Friedrich Kruse (Germany), Markus Rudolf (Germany)

The economic value of nonlinear predictions in asset allocation

Abstract

Predictions of asset returns and volatilities are heavily discussed and analyzed in the finance research literature. This paper compares linear and nonlinear predictions for stock and bond index returns and their covariance matrix. The authors show in-sample and out-of-sample prediction accuracy as well as their impact on asset allocation results for short-horizon investors. The data comprises returns from the German DAX stock market index and the REXP bond market index as well as their joint covariance matrix over the period of January 1988-December 2007. The comparison of a linear and nonlinear prediction approach is the focus of this study. The results show that while out-of-sample prediction accuracies are weak in terms of statistical significance, asset allocation performances based on linear predictions result in significant Jensen's alpha measures and Sharpe ratio and are further improved by nonlinear predictions.

Keywords: nonlinear prediction, neural networks, asset allocation.
JEL Classification: C32, C45, C53, G11.

Introduction

Predicting economic time series plays a crucial role in finance. It is essential for mean-variance efficient asset allocation to estimate expected returns, correlations and volatilities. There is a vast number of literature dealing with linear prediction of risk and return. But also nonlinear prediction attains growing interest. As such, nonlinear models become more and more important in economic forecasting[1]. Comparative studies of linear and nonlinear models for economic predictions were performed, for example, by Swanson and White (1995; 1997). They apply nonlinear neural network models to predict future spot rates and macroeconomic variables. Many other studies testify that economic time series are nonlinear in nature. However, the effect of nonlinear predictions on asset allocation and the question whether nonlinear predictions are economically exploitable has been neglected in the literature.

There exist several possibilities to build a nonlinear forecasting model[2]. But due to their flexibility, artificial neural networks can be a powerful method for predictions, especially if we have little prior knowledge about the data generating process. Therefore, we use a neural network approach for nonlinear forecasting in this study. We consider a mean-variance investor with one month investment horizons who allocates his wealth to stocks, bonds and a risk-free asset. The investor updates his opinion about conditional returns and the conditional covariance matrix on the basis of a linear or a nonlinear prediction model, respectively. Predictions for monthly excess returns and the conditional covariance matrix are performed and plugged in a mean-variance asset allocation strategy. The asset universe consists of the German stock market index DAX, the German bond market index REXP and a risk-free rate using data over the period of January 1988-December 2007. We report traditional portfolio performance measures such as Sharpe ratio and Jensen's alpha as well as the Treynor Mazuy measure (Treynor and Mazuy, 1966) and a utility based measure of portfolio performance inspired by West, Edison and Cho (1993)[3]. Our results show that linear and nonlinear predictions are economically relevant for mean-variance investors. Nonlinear prediction models outperform linear prediction models in terms of Sharpe ratio and Jensen's alpha. In terms of the utility based performance measure and the Treynor and Mazuy measure, nonlinear predictions result in better asset allocation performances when we impose short-sales constraints.

1. Literature review

1.1. Predictability of asset returns. The question whether asset returns are predictable triggered a controversial debate in the finance literature. Cochrane (1999) reviews the research on predictability of stock and bond returns, and concludes that predictability can be seen as "new fact in finance". In earlier studies, during the 1980s, valuation ratios were used to predict future returns, starting with dividend yields by Rozeff (1984) as well as Fama and French (1988). Also, Campbell and Shiller (1988a; 1988b) found that dividend yields are positively correlated with future returns. More recently, Kothari and Shanken (1997), Pontiff and Schall (1998), Lamont (1998), Stambaugh (1999), Lewellen (2004), and Campbell and Yogo (2006) examined the predictability of returns by financial ratios. They show that book-to-market ratios and dividend yields have predictive power for subsequent stock market returns.

We would like to thank Tim Kruse for helping us to improve the paper by his valuable comments.
[1] They increasingly appear in standard econometric handbooks. See, for example, Tsay (2002) and TerËasvirta (2006).
[2] See for example TerËasvirta (2006).
[3] This measure has been used, for example, by Fleming, Kirby and Ostdiek (2001) and Marquering and Verbeek (2004) for empirical asset allocation evaluation.

Research on other predictive variables include short-term interest rates (e.g., Fama and Schwert, 1977), spreads between long-term and short-term interest yields (e.g., Keim and Stambaugh, 1986; Campbell, 1987), spreads between corporate bonds and the one-month bill rate (e.g., Fama and French, 1989), stock market volatility (e.g., French, Schwert and Stambaugh, 1987), the level of consumption in relation to wealth (Lettau and Ludvigson, 2001)[1], and others (see, e.g., Ferson and Harvey, 1993; Pesaran and Timmermann, 1995). A critical view on return predictability of risky assets is taken by Valkanov (2003), Ang and Bekaert (2007), and Goyal and Welch (2003; 2008). For long-horizon predictive regressions, Valkanov (2003) shows that *t*-statistics do not converge to well-defined distributions and the R^2 is in some cases an inadequate measure of the goodness of fit. Ang and Bekaert (2007) correct for heteroscedasticity and find that long-horizon predictability vanishes and is not robust across sample periods. Their results suggest that predictability is mainly a short-horizon phenomenon. These papers rely on econometric arguments using several statistical tests. The large variety of test procedures that have been proposed for return predictability, which have led to different conclusions, has hampered the understanding of the rather large literature on predictability[2]. In a comprehensive study, Goyal and Welch (2008) analyze predictive variables in linear regression models out-of-sample. They argue that the historical average excess stock return has more predictive power than most of the tested linear regression models. A critical view on Goyal and Welch's (2008) conclusion is taken by Cochrane (2008) and Campbell and Thompson (2008). Cochrane (2008) points out that a poor out-of-sample R^2 does not reject the null hypothesis that returns are predictable. The out-of-sample R^2 is not a test statistic that gives stronger evidence about return predictability than regression coefficients or other standard hypothesis tests. However, Cochrane (2008) argues that the absence of out-of-sample performance is not likely to be useful in forming real-time forecasts or market-timing portfolios. In addition, Campbell and Thompson (2008) impose simple restrictions on predictive regressions and consequently improve the out-of-sample performance.

Interestingly, most of the studies rely on linear regression models. These models imply that predictive variables are subject to linear dependencies on predicted returns. However, Lee, White and Granger (1993) show that economic time series can be subject to nonlinear dependencies and they establish a test to detect neglected nonlinearities in time series models based on neural networks. In their empirical study, Abhyankar, Copeland and Wong (1997) confirm nonlinear dependence for the stock indexes FTSE-100, the S&P 500, the DAX, and Nikkei[3]. Also, Desai and Bharati (1998) find evidence of nonlinear dependencies between explanatory variables and returns of large stocks and corporate bonds[4].

On the basis of monthly excess returns, the analysis of Qi (1999) as well as Qi and Maddala (1999) show that a nonlinear neural network approach outperforms a linear regression model in forecasting S&P 500 returns. Qi (1999) mainly uses the same variables as Pesaran and Timmermann (1995) in their linear regressions. As a result, Qi (1999) reports better in-sample fit and out-of-sample forecasts using a neural network model compared to its linear counterpart. More recently, Ince and Trafalis (2007) and Bekiros and Georgoutsos (2008) show that neural network models can be successfully implemented for return predictability.

1.2. Predictability of volatilities and correlations. For their portfolio construction, mean-variance investors need estimates of asset return volatilities. The application of volatility in optimal portfolio selection – among other important applications – has motivated numerous studies on volatility modeling[5]. In their analysis, Schwert (1989) and French et al. (1987) based the volatility measure on the sum of squared daily returns[6]. This approach is commonly referred to as *realized volatility* in the literature[7]. Another approach to modeling volatility is spurred by Engle's (1982) model of autoregressive conditional heteroscedasticity (ARCH) and its generalized form (GARCH) (Bollerslev, 1986). A vast number of researchers have surveyed volatility forecasts in this framework[8].

However, as Andersen, Bollerslev, Diebold and Labys (2003) and Andersen, Bollerslev and Diebold (2009) argue, the measure of *realized volatility* benefits from being free of parametric functional form assumptions in contrast to the ARCH and GARCH approaches. Realized volatility is a consistent esti-

[1] In a more recent study, Guo (2006) supports their findings and confirms out-of-sample predictability of stock returns.

[2] Campbell and Yogo (2006) provide an understanding of the various test procedures and their empirical implications.

[3] They use intra-day data between 15-seconds, 1-minute, and 5-minute intervals. They find that their results are also consistent for 15-minute, 30-minute, and 1-hour intervals.

[4] Among other tests, they use the test for neglected nonlinearity proposed by Lee et al. (1993).

[5] For a comprehensive review we refer to Andersen, Bollerslev, Christoffersen and Diebold (2006) and Poon and Granger (2003).

[6] The model is based on an earlier work by Merton (1980). See Merton (1980), Appendix A.

[7] See, for example, Andersen, Bollerslev, Diebold and Labys (2001).

[8] The ARCH and GARCH class of models have been surveyed, for example, by Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), Engle (2001; 2004), and Diebold (2004).

mate of ex-post return variability and provides a benchmark for out-of-sample volatility forecasts[1]. As such, Ludvigson and Ng (2007) perform volatility forecasts using time-series variation of daily returns, following Schwert (1989) and French et al. (1987). They use a factor analysis approach to predict quarterly volatility of excess returns. Also, Whitelaw (1994) and Ghysels, Santa-Clara and Valkanov (2005) use this measure of volatility in their studies.

Hamid and Iqbal (2004) perform volatility forecasts of S&P 500 Index future prices using a neural network approach. They compare the forecast performance to Barone-Adesi and Whaley's (1987) model of implied volatility forecasts of American futures option pricing. The reported results show that the nonlinear neural network approach outperforms implied volatility forecasts when compared to realized volatility[2].

Ferland and Lalancette (2006) use the measure of realized volatility to compare neural network, ARMA, GARCH-BEKK, and naive volatility and correlation forecasts of Bax and Eurodollar futures. They measure realized correlation similar to realized volatility. As such, forecasts are performed for weekly volatilities and correlations with an out-of-sample period of 98 weeks. The neural network model and the ARMA model provide the best forecasting performance. Furthermore, the results show, that the neural network approach has additional explanatory power beyond the ARMA-based model.

As for return predictability, most of the literature focuses on evaluating the statistical performance of volatility models rather than the economic relevance of predictive volatility. For predictions of forecasting covariances or correlations, the impact on optimum portfolios has been discussed by Chan, Karceski and Lakonishok (1999), Ledoit and Wolf (2003), and Elton, Gruber and Spitzer (2006). Chan et al. (1999) use a linear factor model with a number of factors ranging from one to ten in order to forecast covariances. Whereas they find that predictions do not significantly outperform a constant covariance model and consequently do not lead to better asset allocation, Elton et al. (2006) show that correlation forecasts that outperform constant historical correlations in terms of root mean squared errors also result in better portfolio decisions. They calculate minimum variance portfolios based on the forecast of correlations and historic variances. Subsequently they examine the actual risk of the portfolio. The results show that their correlation forecasts lead to lower portfolio variance.

However, these studies on correlation and covariance matrix forecasts are restricted to linear models and do not take into account estimates of returns in the calculation of portfolios.

2. Mean-variance investors and optimization

To quantify the value of nonlinear predictions, we measure the impact on the performance of asset allocation strategies of an investor with one-month investment horizons. In order to implement the asset allocation strategy we use mean-variance analysis. In period t, the investor selects the asset fractions $w_t \in R^N$ from a given set of N risky assets and allocates the remainder $(1 - w_t'\mathbf{1})$, where $\mathbf{1}$ is an N x 1 vector of ones, to the risk-free asset. Let r_{t+1} be an N x 1 vector of excess returns over the risk-free asset return R_t^f in period $t + 1$. The portfolio return is then $R_{t+1}^p = R_t^f + w_t' r_{t+1}$. The N-dimensional vector of conditional expected value of r_{t+1} is given by $\hat{r}_{t+1}$ and the expected conditional covariance matrix is $\hat{\Omega}_{t+1} \in R^{N \times N}$. For each month t, the investor solves the optimization problem[3]:

$$\begin{aligned} \min_{w_t} \quad & var_t\left[r_{t+1}^p\right] = w_t' \hat{\Omega}_{t+1} w_t \\ s.t. \quad & E_t\left[r_{t+1}^p\right] = w_t' \hat{r}_{t+1}, \end{aligned} \tag{1}$$

where $E_t\left[r_{t+1}^p\right]$ is the expected target excess return of the portfolio and r_{t+1}^p is the portfolio excess return in period t+1. The solution to the problem is:

$$w_t = \frac{w_t' \hat{r}_{t+1}}{\underbrace{\hat{r}_{t+1}' \hat{\Omega}_{t+1}^{-1} \hat{r}_{t+1}}_{=\lambda}} \cdot \hat{\Omega}_{t+1}^{-1} \hat{r}_{t+1}, \tag{2}$$

where λ is a constant that scales the resulting vector of $\hat{\Omega}_{t+1}^{-1} \hat{r}_{t+1}$ to the expected target excess return of the portfolio $E_t\left[r_{t+1}^p\right]$. As such, the solution is a minimum-variance portfolio to the pre-determined expected excess portfolio return. However, the determination of the portfolio risk premium $E_t\left[r_{t+1}^p\right]$ depends on the investor's tolerance for risk.

Therefore, we consider an investor with a particular risk aversion and assess the asset weights that are optimal for the investor's risk-return trade-off. More specifically, we impute a utility function and reformulate the optimization problem in terms of maximizing the investor's utility:

[1] See the discussion in Andersen et al. (2006), p. 830 and the literature stated above. Also Hansen and Lunde (2006) discuss realized volatility as a proxy for "true" volatility to evaluate the performance of volatility models.

[2] In their study, *realized volatility* is defined slightly different, since the standard deviation is divided by the root number of days of the volatility window. See Hamid and Iqbal (2004, p. 1121).

[3] For a comprehensive discussion see Rudolf (1994).

$$\max_{w_t} E_t\left[r_{t+1}^p\right] - \frac{\gamma}{2} var_t\left[r_{t+1}^p\right], \quad (3)$$

where γ is a coefficient representing the investor's level of relative risk aversion. The solution to the above maximization is:

$$w_t = \frac{1}{\gamma}\hat{\Omega}_{t+1}^{-1}\hat{r}_{t+1}, \quad (4)$$

which is equivalent to equation (2) with $\gamma = 1/\gamma$ as a measure of the investor's tolerance for risk. We use this approach to determine the optimal asset weights and show the impact on altering the investor's risk aversion γ.

The calculation of optimal asset weights requires one-step-ahead knowledge about the vector of conditional excess returns $\hat{r}_{t+1}$ and the conditional covariance matrix $\hat{\Omega}_{t+1}$. In general, the parameters are unknown and have to be estimated. In the next section we develop linear and nonlinear models to predict the required parameters.

3. Methodology of predictions

3.1. Linear and nonlinear models for asset returns and correlations. For mean-variance efficient asset allocation we need estimations for expected returns, correlations and volatilities of the asset universe. Consequently, we have to establish three prediction models. In their paper, Marquering and Verbeek (2004) use linear models for the conditional expectations of excess returns and volatilities[1]. We will use these models as benchmark to our extended, nonlinear approach.

Linear regression modeling is well known and most widely used for prediction problems. Although a linear model is powerful and has convenient properties, it still rules out many useful *nonlinear* functional forms. A linear regression model assesses the relationship between a dependent variable y_{t+1} and one or a vector of more independent variables $x_t \in R^{k+1}$, where k is the number of independent variables, including a constant. The functional relationship is assumed to be linear. As such, the generic form of a linear regression model (including a constant β_0) is given by:

$$\begin{aligned} y_{t+1} &= f(x_t) + \varepsilon_{t+1} = \beta_0 + \beta_1 x_{t,1} + \cdots + \beta_k x_{t,k} + \\ &+ \varepsilon_{t+1} = x_t'\beta + \varepsilon_{t+1}, \end{aligned} \quad (5)$$

where $\beta \in R^{k+1}$ is a vector of to be estimated parameters and ε_{t+1} are residuals not explained by the model[2]. Most studies limit the set of predictive functions to the linear relation $f(x_t) = l(x_t, \beta) \equiv x'\beta$, with a vector $x_t \in R^{k+1}$ of k independent variables including a constant (i.e., $x_t = (1, x_{t,1}, \ldots, x_{t,k})'$), and a vector $\beta \in R^{k+1}$ of to be estimated linear parameters. Generally, the true functional relationship is, however, unknown.

Marquering and Verbeek (2004) use those simple linear regressions to predict excess returns with a set of k predictive variables $x_{r,t}$ available at time $t = 1,...,T$. Thus, the expected value of r_{t+1}, given the information of predictive variables $x_{r,t}$, is assumed to be[3]:

$$r_{t+1} = x_{r,t}'\beta_r + \varepsilon_{r,t+1}, \quad (6)$$

where $x_{r,t} = (1, x_{r,t,1}, \ldots, x_{r,t,k})' \in R^{k+1}$. For predictive variables in $x_{r,t}$, lagged financial ratios and macroeconomic data is used. Thus, k indicates the number of independent variables that are used in the model. Since we include a constant in our prediction model, β_r is a $(k + 1)$ vector of to be estimated parameters. The parameters $\hat{\beta}_r$ are determined by an OLS estimation, so that the forecasts for excess returns in $t + 1$ are:

$$\hat{r}_{t+1} = x_{r,t}'\hat{\beta}_r. \quad (7)$$

In the same way, we consider a linear model for predicting conditional correlations $p_{ij,t+1}$ between asset i and j $(i \neq j)$:

$$\hat{p}_{ij,t+1} = x_{\rho,t}'\hat{\beta}_\rho \qquad \hat{p}_{ij,t+1} \in [-1; 1], \quad (8)$$

where $x_{\rho,t} = (1, x_{\rho,t,1}, \ldots, x_{\rho,t,p})' \in \mathrm{R}^{p+1}$ is a vector of predictive variables for correlations, $\hat{\beta}_\rho$ is a $(p + 1)$ vector of estimated parameters (including a constant) and we limit $\hat{\rho}_{ij,t}$ to be in the interval [-1; 1]. ρ indicates the number of predictive indicators that are used in the model for correlation predictions.

We now extend the linear prediction models for asset returns and correlation to capture *nonlinear* dependencies. In order to assure that the nonlinear model nests the linear model, we follow White's (2006) proposition by including a linear component in $f(\cdot)$ and specify:

$$y_{t+1} = f(x_t, \beta, \psi) + \epsilon_{t+1} = \underbrace{x_t'\beta}_{\text{linear part}} + \underbrace{n(x_t, \psi)}_{\text{nonlinear part}} + \epsilon_{t+1}, \quad (9)$$

where $n(x_t, \psi)$ is a nonlinear function with to be estimated free parameters ψ. For estimation purpos-

[1] They omit the problem of estimating correlations since they use only one risky asset in their analysis.

[2] Residuals are usually assumed to be uncorrelated with x_t, have conditional expected value zero and constant variance.

[3] The predictions are done for every asset $i, i = 1,...,N$. However, we do not state i in the following equations for the virtue of simplicity. The sub-script r indicates that the variable is part of the model for predicting returns r.

es, we use the OLS solution to assess the linear parameters β (see the general linear equation (5)). From this linear regression model, we further analyze the resulting vector of residuals $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$. We uncover the purely nonlinear dependency between independent variables x_t and the dependent variable y_{t+1} by using the residuals as dependent variable for the nonlinear model. The regression parameters ψ from the nonlinear part of the prediction model are assessed by a neural network approach. As such, we use the linear regression method for estimating parameters in the linear part of the model and the neural network approach to estimate parameters in the nonlinear part represented by $n(x_t, \psi)$.

The implication of this procedure is as follows. The errors ε_{t+1} in the general linear prediction model (5) can be random noise, which is unrelated to the predictive independent variables x_t. But they might also include hidden dependencies which cannot be mapped by the linear structure of the regression model. If there are nonlinear dependencies, the neural network "adds value" to the prediction model which will be reflected in a reduction of the squared error term. Consequently, we try to explain the errors ε_{t+1} from the linear model by a nonlinear neural network model:

$$\varepsilon_{t+1} = n(x_t, \psi) + \epsilon_{t+1}, \tag{10}$$

where ψ are to be estimated parameters and ϵ_{t+1} is random noise. If the mean-squared-error over the vector of all errors ϵ is smaller compared to ε, we can expect some nonlinear dependencies between r_{t+1} and x_t.

Hence, we write the combined model for return predictions for the nonlinear model following equation (9):

$$\hat{r}_{t+1} = x'_{r,t}\hat{\beta}_r + n(x_{r,t}, \hat{\psi}_r), \tag{11}$$

where the model parameters $\hat{\beta}_r$ are OLS estimates. The parameters $\hat{\psi}_r$ are obtained from neural network training with the linear residuals as network targets (equation (10)). By analyzing the errors ε and ϵ, we can separate the prediction performance due to the linear regression and assess the added value of the neural network.

For realized correlation consider the same model structure as presented for asset returns. That is for the nonlinear predictions:

$$\hat{\rho}_{ij,t+1} = x'_{\rho,t}\hat{\beta}_\rho + n(x_{\rho,t}, \hat{\psi}_\rho), \tag{12}$$

where $\hat{\beta}_\rho$ are again estimated OLS parameters and $\hat{\psi}_\rho$ are estimated neural network parameters. For an in depth overview on neural network modeling we refer to Kuan and White (1994) and Bishop (1995). The following specifications are made to the neural network which are used for estimating nonlinear parameters $\hat{\psi}$. We use Bayesian regularization for parameter estimation based on the work of MacKay (1992)[1]. This is a very sophisticated and powerful algorithm to achieve network models with good generalizing abilities. The algorithm ensures that the network is kept at the necessary simplicity and is not overparameterized. We use a computationally efficient method which is described by Foresee and Hagan (1997) by integrating the Levenberg-Marquardt algorithm to the neural network estimation process. All neural networks we use throughout this paper are trained with seven neurons in a single hidden layer with logistic activation function[2].

3.2. Linear and nonlinear prediction models for volatilities. The prediction model for conditional volatilities is explained by a vector of predictive variables $\mathrm{x}_{VOL,t} = (1, x_{VOL,t,1}, \ldots, x_{VOL,t,q})' \in \mathrm{R}^{q+1}$ and for the linear part by a $(q + 1)$ vector β_{VOL} of coefficients (again including a constant) which has to be estimated. q states the number of independent variables that are used to predict volatility. Similar to Marquering and Verbeek (2004) and in the fashion of Harvey's (1974) model of multiplicative heteroscedasticity, we assume for the linear model:

$$\begin{aligned} VOL_{t+1} &= \exp(\mathrm{x}'_{VOL,t}\beta_{\mathrm{VOL}} + \varepsilon_{VOL,t+1}) = \\ &= \exp(x'_{VOL,t}\beta_{VOL}) \cdot \exp(\varepsilon_{VOL,t+1}), \end{aligned} \tag{13}$$

where we denote VOL_{t+1} to be the asset's conditional volatility. Again, the parameters β_{VOL} are estimated via *OLS*. The linearly predicted value $\hat{VOL}_{t+1}$ in period $t + 1$ for conditional volatility is then given by:

$$\log\hat{VOL}_{t+1} = x'_{VOL,t}\hat{\beta}_{VOL}. \tag{14}$$

To establish the nonlinear model of realized conditional volatilities, we extend the prediction model for VOL_{t+1} by forecasting $\exp(\varepsilon_{VOL,t+1})$ from equation (13) through a neural network approach. We therefore write:

$$\exp(\varepsilon_{VOL,t+1}) = n(x_{VOL,t}, \psi_{VOL}) + \epsilon_{VOL,t+1}, \tag{15}$$

where estimates for the parameters ψ_{VOL} are again obtained from neural network learning. Given the linear OLS estimates $\hat{\beta}_{VOL}$ from (13) and the network estimates $\hat{\psi}_{VOL}$ from (15), the predicted values from the nonlinear model for $\hat{VOL}_{t+1}$ are derived from:

$$\hat{VOL}_{t+1} = \exp(\mathrm{x}'_{VOL,t}\beta_{VOL}) \cdot n(\mathrm{x}_{VOL,t}, \hat{\psi}_{VOL}). \tag{16}$$

[1] A review of Bayesian techniques for neural networks is also given by Bishop (1995, pp. 385-433).

[2] Due to the properties of Bayesian regularization, an increasing number of neurons in the hidden layer does not change the results severely.

Predictions for realized conditional volatilities for all assets in combination with predictions for all assets correlations allow us to construct a prediction of their joint covariance matrix. This will be useful for calculating optimal portfolios.

In summary, the *linear* prediction models are described by equation (7) (excess returns), (8) (correlations), and (14) (volatilities). The *nonlinear* models are represented by equation (11) (excess returns), (12) (correlations), and (16) (volatilities).

3.3. Realized volatility and correlation. Defining realized returns is straight forward. We compute monthly returns r_t from prices P as $r_t = \log P_t - \log P_{t-1}$. We subtract the risk-free rate to obtain excess returns. However, for assessing the covariance matrix, we have to rely on volatility modeling. In our analysis, we use daily returns to calculate monthly realized volatility. This approach is discussed by Merton (1980), French, Schwert and Stambaugh (1987), Schwert (1989), and more recently by Andersen, Bollerslev, Diebold and Labys (2003) and Andersen, Bollerslev and Diebold (2009). Similar to Schwert (1989), we measure volatility of monthly returns as the sum of squared daily returns:

$$VOL_t = \sum_{d=1}^{D_t} r_{d,t}^2, \tag{17}$$

where $r_{d,t}$ is the return on day d and D_t is the number of trading days in month t.

In a similar fashion, we measure the correlation between asset returns. We denote the correlation between asset i and j as[1]:

$$\rho_{ij,t} = \frac{\sum_{d=1}^{D_t} r_{i,d,t} r_{j,d,t}}{\sigma_{i,t}\sigma_{j,t}} \qquad \rho_{ij,t} \in [-1;1]. \tag{18}$$

Subsequently, we can generate the covariance matrix of two assets i and j in month $t + 1$ as

$$\Omega_{t+1} = \begin{pmatrix} VOL_{i,t+1} & \rho_{ij,t+1} \cdot \sqrt{VOL_{i,t+1}} \cdot \sqrt{VOL_{j,t+1}} \\ \rho_{ij,t+1} \cdot \sqrt{VOL_{i,t+1}} \cdot \sqrt{VOL_{j,t+1}} & VOL_{j,t+1} \end{pmatrix}. \tag{19}$$

We predict the individual elements in Ω_{t+1} with a linear and a nonlinear model as described in the previous sections to assess the expected covariance matrix $\hat{\Omega}_{t+1}$. The predictions are restricted to be positive for values of VOL_{t+1} and between -1 and 1 for values of $\rho_{ij,t+1}$. As such, we assure that the covariance matrix is positive semidefinite.

4. Empirical results

4.1. Data summary. In this section, the forecasting models are applied to empirical data. The asset universe consists of two popular risky asset classes, i.e., equities and bonds, and a risk-free asset. The investor allocates his wealth each month to the German stock market performance index DAX, the bond performance index REXP and the one-month Frankfurt banks middle rate (as a proxy for the risk-free asset)[2]. We use monthly data for returns covering the time period from January 1988 to December 2007 and we use daily returns of the DAX and REXP to measure realized volatility and correlation.

Table 1. Sample summary statistics

The table provides summary statistics for excess returns of the German stock market DAX 30 performance index and the REXP performance bond index for the period from 1988 to 2007 and two sub-periods from 1988 to 1997 and from 1998 to 2007. The table also provides summary statistics of the mean of the 30-days average realized volatility and correlation measure based on daily returns.

	January 1988-December 2007 $T = 240$	January 1988-December 1997 $T = 120$	January 1998-December 2007 $T = 120$
DAX excess returns			
Mean (%)	0.492	0.709	0.275
Mean realized volatility	6.386	5.369	7.262
REXP excess returns			
Mean (%)	0.113	0.097	0.130
Mean realized volatility	0.898	0.873	0.922
DAX, REXP return correlation			
Mean realized correlation	0.163	0.432	-0.106

[1] See also Ferland and Lalancette (2006) for a similar approach.
[2] The rate has a correlation to the one-month EURIBOR of practically one. However, we use the Frankfurt rate due to a longer time-series history.

Table 1 presents summary statistics of the stock- and bond index excess returns. The table provides sample statistics for the entire data sample of 20 years and two sub-samples of 10 years from 1988 to 1997 and 1998 to 2007. The sample statistics show the typical characteristics of the two asset classes, where the equity index yields higher excess returns compared to bond index excess returns, but is associated with higher risk (i.e., realized volatility). For the entire sample period, January 1988 to December 2007, the mean of monthly DAX excess returns is 0.49% with a corresponding mean realized volatility of 6.39%. For the same sample period bonds have a mean of monthly excess returns of 0.11%. The lower returns also correspond to lower volatility of about 0.90% per month. The mean realized correlation of excess returns is 0.16 for the entire sample. The sub-samples reveal that the correlation is clearly positive during the period from 1988 to 1997 (0.43) and negative in the subsequent period from 1998 to 2007 (-0.11).

For our choice of predictive variables for asset returns we refer to Keim and Stambaugh (1986), Fama and French (1988), Campbell and Shiller (1988a; 1988b), Fama and French (1989), and Campbell and Yogo (2006). For DAX excess return predictions, the dividend yield (*DY*), earnings-price ratio (*E/P*) and term-structure (*TERM*) are included as predictive variables. All variables are included with a one-month lag in order to perform true ex ante predictions. The term structure is defined as the spread between long-term German government bond yields (9-10 years) and the one-month short-term interest. For REXP bond excess return predictions the term-structure (*TERM*) is used and the growth rate of *M*3 money supply (contribution to Euro basis) denoted by ΔMNY. The choice of predictive variables is motivated by the literature: The dividend yield and earnings-price ratio as predictive variables for equity returns have been used, for example, by Lewellen (2004) and Lettau and van Nieuwerburgh (2007), among many others. Predictions of bond returns are discussed, for example, by Keim and Stambaugh (1986) and Fama and French (1989).

For DAX and REXP volatility predictions the same predictive variables are used as for their return predictions, but also one-month lagged realized conditional log volatility is included to account for volatility clusters[1]. Correlations are predicted by one-month lagged realized correlations and one-month lagged realized conditional log volatilities of DAX and REXP returns. This choice of variable selection is motivated by Elton, Gruber and Spitzer (2006), who emphasize the autoregressive character of correlations and note that variances can be useful for correlation forecasting. All data is taken from Thomson Reuters Datastream.

4.2. Prediction results. The sample period goes from January 1988 to December 2007 and contains 240 months of data. Predictions are made in recursive model estimations and proceeds as follows: 120 observations from the beginning of the time series (i.e., from January 1988 to December 1998) are taken as base period. The linear and nonlinear models are estimated from these observations and the parameter values are used to make predictions for January 1999. In order to make predictions for Fenruary 1999, the sample is increased to the period from January 1988 to January 1999 and we re-estimate the model parameters each month. We proceed as such for each subsequent month. Consequently, we employ an expanding data window for parameter estimation, where the final parameters are estimated from the data sample containing data from January 1988 to November 2007 in order to make predictions for December 2007[2]. We can evaluate the predictions for each month as out-of-sample results since the recursive procedure ensures that the predicted month is not used for model parameter estimation.

We report both, in-sample and out-of-sample prediction results in order to assess the quality of the predictive models. In-sample results are derived from a prediction model whose parameters are estimated from a data set and the same data is used for in-sample performance measures. By contrast, for out-of-sample results excluded data from the estimation period is reserved for out-of-sample performance measurement. For linear regressions it is common practice to report only in-sample results. However, in this analysis it is also important to consider out-of-sample results due to two reasons: First, the in-sample and out-of-sample performance of neural network models can severely differ because of over-fitting. Second, it is a general problem to achieve good out-of-sample performances for financial market predictions, even though when in-sample results are promising, as shown by Bossaerts and Hillion (1999) and Goyal and Welch (2003; 2008).

In particular, the following measures of forecasting accuracy are used to compare the fit of the models. When $\hat{y}_{t+1}$ is a fitted (in the case of in-sample analysis) or predicted (in the case of out-of-sample analysis) value and y_{t+1} is the true value, the root mean squared error (RMSE) of prediction is given by:

[1] See Marquering and Verbeek (2004) for a similar procedure.

[2] This recursive method of model estimation is also chosen by Qi (1999) and Marquering and Verbeek (2004).

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=0}^{T-1}(y_{t+1} - \hat{y}_{t+1})^2}, \quad (20)$$

where T is the number of months which are used for parameter estimation. Furthermore, the mean absolute error (*MAE*) is defined as the absolute deviation between the true value and the fitted or predicted value:

$$MAE = \frac{1}{T}\sum_{t=0}^{T-1}\left|y_{t+1} - \hat{y}_{t+1}\right|. \quad (21)$$

The correlation (*CORR*) between the true value and the fitted or predicted value by the model is given by:

$$CORR = \frac{\sum(y_{t+1} - \bar{y}_{t+1})(\hat{y}_{t+1} - \bar{\hat{y}}_{t+1})}{\sqrt{\sum(y_{t+1} - \bar{\hat{y}}_{t+1})^2}\sqrt{\sum(\hat{y}_{t+1} - \bar{\hat{y}}_{t+1})^2}}, \quad (22)$$

where $\bar{y}_{t+1}$ is the mean of the true values and $\bar{\hat{y}}_{t+1}$ is the mean of the fitted or predicted values. Firstly, we concentrate on the (in-sample) *fitted* values rather than on the (out-of-sample) *predicted* values, which are analyzed subsequently.

Table 2. In-sample model fit for DAX and REXP excess returns

The table provides average performance measures, given by (20), (21) and (22), for fitted values of DAX and REXP excess return. The linear model is based on OLS estimates for $\hat{r}_{t+1} = \mathrm{x}'_{r,t}\hat{\beta}_r$ and the nonlinear model is based on OLS and neural network estimates for $\hat{r}_{t+1} = \mathrm{x}'_{r,t}\hat{\beta}_r + n(\mathrm{x}_{r,t}, \psi_r)$. * Indicates that the model is significantly different to the performance measure of the historical mean model, based on a 5% significance level *t*-test. † indicates that the performance measure for the nonlinear model is significantly different to the performance measure of the linear model at the 5% significance level.

	RMSE	*MAE*	*CORR*
DAX excess returns (%)			
Historical mean	5.924	4.596	-
Linear model	5.854	4.536	0.145
Nonlinear model	5.433*†	4.171*†	0.392†
REXP excess returns (%)			
Historical mean	1.011	0.815	-
Linear model	1.008	0.818	0.064
Nonlinear model	0.939*†	0.753*†	0.368†

For excess return predictions, we use lagged predictive variables (at time *t*) $\mathrm{x}_{rDAX,t} = \left[DY_t, E/P_t, TERM_t\right]'$ for predicting DAX excess returns and $\mathrm{x}_{r_{REXP},t} = \left[\mathrm{TERM}_t, \Delta MNY_t\right]'$ for predicting REXP excess returns. All networks are trained with seven neurons in a single hidden layer with logistic activation function and linear activation function for the output unit. Due to a total of 120 recursive estimations it is infeasible to report all details of every single estimation. In Table 2 the average of each measure across all estimations is reported. In addition to the linear and nonlinear model, the table reports performance measures if the historical mean is taken as prediction value. Using the historical mean as prediction value corresponds to the assumption that there is no dependence whatsoever, neither linear nor nonlinear, between the dependent and independent variables[1].

Table 2 shows that the nonlinear model provides the best fit to the data across all measures. Obviously, the nonlinear model yields the smallest *RMSE* and *MAE* performance values, and the correlation measure *CORR* is the largest. For *DAX* excess returns the *RMSE* reduces from 5.92 for the historical mean model to 5.85 for the linear model. This error measure decreases to 5.43 for the nonlinear model. The differences might seem small but they are significantly different based on a *t*-test with 5% significance level. Also, the correlation between the true values and predicted values significantly improve when we turn from the linear model (*CORR* = 0.15) to the nonlinear model (*CORR* = 0.39)[2].

Table 3. In-sample model fit for DAX and REXP volatilities and correlation

The table provides average performance measures, given by (20), (21) and (22), for fitted values of DAX and REXP volatilities and correlation. The linear model is based on OLS estimates and is given for correlation forecasting by equation (8) and for volatility forecasting by (14). The nonlinear model is based on OLS and neural network estimates and is given for correlation forecasting by equation (12) and for volatility forecasting by (16). * Indicates that the model is significantly different to the performance measure of the historical mean model, based on a 5% significance level *t*-test. † indicates that the performance measure for the nonlinear model is significantly different to the performance measure of the linear model at the 5% significance level.

	RMSE	*MAE*	*CORR*
DAX variances (%)			
Historical mean	7.032	5.085	-
Linear model	6.459*	4.475*	0.501
Nonlinear model	6.171*†	4.619*	0.569†
REXP variances (%)			
Historical mean	0.815	0.652	-
Linear model	0.767*	0.604*	0.456
Nonlinear model	0.693*†	0.584*†	0.658†
DAX, REXP correlation			
Historical mean	0.357	0.294	-
Linear model	0.281*	0.222*	0.603
Nonlinear model	0.277*†	0.220*	0.620

[1] Calculating the correlation measure (*CORR*) for the historical mean model for in-sample performance is infeasible because the prediction value corresponds to the historical mean, leading to a nominator and denominator in equation (22) being close to zero. The values are therefore not reported.

[2] Significance for *CORR* is based on a hypothesis test for equal correlation coefficients at the 5% significance level.

In addition to excess return predictions, we perform forecasts for realized conditional volatilities and correlations of the DAX and REXP indexes. For volatility forecasts the same independent variables are included as for their excess return predictions. In addition, one-month lagged conditional volatilities are included to account for volatility clustering. I.e., the vectors of independent predictive variables are given by $x_{VOL_{DAX},t} = \left[\log VOL_{DAX,t}, DY_t, E/P_t, TERM_t\right]'$ and $x_{VOL_{REXP},t} = \left[\log VOL_{REXP,t}, TERM_t, \Delta MNY_t\right]'$. For predicting correlations, lagged correlation and lagged log variances are used as vector of independent variables, i.e., $x_{\rho,t} = \left[\rho_t \log VOL_{DAX,t} \log VOL_{REXP,t}\right]'$. The in-sample performance results of the fitted values are presented in Table 3.

We observe that the nonlinear model performs again best according to the given performance measures. A single exception is the *MAE* for DAX volatilities, where the *MAE* = 4.48 for the linear model yields a better performance than the *MAE* = 4.62 for the nonlinear model. We find that the linear and nonlinear model is always significantly better than the historical mean. In-sample results presented so far indicated considerable improvement by implementing neural networks for predicting excess returns and conditional volatilities. We now turn to the out-of-sample performances. It is well known from Bossaerts and Hillion (1999), Goyal and Welch (2003; 2008) and Lettau and van Nieuwerburgh (2007) that for linear regressions, even with significant in-sample predictability of financial market data, the out-of-sample prediction properties are generally poor.

Out-of-sample results of predicted values are again evaluated based on the measures *RMSE*, *MAE* and *CORR*. The measures are calculated over all *out-of-sample* predictions – a total of 120 observations.

Table 4. Out-of-sample prediction performance for DAX and REXP excess returns

The table provides average performance measures, given by (20), (21) and (22), for DAX and REXP excess return predictions. The linear model is based on OLS estimates for $\hat{r}_{t+1} = x'_{r,t}\hat{\beta}_r$ and the nonlinear model is based on OLS and neural network estimates for $\hat{r}_{t+1} = x'_{r,t}\hat{\beta}_r + n(x_{r,t}, \psi_r)$. * Indicates that the performance measure for the nonlinear model is significantly different to the performance measure for the historical mean model at the 5% significance level.

	RMSE	*MAE*	*CORR*
DAX excess returns (%)			
Historical mean	6.862	5.249	-0.089
Linear model	6.873	5.216	0.101
Nonlinear model	7.338	5.747	0.152*
REXP excess returns (%)			
Historical mean	0.905	0.741	-0.162
Linear model	0.903	0.741	0.039
Nonlinear model	0.988	0.801	0.138*

Table 4 reports the out-of-sample performance measures. The *RMSE* for DAX excess returns for the nonlinear model is 7.34, for the linear model 6.87 and 6.86 if we take the historical mean as predictive value[1]. Consistent with the literature, we find that the out-of-sample prediction power vanishes. In terms of *RMSE* and *MAE* the nonlinear model underperforms the benchmark models. However, the correlation of predicted values and true values increases. The nonlinear *CORR* results are significantly better than the historical mean model, whereas we find no significant improvement compared to the linear model.

Table 5. Out-of-sample prediction performance for DAX and REXP volatilities and correlation

The table provides average performance measures, given by (20), (21) and (22), for DAX and REXP volatilities and correlation predictions. The linear model is based on OLS estimates and is given by equation (8) for correlation forecasting and by (14) for variance forecasting. The nonlinear model is based on OLS and neural network estimates and is given by equation (12) for correlation forecasting and by (16) for variance forecasting. * Indicates that the performance measure for the model is significantly different to the performance measure for the historical mean model at the 5% significance level. † indicates that the *CORR* performance measure is significantly greater than zero based at the 5% significance level.

	RMSE	*MAE*	*CORR*
DAX volatility (%)			
Historical mean	8.207	6.059	-0.103
Linear model	7.368	5.169	0.616*†
Nonlinear model	6.742	5.086	0.668*†
REXP volatility (%)			
Historical mean	0.745	0.615	-0.186
Linear model	0.684	0.569	0.456*†
Nonlinear model	0.710	0.612	0.463*†
DAX, REXP correlation			
Historical mean	0.467	0.399	0.248†
Linear model	0.316	0.250	0.383†
Nonlinear model	0.315	0.254	0.371†

The out-of-sample prediction results for volatilities and correlation are presented in Table 5. DAX volatilities are best predicted by the nonlinear model as all three performance measures indicate. The *RMSE* is 6.74, compared to 7.37 for the linear model and 8.21 for the historical mean. A joint *F*-test shows that the nonlinear model outperforms the historical mean model at the 5% significance level.

For REXP volatility predictions we see that the *RMSE* and *MAE* are the lowest for the linear model. The correlation is the largest for the nonlinear model (*CORR* = 0.46) showing some ambiguity to rank the

[1] Note that the performance measures for DAX excess returns of the nonlinear model are based on two removed outliers, which gave clearly unrealistic return predictions of 60% and -42%.

models even though the linear and nonlinear models outperform the historical mean. We find the same result for correlation forecasts. The linear and nonlinear models outperform the historical mean.

Before we discard our prediction results as their statistical out-of-sample accuracy is weak and somewhat ambiguous (i.e., weak *RMSE* and *MAE*, but improving *CORR* measure) we analyze the contribution to the allocation of assets by a mean-variance investor. Therefore, in the next section we let a mean-variance investor use our prediction results to compute optimal asset weights and we analyze whether the predictions are economically exploitable.

4.3. Asset allocation results. To evaluate the ultimate value of our predictions, we use our out-of-sample predictions as inputs for portfolio construction. Consider the mean-variance investor from section 3 with one-month investment horizons. The asset weights w_t are the solution of equation (4). Those mean-variance asset allocations are solely based on out-of-sample predictions. The predictions are used as conditional expectations for the necessary parameters in mean-variance optimization. The conditional expectations for excess returns $\hat{r}_{t+1}$ and covariances $\hat{\Omega}_{t+1}$ are plugged into the mean-variance optimization.

We use our 120 out-of-sample monthly predictions for 120 out-of-sample portfolios covering the period from January 1998 to December 2007. We focus on three kinds of investors. The first investor is an "uninformed" investor who does not use prediction models for conditional expectations. This investor uses the historical mean for expected returns and the covariance matrix. The second investor relies on recursive *linear* prediction models for conditional parameter expectations and the third investor uses recursive *nonlinear* prediction models. For comparison, we additionally evaluate equally weighted portfolios.

We assume $\gamma = 6$ for the investor's risk aversion, which represents a moderately risk averse investor. We also allow or exclude short-sales constraints, i.e. we restrict the portfolio weights to be positive and the sum of weights to be less or equal to one.

For evaluating portfolio performances, we calculate Sharpe ratio as the ratio of mean excess portfolio returns and the portfolio's standard deviation. However, for time-varying volatility, Sharpe ratio overestimates portfolio risk. We report Sharpe ratio as a common measure of portfolio performance as well as average realized utility $\bar{U}$. The utility-based performance measure is an estimate of the portfolio's economic value. It states the investor's certainty equivalent, i.e., the certain return that provides the same utility to the investor as the risky portfolio[1]. Average realized utility $\bar{U}$ is obtained by:

$$\bar{U} = \frac{1}{T}\sum_{t=1}^{T-1}\left(R_{t+1}^{p} - \frac{\gamma}{2} w_t' \Omega_{t+1} w_t\right), \tag{23}$$

where R_{t+1}^{p} is the realized return of the portfolio and Ω_{t+1} is the realized covariance matrix in month $t + 1$. A utility-based measure of portfolio performance in a linear setting is used by Marquering and Verbeek (2004) and for volatility timing by Fleming et al. (2001). This approach provides a good measure to compare different portfolio strategies. Furthermore, we report a common measure from the literature to evaluate the outperformance of a portfolio strategy which is provided by Jensen's alpha (Jensen, 1969). Alpha represents outperformance and is calculated as the estimated intercept in a regression of excess returns upon the market's excess returns:

$$r_t^p = \alpha + \beta r_t^m + \varepsilon_t. \tag{24}$$

An advantage of the performance measure α is that it does not include a simple leverage of the market portfolio (which is absorbed by β). In addition, we can use standard regression analysis to evaluate this measure. A standard *t*-test can be employed for testing whether *a* is significantly larger then zero. It is reasonable for an investor to invest a part of his wealth in the portfolio when *a* has a positive value.

However, market-timing implies that beta varies over time. In this case, Jensen's alpha is a biased measure because beta is not constant and depends on market excess returns. Treynor and Mazuy (1966) suggest to use a measure that adds a quadratic term in the fitting formula[2]. The Treynor and Mazuy measure is thus given by:

$$r_t^p = \alpha + \beta_0 r_t^m + \beta_1 (r_t^m)^2 + \varepsilon_t. \tag{25}$$

Including the quadratic term of market excess returns implies that beta depends linearly on market excess returns. Consequently, we can explain the Treynor and Mazuy measure by plugging $\beta = \beta_0 + \beta_1 r_t^m$ into equation (24). If β_1 has positive value, the equation describes a convex and upward-sloping regression line. Positive slope of squared market excess returns (i.e., a significant positive value of β_1) indicates successful market timing. That means a positive coefficient indicates higher returns in up markets and less negative returns in down markets. We can also use *t*-statistics to evaluate the

[1] This criterion for ranking the performance of prediction models on a utility-based measure is suggested by West, Edison and Cho (1993). The proposed measure utilizes the close relationship between mean-variance analysis and quadratic utility.

[2] See Treynor and Mazuy (1966, p. 134).

regression model. We also report asset allocation results with short selling constraints, i.e. restricting the sum of risky asset weights to the interval [0; 1]. Imposing short selling restrictions avoids extreme asset weights and is common practice for portfolio optimization. Markowitz (1959) and Rudolf (1994) discuss techniques to optimize portfolios including asset weights being subject to specific constraints.

Table 6 presents out-of-sample asset allocation results when excess returns and the elements in the covariance matrix are predicted by the respective forecasting models. Using historical means of moments for portfolio construction (strategy 3 and 4) yields a Sharpe ratio of 0.08 for the unconstrained case and 0.02 when asset weights are restricted to [0; 1].

Table 6. Asset allocation evaluation: equities, bonds and risk free asset

The table provides out-of-sample results of asset allocation performances including monthly excess returns, monthly standard deviation, Sharpe ratio, monthly average utility assuming a relative risk aversion of $\gamma = 6$, Jensen's alpha, its *t*-statistics and *t*-statistics for the Treynor Mazuy measure. The investor can allocate his wealth to the DAX equity index, the REXP bond index and a risk free asset. Excess returns, standard deviation and Jensen's alpha are in percent.

	Avg. exc. return	Std. dev.	Sharpe ratio	Avg. utility	J's alpha	(*t*-stat.)	TM (*t*-stat.)
1. 33/33/33	0.1335	2.1614	0.0618	-0.0004			
2. 50/50	0.2023	3.2749	0.0618	-0.0019			
3. Hist. moments	0.2488	3.1643	0.0786	-0.0010			
4. Hist. moments [0; 1]	0.0586	2.2597	0.0226	-0.0016			
5. μ_{linear}	1.2569	12.2141	0.1029	-0.0255	0.7208	0.7735	-1.1764
6. $\mu_{nonlinear}$	9.2956	32.7581	0.2838	-0.8492	8.7567	2.9727	-0.5098
7. μ_{linear} [0;1]	0.5685	3.9226	0.1449	0.0005	0.4955	2.4363	-0.3365
8. $\mu_{nonlinear}$ [0;1]	1.1473	4.2794	0.2681	0.0056	1.0836	3.6751	3.8734
9. $(\mu, \sigma)_{linear}$	1.6845	7.8473	0.2147	-0.0005	1.3848	2.1953	-1.6475
10. $(\mu,\sigma)_{nonlinear}$	4.1675	16.6644	0.2501	-0.1386	3.7143	2.5832	-1.0147
11. $(\mu, \sigma)_{linear}$ [0;1]	0.5367	3.2275	0.1663	0.0022	0.4826	2.4372	-2.0535
12. $(\mu, \sigma)_{nonlinear}$ [0;1]	1.0685	3.9115	0.2732	0.0056	1.0109	3.7190	3.9119

In terms of Sharpe ratio, the unconstrained approach produces slightly better performances than the naive equally weighted strategies 1 (denoting one third of wealth to each asset) and 2 (denoting 50% of wealth to each risky asset). Average realized utility is -0.0004 and -0.0019 for strategies 1 and 2 respectively. In terms of this measure, strategies 3 and 4 perform better than strategy 2, but not in comparison to strategy 1. Strategies 5 to 8 represent results for asset allocations where excess returns are predicted by the linear and nonlinear models. Strategies 9 to 12 also include predictions for volatility. An investor who times the market by a nonlinear model generates significant positive alpha for constrained and unconstrained asset allocation strategies. Also, Sharpe ratio is the highest when expectations are formed by nonlinear predictions. This is true for constrained and unconstrained asset allocations compared to linear predictions and historical means. If for portfolio optimization only excess return predictions are considered (and the sample mean covariance matrix), Sharpe ratios for nonlinear predictions are 0.28 for the unconstrained case and 0.27 for the constrained case. For linear predictions, Sharpe ratio is 0.10 for the unconstrained case and 0.14 for the constrained case. For linear predictions, Jensen's alpha is significant only for constrained portfolios, whereas for nonlinear return predictions alpha of 8.75% (unconstrained) and 1.08% (constrained) are both significant. The Treynor and Mazuy measure is significantly positive only for the constrained case (strategy 8).

If we include volatility and correlation forecasts (strategies 9 to 12), Jensen's alpha is significantly positive for the constrained and unconstrained cases. In this case, however, the Treynor and Mazuy measure is positively significant only for nonlinear predictions and constrained asset weights (strategy 12).

The average realized utility is strongly affected by portfolio constraints. Unconstrained portfolios based on conditional predictions perform poorly in terms of average realized utility. However, if asset weights are restricted, the performance is superior compared to the uninformed strategies 1 to 4. Constrained portfolios based on nonlinear predictions lead to the best performances. Switching from strategy 7 to 11 by including linear covariance predictions increases average realized utility from 0.0005 to 0.0022. By contrast, average realized utility remains 0.0056 when an investor switches from strategy 8 to 12 by including nonlinear covariance predictions. This indicates that nonlinear covariance predictions have no economic value to mean-variance investors in terms of average realized utility.

Imposing constraints on asset weights seems reasonable in the light of extreme unconstrained asset weights[1]. Constrained portfolios based on nonlinear predictions have larger Sharpe ratios than their linear and historical mean counterparts, larger average realized utility, larger alpha values and have significantly positive Treynor and Mazuy measures in contrast to the linear predictions. The asset allocations returns for strategy 7 are plotted against constrained historical mean portfolio returns in Figure 1. The fitted Treynor and Mazuy regression is also shown. The graph indicates that the nonlinear investment strategy leads to favorable market timing. The convex regression line shows that asset allocations returns are higher in up markets and less negative in down markets.

Whereas Marquering and Verbeek (2004) find that linear prediction models for S&P 500 returns and volatility lead to a significant coefficient in the Treynor and Mazuy measure, we find the contrary in our linear prediction models for DAX and REXP returns and volatilities. The Treynor and Mazuy measure is only significantly positive if returns are predicted by the nonlinear model.

In order to analyze to which period we can attribute the asset allocation performances, it is helpful to look at out-of-sample portfolios for several subsamples. This question is investigated by plotting the capital market line (CML) based on portfolios constructed by using historical mean estimates. Portfolio standard deviations and returns are then calculated based on average *realized* returns and risks. We also plot the average *realized* portfolio standard deviation and return for constrained linear and nonlinear portfolios (i.e., strategies 11 and 12).

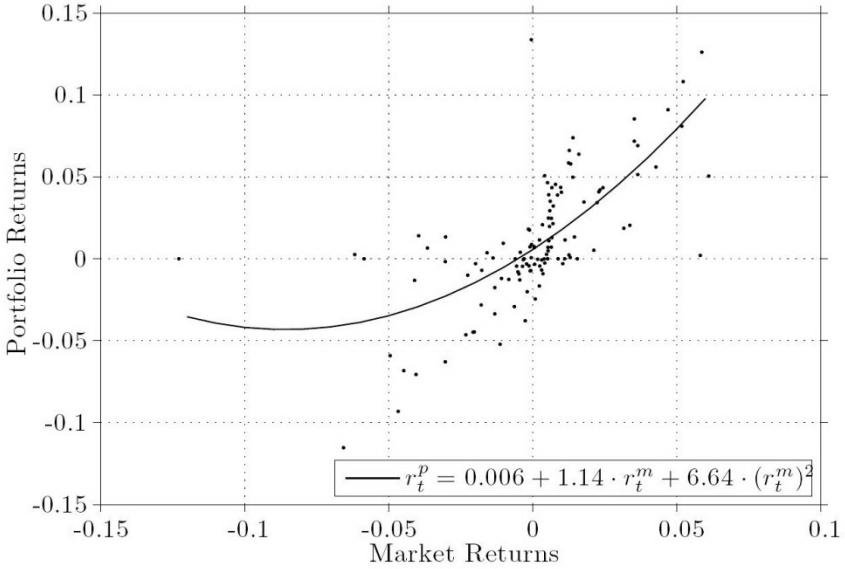

Fig. 1. Empirical Treynor and Mazuy measure for strategy 12

The sample is equally divided into six sub-samples, such that each sub-sample includes a period of 20 months. The results are graphically shown in Figure 2. Portfolios based on linear predictions are indicated by a cross, portfolios based on nonlinear predictions are indicated by a circle. A portfolio located above the historical mean CML indicates an outperforming risk-return-profile. Figure 2a reveals that the constrained nonlinear portfolio delivers a better out-of-sample risk-return-profile than the historical mean portfolios during the period from January 1998 to August 1999. By contrast, the linear portfolio performs worse in this period. We observe that the out-of-sample nonlinear portfolios are superior during the period from January 1998 to August 1999 and during January 2003 to December 2007. The constrained linear portfolio outperforms the historical mean only during the last two sub-periods (September 2004 to December 2007). The linear portfolio strategy outperforms the nonlinear strategy only during the period from May 2001 to December 2002. The linear portfolio strategy is dominated by the nonlinear strategy during all other periods.

Conclusion

The aim of our paper is to scrutinize, whether nonlinear predictability of returns and volatilities has economic relevance. Several papers present artificial neural networks as promising nonlinear prediction method and find some nonlinear dependencies in economic time series. We used this approach to perform one-month ahead predictions for excess returns and the covariance matrix of the German equity index DAX and the bond index REXP.

[1] Unconstrained asset weights based on linear predictions range from -280% to 803% and asset weights based on nonlinear predictions range from -6,192% to 2,202%.

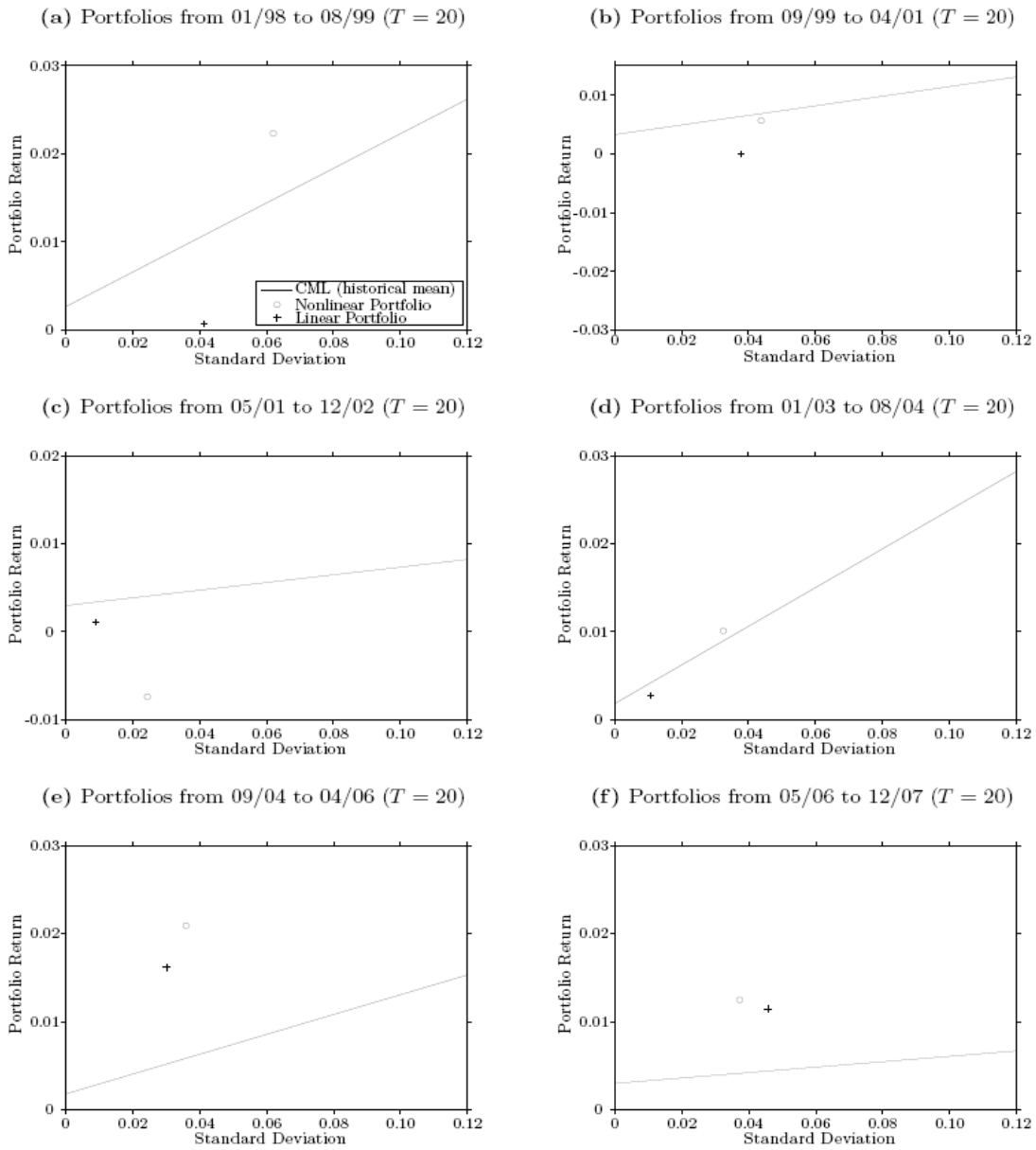

Fig. 2. Out-of-sample portfolios of strategy 11 as well as 12 and CML

We considered unconditional asset allocations and linear predictions as benchmark. Our results suggest that accounting for nonlinearities in prediction models for conditional asset allocations lead to exploitable improvements to risk-averse investors.

In summary, several conclusions can be drawn. Prediction accuracies from in-sample statistics vanish for out-of-sample data. Despite weak proof for out-of-sample prediction accuracy, Table 6 shows that out-of-sample asset allocation results improve significantly in terms of Jensen's alpha and Sharpe ratio for linear and nonlinear predictions. The Treynor Mazuy measure is significantly positive and average utility improves when asset weight constraints are imposed. Overall, the results from Figure 2 indicate that most of the success of the timing strategies are attributed to the time period from September 2004 to December 2007, and, in addition for the nonlinear strategy, from January 1998 to August 1999.

The results show that detecting nonlinear dependencies performed by artificial neural networks are exploitable and contribute to improved asset allocation performances. They further improve asset allocation results compared to linear predictions and historical moments in terms of Sharpe ratio and Jensen's alpha. Furthermore, when portfolio weights are constrained, nonlinear predictions lead to higher average realized utility for mean-variance investors and to a significantly positive Treynor Mazuy measure.

References

1. Abhyankar, A., Copeland, L.S. and Wong, W. (1997). Uncovering nonlinear structure in real-time-stock-market indexes: the S&P 500, the DAX, the Nikkei 225, and the FTSE-100, *Journal of Business & Economic Statistics*, 15, pp. 1-14.
2. Andersen, T.G., Bollerslev, T., Christoffersen, P.F. and Diebold, F.X. (2006). Volatility and correlation forecasting, *in* G. Elliot, C.W.J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 1, Elsevier, Chapter 15, pp. 777-0878.
3. Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2009). Parametric and nonparametric measurements of volatility, *in* Y. Ait-Sahalia and L.P. Hansen (eds), *Handbook of Financial Econometrics*, Elsevier.
4. Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2001). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, 96, pp. 42-55.
5. Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2003). Modeling and forecasting realized volatility, *Econometrica*, 71, pp. 579-625.
6. Ang, A. and Bekaert, G. (2007). Stock return predictability: is it there?, *Review of Financial Studies*, 20, pp. 651-707.
7. Barone-Adesi, G. and Whaley, R.E. (1987). Efficient analytic approximation of american option values, *Journal of Finance*, 42, pp. 142-155.
8. Bekiros, S.D. and Georgoutsos, D.A. (2008). Direction-of-change forecasting using a volatility based recurrent neural network, *Journal of Forecasting*, 27, pp. 407-417.
9. Bishop, C.M. (1995). *Neural networks for pattern recognition*, Clarendon Press, Oxford.
10. Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, 31, pp. 307-327.
11. Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992). Arch modeling in finance: A selecctive review of the theory and empirical evidence, *Journal of Econometrics*, 52, pp. 5-59.
12. Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994). Arch models, *in* R.F. Engle and D. McFadden (eds), *Handbook of Econometrics*, Vol. 4, North-Holland, Amsterdam.
13. Bossaerts, P. and Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: what do we learn? *Review of Financial Studies,* 12 (2), pp. 405-428.
14. Campbell, J.Y. (1987). Stock returns and the term structure, *Journal of Financial Economics*, 18, pp. 373-399.
15. Campbell, J.Y. and Shiller, R.J. (1988a). The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1, pp. 195-228.
16. Campbell, J.Y. and Shiller, R.J. (1988b). Stock prices, earnings, and expected dividends, *Journal of Finance*, 43, pp. 661-676.
17. Campbell, J.Y. and Thompson, S.B. (2008). Predicting excess returns out of sample: can anything beat the historical average? *Review of Financial Studies*, 21, pp. 1509-1531.
18. Campbell, J.Y. and Yogo, M. (2006). Efficient tests of stock return predictability, *Journal of Financial Economics*, 81, pp. 27-60.
19. Chan, L.K.C., Karceski, J. and Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model, *Review of Financial Studies*, 12, pp. 937-974.
20. Cochrane, J.H. (1999). New facts in finance, *Economic Perspectives*, 23, pp. 36-58.
21. Cochrane, J.H. (2008). The dog that did not bark: A defense of return predictability, *Review of Financial Studies*, 21, pp. 1533-1575.
22. Desai, V.S. and Bharati, R. (1998). The efficacy of neural networks in predicting returns on stock and bond indices, *Decision Science*, 29, pp. 405-425.
23. Diebold, F.X. (2004). The Nobel Memorial Prize for Robert F. Engle, *Scandinavian Journal of Economics*, 106, pp. 165-185.
24. Elton, E.J., Gruber, M.J. and Spitzer, J. (2006). Improved estimates of correlation coefficients and their impact on optimal portfolios, *European Financial Management*, 12, pp. 303-318.
25. Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of the U.K inflation, *Econometrica*, 50, pp. 987-1008.
26. Engle, R.F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics, *Journal of Economic Perspectives*, 15, pp. 157-168.
27. Engle, R.F. (2004). Nobel lecture. Risk and volatility: Econometric models and financial practice, *American Economic Review*, 94, pp. 405-420.
28. Fama, E.F. and French, K.F. (1988). Dividend yields and expected stock returns, *Journal of Financial Economics*, 22, pp. 3-25.
29. Fama, E.F. and French, K.R. (1989). Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics*, 25, pp. 23-49.
30. Fama, E.F. and Schwert, G.W. (1977). Asset returns and inflation, *Journal of Financial Economics*, 5, pp. 115-146.
31. Ferland, R. and Lalancette, S. (2006). Dynamics of realized volatilities and correlations: an empirical study, *Journal of Banking & Finance*, 30, pp. 2109-2130.

32. Ferson, W.E. and Harvey, C.R. (1993). The risk and predictability of international returns, *Review of Financial Studies*, 6, pp. 527-566.
33. Fleming, J., Kirby, C. and Ostdiek, B. (2001). The economic value of volatility timing, *Journal of Finance*, 56, pp. 329-352.
34. Foresee, D.F. and Hagan, M. (1997). *Gauss-Newton approximation to Bayesian learning*, pp. 1930-1935.
35. French, K.R., Schwert, G.W. and Stambaugh, R.F. (1987). Expected stock returns and volatility, *Journal of Financial Economics*, 19, pp. 3-29.
36. Ghysels, E., Santa-Clara, P. and Valkanov, R. (2005). There is a risk-return trade-off after all, *Journal of Financial Economics*, 76, pp. 509-548.
37. Goyal, A. and Welch, I. (2003). Predicting the equity premium with dividend ratios, *Management Science*, 49, pp. 639-654.
38. Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies*, 21, pp. 1455-1508.
39. Guo, H. (2006). On the out-of-sample predictability of stock market returns, *Journal of Business*, 79, pp. 645-670.
40. Hamid, S.A. and Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 Index future prices, *Journal of Business Research*, 57, pp. 1116-1125.
41. Hansen, P.R. and Lunde, A. (2006). Consitent ranking of volatility models, *Journal of Econometrics*, 131, pp. 97-121.
42. Harvey, A.C. (1974). Estimating regression models with multiplicative heteroscedasticity, *Econometrica*, 44, pp. 461-465.
43. Ince, H. and Trafalis, T.B. (2007). Kernel principal component analysis and support vector machines for stock price prediction, *IEE Transactions*, 39, pp. 629-637.
44. Jensen, M.C. (1969). Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios, *Journal of Business*, 42 (2), pp. 167-247.
45. Keim, D.B. and Stambaugh, R.F. (1986). Predicting returns in the stock and bond market, *Journal of Financial Economics*, 17, pp. 357-390.
46. Kothari, S. and Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time series analysis, *Journal of Financial Economics*, 44, pp. 169-203.
47. Kuan, C.-M. and White, H. (1994). Artificial neural networks: An econometric perspective, *Econometric Reviews*, 13, pp. 1-94.
48. Lamont, O. (1998). Earnings and expected returns, *Journal of Finance*, 53, pp. 1563-1587.
49. Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance*, 10, pp. 603-621.
50. Lee, T.-H., White, H. and Granger, C.W.J. (1993). Testing for neglected nonlinearity in time series models: a comparison of neural network methods and alternative tests, *Journal of Econometrics*, 56, pp. 269-290.
51. Lettau, M. and Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns, *Journal of Finance*, 56, pp. 815-849.
52. Lettau, M. and van Nieuwerburgh, S. (2007). Reconciling the Return Predictability Evidence, *Review of Financial Studies*, 21 (4), pp. 1607-1652.
53. Lewellen, J. (2004). Predicting returns with financial ratios, *Journal of Financial Economics*, 74, pp. 209-235.
54. Ludvigson, S.C. and Ng, S. (2007). The empirical risk-return relation: A factor analysis approach, *Journal of Financial Economics*, 83, pp. 171-222.
55. MacKay, D.J. (1992). Bayesian interpolation, *MIT Press*, pp. 415-447.
56. Markowitz, H. (1959). *Portfolio selection; efficient diversification of investments*, Wiley, New York.
57. Marquering, W. and Verbeek, M. (2004). The economic value of predicting stock index returns and volatility, *Journal of Financial and Quantitative Analysis*, 39, pp. 407-429.
58. Merton, R.C. (1980). On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics*, 8, pp. 323-361.
59. Pesaran, M.H. and Timmermann, A. (1995). Predictability of stock returns: robustness and economic significance, *Journal of Finance*, 50, pp. 1201-1228.
60. Pontiff, J. and Schall, L.D. (1998). Book-to-market ratios as predictors of market returns, *Journal of Financial Economics*, 43, pp. 141-160.
61. Poon, S.-H. and Granger, C.W.J. (2003). Forecasting volatility in financial markets: A review, *Journal of Economic Literature*, 41, pp. 478-539.
62. Qi, M. (1999). Nonlinear predictability of stock returns using financial and economic variables, *Journal of Business & Economic Statistics*, 17, pp. 419-429.
63. Qi, M. and Maddala, G.S. (1999). Economic factors and the stock market: A new perspective, *Journal of Forecasting*, 18, pp. 151-166.
64. Rozeff, M.S. (1984). Dividend yields are equity risk premiums, *Journal of Portfolio Management*, 11, pp. 68-75.
65. Rudolf, M. (1994). *Algorithms for portfolio optimization and portfolio insurance*, Verlag Paul Haupt, Bern u.a.
66. Schwert, G.W. (1989). Why does stock market volatility change over time? *Journal of Finance*, 44, pp. 1115-1153.
67. Stambaugh, R. (1999). Predicitve regressions, *Journal of Financial Economics*, 54, pp. 375-421.

68. Swanson, N.R. and White, H. (1995). A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks, *Journal of Business and Economic Statistics*, 13, pp. 265-275.
69. Swanson, N.R. and White, H. (1997). A model selection approach to real-time macroeconic forecasting using linear models and artificial neural networks, *Review of Economics and Statistics*, 79, pp. 540-550.
70. TerËasvirta, T. (2006). Forecasting economic variables with nonlinear models, *in* G. Elliott, C.W.J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 1, Elsevier, Chapter 8, pp. 413-457.
71. Treynor, J.L. and Mazuy, K. (1966). ”Can mutual funds outguess the market?” *Harvard Business Review*, 44, pp. 131-136.
72. Tsay, R.S. (2002). Nonlinear models and forecasting, *in* M. Clements and D. Hendry (eds), *A Companion to Economic Forecasting*, Blackwell, pp. 453-484.
73. Valkanov, R. (2003). Long-horizon regressions: theoretical results and applications, *Journal of Financial Economics*, 68, pp. 201-232.
74. West, K.D., Edison, H.J. and Cho, D. (1993). A utility-based comparison of some models of exchange rate volatility, *Journal of International Economics*, 35 (1/2), pp. 23-45.
75. White, H. (2006). Approximate nonlinear forecasting methods, *in* G. Elliott, C.W.J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 1, Elsevier, chapter 9, pp. 459-512.
76. Whitelaw, R.F. (1994). Time variations and covariations in the expectation and volatility of stock market returns, *Journal of Finance*, 49, pp. 515-541.