inflationary spirals, wages-prices, caused by the oil price.

The macroeconomic effect of an increase in the oil price can potentially result in stagflation. This phenomenon is particularly important as an explanation of the crisis of the 1970s. As the rate of inflation is linked to monetary policy, the impact of an oil crisis depends primarily on the reaction of central banks to this economic shock. Hooker (2002) illustrates this phenomenon by showing that oil crises contributed significantly to inflation in the US up to 1981, the year in which the question of inflation became a priority of monetary policy.

Modifications of Channel Transmissions

While the economies of OECD countries have seen real variations in oil price in 2000 and 2003—which were as serious as the oil crises of 1973 and 1979—no variations in GDP or inflation were recorded. This calls into question the mechanisms put in place with regard to the relationship between oil prices and macroeconomic conditions.

Blanchard and Gali (2007) confirm the hypothesis that oil prices influenced the stagflation of the 1970s, but point out that other effects were at work. Globally, economies—notably those of the OECD—are far less sensitive to fluctuations in the oil price. The impact of variations in the oil price on inflation has weakened. There may be several explanations for this. First, wage inflexibility has increased, which partly explains the reduced impact of oil price fluctuations. Second, central banks have actively adopted a policy of maintaining a low rate of inflation since the beginning of the 1980s (Herrera & Pesavento, 2009). Finally, energy consumption and economies' dependence on oil have both dropped, even though there are disparities according to the country concerned.

2.2.3. Oil Prices and Economic Risk: The Case of Oil-Producing Countries

Most studies have focused on oil importing countries, but the relationship between economic growth and oil prices is radically different in an oil-producing country. Mechanically, an increase in the oil price should lead to an increase in GDP. This connection is, however, questioned by the efficiency of wealth redistribution systems and economic development models. Corden and Neary's (1982) "Dutch disease" model predicted that an important increase in oil revenues could damage the GDP of certain developing countries. This model was empirically backed up in oil-producing countries during the 1970s—the production of oil was developed to the detriment of the manufacturing and agricultural sectors.

It is certain that oil prices affect the credit risk of all governments. If the relationship between oil prices and macroeconomic performance has been developed since 1980, it remains significant. However, no empirical study has specifically examined the impact of oil prices on the spread of government bonds. This is the subject of the second part of our article.

3. Empirical Study

Our objective is to quantify the impact of the development of oil prices on the risk premium of government bonds. To do this, after presenting some descriptive statistics, we will proceed with an analysis in time series on each country considered, then to a panel analysis.

3.1. Data Sources

We use daily data on government bond spreads across 14 emerging countries and four regions (Latin America, the Middle East, Asia, and Africa). The data is obtained from DataStream and Reuters, and spans almost ten years from 1 January 1998 to 30 May 2008.

3.1.1. Explained Variables

The spread of government bonds refers to the risk premium that the bond-holder demands from the seller to hold the bond. For bonds issued at par, the spread corresponds to the difference between the bond's interest rate and the no-risk interest rate, here, the interest rate of bonds from the American treasury.

We use the EMBIG published by JP Morgan—an index of the spread of emerging countries' government bonds made out in US dollars. This index measures the difference between the premium paid by an emerging country and that of an American treasury bond of similar maturity. It is calculated from the average of all the bonds weighted by the capitalization of the bonds market. In contrast, the emerging markets bond index (EMBI) includes only liquid bonds, including Eurobonds and Brady bonds, whose minimum face value is USD500 million. The EMBIG's maturity period is longer than two and a half years, and covers more than 27 countries since 1998 (the EMBI covers only five countries from 1991 to 1995 and 11 since 1995).

Including the two series in the same empirical study creates a selection bias because the EMBI covers only Brady-type bond yields and the yields on certain structured instruments. Moreover, the two indices may yield different risk measurements because the composition of the two bond portfolios is different. We consider only the EMBIG, which is largely sufficient to bring the empirical estimation to a satisfactory conclusion. The EMBIG allows a more pertinent geographical analysis by region rather than by country. The indices by region are calculated as a geometrical average of the country indices.

3.1.2. Explanatory Variables

Our study draws on the models studied in the first part. The EMBIG makes it possible to obtain data gathered over several decades, which is not the case in most macroeconomic series. Following Longstaff et al. (2007), we include three sets of independent variables in our model: (i) market risk variables, (ii) exchange rates, and (iii) external shocks.

The market risk is interpreted by two indices: the Chicago Board Options Exchange VIX and the S&P500 index. The VIX is a measure of the implicit volatility of a bond in the S&P500 index, and concerns investors' perceptions of risk. An increase in the VIX is explained by an increase in the bonds spread. Pan and Singleton (2007) show that the premium bonds of different countries are strongly correlated to the VIX and, more generally, to the S&P500 index, which is a factor of global risk. We assume that these two indices have positive impacts on the spread.

We include in our model the country's interest rate in dollar units (USD)—noted as FX—which reflects both the economy's competitiveness and the country's solvency. These two factors have a distinct impact on the spread. An increase in the exchange rate reduces the competitiveness of a country and increases the EMBIG, while an increase in the exchange rate increases the country's ability to fulfill its contract and thus reduce the bond risk premium.

External shocks are highlighted in this study by international liquidity and oil price. An increase in the interest rate increases the cost of new finance and debts that have already been contracted—a result confirmed by Eichengreen and Mody (1998). The impact of international liquidity is analyzed by the interest rate over a three-month period of US treasury bills, i.e., the short-term interest (STI), and by the ten-year or long-term interest (LTI) rate of US treasury bonds. As explained by the liquidity preference theory and shown empirically by Ferrucci (2003) and Kamin and von Kleist (1999), the impact of STI should be positive while that of LTI should be negative.

3.1.3. Impact of Oil Prices

We use the West Texas Intermediate (WTI)—also known as the Texas Light Sweet—to represent variations in the oil price. The WTI index is an index of light crude oil that serves as a yardstick for establishing the average oil price from the US. Economic theory suggests studying the real oil price rather than its nominal price. Nevertheless, while taking into account the wide range of fluctuations in oil price and the low inflation rate over the period considered, the choice to use the real price or nominal oil price does not interfere with the estimation of the spread. In the tradition of most empirical studies, our estimation adopts a nominal price level in logarithm form, noted as LN (WTI).

3.2. Descriptive Statistics

Tables 1 and 2 present a statistical description of the variables.

Table 1: Statistical description of dependent variables

Country or region	Mean	Median	Max.	Min.	Standard deviation	Observations
Africa	339.79	344.00	530.00	130.00	115.15	2,500
Argentina	111.07	104.00	194.00	46.00	41.62	2,500
Asia	218.45	219.00	340.00	100.00	66.41	2,500
Brazil	339.74	271.00	665.00	111.00	165.38	2,500
Bulgaria	505.59	542.00	745.00	182.00	150.12	2,500
China	217.84	233.00	301.00	136.00	47.72	2,500
Colombia	183.31	172.00	327.00	69.00	72.25	2,500
Latin America	239.77	211.00	382.00	116.00	75.03	2,500
Mexico	253.46	261.00	394.00	117.00	79.34	2,500
Middle East	162.99	165.00	255.00	99.00	46.12	2,499
Nigeria	344.26	328.00	539.00	136.00	144.25	2,500
Panama	428.93	408.00	712.00	198.00	144.49	2,500
Poland	269.33	282.00	387.00	162.00	62.23	2,500
Russia	342.74	379.00	627.00	25.00	185.45	2,500
South Africa	260.36	282.00	379.00	109.00	78.84	2,500
Thailand	160.25	175.00	194.00	74.00	34.35	2,500
Turkey	233.17	213.00	396.00	91.00	93.68	2,500
Venezuela	355.40	281.00	637.00	96.00	163.75	2,500

Source: Authors' calculations.

The average differs greatly according to the country and the region. Therefore, with an average EMBIG of more than 509.59, the bonds issued by Bulgaria show the highest spread in the sample. Contrary to this, with an EMBIG average of 111.07, Argentina has the lowest spread of the countries and regions covered in this study. The variance is also very different according to the country. Argentina shows the weakest standard deviation, 41.62, while Russia shows the strongest at 185.45. Evidently, countries that have the highest average level of government bond spread have a particularly high variance.

Table 2: Statistical description of independent variables

Variable	Mean	Median	Max.	Min.	Standard deviation	Observations
WTI	41.16	31.85	121.57	10.73	22.46	2,610
VIX	20.29	20.00	45.00	9.00	6.86	2,717
S&P500	1,217.88	1,215.81	1,565.15	776.76	176.77	2,618
STI	3.53	3.94	6.42	0.61	1.71	2,584
LTI	6.62	6.46	9.09	4.16	1.02	2,717

Note: LTI = long-term (30-year) interest rate, STI = short-term (three-month) interest rate.
Source: Authors' calculations.

The short-term risk-free interest rate (STI, three-month maturity) is lower than the long-term risk-free interest rate (LTI, 30-year maturity). Carrying out a unit root test on each time series sample reveals that the government bonds spread is part of a first-order integration (1).

3.3. *Estimation and Interpretation*

3.3.1. *Analysis of Country/Regional Time Series*

Using a linear model with least squares estimation, we carry out a time series analysis of each country/region's government bond spread. This makes it possible to rely on the impact of the oil price on the EMBIG index on the country/region's idiosyncratic situation.

$$\begin{aligned}
&\log(EMBIG_{it}) - \log(EMBIG_{it-1}) = \\
&\beta_{i1}(STI_t - STI_{t-1}) + \beta_{i2}(LTI_t - LTI_{t-1}) + \beta_{i3}(VIX_t - VIX_{t-1}) + \\
&\beta_{i4}(SP500_t - SP500_{t-1}) + \beta_{i5}(\log(WTI)_t - \log(WTI)_{t-1}) + \varepsilon_{it} \qquad (1)
\end{aligned}$$

LTI is the long-term interest rate, STI is the short-term interest rate, WTI is the West Texas Intermediate index, and VIX is the Chicago Board Options Exchange volatility index. Table 3 presents the results of this model, that is to say, the impact of independent variables on the EMBIG logarithm.

The White test has a very weak p-value—we can thus reject the null hypothesis of homoscedasticity. The t-statistics presented are, therefore, adjusted by the White correction, making it possible to have a consistent covariance matrix estimator and a direct test for heteroscedasticity. The F-tests show that the groups of coefficients are significant at the threshold of 5 percent.

The oil price is a significant indicator of the global risk of external factors (Table 3). Most of the coefficients are positive and significant at the 5-percent level. Any increase in global risk has a knock-on effect on the bonds market. For example, the spread of a country such as Mexico increases in the case of a rise in the oil price. An increase in the oil price increases investors' perception of global risk, whatever the country's individual characteristics and EMBIG level.

In addition, the effect of the oil price differs greatly according to the country. Russia, Argentina, and Venezuela are three countries for which the impact of the oil price on spread is highest. An increase of 1 percent in the oil price manifests itself as an increase in the EMBIG of 0.04 percent in Russia and 0.03 percent in Venezuela. The oil price has a negative impact on the EMBIG for Asia, the Middle East, and countries such as China.

The results show that the development of oil prices has a different impact depending on the country studied. This difference could be explained by the fact that that certain countries import oil while others export it. An increase in the oil price constitutes a financial burden for the former and a benefit for the latter. For the borrowing countries, this transfer of wealth could have an impact on their default probability and the losses associated with it.

With regard to the other explanatory variables (VIX, S&P500, STI, and LTI), our estimate partly confirms the empirical results reviewed in the first part of our study, with the exception of the VIX. This index of market risk has a coefficient that is negative most of the time and significantly so on the threshold of 5 percent. This result could be explained by the migration of investors toward government bonds, which are relatively low-risk following an increase in risk in the stock market. The influence of the S&P500 index is positive and significant. The effect of the exchange rate on the EMBIG is negative and significant, seeming to indicate that an increase in the exchange rate is synonymous with an increase in solvency rather than a drop in competitiveness. Finally, the sign of the long-term interest rate concurs with Ferrucci's (2003) results.

Table 3: Impact of independent variables on EMBIG

Country or region	Market risk			External shock			R^2
	S&P500	VIX	Exch. rate	WTI	LTI	STI	adj.
Africa	0.031***	-0.001***	-	0.006*	-0.004	-0.002	0.044
	(2.80)	(-7.62)		(1.86)	(-1.58)	(-0.99)	
Argentina	0.048*	-0.002***	-0.038***	0.032**	0.007	-0.015***	0.046
	(1.86)	(-9.42)	(-3.65)	(2.55)	(1.45)	(-3.06)	
Asia	0.017***	0.000***	-	-0.002	-0.002	-0.002*	0.004
	(2.58)	(-4.75)		(-0.52)	(-1.59)	(-1.95)	
Brazil	0.021	-0.002***	-0.156***	0.012**	-0.004	-0.007**	0.211
	(1.15)	(-14.04)	(-20.57)	(2.33)	(-1.25)	(-2.06)	
Bulgaria	0.013	-0.001***	-	0.010	-0.002	-0.003	0.045
	(0.87)	(-10.80)		(1.33)	(-0.77)	(-1.18)	
China	0.015**	0.000***	-0.025	0.006*	-0.003**	0.001	0.009
	(2.35)	(4.44)	(-1.57)	(1.81)	(-2.53)	(0.66)	
Colombia	0.027**	-0.001***	-0.010	-0.010	-0.002	-0.003	0.043
	(2.03)	(-10.62)	(-1.48)	(-1.48)	(-0.75)	(-1.22)	
Latin America	-0.002***	-0.015	-	0.015**	-0.004	-0.007***	0.081
	(-14.38)	(-1.11)		(2.24)	(-1.42)	(-2.61)	
Middle East	-0.014**	0.000	-	-0.021***	-0.002	-0.002**	0.030
	(-2.18)	(-1.28)		(-2.59)	(-1.61)	(-2.03)	
Mexico	-0.010	-0.001***	-0.034***	0.007*	-0.003*	-0.003**	0.156
	(-1.16)	(-8.06)	(-17.51)	(1.69)	(-1.95)	(-2.05)	
Nigeria	0.017	-0.001***	-	0.013	-0.007**	-0.004	0.015
	(1.10)	(-5.78)		(1.64)	(-2.35)	(-1.25)	
Panama	0.012	-0.001***	-	0.021***	0.002	0.002	0.044
	(1.14)	(-10.31)		(4.03)	(1.16)	(0.96)	
Poland	0.012	0.000***	-0.029***	-0.008**	-0.002	0.002**	0.037
	(1.58)	(-4.95)	(-8.17)	(-2.13)	(-1.38)	(2.29)	
Russia	-0.038	-0.002***	-0.011***	0.040**	-0.006	0.008	0.058
	(-1.18)	(-8.09)	(-8.76)	(2.54)	(-0.88)	(1.26)	
S. Africa	0.038***	0.000	-0.010***	0.012**	0.001	0.003	0.018
	(3.32)	(1.57)	(-5.96)	(2.14)	(0.41)	(1.46)	
Thailand	0.024**	0.000***	-0.002***	-0.007	-0.006***	-0.001	0.014
	(2.41)	(-3.65)	(-3.02)	(-1.33)	(-2.86)	(-0.59)	
Turkey	0.036**	-0.001***	-0.177***	0.010	-0.008**	0.007**	0.099
	(2.29)	(-5.11)	(-14.84)	(1.28)	(-2.44)	(2.42)	
Venezuela	-0.031	-0.002***	-	0.026***	-0.006	-0.006	0.060
	(-1.66)	(-12.04)		(2.93)	(-1.57)	(-1.61)	

Note: LTI = long-term (30-year) interest rate, STI = short-term (three-month) interest rate. Adjusted t-statistics of White correction are given in parentheses below coefficients. Asterisks *, **, and *** indicate significance at 90, 95, and 99 percent, respectively.
Source: Authors' calculations.

3.3.2. Panel Analysis

The previous study gives an insight into the individual impact of oil prices on a country's EMBIG. It would be interesting to design a model that could quantify this impact from a global point of view. The question would be, therefore, to know what effect an increase in the oil price could have on the spread of government bonds.

The data we have used for this study includes more than 17 countries and regions, and has been recorded on a daily basis over almost ten years from January 1998 to April 2008. A panel analysis was necessary to improve the results. Model 2 takes the following form:

$$\log[EMBIG_{it}] = \alpha_i + \beta_1 STI_{it} + \beta_2 LTI_{it} + \beta_3 VIX_{it} +$$

$$\beta_4 SP500_{it} + \beta_5 \log(WTI)_{it}) + \varepsilon_{it} \qquad (2)$$

The number of observations is 25,669 and the number of groups, 17. Table 4 presents the results of our panel analysis. All the coefficients are significant at 5 percent. In the case of a fixed effects model, the most relevant R^2 value is the R^2 within because it gives an idea of the intra-individual share of the dependent variable explained by the explanatory variables. The R^2 within is 0.6396, which is very satisfactory. The R^2 between (0.1044) gives an idea of the contribution of fixed effects to the model.

In order to determine which of the two models is the most relevant, we revert to the Haussmann test's hypotheses:

- H0: The random models are equivalent.
- Ha: The fixed effects model is better than the random effects model.

Applying this test makes it possible to reject the null hypothesis according to which the models are equivalent. Here, the most relevant model is the fixed effects model. If the test makes it possible to categorize between the two models, the model carried over must depend on other, more theoretical, considerations. Allowing for the existence of random effects returns to the supposition that the factor representing the individual effects is not correlated with the explanatory variables. This hypothesis is particularly strong for our model and, consequently, the fixed effects model is the more appropriate one for our study.

The panel analysis shows that the oil price has a significant effect on the EMBIG. An increase of 1 percent in the former brings about an increase of 0.298 percent in the spread. The coefficient of the oil price is significant at 1 percent. The development of the oil price is a factor of global and external risk, which influences the cost of credit for governments.

Table 4: Results of panel analysis 2(a)

Explained variable: (EMBIG) log		
Explanatory variables	**Random effects model**	**Fixed effects model**
Constant	4.90169***	4.90171***
	(-58.47)	(150.63)
STI	-0.02093***	-0.02093***
	(-11.32)	(-11.32)
LTI	-0.04473***	-0.04473***
	(-7.36)	(-7.36)
S&P500	0.00016***	0.00016***
	(8.48)	(8.48)
VIX	0.00265***	0.00265***
	(6.45)	(6.45)
Log (WTI)	0.29838***	0.29838***
	(96.68)	(96.69)
σ	0.31851	0.32599
Error	0.29196	0.29196
R^2 within	0.63960	0.63960
R^2 between	0.10440	0.10440
Observations	25,669	25,669

Note: LTI = long-term (30-year) interest rate, STI = short-term (three-month) interest rate, σ = random effects/fixed effects. T-statistics are given in parentheses below coefficients. Asterisks *** indicate significance at 1 percent.
Source: Authors' calculations.

With regard to the other explanatory variables, the empirical study rejoins the results suggested by theory and empirical literature. As Ferrucci (2003) shows, long-term interest rates have a weaker impact than short-term rates. The VIX has a positive and significant effect on the EMBIG, while the S&P500 index has a relatively weak effect on the explanatory variable. This can be explained by the fact that market risk is already represented by the volatility of the options of the S&P500 index, graded VIX. Finally, σ corresponds, respectively, to the fixed effects and random effects of the models.

Numerous empirical studies, such as that of Blanchard and Gali (2007), have shown that the effect of oil prices on macroeconomic performance has dropped significantly since the 1980s. Some studies explain this development as being the result of an increase in the variance of the oil price. In the following panel analysis, the empirical study constructs an index of oil price corrected for volatility. The literature review has shown that the impact of a rise in oil prices on economic activity has dropped since the 1980s. In fact, it is the mechanisms of transmission of oil crises on economic performance that have changed. The literature shows the important role of the volatility of oil prices on a country's economic growth.

Lee et al. (1995) show that the relationship between economic growth and oil price was no longer significant after 1986. The authors defend the idea that oil prices have not lost their effect on GDP if one takes into account the extent of variations in oil price. Crises relating to oil prices are more likely to affect economic performance in an environment of stable oil prices than in one where oil price movements are erratic.

Lee et al. (1995) develop an indicator of the oil price corrected of its variance, and note that this indicator significantly influences macroeconomic performances regardless of the period under consideration. This result concurs with the idea that the impact of oil prices on economic activity is different according to whether or not it is anticipated. Following the example of empirical studies such as Hamilton (1997), Lee et al. state that the effect of oil price is asymmetrical: an increase in oil price is associated with a significant drop in GNP, while a drop in price is not significantly linked to economic activity. This is why we use a positive semi-variance.

In our study, we put in place an indicator that concurs with the empirical results on the nature of the impact of oil prices on economic activity. From this point of view, the indicator appears as a risk premium of oil that takes into account the volatility of oil prices. This variable, denoted as "prime WTI," is defined by the relationship between the yield of the WTI index and a positive three-month semi-variance.

$$\mathrm{Pr}\,imeWTI = \frac{WTI}{T^{-1}\sum\left(\left(WTI_t - \overline{WTI}\right)^2\right).I\left(WTI_t > \overline{WTI}\right)} \tag{3}$$

An increase in the oil price is just as likely to influence the risk premium of government bonds because its variance is weak. This leads us to conclude that such a variable has a positive and significant impact on the measurement of government bonds spread.

We then proceed to a panel analysis, the results of which are shown in Table 5. The method of estimation adopted is similar to that of the analysis in the previous panel. The R^2 within and between indicates that the independent variables explain more than 66 percent of the fluctuation in the EMBIG. The R^2 both within and between is higher than in model 2(a). The coefficients are all individually significant at 5 percent except the S&P500. The latter has no effect on the risk perception of sovereign bonds. The coefficient of the new indicator of the variations in oil price, denoted by "prime WTI," is particularly high and significant on the threshold of 1 percent.

Table 5: Results of panel analysis 2(b)

Explained variables: (EMBIG) log		
Explanatory variables	**Random effects model**	**Fixed effects model**
Constant	4.20574***	4.20576***
	(122.37)	(49.74)
STI	-0.02365***	-0.02365***
	(-13.23)	(-13.23)
LTI	-0.12707***	-0.12707***
	(-24,99)	(-24,99)
S&P500	0.00001	0.00001
	(0.39)	(0.39)
VIX	0.00352***	0.00352***
	(7.69)	(7.69)
WTI premium	0.56223***	0.56223***
	(108.19)	(108.19)
σ	0.31852	0.32600
Error	0.28260	0.28260
R^2 within	0.66240	0.66240
R^2 between	0.10440	0.10440
Observations	25,669	25,669

Note: LTI = long-term (30-year) interest rate, STI = short-term (three-month) interest rate, σ = random effects/fixed effects. T-statistics are given in parentheses below coefficients. Asterisks *** indicate significance at 1 percent.
Source: Authors' calculations.

According to model 2(b), a 1-percent increase in the ratio of the oil price to its variance causes an increase of 0.5622 percent on the EMBIG. This model validates the theory according to which the variance of the oil price plays a determining role in measuring the risk associated with government bonds. Correcting the oil price by its semi-variance makes it possible to obtain better estimates.

4. Conclusion

Oil prices represent a global risk factor likely to influence countries' credit risk. This analysis is the first to demonstrate the significant effect of oil price variations on the bond spread of a country.

The empirical study proceeds firstly with an analysis in time series of each of the 17 countries from January 1998 to 2008. The models developed concur with the theoretical models. Also, the impact of the oil price on the risk associated with government bonds depends on the individual characteristics of the country considered.

Second, we analyze the impact of oil prices on the EMBIG as a factor of global risk. This explains the use of a panel analysis. The estimate means used is a fixed effects model. Oil prices have a positive and significant impact on the risk premiums of government bonds. An increase of 1 percent in oil price increases the EMBIG by 0.26 percent.

Third, the panel analysis uses an indicator that corrects oil price volatility. This indicator is justified by the argument that many empirical studies underline the importance of oil price volatility to the relationship between oil prices and macroeconomic performance. An increase of 1 per cent in the oil price leads to an increase of 0.56 per cent in the EMBIG. From these results, it seems that the first panel analysis underestimates the real impact of the oil price on government bond risks.

In terms of the overall picture, we show that oil prices significantly influence the risk premiums of sovereign obligations. Including this variable in the measurement models of risk associated with government bonds is, therefore, clearly justified.

References

Ades, A., Kaune, F., Leme, P., Masih, R., & Tenengauzer, D. (2000). *A new framework for assessing fair value in emerging market hard currency debt* (Global Economics Paper No. 45). New York, NY: Goldman Sachs Group.

Barsky, R., & Kilian, L. (2004). *Oil and the macroeconomy since 1970* (Discussion Paper No. 4496). London, UK: Centre for Economic Policy Research.

Blanchard, O., & Gali, J. (2007). *Monetary policy, labor market, rigidity and oil price shock*. Paris: Banque de France Foundation for Monetary, Financial and Banking Economic Research.

Bohi, D. R. (1991). On the macroeconomic effects of energy price shocks. *Resources and Energy*, *13*, 145–162.

Cantor, R., & Packer, F. (1996). Determinants and impact of sovereign credit ratings. *Federal Reserve Bank of New York Economic Policy Review*, *2*, 37–53.

Corden, W. M., & Neary, J. P. (1982). Booming sector and de-industrialization in a small open economy. *Economic Journal*, *92*(368), 825–848.

Edwards, S. (1985). LDCs' foreign borrowing and default risk: An empirical investigation, 1976–1980. *American Economic Review*, *74*(4), 726–734.

Eichengreen, B., & Mody, A. (1998). *What explains changing spreads on EM debt: fundamentals or market sentiment?* (Working Paper No. 6408). Cambridge, MA: National Bureau of Economic Research.

Ferrucci, G. (2003). *Empirical determinants of emerging market economies' sovereign bond spreads* (Working Paper No. 205). London, UK: Bank of England.

Hamilton, J. D. (1983). Oil and the macroeconomy since World War II. *Journal of Political Economy*, *91*, 228–248.

Hamilton, J. D. (1997). Measuring the liquidity effect. *American Economic Review*, *87*(1), 80–97.

Herrera, A. M., & Pesavento, E. (2009). Oil price shocks, systematic monetary policy, and the "Great Moderation." *Macroeconomic Dynamics*, *13*(1), 107–137.

Hooker, M. A. (2002). Are oil shocks inflationary? Asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit and Banking*, *34*, 540–561.

Kamin, S. B., & von Kleist, K. (1999). *The evolution and determinants of EM credit spreads in the 1990s* (International Finance Discussion Paper No. 653). New York, NY: Federal Reserve Bank of New York.

Lee, K., Ni, S., & Ratti, R. A. (1995). Oil shocks and the macroeconomy: The role of price variability. *Energy Journal*, *16*, 39–56.

Longstaff, F. A., Pan, J., Pedersen, L. H., & Singleton, K. J. (2007). *How sovereign is sovereign credit risk?* (Working Paper No. 13658). Cambridge, MA: National Bureau of Economic Research.

Min, H.-G. (1998). *Determinants of emerging market bond spread: Do economic fundamentals matter?* (Policy Research Working Paper No. 1899). Washington, DC: World Bank.

Pan, J., & Singleton, K. J. (2007). *Default and recovery implicit in the term structure of sovereign CDS spreads*. Retrieved from http://www.mit.edu/~junpan/sovrev.pdf.

Rotemberg, J. J., & Woodford, M. (1996). Imperfect competition and the effects of energy price increases. *Journal of Money, Credit and Banking*, *28* (Part 1), 549–577.

The Lahore Journal of Business
1:1 (Summer 2012): pp. 23–36

Organizational Role Stress Among Public and Private Sector Employees: A Comparative Study

Bushara Bano and Rajiv Kumar Jha*

Abstract

The aim of this study is to explore the differences in job-related stress, if any, between public and private sector employees, based on ten role stressors. It also examines the role of demographic variables on the stress levels of both public and private sector groups. Our methodology entails a survey of 182 public and 120 private sector employees in Uttar Pradesh, India, whose responses are measured according to an occupational role stress scale. We also use secondary data provided by the literature review. The sample was collected through convenience sampling. On applying the t-test and ANOVA test to the data, we find that both public and private sector employees face moderate levels of stress. While there is no significant difference overall between public and private sector employees in terms of total stress levels, certain individual stressors—such as work experience and educational qualifications—do yield differences. The major limitation of this study is that it was conducted in Uttar Pradesh alone, while the work culture of organizations other than in Uttar Pradesh may be different.

Keywords: Role stress, public sector, private sector.

Classification: M10, M12, M14

1. Introduction

Stress has become a very common phenomenon of routine life, and an unavoidable consequence of the ways in which society has changed. This change has occurred in terms of science and technology, industrial growth, urbanization, modernization, and automation on one hand; and an expanding population, unemployment, and stress on the other. The term "stress" was first used by Selye (1936) in the literature on life sciences, describing stress as "the force, pressure, or strain exerted upon a material object or person which resist these forces and attempt to maintain its original state." Stress can also be defined as an adverse

---

* Ms Bano is a research scholar at the Department of Business Administration at Aligarh Muslim University in India; she can be contacted at syedabushara@gmail.com. Mr Jha is a quality manager at Medica Synergie Pvt. Ltd. in Aligarh, India.

reaction that people experience when external demands exceed their internal capabilities (Waters & Ussery, 2007).

Organizations are an important source of stress, and employees' workloads and professional deadlines have increased manifold. These advancements have created stress among employees in the form of occupational stress, which Sauter, Lim, and Murphy (1996) define as the harmful physical and emotional responses that arise when the demands of a job do not match the worker's abilities, resources, or needs. Occupational stress is further defined as a condition arising from the interaction of people and their jobs, and characterized by changes within people that force them to deviate from their normal functioning (Beehr & Newman, 1978).

The perception of the effects of stress on an individual has changed. Stress is not always dysfunctional in nature, and, if positive, can prove one of the most important factors in improving productivity within an organization (Spielberger, 1980). If not positive, stress can create a number of physical and psychological disorders among employees, and can be responsible for frustration, haste, and job dissatisfaction. As a result, the lack of work may cause complacency within the organization. Stress is, therefore, multidimensional, and its results depend on whether employees perceive it as a problem or a solution.

For our purposes, public sector organizations are considered those that are government-owned and -operated. Such organizations are considered to focus primarily on the administration of essential services and the control and maintenance of a country's social and economic conditions. In contrast, private sector organizations are considered either profit-making enterprises or community service groups that operate independently of the government (Macklin, Smith, & Dollard, 2006).

Different studies have classified occupational stress in terms of physical environment, role stressors, organizational structure, job characteristics, professional relationships, career development, and work-versus-family conflict (see Burke, 1993). Cooper and Marshall (1976) add to this list factors intrinsic to a job, the management's role, and professional achievements. Based on these complexities, stressors can be grouped into two main categories: (i) job-related stressors, and (ii) individual-related stressors.

Stress is measured using a number of instruments. Our focus, however, is organizational role stress (ORS), which measures total role stress. We use Pareek's (1983) scale, which evaluates respondents' quantum of stress in terms of total ORS scores. It also measures the intensity of the following ten role stressors that contribute to the total ORS score:

1. Inter-role distance (IRD): Conflict between organizational and nonorganizational roles.
2. Role stagnation (RS): The feeling of being "stuck" in the same role.
3. Role expectation conflict (REC): Conflicting expectations and demands between different role senders.
4. Role erosion (RE): The feeling that functions that should belong to the respondent's role are being transformed/performed or shared by others.
5. Role overload (RO): The feeling that more is expected from the role than the respondent can cope with.
6. Role isolation (RI): Lack of linkages between the respondent's role and that of other roles in the organization.
7. Personal inadequacy (PI): Inadequate knowledge, skills, or preparation for a respondent to be effective in a particular role.
8. Self-role distance (SRD): Conflict between the respondent's values/self-concepts and the requirements of his or her organizational role.
9. Role ambiguity (RA): Lack of clarity about others' expectations of the respondent's role, or lack of feedback on how others perceive the respondent's performance.
10. Resource inadequacy (RIn): Nonavailability of resources needed for effective role performance.

2. A Review of the Literature

2.1. Studies at the National Level

Sharma (1987) focuses on the managers and supervisors of public and private pharmaceutical organizations to ascertain the role of a motivated climate on four psychological variables: (i) job satisfaction, (ii) participation, (iii) alienation, and (iv) role stress. The study's sample comprises 150 respondents, including 75 managers and 75 supervisors. Sharma's findings indicate that employees of public sector organizations score lower than and differ significantly from those of private sector organizations. However, public sector employees score significantly higher in terms of role stagnation.

Ahmad, Bharadwaj, and Narula (1985) assess stress levels among 30 executives from both the public and private sector, using an ORS scale to measure ten dimensions of role stress. Their study reveals significant differences between public and private sector employees in three dimensions of role stress—role isolation, role ambiguity, and self-role distance. The authors also establish the insignificant effect of several background factors, such as age, level of education, income, marital status, and work experience.

Jha and Bhardwaj's (1989) empirical study of job stress and motivation among 120 frontline managers from both the public and private sector finds that the latter score more than the former in factors such as the need for achievement and total motivation. Chaudhary (1990) probes the relationship between role stress and job satisfaction among bank officers. The author's results indicate that role erosion and resource inadequacy act as dominant stressors while role ambiguity and role expectation conflict are remote contributors to role stress in the sample population.

Srivastava (1991) surveys 300 employees of the Life Insurance Corporation and reports that there is a significant positive correlation between various dimensions of role stress and symptoms of mental ill health. Stress arising from role ambiguity and role stagnation is the most intensively correlated with anxiety. Finally, Dwivedi (1997) assesses the magnitude of trust, distrust, and ORS to determine the extent of this relationship among public and private sector organization. Surveying 55 executives from the public sector and 62 from the private sector, the author finds that stress levels are low in high-performance organizations and high in low-performance organizations.

2.2. Studies at the International Level

Lewig and Dollard (2001) find that public sector employees are subject to greater work-related stress than private sector employees. Dollard and Walsh (1999), however, report that private sector workers in Queensland, Australia, had made twice as many stress claims as public sector workers. Macklin et al. (2006) survey 84 public and 143 private sector employees to assess any significant difference in their stress levels. They conclude that there is no significant difference between employees on the basis of sector, but that there is a significant difference between genders, i.e., female employees are subject to greater stress than males.

D'Aleo, Stebbins, Lowe, Lees, and Ham (2007) examine a sample of 559 public and 105 private sector employees to assess their respective risk profiles. They find that public sector employees face more stress than private sector employees. Malik (2011) collects data on 200 bank employees in Quetta, Pakistan, of which 100 work in public sector banks and the remaining 100 in private sector banks. The author finds that there is a significant difference in the level of stress to which both groups are subject, and that public sector bank employees face a high level of occupational stress.

It is clear that different studies have generated different results on the basis of their particular contexts. Some studies argue that public sector employees are subject to greater stress while others argue the opposite. The literature review shows that work-related stress is almost equal in both the public and private sector, and that research on this topic remains a popular field of enquiry.

3. Objectives and Hypotheses

This study's aims are to (i) examine the difference in stress levels between public and private sector employees, and (ii) assess the impact of socio-demographic factors on employees' stress levels. To do so, we propose the following hypotheses:

- H01: There is no significant difference in ORS among different age groups of employees.
- H02: There is no significant difference in ORS among employees of different marital status.
- H03: There is no significant difference in ORS among employees with different levels of work experience.
- H04: There is no significant difference in ORS among employees with different educational qualifications.
- H05: There is no significant difference in ORS between public and private sector employees.

4. Research Methodology

The sample population for this study comprises a total of 302 employees drawn from different public and private organizations—182 from the former and 120 from the latter. The public organizations sampled include the Archaeological Survey of India, the District Treasury Board, and Hindalco; the private organizations sampled include Tata Motors, TELCO, and Pashupati Oil Mills. The sample was collected on the basis of convenience sampling, and is located in the Agra and Aligarh districts of Uttar Pradesh in India.

4.1. *Reliability of ORS Scale*

ORS is measured on a five-point Likert scale with values ranging from 0 to 4. The scale is used to investigate the ORS arising from ten different role stressors. Table 1 shows that the Cronbach's alpha value of the ORS scale is 0.932, indicating that the scale is highly reliable for this particular study. The table also gives Cronbach's alpha values for the different dimensions of ORS, showing that all the stressors, apart from SRD, have a high Cronbach's alpha value. We can thus eliminate SRD from further study, and examine the remaining nine dimensions of the ORS scale.

Table 1: Cronbach's alpha value of stressors

No.	Variable	Coefficient
1.	Inter-role distance (IRD)	0.800
2.	Role stagnation (RS)	0.717
3.	Role expectation conflict (REC)	0.719
4.	Role erosion (RE)	0.719
5.	Role overload (RO)	0.812
6.	Role isolation (RI)	0.617
7.	Personal inadequacy (PI)	0.720
8.	Self-role distance (SRD)	0.592
9.	Role ambiguity (RA)	0.767
10.	Resource inadequacy (RIn)	0.760
	ORS	0.932

Source: Authors' calculations.

4.2. *Factor Analysis*

The Kaiser-Meyer-Olkin (KMO) test provides a measure of sampling adequacy in which, generally, a value greater than 0.4 is desirable. In this case, the KMO measure is 0.812 (Table 2), implying that the correlation between pairs of variables can be explained to a great degree by other variables. The Bartlett's test value is 0.000, indicating that the value is highly significant.

Table 2: Results of KMO and Bartlett's test

Test	Test statistic	df	Significance value
KMO measure of sampling adequacy	0.812	-	-
Bartlett's test of sphericity	8.619	1225	0.000

Source: Authors' calculations.

Table 3 shows that the value of all components is far higher than 1, implying that they all converge on one overall stressor, i.e., ORS. We can, therefore, conclude that the scale is convergent.

Table 3: Eigenvalue of components

Component	Initial Eigenvalue
1	12.909
2	3.228
3	2.751
4	2.432
5	1.910
6	1.758
7	1.609
8	1.338
9	1.244

Source: Authors' calculations.

We use varimax rotation to carry out a factor analysis of the refined data. Factor loadings indicate the strength of the relationship between a particular factor and a particular variable. In a simple-component matrix, a particular variable may show higher loadings for many factors, making it difficult to determine the variables under any given factor. We solve this problem by rotating the matrix, making it easier to assign a number of variables with greater loading for a particular factor. The rotated-component matrix shows that most of the items load well (> 0.4) on nine factors of the ORS scale. Akinyokun, Angaye, and Ubaru (2009) argue that a value greater than 0.4 should be considered meaningful, allowing us to conclude that there is a strong relationship between the factors and variables on this scale.

5. Data Analysis

The data is analyzed in the form of variables such as ORS scores for public and private sector employees, in which we consider low, medium, and high levels of stress among public and private sector employees, their educational qualifications, duration of service, marital status, and age. Table 4 groups employees by different variables. Using SPSS 16.0 to analyze the results, we tabulate our findings separately.

Table 4: Demographic profile of respondents

Variable	Description	Respondents
Educational qualifications	Group A (up to 12th standard)	56
	Group B (graduate and postgraduate)	232
	Group C (doctorate)	14
Age	Group A (up to 35 years)	176
	Group B (36–50 years)	102
	Group C (more than 50 years)	24
Work experience	Group A (1–10 years)	164
	Group B (11–20 years)	84
	Group C (21–30 years)	42
	Group D (31–36 years)	12
Sector	Group A (public sector employee)	182
	Group B (private sector employee)	120
Marital status	Group A (unmarried)	80
	Group B (married)	222

Source: Authors' calculations.

6. Results and Discussion

In order to rank various stressors, we calculate their mean values and standard deviations, followed by those of the total ORS scale. Table 5 shows that all nine individual stressors give rise to moderate levels of stress among the employees sampled. The mean value of total role stress is 1.4913, implying that employees face moderate levels of total ORS. The highest mean value of role erosion is 1.778, implying that employees are subject to this stressor the most. The highest standard deviation value of role overload is 1.009, indicating that some groups experience role overload more than others.

In order to analyze the role of socio-demographic factors on employees' stress levels, we run a t-test and ANOVA test on the sample. The latter helps assess the difference in total stress between age groups. Table 6 indicates that the age factor is not significant. H01, which states that there is no significant difference in the stress levels of employees of different age groups, is therefore an acceptable hypothesis.

Table 5: Status of stressors

Stressor	Mean	Standard deviation	Rank	Status
IRD	1.675	0.972	2	Moderate
RS	1.597	0.931	4	Moderate
REC	1.358	0.820	8	Moderate
RE	1.778	0.890	1	Moderate
RO	1.365	1.009	7	Moderate
RI	1.562	0.820	5	Moderate
PI	1.393	0.911	6	Moderate
RA	1.112	0.926	9	Moderate
RIn	1.663	0.990	3	Moderate
ORS	1.491	0.654		Moderate

Note: We have calculated the mean score on a scale of 0 to 4, and divided stress levels into ″low“ (0–1), "moderate" (1–2), and "high" (more than 2 and up to 4).
Source: Authors' calculations.

Table 6: Impact of socio-demographic factors on ORS

Hypothesis	Stress	Demographic	Significance value	Remarks
H0 1	ORS	Age	0.280	Accepted
H0 2		Marital status	0.282	Accepted
H0 3		Work experience	0.005**	Not accepted
H0 4		Qualifications	0.002**	Not accepted

Note: ** = significant at 99-percent confidence level.
Source: Authors' calculations.

We use the t-test to analyze the role of marital status on employees' stress levels, and, again, find no significant value. Table 6 also shows that there is no significant difference in ORS among employees of a different marital status. Thus, H02, which states that there is no significant difference in ORS among employees of a different marital status, is an acceptable hypothesis.

Work experience, the third socio-demographic factor, does, however, affect employees' stress levels. Running an ANOVA test on the sample reveals that there is a significant difference in ORS between groups with different degrees of work experience. This implies that H03, which states that there is no significant difference in ORS among groups with different levels of work experience, is not an acceptable hypothesis.

Similarly, we use the ANOVA test to analyze the impact of educational qualifications on employees' stress levels. As Table 6 shows, there is a significant difference in ORS among groups with different levels of educational qualification groups. Thus, H04, which states that there is no significant difference in ORS among groups with different qualifications, is not an acceptable hypothesis.

Calculating the mean, standard deviation, and t-test values for different stressors allows us to compare role stress between the public and private sector. Table 7 shows that there is no significant difference between the two sectors in terms of employees' total stress level. H, which states that there is no significant difference between the two sectors with regard to total role stress, is an acceptable hypothesis.

Table 7: Comparative levels of stress among public and private sector employees

Stressor	Public sector Sample = 182		Private sector Sample = 120		Significance value
IRD	Mean	1.613	Mean	1.770	0.029*
	SD	0.911	SD	1.054	
RO	Mean	1.228	Mean	1.573	0.843
	SD	1.008	SD	0.980	
RI	Mean	1.534	Mean	1.606	0.000**
	SD	0.882	SD	0.718	
RE	Mean	1.806	Mean	1.736	0.441
	SD	0.919	SD	0.846	
REC	Mean	1.312	Mean	1.430	0.536
	SD	0.835	SD	0.795	
PI	Mean	1.470	Mean	1.276	0.000**
	SD	0.990	SD	0.765	
RS	Mean	1.492	Mean	1.756	0.698
	SD	0.909	SD	0.944	
SRD	Mean	1.362	Mean	1.420	0.788
	SD	0.788	SD	0.759	
RA	Mean	1.076	Mean	1.166	0.815
	SD	0.948	SD	0.893	
RIn	Mean	1.742	Mean	1.543	0.156
	SD	1.026	SD	0.923	
ORS	Mean	1.464	Mean	1.532	0.687
	SD	0.677	SD	0.618	

Note: ** significant at 99-percent confidence level, * significant at 95-percent confidence level.
Source: Authors' calculations.

However, on applying the t-test separately to different dimensions of ORS, we find that three factors reflect a significant difference among public and private sector employees. These factors include role isolation, personal inadequacy, and inter-role distance. Table 7 also shows that employees face a moderate level of total role stress, but that the mean values of most of the stressors—apart from role erosion, personal inadequacy, and resource inadequacy—to which private sector employees are subject, is greater than that of public sector employees.

7. Regression Analysis

We find that total role stress, i.e., ORS, is a dependent variable while its other dimensions—IRD, RS, REC, RO, RE, RI, PI, RA, and RIn—are independent variables, which generates total ORS. A regression analysis of the sample reveals that the adjusted R^2 value is 99.3, i.e., 99.3 percent of the variation in the dependent variable ORS is explained by independent variables (stressors). Further, the significant coefficient value of all the dimensions is 0.000, showing that the independent variables all have a significant impact on the dependent variable ORS.

The regression equation takes the form

$$y = ax_1 + bx_2 + cx_3 + \ldots + jx_{10}$$

Based on the analysis, total stress (ORS) is written as

$$ORS = 0.158\ IRD + 0.137\ RS + 0.127\ REC + \ldots + 0.150\ RIn$$

Table 8: Regression results

Stressor	Beta value	Significance value
IRD	0.158	0.000**
RS	0.137	0.000**
REC	0.127	0.000**
RE	0.146	0.000**
RO	0.162	0.000**
RI	0.106	0.000**
PI	0.134	0.000**
RA	0.143	0.000**
RIn	0.150	0.000**

Note: ** = significant at 99-percent confidence level.
Source: Authors' calculations.

8. Conclusion

Our study has led us to conclude that employees in both the public and private sectors face moderate levels of stress, of which they are subject to role erosion the most and resource inadequacy the least. Further, there is no significant difference in total role stress among public and private sector employees. These results support the findings of a number of earlier studies, e.g., Macklin et al. (2006), although we have noted that private sector employees facing slightly more stress than those in the public sector. Our analysis of the impact of various socio-demographic factors on stress level reveals that educational qualifications and work experience have a significant impact on employees' stress levels.

References

Agarwala, U. N., Malhan, N. K., & Singh, B. (1979). Some clarifications of stress and its applications at work. *Indian Journal of Industrial Relations*, *15*(1), 41–50.

Ahmad, S., Bharadwaj, A., & Narula, S. (1985). A study of stress among executives. *Journal of Personality and Clinical Studies*, *1*(1–2), 47–50.

Akinyokun, C. O., Angaye, C. O., & Ubaru, M. O. (2009). Factor analysis of the performance indices of information and communications technology projects in the public sector of the Nigerian economy. *Journal of Technology Research*, *1*, 1–15.

Beehr, T. A., & Newman, J. E. (1978). Job stress, employee health, and organizational effectiveness: A facet analysis model and literature review. *Personnel Psychology*, *31*(4), 665–699.

Burke, R. J. (1993). Toward an understanding of psychological burnout among police officers. *Journal of Social Behavior and Personality*, *8*(3), 425–438.

Chaudhary, A. (1990). *A study of the relationship between job satisfaction and role stress of bank officers*. Unpublished Master's thesis, University of Rajasthan, India.

Cooper, C. L., & Marshall, J. (1976). Occupational sources of stress: A review of the literature relating to coronary heart disease and mental ill-health. *Journal of Occupational Psychology*, *49*(1), 11–28.

D'Aleo, N., Stebbins, P., Lowe, R., Lees, D., & Ham, D. (2007). Managing workplace stress: Psychosocial hazard risk profiles in public and private sector Australia. *Australian Journal of Rehabilitation Counselling*, *13*(2), 68–87.

Dollard, M. F., & Walsh, C. (1999). Illusory correlation: Is work stress really worse in the public sector? *Journal of Occupational Health and Safety*, *15*(3), 219-231.

Dwivedi, R. K. (1997). Trust and role stress. In D. M. Pestonjee & U. Pareek (Eds.), *Studies in organizational role stress and coping*. New Delhi: Rawat.

Jha, P. K., & Bhardwaj, G. (1989). *Stress and motivation: An empirical study on front line managers*. Unpublished manuscript.

Lewig, K. A., & Dollard, M. F. (2001). Social construction of work stress: Australian newsprint media portrayal of stress at work, 1997–98. *Work and Stress*, *15*(2), 179–190.

Macklin, D. S., Smith, L. A., & Dollard, M. F. (2006). Public and private sector work stress: Workers' compensation, levels of distress and job satisfaction, and the demand-control-support model. *Australian Journal of Psychology*, *58*(3), 130–143.

Malik, N. (2011). A study on occupational stress experienced by private and public sector bank employees in Quetta city. *African Journal of Business Management*, *5*(8), 3063–3070.

Pareek, U. (1983). *Role stress scale: ORS scales booklet, answer sheet, and manual*. Ahmadabad: Naveen Publications.

Sauter, S. L., Lim, S.-Y., & Murphy, L. R. (1996). Organizational health: A new paradigm for occupational stress research at NIOSH. *Japanese Journal of Occupational Mental Health*, *4*, 248–254.

Selye, H. (1936). A syndrome produced by diverse noxious agents. *Nature*, *138*, 32–35.

Sharma, T. (1987). *Differential effects of organizational climate on job satisfaction, sense of participation, alienation and role stress*. Unpublished doctoral thesis, Gujrat University, India.

Spielberger, C. (1980). *Preliminary manual for the state-trait anger scale*. Tampa, FL: University of South Florida.

Srivastava, A. K. (1991). A study of the role stress-mental health relationship as a moderator by adopting coping strategies. *Psychological Studies*, *3*, 192–197.

Waters, J. A., & Ussery, W. (2007). Police stress: History, contributing factors, symptoms, and interventions. *Policing: An International Journal of Police Strategies and Management*, *30*(2), 169–188.

The Lahore Journal of Business
1:1 (Summer 2012): pp. 37–58

The Forecasting Ability of GARCH Models for the 2003–07 Crisis: Evidence from S&P500 Index Volatility

Mahreen Mahmud*

Abstract

This article studies the ability of the GARCH family of models to accurately forecast the volatility of S&P500 stock index returns across the financial crisis that affected markets in 2003–07. We find the GJR-GARCH (1,1) model to be superior in its ability to forecast the volatility of the initial crisis period (2003–06) compared to its realized volatility, which acts as a proxy for the actual. This model is then extended to make forecasts for the crisis period. We conclude that the model's ability to forecast volatility across the crisis is not substantially affected, thus supporting the use of the GARCH family of models in forecasting volatility.

Keywords: Forecasting, volatility clustering, financial crisis.

JEL classification: G01, G12, G17.

1. Introduction

Volatility modeling and forecasting has received enormous attention in the last two decades, driven by its importance to the financial sector. Many studies have tried to obtain accurate estimates of volatility, which is a key input into the pricing of options and assets and in hedging strategies.

There are several approaches to forecasting volatility. The options-based approach extracts a volatility estimate from the price of traded options. Another approach is to look at the past prices of financial securities, i.e., historical volatility. A third approach—which we employ in this study—makes use of the generalized autoregressive conditional heteroscedastic (GARCH) family of models. These have been specifically developed to model volatility in financial time series, and the basic model's extensions are able to take into account the asymmetric effects of good and bad news on volatility.

* The author is a research and teaching fellow at the Centre for Research in Economics and Business (CREB) based at the Lahore School of Economics. She can be contacted at mahreenm@gmail.com.

Our aim is twofold. The first is to determine which of the GARCH family models performs best at the out-of-sample forecasting of stock index returns volatility during the initial sample period (2003–06). The second aim relates to the recent subprime mortgage crisis that hit the US market. We will use the selected model to make forecasts for the crisis period and assess its performance for this period relative to before.

The trigger for the recent financial crisis was a shift in how mortgages were issued in the US. The crisis, the effects of which began to show in early 2007, had a major adverse impact on banks and financial markets in the country and around the world. The Standard and Poor's (S&P) 500 composite index, the leading indicator of the US economy's performance, went down 45 percent between 2007 and 2008. It is of interest to examine how well the GARCH models are able to forecast the returns volatility of the index during this turbulent time relative to the preceding tranquil time.

We need measures of true volatility in order to assess the quality of the forecasts made for both periods. Volatility is, however, a latent variable in that it is unobserved and develops stochastically over time. While squared returns are commonly used as a proxy for true volatility, they have proved a noisy estimator. Instead, we will use realized volatility (RV), a proxy that utilizes the extra information provided by intraday returns.

The following sections are organized as follows. The existing literature is outlined in Section 2, the dataset used is described in Section 3, and our proposed methodology explained in Section 4. Section 5 presents our empirical findings and Section 6 provides concluding remarks.

2. A Review of the Literature

2.1. Modeling and Forecasting Volatility

The history of the GARCH models originates in Engle's (1982) seminal study, followed by the more popular generalization proposed by Bollerslev (1986).[1] In one of the earliest studies on the topic, Akgiray (1989) found support for the GARCH (1,1) model's ability to better forecast monthly return variances (using CRSP value-weighted indices for 1983–86) than the ARCH model, historical volatility estimates, and exponentially weighted moving averages.

[1] Bollerslev, incidentally, used the same index as we do—the S&P500—to introduce the popular GARCH specification for modeling financial time series volatility.

Poon and Granger's (2003) extensive review of the literature on financial forecasting spans 93 published and working papers, providing a detailed analysis of the various techniques used in financial forecasting and of the quality of results obtained from each. They conclude that, from within the GARCH family, asymmetric models yield superior forecasts because they factor in the more pronounced effect of a negative shock to volatility than a positive one of same magnitude. In particular, for stock index data, Brailsford and Faff (1996) find evidence in favor of the Glosten-Jagannathan-Runkle (GJR)-GARCH (1,1) model when applied to Australian data, and Engle and Ng (1993) the same for Japanese daily stock index returns (the Japanese TOPIX index) during 1980–88.

Using late 19th- and early 20th-century US data, Pagan and Schwert (1990) propose the exponential GARCH (EGARCH) to be the best fit. Kim and Kon (1994) examine data on 30 individual stocks and three stock indices in the US over 1962–90, and find that the GJR-GARCH (1,3) model performs well for stocks while the EGARCH (1,3) best models stock index volatility. Thus, overall, there is mixed evidence on which specific model is superior to the other. All evidence, however, points toward the superiority of asymmetric GARCH models for stock index returns volatility relative to their symmetric counterparts.

Taylor (2004) employs eight different stock indices from across the world—including the S&P500 for New York—to concentrate on one-step-ahead forecasting. Using weekly data for the period 1987–1995, the author finds that the GJR-GARCH model performs best when the regression analysis uses RV as a proxy for actual unobserved volatility. The GJR-GARCH (1,1) model "estimated using daily returns outperforms all five GARCH models estimated using weekly returns. The extra information supplied by the higher frequency data is clearly beneficial for the GJR-GARCH model."

Corradi and Awartani (2005) use S&P500 index daily data for the period 1990–2001 to study the forecasting ability of several GARCH models. As mentioned in Section 1, a measure of true variance is required in order to evaluate the quality of the forecasts made. Since the true variance based on the population is latent, a proxy is used—in this case, the authors adopt the conventional approach of using squared returns as a proxy for unobservable volatility process since their aim is merely to rank the models. They find that the asymmetric GARCH models are better than the GARCH (1,1), although this dominance is smaller for forecasts of longer horizons.

2.2. RV as a Proxy for True Volatility

The burgeoning literature on time-varying financial market volatility abounds with empirical studies in which competing models are evaluated and compared on the basis of their forecast performance. The variable of interest (volatility) is not directly observable, rather being inherently latent. As a consequence, any assessment of forecast precision is plagued by problems associated with its measurement. Recognition of the importance of this issue led to a number of studies conducted in the late 1990s that advocated the use of so-called RV, constructed from the summation of squared high-frequency returns, as a method for improving the volatility measure (Anderson, Bollerslev, & Meddahi, 2005).

Groundbreaking work by Anderson and Bollerslev (1997) using two series of spot exchange rates (DM-$ and ¥-$ spot exchange rates from 1 October 1987 through 30 September 1992) shows that an alternate proxy, i.e., RV, helps the GARCH model explain more than half of true volatility. The basis of RV is found in continuous time whereby the extra information contained in intraday data reduces the sampling error, yielding better estimates of true unobserved volatility. Anderson and Bollerslev use five-minute frequency data to show improved out-of-sample forecasting as opposed to when squared returns are used.

Subsequent studies, such as that by Hansen and Lunde (2006), establish that the use of squared returns worsens the predictive ability of GARCH models out of sample even when they perform extremely well within the sample. McMillan and Speight (2004) lends further support to the use of RV, concluding that GARCH models can successfully model the conditional variance of financial time series but that their forecasting ability is adversely affected when compared with a fallacious estimate of volatility. Using RV as a proxy, they use data from 17 daily exchange rate series relative to the US dollar for the period 1990–1996 to prove that GARCH models do better at forecasting volatility than the smoothing and moving-average models that had earlier been thought superior (see Figlewski, 1997).

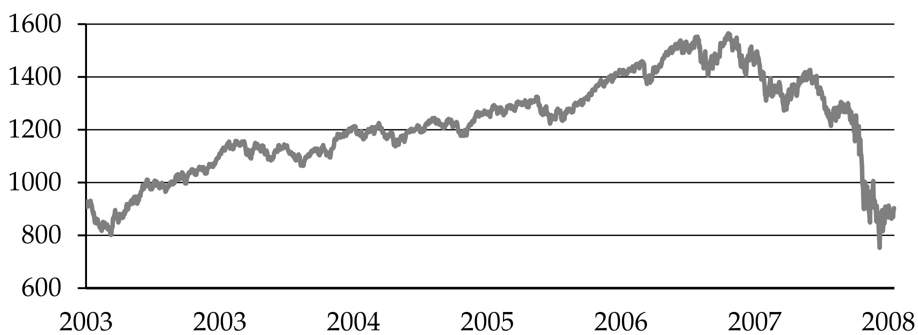
3. Description of Dataset

3.1. In-Sample Data

The first part of our empirical analysis is based on the S&P500 composite index for the period 2 January 2003 to 29 December 2006 (data obtained from Yahoo Finance). This in-sample period consists of 1,007 daily observations for the four-year trading period. The S&P500 is a value-weighted index of the stocks of 500 leading industries traded on the US stock exchanges based on their market capitalization (Standard and Poor's Online) and a leading indicator of the US's equity market. The sample period is chosen as such to avoid the effects of the "dot.com" crash of the early 2000s and to end just before the latest crisis, the study of which is our objective.

The second part of the analysis extends this sample period from January 2003 to December 2007, by which point the effects of the crisis had begun to show (note the downward trend toward the end of the plotted graph in Figure 1), increasing number of observations to 1,257 for the five-year period.

Figure 1: S&P500 index (2003–07)



Source: Author's calculations.

The daily stock index value (P_t) series is nonstationary; it follows an upward trend and no mean reversion (Figure 1). This is formally confirmed by using the Dickey-Fuller test to test for the presence of a unit root, which is an indicator of the nonstationarity of the series:

$$\Delta y_t = \delta y_{t-1} + u_t \tag{1}$$

y_t is the P_t series and u_t the error term. The null hypothesis proposes that there is a unit root, $\delta = 0$. Regressing the first difference of the P_t series on its own lag yields a p-value of 0.8105 and so the null hypothesis cannot be rejected at a 5-percent confidence level, confirming that the series is non-stationary. Given that the P_t series is nonstationary, we use daily stock returns for analysis:

$$R_t = \ln\,(P_t/P_{t-1}) \tag{2}$$

Converting the index value series to a returns series results in a graph that reverts to its long-run mean instead of following an upward trend. The mean of the series is close to 0, and imposing a normal distribution line on its histogram shows that the R_t series is characterized by thicker tails than normal (Figure A1 in the Appendix).

The distribution is negatively skewed and, furthermore, more peaked than the normal curve, indicating excess kurtosis (Table 1). The joint skewness/kurtosis test for normality yields a p-value of 0.00, allowing us to reject the null of normality. The series also demonstrates another characteristic common to financial time series—volatility clustering, i.e., periods of well-defined high and low volatility.

Table 1: Summary statistics for R_t series

	Observations	Mean	Standard deviation	Skewness	Kurtosis	Min.	Max.
R_t	1,007	0.0004	0.0078	–0.11	4.86	–0.036	0.035

Source: Author's calculations.

3.2. Out-of-Sample Data

If true volatility is, as discussed above, latent, then in order to evaluate out-of-sample forecasts, it is important to find a proxy for it. Anderson, Bollerslev, Diebold, and Christoffersen (2006) define the R_t series as comprising an expected conditional mean return term (μ_t) and another term (ε_t) that comprises the standard deviation and an idiosyncratic error term (z_t) such that

$$R_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t \tag{3}$$

The one-step-ahead volatility forecast can therefore be compared with squared returns:

$$R_t^2 = \sigma_{t|t-1}^2 z_t^2 \tag{4}$$

However the variance of z_t results in a great deal of noise when squared returns are used as the true underlying volatility. We therefore propose considering the R_t series as a continuous time process so that true volatility, referred to as integrated volatility (IV), is given by

$$IV(t) = \int_{t-1}^{t} \sigma^2(s) ds \tag{5}$$

If the returns series is sampled discretely and its variance taken for infinitesimally small periods, it will give an approximate measure of IV(t) (see Andersen, Bollerslev, Diebold, & Ebens, 2001; Andersen, Bollerslev, Diebold, & Labys, 2001, 2003). This approximate measure, RV, is not affected by the idiosyncratic error (z_t) as above, and is thus a far superior alternative that utilizes the additional information that intraday data has to offer. Hansen and Lunde's (2006) method is used to construct RV estimates as follows:

$$RV_t = \sum_{i=1}^{m} y_i^2 \tag{6}$$

y_i, i = 1,…, m are intraday returns, m being the number of returns in one trading day. The idea is that, with increased sampling frequency, this measure is a better approximation of true volatility ($t \to \infty$, $RV_t \to IV_t$). The important question that arises is the frequency at which the data should be sampled. At the highest frequencies, tick-by-tick returns violate the restrictions implied by the no-arbitrage assumptions in continuous-time asset pricing models. These same features also bias empirical RV measures constructed directly from ultra-high-frequency returns,[2] so in practice the measures are instead constructed from intraday returns sampled at an intermediate frequency (Anderson et al., 2005).

The S&P500 index is based on five-days-a-week trading starting at 0830 and ending at 1500. Market microstructure frictions can cause problems with very high-frequency data, so despite the availability of one-minute-interval data, we will use intermediate-frequency data sampled at five-minute intervals (our data source is TickData Online). This generates 78 observations for each day, which are then used to construct RV estimates for the two out-of-sample periods, each spanning six months beginning in January 2007 and January 2008, respectively.

4. Methodology

First, we model the volatility of the S&P500 index daily returns in order to predict their future values. Second, we compare the out-of-sample forecasting ability of the fitted models and select whichever model produces superior forecasts. The sample period is then extended up to December 2007, and the selected model re-estimated and used to produce crisis-period forecasts. These forecasts are then compared with the earlier

[2] This is due to market microstructure noise—bid-ask price spreads, jumps, and formation of patterns.

ones to establish the impact of the crisis period on their accuracy. We formally check for the presence of ARCH effects (conditional heteroscedasticity), using Engle's (1982) Lagrange Multiplier test.

The stylized facts concerning financial time series—persistence in volatility, mean-reverting behavior, and the asymmetric impact of negative- versus positive-return innovations—may significantly influence volatility. Among others, Engle and Patton (2001) illustrate these stylized facts and the GARCH models' ability—evaluated by their forecasting ability—to capture these characteristics. The sample employed in this study displays similar characteristics, thus the next logical step is to estimate the GARCH family of models:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{7}$$

This GARCH process does not differentiate between the impact of a positive and negative unexpected change in returns. It is therefore unable to capture the asymmetric effect of good or bad news on the volatility of the financial time series—a phenomenon termed the "leverage effect."

Anderson, Bollerslev, Diebold, and Ebens (2001) present two explanations for this so-called leverage effect. The first is that, when there is a negative shock, i.e., a negative return, it increases financial and operating leverage, which causes volatility to rise. The second is that, "if the market risk premium is an increasing function of volatility, large negative returns increase the future volatility by more than positive returns due to a volatility feedback effect." This means that the effect on volatility of unexpected bad news in the market would be higher than that of unexpected good news of the same magnitude. This renders the symmetry constraint imposed on the conditional variance equation in the GARCH process invalid. It is imperative to take account of this characteristic in order to make effective forecasts.

The presence of asymmetry in a financial time series necessitates the use of variants of the GARCH model that capture this phenomenon. Our review of the literature has established these models' goodness of fit as well as their forecasting ability. The mixed response concerning the superiority of a particular kind of model makes it difficult to allow a single choice and, so, we estimate the basic asymmetric GARCH (AGARCH) and the popular EGARCH and GJR-GARCH models.

Once the models have been fitted to the data, the next step is to generate forecasts using the estimated models. In order to select the best model, we evaluate the forecasts made by each based on the criteria detailed below. An alternative method would be to select the best model based on information criteria such as the Akaike or Bayesian, which would indicate the best in-sample fit. However, our aim is not to select the specification that best models the sample's volatility, but to select whichever makes the best forecasts out of sample. We therefore concentrate on the out-of-sample predictive capability of the models fitted and not on their in-sample fit.

4.1. *Out-of-Sample Forecast Evaluation*

For the purposes of forecast evaluation, we use each model to generate one-step-ahead forecasts of conditional volatility for the six-month period beginning in January 2007. Once each model has been estimated, in making each forecast we use the actual data available up to that point as an input into the equation estimated for conditional volatility by that model. The choice of the number of forecasts made is such that it ensures an adequate number of forecast observations for the analysis that is to be carried out. The quality of these forecasts is evaluated through the standard evaluation technique of employing loss-based functions and regression analysis. The chosen model is then re-estimated using the extended sample (now including 2007), and subsequently used to make predictions for the same horizon for the following year, i.e., one-step-ahead predictions for the six-month period beginning in January 2008 when the crisis hit the US equity market.

4.1.1. Regression-Based Evaluation

The first step in evaluating the quality of the forecasts made would be to regress the proxy for conditional volatility (RV_t) on the predicted volatility (h_t) from each model. This regression-based approach to evaluating out-of-sample forecasts—proposed by Mincer and Zarnowitz (1969)—has, however, been criticized in the literature. Pagan and Schwert (1990) note that, if the proxy RV_t contains large observations (outliers), problems arise when these regressions are run using ordinary least squares (OLS) because the OLS estimates are disproportionately affected by the larger values. Additionally, it "measures the level of variance errors rather than the more realistic proportional errors" thereby mainly assessing the performance of high values (Engle & Patton, 2001).

One solution to these two problems is to use the log of RV_t. Such log regressions are established as being less sensitive to the problems posed by larger observations. Thus, we run the following regression:

$$lnRV_t = \alpha + \beta * ln\hbar_t + u_t \tag{8}$$

If the forecasts (h_t) are perfect, the intercept (α) should equal 0 and the slope (β), 1. A model's superiority can be established by comparing the R^2 term—the higher the R^2 the better the forecasts explain the actual volatility

4.1.2. Loss Functions-Based Evaluation

An alternative to the regression analysis above is to assess how different the model's conditional variance predictions are from the proxy being used for the true variance. The simplest way of doing this is to calculate the mean forecast error (ME), which is

$$= \left(\frac{1}{m}\right)\sum_{t=1}^{m}(\widehat{y}_t - y_t) \tag{9}$$

The term m represents the number of forecasting observations, $\hat{y}_t$ is the predicted volatility, and y_t is the value of RV_t being used as a proxy for actual volatility. The lower the value, the better the forecast. Other, more sophisticated statistics that have been developed include a common forecast evaluation statistic, the mean squared error (MSE), which is

$$= \left(\frac{1}{m}\right)[\sum_{t=1}^{m}(\widehat{y}_t - y_t)^2] \tag{10}$$

The MSE squares the forecast errors ($\hat{y}_t + h - y_t + h$) and so penalizes larger errors more than smaller ones. Corradi and Awartani (2005) note that, since RV_t is measure-free and an unbiased estimator, it allows one to compare models in terms of loss functions other than quadratic. Thus, we can make use of the mean absolute percentage error (MAPE), which is

$$= (\frac{1}{m})[\sum_{t=1}^{m}\left|\frac{\widehat{y}_t - y_t}{y_t}\right|] \tag{11}$$

Unlike the MSE, the mean absolute error (MAE) does not penalize larger forecast errors more heavily than smaller ones. However, since it takes the absolute value, it does not allow the effect of under- and over-predictions of the same magnitude but carrying opposite signs to be cancelled out.

4.1.3. Diebold-Mariano Test

Diebold and Mariano's (DM) (1995) test allows one to compare the forecasting ability of two models. For one-step-ahead forecasts, let the forecast error ($\hat{y}_t + h - y_t + h$) be denoted by $g(e)$. The difference in loss in period i from using model 1 versus model 2 is defined as $d_i = g(e1_i - g(e2_i)$. The mean loss is given by

$$\bar{d} = 1/H\sum_{i=1}^{H}\left[g\left(e_{1i}\right) - g(e_{2i})\right] \tag{12}$$

H is the number of forecast errors. The DM statistic is asymptotically standard normal when applied to non-nested forecasts, so that the t-test can be used to test the null hypothesis that any two fitted models have equal predictive abilities, which is when $\bar{d} = 0$.

$$DM = \bar{d} / \sqrt{\mathrm{var}(\bar{d})} \tag{13}$$

This test is used to determine if there is any statistical difference between the forecasts generated by the chosen model before and during the crisis. Applying the test requires $var(\bar{d})$. If the d_i series is uncorrelated, $var(\bar{d})$ is given by $\gamma_0/(H - 1)$, else Enders' (2004) specification is followed where $var(\bar{d}) = (\gamma_0 + 2\gamma_1 + \ldots + 2\gamma_q)/H - 1$, and γ_i denotes the ith autocovariance of d_i where the first q values of γ_i are significant.

5. Empirical Analysis

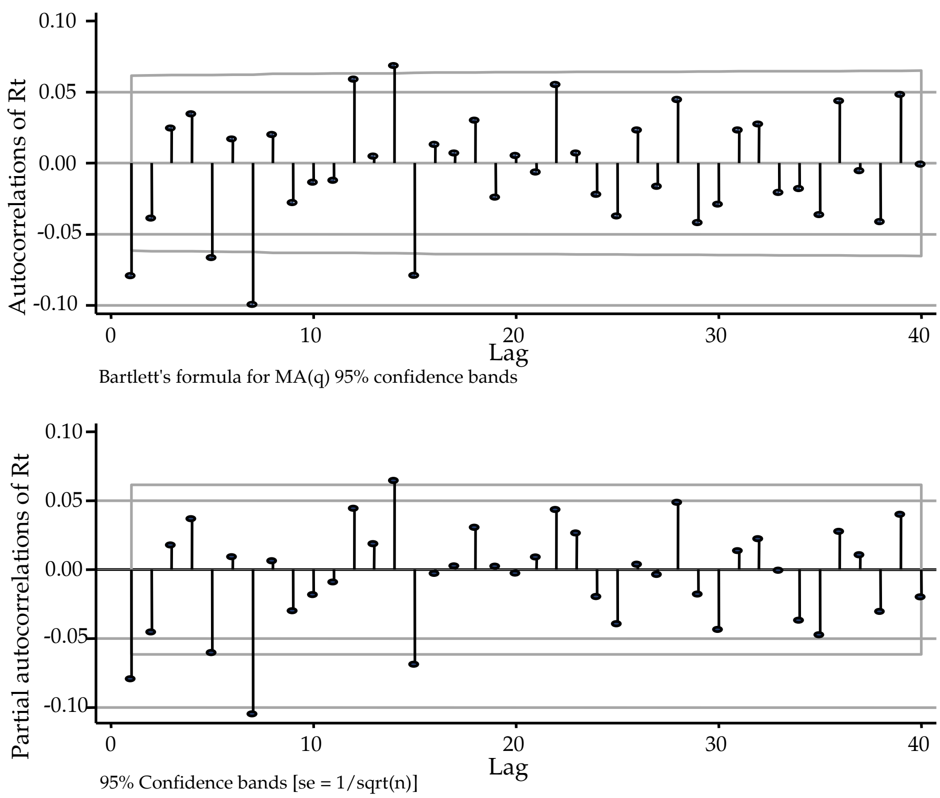
5.1. Modeling the Conditional Mean

The first step in our empirical analysis is to estimate the mean equation of returns. In order to identify the best-fitting model, the autocorrelogram (AC) and partial autocorrelogram (PAC) are plotted (Figure 2) to determine which lags are statistically significant at 5 percent ($\pm 1.96/\sqrt{T}$). Lags 1, 5, 7, 14, and 15 are found to be significant (Table A1 in the Appendix). Thus, the autoregressive moving average (ARMA) (15,15) is estimated using OLS.

Next, the estimated model's standardized residuals[3] are verified using the Portmanteau/Box-Pierce/Ljung-Box test to evaluate the adequacy of the fitted model. The null hypothesis for this test is that the errors are white noise. The p-value for the Q-statistic is 0.7126, thus the null hypothesis cannot be rejected at a 5-percent level of significance. The mean equation, therefore, fits the data.

Since it is important to check for second-order dependence in the residuals (conditional heteroscedasticity), we use Engle's (1982) Lagrange Multiplier test to check for serial correlation. This entails running OLS regression of the squares of the residuals on its own lags. The null hypothesis proposes that there is no ARCH effect, i.e., that the coefficients should all be jointly 0. The p-value yielded is 0.00, thus we can confidently reject the null hypothesis and conclude that there are, in fact, ARCH disturbances in the returns series.

Figure 2: Autocorrelogram and partial autocorrelogram of R_t series



Source: Author's calculations.

[3] Residuals divided by standard deviation.

A further check is to apply the Portmanteau test to the squared standardized residuals. The null hypothesis proposes that the errors are not serially correlated, i.e., that there are no ARCH effects. As above, a p-value of 0.000 allows us to confidently reject the null hypothesis and confirm the presence of ARCH disturbances in the data. Both tests prove that the variance is conditional on the past period, implying that we need to fit a model that can account for this effect.

5.2. *Modeling the Conditional Volatility*

The GARCH family of models adequately takes into account the presence of conditional heteroscedasticity, and these models are estimated in order to model volatility. Table 2 reports estimates of the maximum likelihood estimator parameters.

Table 2: Maximum likelihood estimation results

	GARCH (1,1)	EGARCH (2,1)	AGARCH (1,1)	GJR-GARCH (1,1)
α_0	7.74e-07 (4.39e-07)	–0.0760587 (0.0370346)*	8.48e-07 (3.19e-07)*	5.03e-07 (2.52e-07)*
β_1	0.9376958 (0.0179654)*	0.9922645 (0.0037748)*	0.9487258 (0.0148996)*	0.9562212 (0.0139044)*
α_1	0.0470485 (0.0124951)*	–0.239826 (0.0430449)*	0.0356606 (0.0113298)*	0.0690447 (0.0150748)*
γ_1	-	–0.2274885 (0.0781236)*	–0.0005019 (0.0001111)*	0.0718498 (0.0165285)*
α_2	-	0.1741572 (0.0421315)*	-	-
γ_2	-	0.2901534 (0.079057)*	-	-

Note: * = statistically significant at 1 percent. Standard errors are given in parentheses.
Source: Author's calculations.

We begin estimating the model using the most parsimonious specification, the GARCH (1,1). If the model is a good fit, it should be able to capture the serial correlation and no ARCH effects should remain. The p-value of the Q-statistic for the squared standardized residuals is 0.1284. Thus, we cannot reject the null hypothesis, implying that they are no remaining ARCH effects, and that the model of variance has been adequately fitted. A p-value of 0.3934 yielded by the Lagrange Multiplier test further confirms this since it allows us to accept the null hypothesis of no remaining ARCH effects. The sum $\beta_1 + \alpha_1$ equals 0.96, which is less than 1 and satisfies the condition for stationarity. The results show that the coefficient of the lag of conditional variance, β_1 (0.94), is quite high, indicating the persistence of past effects.

As noted earlier, financial time series are characterized by the presence of leverage effects. Testing for this phenomenon entails regressing squared standardized residuals on the lags of standardized residuals, resulting in a p-value of 0.0007. This allows us to confidently reject the null hypothesis and conclude that the coefficients are not jointly equal to 0. It thus confirms the presence of leverage effects in the data. A further test involves the use of a dummy to signal any negative shocks that may have occurred in the previous period. When the squared standardized residuals are regressed on the dummy, its coefficient turns out to be significant (a p-value of 0.020). This is conclusive proof that negative shocks do, in fact, increase the conditional variance, as reported in the existing literature (see Corradi & Awartani, 2005; Taylor, 2004).

The next three models to be fitted formally account for this asymmetric effect in addition to the phenomena of volatility clustering and excess kurtosis. The first is the GJR-GARCH (1,1), which employs an indicator function that emerges when there is a negative shock in the past to account for the asymmetries. Table 2 shows that the coefficient of the indicator function γ_1 is significant and positive, implying that there are asymmetric effects. The p-value of the Q-statistic for the squared standardized residuals of this model is 0.9445. This signals that the null hypothesis cannot be rejected at a 5-percent level of significance. The GJR-GARCH (1,1) thus adequately models the second-order moment of the series.

The EGARCH (1,1) model is fitted next (Table A2 in the Appendix). The γ_1 coefficient appears to be insignificant, and the p-value yielded by the test that is applied to the model's squared standardized residuals is 0.000, implying that the null hypothesis can be rejected. Hence, the residuals are not white noise, and the model is not deemed an adequate fit. When a higher-order specification, the EGARCH (2,1), is fitted (Table 2), however, all the coefficients emerge as significant. The Portmanteau test confirms that the residuals are white noise (the p-value is 0.34). The parameter β_1 is equal to 0.99, i.e., less than 1, thus satisfying the condition for the process being stationary.

The AGARCH (1,1) model, which is fitted next, modifies the term that captures shocks that have occurred in previous periods. The estimates yielded confirm the presence of leverage effects since γ_1 is significant and negative. The standardized squared residuals are checked to ensure that no autocorrelation remains, and thus no additional lags are required. The p-value is 0.54, which does not allow us to reject the null hypothesis at a 5-percent level of significance, implying that a higher-order specification is not needed. All the coefficients are statistically significant with a β_1 that is close to 1, indicating persistent volatility.

5.3. *Forecast Evaluation*

Using five-minute-interval intraday return data, we construct RV_t estimates for six months that will act as a proxy for true (unobserved) volatility. This entails making one-step-ahead forecasts using each of the four models in order to evaluate the out-of-sample forecasts. The first stage of the evaluation process employs the loss function-based evaluation technique, with RV_t acting as a benchmark (Table 3). The lower the value of the criteria estimated using a particular model's forecasts, the better the forecasts. In the second stage, we carry out a regression analysis to verify the results obtained in the first stage.

Table 3: Loss function values for one-step-ahead predictions

Criterion	GARCH	EGARCH	AGARCH	GJR-GARCH
ME	–9.44887E-06	–2.106E-05	–1.15222E-05	–3.61821E-06
MSE	4.93968E-09	5.71367E-09	5.10746E-09	4.81257E-09
MAPE	0.868947666	0.626980967	0.795054713	0.983587439

Source: Author's calculations.

The value of the ME and MSE criteria is lowest in the case of the GJR-GARCH (1,1) model, with the GARCH (1,1) a close second in both. However, when the MAPE is taken into account, the EGARCH (2,1) emerges as the superior model. Thus, all three criteria indicate that the asymmetric models are superior. These results are in accordance with Corradi and Awartani (2005) who found that the asymmetric models dominated the GARCH (1,1) specification in making one-step-ahead forecasts.

After running log regressions of RV_t on the forecasted variance series[4] (see results in the Appendix), we check to see if the coefficients are statistically close to 1. The p-values of the GARCH (1,1) and AGARCH (1,1) coefficients are 0.02 and 0.00, respectively, which allows the null hypothesis to be rejected at a 5-percent level of significance, and implies that the coefficients are different from 1. However, the p-value of the GJR-GARCH (1,1) coefficient is 0.31 and that of the EGARCH (2,1) is 0.3559, which does not allow us to reject the null hypothesis, and suggests that these coefficients are not statistically different from 1. This is in synch with the results obtained from the loss function criteria, which also showed that the estimates of these two models are superior to those of the other two.

The adjusted R^2 values (Table 4) of the regression of log RV_t on the log of the forecasts show that the R^2 of the AGARCH (1,1) model is the lowest. The model's estimates appear to perform poorly on all criteria, confirming its poor predictive ability. The R^2 of both the GARCH (1,1) and GJR-GARCH (1,1) models is close to 42 percent, with the EGARCH (2,1) at 36 percent. This higher R^2 further supports the superiority of the GJR-GARCH (1,1) model over the EGARCH (2,1), which has a better MAPE measure. We can therefore proceed with the GJR-GARCH (1,1) model as the model with the best forecasting ability.

Table 4: Adjusted R^2 from regressions of log RV_t on log of predicted values

	AGARCH	GARCH	GJR-GARCH	EGARCH
R^2	0.3062	0.4255	0.4119	0.3603

Source: Author's calculations.

The next step is to extend the in-sample period from December 2006 to December 2007, and estimate the chosen model, the GJR-GRCH (1,1), based on this sample (Table A2 in the Appendix). The effect of positive news on conditional volatility is given by $\alpha_1 + \gamma_1$, the previous value of which was 0.0028 and is now 0.0057. All coefficients are still statistically significant. It is important to check if the model is an adequate fit. The p-value of the Q-statistic is 0.8276, verifying that the squared standardized residuals are white noise and that the model, therefore, fits the data.

[4] The volatility forecast series is tested for the presence of a unit root. If the series is nonstationary, the regression is spurious and yields meaningless coefficients. The null of nonstationarity is rejected at a 5-percent level of significance.

The selected model is then used to generate one-step-ahead forecasts for six months, which are compared with the out-of-sample forecasts that were made for the pre-crisis period. As before, RV_t is used as a benchmark against which to evaluate the forecasts. By January 2008, the impact of the subprime mortgage crisis had begun to show in the equity market, and the S&P 500 index had started to decline. As expected, Table 5 shows that the values of all three evaluation criteria for all horizons have increased relative to 2007.

Table 5: Loss function values for 2007 and 2008 forecasts from GJR-GARCH model

Year	ME	MSE	MAPE
2007	–3.61821E-06	4.81257E-09	0.983587
2008	–1.9694E-05	4.23284E-08	1.324632

Source: Author's calculations.

This signals worsened forecasting, which is not surprising given that crisis periods are more volatile than usual, which is why predicting becomes more difficult. However, a regression analysis of these forecasts yields an R^2 that has increased to 43 percent, lending support to the model's forecasting ability during the turbulent period.

The DM test is used to determine if there is any statistically significant difference between the model's forecasting ability in the two periods. The DM statistic yielded is 0.875, which is less than 1.96, implying that we cannot reject the null hypothesis at a 5-percent level of significance. Thus, there is no statistical difference between the model's forecasting ability in terms of period type—it is equally capable of predicting volatility for a crisis period and a normal period. The increased volatility that characterizes a crisis period is adequately accounted for. Moreover, the model fitted takes special account of leverage effects and thus effectively handles the downturn in index returns.

6. Conclusion

Based on an out-of-sample evaluation of the forecasts made by the four GARCH models, the model that best estimates daily returns volatility is the GJR-GARCH (1,1) model—in accordance with the findings of Corradi and Awartani (2005) and Taylor (2004)—when applied to the in-sample period from January 2003 to December 2008, which is characterized by relative tranquility.

Having selected the best model based on several evaluation criteria, the in-sample period is then extended up to December 2007 and the model is re-estimated. The one-step-ahead forecasts for six months obtained from this re-estimated model are compared to the prior forecasts obtained. This meets our second aim—to assess the model's ability to cope with the pronounced volatility characterizing the recent crisis that hit the US equity market. We have found that, while the model's predictive ability decreases, there is no substantial change. This supports the ability of the GJR-GARCH model in particular and of the asymmetric GARCH family of models in general to remain relatively robust across periods of pronounced volatility.

While we have used high-frequency data to construct an RV measure as a proxy for unobserved true volatility, to the presence of market microstructure noise meant that intraday returns were not aggregated at greater-than-five-minute intervals. This proxy could thus be further refined by increasing the sampling frequency and by explicitly accounting for the jumps and patterns that arise during the day when intraday data is used.

Other avenues of research could make use of the implied volatility that is extracted from options written on the index for further insight into how volatility forecasts are affected during crisis periods. Additionally, the stochastic volatility model, which has been found to be more flexible than ARCH-class models and to "fit financial market returns better and have residuals closer to standard normal" (Poon & Granger, 2003), has not been estimated here due to computational difficulties. Further research could use this model for detailed analyses analysis based on several modeling techniques.

References

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of Business*, *62*, 55–80.

Andersen, T. G., & Bollerslev, T. (1997). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*(4), 885–905.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2006). Volatility and correlation forecasting. In G. Elliot, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting* (pp. 778–787). Amsterdam: North-Holland.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of stock return volatility. *Journal of Financial Economics*, *61*, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*, 579–625.

Andersen, T. G., Bollerslev, T., & Meddahi, N. (2005). Correcting the errors volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica*, *73*, 279–296.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Brailsford, T. J., & Faff, R. W. (1996). An evaluation of volatility forecasting techniques. *Journal of Banking and Finance*, *20*, 419–438.

Corradi, V., & Awartani, B. M. A. (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: The role of asymmetries. *International Journal of Forecasting*, *21*, 167–183.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–265.

Enders, W. (2004). *Applied econometrics time series* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, *48*(5), 1749–1778.

Engle, R. F., & Patton, A. J. (2001). What good is a volatility model? *Journal of Banking and Finance*, *20*, 419–438.

Fair, R. C., & Shiller, R. J. (1990). Comparing information in forecasts from econometric models. *American Economic Review*, *80*(3), 375–380.

Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions and Instruments*, *6*(1), 1–88.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *48*(5), 1779–1801.

Hansen, P. R., & Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics*, *131*(1–2), 97–121.

Kim, D., & Kon, S. J. (1994). Alternative models for the conditional heteroscedasticity of stock returns. *International Journal of Forecasting*, *67*(4), 563–598.

McMillan, D. G., & Speight, A. E. H. (2004). Daily volatility forecasts: Reassessing the performance of GARCH models. *Journal of Forecasting*, *23*, 449–460.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts and expectations. In J. A. Mincer (Ed.), *Economic forecasts and expectations*. New York: National Bureau of Economic Research.

Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: A new approach. *Econometrica*, *59*(2), 347–370.

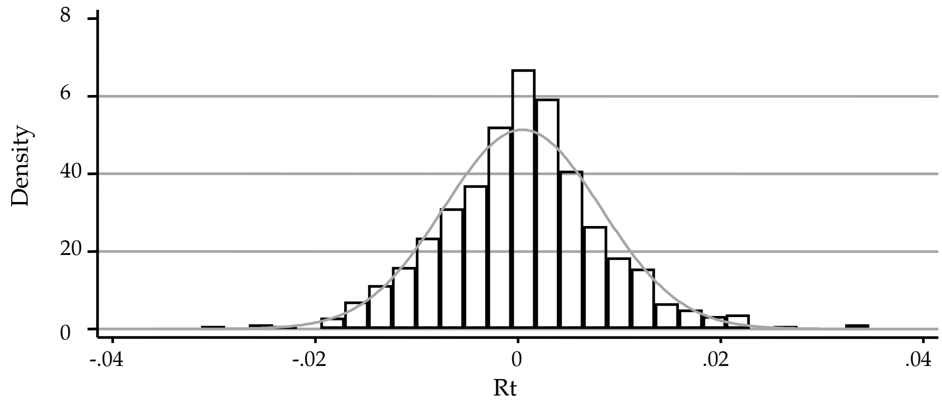
Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional volatility. *Journal of Econometrics*, *45*, 267–290.

Poon, S.-H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478–539.

Taylor, J. W. (2004). Threshold volatility forecasting with smooth transition exponential smoothing. *International Journal of Forecasting*, *20*, 237–286.

Appendix

Figure A1: Histogram of R_t series with normal curve imposed



Source: Author's calculations.

Table A1: AC and PAC values for R_t series

LAG	AC	PAC
1	–0.0801	–0.0802
2	–0.0395	–0.0462
3	0.0241	0.0174
4	0.0349	0.0371
5	–0.0662	–0.0598
6	0.0171	0.0171
7	–0.0999	–0.1054
8	0.0200	0.0062
9	–0.0289	–0.0312
10	–0.0395	–0.0462
11	0.0241	0.0174
12	0.0600	0.0455
13	0.0059	0.0202
14	0.0683	0.0644
15	–0.0793	–0.0689
16	0.0129	–0.0034

Source: Author's calculations.

Table A2: Estimation results

Parameter	EGARCH (1,1)	GJR-GARCH (1,1)
α_0	–18.48361	1.05e-06
	(0.000)*	(0.000)*
β_1	0.9376958	0.9439015
	(0.000)*	(0.000)*
α_1	–0.0330587	0.0808454
	(0.122)	(0.000)*
γ_1	–0.0181544	–0.0865688
	(0.455)	(0.000)*

Note: * = significant at 1 percent; p-values are given in parentheses.
Source: Author’s calculations.

The Lahore Journal of Business
1:1 (Summer 2012): pp. 59–77

Job Satisfaction and Women's Turnover Intentions in Pakistan's Public Universities

Aliya Bushra*

Abstract

The aim of this study is to test the impact of women's job satisfaction on their turnover intentions, specifically for those employed in the education sector. Using a sample drawn from two different universities in Lahore, Pakistan, we measure their levels of job satisfaction by evaluating their general working conditions, pay and potential for promotion, professional relationships, use of skills and abilities, and activities assigned. We find that flexible working hours, workplace location, performance appraisal, and skills utilization have a highly positive significance on turnover intentions, while professional autonomy, job security, and promotion have an inverse impact on job satisfaction and turnover intentions.

Keywords: Turnover intentions, job characteristics, job independence, job involvement.

JEL classification: M10, M12, M19

1. Introduction

Job satisfaction, or a feeling of contentment with one's present job, lowers the chances of one's quitting that job. This degree of satisfaction is determined by various factors, such as pay scale, employer's attitude, skill variety, motivation, job security, working environment, task identity, and feedback. O'Reilly (1989) points out that these expectations on the part of an employee produce organizational norms that shape employees' behavior in that organization. High levels of motivation lead to better performance and eventually make the organization more effective (Tietjen & Myers, 1998).

An employee's daily work routine, under favorable working conditions, leads to job satisfaction; many studies have empirically tested this idea. Katz (1978) examines the relationship between job satisfaction and five task characteristics: (i) skill variety, (ii) task identity, (iii) task

---

* The author is a teaching fellow at the Lahore School of Economics, and can be contacted on a_aylz@hotmail.com.

significance, (iv) autonomy, and (v) feedback on performance. The author's results show that each of these characteristics is directly related to job satisfaction.

Our aim is, specifically, to determine the relationship between job satisfaction and turnover intentions among women in public sector universities/colleges in Pakistan. The rest of this article is organized as follows: Section 2 provides a review of the literature, Section 3 explains the methodology used, Section 4 analyzes the data collected, Section 5 presents the results of our statistical analysis, and Section 6 concludes the study.

2. A Review of the Literature

Griffin, Patterson, and West (2001) explore the ways in which overall job satisfaction in teams is influenced by changes in leadership roles. Teamwork involves leadership and employees' experience of supervisory support and encouragement. A team leader's role, therefore, produces outcomes on the part of his or her employees, and is shown to be a strong determinant of job satisfaction. The relationship between employees and their supervisors also plays an important role in retaining the former. Gilstrap (2009) emphasizes the importance of a leader's role in motivating his or her employees with rewards and monetary benefits.

Many other factors lead to work satisfaction, including bonuses, fringe benefits, job enrichment, clarity of roles, and met expectations. Howard (1966) points out the importance of fringe benefits in retaining employees. These benefits can include paid leisure leave, paid sick leave, and health insurance, all of which help develop an employee's bond with the organization.

Khandwalla and Jain (1984) conclude that employee morale, customer loyalty, and customer goodwill can contribute strongly to job satisfaction among lower management. Frameworks developed by Goffee and Jones (1996) and Alvesson (2002) show that motivation exists where employees' values are aligned to the values of their employer organization. Miller and Wheeler (1992) find that gender differences in job satisfaction disappear if employees are satisfied with their jobs. Bordia, Hobman, Jones, Gallois, and Callan (2004) point out that uncertainty among employees can cause emotional stress, lack of motivation, and lower concentration, which in turn leads to poorer job performance. It is, therefore, very important for an employee to feel mentally and physically satisfied with his or her job.

Some employees may be subject to stress if they perceive unfairness on the part of their employer. This can lead to job dissatisfaction and, in turn, cause employees to search for alternative jobs. For example, employees needing to take maternal leave may decide to quit their jobs if their employer does not provide support. Odom, Boxx, and Dunn (1990) investigate the relationship between organizational culture and three elements of employee behavior: (i) commitment, (ii) group work cohesion, and (iii) job satisfaction, all of which are found to be significant with respect to turnover intentions.

Certain aspects of organizational politics can be detrimental to job satisfaction, such as mistrust between employees and their employers, and between coworkers—this results in high turnover in most organizations. Ilgen and Favero (1985) emphasize the usefulness of performance appraisal, which an essential factor for motivating employees and, in turn, leading to low turnover.

It is necessary that an organization retain its employees because it affects the performance of both. In many organizations, an employee's performance is measured by the number of years that he or she has worked. Employees may leave organizations for reasons such as perceived unfair treatment when promotions are not merit-based, or if they are required to carry out work that is against their values. Such instances create frustration and result in employees leaving their jobs.

Mitchell, Holtom, Lee, and Graske (2001) argue that employees who leave their organization create a problem for other employees and for the organization itself. The latter has to bear the extra cost of hiring and training new workers from scratch, while other employees also begin to feel discontent with their jobs.

Organizations may undergo different changes to be successful. Frahm and Brown (2007) examine employees' rate of adaptability to such changes. Lund (2003) shows that there is a positive relationship between organizational culture and job satisfaction. Thus, both organizational culture and communication can have a significant impact on an employee's satisfaction level.

Organizations such as educational institutions need to provide their employees with equal, merit-based promotion opportunities. Price and Mueller (1981) find that organizations that provide such opportunities and define employees' working hours face lower turnover intentions.

While there is a body of literature on job satisfaction and turnover intentions—as reviewed above—there is still need for further research because no significant theory has yet clearly identified the job satisfaction factors that lead to voluntary turnover among women. Our aim is to discuss the determinants of job satisfaction and assess how they can reduce turnover intentions. The independent variables used are (i) work satisfaction and (ii) job characteristics. Specifically, we focus on women employees' performance and satisfaction in relation to changes in a university/college setting.

Women employees who feel professionally enriched and are given opportunities to advance in their institution are more likely to be retained by it. Age, wage level, tenure, and recognition of accomplishments also play an important part in contributing to job satisfaction. Thus, organizations that fail to implement the appropriate policies to retain their employees may eventually lose their assets.

3. Methodology

3.1. Objectives

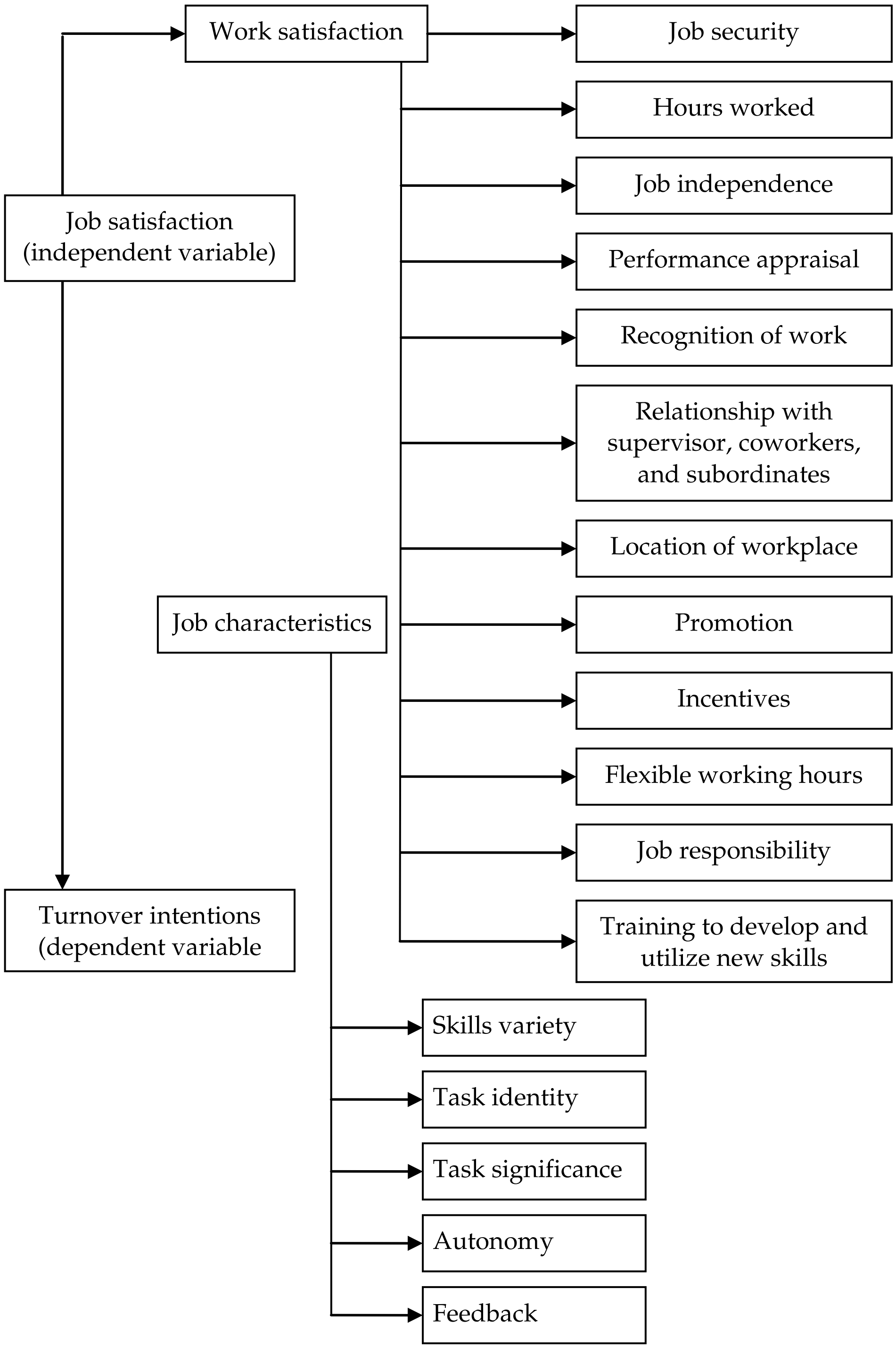
The study's objectives are to determine the relationship between job satisfaction and turnover intentions—among women in public sector universities/colleges—in terms of the following questions:

1. What is the relationship between the factors that affect job satisfaction (working hours, location of workplace, paid leave, incentives, promotion, performance appraisal, recognition of work or task significance, relationship with colleagues (supervisors, coworkers, and subordinates), autonomy, responsibilities (task identity/job involvement), skill variety, job security, and turnover intentions)?

2. What is the relationship between the linear combination of these factors (as identified above)?

Figure 1 illustrates our theoretical framework of job satisfaction, which is further divided into (i) work satisfaction characteristics, and (ii) job characteristics, with respect to turnover intentions.

Figure 1: Theoretical framework of job satisfaction



3.2. *Data Sample and Variables*

Our sample population for the study was drawn from among the faculty of two women's colleges in Lahore, Pakistan—Kinnaird College and Lahore College. Of a total of 750 faculty members, we selected 100 respondents, including lecturers, assistant/associate professors, and heads of departments. Fifty questionnaires each were distributed between the two colleges, to collect data on the variables shown in Table 1.

Table 1: Categories of study variables

Work satisfaction characteristics: Level of satisfaction with	**Job characteristics**
Salary and bonuses	Skills variety (new skills)
Promotion	Task significance (skill utilization)
Job security	Task identity (recognition)
Location of workplace	Autonomy
Working conditions: Flexible hours	Feedback (performance appraisal)
Working conditions: Paid leave	
Working conditions: Hours worked	
Utilization of skills	
Training and learning new skills	
Supervisor	
Coworkers	
Subordinates	
Recognition of work	
Job responsibility	
Performance appraisal	

3.3. *Instrumentation*

We have used a structured questionnaire comprising a combination of instruments to collect the data required. All the instruments used were developed specifically to test the impact of job satisfaction on turnover intentions, and their reliability has been confirmed by a number of studies (see Kanungo, 1982; Lin, 1999; Hung & Tsai, 2008). Table 2 presents the four-item scale used by Kelloway, Gottlieb, and Barham (1999) to measure turnover intentions as a dependent variable.

Table 2: Turnover intentions as dependent variable

Questions	Four-item scale
I am thinking about leaving this organization	Measured by Likert five-point scale from 1 = strongly disagree to 5 = strongly agree
I am planning to look for a new job	As above
I intend to ask others about new job opportunities	As above
I do not plan to stay with this organization much longer	As above

The independent and dependent variables were combined into one comprehensive questionnaire for participants to complete. The questionnaire is based on the Likert five-point scale, which measures both the high and low dimensions of all the variables (1 = strongly disagree to 5 = strongly agree). It consists of two sections: (i) turnover intentions (dependent variable), and (ii) job satisfaction (independent variable), which is further subdivided into (a) work satisfaction characteristics, and (b) job characteristics. Table 3 describes the independent variables used.

Table 3: Definitions of independent variables

Independent variable: Level of satisfaction with	Constitutive definition	Operational definition
Work satisfaction characteristics	Minnesota Satisfaction Questionnaire (Leung, 1996)	12-item scale
Incentives (salary and bonuses)	Amount of financial remuneration received	Measured by Likert five-point scale where 1 = strongly disagree and 5 = strongly agree
Promotion	Satisfaction with career advancement	As above
Job security	Expectations of job continuity	As above
Training and learning new skills	Satisfaction with on-the-job and other training and learning new skills	As above
Location of workplace	Satisfaction with location of university relative to place of residence	As above
Working conditions: Flexible hours	Satisfaction with flexibility of working hours	As above

Independent variable: Level of satisfaction with	**Constitutive definition**	**Operational definition**
Working conditions: Paid leave	Satisfaction with paid leave	As above
Working conditions: Hours worked	Satisfaction with average number of hours worked per day	As above
Utilization of skills	Satisfaction with skill utilization in present job	As above
Supervisor	Relationship with supervisor has an impact on job satisfaction	As above
Coworkers	Relationship with coworkers has an impact on job satisfaction	As above
Subordinates	Relationship with subordinates has an impact on job satisfaction	As above
Recognition of work	Satisfaction with appreciation shown and rewards for good work	As above
Job responsibility	Satisfaction with tasks performed and accomplishment of tasks	As above
Performance appraisal	Satisfaction with own evaluation	As above
Job characteristics	Job Diagnostic Survey (Hackman & Oldham, 1975)	Five-point scale
Skills variety	Degree to which job involves different activities, or use of different skills/talents in carrying out the work	Mean of items 4, 6, and 8 in Section 3, Part A, of questionnaire Measured by five-point scale where 1 = strongly disagree and 5 = strongly agree
Task significance	Degree to which job has substantial impact on the lives or work of others	Mean of items 2, 10, and 15 in Section 3, Part A, of questionnaire Measured by five-point as above

Independent variable: Level of satisfaction with	Constitutive definition	Operational definition
Task identity	Degree to which job requires completion of a whole and identifiable piece of work with visible results	Mean of items 3, 7, and 13 in Section 3, Part A, of questionnaire Measured by five-point as above
Autonomy	Degree to which job provides freedom, independence, and self-discretion in scheduling work and determining which procedures to use	Mean of items 1, 11, and 14 in Section 3, Part A, of questionnaire Measured by five-point as above
Feedback	Degree to which employee receives clear information about his/her performance from supervisor/coworkers	Mean of items 5, 9, and 12 in Section 3, Part A, of questionnaire

4. Data Analysis

Using EViews and SPSS software, we apply the t-test to determine the difference between the mean of the independent variable and that of the dependent variable. A series of ANOVA tests is run to verify the study's hypothesis, where the alpha term represents a 0.01 (extremely significant) and 0.05 (highly significant) level of significance. We carry out multiple linear regressions to compute the significance of the factors affecting job satisfaction and turnover intentions. In the following analysis, the coefficient of determination, R^2 (adjusted R^2), explains data variations caused by turnover intentions. The most significant independent variable is easily identified as the p-value gives the same results as above when compared with the significance level.

Model 1 is written as

$$TI_t = \beta_0 + \beta_{1t}LWP + \beta_{2t}PV_t + \beta_{3t}FLXT_t + \beta_{4t}ISB_t + \beta_{5t} + PRM_t + \beta_{6t}PMAP_t + \beta_{7t}JS_t + \beta_{8t}RW_t + \beta_{9t}RSWSCS_t + \beta_{10t}NJT_t + \beta_{11t}SU_t + \beta_{12t}JI_t + \beta_{13t}JR_t + \beta_{14t}HW_t + \varepsilon \tag{1}$$

TI_t is turnover intention at time t; *LWP* is workplace location; *PV* is paid leave; *FLXT* is flexible working hours; *ISB* is incentives such as salary and bonuses; *PRM* is promotion; *PMAP* is performance appraisal; *JS* is job security; *RW* is recognition of work; *RSWSCS* is employees' relationship with their supervisors, coworkers, and subordinates; *NJT* is on-the-job training; *SU* is skill utilization; *JI* is job independence; *JR* is job responsibility; and *HW* is the number of hours worked at time *t*. β_0 and ε represent the constant and error term, respectively.

Model 2 is written as

$$TI_t = \beta_0 + \beta_{1t}LWP + \beta_{2t}FLXT_t + \beta_{3t}ISB_t + \beta_{4t} + PRM_t + \beta_{5t}PMAP_t + \beta_{6t}JS_t + \beta_{7t}RW_t + \beta_{8t}RSWSCS_t + \beta_{9t}NJT_t + \beta_{10t}SU_t + \beta_{11t}JI_t + \varepsilon \quad (2)$$

TI_t is turnover intention at time *t*; *LWP* is workplace location; *FLXT* is flexible working hours; *ISB* is incentives such as salary and bonuses; *PRM* is promotion; *PMAP* is performance appraisal; *JS* is job security; *RW* is recognition of work; *RSWSCS* is employees' relationship with their supervisors, coworkers, and subordinates; *NJT* is on-the-job training; *SU* is skill utilization; and *JI* is job independence at time *t*. β_0 and ε represent the constant and error term, respectively.

Model 3 is written as

$$TI_t = \beta_0 + \beta_{1t}LWP + \beta_{2t}FLXT_t + \beta_{3t}PRM_t + \beta_{4t}PMAP_t + \beta_{5t}JS_t + \beta_{6t}NJT_t + \beta_{7t}SU_t + \beta_{8t}JI_t + \varepsilon \quad (3)$$

TI_t is turnover intention at time *t*; *LWP* is workplace location; *FLXT* is flexible working hours; *PRM* is promotion; *PMAP* is performance appraisal; *JS* is job security; *NJT* is on-the-job training; *SU* is skill utilization; and *JI* is job independence at time *t*. β_0 and ε represent the constant and error term, respectively.

5. Results of Analysis

The results tabulated in Tables 4, 5, and 6 correspond to Models 1, 2, and 3, respectively, and are somewhat consistent with the findings reported in the existing literature. In line with Ilgen and Favero's (1985) findings, our results show that the factors affecting job satisfaction—number of hours worked, location of workplace, paid leave, financial incentives, promotion, performance appraisal, recognition of work, relationship with supervisors/coworkers/subordinates, job independence, level of responsibility, and job security—have a positive and significant relationship with turnover intentions. We also find that flexible working hours have a positive and significant impact on turnover intentions, which is consistent with Griffin et al. (2005).

As Table 4 shows, a faculty member who feels that her job is secure will have increased job satisfaction. While job responsibility is not found to be significant, job security is highly significant. Women faculty members feel more job-secure when they are promoted, given positive feedback, or if their institution is relatively near their residence.

Table 4: Regression results for Model 1

	β_0	β_t	$t(\beta)$	Adj. R^2
β_0	0.162		0.798	
LWP	0.155	0.216	0.016**	
PV	–0.041	–0.04	0.664	
FLXT	0.233	0.301	0.001***	
ISB	0.009	0.009	0.922	
PRM	–0.209	–0.209	0.184	
PMAP	0.508	0.643	0.000***	0.392
JS	–0.194	–0.247	0.076	
RW	–0.126	–0.149	0.372	
RSWSCS	–0.033	–0.027	0.758	
NJT	0.139	0.156	0.233	
SU	0.453	0.408	0.000***	
JI	–0.314	–0.316	0.008***	
JR	–0.099	–0.079	0.406	
HW	0.131	0.179	0.067	

Note: ** = significant at 95 percent, *** = significant at 99 percent.

ANOVA results for Model 1

	Sum of squares	df	Mean square	F	Significance
Regression	32.119	14	2.294	5.565	0.000
Residual	35.041	85	0.412		
Total	67.160	99			

Source: Author's calculations.

Job satisfaction has a negative and insignificant relationship with the paid-leave variable, which is not consistent with studies such as that of Katz (1978). This may be because faculty members at the sample institutions—Kinnaird College and Lahore College—face a low probability of having applications for paid leave approved; even when such applications are approved, faculty members may still be required to attend official meetings.

The results also show that workplace location is an important consideration for women faculty members who are married and/or have domestic responsibilities. Osterioh and Frey (2000) similarly suggest that there is positive and extremely significant relationship between workplace location and job performance and motivation.

While Roberson (1990) finds that the time spent at a workplace is positively correlated with satisfaction, our results give a different picture with regard to women faculty members. We find that they prefer shorter working hours that allow them to meet any domestic/family-related responsibilities. Additionally, Hung and Tsai's (2008) findings on the relationship between turnover intentions and supervision are also inconsistent with the results yielded by our sample, where the relationship between supervisors/coworkers/subordinates appears to be insignificant (Tables 4, 5, and 6).

The variations explained by Model 1—39.2 percent as the adjusted R^2— show that the job satisfaction variables are well explained by their dependent variable, i.e., turnover intentions. Skill utilization, performance appraisal, and flexible working hours have an extremely positive significance for turnover intentions. This result is consistent with Price and Mueller (1981) who find that positive feedback motivates employees to adapt to change and to learn new skills.

We have already noted that women faculty members who need to manage their careers along with domestic/family-related responsibilities prefer flexible working hours. Most of the women in our sample are married, implying that, for them, flexible hours result in higher job satisfaction and, consequently, lower turnover intentions. Moreover, skill utilization has a positive significance for the dependent variable since the more an employee is able to utilize her skills, the more she is likely to be satisfied with her job. This will eventually better her job performance, a result that is supported by Katz (1978), Ilgen and Favero (1985), and Alvesson (2002).

Table 5: Regression results for Model 2

	β_0	β_t	$t(\beta)$	Adj. R^2
β_0	0.186		0.731	
LWP	0.142	0.198	0.025**	
FLXT	0.258	0.333	0.000***	
ISB	0.018	0.016	0.845	
PRM	–0.319	–0.318	0.020**	
PMAP	0.528	0.669	0.000***	0.415
JS	–0.237	–0.302	0.017**	
RW	–0.126	–0.062	0.365	
RSWSCS	–0.039	–0.031	0.714	
NJT	0.157	0.175	0.115	
SU	0.409	0.369	0.000***	
JI	–0.351	–0.353	0.002***	

Note: ** = significant at 95 percent, *** = significant at 99 percent.

ANOVA results for Model 2

	Sum of squares	df	Mean square	F	Significance
Regression	31.417	8	3.491	8.790	0.000
Residual	35.743	91	0.397		
Total	67.160	99			

Source: Author's calculations.

As Tables 5 and 6 show, the variable *JI*, job independence, is extremely significant but has an inverse relationship with turnover intentions. This result is different from that of studies such as Khandwalla and Jain (1984), who find that task significance and job independence decrease turnover intentions. According to our results, job independence increases job satisfaction but does not decrease turnover intentions since it has a negative slope.

This relationship could be explained by the fact that the women faculty members in our sample are required to follow a given course structure and not given the independence to include/exclude topics of their interest. This repetitive structure may increase boredom and eventually de-motivate an employee, giving her reason to leave her job. As Model 3 (Table 6) shows, an R^2 value of 41.5 percent indicates that turnover intentions are well explained by the independent variables we have used.

Table 6: Regression results for Model 3

	β_0	β_t	$t(\beta)$	Adj. R^2
β_0	0.211		0.687	
LWP	0.144	0.201	0.021**	
FLXT	0.259	0.333	0.000***	
PRM	–0.321	–0.320	0.018**	
PMAP	0.532	0.673	0.000***	0.415
JS	–0.237	–0.301	0.017**	
NJT	0.157	0.175	0.114	
SU	0.408	0.367	0.000***	
JI	–0.349	–0.351	0.002***	

Note: ** = significant at 95 percent, *** = significant at 99 percent.

ANOVA results for Model 3

	Sum of squares	df	Mean square	F	Significance
Regression	31.042	8	3.880	9.776	0.000
Residual	36.118	91	0.397		
Total	67.160	99			

Source: Author's calculations.

Fisher (2000) concludes that employees who are allowed to work independently are more productive, leading to greater job satisfaction. Our findings, however, are different because the employees in our sample are not able to work independently and are required to coordinate their work with other faculty members who teach the same courses. When their institution's administration gives this authority to its faculty members, they are likely to relax their productivity, while absenteeism increases. This, in turn, eventually decreases employees' productivity to the point that the rate of turnover rises.

While Lindbeck and Snower (1988) greatly stress the significance of paid leave and job satisfaction for turnover intentions, our study finds that paid leave and turnover intentions have an insignificant relationship since the former is not often approved. Further, the number of hours worked, job responsibility, recognition of work, and salary and bonus incentives do not appear to have a significant impact on employee retention.

The promotions variable and positive feedback variable are both shown to have a significant impact on turnover intentions. The relationship is an inverse one, indicating that the more women employees are given chances for promotion, the more they are satisfied with their current jobs, and the less likely are they to leave. Job security also has a highly significant negative impact on turnover intentions, and helps retain employees in that women faculty members who are assured job security report lower turnover intentions. Gilstrap (2009) and Mitchell et al. (2001) give similar results but using a different sample size. The expected signs of these variables are, therefore, consistent with the findings of the existing literature.

6. Conclusion

The study's aim was to assess the impact of different factors of job satisfaction on women employees' turnover intentions in public sector universities in Pakistan. We have found that flexible working hours, workplace location, performance appraisal, and skill utilization help determine these employees' level of job satisfaction and, in turn, their turnover intentions. This can be explained by the argument that many women employees have to manage domestic responsibilities along with their careers, which the teaching profession allows them to do. The factors mentioned above increase employees' satisfaction and give them incentive to retain their posts at the employing university.

While most faculty members possess a level of competence that matches their ability to work independently, our study has shown that there can also be an inverse relationship between the two. Most women faculty members in the sample institutions are not authorized to frame new syllabi or teach new topics, but instead are required to comply with specific course outlines, all of which decrease their satisfaction level. The lack of independence to include or exclude topics of their interest eventually de-motivates such employees and leads them to leave the institution.

Skill utilization has a highly significant impact on turnover intentions since it gives employees a better chance to enhance their selves. Performance appraisal is also extremely significant as a factor of job satisfaction. University administrations must focus on these variables to reduce turnover among their women faculty members, and to avoid the consequent loss of experience, knowledge, and motivation among other faculty members.

References

Alvesson, M. (2002). *Understanding organizational culture*. London: Sage Publications.

Bordia, P., Hobman, E., Jones, E., Gallois, C., & Callan, V. (2004). Uncertainty during organizational change: Types, consequences, and management strategies. *Journal of Business and Psychology*, *18*(4), 507–532.

Byrne, Z. S. (2005). Fairness reduces the negative effects of organizational politics on turnover intentions, citizenship behavior, and job performance. *Journal of Business and Psychology*, *20*(2), 175–200.

Cansu, (2006). Components of job satisfaction on the organization communication. Journal of Occupational Health Psychology, 6, 377–396.

Cropanzano, R., Howes, J. C., Grandey, A. A., & Toth, P. (1997). The relationship of organizational politics and support to work behaviors, attitudes, and stress. *Journal of Organizational Behavior*, *18*, 159–180.

Denison, D. R., & Mishra, A. K. (1995). Toward a theory of organizational culture and effectiveness. *Organization Science*, *6*(2), 204–223.

Diefendorff, J. M., Brown, D. J., Kamin, A. M., & Lord, R. G. (2002). Examining the roles of job involvement and work centrality in predicting organizational citizenship behaviors and job performance. *Journal of Organizational Behavior*, *23*(1), 93–108.

Dormann, C., & Zapf, D. (2001). Job satisfaction: A meta-analysis of stabilities. *Journal of Organizational Behavior*, *22*, 483–504.

Dougherty, T. W., Turban, D. B., Olson, D. E., Dwyer, P. D., & Lapreze, M. W. (1996). Factors affecting perceptions of workplace sexual harassment. *Journal of Organizational Behavior*, *17*, 489–501.

Fey, C. F., & Denison, D. R. (2003). Organizational culture and effectiveness: Can American theory be applied in Russia? *Organization Science*, *14*(6), 686–706.

Fisher, C. D. (2000). Mood and emotions while working: Missing pieces of job satisfaction? *Journal of Organizational Behavior*, *21*, 185–202.

Frahm, J., & Brown, K. (2007). First steps: Linking change communication to change receptivity. *Journal of Organizational Change Management*, *20*(3), 370–387.

Gandhi, P. (1992). Impact of job enrichment on work and organizational identification. *Indian Journal of Industrial Relations*, *27*(3), 289–298.

Gandz, J., & Murray, V. (1980). The experience of workplace politics. *Academy of Management Journal*, *23*, 237–251.

Gilstrap, D. L. (2009). A complex systems framework for leadership and organizational dynamics in academic libraries. *Portal: Libraries and the Academy*, *9*(1), 57–77.

Glance, N. S., Hogg, T., & Huberman, B. A. (1997). Training and turnover in the evolution of organizations. *Organization Science*, *8*(1), 84–96.

Goffee, R., & Jones, G. (1996). What holds the modern company together? *Harvard Business Review*, November/December, 133–148.

Griffin, M. A., Patterson, M. G., & West, M. A. (2001). Job satisfaction and teamwork: The role of supervisor support. *Journal of Organizational Behavior*, *22*, 537–550.

Hackman, J. R., & Oldham, G. R. (1975). Development of the job diagnostic survey. *Journal of Applied Psychology*, *60*, 159–170.

Harris, K. J., Andrews, M., & Kacmar, K. M. (2007). The moderating effects of justice on the relationship between organizational politics and workplace attitudes. *Journal of Business and Psychology*, *22*(2), 135–144.

Howard, A. W. (1966). Salaries are not enough: Fringe benefits also count. *The Clearing House*, 40, 525–528.

Hung, T.-K., & Tsai, I.-J. (2008, July). *The effect of confidant relationship on turnover intention and moderated by employee's job involvement*. Paper presented at the International Conference on Business and Information, Seoul, Korea.

Huselid, M. (2005). Impact of human resource management practices on turnover, productivity and corporate financial performance. *Academy of Management Journal*, *38*, 635–672.

Ilgen, D. R., & Favero, J. L. (1985). Limits in generalization from psychological research to performance appraisal processes. *Academy of Management Review*, *10*(2), 311–321.

Kanungo, R. N. (1982). Measurement of job and work involvement. *Journal of Applied Psychology*, *67*(3), 341–349.

Katz, R. (1978). Job longevity as a situational factor in job satisfaction. *Administrative Science Quarterly*, *23*, 204-223.

Kelloway, E. K., Gottlieb, B. H., & Barham, L. (1999). The source, nature, and direction of work and family conflict: A longitudinal investigation. *Journal of Occupational Health Psychology*, *4*(4), 337–346.

Khandwalla, P. N., & Jain, G. R. (1984). Organizational goals and lower management job satisfaction. *Indian Journal of Industrial Relations*, *20*(2), 111–136.

Kotter, J. P., & Heskett, J. L. (1992). *Corporate culture and performance* (12th ed.). New York: Free Press.

Lin, H.-M. (1999). The construction of the Chin-Shin relationship inventory. *Journal of Education and Psychology*, *22*, 323–354.

Lindbeck, A., & Snower, D. J. (1988). Job security, work incentives and unemployment. *Scandinavian Journal of Economics*, *90*(4), 453–474.

Lund, D. (2003). Organizational culture and job satisfaction. *Journal of Business and Industrial Marketing*, *18*(3), 219–236.

Miller, K. B., Rutherford, A. M., & Kolodinsky, W. R. (2008). Perceptions of organizational politics: A meta-analysis of outcomes. *Journal of Business Psychology*, *22*, 209–222.

Miller, J. G., & Wheeler, K. G. (1992). Unraveling the mysteries of gender differences in intentions to leave the organization. *Journal of Organizational Behavior*, *13*(5), 465–478.

Mitchell, T. R, Holtom, B. C., Lee, T. W., & Graske, T. (2001). How to keep your best employees: Developing an effective retention policy. *Academy of Management Executive*, *15*(4), 96–109.

Nanda, R., & Browne, J. J. (1977). Hours of work, job satisfaction and productivity. *Public Productivity Review*, *2*(3), 46–56.

O'Reilly, C. A. (1989). Corporations, culture, and commitment motivation and social control in organizations. *Management California Review*, *31*(4), 9–25.

Odom, R. Y., Boxx, W. R., & Dunn, M. (1990). Organizational cultures, commitment, satisfaction and cohesion. *Public Productivity and Management Review*, *14*(2), 157–169.

Osterioh, M., & Frey, B. S. (2000). Motivation, knowledge transfer and organizational forms. *Organization Science*, *11*(5), 538-–550.

Price, J. L., & Mueller, C. W. (1981). A causal model of turnover for nurses. *Academy of Management Journal*, *24*(3), 543–565.

Roberson, L. (1990). Prediction of job satisfaction from characteristics of personal work goals. *Journal of Organizational Behavior*, *11*(1), 29–41.

Saari, L. M., & Judge, T. A. (2004). Employee attitudes and job satisfaction. *Human Resource Management*, *43*(4), 395–407.

Sheridan, J. E. (1992). Organizational culture and employee retention. *Academy of Management Journal*, *35*(5), 1036–1056.

Tietjen, M. A., & Myers, R. M. (1998). Motivation and job satisfaction. *Management Decision*, *36*(4), 226–231.

Volkwein, J. F., & Zhou, Y. (2003). Testing a model of administrative job satisfaction. *Research in Higher Education*, *44*(2), 149–171.

Yousef, D. A. (1998). Satisfaction with job security as a predictor of organizational commitment and job performance in a multicultural environment. *International Journal of Manpower*, *19*(3), 184–194.

The Lahore Journal of Business
1:1 (Summer 2012): pp. 79–108

Modeling and Forecasting the Volatility of Oil Futures Using the ARCH Family Models

Tareena Musaddiq*

Abstract

This study attempts to model and forecast the volatility of light, sweet, crude oil futures trading at the NYMEX during 1998–2009, using various models from the ARCH family. The results reveal that the GJR-GARCH (1,2) model is best suited to forecast purposes. The fitted models also suggest the presence of asymmetric effects in the data. The study also reveals that trading volume and open interest do not reduce the persistence of volatility for these oil futures.

Keywords: Modeling volatility, forecasting, oil futures.

Classification: C53, C58, G17.

1. Introduction

Extensive research has been conducted on modeling volatility clustering in financial markets, using different econometric techniques. Volatility modeling is a key area of interest to researchers because it plays an important role in managing risk, pricing derivatives, hedging, selecting portfolios, and policymaking. "Investors and portfolio managers have certain levels of risk which they can bear. A good forecast of the volatility of asset prices over the investment holding period is a good starting point for assessing investment risk" (Poon & Granger, 2003). Accurate volatility forecasts are thus very important and, over time, have motivated new approaches to volatility modeling to help forecast future volatility for asset pricing and risk management purposes.

Although most research on volatility modeling has focused on equity markets (see, for instance, Bollerslev, Chou, & Kroner, 1992; Pagan & Schwert, 1990) and foreign exchange markets, the success of a particular type of forecasting model applied to one type of market cannot be generalized across other markets (see Sadorsky, 2006). There has been relatively less research on futures markets, and only in recent years has volatility modeling for these markets gained popularity. With futures

* The author is a teaching fellow at the Lahore University of Management Sciences in Pakistan.

gaining increasing importance in terms of assessing and managing risk, it has become important to work with the relevant models to forecast volatility. Futures are especially seen to display particular features that differentiate them from other financial tools, in that they are high-risk volatile investment tools in which a small price movement can have a huge impact on trading (see Carvalho, da Costa, & Lopes, 2006).

Futures markets merit further attention, especially given their growing use for hedging purposes. Investors and economic agents who trade in physical spot commodities may wish to hedge the price risk of commodities, and this is one of the primary reasons for the development of commodity futures. Although futures markets exist for all sorts of commodities—including metals, agricultural goods, and animal products—their most actively traded commodity is crude oil followed by its various derivatives such as heating oil and gas, etc. This is not surprising since oil and its derivatives are important factors of production in the world's economies, and oil price fluctuations can significantly affect their performance.

In the past, the spot prices of crude oil have been affected both by economic and geopolitical events. Examples include the price falls in 1998 that occurred due to a slowdown in Asian economic growth, and the price rise caused by OPEC's curtailed oil supply in 2000/01 and by US military action in Iraq in 2003 and after (see Kang, Kang, & Yoon, 2009). Ample research shows that oil price volatility has significant macroeconomic effects. While Ferderer (1996) and Lee, Ni, and Rutti (1995) investigate these macroeconomic impacts without stock market variables, Sadorsky (1999) uses vector autoregression (VAR) to show that oil price volatility has a significant impact on stock price volatility as well. Hence, the study of oil price volatility is important because it impacts macroeconomic variables such as aggregate output and employment both in countries and financial markets worldwide.

With developments in financial markets and the increased use of hedging techniques to manage risk, there has been tremendous growth in the use of both over-the-counter and exchange-traded derivatives to manage risk related to the volatile energy sector. Oil futures are one such example, the trading of which began in 1978 on the New York Mercantile Exchange (NYMEX). The light, sweet, crude oil futures contract traded on the NYMEX is used as a key international pricing benchmark due to its liquidity and price transparency.

With crude oil being the world's most actively traded commodity, its futures on the NYMEX provide the world's most liquid forum for crude oil trading and account for the largest futures contract trading on a physical commodity in terms of volume. Owing to the importance both of oil and emerging futures markets, and given that futures markets are relatively less well researched than others, modeling and forecasting the volatility of these futures can prove a worthwhile exercise.

The structure of this article is as follows. Section 2 provides a brief overview of the existing literature, Section 3 describes the data used in this study, Section 4 details the methodology, Section 5 presents empirical results, and Section 6 concludes the study.

2. A Review of the Literature

Engle's (1982) classic study was the first to distinguish between unconditional and conditional variance, and introduced a technique for simultaneously modeling both the mean and variance of an economic or financial time series, using the autoregressive conditional heteroscedastic (ARCH) model. Subsequently, Bollerslev (1986) introduced a parsimonious representation of Engel's model, followed by a number of studies that proposed and tested variants of the ARCH model, while accounting for asymmetries and persistence. Poon and Granger (2003) noted that, at the time of their study, at least 93 published and working papers had studied the forecasting performance of volatility models while many others had not incorporated the forecasting aspect. They also pointed out that models that allowed for asymmetric effects were able to provide better forecasts owing to the negative relationship between volatility and shocks. Other variants of the ARCH model included the multivariate generalized ARCH (GARCH) approach (see Brooks & Persand, 2003).

While these studies focus on equity and foreign exchange markets, others have attempted to model the volatility of commodities in the futures market. Bracker and Smith (1999) study the volatility of the copper futures market, and conclude through its root mean squared error (RMSE) that both the GARCH and exponential GARCH (EGARCH) models best fit the market's volatility, followed by the Glosten-Jagannathan-Runkle (GJR) model.[1] Carvalho et al. (2006) devise a systematic modeling strategy for futures markets in general and apply it to soya beans futures. Of eight different ARCH family models, they find

[1] The choice of model depends on the series of prices used; here, the authors have used daily data for a 10-year period up to 1999.

evidence of asymmetric effects when using the EGARCH as their selected model (according to the mean squared prediction error (MSPE) and mean absolute prediction error (MAPE) criteria). They claim their methodology to be independent of the type of market and applicable to all commodities in the futures market. Application to wheat and corn futures reveals that the quadratic GARCH (QGARCH) and threshold GARCH (TGARCH) specifications, respectively, are best suited. Brooks (1998) models the volatility of Kuala Lumpur crude palm oil futures and, apart from daily and monthly effects, finds significant evidence of the impact of open interest and volume when using GARCH models to estimate volatility.

Some studies have focused on the volatility of crude oil futures traded on the NYMEX. Sadorsky (2006) uses data up to 2003, choosing a TGARCH model for heating oil and gas and a GARCH model for crude oil and unleaded gasoline. The study not only shows that the GARCH family models outperform random walk, historical mean, and exponential smoothing models, but that single-equation GARCH models better model volatility than VAR and bivariate GARCH models. More recently, Kang et al. (2009) study volatility modeling for three crude oil markets—Brent, Dubai, and West Texas Intermediate (WTI). Using data up to 2006, they conclude that the component GARCH (CGARCH) or fractionally integrated GARCH models made better forecasts for the three series than a simple GARCH model.

Agnolucci (2009) uses data up to 2005 to compare the predictive ability of GARCH and implied volatility models for oil futures traded on the NYMEX, and concludes that former seem to perform better than the latter (which are obtained by inverting the Black-Scholes equation). For forecasting purposes, however, Agnolucci suggests that, in contrast to Sadorsky (2006), the CGARCH model performs better than the GARCH model. The difference in their two conclusions could be attributed to the different time frame and forecast evaluation techniques used.

Other studies have included the effect of trading volume and open interest in the GARCH processes as a proxy for the arrival of information. Clark (1973) first introduced the mixture-of-distributions hypothesis (MDH), which explored the role of trading volume in stock price movements. Lamoureux and Lastrapes (1990) use the daily returns and volumes of 20 actively traded stocks in the US market to test the relation between conditional variance and trading volume by deriving a GARCH effect. They find that volatility persistence disappears when daily trading volume is added to the conditional variance equation. Brailsford (1996) uses the GARCH process to investigate the effect of trading volume on the persistence of volatility in the Australian stock market, finding that it significantly reduces persistence. Using data on 10 actively traded US stocks, Gallo and Pacini (2000) put forward similar findings, as do Pyuna, Lee, and Nam (2000) in the case of the Korean Stock Exchange.

In contrast, some studies have found that trading volume has little, if any, effect on the persistence of market volatility. Sharma, Mougous, and Kamath (1996) use data on the New York Stock Exchange index, and argue that that trading volume does not completely explain the GARCH effect for the market index and that volatility persistence did not diminish on adding volume (see also Brooks, 1998). Darrat, Rahman, and Zhong (2003) use the EGARCH model to test Dow Jones industrial average (DJIA) stocks, and find significant contemporaneous correlations between trading volume and volatility in only three of 30 DJIA stocks.

Researchers such as Najand and Yung (1991) recommend adding a lagged volume variable, and testing to ensure avoidance of any specification bias. Bessembinder and Seguin (1993) extend this line of research and analyze the roles of open interest and volume in determining volatility in eight futures markets; they conclude that both variables significantly impact volatility. Foster (1995) examines the volume-volatility relationship for crude oil futures trading on the NYMEX, arguing that volume does not remove the GARCH effect and that previous volatility better explains volatility—this implies that volume does not represent the rate of information arrival for oil futures.

The present study attempts to model the volatility of returns on light, sweet, crude oil futures traded on the NYMEX by employing the ARCH model and its variants and extensions. We present both in- and out-of-sample forecasts of volatility, and use techniques based on past research to assess which model best forecasts volatility. We also attempt to update previous research by using data up to July 2009. The asymmetric modeling of these futures has not been studied in such detail, and thus adds to the existing body of research by assessing dynamic forecasts. We also add volume and open interest as independent variables to ascertain if trading activity has a significant effect on volatility.

3. Data

3.1. Sources of Data

The data required is the returns on the light, sweet crude oil futures traded on the NYMEX. In order to calculate these returns, we use data on the daily price of these futures, obtained from the Bloomberg database (the use of daily data for this study is in line with the literature discussed earlier). Data pertaining to the volume and open interest of the futures contract is also obtained from the Bloomberg database. The price taken as the daily price is the last trading price of the day (see http://www.nymex.com/CL_spec.aspx for details of contract). Data on futures prices spans the period from 23 June 1998 to 16 July 2009—a total of 2,780 observations. Of these, we use the data ranging from 23 June 1998 to 23 February 2009 for modeling purposes, i.e., a total of 2,680 observations, which is sufficient for modeling daily returns. The remaining 100 observations are treated as an out-of-sample period in order to assess the forecasts made.[2]

3.2. Description and Testing

The plotted autocorrelation and partial autocorrelation of the price of the futures contract indicate that the series is nonstationary (Figures A1a and A1b in the Appendix). Applying the Dickey-Fuller test to the series confirms this (Table A1 in the Appendix), and suggests that it cannot be used to model volatility.

Returns rather than prices are more appropriately used here, first, because our aim is to model the volatility of returns on oil futures, and second, because a returns series is more likely to be stationary and thus more suitable for modeling than a price series. The returns are calculated by applying the first difference of the log of prices. Table 1 summarizes the statistics on returns, showing that oil futures have an average daily return of 0.05503 percent and a standard deviation of 0.02616, which indicates an average annualized volatility of 41.53 percent. The skewness coefficient is –0.2203, its sign being common to most financial time series. The kurtosis value is higher than 3, implying that the returns distribution has fat tails. The ARCH family of models should, therefore, be used to account for these characteristics of the data.

[2] We have used STATA software to model all specifications by maximum likelihood, and assume the underlying distribution to be normal.

Table 1: Summary statistics for returns

	Mean	Standard deviation	Variance	Skewness	Kurtosis
Returns	0.0005503	0.0261624	0.0006845	–0.2203233	6.805797

Source: Author's calculations.

It is imperative when modeling such a series that it be stationary and the data mean-reverting. For this purpose, the Dickey-Fuller test is applied to the returns series (Table A2 in the Appendix), and the results show that the series is stationary. On application, the Phillips-Perron test also indicates that the series is stationary and can be used for modeling purposes (Table A3 in the Appendix).

The plotted autocorrelation and partial autocorrelation of squared returns indicate dependence and, hence, imply time-varying volatility (Figures A2a and A2b in the Appendix). This is further supported by the q test for squared returns, which also suggests that is the series is time-dependent.

4. Methodology

In order to model the volatility of the returns, we need to determine their mean equation. The return for today will depend on returns in previous periods (autoregressive component) and the surprise terms in previous periods (moving order component). Plotting the autocorrelation and partial autocorrelation of the returns series can help determine the order of the mean equation.

Like most financial time series, the returns series exhibits what is referred to as "volatility clustering" (Figures A3a and A3b in the Appendix), i.e., it exhibits alternating periods of relative tranquility and unusually large volatility. In order to model such patterns of behavior, the variance of the error term is allowed to depend on its history. The classic model of such behavior is the ARCH model introduced by Engle (1982), which simultaneously models the mean and variance of a series.

For this purpose, if we assume y_t to be the returns series and I_{t-1} to be the information set available, then

$$y_t = E[y_t | I_{t-1}] + \varepsilon_t$$

E [.] is the expectations operator and represents the predictable part of the returns, while ε_t is the unpredictable part and is given by

$$\varepsilon_t = y_t - \beta x_t$$

x_t is the set of explanatory variables and β is the set of parameters from a linear regression in vector form. The expected value of ε_t is 0, and its values are serially uncorrelated. Engle (1982) argued that, if we assume that for a time series the forecast of today's value based on past information is simply $E\ (y_t \mid y_{t-1})$, then the forecast for y_t depends on the value of the conditioning variable y_{t-1} and the variance of this one-period forecast is given by $(y_t \mid I_{t-1},\ y_{t-1})$. Engel therefore proposed a model under these assumptions in which the variance did depend on past information unlike the conventional models present of the time. The ARCH model simultaneously models the conditional mean and variance of a time series with the conditional heteroscedasticity of the unpredictable part of the series modeled as

$$\varepsilon_t = z_t \sqrt{h_t}$$

h_t is a nonnegative function and z_t is an i.i.d stochastic process with a zero mean and unit variance. From this, it follows that the conditional mean of ε_t is 0 and its variance is h_t, implying that ε_t is a heteroscedastic process. Given this information about ε_t, it can easily be seen that the mean of a series y_t, conditional on the past information set is $E\ (y_t \mid I_{t-1})$ and the variance is h_t. Engel proposed the following specification for the process h_t:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_q \varepsilon_{t-q}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2$$

$\alpha_i' \ s \geq 0$ and $i = 1..., q$ are constant parameters. This is the so-called ARCH (q) model. As the primary model introduced for modeling volatility, this will be the first model on which we fit our returns series. However, the ARCH model often needs a higher-order q to capture the volatility of a financial time series and, hence, requires estimating many parameters. As Bollerslev (1986) points out, "In empirical applications of the ARCH model a relatively long lag in the conditional variance equation is often called for, and to avoid problems with negative variance parameter estimates a fixed lag structure is typically imposed." His solution to this problem was the generalized ARCH (GARCH) model, which allows both a more flexible lag structure and a longer memory relative to ARCH specification. As opposed to the ARCH model, the GARCH model's specification also includes lagged conditional variances. In general, a GARCH (*p*, *q*) model is given by

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i}$$

α and β are the parameters to be estimated, p is the number of lags for past variances, and q is the number of lags for past squared residuals. The GARCH model thus allows both autoregressive and moving-average components in heteroscedastic variance. It gives a more parsimonious representation of the ARCH model and is much easier to identify and estimate. The GARCH model is, therefore, the second model that will be fitted to the data.[3]

Realistically speaking, if "bad news" has a more pronounced effect on volatility than "good news" of the same magnitude, then a symmetric specification such as ARCH or GARCH is not appropriate since in standard ARCH/GARCH models the conditional variance h_t is unaffected by the sign of the past periods' errors (it depends only on squared errors). Various extensions have therefore been proposed to capture these asymmetric effects often shown by financial time series.

Before applying the asymmetric models, however, one needs to test for the presence of such effects. Engle and Ng (1993) propose various tests to detect the presence of asymmetric effects, which are run on the standardized residuals of the GARCH model. The sign bias test is given by

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 S_{t-1}^- + error$$

S_{t-1}^- = 1 when $\hat{\varepsilon}_{t-1}^2 < 0$, and 0 otherwise. If the dummy variable's coefficient is significant and positive, this suggests the presence of asymmetric effects. The negative sign bias test determines whether the size of the negative shock also affects the impact it has on conditional variance, and is given by

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 S_{t-1}^- + \hat{\varepsilon}_{t-1} + error$$

[3] We also estimate the GARCH in the mean form of the GARCH model, which allows the ARCH component in the specification of the mean equation.

For the existence of a size effect, the coefficient must be negative and significant. The positive sign bias test determines if the size of a positive shock impacts its conditional variance, and is given by

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 S_{t-1}^{+} + \hat{\varepsilon}_{t-1} + error$$

S_{t-1}^{+} S_{t-1}^{-}. For the size effect to be present, the coefficient must be negative and significant. If the tests above indicate the presence of asymmetric effects, then the ARCH/GARCH models are no longer deemed appropriate and their other variants need to be considered.

The first model to account for such effects was the EGARCH model proposed by Nelson (1991). It uses a logarithmic function to treat asymmetric effects, and is given by

$$Ln(h_t) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \left(\left| \frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}} \right| - \sqrt{\frac{2}{\pi}} \right) - \sum_{i=1}^{q} \gamma_i \left(\frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}} \right) + \sum_{i=1}^{p} \beta_i \ln(h_{t-i})$$

In this case, the logarithmic function ensures that the conditional variance is positive and, therefore, the parameters can be allowed to take negative values. The specification implies that the impact of past errors is exponential, unlike standard GARCH models that imply that the effect is quadratic. If the shock is positive, its effect on the log variance is $\alpha_1 + \gamma$ while the effect is $\alpha_1 - \gamma$ if the shock is negative. For significant asymmetric effects, therefore, the coefficient γ should take a negative sign.

Unlike the EGARCH model, Glosten, Jagannathan, and Runkle's (1993) eponymous model does not look at exponential values but assumes that the impact of squared residuals on the variance depends on whether the residual term is negative or positive. For this purpose, it employs an indicator function as follows:

$$h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i S_{t-i}^{+} \varepsilon_{t-i}^2$$

The indicator function S_{t-i}^{+} takes a value of 1 if $\varepsilon_{t-i} > 0$, and 0 otherwise. For the effect of the previous period's bad news to be greater than the effect of good news of the same magnitude, γ should be significant and have a negative sign.

Zakoïan's (1994) threshold ARCH (TARCH) model is given by

$$h_t^{1/2} = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i}^{1/2} + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^{+} + \gamma_i \left| \varepsilon_{t-i}^{-} \right|$$

$\varepsilon^+ = \max(\varepsilon, 0)$ and $\varepsilon^- = \min(\varepsilon, 0)$. The effect of good and bad news is captured separately through the two coefficients, α and γ, respectively. Unlike the GJR model where the indicator function becomes 0 in the case of a negative shock, the TARCH model adds a separate variable for negative shocks. Another variant of the TGARCH model mentioned by Zakoïan (1994) and applied by Taylor (1986) and Schwert (1989) takes into account the effect of the shock's size on volatility, and is given by

$$h_t^{1/2} = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i}^{1/2} + \sum_{i=1}^{q} \alpha_i \left| \varepsilon_{t-i} \right|$$

Our aim is to determine how well these different models perform in terms of forecasting volatility and will be assessed based on the forecasts they make. The forecasting approach used is such that the last 100 observations of the sample are used to assess out-of-sample forecasts. We will make dynamic forecasts for these models, i.e., estimate the models using the first 2,680 observations and make one-step-ahead predictions for the variance of these observations in a static manner by employing the original value of the variance right up to the point of prediction. For the next 100 observations, we will make dynamic predictions, i.e., the predicted value of the variance will be used recursively to make subsequent observations. As an example, the dynamic forecasts for the GARCH (1,1) model would be

$$\hat{h}_{t+1/t} = E(\varepsilon_{t+1}^2 | I_t) = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 h_t$$

This predicted value of the variance is used to predict the variance of subsequent observations as

$$\hat{h}_{t+2/t} = E(\varepsilon_{t+2}^2 | I_t) = \alpha_0 + \alpha_1 E(\varepsilon_{t+1}^2 | I_t) + \beta_1 \hat{h}_{t+1/t} = \alpha_0 + (\alpha_1 + \beta_1) \hat{h}_{t+1/t}$$

$$\hat{h}_{t+3/t} = E(\varepsilon_{t+3}^2 | I_t) = \alpha_0 + \alpha_1 E(\varepsilon_{t+2}^2 | I_t) + \beta_1 \hat{h}_{t+2/t}$$

$$= \alpha_0 + (\alpha_1 + \beta_1)(\alpha_0 + (\alpha_1 + \beta_1) \hat{h}_{t+1/t})$$

Likewise, we use the other models' respective equations to obtain volatility forecasts. Once the forecasts have been made, the next step is to evaluate them. For comparison purposes, we compare out-of-sample forecasts with historical volatility. Volatility is itself a latent variable and thus its value can only be approximated. Previous studies on oil volatility have used daily squared returns from market prices as a proxy (see Agnolucci, 2009; Sadorsky, 2006). Moreover, since the historical volatility figure is used for comparison purposes only, using it as a proxy does not cause problems because it is unbiased (Lopez, 2001).

We will follow the standard techniques used by earlier studies, including Brailsford and Faff (1996), to assess the models' forecasting performance:

$$\text{RMSE}\,\frac{1}{m} = \sqrt{\sum_{h=1}^{m}(\hat{\sigma}^2 - \sigma^2)^2}$$

$$\text{MAPE } 100\left(\frac{1}{m}\right)\left[\sum_{h=1}^{m}\left|(\hat{\sigma}^2 - \sigma^2)/\sigma^2\right|\right]$$

$$\text{Mean absolute error (MAE)} = \left(\frac{1}{m}\right)\sum_{h=1}^{m}\left|(\hat{\sigma}^2 - \sigma^2)\right|$$

While the first measure depends on the scale of the forecast, the second is scale-invariant. Unlike the MAE and MAPE, the RMSE penalizes larger forecast errors more than smaller ones since it squares them. Since the MAE and MAPE use absolute values, they have the advantage of not letting the effect of under- and over-predictions of the same size cancel out.

Once we have selected the best model based on these measures, we can determine the impact of trading activity on volatility through trading volume and open interest. Following Lamoureux and Lastrapes (1990), we will add volume to the volatility equation as an explanatory variable to help assess the impact. For example, the GARCH model's volatility equation would become

$$h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i} + \gamma V_t$$

In order to avoid the effects of any contemporaneous relationship between volume and volatility, and in line with previous research, we will also test the model with lagged volume.[4] Since open interest is also

---

[4] $h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i} + \gamma V_{t-1}$

considered a proxy for trading activity, the model will also be tested with open interest replacing the volume variable. This approach can be generalized for all the models applied to determine whether trading activity reduces volatility persistence and impacts volatility.

Empirical Findings

As described in Section 4, our first step is to identify the mean equation for the returns. The autocorrelation and partial autocorrelation function for the returns show that autocorrelations and partial autocorrelations up to the fifth lag are significant (Figures A3a and A3b in the Appendix). We, therefore, propose using an autoregressive moving average (ARMA) (5,5) mean equation to model volatility in the ARCH models. The estimated ARMA (5,5) equation for the mean is found to be significant with a Wald statistic of 799.98 and significant t-values for the coefficients. The residuals of the mean equation indicate the absence of autocorrelation through the *q* statistic (Figures A4a and A4b in the Appendix). The ARMA (5,5) model is thus deemed an appropriate model for the mean equation.

The *q* statistic implies that there is second-order dependence in the squared residuals of the mean equation and, hence, the presence of conditional heteroscedasticity in the returns (Figures A5a and A5b in the Appendix). Further, the ARCH-Lagrange Multiplier (LM) test gives a Chi-squared value of 75.46, confirming the presence of ARCH effects and the need to model this conditional heteroscedasticity using the ARCH family models (Table A4 in the Appendix).

Table 2 presents the results of the models fitted to the data on returns. We do not estimate the ARCH model, the idea being that the GARCH model is a more parsimonious version of higher-order ARCH models. With the ARMA (5,5) as the underlying mean equation, estimating the GARCH (1,1) model reveals that the t-statistics for both coefficients are significant. A value of 0.9251 for past variance implies that the shock of past volatility has a persistent effect on future volatility. The sum of the two coefficients is a succinct measure of the persistence of variance, and that its value is close to 1 implies that there is significant persistence in volatility. The unconditional variance is 0.0006414, which is equivalent to an annualized variance of 10.14 percent.

Table 2: Results of models estimated

	GARCH (1,1)	EGARCH (1,2)	GJR (1,2)	TARCH (1,2)	TARCH variant (1,1)
α_0	0.0000117*	0.1267491*	0.0000153*	0.0005391*	0.0004697*
	(3.27e-06)	(0.0430031)	(4.16e-06)	(0.0001609)	(0.0001362)
α_1	0.0566712*	0.1616697*	0.1085503*	–0.0770229*	0.073906*
	(0.0065014)	(0.0166975)	(0.0125366)	(0.0119319)	(0.0066358)
β_1	0.9250869*	0.2922021*	0.1752578	0.2391775*	0.9241005*
	(0.0102548)	(0.0799058)	(0.0863492)	(0.0961673)	(0.0092274)
β_2	-	0.6903359*	0.7222816*	0.673488*	-
		(0.0796525)	(0.0857648)	(0.0951695)	
γ_1	-	–0.0687259*	–0.0634249*	0.1230721*	-
		(0.0113834)	(0.0148894)	(0.0113155)	

Notes: * = significant at 5 percent. Standard errors are given in parentheses.

For GARCH-M (1,1) model α_0 = 0.0000132*(3.48e-06), $\alpha_1$0.0608757*(0.0066205), β_1 = 0.9185359*(0.0107267), and β_2= –1.59738 (1.546592).

Source: Author's calculations.

Next, we predict and test the standardized residuals of the GARCH (1,1) model.[5] Plotting autocorrelation and partial autocorrelation functions for both standardized residuals and squared standardized residuals, and testing them using the q statistic reveals that the errors are white noise (Figure A6 in the Appendix). This means that we do not need higher-order GARCH models and that the GARCH (1,1) model is able to appropriately capture the GARCH effects. Further, the ARCH-LM test yields a p value of 0.2875, which means that the null of homoscedasticity is not rejected. We can therefore conclude that the GARCH (1,1) model is a parsimonious model and there are no remaining ARCH effects that need to be modeled by higher-order GARCH models.[6]

Following this, we test for the presence of asymmetric effects. The sign bias test yields the following results:

$$\hat{\varepsilon}_t^2 = 0.9057469 + 0.0759758 S_{t-1}^{-} + error$$

(16.90) (2.64)

[5] Standardized residuals are defined as $S_t = \hat{\mathcal{E}}_t / \hat{h}_t$.

[6] A GARCH-in-mean model was also estimated but the coefficients of the mean terms were insignificant, and hence the model was dropped.

A positive and significant coefficient indicates the presence of leverage effects, implying that positive and negative shocks do have a different effect on the conditional variance. Estimating the negative sign bias tests yields the following results:

$$\hat{\varepsilon}_t^2 = 0.9629348 - 0.1007951 S_{t-1}^{-}\hat{\varepsilon}_{t-1} + error$$

(21.42) (1.66)

A negative but insignificant coefficient implies that the effect of a negative shock on the variance does not depend on the size of that shock. Finally, the positive sign bias test yields

$$\hat{\varepsilon}_t^2 = 1.029794 - 0.0718445 S_{t-1}^{+}\hat{\varepsilon}_{t-1} + error$$

(22.46) (1.06)

Insignificant coefficients on both the negative and positive sign bias test but a significant coefficient on the sign bias test implies that there are sign effects but no size effects. Positive and negative shocks do have a different effect on the conditional variance but their effect on the variance does not depend on the size of the shocks.

Since the sign bias test indicates the presence of asymmetric effects, we proceed to estimate models from the ARCH family that do take into account asymmetries. As in the case of the GARCH model, we test the standardized residuals of these models in the same fashion using the *q* test to determine if they correctly capture the asymmetric and GARCH effects of the data. In each case, if the null hypothesis of errors being white noise is rejected, a higher-order model will be estimated until the errors tested turn out to be white noise. Table 3 presents the results for the *p* values of these tests.

The first asymmetric model considered is the EGARCH (1,1) model, the residuals of which, when tested, are not white noise, implying that a higher-order model is needed. Estimating the EGARCH (1,2) model reveals that the negative and significant coefficient of the standardized residuals provides evidence for the asymmetric effect of negative shocks on the conditional variance. The coefficient is, however, smaller in absolute value than the symmetric parameter. The results imply that if the shock is positive, its effect on $ln(h_t)$ is (0.1616697-0.0687259). However, if the shock is of the same magnitude but negative, its effect is

(0.1616697+ 0.0687259), which is almost 2.5 times more than the effect of the positive shock. There is thus strong evidence that negative innovations are more destabilizing than positive ones. The effect is, however, smaller than the symmetric effect.

Table 3: Testing for white noise using *q* statistic

***p* value**	**GARCH (1,1)**	**EGARCH (1,2)**	**GJR (1,2)**	**TARCH (1,2)**	**TARCH (1,1)**
For standardized residuals	0.9732	0.8854	0.6718	0.6848	0.8633
For squared standardized residuals	0.4324	0.3509	0.5859	0.3521	0.2753

Note: Null hypothesis: errors are white noise.
Source: Author's calculations.

In the case of GARCH effects, the effect of the two-period lagged value of volatility is greater than the effect of the one-period lagged value, but the sum of both values indicates persistence. In the GJR (1,2) model, the negative and significant coefficient of the indicator variable implies the presence of asymmetric effects. For positive shocks, therefore, the effect on the conditional variance is (0.1085501-0.0634249), while for negative shocks it is greater (0.1085503). The effect of a negative shock is more than twice the effect of the positive, which is consistent with the EGARCH model. As in the case of the EGARCH model, the effect of a two-period lagged value of conditional variance is much higher than the effect of a one-period lagged value.

Of the two variants of the TGARCH model, the first is the standard TGARCH model introduced by Zakoïan (1994). The significant coefficients of the error terms in the TARCH (1,2) model indicate the presence of asymmetric effects. In this case, the effect of a positive shock is given by (0.1230721-0.0770229), which is less than the effect of a negative shock (0.1230721). This is consistent with the results of the previous two asymmetric models. The sum of the coefficients of the GARCH terms indicates volatility persistence.

The second variant of the TARCH model takes into account the effect of the size of the shock rather than its sign. This variant is closer to the GARCH model than other asymmetric models, and shows the persistent effect of past periods' conditional variance. We use the information criteria approach to test the model's goodness of fit (Table A5 in the Appendix). It is, however, important to note that our main aim is to evaluate forecasts and, hence, forecast evaluation measures better serve this purpose.[7]

Although all asymmetric extensions of the GARCH model use different techniques to capture volatility, they produce consistent results for crude oil futures traded at the NYMEX, and imply the presence of asymmetric effects. This is in contrast to the findings of Agnolucci (2009) and Sadorsky (2006), and may be because the time frame we have used is different from the latter two, which employ data only up to 2005 and so do not take into account the recent financial crisis. Moreover, Agnolucci's (2009) asymmetric models do not consider higher-order GARCH effects, while previous research does not take into account some of the variants we have used here.

Having estimated the models, our next step is to assess their forecasts. As discussed in Section 4, we use the models to make dynamic forecasts of volatility for the next 100 observations (Table 4).

Table 4: Forecast evaluation

Model	RMSE	MAE	MAPE
GARCH (1,1)	0.000004287	0.0014844	10,373.92
EGARCH (1,2)	0.000004307	0.0014437	10,426.58
GJR (1,2)	0.000004176	0.0013299	5,956.128
TARCH (1,2)	0.000004410	0.0015198	1,1392.29
TARCH (1,1)	0.000004240	0.0013956	9,538.237

Source: Author's calculations.

All three evaluation statistics indicate that the GJR (1,2) model is best able to forecast volatility. The Diebold-Mariano test, which is applied to the GJR (1,2) and TARCH (1,1) models, ranks the latter as better able to forecast than the GJR (1,2) (see Diebold & Mariano, 1995). Of the three models, Zakoïan's (1994) TARCH makes the least accurate forecasts. Based on the statistics in Table 4, the GJR (1,2) emerges as the best model with which to forecast the volatility of returns on oil futures.

The selected model is further used to test if trading volume has any significant impact on the model itself and the forecasts made. On the lines of Liew and Brooks (1998) and Park, Switzer, and Bedrossian (1999), we add trading volume as an explanatory variable to the GJR (1,2) model

[7] The results obtained from the two techniques may not necessarily be consistent with one another.

(Table A6 in the Appendix). The coefficient of the volume variable is insignificant, indicating that volume does not have any significant impact on estimation and forecasting through the GJR model. The same is the case when using the lagged value of volume (Table A7 in the Appendix). This result is consistent with the findings of Foster (1995) but updates the latter's research by using data up to 2009, and concludes that volume does not reduce persistence in crude oil futures.

Another measure of trading activity is open interest (Bessembinder & Seguin, 1993), which we use as a proxy and test using the GJR model. As with volume, its coefficient is insignificant, indicating that open interest does not impact the estimation of volatility modeling (Tables A8 and A9 in the Appendix).

6. Conclusion

This study has attempted to model the volatility of crude oil futures and assess the forecasting ability of the ARCH family of models. We have used historical volatility for modeling purposes through the ARCH family of models and made dynamic forecasts of future volatility. The study finds the presence of asymmetric effects in the light, sweet, crude oil futures traded on the NYMEX. Of the ARCH models, the GJR (1,2) is able to make the most accurate forecasts with the TARCH (1,1) as a close second. Therefore, when volatility forecasts for oil futures are used for hedging and pricing purposes, asymmetric rather than symmetric models are best used. Additionally, we find that trading volume and open interest are unable to reduce volatility persistence in these futures.

The models' forecasts can be extended for use in asset pricing models. Further improvements to the current study are also possible. First, intraday data and realized volatility could prove a better proxy for actual volatility than squared residuals. This would refine the process of forecast evaluation. Second, asymmetric power models and fractionally integrated models—also of the ARCH family—could be used to analyze volatility behavior.

References

Agnolucci, P. (2009). Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. *Energy Economics*, *31*, 316–321.

Alsubaie, A., & Najand, M. (2009). Trading volume, time-varying conditional volatility, and asymmetric volatility spillover in the Saudi stock market. *Journal of Multinational Financial Management*, *19*, 139–149.

Bessembinder, H., & Seguin, P. J. (1993). Price volatility, trading volume and market depth: Evidence from futures markets. *Journal of Financial and Quantitative Analysis*, *28*, 21–39.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollerslev, T., Chou, R., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*, 5–59.

Bracker, K., & Smith, K. L. (1999). Detecting and modeling volatility in the copper futures market. *Journal of Futures Markets*, *19*(1), 79–100.

Brailsford, T. J. (1996). The empirical relationship between trading volume, returns, and volatility. *Accounting and Finance*, *35*, 89–111.

Brailsfold, T. J., & Faff, R. W. (1996). An evaluation of volatility forecasting techniques. *Journal of Banking and Finance*, *20*, 419–438.

Brooks, C. (1998). Predicting stock index volatility: Can market volume help? *Journal of Forecasting*, *17*, 59–80.

Brooks, C., & Persand, G. (2003). Volatility forecasting for risk management. *Journal of Forecasting*, *22*, 1–22.

Carvalho, A. F., da Costa, J. S., & Lopes, J. A. (2006). A systematic modeling strategy for futures markets volatility. *Applied Financial Economics*, *16*, 819–833.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, *41*, 135–156.

Darrat, A. F., Rahman, S., & Zhong, M. (2003). Intraday trading volume and return volatility of the DJIA stocks: A note. *Journal of Banking and Finance*, *27*, 2035–2043.

Dickey, D., & Fuller, W. (1979). Distributions of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427–431.

Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253–263.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*, 987–1007.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, *48*, 1749–1778.

Ferderer, J. (1996). Oil price volatility and the macro-economy. *Journal of Macroeconomics*, *18*, 1–26.

Foster, A. J. (1995). Volume volatility relationships for crude oil futures market. *Journal of Futures Markets*, *15*, 929–951.

Gallo, G. M., & Pacini, B. (2000). The effects of trading activity on market volatility. *European Journal of Finance*, *6*, 163–175.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stock. *Journal of Finance*, *48*, 1779–1801.

Hull, J. C. (2005). *Options futures and other derivatives* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Kang, S.-H., Kang, S.-M., & Yoon, S.-M. (2009). Forecasting volatility of crude oil markets. *Energy Economics*, *31*, 119–125.

Lamoureux, G. C., & Lastrapes, W. D. (1990). Heteroscedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, *45*, 221–229.

Lee, K., Ni, S., & Rutti, R. A. (1995). Oil shocks and the macro-economy: The role of price variability. *The Energy Journal*, *16*, 39–56.

Liew, K. Y., & Brooks, R. (1998). Returns and volatility in the Kuala Lumpur crude palm oil futures market. *Journal of Futures Markets*, *18*, 985–999.

Lopez, J. A. (2001). Evaluating the predictive accuracy of volatility models. *Journal of Forecasting*, *20*, 87–109.

Najand, M., & Yung, K. (1991). A GARCH examination of the relationship between volume and price variability in futures market. *Journal of Futures Markets*, *11*, 613–621.

Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: A new approach. *Econometrica*, *59*, 347–370.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, *45*, 267–290.

Park, T.-H, Switzer, L. N., & Bedrossian, R. (1999). The interactions between trading volume and volatility: Evidence from the equity options market. *Applied Financial Economics*, *9*, 627–637.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*, 335–346.

Poon, S.-H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*, 478–539.

Pyuna, C. S., Lee, S. Y., & Nam, K. (2000). Volatility and information flows in emerging equity markets: A case of the Korean Stock Exchange. *International Review of Financial Analysis*, *9*, 405–420.

Ragunathan, V., & Peker, A. (1997). Price variability, trading volume and market depth: Evidence from the Australian futures market. *Applied Financial Economics*, *7*, 447–454.

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics*, *21*, 449–469.

Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. *Energy Economics*, *28*, 467–488.

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance*, *44*(5), 1115–1153.

Schwert, G. W. (1990). Stock volatility and the crash of ’87. *Review of Financial Studies*, *3*, 77–102.

Sharma, J. L., Mougous, M., & Kamath, R. (1996). Heteroscedasticity in stock market indicator return data: Volume versus GARCH effects. *Applied Financial Economics*, *6*, 337–342.

Taylor, S. (1986). *Modeling financial time series*. Chichester, UK: Wiley.

Zakoïan, J.-M. (1994). Threshold heteroscedastic models. *Journal of Economic Dynamics and Control*, *18*, 931–955.

Appendix

Figure A1a: AC of prices

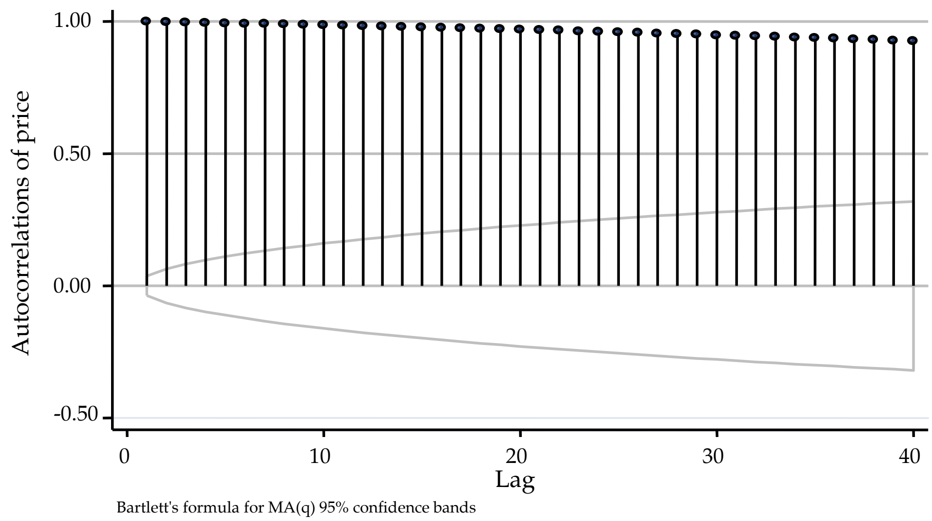
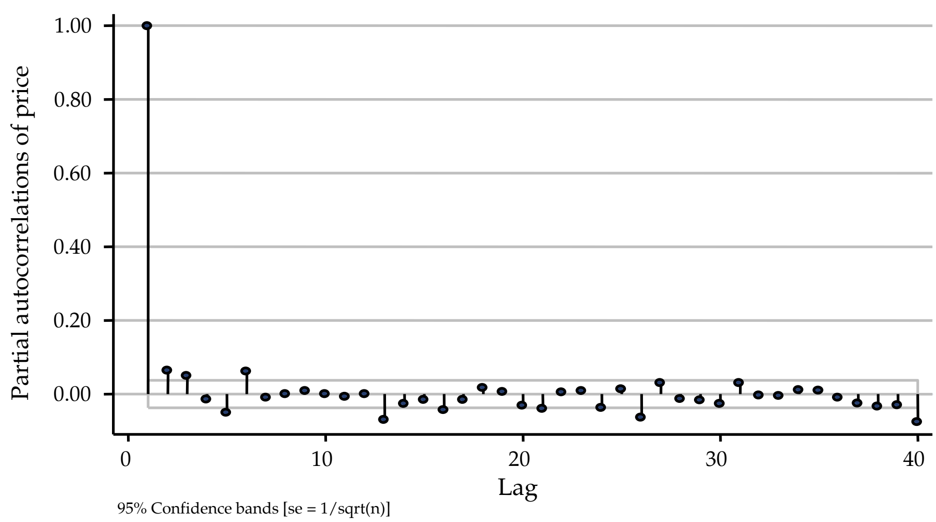


Figure A1b: PAC of prices



Source: Author's calculations.

Table A1: Dickey-Fuller test for price

Test statistic	1% critical value	5% critical value	10% critical value
–1.548	–3.430	–2.860	–2.570

Note: MacKinnon approximate p-value for Z(t) = 0.5081.
Source: Author's calculations.

Table A2: Dickey-Fuller test for returns

Test statistic	1% critical value	5% critical value	10% critical value
–52.743	–3.430	–2.860	–2.570

Note: MacKinnon approximate p-value for Z(t) = 0.0000.
Source: Author's calculations.

Table A3: Phillips-Perron test for returns

	Test statistic	1% critical value	5% critical value	10% critical value
Z(rho)	–2,588.516	–20.700	14.100	–11.300
$Z(t)$	–52.881	–3.430	–2.860	–2.570

Note: MacKinnon approximate p-value for Z(t) = 0.0000.
Source: Author's calculations.

Figure A2a: AC of squared returns

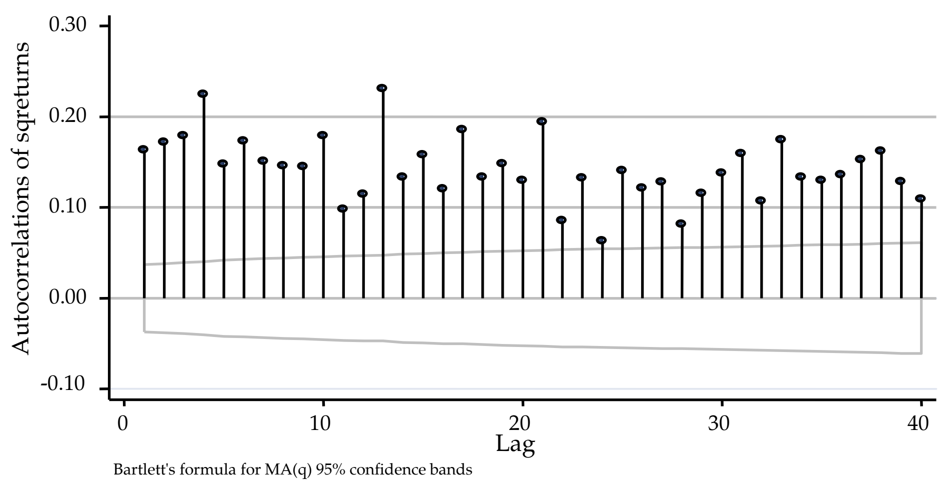
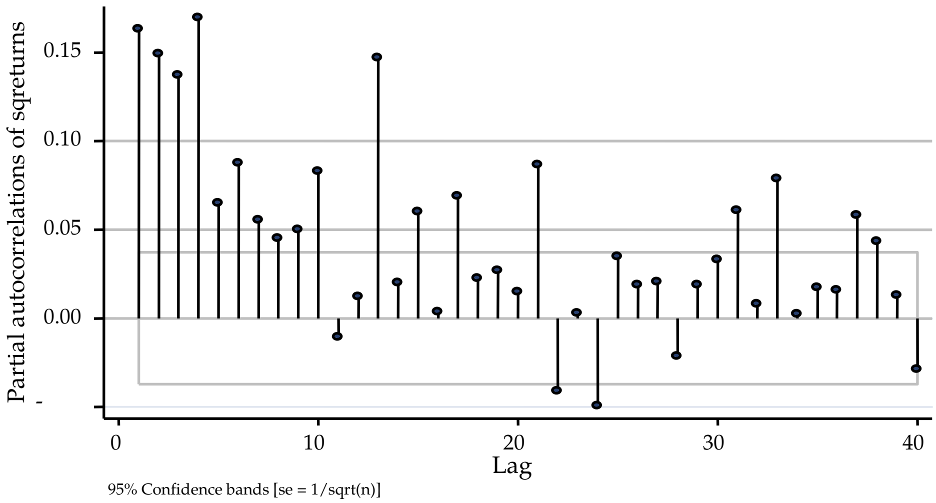


Figure A2b: PAC of squared returns



Source: Author's calculations.

Figure A3a: AC of returns

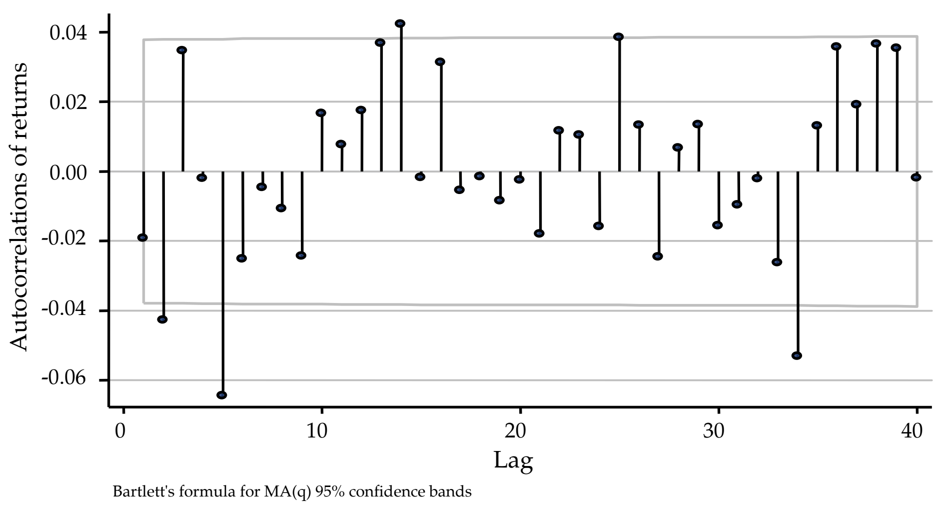
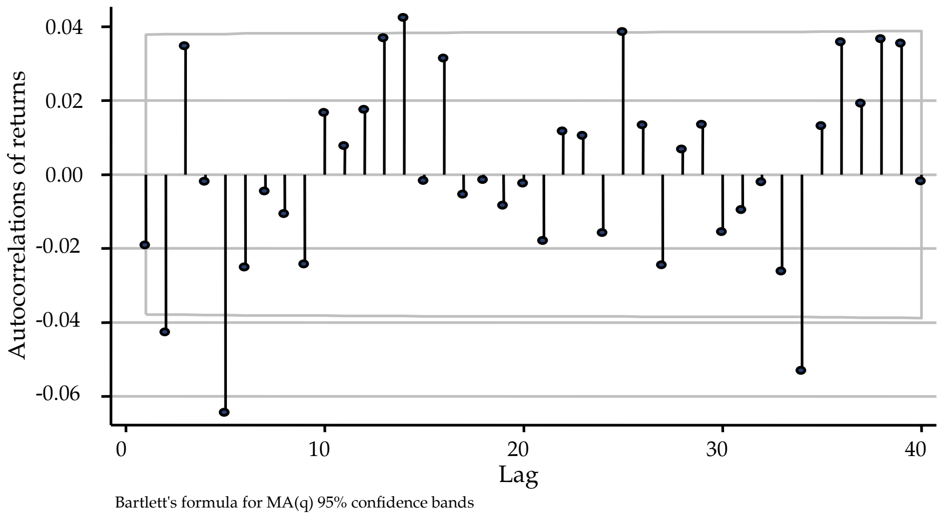


Figure A3b: PAC of returns



Source: Author's calculations.

Figure A4a: AC of residuals of mean equation

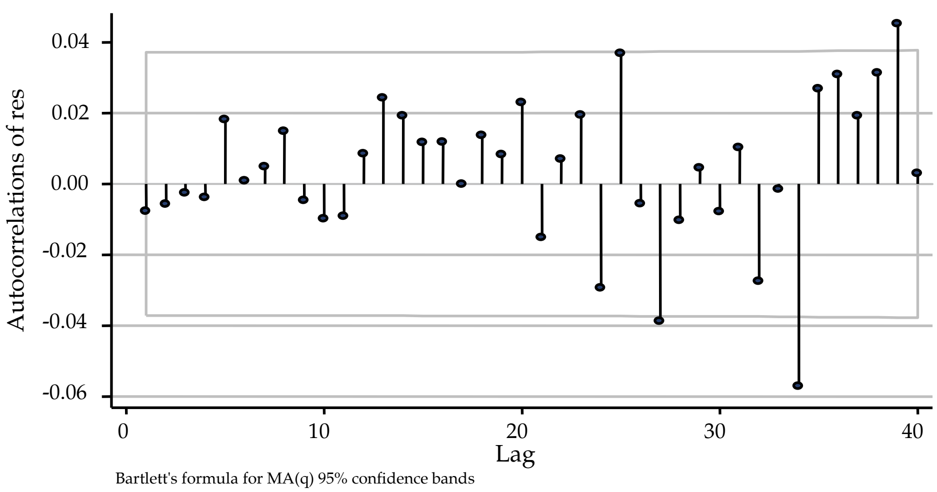
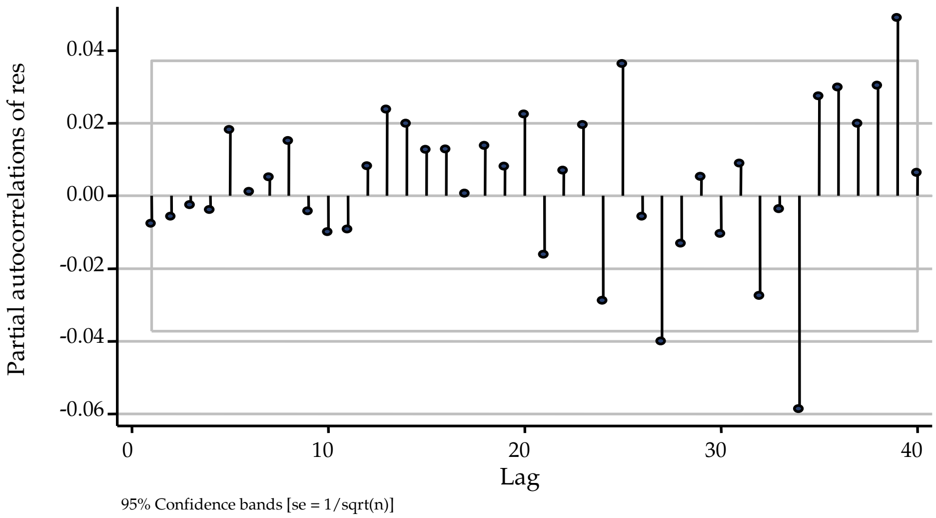


Figure A4b: PAC of residuals of mean equation



Source: Author's calculations.

Figure A5a: AC of squared residuals of mean equation

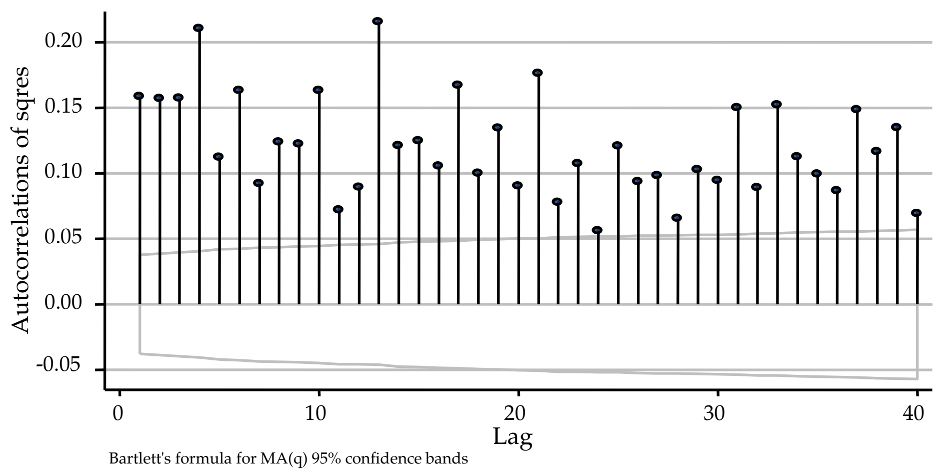
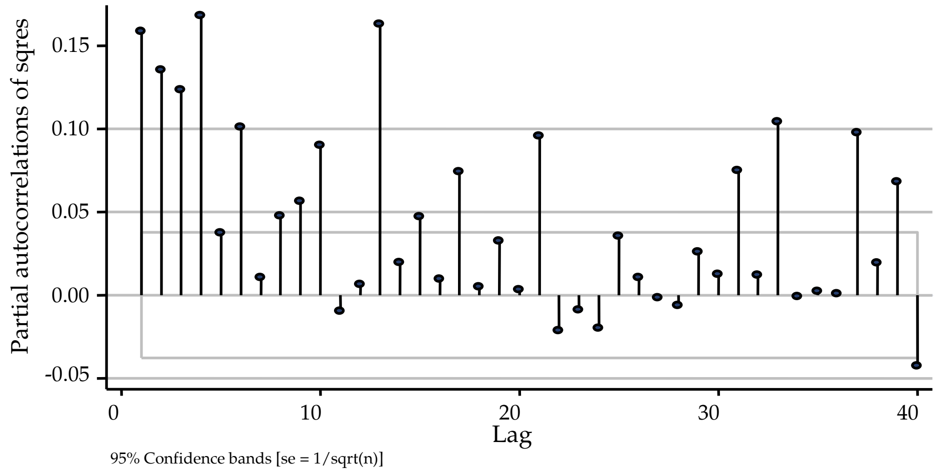


Figure A5b: PAC of squared residuals of mean equation



Source: Author's calculations.

Table A4: LM test for autoregressive conditional heteroscedasticity

Lags (*p*)	**Chi²**	**Df**	**Prob. > Chi²**
1	75.462	1	0.00

Note: H0: no ARCH effects vs. H1: ARCH (*p*) disturbance.
Source: Author's calculations.

Figure A6a: AC of squared residuals of GARCH (1,1)

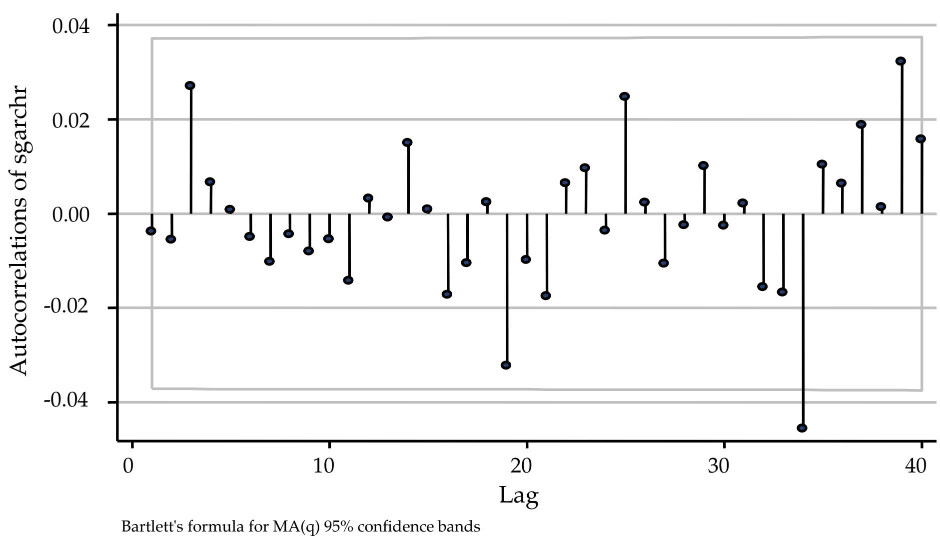
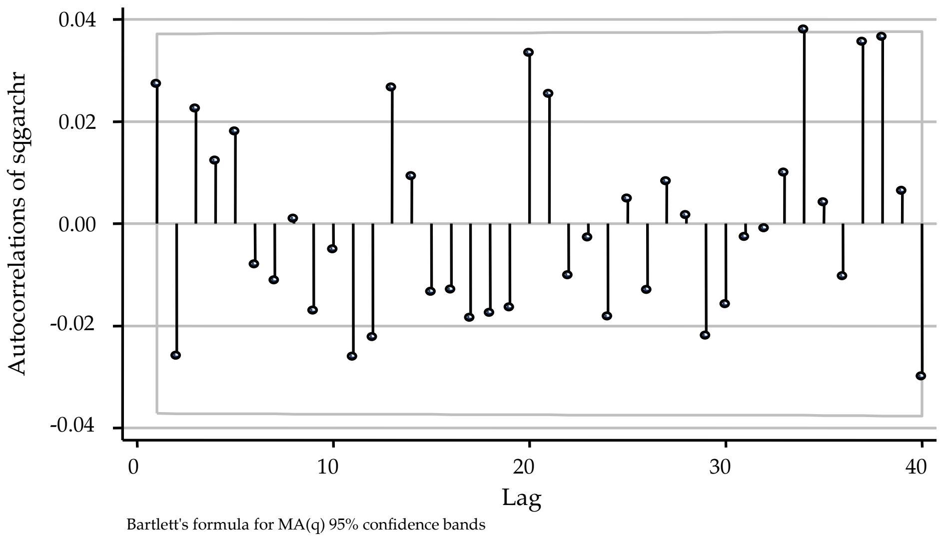


Figure A6b: PAC of squared residuals of GARCH (1,1)



Source: Author's calculations.

Table A5: Goodness-of-fit tests for models

Model	AIC	BIC
GARCH (1,1)	–12761.56	–12678.62
EGARCH (1,2)	–12448.77	–12364.47
GJR (1,2)	–12775.49	–12680.70
TARCH (1,2)	–12802.24	–12707.45
TARCH (1,1)	–12443.82	–12361.32

Source: Author's calculations.

Table A6: GJR model with volume variable

Variable	Coefficient	t-statistic
α_0	0.0000154	3.69
α_1	0.1085968	8.66
β_1	0.1751763	2.03
β_2	0.7223081	8.42
γ_1	–0.0634411	–4.23
Volume	–7.30e-11	–0.01

Source: Author's calculations.

Table A7: GJR model with lagged volume variable

Variable	Coefficient	t-statistic
α_0	0.0000155	3.70
α_1	0.1090224	8.64
β_1	0.1770838	2.03
β_2	0.7195726	8.30
γ_1	–0.0629534	–4.20
Lagged volume	2.86e-10	0.06

Source: Author's calculations.

Table A8: GJR model with open interest variable

Variable	Coefficient	t-statistic
α_0	–2.37e-09	–0.53
α_1	0.1090175	8.68
β_1	0.1740498	2.03
β_2	0.7229511	8.50
γ_1	–0.0638242	–4.25
Open interest	0.0000155	3.71

Source: Author's calculations.

Table A9: GJR model with lagged open interest variable

Variable	Coefficient	t-statistic
α_0	0.0000156	3.72
α_1	0.1093515	8.68
β_1	0.1762446	2.03
β_2	0.7198879	8.50
γ_1	–0.0632256	–4.23
Lagged open interest	–2.53e-09	–0.55

Source: Author's calculations.