



Journal of Applied Economics and Business Studies (JAEBS)

Journal homepage: <https://pepri.edu.pk/jaeps>

ISSN (Print): 2523-2614

ISSN (Online) 2663-693X



An empirical testing of comparative efficiency of static and dynamic factor models towards stock returns' predictability in capital market of Pakistan

Laila Taskeen Qazi^{1*}, Atta Ur Rahman², Shahid Ali³ & Sohail Alam⁴

¹ Assistant Professor, Institute of Management Sciences Peshawar.

² Associate Professor, Institute of Management Sciences Peshawar

³ Associate Professor, Institute of Management Sciences Peshawar

⁴ Lecturer, Department of Economics, Abdul Wali Khan University Mardan

ABSTRACT

Efficient Market Hypothesis has its supporters and critics as it has invited significant attention of research scholarship in recent years. The taxonomy and existence of this hypothesis is widely debated in terms of making economic decisions in the capital markets. Stock returns predictability has galvanized researchers to use forecasting models. Literature shows that forecasting is possible yet it debates problems associated with the techniques used for forecasting from the time series data. The study relies on stock returns for 67 randomly selected companies listed on the Pakistan Stock Exchange. The static and the dynamic factor models are compared in terms of forecast efficiency. The study also uses eight macroeconomic variables to forecast stock returns by including gold prices, crude oil prices, market capitalization, PSX- 100 index, PSX-100 index turnover, KIBOR 1-month rates, KIBOR 3 years rates and Rupee to Dollar rates. The results of the hit rates and out-of-sample forecasting technique suggest that dynamic factor model is the best multivariate time series forecasting model in the Pakistani context.

Keywords

Factor models,
ARX models

JEL

Classification

C32, C38, C52,
C58

1. Introduction

Evolution seeds counterrevolution and the same is found out to be true about the efficient market hypothesis in Finance. The scholarly supremacy of efficient market revolution has been significantly challenged by the economists who claim that the stock market returns to a certain stretch are predictable. This predictability argument of researchers runs against Eugene Fama's (1970) market efficiency proposition. In this study, the strength of the predictability argument has been verified which is found to be aligned with the weak form of market efficiency proposed by Fama (1970). Testing the weak form of market efficiency primarily strive to explore how well the past returns can predict the future returns. Fama (1991), in his study of

* attaurrahman@imsciences.edu.pk

Efficient Capital Markets II has put the weak form of market efficiency tests in more general words as ‘tests of returns predictability’. Provision of such tests towards returns predictability formed the basis for this study. Such findings may not be novel, but researching it in the local context may bring exciting findings towards Efficient Market Hypothesis.

Past studies indicate that the use of large models remains limited due to two main visible reasons. Firstly, the use of large models lead to computational burden and secondly, not every time series is found to be informative. Consequently, a large proportion of research in the field of predictability is more focused on the use of relevant information retrieved from large number of variables. Previous research claims that the data declares a factor structure and depicts a common idiosyncratic decomposition, which forms the primary motivation behind the use of factor models in this study for forecasting (Boivin & Ng, 2006; Pesaran & Timmermann, 1995). The dimension reduction concept forms the methodological basis of this study. The use of factor models is found to overcome the issue of extensive parameterization prevalent in case of forecasting through Vector Autoregressive with Exogenous variables (VARX) models. The idea of dimension reduction achieved through factor modeling facilitates efficient multivariate time series forecasting with the minimal loss of degree of freedom and strives to suggest an alternative multivariate time series forecasting model.

The concept of multivariate time series forecasting is applied to data from Pakistani Stock market, which is exposed to several systematic and unsystematic risk factors inducing major inefficiencies in the market. Therefore, in this study both the systematic as well as unsystematic risk factors have been incorporated for multivariate forecasting of stock returns. In order to account for the systematic risk factors, eight macroeconomic variables have been used including gold prices, crude oil prices, market capitalization, PSX-100 index, PSX-100 index turnover, KIBOR 1-month rates, KIBOR 3 year’s rates and Rupee to Dollar rates. Any changes in these macroeconomic factors create an impact on the whole market and serve as the systematic risk factors. For the unsystematic factors the company’s historical data is used to identify a common factor structure, which entails maximum variation in the historical stock returns. Through the use of factor models this study strives to fill the gap in literature in terms of multivariate time series forecasting. Both static and dynamic factor models are employed to create a comparison between the two approaches. In the light of mixed previous findings, this study works toward fresh evidence towards the forecasting the stock returns. The forecasting results of our suggested models will enable this research effort to conclude whether beating the market through forecasts of the stock returns is possible.

2. Literature review

The time series forecasting is a challenging topic with an abundance of literature supporting or negating the notion of forecasting in the light of Efficient Market Hypothesis (EMH)

proposed by Fama (1965). In order to strengthen the argument of the study, this section provides a comprehensive and critical analysis of the past researches.

Fama (1965) for the first time gave the definition of an efficient market, in his landmark research work based on empirical analysis of stock market prices. Fama (1965) concludes that the stock prices follow a random walk. The argument of random walk is further supported by Samuelson (1965), who provides the economic idea of martingale to support the efficiency of markets. This fundamental ideology of efficient markets forms the grounds for EMH and on the basis of this hypothesis, markets are further segregated as weakly or strongly efficient (Robert, 1967). As a proof of market efficiency, Fama et al., (1969) conduct a study and find considerable support to the EMH and conclude that the markets are efficient. (Fama, 1970) conclude that all supporting arguments lead to the development of a definitive understanding regarding the concept of EMH and the categorization of efficient markets emerge as a taxonomy. The literature associated the concept of EMH with that of random walk according to which flow of information is uninterrupted and any new information is immediately reflected in the stock prices therefore future prices must be unpredictable and random (Charlesa & Darnéb, 2009; Gupta & Basu, 2007; Cootner, 1964).

Kemp and Reid (1971) rejects the EMH concluding that share price movements are non-random proposing that revelation of unequal information prevents risk sharing and allows one party in a trading process to either earn abnormal profits or avoid losses. Similarly, several forms of firm announcements also serve as a threat to EMH (Schole, 1972;). In terms of these economic profits Jensen (1978) attempts to define an efficient market as a market where on the basis of an information set, it is impossible to generate economic profits through trading. However, it is impossible for a market to be informationally efficient as generating information is costly and investors tend to spend resources on gathering and processing the information (Charlesa & Darnéb, 2009; Gupta & Basu, 2007; Grossman & Stiglitz, 1980).

It is not only the information set that makes a market inefficient rather the response of share prices towards any particular information also creates room for arbitrage and economic profits. Literature argues that excessive volatility in stock markets rejects EMH (LeRoy & Porter, 1981; Stiglitz, 1981; De Bondt and Thaler, 1985; Roll, 1986). In addition to the stock market volatility and price overreaction, another factor that adds to the market inefficiencies is the duration of overreaction or market volatility in response to any new information (Fama, 1988; Porterbaand & Summers, 1988). Other similar findings tend to reject the random walk hypothesis for the movements in stock returns and establishes the notion of stochastic trends found in the stock returns depicting mean reverting behaviors (Lo & Mackinlay, 1988; Conrad & Kaul, 1988; Lo and Mackinlay, 2011; Bekaert and Hodrick, 1992).

Lo, Mamaysky and Wang (2000) in their study focus on the use of statistical tools and techniques in pursuit of identifying signals in the stock prices that can be used in the technical

analysis to determine the future stock prices. Their findings support the predictability argument. The predictability argument is further supported by Lim & Brooks, (2010), Charlesa & Darnéb, (2009) and Gupta & Basu, (2007). This argument is further strengthened in the literature by the well-known and heavily cited work of Lo & Mckinlay, (1988) and fama (1991). Lo and Makinlay (1988) clearly assert that Stock prices do not follow a random walk and claimed that the stock prices are partially predictable. Lehmann (1990) finds that the predictability in stock markets is attributable to the predictable changes in equity returns, market inefficiencies, and/or to stock price overreactions. Short time intervals such as a week, induces predictability in the stock returns (Lim & Brooks, 2010). Ball (2009) points out the international financial crisis is the best evidence for market inefficiencies. Moreover, incorporating the markets of developing countries, Lee et al., (2010) investigate the stationarity of the stock prices for 32 developed and 26 developing countries and conclude that the stock markets are inefficient as the return predictability is supported by their findings.

Also, there is abundance of literature that finds the notion of predictability through various macroeconomic variables including stock market returns (Schumacher et al., 2006; Chin and Lin, 2011; Ince & Trafali, 2007; Stock and Watson, 2006; Banerjee, et al., 2006; Giannone, et al., 2008). The literature that rejects EMH creates optimism among the researchers and investors for forecasting of stock returns as well as of several other macroeconomic variables such as Gross Domestic Product (GDP), Inflation and Industrial Production etc (Ince & Trafali, 2007). In terms of macroeconomic variables forecasting becomes important due to the fact that the current state of the economy cannot be fairly assessed as the important economic indicators are released with a time lag (Breitung & Eickmeier, 2006). Similarly, investors in stock markets strive to acquire firsthand information as quickly as possible in order to take the arbitrage advantage. Consequently, forecasting allows for the availability of timely information (Schumacher et al., 2006).

Chin and Lin (2011) believe that forecasting in time series through univariate autoregressive model or vector autoregressive models (VAR) are some common options. Amongst these methods, VAR is preferred in terms of forecasting as it allows for the use of larger information set. The inclusion of larger information sets improve the predictability power of the models (Ince & Trafali, 2007). A univariate model allows for the use of small subset of the whole information set. Therefore, a natural innovation to the univariate models is the VAR models. In pursuit of using larger information sets for predictability, the model does not make use of all the available variables in the given model (Boivin & Ng, 2006). Using such models is considered unwise due to the fact that they decrease the degrees of freedom, making the estimation ineffective (Chin & Lin, 2011).

Apart from the size of information needed for forecasting, another significant issue is the fact that the financial times series are usually found to be non-stationary, complex and deterministically chaotic (Boivin & Ng, 2006). To incorporate the use of a larger information

set, the literature stimulates the use of dimension reduction techniques of forecasting (Bovin & Ng, 2006). According to dimension reduction techniques if movements in all macroeconomic variables are driven by a few common factors then those few common factors must be used to forecast the macroeconomic variables (Chin & Lin, 2011).

In recent studies, factor models are quite successful in forecasting especially in case of panel data (Stock and Watson, 2006). The large number of variables used in the literature for forecasting the stock market returns indicate that failing to incorporate any significant variable might encounter the problem of omitted variables (Marcellino, et al., 2003; den Reijer, 2005; Schumacher, 2007). In this regard, factor analysis allows the use of large data sets enabling the researchers to look at everything necessary for forecasting (Breitung & Eickmeier, 2006). The use of factor analysis for forecasting has been significantly confirmed in literature (Stock and Watson, 2006; Banerjee, et al., 2006; Giannone, et al., 2007). The estimated factors from the factor models are used for estimation and inference (Breitung and Eickmeier, 2006; Reichlin, 2003).

Literature also provides evidence on the segregation of factor models into two main forms including Static and Dynamic factor models (Bovin and Ng, 2005; Stock and Watson, 2003). Bovin and Ng (2005) provide a comparison of the static and dynamic factor models for the purpose of forecasting. Factor models are used in literature for forecasting conditional means of market data (Stock and Watson, 2002; Cristadoro, et al., 2001; Artis, et al., 2005; Marcellino, et al., 2003; den Reijer, 2005; Schumacher, 2007, Alonso et al., 2018). Moreover, these models are also used for forecasting conditional volatility. The literature shows the use of factor models in policy analysis as well (Bernanke and Bovin, 2003; Giannone, et al., 2005; Favero, et al., 2005; Stock and Watson, 2005; Forni, et al., 2009). Ludvigson and Ng, (2007) discusses the use of factor analysis for conditioning information and for the term structure analysis.

Forecasting models based on static factors approach are employed by Brisson et al., (2003) to forecast stock returns using the Canadian data. These models are also employed for forecasting of British time series data and the Spanish macroeconomic data (Artis et al., 2005; Camacho and Sancho, 2003).

Regardless of the initial success of the factor models in the field of forecasting, in the literature authors tend to doubt the forecasting results of static factor models employing the use of static component factors. Giacomini and White (2006) assert in their work that static factor models do not depict superior prediction performance in competition to other models under the moving window simulation studies. Similarly Banerjee et al., (2005) compares forecasts made by static factor model and single indicator regression models for the European time series data. The results of comparison do not depict any supremacy of forecasts made by static factor models over the single indicator models. Schumacher and Dreger (2004) use the German stocks

data to analyze the forecasting capabilities of the static factor models and concluded that the use of factor models do not give any better forecasting results than the usual statistical tests.

The apprehensions pertaining to the forecasting ability of the static factor models paves way for the emergence of the use of dynamic factor models. This notion is presented by Forni et al., (2001, 2003a, b) through the estimation of factor models in frequency domain. Forni et al., (2003b) highlight the theoretical significance of dynamic factor models over the static factor models. The proposition that runs behind the dynamic factor models is the fact that dynamic models connect the variables at different points in time, however, only the contemporaneous variables enter the static models. There is abundance of literature that strives to empirically test the forecasting ability of these models. For instance, Den Reiger (2005) confirms the forecasting supremacy of these models by using them to successfully forecast the Dutch macroeconomic variables. Similarly, Kapetanios (2004) confirms the immense forecasting capability of these models by using them to forecast the core index for UK inflation. The popularity of their use increased manifold due to their computational advancements and availability of large panels of time series. Altissimo et al., (2001) use the dynamic factor models for constructing the coincidence indicator for the Euro business area cycle and found out three common factors. Similarly Kabundi (2004) employ the dynamic factor models for forecasting the French business cycle growth rates. Moreover, Eickmeier and Ziegler (2008) use dynamic factor models for forecasting the output and inflation and also conducted the comparison of forecasts made by the dynamic factor model with other forecasting models. Results of their study significantly support the immense forecasting ability of the dynamic factor models.

In spite of theoretical and empirical support for these models, some researchers still argue their use on grounds of robustness and misspecification errors. For instance, Bovin and Ng (2005) assert the significance of static factor models on static principal components claiming that these models are more robust to misspecification errors. For their use, fewer auxiliary parameters are needed to be estimated as compared to the dynamic models.

Due to the mixed results pertaining to the forecasting ability of both the dynamic and static factor models, this study strives to create a comparison between the more sophisticated dynamic factor models proposed by Forni et al., (2001, 2003a, b), and the static factor models proposed by Stock and Watson (2002). This study intends to assess which of the two econometric models does a better job of forecasting stock returns by using data from Pakistani stock markets. In addition, the study works to identify common factors through factor analysis that does forecast the stock returns better. An attempt is made to incorporate certain significant exogenous variables in order to make more accurate forecasts. Several studies in the literature are found to study the relationship between stock prices and fundamental macroeconomic as well as firm specific variables. In terms of firm specific variables these studies use dividend yield, earnings to price ratio and size and claim the predictive power of these variables (Bassu, 1977; Fama & French, 1992). Within the context of several international stock markets,

variables like size, book to market ratio, cash flow yield and earnings to price ratio are found to be significant variables in terms of predictability (Jaffe & Westerfield, 1985; Kato, Zeimba & Schwartz, 1990).

Factor models are considered preferable over the traditional forecasting regressions due to a number of its advantages (Breitung & Eickmeier, 2006, Hallin, & Lippi, 2013). Firstly, factor models can cope with many variables without sacrificing the degree of freedom which is a frequent problem with regressions including more than usual number of variables. Secondly, since the factor models allow for the use of a large number of variables, they enable researchers to exploit more information for making more precise forecasts (Breitung & Eickmeier, 2006). Thirdly, idiosyncratic movements, which possibly include measurement error and the local shocks, can be eliminated (Breitung & Eickmeier, 2006, Qazi et al., 2015). Fourth, forecasting through factor modeling allows the investors to have a more in-depth knowledge over the stock market and relevant strategies can therefore be developed.

In addition to the company specific variables, literature categorizes several macroeconomic variables as the determinants of stock market returns (Humpe & Macmillan, 2009; Paye, 2012; Chen, 2009; Patelis, 2012; Balvers, Cosimano & McDonald, 2012). The most prominent of these variables include oil prices, gold prices, cash reserve ratio, food price inflation, call money rates, dollar rates, foreign direct investments (FDI), foreign portfolio investments and foreign exchange reserves, stock market trade volume, market capitalization and long term and short term interbank offer rates (Ghosh, Bandyopadhyay & Choudhuri, 2011). In the same context, our study checks the predictive power of macroeconomic variables including historical gold rates, dollar rates, 1 month KIBOR rates, 3 years KIBOR rates, PSX-100 Index, PSX market capitalization, PSX Index Turnover and Crude Oil Prices.

The use of factor models to make forecasts is supported by the literature, however, there is little evidence for implementing these models for the stock returns in Pakistan and making less erroneous forecasts. Our study attempts to fulfill this gap, it also creates a comparison to evaluate which of the two types of factor models is more effective in forecasting.

3. Research methods

This study uses exploratory factor analysis for the given time series to identify the most significant factors and to make predictions. The work is relying on meticulously derived econometric models by using historical stock returns data from Pakistan Stock Exchange.

3.1 Sample design

Finding data about stock prices is not a novel thing anymore, as there is abundance of data from stock exchanges around the world. Our econometric model setting demands that our unique approach of forecast efficiency is measured using local data which can be taken to broader data sets in later studies. Pakistan Stock Exchange publishes data on listed Pakistani

companies in routine and has historical datasets that can be used by researchers. At the time of our data acquisition, Pakistan Stock Exchange has 651 listed companies. To assure the selection of a random sample, a pilot sample of 50 stock prices is taken and standard deviation in their stock prices is measured. The standard deviation from stock prices of this pilot sample is recorded as 51.478 which is used along a 95% confidence coefficient and allowable error of 1% to estimate the size of the sample.

The following formula is used to estimate the sample size;

$$n = \left(\frac{z_{\alpha/2} \sigma}{e} \right)^2$$

Our resultant objective sample is of 106 companies, which is initially selected randomly using software from 651 listed companies. But during stock data acquisition it was realized that 39 companies are having significant missing values, which are ultimately dropped from the sample. The final obtained sample is found to be 10% of the population, which is deemed acceptable for further analysis. Next, there is an important consideration of choosing the timeline for the dataset which is selected from December 15, 2016 to October 2018. This timeline was used to minimize the chances of any time related extraneous effects that may affect our modeling and can affect our forecasts as the effort was to see the efficiency of the different factor models in terms of forecasts. Our final sample consists of 489 observations for each company's stock. This enabled our research to rely on 32763 stock return observations for our sample of 67 randomly selected firms.

Relative returns r_t have been calculated at time t and $t-1$ in this study using the following formula,

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$$

The study employs the use of relative returns due to its empirical nature. The relative differences give interest or percentage yield obtained within period $t-1$ to t using daily compounding. Therefore the average returns here are computed as geometric mean of the r_t with daily compounding.

Next, exogenous variables are gathered for the selected timelines using publications of different resources like Pakistan Stock Exchange, Daily Dawn, Ministry of Finance, and State Bank of Pakistan. These include eight independent variables consisting of PSX 100 index, PSX 100 index turnover, PSX 100 index market Capitalization, rupee to dollar rates, gold rates, 1 month KIBOR rates, 3 years KIBOR rates and crude oil prices.

The final sample dataset is organized in spreadsheets for further statistical analysis.

3.2 Static factor models

There is abundance of literature in the area of static factor models (Lawley & Maxwell 1971; Stock & Watson, 1998; Croux, et al. 2004; Deistler & Zinner, 2007). A static factor model may take the following form;

$$y_{it} = \lambda_i F_t + e_{it} \quad (1)$$

In equation. (1) for static factor analysis e_{it} refers to the idiosyncratic error and λ_i means factor loadings.

The model can be further simplified and presented as follows;

$$y_t = \Lambda F_t + e_t \quad (2)$$

Where y_t is the n-dimensional of vector of returns, F_t are the important Principal components (PC), e_t the n-dimensional vector of noise, and $\Lambda \in R^{n \times r}$ is the loading matrix.

The F_t and e_t are uncorrelated, $E(F_t e_t) = 0$

As an autoregressive process with input (ARX) the one step ahead forecast of factor process F_t can be made as under.

$$F_{t+1} = A(z)F_t + D(z)x_t + u_{t+1} \quad (3)$$

$A(z)$ and $D(z)$ are polynomial matrices in the backward shift operator z of order p and q , respectively and the stability condition $\det(I - zA(z)) \neq 0$ for all $|z| \leq 1$ holds.

We assume that u_t is an r -dimensional white noise, x_t is an m -dimensional linearly regular, stationary and ergodic process with mean zero and $E(x_t \epsilon_s) = 0$ for all $t, s \in Z$.

One step-ahead forecasts of y_{t+1} are then obtained as

$$\hat{y}_{t+1/t} = \hat{\Lambda} \hat{F}_{t+1/t} \quad (4)$$

Where $\hat{F}_{t+1/t}$ is the one step ahead forecast of F_t based on equation (3).

3.3 Dynamic factor model

This study investigates the viability of dynamic factor models by comparing two approaches of dynamic factor models including the method of Stock and Watson (2002) and the method of Forni, Hallin, Lippi, and Reichlin (FHLR; 2005).

According to the dynamic factor model proposed by Stock and Watson (2002), X_{it} is the observed data for the financial time series under study at time t , for $i=1,2,3 \dots N$ and $t=1,2,3 \dots t$. here y_t is the time series variable to be forecasted. Considering a dynamic factor model,

$$X_t = b(z)f_t + e_t \quad (5)$$

$$y_{t+h} = \beta_f(z)'f_t + \beta_w'w_t + \epsilon_{t+h} \quad (6)$$

In equation (5) e_t is the idiosyncratic disturbance, h is the forecast horizon, w_t is an $m \times 1$ vector of observed variables that together with the q -dimensional dynamic factors f_t are useful for forecasting the y_{t+h} and ϵ_{t+h} is the resulting forecast error.

According to the method of Forni, Hallin, Lippi, and Reichlin (FHLR; 2005) in the dynamic factor model the observations y_{it} are written as the sum of common components X_{it} and the idiosyncratic component ϵ_{it} . The common component is driven by a q -dimensional vector of common dynamic factors $f_t = (f_{1t}f_{2t} \dots f_{qt})'$ which are loaded with different coefficients and lags. Therefore,

$$y_t = CF_t + \epsilon_t \quad (7)$$

Where $F_t = (f_t, f_{t-1}, \dots, f_{t-s})$ and $C = (B_0, B_1, \dots, B_s)$, and in the above equation (7) is a static factor model with $r = q(s+1)$ dimensional static common factors F_t and q -dimensional dynamic factors f_t .

For the purpose of forecasting let us assume y_t as a stochastic process for which the X_t is a vector of common components and can be assumed with the following equation,

$$\widehat{X}_t = \left[\widehat{\Gamma}_0^X \widehat{Z} (\widehat{Z}' \widehat{\Gamma}_0 \widehat{Z})^{-1} \right] \widehat{Z}' y_t \quad (8)$$

The equation (8) provides consistent in-sample estimators for X_t . The dynamic factor model employs the use of spectral density matrix for in-sample estimators. Moreover out of sample predictors for X_{T+h} are given by the following equation.

$$\widehat{X}_{T+h/T} = \left[\widehat{\Gamma}_h^X \widehat{Z} (\widehat{Z}' \widehat{\Gamma}_0 \widehat{Z})^{-1} \right] \widehat{Z}' y_T \quad (9)$$

Here $\widehat{Z} = (\widehat{Z}_1 \dots \widehat{Z}_r)$ are the r generalized principal components of y_t . Therefore under the dynamic factor model one step ahead forecast of y_{t+1} is given by \widehat{X}_{T+h} .

Providing a brief description of both forms of factor models the study utilized both the models to identify the most relevant common factors creating major impact on the stock prices and on the basis of these factors forecasting for the future stock market were performed.

Under the step 1, model is specified through two simultaneous activities being the estimation of the common components or factors r and the inclusion of the explanatory variables. Here the estimation of the number of common factors is done through two variants of factor analysis being static and dynamic factor models. Within the two models the number of common components r is determined through using the Kaiser Criteria. According to the Kaiser Criteria r is determined on the basis of the number of Eigen values in the correlation matrix having vales greater than 1. The number of r determined in the first step through both static and dynamic factor models are then transferred to step 2.

For the given number of r in step 2, loading and factors are estimated for achieving the one step ahead forecasts for y_t . In this step first PCA is calculated as a static factor model which gives the static factors and spectral density matrix gives the dynamic factors. These factors are then estimated through the ARX model. The explanatory variables inclusion as input to the ARX model and specification of other dynamics are the most important issues in the ARX models. In the present study the selection of explanatory variables is made on the basis of AIC-BIC type information criteria. Those explanatory variables for which the model's AIC-BIC values were found to be the smallest were considered for the final selection for the purpose of forecasting.

After the selection of a final ARX model forecasting has been done in step 3 on the basis of equation (10) given as under,

$$F_{t+1} = A(z)F_t + D(z)x_t + u_{t+1} \quad (10)$$

This one step ahead forecasting has been done using both dynamic as well as static factors. In step 4 the comparison of the forecasting results from different models using the out of sample forecasting technique. According to the out-of-sample forecasting method for y_{t+1} , the sample is divided into two non-overlapping parts i.e. $1, \dots, s$ and $s + 1, \dots, T$. The data falling in the first part is termed as in sample and is used for the estimation and the specification of the model. The second part of the data is termed as out-of-sample and has been used for the evaluation of the models and their forecast quality. The one-step ahead predictor of $y_{i,t+1}$ is $\hat{y}_{i,t+1/t}$ and the equivalent prediction error is $\hat{\epsilon}_{i,t+1/t} = y_{i,t+1} - \hat{y}_{i,t+1/t}$. Following are given different techniques for assessing the forecast quality.

3.4 Hit rate

A simple concept using the following hit rule is applied to the sampled stock return data of 67 companies. Since the observed stock returns data of 489 observations is at hand, which is halved as in-sample and out-of-the sample to monitor the forecast efficiency of our models. The in-sample data is used for estimating the out of sample stock returns at one step only. This suggests that the following conceptual table can be thought of;

Table 1: Hit rate

	Forecasts	
Actual Observations	Profit	Loss
Profit	a	b
Loss	c	d

The accuracy of the forecast is determined by the forecast hit rate which is measured as $(a+d)/(a+b+c+d)$ and expressed as a percentage in the findings section; the following formula is used to measure the same;

$$h_i = \frac{1}{T-s} \sum_{t=s+1}^T \mathbb{1}_{\text{sign}(y_{i,t+1})=\text{sign}(\hat{y}_{i,t+1/t})}$$

Where $\hat{y}_{i,t+1/t}$ is the one-step ahead forecast of $y_{i,t+1}$.

3.5 Coefficient of determination

This relative statistical measure is a basic computation to check for model fit. A model is better fitted if the coefficient of determination is higher having a value from 0 to 100%. The following rule is used to compute its value in our models;

$$R_i^2 = \text{Cor}(y_i \hat{y}_i)^2 * \text{Sign}(\text{Cor}(y_i \hat{y}_i))$$

4. Empirical results

4.1 Comparison of forecasting results

The primary objective of this study is to investigate stock return predictability. Factor models are used for predictability. Furthermore, the study also strives to establish that which factor modeling technique does a better job in forecasting stock returns. Since methodologically the study focuses on testing the forecasting power of factor modeling therefore the study employs the use of two main variants of factor models including static factor models and dynamic factor models. Here the forecasting results of these models have been compared using the technique of out-of-sample forecasting. In the out-of-sample forecasting procedure of y_{t+1} , the sample time period has been divided into two non over lapping time periods. The first time period from December 15, 2016 to April 5, 2018 is the in sample which is 70% of the total sample and is used for the estimation of common factors and for model specification. The remaining 30% of the sample time period, is out-of-sample used for evaluation of the forecasting results. Based on a moving window of 5 days, out-of-sample 1-step-ahead forecasts from April 6, 2018 to October 28, 2018 have been generated.

In the static forecasting model the number of principal components or factors, the lag of factors and the number of lags of exogenous variables have been selected on the basis of Bayesian Information Criterion (BIC). On the basis of BIC criteria the static factor model identifies three principal components significant for forecasting the stock returns. Moreover, the 1-step-ahead out-of-sample forecast generated through static factor model incorporates 3 lags of principal components and exogenous variables basing this selection on the BIC criteria.

Within the dynamic factor models the study tested two approaches, one suggested by Stock and Watson (2002) and the other suggested by Forni et al., (2009). In the dynamic factor models the in-sample estimates of dynamic factors are calculated through the spectral density matrix for each data group and the selection of the common dynamic factors is made on the basis of BIC criteria.

The forecasting quality of the understudy models has been evaluated with respect to out-of-sample hit rate and R^2 . Here the forecasting quality has been evaluated on the basis of positive out-of-sample R^2 and Hit rates greater than 0.5 for all the under study forecasting models. Following a discussion has been provided pertaining to the forecasting results for all 67 companies returns time series arranged in seven groups employed in the study in order to check the multivariate time series forecasting power of the models under study. The forecasting results of each group are provided below.

Table 2: Forecasting results for group 1

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R^2	Hit Rate	SW		FLHR	
				R^2	Hit Rate	R^2	Hit Rate
Financial Services	ARIF HABIB LIMITED	0.000	0.16	0.01	0.36	0.1	0.46
	BIPL SECURITIES LTD.	0.01	0.40	0.01	0.50	0.05	0.52
	DAWOOD EQUITIES LTD.	0.04	0.41	0.00	0.53	0.03	0.53
Banks	BANK AL HABIB	0.06	0.50	0.01	0.55	0.05	0.56
	BANK AL-FALAH LIMITED	0.01	0.49	0.02	0.51	0.04	0.51
	BANK OF PUNJAB	0.04	0.59	0.01	0.56	0.07	0.58
	ASKARI BANK	0.01	0.51	0.03	0.53	0.05	0.55
	FAYSAL BANK	0.02	0.54	0.03	0.54	0.01	0.57
	HABIB METROPOLITAN BANK	0.01	0.53	0.04	0.53	0.01	0.55
	Average	0.0222	0.459	0.0178	0.512	0.0456	0.537

(a) static factor 3.5, mva BIC moving Window, size = 5 (b) SW: q=1, s=2, BIC moving Window, size = 5 (c) FLHR: q=1, s=2, BIC moving Window, size = 5.

Table 2 presents a comparison of forecasting results of both static and dynamic factor models for Group 1. According to the criteria defined for analyzing out-of-sample forecasting quality of a model, the results depict the fractions of positive out-of-sample R^2 for all the companies of Financial Services and Banking sector included in Group 1. The highest average R^2 of 0.0456 has been reported by the FLHR method of dynamic factor models. The average R^2 of static factor (0.0222) model is although better than the Stock and Watson method of dynamic factor models yet it's less than the FLHR method. However in terms of out-of-sample hit rates forecasting quality of the models understudy, both the methods of dynamic factor models outperform the static factor model (Table 1).

The average hit rate of 53.7% under the FLHR method, for the companies of financial services and banking sectors it indicates the supremacy of dynamic factor models in general and the FLHR method in particular over the static factor models. Static factor model in Group 1 reported average hit rate of 45.9% which is indicative of poor forecasting ability of the model. Therefore it can be concluded that in Group 1 for financial services and banking sector companies the dynamic factor model proposed by FLHR method is the best alternate multivariate time series forecasting model.

The forecasting results of Group 2 are reported in Table 3, it can be inferred that again the highest average R^2 of 0.04 has been reported for FLHR method (Table 2). The average R^2 for

the static factor model is 0.023 which is better than the Stock and Watson method of dynamic factor models for which the average R^2 is 0.012. The highest average hit rate of 53.7% for life and non life insurance companies is reported by the static factor models which is greater than the 52.2% average hit rate of dynamic factor models (FLHR method). Therefore R^2 and hit rate criteria provide contradictory results in Group 2. According to R^2 criteria the dynamic factor models give better forecasts as compared to static models whereas hit rates illustrate dominance of static factor models (Table 3).

Table 3: Forecasting results for group 2

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R^2	Hit Rate	SW		FLHR	
				R^2	Hit Rate	R^2	Hit Rate
Non Life Insurance	ADAMJEE INSURANCE	0.01	0.57	-0.01	0.53	0.02	0.53
	CENTRAL INSURANCE	0.04	0.56	0.02	0.51	0.06	0.54
	CENTURY INSURANCE CO.	0.00	0.46	0.02	0.53	0.03	0.53
	EFU GENERAL INSURANCE	0.01	0.57	0.02	0.51	0.06	0.54
	ATLAS INSURANCE	0.02	0.51	0.01	0.43	0.02	0.49
Life Insurance	EFU LIFE ASSURANCE	0.06	0.55	0.01	0.49	0.05	0.5
	Average	0.023	0.537	0.012	0.500	0.040	0.522

(a) static factor 3.5, mva BIC moving Window, size = 5 (b) SW: q=1, s=2, dynamic factors = 2, BIC moving Window, size = 5. (c) FLHR: q=1, s=2, dynamic factors = 3, BIC moving Window, size = 5.

The forecasting results from factor models of Group 3 are reported in Table 4. According to the forecasting results for Group 3, worst forecasting quality is shown by static factor models (Table 4). The average R^2 for static factor model is 0.007 which is far below the average R^2 of 0.016 reported by the dynamic factor models under the FLHR method. In addition in terms of hit rates the static factor model failed to satisfy the minimum criteria of hit rates being greater than 0.5 for quality forecasts (Table 3). Both the variants of dynamic factor models satisfy the hit rate criteria as for Stock and Watson approach the average hit rate for group 3 companies is 50.8%, however the FLHR method does better than the static model and Stock and Watson method of dynamic factors, depicting average hit rate of 52.4%. Therefore the forecasting results of Group 3 categorize dynamic factor models under the FLHR approach as the best alternate multivariate time series forecasting model for predicting stock returns in the PSX.

Table 4: Forecasting results for group 3

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R ²	Hit Rate	SW		FLHR	
				R ²	Hit Rate	R ²	Hit Rate
Personal Goods	GADOON TEXTILE	0.00	0.50	0.01	0.50	0.02	0.52
	COLGATE PALMOLIVE PAK.	0.01	0.50	0.00	0.49	0.01	0.51
	BATA PAKISTAN	0.01	0.50	0.01	0.50	0.01	0.53
	AZGARD NINE	0.00	0.52	0.01	0.51	0.01	0.52
Leisure Goods	GRAYS OF CAMBRIDGE	-0.01	0.49	0.01	0.54	0.01	0.54
Households Goods	AL ABID SILK	0.00	0.36	-0.01	0.51	0.01	0.51
Food Producers	HABIB SUGAR	0.00	0.51	0.02	0.51	0.01	0.54
	HABIB ADM LIMITED	0.01	0.48	0.01	0.50	0.01	0.51
	DEWAN SUGAR	0.04	0.50	0.01	0.51	0.05	0.54
	Average	0.007	0.484	0.008	0.508	0.016	0.524

(a) static factor 3.5, mva BIC moving Window, size = 5, (b) SW: q=1, s=2, dynamic factors = 3, BIC moving Window, size = 5. (c) FLHR: q=1, s=2, dynamic factors = 4, BIC moving Window, size = 5.

Similar forecasting results have been reported for Group 4 (Table 5). On the basis of both the R^2 as well as the hit rate criteria, dynamic factor model under the FLHR method outperforms the static factor model and the dynamic factor model of Stock and Watson (2002). The average R^2 for FLHR method is 0.017 as compared to the average R^2 of 0.01 for static factor model and 0.008 for dynamic factor model of Stock and Watson (2002) (Table 5). Although the R^2 for all the three models under study are positive yet the highest average R^2 has been reported by FLHR method. Likewise the highest average hit rate is reported for FLHR dynamic model being 53.4%. The hit rate report by FLHR dynamic model is evidently better than those reported for static (51.2%) and stock and Watson dynamic factor model (50.3%). On the basis of these findings, for the companies in Group 4, primarily representing the manufacturing sector, it can be inferred that FLHR method of dynamic factor modeling is the best alternate for forecasting the multivariate time series data in the Pakistani Equity market.

Table 5: Forecasting results for group 4

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R ²	Hit Rate	SW		FLHR	
				R ²	Hit Rate	R ²	Hit Rate
Pharma and Bio Tech	ABBOTT LABS.(PAK.)	0.02	0.56	0.01	0.5	0.03	0.56
	FEROZSONS LAB	-0.01	0.44	0	0.45	0.01	0.53
	GLAXOSMITHKLINE PAK.	0.02	0.58	0.02	0.51	0.02	0.58
Chemicals	ARIF HABIB CORPORATION	0.00	0.52	0.01	0.49	0.01	0.51
	CLARIANT PAKISTAN	0.01	0.51	0.01	0.51	0.01	0.54
	DAWOOD HRC.CHEMS.CORP.	0.01	0.53	0.02	0.53	0.02	0.55
	DEWAN SALMAN FIBRE	0.01	0.56	0.01	0.51	0.05	0.56
	ENGRO	0.00	0.5	0.01	0.5	0.02	0.52
	FAUJI FERTILIZER	0.02	0.43	-0.01	0.51	0.01	0.51
	FAUJI FTLZ.BIN QASIM	0.03	0.53	0.01	0.51	-0.01	0.52
	GATRON INDUSTRIES	0.02	0.47	-0.01	0.49	0	0.49
ICI PAKISTAN	-0.01	0.51	0.01	0.52	0.03	0.54	
Average	0.010	0.512	0.008	0.503	0.017	0.534	

(a) static factor 3.5, mva BIC moving Window, size = 5, (b) SW: q=1, s=2, dynamic factors = 3, BIC moving Window, size = 5. (c) FLHR: q=1, s=2, dynamic factors = 5, BIC moving Window, size = 5.

The forecasting results for Group 5 exhibit complete dominance of FLHR dynamic model over the static as well as Stock and Watson dynamic factor model (Table 6). Although the average out-of-sample R^2 for all the three models is positive yet the highest is reported by FLHR dynamic model i.e. 0.02 (Table 6). However, in terms of hit rate criteria, the average hit rates for static factor model as well as for the Stock and Watson dynamic model, is less than 50% which categorizes the forecast quality of these models as lower to FLHR. The average hit rate for FLHR dynamic model is 51.3% (Table 6). On the basis of out-of-sample hit rate and out-of-sample R^2 , FLHR dynamic factor model is sorted out as the best alternate multivariate time series forecasting model for Pakistani Equity market.

Table 6: Forecasting results for group 5

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R^2	Hit Rate	SW		FLHR	
				R^2	Hit Rate	R^2	Hit Rate
Construction and Materials	GHARIBWAL CEMENT	0.01	0.50	0.00	0.50	0.01	0.54
	GAMMON PAKISTAN	-0.01	0.22	-0.01	0.34	-0.01	0.41
	PECTO CEMENT	0.01	0.38	0.01	0.5	0.01	0.5
	FAUJI CEMENT COMPANY	0.01	0.51	0.01	0.5	0.05	0.53
	DG KHAN CEMENT COMPANY	0.02	0.54	0.06	0.54	0.09	0.6
	DEWAN CEMENT	0.00	0.5	0.01	0.51	0	0.5
	DADEX ETERNIT	0.01	0.45	0.01	0.5	0.03	0.53
	DANDOT CEMENT	0.01	0.36	0.02	0.5	0.02	0.5
	DADABHOY CEMENT	0.02	0.52	0	0.5	0.03	0.54
	CHEARAT CEMENT COMPANY	0.01	0.5	0.01	0.54	0.01	0.52
	BESTWAY CEMENT	0.00	0.21	0.03	0.47	0.03	0.49
	BALUCHISTAN GLASS	0.06	0.35	0.01	0.41	0.01	0.45
	ATTOCK CEMENT PAKISTAN	0.03	0.55	0.01	0.56	0.01	0.56
	AL-ABBAS CEMENT	0.04	0.54	0.03	0.5	0.01	0.52
Industrial Metals and Mining	CRESCENT STEEL	0.02	0.55	0.01	0.55	0.02	0.54
	HUFFAZ SEAMLESS PIPE	0.03	0.46	-0.01	0.5	0	0.52
	Average	0.017	0.446	0.013	0.495	0.020	0.513

(a) static factor 3.5, mva BIC moving Window, size = 5, (b) SW: q=1, s=2, dynamic factors = 3, BIC moving Window, size = 5. (c) FLHR: q=1, s=2, dynamic factors = 6, BIC moving Window, size = 5.

The one-step-ahead forecast quality of the understudy models in Group 6 depicts significant win of dynamic factor models over the static factor model (Table 7). The average out-of-sample R^2 for both the variants of dynamic factor models i.e. FLHR method (0.015) and Stock and Watson method (0.017) are positive and greater than that of static factor model (0.011). The average hit rate for the static factor model is 42.1% which is less than 50% therefore the forecast quality of static model is found out to be poor for the return time series of companies belonging to industrial engineering and automobiles and parts sectors (Table 7). In forecasting the multivariate time series for Group 5 the FLHR dynamic model outperformed

the static and Stock and Watson dynamic factor model as the average hit rate for FLHR dynamic factor model is reported to be 53.7% (Table 7).

Table 7: Forecasting results for group 6

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R ²	Hit Rate	SW		FLHR	
				R ²	Hit Rate	R ²	Hit Rate
Industrial Engineering	HINOPAK MOTORS	0.02	0.37	0	0.45	0	0.47
	DEWAN AUTV.ENGR.	0.04	0.25	0.02	0.45	0.02	0.49
	AL-GHAZI TRACTORS	0.01	0.52	0.01	0.53	0.01	0.56
	AL-KHAIR GADOON	0.00	0.14	0.01	0.48	0	0.49
	BOLAN CASTINGS	0.01	0.41	0.04	0.52	0.02	0.54
Automobiles and Parts	AGRIAUTO INDUSTRIES	0.01	0.47	0.02	0.53	0.01	0.57
	ATLAS HONDA	0.01	0.50	0.01	0.51	0.02	0.56
	DEWAN FAROOQUE MOTORS	0.00	0.50	0.01	0.58	0.01	0.58
	GENERAL TYRE & RUBBER	0.02	0.53	0.01	0.5	0.01	0.53
	HONDA ATLAS CARS (PAK.)	-0.01	0.52	0.04	0.56	0.05	0.58
	Average	0.011	0.421	0.017	0.511	0.015	0.537

(a) static factor 3.5, mva BIC moving Window, size = 5. (b) SW: q=1, s=1, dynamic factors = 2, BIC moving Window, size = 5. (c) FLHR: q=1, s=2, dynamic factors = 4, BIC moving Window, size = 5.

The forecasting results of Group 7 are given in Table 8 below. The forecasting result for Group 7 also categorize FLHR dynamic model as the best alternate multivariate time series forecasting model which is attributable to the fact that the average hit rate of FLHR model in Group 7 is 51% (Table 8). The average hit rate reported by the FLHR model is better than the average hit rate of 50% reported by the Stock and Watson dynamic model and 48.2% reported by the Static Factor model (Table 8). Moreover these results are further confirmed by the positive fractions of out-of-sample R^2 . The average out-of-sample R^2 is also reported highest by the FLHR dynamic factor model in Group 7.

Table 8: Forecasting results for group 7

Sector	Company Names	Static Factor Models		Dynamic Factor Models			
		R ²	Hit Rate	SW		FLHR	
				R ²	Hit Rate	R ²	Hit Rate
Electricity	HUB POWER COMPANY	0.04	0.50	0.01	0.53	0.02	0.52
Forestry and Paper	CENTURY PAPER	0.01	0.47	0.01	0.47	0.06	0.5
Oil and Gas Devp.	ATTOCK PETROLEUM	0.02	0.51	0.00	0.49	0.02	0.5
	ATTOCK REFINERY	0.01	0.50	-0.01	0.51	0	0.53
	BYCO PETROLEUM PAKISTAN	0.02	0.43	0.00	0.50	0.01	0.5
	Average	0.020	0.482	0.002	0.500	0.028	0.510

(a) static factor 3.5, mva BIC moving Window, size = 5. (b) SW: q=1, s=1, dynamic factors = 2, BIC moving Window, size = 5. (c) FLHR: q=1, s=2, dynamic factors = 4, BIC moving Window, size = 5.

Overall the forecasting results for all the groups suggest the dynamic factor models in general and the FLHR method of dynamic factor modeling as the best alternate model for the multivariate time series forecasting. For all the groups the FLHR model reported positive average out-of-sample R^2 and hit rates greater than 0.5 or 50%. Therefore on the basis of these

two criteria the forecast quality of dynamic factor models under the FLHR method is found out to dominate the other two models under study.

Eight variables are included in this study consisting of historical gold rates, dollar rates, 1 month KIBOR rates, 3 years KIBOR rates, PSX-100 Index, PSX market capitalization, PSX Index Turnover and Crude Oil Prices. Along with the common factors these variables are incorporated in the forecasting model for the out-of-sample data. The out-of-sample forecasting results depict convergence between the literature and the empirical findings of this study. Several variables have been found significant in the forecasting of stock returns. The 3-yrs KIBOR rates are found to be significant on the 1st lag whereas 1-month KIBOR rates are found significant on the 3rd lag. The returns of PSX-100 index are found significant in forecasting the company's stock returns on the 4th lag. Similarly crude oil prices, PSX-100 index turnover and dollar rates become significant on the 4th lag which means the impact of these variables become evident on stock returns over time period. In addition the historical Gold rates are found to be significant on the 2nd lag. Therefore it can be concluded that out of the 8 exogenous variables incorporated in the study seven are found to be significant in forecasting the stock return.

5. Conclusion

This study utilizes factor models within the framework of dimension reduction in forecasting. Within this context, this study focuses on creating a comparison of various forms of factor models in order to suggest a best fit model for forecasting multivariate time series data. The study tests forecasting power of static factor models and dynamic factor models and within the dynamic factor models the study compares two approaches, one proposed by Stock and Watson (2002) and the other proposed by Forni, et al., (2005). In this study, the forecasting quality of each model is analyzed through the use of out-of-sample forecasting technique. Two main criteria were applied for analyzing the forecast quality of all the considered models including positive out-of-sample R^2 and out-of-sample hit rates greater than 50%.

The empirical findings of the study categorizes dynamic factor models under the method proposed by Forni, et al., (2005) (FLHR method). The average R^2 for all the data groups for dynamic models were positive fractions with all the average hit rates greater than 50%.

Overall this study can prove to be helpful for the stock market investors and analysts as it attempts to provide a robust forecasting model for the purpose of effective forecasting of financial time series data. Moreover, the study also includes the macroeconomic variables for the purpose of forecasting which broadens the scope of the study and the applicability of the forecasting results of the study. The empirical findings of the study challenges the Efficient Market Hypothesis and finds it less useful in the Pakistani context. The study supports the weak form of EMH proposed by Fama (1991) which entails predictability and also recommends dynamic factor models as an efficient tool to forecast the stock returns.

References

- Alonso, A. M., Bastos, G., & García-Martos, C. (2018). Bias correction for time series factor models. *Journal of Statistical Computation and Simulation*, 88(8), 1576-1602.
- Artis, M., Banerjee, A. & Marcellino, M. (2005). 'Factor forecasts for the UK', *Journal of Forecasting*, 24(4), . 279–298.
- Ball, R. (2009). The global financial crisis and the efficient market hypothesis: What have we learned? *Journal of Applied Corporate Finance*, 21(4), 8–16.
- Balvers, R. J., Cosimano, T. F. & McDonald, B. (2012). Predicting Stock Returns in an Efficient Market, *The Journal of Finance*, 45(4), 1109-1128..
- Banerjee, A., Marcellino, M., & Masten, I. (2005). Leading indicators for euro-area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics*, 67, 785-813.
- Banerjee, A., Masten, I. & Massimiliano, M. (2006). Forecasting macroeconomic variables using diffusion indexes in short samples with structural change,
- Bekaert, G. & Hodrick, R. J. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets, *The Journal of Finance*, 47(2), 467–509.
- Bernanke, B. & Boivin, J. (2003). Monetary policy in a data rich environment, *Journal of Monetary Economics*, 50(3), 525–546.
- Boivin, J. & Ng, S. (2006). Are more data always better for factor analysis?. *Journal of Econometrics*, 132(1) 169–194.
- Breitung, J., & Eickmeier, S. (2006). Dynamic factor models. *Allgemeines Statistisches Archiv*, 90(1), 27-42.
- Brisson, M., Campbell, B., & Galbraith, J. W. (2003). Forecasting some low-predictability time series using diffusion indices. *Journal of Forecasting*, 22(6-7), 515-531.
- Business Recorder, (2012). Karachi Stocks, *Business Recorder*, retrieved on July 15, 2012 from, <http://www.brecorder.com/market-data/karachi-stocks/>
- Camacho, M., & Sancho, I. (2003). Spanish diffusion indexes. *Spanish economic review*, 5(3), 173-203.
- Charles, A., & Darné, O. (2009). The random walk hypothesis for Chinese stock markets: Evidence from variance ratio tests. *Economic Systems*, 33(2), 117-126.
- Chen, S. S. (2009). Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking & Finance*, 33(2), 211-223.
- Chen, C. C., & Lin, H. W. (2011). The advantages of dynamic factor models as techniques for forecasting: evidence from taiwanese macroeconomic data. *International Journal of Economics and Finance*, 3(5), 208-215.
- Cootner, P. H. (ed.) (1964). *The Random Character of Stock Market Prices*, The MIT Press, Cambridge, MA.
- Cristadoro, R., Forni, M., Reichlin, L. & Giovanni, V. (2001). A core inflation index for the euro area,
- Croux, C., Renault, E., & Werker, B. (2004). Dynamic factor models. *Journal of econometrics*, 119(2).

- Conrad, J., & Kaul, G. (1988). Time-variation in expected returns. *Journal of business*, 409-425.
- De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, 40(3), 793-805.
- Deistler, M., & Zinner, C. (2007). Modelling high-dimensional time series by generalized linear dynamic factor models: An introductory survey. *Communications in Information & Systems*, 7(2), 153-166.
- Eickmeier, S., & Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27(3), 237-265..
- Fama, E. (1965). The behavior of stock market prices *Journal of Business*, vol. 38. 34-105
- Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International economic review*, 10(1), 1-21.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.
- Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2), 246-273.
- Favero, C. A., Marcellino, M., & Neglia, F. (2005). Principal components at work: the empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, 20(5), 603-620.
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2001). Coincident and leading indicators for the euro area. *The Economic Journal*, 111(471), C62-C85.
- Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2003a). Do financial Variables help forecasting inflation and real activity in Euro Area? *Journal of Monetary Economics*, 50(6), 1243-1255.
- Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2003b). The Generalized Dynamic Factor Model: One sided Estimation and forecasting, Mimeo.
- Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2005). The generalized dynamic factor model: one-sided estimation and forecasting, *Journal of the American Statistical Association*, 100(471), 830-840.
- Forni, M., Giannone, D., Lippi, M., & Reichlin, L. (2009). Opening the black box: identifying shocks and propagation mechanisms in VAR and factor models. *Econometric Theory*, 25(1319), U1347.
- Giannone, D., Reichlin, L., & Sala, L. (2004). Monetary policy in real time. *NBER macroeconomics annual*, 19, 161-200.
- Giannone, D., Reichlin, L., & Sala, L. (2006). VARs, common factors and the empirical validation of equilibrium business cycle models. *Journal of Econometrics*, 132(1), 257-279.
- Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665-676.

- Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6), 1545-1578.
- Grossman, S. J. & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets, *The American Economic Review*, 70(3), 393–408.
- Gupta, R., & Basu, P. K. (2007). Weak form efficiency in Indian stock markets. *International Business & Economics Research Journal (IBER)*, 6(3)..
- Hallin, M., & Lippi, M. (2013). Factor models in high-dimensional time series—A time-domain approach. *Stochastic processes and their applications*, 123(7), 2678-2695.
- Humpe, A., & Macmillan, P. (2009). Can macroeconomic variables explain long-term stock market movements? A comparison of the US and Japan. *Applied Financial Economics*, 19(2), 111-119..
- Ince, H., & Trafalis, T. B. (2007). Kernel principal component analysis and support vector machines for stock price prediction. *Iie Transactions*, 39(6), 629-637.
- Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of financial economics*, 6(2/3), 95-101.
- Kabundi, A., (2004). Estimation of Economic Growth Using Business Survey Data. *International Monetary Fund*, 22 Working Paper No. 04/69.
- Kemp, A. G. & Reid, G. C. (1971). The random walk hypothesis and the recent behavior of equity prices in Britain, *Economica*, 38(149), 28–51.
- Lawley, D. N. & Maxwell, A. E. (1971). *Factor Analysis as a Statistical Method*, 2nd ed., London: Butterworth and Co.
- LeRoy, S. F. & Porter, R. D. (1981). The present-value relation: Tests based on implied variance bounds, *Econometrica*, 49(3), 555–574.
- Lee, C.-C., Lee, J.-D. & Lee, C.-C. (2010). Stock prices and the efficient market hypothesis: Evidence from a panel stationary test with structural breaks, *Japan and the World Economy*, 22(1), 49–58.
- Lehmann, B. N. (1990). Fads, Martingales, and Market Efficiency, *The Quarterly Journal of Economics*, 105 (1), 1-28.
- Lo, A. W. & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test, *The Review of Financial Studies*, 1(1), 41–66.
- Lo, A. W. & MacKinlay, A. C., (2011). *A Non-Random Walk Down Wall Street*, Princeton: Princeton University Press.
- Ludvigson, S. & Ng, S. (2007). The empirical risk return relation: A factor analysis approach. *Journal of Financial Economics*, 83(1), 171–222.
- Marcellino, M., Stock, J. H. & Watson, M. (2003). Macroeconomic forecasting in the Euro area: Country specific versus Euro wide information. *European Economic Review*, 47(1), 1–18.
- Paye, B. S. (2012). ‘Déjà vol’: Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics*, 106(3), 527-546.

- Pesaran, M. H. & Timmermann, A. (1995). Predictability of Stock Returns: Robustness and Economic Significance, *The Journal of Finance*, 50(4), 1201-1228.
- Qazi, L. T., Rahman, A. U., & Gul, S. (2015). Which pairs of stocks should we trade? Selection of pairs for statistical arbitrage and pairs trading in Karachi Stock Exchange. *The Pakistan Development Review*, 215-244.
- Reichlin, L. (2003). Factor models in large cross sections of time series. In: S. T. M. Dewatripoint and L. P. Hansen (eds.): *Advances in Economics and Econometrics: Theory and Applications*, Vol. 111, 8th World Congress of the Econometric Society. Cambridge: Cambridge University Press.
- Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly, *Industrial Management Review*, 6(2), 41-49.
- Schumacher, C. & Dreger, C. (2004). Estimating Large Scale factor models for economic activity in Germany; do they outperform simpler models? *Journal of National Economy and Statistics*, 224(6), 731-750.
- Schumacher, C. & Breitung, J. (2006). Real-Time forecasting of GDP based on a large factor Model with Monthly and Quarterly Data,
- Schumacher, C. (2007). Forecasting German GDP using alternative factor models based on large datasets. *Journal of Forecasting*, 26(4), 271-302.
- Stock, J. H., & Watson, M. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3), 788-829.
- Stock, J. H. & Watson, M. (1998). Diffusion indexes. NBER working paper 6702.
- Stiglitz, J. E. (1981). The allocation role of the stock market: Pareto optimality and competition, *The Journal of Finance*, 36(2), 235-251.
- Stock, J. H., & Watson, M. W. (2006). Forecasting with many predictors. *Handbook of economic forecasting*, 1, 515-554.
- Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147-162.
- Stock, J. H. & Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis. NBER Working Paper, 11467.