Predicting Returns with Financial Ratios: Evidence from Greece

Christos Floros[1], Shabbar Jaffry and Yasseen Ghulam

Department of Economics, University of Portsmouth, Portsmouth Business School, Portland Street, Portsmouth, PO1 3DE, UK.

Abstract

This article investigates whether financial ratios (dividend yield, price earnings ratio, book to market value) can predict aggregate stock returns. We report a forecasting competition between single and multiple OLS, GARCH and ECM-GARCH regressions of the Greek return series. First, we test the out-of-sample forecasting accuracy, and then we compare the forecasting techniques based on the symmetric error statistics under both static and dynamic methods. The results show that ECM-GARCH(1,1) model has significant coefficients. Both static and dynamic forecasts confirm that ECM-GARCH(1,1) is the most appropriate model for forecasting returns during the period January 2003 - December 2004.

JEL classification: C32; G12.

Keywords: Financial Ratios, GARCH, ECM, Forecasting

[1] Author for correspondence: Christos.Floros@port.ac.uk, Tel.: +44 2392 844244, Fax: +44 2392 844037.

1. Introduction

Investor ratios are used in finance for assessing the effects of proposed financing, valuing a target company in a takeover target, analysing dividend policy, but also predicting the effect of a rights issue. The ratios that relate company fundamentals to market prices - such as dividend yield (DY), price earnings ratio (P/E), book to market value (PBV), or sales to market value (SMV), are helpful for companies that do not show positive financial results but investors are already discounting future earnings according to their own expectations. For companies belonging to the same activity sector, high ratios represent a good buying opportunity since the stock is apparently under-priced in comparison to other companies. On the contrary, low ratios suggest that the stock is over priced and it may incentive investors to sell their positions.

Since some investors use the relative pricing feature on stock pricing, it is expected that financial returns can be related to these ratios. It is also likely that investors may have individual thresholds or ratio levels above (under) which they start buying (selling) the stocks.

Financial researchers use these ratios to predict stock returns under different methods. Recent studies show that financial ratios share many time-series statistical properties with the financial returns series (because prices are included on the denominator). This fact has implications on the quality of the estimators given by the simple Ordinary Least Squares (OLS) method.

The existing studies focus mainly on the US and other European markets. Ang and Bekaert (2002) analyse US, UK, French, German and Japanese markets. They perform a cross sectional analysis and suggest that predictability of financial returns through dividend yields, earnings yield and short term bonds yields, is mainly a short term phenomenon. They find that the strongest predictor is the short-term interest rate. In addition their results show that cross-country predictability is stronger than domestic predictability.

Bekaert and Hodrick (1992) find evidence of predictability of excess returns through yield variables in the foreign exchange market. Campbell and Shiller (1988) find that the logged dividend yield moves with expected future growth of dividends, while that it granger causes the real dividend growth.

Goyal and Welch (2003) conclude that out-of-the-sample sum squared residuals can be a powerful diagnostic check for equity premium and stock price prediction. For simple dividend yield models, their conclusion is that good in-sample performance does not ensure a good out of sample performance. To justify the poor prediction, they identify that the primary cause is due to parameter instability. In addition the dividend yield fails to forecast one-year ahead returns and dividend growth changes because first it forecasts its own change. When they develop the model to account for the time-varying pattern of the dividend ratio and the dividend growth properties, it does not increase the predictability of stock market levels through the dividend yield.

Lewellen (2004) suggests a new test of predictability of financial ratios that controls for the strong autocorrelation on the predictive variables. He finds strong evidence that the dividend yield predicts the value and equal weighted NYSE returns for the period 1946-2000. Nelson and Kim (1993) argue that predictive regressions are subject to small sample biases, and the coefficients become biased whenever the predictor is endogenous, while the standard errors are under estimated in case of overlapping periods. Both biases result in increased t-ratios that suggest predictability (even when they are not present).

This paper investigates whether financial ratios (DY, PBV, P/E) can predict stock returns using recent data from Greece. To the best of our knowledge, this is the first empirical paper on this topic for Greece. The paper proceeds as follows. Section 2 presents the methodology and describes the data, while Section 3 presents the empirical results. Section 4 summarizes the main findings and concludes the paper.

2. Methodology and Data

Suppose *X* denotes a financial ratio- DY, PBV or P/E- we employ three different models for predicting returns. Model 1 shows that the predicted variable is assumed to follow a stationary AR(1) process. Random walk model (Model 1) is given by:

$$\log(X_t) = c + b\log(X_{t-1}) \qquad (1)$$

where b<1. Under the same assumptions[2], model 2 presents the regression of returns (R) on the predictive variable. Here, β should be zero if expected returns are constant (alternatively, β>0).

$$R_t = \alpha + \beta\log(X_{t-1}) \qquad (2)$$

GARCH methodology

Furthermore, we model the non-constant volatility parameter using a Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model. GARCH(1,1) is parsimonious (the coefficients of the model are easily interpreted) and gives significant results, since it allows the conditional variance of a stock price or index to be dependent upon previous own lags. The GARCH (1,1) model is given by

$$\begin{aligned} R_t &= \mu + c\log(X_{t-1}) + \varepsilon_t && \text{(Mean Equation)} \\ \sigma_t^2 &= \omega + a\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 && \text{(Variance Equation)} \end{aligned} \qquad (3)$$

where the error, ε_t, is assumed to be normally distributed with zero mean and conditional variance, σ_t^2. R_t are returns, so we expect their mean value (which will be given by μ) to be positive and small. We also expect the value of ω again to be small. All parameters in variance equation must be significant, and the sum of α and β is expected to be less than, but close to, unity, with $\beta > \alpha$. News about volatility from the previous period can be measured as the lag of the squared residual from the mean equation (ARCH term). Also, the estimate of β shows the persistence of volatility to a shock or, alternatively, the impact of old news on volatility.

VAR and Cointegration

Recent empirical methods on testing the interdependence among the financial series include the Vector Autoregressive (VAR), Error Correction Model (ECM) and Johansen cointegration technique. If series are cointegrated, then they exhibit a long run relationship. We use the Johansen maximum likelihood approach (Johansen, 1988) to test for the evidence of cointegration among the variables. The basis of this approach is to estimate by maximum likelihood methods an equation of the form:

$$\begin{aligned} \Delta X_t &= \Pi X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} + \mathrm{B} y_t + \varepsilon_t \\ \Pi &= \sum_{i=1}^{p} A_i - I, \Gamma_i = -\sum_{j=i+1}^{p} A_j \end{aligned} \qquad (4)$$

where the i determines the number of lags specified in the dynamic VAR relationship. X_t is a column vector of the three indices. If Π has zero rank, no stationary linear combination can be identified, so the variables in X_t are non-cointegrated. However, if the rank is r, there will be r possible stationary linear combinations (relationships). According to Johansen (1988), when there are k series, then up to k-1 cointegrating vectors can exist. There are two test statistics for testing the null hypothesis of no cointegration under the Johansen approach, the trace and maximum eigenvalue statistic tests, see Floros (2003, 2005a, 2005b).

If series are functioning properly, price movements should be best described by error correction model (ECM) with the error correction term being the price differential between the series. The ECM is a dynamic correlation model of returns. It reveals the causality that must be present when there is cointegration. The ECM is a VAR on first differences, augmented with the stationary disequilibrium term. Consider X = ln(ratio) and P = ln(Price), then

$$R_t = a + \sum \beta_i \Delta X_{it} + \beta_2 (P_{t-1} - \sum \beta_i (X_{t-1})) + v_t \quad (5)$$

[2] For properties (assumptions) of OLS, see Lewellen (2004).

In the ECM equation, the differenced terms describe the short term dynamics while the lagged value of the error of the cointegration regression describes the long run equilibrium relationship. The inclusion of equilibrium terms in ECMs ensures that no information on the levels of variables is ignored. Also, the clear distinction between short- and long-run effects is itself a further advantage of the ECM approach, while it always reduce problems of multicollinearity.

Our (aggregate) data is collected from *DataStream*. The monthly observations of the Greek stock index (FTSE/ASE-20) start in 01/02/1992 and end at 01/12/2004, resulting in 155 observations. The Financial ratios, P/E, DY and PBV were computed through the value weighted-average of each stock listed on each index. Table 1 shows the descriptive statistics for DY, PBV, P/E and price stock index (Price). The Normality tests (Jargue-Bera) show that all series have skewed distributions. The stationarity results from ADF tests (not reported here[3]) confirm that all variables are integrated of order one, *I(1)*.

Table 1. Descriptive Statistics for Series (in Levels)

	DY	PBV	P/E	PRICE
Mean	3.073801	4.401106	35.18618	12.61784
Median	3.160241	3.783710	20.63785	12.70947
Maximum	3.203078	15.09553	272.0319	13.36743
Minimum	2.066556	1.855827	6.207808	6.951250
Std. Dev.	0.216319	2.472265	53.94945	0.913878
Skewness	-2.687932	2.022676	3.423747	-3.892747
Kurtosis	10.23383	7.430803	13.34714	20.40034
Jarque-Bera	524.5987	232.4799	994.2694	2346.867
Probability	0.000000	0.000000	0.000000	0.000000
Observations	155	155	155	155

3. Empirical Results

To get a clear view and in-depth comparison of forecasting models, we divide the full sample (March 1992 - December 2004) into in-sample (March 1992-December 2002) and out-of-sample (forecasting) period (January 2003 - December 2004). First, we apply two simple models using the Least Squares method, i.e. model 1 and model 2. We also estimate a GARCH(1,1) using the Marquardt algorithm and Heteroskedasticity consistent covariance option. Furthermore, we test if series are cointegrated (see Lewellen, 2004) using the test of Johansen (1988). Finally, a simple ECM-GARCH(1,1) model is applied in order to capture both equilibrium relationship and volatility in the series examined.

[3] These results are available upon request.

Predicting with DY

Dividend yield gives a measure of how much an investor expects to receive in exchange for purchasing a given share (Dividend per share/Market price of share). DY gives some idea of the rate of return that the dividend represents. Table 2 shows the predictive power of DY using in-sample (1992-2002). Equation 1 shows an OLS slope of 0.91 (part A) with a standard error of 0.002, while R-sq is close to unity providing strong evidence of predictability. Also, the slope coefficient in equations (2) and (3) is positive and significant (i.e. there is a positive effect of lagged DY to return, R). GARCH coefficient is significant , so there is impact of old news on volatility.

Table 2 Dividend yield and returns, 1992-2002

Part A. Equation (1)

Dependent Variable: LOG(DY)
Method: Least Squares
Sample: 1992M03 -2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
c	0.099202	0.002330	42.56883*	0.0000
LOG(DY_{t-1})	0.913790	0.002060	443.6102*	0.0000
R-squared	0.999822			

PART B. Equation (2)

Dependent Variable: R
Method: Least Squares
Sample: 1992M03 -2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
a	-0.183694	0.101633	-1.80742**	0.0730
LOG(DY_{t-1})	0.168068	0.092434	1.818251**	0.0714

PART C. Equation (3)

Dependent Variable: R
Method: ML - ARCH (Marquardt) - Normal distribution
Sample: 1992M03 -2002M12
Bollerslev-Wooldrige robust standard errors & covariance

	Coefficient	Std. Error	z-Statistic	Prob.
μ	-0.202401	0.081485	-2.483899*	0.0130
LOG(DY_{t-1})	0.183715	0.074119	2.478651*	0.0132
	Variance Equation			
ω	0.000696	0.001836	0.379346	0.7044
ARCH	0.084298	0.119295	0.706638	0.4798
GARCH	0.854494	0.270181	3.162678*	0.0016

* Significant at 5% level, ** Significant at 10% level

Predicting with PBV

Table 3 presents the results from in-sample. Equation 1 shows an OLS slope coefficient of 0.97 (part A) with a standard error of 0.028, while R-sq is close to unity providing strong evidence of predictability. The slope coefficient of equation (2) and equation (3) is negative but not significant. GARCH (but not ARCH) coefficient is significant at 5% level, so GARCH(1,1) model captures volatility in returns.

Table 3. Book-to-market and returns, 1992-2002

PART A. Equation (1)

Dependent Variable: LOG(PBV)
Method: Least Squares
Sample: 1992M03 -2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
c	0.047401	0.039814	1.190552	0.2360
LOG(PBV_{t-1})	0.966546	0.027629	34.98298*	0.0000
R-squared	0.926742			

PARTB. Equation (2)

Dependent Variable: R
Method: Least Squares
Sample: 1992M03 -2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
a	0.006600	0.028101	0.234888	0.8147
LOG(PBV_{t-1})	-0.002361	0.019607	-0.120436	0.9043

PARTC. Equation (3)

Dependent Variable: R
Method: ML – ARCH (Marquardt) - Normal distribution
Sample: 1992M03 -2002M12
Bollerslev-Wooldrige robust standard errors & covariance

	Coefficient	Std. Error	z-Statistic	Prob.
μ	0.015726	0.022318	0.704613	0.4811
LOG(PBV_{t-1})	-0.010357	0.015840	-0.653807	0.5132
	Variance Equation			
ω	0.000567	0.001552	0.365144	0.7150
ARCH	0.121244	0.210229	0.576722	0.5641
GARCH	0.831857	0.305299	2.724726*	0.0064

* Significant at 5% level

Predicting with P/E

Price/Earnings ratio is a key ratio by stock market investors. It shows how much an investor is prepared to pay for a company's shares, given its current earnings per share (Market price per share/Earnings per share). The ratio indicates the confidence of investors in the expected future performance of a company. The higher P/E ratio, the more confident the market is that future earnings will increase (shares with large P/E are those that are highly priced for their historical earnings level). Table 4 shows the results using in-sample. Equation 1 shows a positive effect on R (significant slope coefficient), while equation 2 and equation 3 show negative slope coefficients. GARCH coefficient in equation 3 is positive and significant.

Table 4. Price-Earnings ratio and returns, 1992-2002

PART A. Equation (1)

Dependent Variable: LOG(P/E)
Method: Least Squares
Sample: 1992M03 -2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
c	0.171657	0.130602	1.314347	0.1911
LOG(P/E$_{t-1}$)	0.950806	0.036848	25.80337*	0.0000
R-squared	0.800191			

PARTB.Equation(2)

Dependent Variable: PRICE-PRICE(-1)
Method: Least Squares
Sample: 1992M03 -2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
a	0.023650	0.040629	0.582109	0.5615
LOG(P/E$_{t-1}$)	-0.006888	0.013844	-0.497545	0.6197

PART C. Equation (3)

Dependent Variable: R
Method: ML - ARCH (Marquardt) - Normal distribution
Sample: 1992M03 -2002M12
Bollerslev-Wooldrige robust standard errors & covariance

	Coefficient	Std. Error	z-Statistic	Prob.
μ	0.046712	0.033298	1.402861	0.1607
LOG(P/E$_{t-1}$)	-0.015616	0.011350	-1.375841	0.1689
	Variance Equation			
ω	0.000485	0.001197	0.405342	0.6852
ARCH	0.145398	0.237192	0.612999	0.5399
GARCH	0.819602	0.285052	2.875271*	0.0040

* Significant at 5% level

Predicting Returns using Multiple Regressions

We also run multiple regressions to test if financial ratios have an effect on the returns. Equation (2) can be written as

$$R_t = \alpha + \beta \sum \log(X_{t-1}) \qquad (6),$$

while equation (3) is now given by

$$R_t = \mu + c\sum \log(X_{t-1}) + \varepsilon_t \quad \text{(Mean Equation)}$$
$$\sigma_t^2 = \omega + a\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \quad \text{(Variance Equation)} \qquad (7)$$

The results from equation (6) and equation (7) are presented in Table 5. Multiple OLS regression (Part A) shows significant coefficients for DY and P/E, but the coefficient of PBV is negative and not significant. (i.e. there is no effect on returns). Multiple OLS with GARCH errors (Part B) show similar results, while GARCH coefficient is positive and significant.

Table 5. Multiple Regression results

Part A

Dependent Variable: R
Method: Least Squares
Sample (adjusted): 1992M03 2002M12
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.257726	0.096530	-2.669890	0.0086*
LOG(GDY $_{t-1}$)	0.325661	0.104431	3.118417	0.0023*
LOG(GPBV $_{t-1}$)	-0.013463	0.020211	-0.666148	0.5065
LOG(GPE $_{t-1}$)	-0.027415	0.014164	-1.935502	0.0552*

PART B

Dependent Variable: R
Method: ML - ARCH
Sample (adjusted): 1992M03 2002M12
Bollerslev-Wooldrige robust standard errors & covariance

	Coefficient	Std. Error	z-Statistic	Prob.
μ	-0.268182	0.076905	-3.487181	0.0005
LOG(GDY $_{t-1}$)	0.346724	0.082851	4.184922	0.0000*
LOG(GPBV $_{t-1}$)	-0.022845	0.018848	-1.212046	0.2255
LOG(GPE $_{t-1}$)	-0.028630	0.012986	-2.204683	0.0275*
Variance Equation				
ω	0.000467	0.000915	0.510852	0.6095
ARCH	0.154377	0.193544	0.797632	0.4251
GARCH	0.822629	0.224565	3.663210	0.0002*

* Significant at 5% level

Johansen and ECM results

VAR selects a system of equations with 2 lags for each variable[4]. Table 6 presents the results from the Johansen test. The test of Johansen confirms that series are cointegrated, with 1 cointegrating vector (CV), implying that in the long run financial ratios and prices have a long-run steady state equilibrium relationship.

Table 6. Johansen results

Sample (adjusted): 1992M05 2002M12
Trend assumption: No deterministic trend (restricted constant)
Series: LOG(PRICE) LOG(GPE) LOG(GPBV)
LOG(GDY)
Lags interval (in first differences): 1 to 2
Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.856315	269.3581	54.07904	0.0000
At most 1	0.081722	21.02145	35.19275	0.6609
At most 2	0.058261	10.10885	20.26184	0.6291
At most 3	0.018770	2.425337	9.164546	0.6924

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Furthermore, we also run a Error Correction model (ECM) to capture both the short-run dynamics and the long-run relationship between these variables, and predict returns using financial ratios. ECM confirms cointegration between variables under examination. However, to capture leptokurtosis, skewness and volatility clustering, we re-estimate ECM with GARCH errors. In particular, we employ ECM as mean equation and GARCH model as conditional variance equation (Floros, 2005c). We run the simple ECM-GARCH(1,1) model using R (returns) as the dependent variable and ΔX_{t-1} (for each of the ratios) as independent variables. In Table 7, we present the results for ECM-GARCH(1,1) model. Note that all the diagnostic tests show that there is no autocorrelation.

All coefficients of financial ratios are statistically significant, except P/E coefficient. The coefficients of two financial ratios (DY and PBV) show a significant influence on the returns. Also, the coefficient of error correction term (ECT) indicates that a downwards adjustment during the next period is expected (this can also explain price movements in financial returns).

[4] VAR results are available upon request.

However, this adjustment is very slow. Both ARCH and GARCH parameters are significant, indicating that ECM-GARCH(1,1) model fully captures volatility clustering.

Table 7. ECM-GARCH(1,1) results

Dependent Variable: R
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1992M03 2002M12
Bollerslev-Wooldrige robust standard errors & covariance

	Coefficient	Std. Error	z-Statistic	Prob.
μ	0.004786	0.006444	0.742657	0.4577
ΔP/E	0.008478	0.017671	0.479759	0.6314
ΔDY	-2.184941	1.270284	-1.72004**	0.0854
ΔPBV	0.529889	0.061717	8.585787*	0.0000*
ECT	-0.029945	0.011982	-2.499230*	0.0124*
	Variance Equation			
ω	0.001361	0.000797	1.707738**	0.0877
ARCH	0.353811	0.207657	1.703822**	0.0884
GARCH	0.479611	0.137275	3.493805*	0.0005*

* Significant at 5% level, ** Δ denotes first difference.

Forecasting results

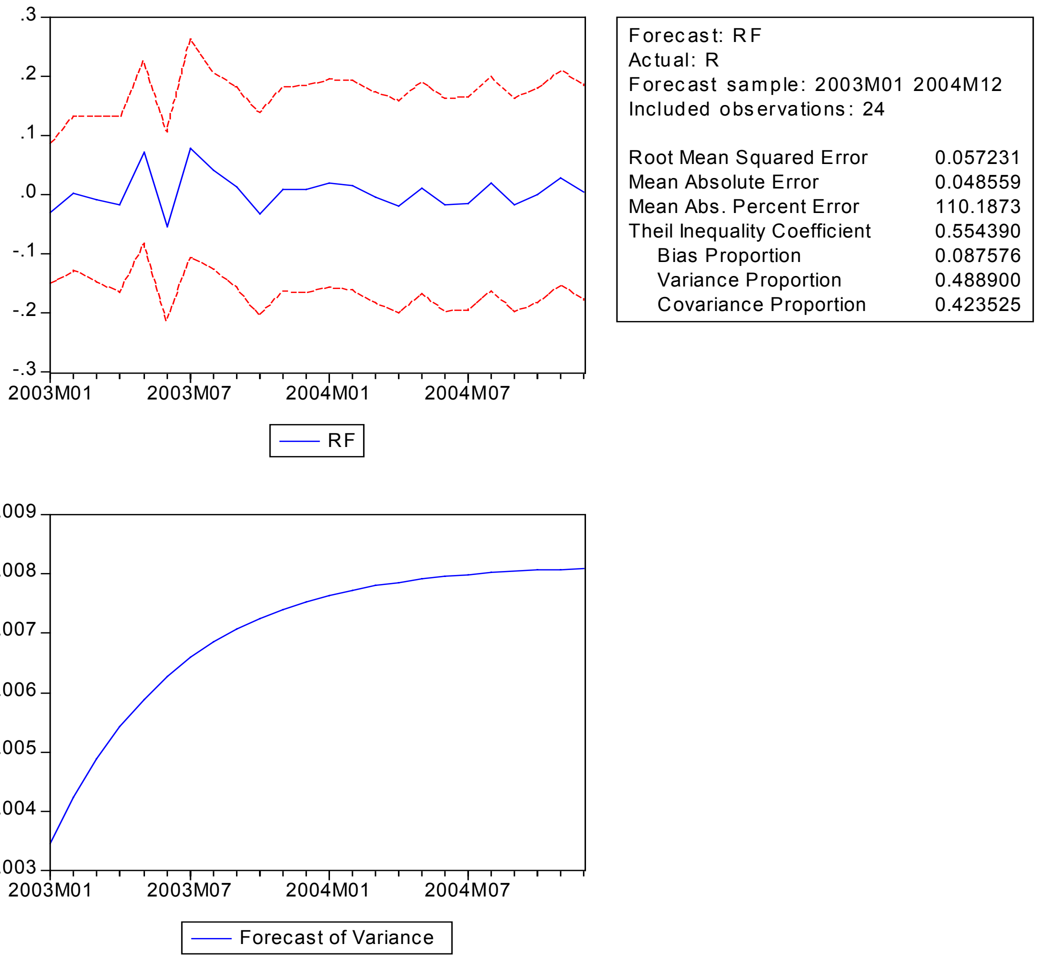
Appendix 1 explains the forecasting theory (error statistics) and describes both static and dynamic forecasts. Dynamic forecasting is performed a multi-step forecast of returns (*R)*, while static forecasting is performed as a series of one-step ahead forecasts of the dependent variable. Table 8 shows the forecasting performance of the models employed based on the error statistics: root mean square (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The Theil inequality coefficient (Theil) always lies between zero and one, where zero indicates a perfect fit. The results show that ECM-GARCH(1,1) model gives the smallest error statistics. ECM-GARCH(1,1) shows small RMSE and MAE errors, while GARCH(1,1) with PBV has the smallest MAPE. One the basis of these error statistics the selected model is the ECM-GARCH(1,1) because the Theil inequality coefficient is also small (close to zero). Hence, this selected model predicts market returns better. Two panels in Figure 1 show the forecasting performance of returns and their variance using the dynamic specification. For comparison purposes we also present (in Figure 2) the forecasting performance using the static specification of the model. The dynamic forecasts show that the variance of the forecast increases in the early years and then stabilises soon. In the static model, volatility (variance) persists. The second panel of the graph shows variance forecast.

Table 8: Forecasting Performance (*Forecasting Period: January 2003-December 2004*)

Model		RMSE	MAE	MAPE (THEIL)
PART A. DY and Returns				
OLS (model 2)-	Static	0.070437	0.055619	101.2353 (0.829563)
GARCH(1,1)-	Dynamic/Static	0.070511	0.055701	100.4640 (0.835429)
PART B. PBV and Returns				
OLS (model 2)-	Static	0.071868	0.057551	*98.22240* (0.925547)
GARCH(1,1)-	Dynamic/Static	0.071379	0.057018	98.33809 (0.893345)
PART C. P/E and Returns				
OLS (model 2)-	Static	0.073002	0.058553	100.8092 (0.904795)
GARCH(1,1)-	Dynamic/Static	0.076979	0.061719	122.5950 (0.817853)
PART D. PBV, DY, P/E and Returns		Multiple	Regressions	
OLS (model 2)-	Static	0.103481	0.088727	277.1156 (0.816197)
GARCH(1,1)-	Dynamic/Static	0.102223	0.087518	270.1241 (0.816624)
PART E. PBV, DY, P/E and Returns				
ECM-GARCH(1,1)-	Dynamic/Static	*0.057231*	*0.048559*	110.1873 (*0.554390*)

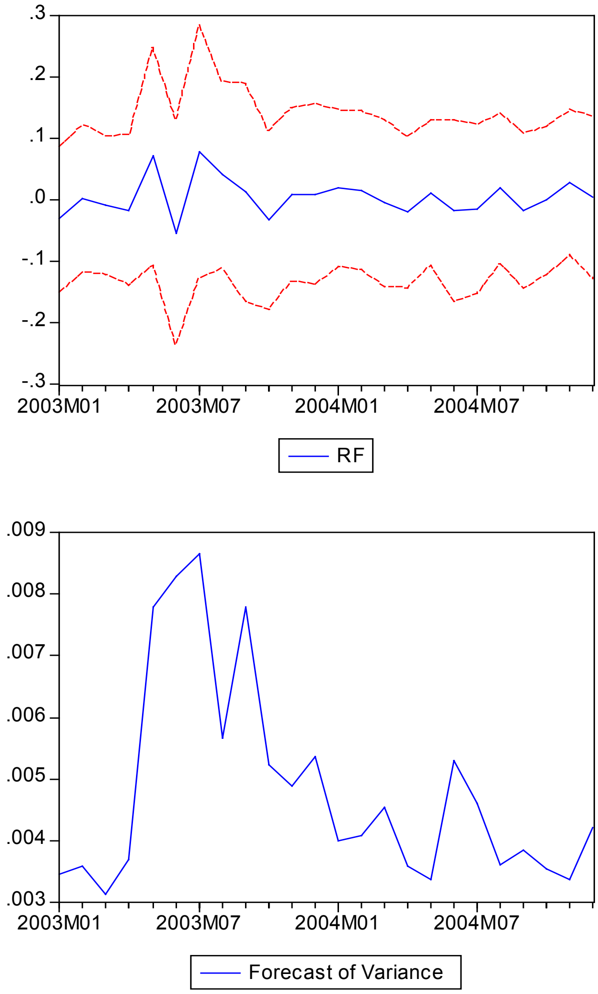
* Denotes the smaller error statistics (Theil coefficient in parentheses).
** We report forecasting results from OLS (model 2) only, because OLS (model 1) has different dependent variable.

Figure 1: Dynamic Forecasting (January 2003 - December 2004) Using ECM-GARCH(1,1)*



* Note: RF denotes the forecasting returns.

Figure 2: Static Forecasting (January 2003 - December 2004) Using ECM-GARCH(1,1)*



Forecast: RF
Actual: R
Forecast sample: 2003M01 2004M12
Included observations: 24

Root Mean Squared Error	0.057231
Mean Absolute Error	0.048559
Mean Abs. Percent Error	110.1873
Theil Inequality Coefficient	0.554390
Bias Proportion	0.087576
Variance Proportion	0.488900
Covariance Proportion	0.423525

* Note: RF denotes the forecasting returns.

4. Summary and Conclusions

The analysis of financial performance is a key activity providing financial information for financial managers and investors. Financial forecasting is a widely researched area in the applied economic literature. Predicting returns with financial ratios is one of the most important applications for financial economists. The accuracy of different forecasting methods is a topic of continuing interest and research. In this paper, we report the forecasting competition between OLS, GARCH and ECM-GARCH models of the Greek return series. We test the out-of-sample forecasting accuracy of several models using three cointegrated financial ratios, namely dividend yield, book-to-market and earnings-price ratios. Specifically, we compare the forecasting techniques based on the symmetric error statistics, while we produce the forecasting results under both static and dynamic methods.

The results from the comparisons of time-series models show that ECM-GARCH(1,1) model has significant coefficients (selected model). Both static and dynamic forecasts confirm that ECM-GARCH(1,1) is the most appropriate model for forecasting. Our findings suggest that error correction model (with lagged P/E, PBV and DY) with GARCH errors predicts market returns better.

This result is not in line with Lewellen (2004). He finds that book-to-market and earnings-price ratio predict returns during shorter sample[5]. He also finds that dividend yield predicts market returns during a longer period.

One possible explanation for the forecasting superiority of GARCH models is the fact that time-series models capture the dynamical structure generating the market returns. Forecasting results may change due to the forecasting periods (horizon) as well as the selection of in-sample and forecast data[6]. Our findings bring econometric theory nearer to the realities of financial market. Further research is required to investigate other financial forecasting methods to predict market returns.

Acknowledgments

We are grateful to Nitin Parekh and Goncalo Valle Lima for providing us with the data.

[5] From our results, only MAPE coefficient shows that book-to-market predicts market returns better.

[6] A longer data set is possible to affect the quality of the forecasts. The selection of the start of the forecast sample is mainly important for dynamic forecasting.

References

Ang, A., & Bekaert, G. (2002). Stock Return Predictability: Is it There? *Working Paper, Columbia University*.

Bekaert, G., Hodrick, R. (1992). Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets. *Journal of Finance*, Vol.47, 2, pp. 467-509.

Campbell, J., & Shiller, R. (1988). Stock Prices, Earnings and Expected Dividends. *Journal of Finance*, Vol. 43, 3, pp. 661-676.

Floros, C. (2003). Stock returns and inflation in Greece, *Applied Econometrics and International Development*, 4-2, pp. 55-68.

Floros, C. (2005a). Testing for cointegration and Granger causality between the US and European financial markets, *European Review of Economics and Finance*, 4(2), in press.

Floros, C. (2005b). Price linkages between the US, Japan and UK stock markets, *Financial Markets and Portfolio Management*. 19-2, pp. 169-178.

Floros, C. (2005c). *Essays on quantitative analysis of the Greek futures markets*, PhD Thesis, University of Wales Swansea, Department of Economics, 353p.

Goyal, A., & Welch, I. (2003). Predicting the Equity Premium with Dividend Ratios. *Management Science*, Vol. 49, 5, pp. 639-654.

Johansen, S. (1988). Statistical analysis of cointegrating vectors, *Journal of Economic Dynamics and Control*, 12, pp. 231-254.

Lewellen, J. (2004). Predicting Returns with Financial Ratios. *Journal of Financial Economics*, Vol. 74, 2, pp. 209-235.

Nelson, C., & Kim, M. (1993). Predictable Stock Returns: The Role of Small Sample Bias. *Journal of Finance*, Vol. 48, 2, pp. 641-661.

Appendix 1: Forecasting method

We produce both dynamic and static forecasts using the selected models over the sample period. Dynamic method calculates multi-step forecasts starting from the first period in the forecast sample. Static method calculates a sequence of one-step ahead forecasts, using actual rather than forecasted values for lagged dependent variables. If *S* is the first observation in the forecast sample, then the dynamic forecast is given by: $\hat{y}_s = \hat{c}(1) + \hat{c}(2)x_s + \hat{c}(3)z_s + \hat{c}(4)y_{s-1}$. On the other hand, static forecast is calculated using the actual value of the lagged endogenous variable as: $\hat{y}_{S+k} = \hat{c}(1) + \hat{c}(2)x_{S+k} + \hat{c}(3)z_{S+k} + \hat{c}(4)y_{S+k-1}$.

We compare the forecast performance of each time-series model through the error statistics (criteria). Three error statistics are employed to measure the performance of the forecasting models. Namely, the Root Mean Squared Error (*RMSE*), the Mean Absolute Error (*MAE*), and the Mean Absolute Percent Error (*MAPE*).

Suppose that the forecast sample is $t = S, S+1, .., S+h$ and denote the actual and forecasted value in period *t* as y_t and $\hat{y}_t$, respectively. The reported forecast error statistics are computed as follows:

$$RMSE = \sqrt{\frac{1}{h+1}\sum_{t=S}^{S+h}(\hat{y}_t - y_t)^2}$$

$$MAE = \frac{1}{h+1}\sum_{t=S}^{S+h}\left|\hat{y}_t - y_t\right|$$

$$MAPE = \frac{1}{h+1}\sum_{t=S}^{S+h}\left|\frac{\hat{y}_t - y_t}{y_t}\right|$$

The *RMSE* and *MAE* error statistics depend on the scale of the dependent variable. We use them to compare forecasts for the same series and sample across different time series models. The better forecasting ability of the model is that with the smaller *RMSE* and *MAE* error statistics.