International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

EXAMINING VALUE AT RISK IN MULTI FACTOR MODEL: THE CASE OF MARKETS GOVERNED BY RETAIL INVESTORS

Dr. Amina Amirat, Dr. Makram Zaidi

[1] Assistant professor at college of administrative sciences – Najran University (KSA)
email: amina.amirat@yahoo.fr (corresponding author)
[1] Assistant professor at college of administrative sciences – Najran University (KSA)
email: zaydi.makram@gmail.com

ABSTRACT

In this article we examine markets dominated by retail investors where herding behavior can be prevalent. We consider a multi factor model to forecast stock returns that we suppose affected by size, book to market and herding behavior. Applying the model to Saudi stock market on daily data from 7th January 2007 to 1st March 2016, we construct three type of weighted portfolio: large, mid and small capitalization. The result of a logistic regression shows that our model can estimate stock returns with a higher precision of more than 70%. Using our model we estimate the out-of-sample Value at Risk using historic simulation. Finally we conduct a back testing which confirm the precision of the forecasted VaR.

Keywords: Back testing, Herding behavior, logistic regression, retail investors, Value at Risk.

JEL classification: G02 – G1

I. INTRODUCTION

The bulk of research shows that retail investors behave differently from institutional investors in stock market. They can hold under-diversified portfolios or trade actively, speculatively, and to their detriment. And, as a crowd, retail investors make systematic, not random, trading decisions.

Literature has recently presented a varied behaviors among retailers that is hard to settle with the traditional vision of a rational homo economicus. In contrast to the "smart money" controlled by institutional investors, it appears believable that retail investors may consistently suffer from costly behavioral biases. These biases may affect not only the performance of retail investors, but also their perception of risk.

Usually retail investors herd more towards growth stocks or glamour stocks which capture newspaper and television headlines. Irrational views drive this herding behavior. Academics have demonstrated that investors disregard their prior beliefs or forego rational analysis, and often follow other investors blindly. Retail Investors inevitably suffer huge economic losses after the burst of the bubbles formed by herding. For this reason we will include the herding behavior in a multi factor model to forecast stock returns and Value at Risk.

Due to increasing importance; the aim of this study is to forecast the stock market returns trends by using logistic regression in a multi factor model including size, book to market and herding behavior. Then, we will use our model to forecast the out-of-sample Value at Risk. Our contribution is twice, the first concern the estimation of stock returns where we include herding behavior as the most important behavior of retail investors. The second is at the level of forecasting the Value at Risk generated by the one-day ahead estimation with rolling window of 5 days. The model has used the preprocessed data set of 17 stocks from seven different sectors of Saudi stock market. The data set encompassed the daily data from 7th January, 2007 to 1st March, 2016.

This paper contains four sections. The first one introduce the study, in the second we review literature linked to our paper, then we explain methodology of research and present results, the forth section concludes the paper.

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

II. LITERATURE REVIEW

Our literature review will be divided into three parts. The first summarizes studies on retail investors behavior. The second recaps works on multi factor models to estimate stock returns. The third outlines papers related to Value at risk forecast.

It is frequently discussed in the literature that retail investors are more expected to be noise traders, traders who involve in speculative trading based on technical information. A key empirical part of this argument is built on an analysis of individual investors trading behavior (barber et al., 2009b; Kumar & Lee, 2006) and its consequences (Barber & Odean, 2000; Barber, Lee, Liu & Odean, 2009;Odean, 1999). Yet, since different clusters of investors inhabit the real world stock market, any ample analysis of the trading behavior of a specific group of traders would ideally rely on the transaction histories of the group members (Kumar & Lee, 2006; Lakonishok et al., 1992).

In the study of Balcilar, Demirer, and Hammoudeh (2013), they examined on the presence of herd behavior in five Gulf Arab stock markets including that of Saudi Arabia. They analysis herding under diverse market conditions, defined as low, high and extreme volatility regimes, by estimating the Chang et al. (2000) model using a three-state Markov switching framework.

Rahman, Chowdhury & Sadique (2015) investigate herding in the Saudi stock market, where more than 95% of the total trading is initiated by the individual investors. Based on readily available stock data, they find evidence of pervasive herding among the market participants. Although herding is prevalent irrespective of market conditions, it tends to get stronger in periods when the market rises and the trading activity intensifies. Traders are found to be indifferent to important stock categories in their herd behavior.

It is notable that almost all preceding research of market-wide herd behavior rely on a specific testing method, which often leads to opposing results. For example, using their respective methodologies, Chang et al.(2000) and Hwang and Salmon (2004) report completely opposite findings of herd behavior in the US stock market. Similarly, while

Demirer and Kutan (2006) using the Christie and Huang (1995) method find no evidence of herding in Chinese stock market, Tan et al. (2008) report significant evidence of herding in that market based on the Chang et al. (2000) methodology.

Concerning the second part of our literature review, there is considerable empirical proof that stock returns are correlated to both market risk factors and factors that are founded on firm-level characteristics like firm size, book-to-market ratio and momentum (Gong & Weng, 2016). For example, the classical capital asset pricing model (CAPM) developed by Sharpe (1964) and Lintner (1965) shows that stock returns can be determined by the market risk factors. Subsequent work by Fama and French (1993) demonstrated that the cross-sectional variations in average equity returns can be better explained by a combination of risk factors such as market risk, firm-level market capitalizations and book-to-market effects than by the one-factor model. Moreover, numerous studies indicate that stock returns may be affected by industry factors, trading volumes, investors' trading behaviors and so on (Barber & Odean, 2008; Huddart, Lang & Yetman, 2009).

Both the arbitrage pricing theory (APT) of Ross (1976) and the multifactor models of asset returns, plays a vital role in modern finance theory. Under a multifactor model, the return of each security is expressed as a linear combination of a small number of factor returns and an asset-specific return. Goyal et al. (2008) argued that the assumption that all factors influence a large number of assets, so-called pervasive factors, are too strong if an economy is partitioned into several groups. Connor and Korajczyk (1993) pointed out that industry-specific components may not be pervasive sources of uncertainty for the entire economy (see also Cho et al. (1986) and Bekaert et al. (2009)). Ando and Bai (2014) analyzes multifactor models in the presence of a large number of potential observable risk factors and unobservable common and group-specific pervasive factors. We show how relevant observable factors can be found from a large given set and how to determine the number of common and group-specific unobservable factors.

The last part of our literature review concerns the forecast of Value at Risk. First stage in calculating VaR is denoted by volatility's estimation. Engle (1982) proposed ARCH model to calculate VaR which is further generalized by Bollerslev (1986) into a GARCH model,

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

which was improved over time. There is a lot of investigation which compare Exponentially Weighted Moving Average (EWMA) performance with different types of ARCH/GARCH models in estimating volatility. According to Hull (2008) the key advantage of EWMA is characterized by relatively little data needs to be stored. Additionally, Tse (1991) and Tse and Tung (1992) highlight the fact that EWMA model over performed ARCH models in estimating the risk for Japanese and Singaporean financial markets.

But not all authors proved that EWMA is the best model in forecasting the volatility. Regarding this, Hammoudeh et al. (2011) found that GARCH-t model over perform EWMA in estimating the risk involved in commodities market. Moreover, Degiannakis et al. (2011) proves that ARCH framework is better in assessing risk compared to RiskMetrics model.

Moreover, several researchers as McMillan and Kambouroudis (2009), and Pafka and Kondor (2001) emphasized the fact that the RiskMetrics performance in forecasting the risk directly depends on the choice of the significance level (of 90%, 95% or 99%). There are some papers as Fan et al. (2004) and Gonzalez-Riviera et al. (2007), which stated that decay factor's value is not necessary equal with 0.94, value imposed by RiskMetrics methodology for daily data.

III. DATA, METHODOLOGY AND RESULTS

3-1. Data and variables

The data consists of daily stock returns of 17 firms from different sectors cited in the Saudi stock market (Tadawul). The period under consideration is from 07/01/2007 to 01/03/2016. The data set consists of 2287 data points for each firm. The data has been obtained from the official web site of Saudi stock market that provides daily stock market data (*www.tadawul.com.sa).*

Table 1: Description of sample data

Sector	Symbol	Name
Banks & Financial Services	1060	Saudi British Bank
	1120	Al Rajhi Bank
	1090	Samba Financial Group

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

Petrochemical Industries	2010	Saudi Basic Industries Corp.
	2330	Advanced Petrochemical Co.
	2250	Saudi Industrial Investment Group
Cement	3010	Arabian Cement Co.
	3040	Qassim Cement Co.
Agriculture & Food Industries	6050	Saudi Fisheries Co.
	2280	Almarai Co.
	6010	National Agricultural Development Co.
Energy & Utilities	5110	Saudi Electricity Co.
	2080	National Gas and Industrialization Co
Telecommunication & Information Technology	7010	Saudi Telecom Co.
Real Estate Development	4020	Saudi Real Estate Co.
	4100	Makkah Construction and Development Co.
	4090	Taiba Holding Co.

The first step in our study is to construct 3 portfolios by size. That designates we choose the stocks that based on the company's market value to escape the size influence since people always select the stocks by the size trading strategy. We divide our 17 securities into 3 parts according to their size, so we get 3 kinds of type stocks and the number of each part is 5, 7 and 5, respectively : large capitalization, medium capitalization and small capitalization.

Table 2: Composition of portfolio

Portfolio	**Symbol**	**Name**	**Sector**
Portfolio 1 Large capitalization	1120	Al Rajhi Bank	Banks & Financial Services
	2010	Saudi Basic Industries Corp	Petrochemical Industries
	2330	Advanced Petrochemical Co.	
	6010	National Agricultural Development Co.	Agriculture & Food Industries
	7010	Saudi Telecom Co.	Telecommunication & Information Technology
Portfolio 2 Medium capitalization	6050	Saudi Fisheries Co.	Agriculture & Food Industries
	2280	Almarai Co.	
	4020	Saudi Real Estate Co.	Real Estate Development
	3010	Arabian Cement Co.	Cement

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

	3040	Qassim Cement Co.	
	5110	Saudi Electricity Co.	Energy & Utilities
	2250	Saudi Industrial Investment Group	Petrochemical Industries
Portfolio 3 Small capitalization	4090	Taiba Holding Co.	Real Estate Development
	4100	Makkah Construction and Development Co.	
	1090	Samba Financial Group	Banks & Financial Services
	1060	Saudi British Bank	
	2280	Almarai Co.	Energy & Utilities

Then, we attempt to optimize each portfolio. The main objective is to maximize of the returns and to minimize the risk of portfolio. We use the Markowitz model (1991) to optimize our three portfolios:

The expected return for portfolio is:

$$R_p = \sum_{i=1}^{N} w_i r_i \qquad (1)$$

The standard deviation of the portfolio is given by:

$$\sigma_p = \sqrt{\sum_{i,j=1}^{N} \sigma_{ij} w_i w_j} \qquad (2)$$

Where:

w_i, w_j is the weighting of asset *i* and *j* respectively.

r_i is the expected return of asset *i*.

σ_{ij} is the covariance between assets *i* and *j*.

N is the number of assets available.

A common risk measure that one might consider maximizing in the allocation procedure is the Sharpe ratio:

$$Sharpe\ ratio\ (Sp) = \frac{Rp}{\sigma_p} \qquad (3)$$

Indeed, Markowitz/Sharpe portfolio theory is built on the result that the weights $w_i$obtained from the optimization problem:

$$\max_{w_i} \frac{\sum_{i=1}^{N} w_i r_i - r_f}{\sqrt{\sum_{i,j=1}^{N} \sigma_{ij} w_i w_j}} \qquad (4)$$

Subject to:

$$\sum_{i=1}^{N} w_i = 1,\ a_i z_i \leq w_i \leq b_i z_i,\ 0 \leq a_i \leq 1,\ 0 \leq b_i \leq 1 \qquad (5)$$

Where, $z_i = \begin{cases} 1, & for\ w_i > 0 \\ 0, & otherwise \end{cases}$ (6)

Equation (5) shows the budget constraint, which ensure that the sum of weights associated with each asset is equal to one; i.e. all the available money is invested in the portfolio. a_i is the floor constraint and it is the lowest limit on the proportion of any asset that can be held in a single portfolio. It prevents an excessive administrative cost for very small holdings, which have an insignificant influence on the performance of the portfolio. b_i is the ceiling constraint and is the maximum limit on the proportion of any asset that can be held in a single portfolio. It prevents the excessive exposure to any portfolio, which is the part of the institutional diversification policy (Mishra et al., 2016).

By applying this theoretical framework, we obtain the following weighted portfolio:

Table 3: Portfolio selection

Portfolio	Symbol	w_i	Rp	σ_p	Sp
Portfolio 1 Large capitalization	1120	0.2	0.000222851	0.021291054	-1.013658969
	2010	0.2			
	2330	0.2			
	6010	0.2			
	7010	0.2			
Portfolio 2 Medium	6050	0.1	0.000171351	0.025089812	-0.862049060
	2280	0.1			
	4020	0.1			
	3010	0.4			

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

capitalization	3040	0.1			
	5110	0.1			
	2250	0.1			
Portfolio 3 Small capitalization	4090	0.2	0.000166647	0.019507297	-1.109230276
	4100	0.2			
	1090	0.2			
	1060	0.2			
	2280	0.2			

These portfolio will be the base of our study.

3-2. Methodology

3-2-1- Multi model factor

In this study we try to generalize the Fama and French model (1993) by adding reasonable exogenous explanatory variables to the regression model. We use traditional factors such as market factor, size factor, book-to-market ratio factor and we introduce a behavior explanatory variables.

Because of the Saudi market is dominated by the retail investors, we choose a behavioral variable which is herding behavior. The behavioral finance theory uses herding to describe the correlation in trades ensuing from investors' interactions. This concept suggests that it is reasonable for less sophisticated investors to imitate market gurus or to seek advice from victorious investors, since using their own information will incur less benefice and more cost. The consequence of this herding behavior is, as Nofsinger and Sias (1999) noted, "A group of investors trading in the same direction over a period of time." Empirically, this may lead to observed behavior patterns that are linked across individuals and that fetch about systematic, erroneous decision-making by all populations (Bikhchandani, Hirshleifer, and Welch, 1992). So, in addition to newspapers, investors' trading behavior can cause stock prices to deviate from their fundamentals. As a result, stocks are not appropriately priced. There are varieties of herding models that have been presented. Generally, we separate models that produce rational prices in efficient markets from those that can potentially give rise to price bubbles and crashes that are due to temporary price pressures pushing market prices away from fundamental values.

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

Our multi factor model is defined as follow:

$$R_{i,t} - RF_t = \beta_0 + \beta_1[RM_t - RF_t] + \beta_2 SIZE_{i,t} + \beta_3 BM_{i,t} + \beta_4 HERD_{i,t} + \varepsilon_{i,t} \quad (7)$$

Where:

$R_{i,t}$ the daily returns from individual stock i.

RF_t the risk-free rate.

RM_t market daily return, which is the proxy for the market factor.

$SIZE_{i,t}$ is the proxy for firm size, which is computed as the product of the number of outstanding stocks and the closing price at the end of each trading day and then is translated to the natural algorithm.

$BM_{i,t}$ is the proxy for the book to market ratio factor and is calculated as the reciprocal of price to book (PB).

$HERD_{i,t}$ is the proxy for herding behavior factor and is calculated using cross-sectional standard deviation of returns (CSSD) as a measure of the average proximity of individual asset returns to the realized market average (Christie and Huang, 1995):

$$CSSD_t = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(R_{i,t} - RM_t\right)^2} \quad (8)$$

This model is estimated using logistic regression. To carry out the logistic regression, first a method is needed for classifying a company as a “good” or “poor” investment choice for a given time. In this study we use a method that is simple, if the value of a company’s stock over a given time rose above market return, it is classified as a “good” investment option; otherwise, it is classified as a “poor” investment option. Here, the TASI (Tadawul All Share Index) return has been taken as proxy for market return. The return was calculated using the following formula:

$$return = \frac{p_j - p_{j-1}}{p_{j-1}} \times 100 \quad (9)$$

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

Where:

p_j is the closing price for week j

p_{j-1} is the closing price for week j-1

3-2- Application of Panel logistic regression

Logistic regression is used in our study because we assume that the relation between variables is non-linear. Also this type of regression is preferred when the response variable is binary which means that can take only two values 1 or 0.

Logistic regression could forecast the likelihood, or the odds ratio, of the outcome based on the predictor variables, or covariates. The significance of logistic regression can be evaluated by the log likelihood test, given as the model chi-square test, evaluated at the $p < 0.05$ level, or the Wald statistic. Logistic regression has the advantage of being less affected than discriminant analysis when the normality of the variable cannot be assumed.

In the logistic regression model, the relationship between Z and the probability of the event of interest is described by this link function.

$$p_{ij} = \frac{e^{zij}}{1+e^{zij}} = \frac{1}{1+e^{-zij}} \quad (10)$$

$$z_{ij} = log\left(p_{ij}/1 - p_{ij}\right) \quad (11)$$

Where

p_{ij} is the probability the j^{h} case experiences the event of interest at time i

z_{ij} is the value of the unobserved continuous variable for the j^{th} case at time t

The z value is the odds ratio. It is expressed by

$$z_{ij} = c + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} \quad (12)$$

Where

x_{ij} is the ratio of the j^th firm at time i^th

β_j is the j^th coefficient

P is the number of firms

i is time

βj are the regression coefficients that are estimated through an iterative maximum likelihood method. However, because of the subjectivity of the choice of these misclassification costs in practice, most researchers minimize the total error rate and, hence, implicitly assume equal costs of type I and type II errors [Ohlson, 1980; Zavgren, 1985].

In order to carry out logistic regression analysis, first a method is needed for classifying returns as a “1” or “0” for a given day. In this study we use a method that is simple and objective, if the value of a return in week j is above the market return, it is noted as a “good” (mentioned "1"); otherwise, it is classified as a “poor” (mentioned "0").

3-2-2- Value at Risk forecast

Value-at-Risk (VaR), is a extensively used measure of financial risk, which offers a way of quantifying and managing the risk of a portfolio. VaR was considered in 1993 partly in response to numerous financial tragedies. The key use of VaR is for measuring market risk (exposure to losses in the market place through physical and derivative positions) although VaR is being used more frequently to assess credit risk (credit VaR modelling). However, VaR does not give a reliable method for measuring risk, as different VaR models will come up with different VaR results. It should also be noted that VaR only measures quantifiable risks; it cannot measure risks such as liquidity risk, political risk, or regulatory risk. In times of great volatility, such as war, it may also not be reliable. For this reason, VaR models should always be used alongside stress testing. VaR is the assessment of the worst loss in value of portfolio over a target horizon within a given probability level (Jorion, 2007). Mathematically:

$$p(R_t \leq VaR_t^{\alpha} | F_{t-1}) = \alpha \qquad (13)$$

α is the probability level.

F_{t-1} is the information set at time *t-1*.

For financial experts and statisticians, it is an constant challenge to catch an suitable models and methods for VaR estimating, modeling, and forecasting. Over the years, both parametric and non-parametric estimation approaches have been suggested to estimate VaR in the literature (Wei, Chen and Lin, 2013; Zhang, Wei, Yu, Lai, and Peng, 2014). In our study, we will use the non-parametric method which is represented by a simple historical simulation (SHS) approach that does not use conditioning information. Within the SHS standard method, an empirical loss distribution of the portfolio returns must be derived using the past historical sample data before the portfolio VaR is computed. Then the requested portfolio VaR can be estimated from the maximum loss in this distribution that is associated with a given probability level.

For each portfolio considered in this study, we generate one-day-ahead forecasts of VaR for each day in the sample period. A rolling window approach is used again, where a fixed sample size of 5 trading days is employed for estimation, to calculate the daily updated predicted VaR for each given α and each given portfolio in the observation period.

3-2-3- Value at Risk Back testing

Backtesting is a technique for simulating a model or strategy on past data to gauge its accuracy and effectiveness. Backtesting in value at risk is used to compare the predicted losses from the calculated value at risk with the actual losses realized at the end of the specified time horizon. This comparison identifies the periods where the value at risk is underestimated or where the portfolio losses are greater than the original expected value at risk. The value at risk predictions can be recalculated if the backtesting values are not accurate, thereby reducing the risk of unexpected losses.

In our study we will use Kupiec (1995) coverage test. Kupiec's "proportion of failures" (PF) coverage test takes a circuitous—and approximate—route to an answer, offering no particular advantage over our recommended standard coverage test. Comparing the two tests

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

can be informative, illustrating the various respects in which test designs may differ. As the first published backtesting methodology, the PF test has been widely cited.

As with the recommended standard test, a VaR measure is observed for α + 1 periods, experiencing X exceedances. We adopt the same null hypothesis $\mathcal{H}_0$ that q = q*. Rather than directly calculate probabilities from the B(α + 1, 1 – q) distribution of X under $\mathcal{H}_0$, the PF test uses that distribution to construct a likelihood ratio:

$$\Lambda = \frac{q^{\alpha+1-X}\,(1-q)^X}{\left(\frac{\alpha+1-X}{\alpha+1}\right)^{\alpha+1-X}\left(\frac{X}{\alpha+1}\right)^X} \qquad (14)$$

For a given significance level ε, we construct a non-rejection interval [x_1, x_2] such that

$$Pr(X < x_1) \leq \varepsilon/2 \ and \ Pr(x_2 < X \ ) \leq \varepsilon/2 \qquad (15)$$

under $\mathcal{H}_0$. To do so, calculate the ε quantile of the χ2(1,0) distribution. Setting this equal to [14.7], solve for X. There will be two solutions. Rounding the lower one down and the higher one up yields x_1 and x_2.

3-3- Results

3-3-1- Multi factor model estimation

As mentioned, the study contains 2287 data where 1525 are used for estimating and 762 used for validating the model. For variables, we have the portfolio excess return as dependent variable and four independent variables.

Before presenting the results of our logistic regression, we summarize the descriptive statistics for individual stocks and for portfolio.

Table 4: Descriptive statistics

Stocks	Minimum	Maximum	Mean	Std. Deviation	VaR (95%)
1060	-.1000000000-	.0994845360	.000045319010	.0190738970236	-2.90%
1090	-.1000000000-	.0979827090	-.000142934627-	.0194892458336	-2.90%

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

1120	-.0989654500-	.0985887500	-.000084108495-	.0180465950351	-2.79%
2010	-.0997624700-	.1000000000	.000178680010	.0226727369667	-3.42%
2330	0E-10	.1003861000	.008907664456	.0169039821655	-3.69%
2250	-.1000000000-	.0989304810	.000034942308	.0242556795356	-3.72%
3010	-.1009900990-	.0995024880	.000118872357	.0189050959808	-2.68%
3040	-.0988142290-	.1009174310	.000046794305	.0162307013641	-2.24%
6050	0E-10	.1037735850	.010743386254	.0217649578563	-5.12%
2280	-.1029962550-	.1016166280	.000853071459	.0190693955233	-2.48%
6010	0E-10	.1004608290	.007404241411	.0146524075400	-3.36%
5110	-.0983606560-	.0986842110	.000190262386	.0167523716311	-2.17%
2080	-.1000000000-	.1000000000	-.000032640035-	.0189290819197	-2.73%
7010	-.1000000000-	.1000000000	.000044120337	.0166591794464	-2.53%
4020	0E-10	.1000000000	.008146950341	.0154317058768	-3.92%
4100	-.1000000000-	.1000000000	.000504360439	.0197962231593	-2.92%
4090	-.1000000000-	.0996441280	.000415364847	.0204240765148	-2.92%

The estimation of logistic regression is done using the software Eviews 7 and the results are summarized in the following table:

Table 5: Estimation for Portfolio 1 (large capitalization)

Variables	Coefficient	Std. error	z-statistic	Prob.
Constant	-7.444959	0.755091	-9.859680	0.0000
RM	-113.0188	10.53539	-10.72754	0.0000
Size	1.476754	0.492525	2.998331	0.0027
BM	-1.549964	1.041450	-1.488275*	0.1367
Herd	147.5633	12.12592	12.16925	0.0000

(*): is significantly non-significant at 5% level

The final panel logistic regression equation is estimated in general model by using the maximum likelihood estimation:

Zij = -7.444959 - 113.0188***(RM –RF)** + 1.476754***SIZE** - 1.549964 * **BM** + 147.5633***HERD**

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

We note that the statistic Log likelihood is equal to -207.4141. This statistic suppose in the null hypothesis that all coefficients are equal to zero except the constant. Here we reject this hypothesis with zero probability to be wrong. This means that our model is globally significant. To enforce the results we make the wald test which study the same hypothesis. We found Wald chi2(10) equal to 917.0351 (prob =0). So we reject the null hypothesis.

Table 6: Estimation for Portfolio 2 (medium capitalization)

Variables	Coefficient	Std. error	z-statistic	Prob.
Constant	-2.364655	0.401173	-5.894350	0.0000
RM	-141.9050	8.945620	-15.86308	0.0000
Size	0.854690	0.179453	4.762744	0.0000
BM	-0.755664	0.481518	-1.569336*	0.1166
Herd	0.571705	6.881024	0.083084*	0.9338

(*): is significantly non-significant at 5% level

The final panel logistic regression equation is estimated in general model by using the maximum likelihood estimation:

Zij = --2.364655- 141.9050*(**RM –RF)** + 0.854690***SIZE** - 0.755664* **BM** + 0.571705***HERD**

We note that the statistic Log likelihood is equal to -791.8655. Here we reject this hypothesis with zero probability to be wrong. This means that our model is globally significant. To enforce the results we make the wald test which study the same hypothesis. We found Wald chi2(10) equal to 33.0720 (prob =0.0003). So we reject the null hypothesis.

Table 7: Estimation for Portfolio 3 (small capitalization)

Variables	Coefficient	Std. error	z-statistic	Prob.
Constant	-2.392804	0.451838	-5.295711	0.0000
RM	-146.5263	9.079475	-16.13819	0.0000
Size	1.272533	0.203507	6.253010	0.0000
BM	-0.743440	0.390819	-1.902262*	0.0571

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

Herd	-15.77850	7.147378	-2.207592	0.0273

(*): is significantly non-significant at 5% level

The final panel logistic regression equation is estimated in general model by using the maximum likelihood estimation:

Zij = -2.392804- 146.5263***(RM –RF)** + 1.272533***SIZE -** 0.743440* **BM** - 15.77850***HERD**

We note that the statistic Log likelihood is equal to -777.2152. Here we reject this hypothesis with zero probability to be wrong. This means that our model is globally significant. To enforce the results we make the wald test which study the same hypothesis. We found Wald chi2(8) equal to 17.7645 (prob =0.0231). So we reject the null hypothesis.

Using this result we will estimate the value of our dependent variable for the left 762 data in order to test the performance of our model. The result are shown in the following table:

Table 8: Expectation-Prediction Evaluation

	Portfolio 1	Portfolio 2	Portfolio 3
Precision	**0.968504**	**0.703412**	**0.734908**

Here we validate the results of the LR Test and we use the results of our multi factor model for estimation with accuracy of our model is of 96.85% for large capitalization, 70.34% for medium capitalization and 73.49% for small capitalization, which is very important and can help investor to implement the best investment strategy.

3-3-2- VaR forecasting and backtesting

The following figures show the estimated VaR at 95% for the three portfolio.

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

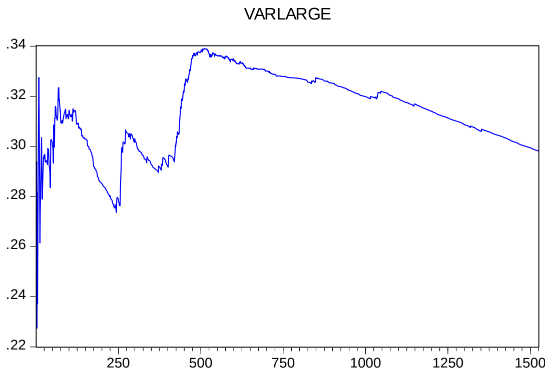

Fig. 1: VaR estimation portfolio 1

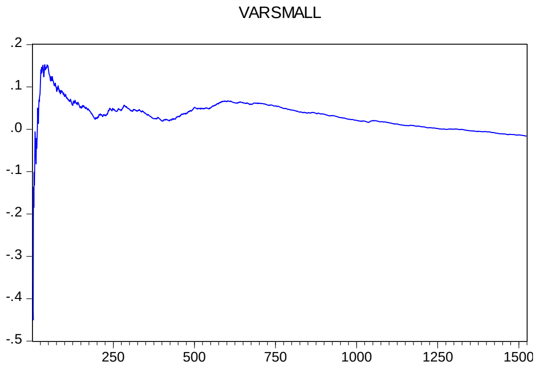

Fig. 2: VaR estimation portfolio 3

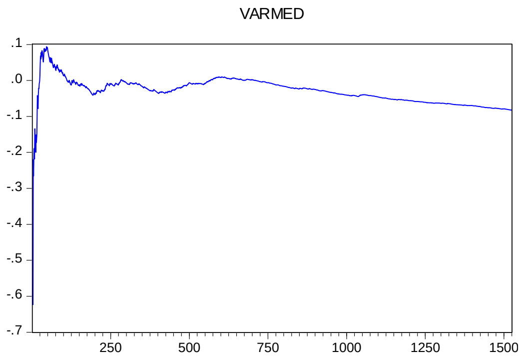

Fig. 3: VaR estimation portfolio 2

In the following table we can compare the VaR of one factor model with the VaR of multi factor model:

Table 9: VaR forecast

	Portfolio 1	**Portfolio 2**	**Portfolio 3**
VaR Markovitz	-5.35%	-5.14%	-4.80%
VaR Multi factro	-3.19%	-3.42%	-3.48%

We can see that the improvement of the VaR forecast is due fact of adding reasonable explanatory variables into the multi factor model.

International Journal of Business Quantitative Economics and Applied Management Research
ISSN: 2349-5677
Volume 2, Issue 11, April 2016

The following table report the results of VaR backtesting:

Table 10: VaR backtesting

	Portfolio 1	Portfolio 2	Portfolio 3
VaR Multi factro	0.0671 accept	0.0442 accept	0.0578 accept

From the table above we conclude that our model is accurate to for the application of stock portfolio VaR forecasting at the quantile level of 5%.

IV. CONCLUSION

In this paper, an attempt is made to explore the use of logistic regression to forecast stock returns for three types of portfolio classified according to size: large capitalization, medium capitalization and small capitalization. We include for variables: market return, size, book to market and herding behavior. Our findings is that our estimation give high degree of precision: 96% for large cap, 73% for medium cap and 70% for small cap.

After that we forecast the VaR using historical simulation for a rolling window of 5 days using one model and multi factor model. We find that the use of more reasonable explanatory variables ameliorate the VaR forecasting. To confirm this result we perform a Backtesting using Kupiec test and we confirm the accuracy of our forecasting.

REFERENCES

[1] Balcilar, M., Demirer, R., & Hammoudeh, S. (2013). Investor herds and regime-switching: Evidence from Gulf Arab stock markets. Journal of International Financial Markets Institutions and Money, 23, 295–321.

[2] Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. Journal of Finance, 55,773–806.

[3] Barber, B. M., Lee, Y.-T., Liu, Y.-J., & Odean, T. (2009). Just how much do individual investors lose by trading? Review of Financial Studies, 22, 609–632.

[4] Barber, B. M., Odean, T., & Zhu, N. (2009). Systematic noise. Journal of Financial Markets, 12, 547–569.

[5] Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom and cultural change as informational cascades. Journal of Political Economy, 100,992–1026.

[6] Bollerslev, T., 1986. Generalized autoregressive heteroskedasticity. Journal of Econometrics, 31, 307-327.

[7] Chang, E. C., Cheng, J. W., & Khorana, A. (2000). An examination of herd behavior inequity markets: An international perspective. Journal of Banking and Finance, 24,1651–1679.

[8] Choi, N., & Sias, R. W. (2009). Institutional industry herding. Journal of Financial Economics, 94, 469–491.

[9] Christie, W. G., & Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market? Financial Analysts Journal, 51, 31–37.

[10] Demirer, R., & Kutan, A. M. (2006). Does herding behavior exist in Chinese stock markets? Journal of International Financial Markets, Institutions and Money, 16,123–142.

[11] Fama, E. F.*;* French, K. R. *(1993). "Common risk factors in the returns on stocks and bonds".* Journal of Financial Economics *33: 3.* doi*:*10.1016/0304-405X(93)90023-5. CiteSeerX*:* 10.1.1.139.5892*.*

[12] Gonzales-Rivera, G., Lee, T.H., Yoldas, E., 2007. Optimality of the RiskMetrics VaR model. Finance Research Letters, 4, 137-145.

[13] H. Markowitz, Portfolio selection*, J. Finance 7 (1952) 77–91.

[14] Hammoudeh, S., Malik, F., and McAleer, M., 2011. Risk management of precious metals. The Quarterly Review of Economics and Finance, 51, 435-441.

[15] Hull, J. C., 2008. Options, Futures, and Other Derivatives, Seventh Edition, (Pearson Education).

[16] Hwang, S., & Salmon, M. (2004). Market stress and herding. Journal of Empirical Finance, 11, 585–616

[17] J. Lintner, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Rev. Econ. Stat. 47 (1965) 13–37.

[18] Kumar, A., & Lee, C. M. C. (2006). Retail investor sentiment and return co-movements. Journal of Finance, 61, 2451–2486.

[19] Kupiec, P.H., 1995. Techniques for Verifying the Accuracy of Risk Measurement Models. The Journal of Derivatives, 3, 73-84.

[20] Lakonishok, J., Shleifer, A., & Vishny, R. W. (1992). The impact of institutional trading on stock prices. Journal of Financial Economics, 32, 23–43.

[21] Nofsinger, J. R., & Sias, R. W. (1999). Herding and feedback trading by institutional investors. Journal of Finance, 54, 2263–2316.Odean, T. (1999). Do investors trade too much?

[22] Odean, T. (1999). Do investors trade too much? American Economic Review, 89,1279–1298.

[23] Ohlson, J. (1980). “Financial ratios and the probabilistic prediction of bankruptcy”. Journal of Accounting Research, 18:109-31.

[24] P. Jorion, Value at Risk: The New Benchmark for Managing Financial Risk, McGraw-Hill, New York, 2007.

[25] Pu Gong, Yingliang Weng, Value-at-Risk forecasts by a spatiotemporal model in Chinese stock market, Physica A 441 (2016) 173–191

[26] R. Engle, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, J. Bus. Econ. Stat. 20 (2002) 339–350.

[27] Rahman, Chowdhury & Sadique (2015), Herding where retail investors dominate trading: The case of Saudi Arabia, The Quarterly Review of Economics and Finance 57 (2015) 46–60

[28] S. Huddart, M. Lang, M.H. Yetman, Volume and price patterns around a stock’s 52-week highs and lows: theory and evidence, Manage. Sci. 55 (2009) 16–31.

[29] S.K. Mishra, et al., Prediction based mean variance model for constrained portfolio assets selection using multi objective evolutionary algorithms, Swarm and evolutionary computation (2016).

[30] Tan, I., Chiang, T. C., Mason, J. R., & Nelling, E. (2008). Herding behavior in Chinese stock markets: An examination of A and B shares. Pacific-Basin Finance Journal,16, 61–77.

[31] Tomohiro Ando and Jushan Bai (2014), Asset Pricing with a General Multifactor Structure, Journal of Financial Econometrics

[32] Tse, Y. K. 1991. Stock Returns Volatility in the Tokyo Stock Exchange. Japan and the World Economy, 3, 285-298.

[33] Tse, Y. K., and Tung, S. H. 1992. Forecasting Volatility in the Singapore Stock Market. Asia Pacific Journal of Management, 9, 1-13.

[34] W.F. Sharpe, Capital asset prices: a theory of market equilibrium under conditions of risk*, J. Finance 19 (1964) 425–442.

[35] X. Zhang, B. Podobnik, D.Y. Kenett, H.E. Stanley, Systemic risk and causality dynamics of the world international shipping market, Physica A 415 (2014) 43–53.

[36] Y. Wei, W. Chen, Y. Lin, Measuring daily value-at-risk of SSEC index: a new approach based on multi-fractal analysis and extreme value theory, Physica A 392 (2013) 2163–2174.

[37] Zugravu, B.; Oanea, D. C.; Anghelache, V. G. 2013, Analysis Based on the Risk Metrics Model, Romanian Statistical Review, Supplement no.2, pp. 145-154.