



**COLLANA DEL  
DIPARTIMENTO DI ECONOMIA**

**MULTIVARIATE STATISTICAL ANALYSIS FOR  
PORTFOLIO SELECTION OF ITALIAN STOCK MARKET**

Alessia Naccarato - Andrea Pierini

**ISSN 2279-6916** Working papers

(Dipartimento di Economia Università degli studi Roma Tre) (online)

---

Working Paper n° 166, 2012

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

esemplare fuori commercio  
ai sensi della legge 14 aprile 2004 n.106

**REDAZIONE:**

Dipartimento di Economia

Università degli Studi Roma Tre

Via Silvio D'Amico, 77 - 00145 Roma

Tel. 0039-06-57335655 fax 0039-06-57335771

E-mail: [dip\\_eco@uniroma3.it](mailto:dip_eco@uniroma3.it)



**DIPARTIMENTO DI ECONOMIA**

**MULTIVARIATE STATISTICAL ANALYSIS FOR  
PORTFOLIO SELECTION OF ITALIAN STOCK MARKET**

Alessia Naccarato - Andrea Pierini

*Comitato Scientifico:*

*Fabrizio De Filippis*

*Anna Giunta*

*Paolo Lazzara*

*Loretta Mastroeni*

*Silvia Terzi*

# Multivariate Statistical Analysis for Portfolio Selection of Italian stock market

Alessia Naccarato

*Department of Economics, University of Roma Tre*

*E-mail: alessia.naccarato@uniroma3.it*

Andrea Pierini

*Department of Economics, University of Roma Tre*

*E-mail: apierini@uniroma3.it*

*Abstract:* The use of bivariate cointegrated vector autoregressive models and Baba-Engle-Kraft-Kroner models ( Engle *et al.* 1995), is proposed for the selection of a stock portfolio (Markowitz type portfolio) based on estimates of average returns on shares and the volatility of share prices. The model put forward envisages the use of explicative variables. This article employs the *intrinsic value of shares* as a variable, which will make it possible to take the theory of value into account. The model put forward is applied to a series of data regarding the prices of 150 shares traded on the Italian stock market.

*Keywords:* Markowitz Portfolio, Cointegrated Vector Autoregressive Models, BEKK Model.

*JEL code:* C580

## ***1. Introduction***\*

The selection of a stock portfolio is broadly discussed in the literature, generally with reference to heteroskedastic regression models (Bollerslev *et al.*, 1994). The model used in the case of multiple time series is of the vector autoregressive (VAR) type and rests on the

---

\* We are grateful to Prof. S. Terzi for valuable suggestions on this article and for her constructive comments

predictability of the average return on shares (Brown and Reily, 2008, Hamilton, 1994).

In particular this paper suggests the use of cointegrated vector autoregressive models (CVAR) and Baba-Engle-Kraft-Kroner models (BEKK) for the selection of a stock portfolio. In other words, it addresses the problem of estimating average returns and the associated risk on the basis of the prices of a certain number of shares over time. This estimate is then used to identify the assets offering the best performance and hence constituting the best investments. While Campbell *et al.* (2003) proposes the use of a VAR(1) model, it is suggested here that use should be made of VEC models, which make it possible to take into account any cointegration between the series employed and the market trend as measured by means of the Thomson Reuters *Datastream Global Equity Italy Index* (Datastream 2008).

Moreover, while Bollerslev, Engle and Wooldridge (1988) employ diagonal vectorization (DVEC) models to estimate share volatility, the use of a BEKK model, as proposed here, makes it possible to extend the estimation procedure based on DVEC models so as to take into account also the correlation between the volatility of the series and the volatility of the market trend.

The series considered regard the Italian stock market (BIT), and specifically the monthly figures for the top 150 shares in terms of capitalization, from 1 January 1975 to 31 August 2011. The estimation procedure proposed for portfolio selection involves two steps.

In the first step, a two-dimensional CVAR model is developed for all of the 150 shares considered in order to obtain an estimate of the average stock market return. A BEKK model is then applied to the series of residuals thus obtained in order to estimate the volatility of the series. The BEKK model appears particularly suitable because it does not entail the condition of normality for the accidental component of the model (Hamilton, 1994).

The second step regards the selection of shares for inclusion in the portfolio. Only those identified as presenting positive average returns during the first phase are considered eligible. For the purpose of selecting the most suitable of these, a new endogenous variable is constructed as the product of two further elements, namely the price-to-

earnings ratio (P/E) and earnings per share (EPS). This variable, which indicates the “intrinsic value” of the share in question, is not constructed for the entire set of 150 shares but only for those presenting positive average returns in the first phase, as it would be pointless in the case of negative returns.

The CVAR-BEKK model is applied once again to this new series in order to estimate the intrinsic value of the shares, and the top 10, (Evans and Archer, 1968), are selected for inclusion in the portfolio on the basis of the difference between this intrinsic value and the price estimated in the first phase (Brown and Reily, 2008).

A quadratic programming model is then employed to determine the quantities to be bought of each of the 10 shares selected.

It should be noted that the variable  $P/E \cdot EPS$  is estimated for each industrial sector (Nicholson, 1960).

## 2. The Cointegrated Vector Autoregressive Models

A concise outline is now given of the phases involved in the selection of shares for inclusion in the portfolio as well as the quantity of shares to be bought for each type selected. The starting point is the  $K = 150$  series, regarding the average returns  $R_{k,t}$  on the shares, and the average return of the market  $R_{M,t}$ ,  $t = t_k, \dots, T$ ,  $k = 1, \dots, K$ . It should be noted in this connection that the length of the series considered is not homogeneous because not all of the joint-stock companies are quoted as from the same point in time. This aspect involves further complications in the estimation procedure.

*Step one* : For each series, the model  $CVAR(p)$  is defined for the random vector  $y_t = [y_{1,t}, y_{2,t}] = [R_{k,t}, R_{M,t}]$  as:

$$y_t = \mu_t + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t \quad (1)$$

with  $\mu_t = \mu_0 + \mu_1 \cdot t$ ,  $A_i$  ( $i=1, \dots, p$ ) is a  $(2 \times 2)$  of the unknown coefficients, and  $u_t = [u_{1,t}, u_{2,t}]'$  is the vector of errors such that  $u_t \approx N(0, \Sigma_u)$ .

Model (1) can be rewritten as follows to take into account a possible cointegration of the variables considered:

$$\Delta y_t = \mu_t + \Pi y_{t-1} + \Gamma_1 y_{t-1} + \Gamma_2 y_{t-2} + \dots + \Gamma_{p-1} y_{t-p+1} + u_t \quad (2)$$

where

$$\Gamma_i = -(A_{i+1} + \dots + A_p) \quad (3)$$

and

$$\Pi = -(I - A_1 - \dots - A_p) \quad (4)$$

It has to be noticed that if  $\Pi$  is singular,  $y_1 = [y_{1,1}, \dots, y_{1,T}]$  and  $y_2 = [y_{2,1}, \dots, y_{2,T}]$  are cointegrated (Johansen, 1995, Lutkepohl, 2007).

In specifying the CVAR model the lag order and the cointegration rank have to be determined. We start by determining a suitable lag length because, in choosing the lag order, the cointegration rank does not have to be known.

The AIC criterion is used to estimate the lag  $\hat{p}$ , with reference to model (1):

$$\hat{p} = \arg \min_m C(m) = \min_m \left\{ \ln \left( \det(\tilde{\Sigma}_u(m)) + \frac{mc_T}{T} : m = 0, 1, \dots, p_{\max} \right) \right\} \quad (5)$$

where  $\tilde{\Sigma}_u(m)$  is the maximum likelihood estimate (MLE) of  $\Sigma_u(m)$  for a VAR( $m$ ) of type (1) with a sample of breadth  $T - t_k + m$  and  $m$  values of initialization, with  $c_T = 8$ ,  $p_{\max} = 10$ .

Notice that, while we have considered specifying the VAR order  $p$ , the criterion is also applicable for choosing the number of lagged differences in a VEC model (2) because  $p-1$  lagged differences in a VEC correspond to a VAR order  $p$  (Lutkepohl, 2007).

In practice it is common to use statistical tests in specifying the cointegration rank. In this framework to ascertain the presence of cointegration in the model (2), the likelihood ratio test (LR) is used:

$$LR(r_0) = -T \sum_{j=r_0+1} \log(1 - \lambda_j) \quad (6)$$

where:  $r_0 = rk(\Pi) = 0, 1$  and  $\lambda_j$  are the eigenvalues of the matrix

$$S_{11}^{-1/2} S_{10} S_{00}^{-1} S_{01} S_{11}^{-1/2} \quad \text{with} \quad S_{00} = \frac{1}{T-p} \sum_{t=p+1}^T \hat{u}_t \hat{u}_t', \quad S_{01} = \frac{1}{T-p} \sum_{t=p+1}^T \hat{u}_t \hat{v}_t',$$

$$S_{11} = \frac{1}{T-p} \sum_{t=p+1}^T \hat{v}_t \hat{v}_t'.$$

The quantities  $\hat{u}_t$ ,  $\hat{v}_t$  are the residuals of the regressions of  $\Delta y_t$  and  $y_{t-1}$  estimated by maximum likelihood (Johansen, 1995).

However, the asymptotic distribution of the LR statistic is nonstandard, in particular it is not a chi-squared distribution. In other words, the limiting distribution is a functional of a standard Wiener process. Percentage points of the asymptotic distribution, thus, critical values for the LR test can be generated considering multivariate random walks (Johansen *et al.*, 1990).

Assessment of the presence of cointegration between the series by means of the LR test is followed by estimation of the parameters of the model. The result of the LR test is considered in deciding whether to adopt the model in form (1) or (2). In particular, if the test shows that

the rank of matrix  $\Pi$  is equal to 0 (hypothesis of stationarity of  $\Delta y_t$ ), then  $\Pi = 0$  and the method of maximum likelihood is applied directly to (2) in order to estimate the parameters  $\mu_0, \mu_1$  and  $\Gamma_1, \dots, \Gamma_{p-1}$ .

If it instead transpires that the rank of  $\Pi$  is equal to 1 (hypothesis of cointegration of  $y_1$  and  $y_2$ ), then  $\Pi = \alpha\beta$ . In this case, it is necessary to estimate model (2) in two stages. First, an MLE of  $\beta$  is obtained by concentrating the log-likelihood with respect to  $\beta$ . Second, this estimate is inserted into (2) in order to obtain the MLE of the other parameters (Johansen, 1995).

If  $rk(\Pi) = 2$ , the method of maximum likelihood is applied directly to (1) in order to obtain estimates of the parameters  $\mu_0, \mu_1$  and  $A_1, \dots, A_p$ .

The portmanteau test is used to ascertain the presence of correlation of residuals, the generalized Lomnicki-Jarque-Bera test for the normality of residuals, and the ARCH test to determine heteroskedasticity

In order to forecast it is convenient in this framework to use the levels VAR representation. Therefore we consider the model type (1) with integrated and possibly cointegrated variables replacing the coefficients with their estimates as calculated before:

$$\hat{y}_{T+1|T} = \hat{\mu}_{T+1} + \hat{A}_1 y_T + \dots + \hat{A}_p y_{T+1-p} \quad (7)$$

### 3. The BEKK model for heteroskedasticity

In the event of the latter test revealing the presence of heteroskedasticity, the *BEKK(1,1)* model is used to estimate the conditional variance-covariance matrix

$$\Sigma_{t|t-1} = \text{cov}(u_t | \text{past}) = (\sigma_{i,j}(t)) \quad i, j = 1, \dots, n$$

which has the following structure:

$$\begin{aligned}
\begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} &= \begin{bmatrix} c_{0,11} & 0 \\ c_{0,21} & c_{0,22} \end{bmatrix} \begin{bmatrix} c_{0,11} & c_{0,21} \\ 0 & c_{0,22} \end{bmatrix} + \\
\begin{bmatrix} c_{1,11} & c_{1,21} \\ c_{1,21} & c_{1,22} \end{bmatrix} &\begin{bmatrix} u_{1,t}^2 & u_{1,t-1}u_{2,t-1} \\ u_{2,t-1}u_{1,t-1} & u_{2,t}^2 \end{bmatrix} \begin{bmatrix} c_{1,11} & c_{1,21} \\ c_{1,21} & c_{1,22} \end{bmatrix} + \quad (8) \\
\begin{bmatrix} c_{2,11} & c_{2,21} \\ c_{2,21} & c_{2,22} \end{bmatrix} &\begin{bmatrix} \sigma_{11,t}^2 & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t}^2 \end{bmatrix} \begin{bmatrix} c_{2,11} & c_{2,21} \\ c_{2,21} & c_{2,22} \end{bmatrix}
\end{aligned}$$

The MLE of parameters  $c_{k,i,j}$  at time  $t=T$  is obtained by maximizing the log-likelihood function:

$$-\ln(2\pi) - \frac{\ln|\Sigma_{t|t-1}|}{2} - u_t' \Sigma_{t|t-1}^{-1} u_t \quad (9)$$

Once the parameters have been estimated, the Generalized Portmanteau test (Hosking, 1980) is applied to ascertain that the model BEKK(1,1) has effectively eliminated the ARCH effects in the residuals of the CVAR model:

$$\tilde{Q}_m = n^2 \sum_{k=1}^m (n-k)^{-1} \hat{g}_k' (\hat{G}_0^{-1} \otimes \hat{G}_0^{-1}) \hat{g}_k \cong \chi_{m-p}^2 \quad (10)$$

where  $p$  is the CVAR order,  $n = T - t_{\max}$ ,  $t_{\max} = \max\{t_i, t_j\}$ ,  
 $\hat{g}_k = \text{vec}(\hat{G}_k)$ ,  $\hat{G}_k = L^T C_k L$ ,  $\hat{C}_k = n^{-1} \sum_{t=k+1}^n \hat{u}_t \hat{u}_{t-k}'$ ,  $LL' = \hat{C}_0^{-1}$ ,  
 $m = 1, \dots, 10$

*Step two:* The estimates obtained in phase one are used to select the shares for which positive average returns are predicted. For the shares thus selected and for each industrial sector (IS), the model CVAR( $p$ )-BEKK( $l, l$ ) is estimated for the random vector

$y_t = [y_{1,t}, y_{2,t}] = [(P/E)(EPS)_{h,t}, (P/E)(EPS)_{IS(h),t}]$ , where  
 $h = 1, \dots, H$  is the index that identifies only the series with positive returns selected out of the initial 150.

On the basis of the  $(P/E)(EPS)_{h,T+1}$  and  $R_{h,T+1}$  forecasts obtained in phase two, the shares are listed for each industrial sector in decreasing order with respect to the values of the difference between *intrinsic value* and *expected price*. The first  $n = 10$  shares are thus selected to make up the portfolio.

It should be stressed that the choice of  $n = 10$  is made on the basis of the assertion of Evans and Archer (1968) that this quantity is sufficient for diversification of portfolio choices. In actual fact, the number of shares selected will vary in further developments of this work.

In order to determine the quantities to be bought of each of the 10 shares selected, it is necessary to solve the Markowitz problem (Markowitz, 1952) by estimating the matrix of share volatility. To this end, let  $\hat{V}_t$  be the estimator of the matrix  $n \times n$  of volatility  $V_t$  for  $t = T + 1$ , the elements of which are  $v_{i,j}(t) = \text{cov}(R_t | \text{past})$ ,  $i, j = 1, \dots, n$ .

The elements of  $\hat{V}_t$  are given by:

$$v_{i,j} = \begin{cases} \hat{\sigma}_{ii,T+1|T}^{bekk} & \text{se } i = j \\ \hat{c}_{i,j} & \text{se } i \neq j \end{cases} \quad (11)$$

with  $\hat{C} = (\hat{c}_{i,j})_{i,j=1,\dots,n} = \sum_{t=t_{\max}}^T \frac{(R_{i,t} - \tilde{R}_i)(R_{j,t} - \tilde{R}_j)}{T - t_{\max}}$ ,  $t_{\max} = \max\{t_i, t_j\}$ ,

$\hat{\sigma}_{ii,T+1|T}^{bekk}$  and  $\hat{\sigma}_{11,T+1}^{(i)}$  in (8),  $i, j = 1, \dots, n, n = 10$ .

On the basis of (11), the solution of the following quadratic problem of Markovitz type, called *global minimum variance portfolio*, for the future time  $T + 1$  can be obtained with the approximation given by the dual method (Goldfarb, 1983, Higham, 2002):

$$\min_{\varpi \in \mathbf{R}} \{ \varpi' \hat{V}_{T+1} \varpi : \varpi' \mathbf{1} = 1, \varpi \geq 0 \} \quad (12)$$

To obtain a better diversification we have also to find the solution of the quadratic problem of Markovitz type (12) without the constraint  $\varpi \geq 0$ , for the future time  $T + 1$  using the explicit solution

$$\hat{\varpi}_{opt,T+1} = \hat{V}_{T+1}^{-1} \mathbf{1} / (\mathbf{1}' \hat{V}_{T+1}^{-1} \mathbf{1}) \quad (13)$$

Then we put to zero the  $\varpi_i < 0$  and reportionate the remaining  $\varpi_i$ ,  $i = 1, \dots, n$ .

We omit the constraint of a fixed value for the expected return to eliminate the sensitiveness of allocation optimization to errors in predicted returns (Hlouskova *et al.*, 2002).

In cases where the matrix  $\hat{V}_{T+1}$  proves positive neither in (12) nor in (13), we approximate it with the closest matrix in the sense of Frobenius possessing the same diagonal given by the elements estimated with the BEKK model.

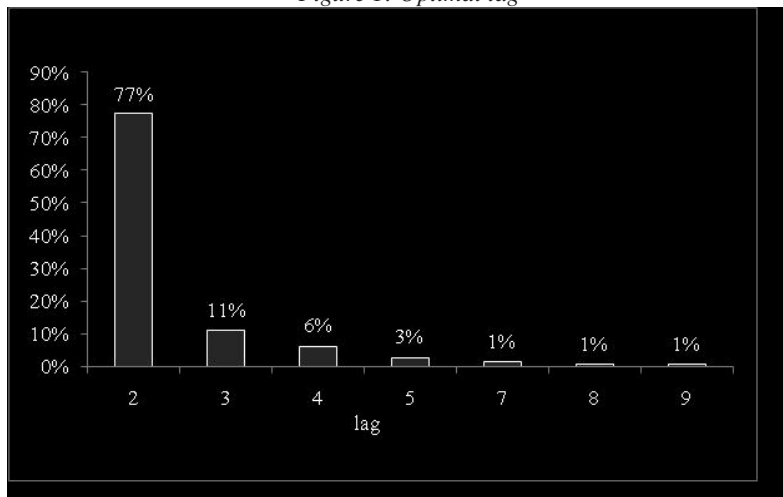
#### 4. Results

Application of the model proposed in this work to the monthly figures for the 150 BIT shares with the highest level of capitalization indicates the following results.

On the basis of the minimum AIC, the optimal lag  $p$  is 2–9 months. Figure 1 shows the empirical distribution of the lag. In particular, while the optimal lag is 2 months for 76% of the entire set of 150 shares it is 2 months for 90% of the shares in the portfolio and 1 month for the remaining 10%. This means that at least 2 months of observations are sufficient to predict the average returns on the vast majority of the shares considered, instead of a random walk model.

It was therefore decided to regard lag 9 as the maximum.

Figure 1: Optimal lag



The results of the LR test for all of the shares considered, yields a degree of cointegration that proves equal to 2 for 91% of the 150 shares, 1 for 7% and 0 for the remaining 2%.

This means that in 7% of cases, when the rank of matrix  $\Pi$  in (2) is equal to 1, use was made of the VEC model, which proves to be estimable (Johansen, 1995) and stationary, unlike the VAR model, which instead proves neither directly estimable nor stationary.

The coefficients of the model estimated in both steps of the procedure prove significant for almost all of the series considered. Table 1 shows, for example, the values of the coefficients estimated for models (1) and (8) for time series number 159.

Table 1: Parameters for models (1) and (8) – stock n°159 selected for the Portfolio.

stock 159					
	VAR		VAR		BEKK
$\mu_{0,1}$	0,014	$a_{5,1,1}$	-1,35	$c_{0,1,1}$	0,037
$\mu_{0,2}$	0,005	$a_{5,2,1}$	-0,604	$c_{0,2,1}$	0
$\mu_{1,1}$	0	$a_{5,1,2}$	0,286	$c_{0,1,2}$	0,023
$\mu_{1,2}$	0	$a_{5,2,2}$	-0,026	$c_{0,2,2}$	0,029
$a_{1,1,1}$	-1,353	$a_{6,1,1}$	-0,935	$c_{1,1,1}$	0,497
$a_{1,2,1}$	-0,472	$a_{6,2,1}$	0,138	$c_{1,2,1}$	-0,107
$a_{1,1,2}$	0,376	$a_{6,1,2}$	0,226	$c_{1,1,2}$	0,122
$a_{1,2,2}$	-0,213	$a_{6,2,2}$	-0,143	$c_{1,2,2}$	0,481
$a_{2,1,1}$	-1,368	$a_{7,1,1}$	-1,108	$c_{2,1,1}$	-0,015
$a_{2,2,1}$	0,023	$a_{7,2,1}$	-0,564	$c_{2,2,1}$	0,001
$a_{2,1,2}$	0,283	$a_{7,1,2}$	0,171	$c_{2,1,2}$	0,033
$a_{2,2,2}$	-0,517	$a_{7,2,2}$	-0,287	$c_{2,2,2}$	-0,001
$a_{3,1,1}$	-1,59	$a_{8,1,1}$	-0,902		
$a_{3,2,1}$	-0,836	$a_{8,2,1}$	-0,692		
$a_{3,1,2}$	0,549	$a_{8,1,2}$	0,391		
$a_{3,2,2}$	0,374	$a_{8,2,2}$	0,265		
$a_{4,1,1}$	-1,522	$a_{9,1,1}$	-0,37		
$a_{4,2,1}$	-0,516	$a_{9,2,1}$	-0,458		
$a_{4,1,2}$	0,203	$a_{9,1,2}$	0,062		
$a_{4,2,2}$	0,162	$a_{9,2,2}$	0,075		

The p value of F statistic is below 0.10 in 83% of cases and below 0.2 in 86%. The model thus proves 90% significant for nearly all of the series. Table 2 shows diagnostic tests for models (1) and (8). The results of the portmanteau Test on the residuals of the CVAR and BEKK models are given in Table 2 (here too, for brevity, they regard only the ten series selected for the portfolio). As regards the presence of ARCH effects in the residuals of the CVAR model (column 5 of Table 2: pAR), the test leads to the conclusion that the hypothesis of the presence of a heteroskedastic component in the model estimated at a confidence level of 99% cannot be ruled out for 60% of the shares chosen. The results of the Hosking test carried out in order to ascertain the presence of ARCH effects in the residuals of the BEKK model, which are shown in column 10 of Table 2 (pHq), lead to the conclusion that the hypothesis of the absence of any heteroskedastic component in the model is to be accepted in 90% of cases. In other words, it can be concluded that the combination of the CVAR and BEKK models picks up the heteroskedastic component and includes the information derived from the same in the procedure.

Column 9 (pH) of Table 2 shows the results of the portmanteau test carried out in order to ascertain the presence of autocorrelation of residuals in the BEKK model. They suggest that the hypothesis of autocorrelation can be ruled out and that the maximum lag considered is sufficient.

Table 2: Test statistics - stocks selected for the portfolio

Series	R2.1	pFvalue	pSC	pAR	pJB
4	0,114	0	0,187	0	0
34	0,085	0,006	0,717	0,475	0
63	0,078	0,08	0,785	0	0
130	0,014	0,894	0,347	0	0
49	0,052	0,007	0,664	0,002	0
109	0,041	0,736	0,502	0	0
159	0,725	0,02	0,049	0,458	0,847
85	0,179	0	0,2	0,056	0,001
54	0,12	0,01	0,838	0,131	0,025
60	0,113	0	0,423	0	0
Series	pSK	pKU	pH	pHq	intervalYN
4	0	0	0,142	0,556	true
34	0,127	0	0,233	0,606	true
63	0,002	0	0,822	0,004	true
130	0,054	0	0,303	0,566	true
49	0,015	0	0,752	0,344	false
109	0	0	0,887	0,266	true
159	0,905	0,553	0,176	0,864	true
85	0,161	0	0,459	0,103	true
54	0,415	0,009	0,924	0,793	true
60	0,108	0	0,877	0,042	true

$R2.1$  = goodness of fit stock equation1;

$pFvalue$  =  $P(F > F_{oss} | H_0: A_i = 0, i=1, \dots, p)$  zero coefficients of CVAR;

$pSC$  =  $P(\chi^2 > \chi^2_{oss} | H_0: E(u_t, u_{t-i}') = 0, i=1, \dots, h > p)$  autocorrelations in VAR residuals test;

$pAR$  =  $P(\chi^2 > \chi^2_{oss} | H_0: \text{no ARCH})$  ARCH in CVAR residuals test;

$pJB$  =  $P(\chi^2 > \chi^2_{oss} | H_0: \text{normality})$  CVAR normality test;

$pSK$  =  $P(\chi^2 > \chi^2_{oss} | H_0: E(u_t^3) = 0)$  CVAR skewness test;

$pKU$  =  $P(\chi^2 > \chi^2_{oss} | H_0: E(u_t^4) = 3)$  CVAR kurtosis test;

$pH$  =  $P(\chi^2 > \chi^2_{oss} | H_0: \text{no autocorrelations})$  autocorrelations in BEKK residuals test;

$pHq$  =  $P(\chi^2 > \chi^2_{oss} | H_0: \text{no ARCH})$  ARCH in BEKK residuals test;

A OLS-based CUSUM test for stability of the market index was also carried out, and the results suggest that the hypothesis of stability of the series over the period considered is acceptable.

The BEKK estimate of *volatility* for each share is between 0.001 and 0.01 for 93% of the series and never above 0.031.

It can be concluded in the light of this result that for most of the series considered, the estimated value of the share does not differ from its real value at a confidence level of 95% by more than  $\pm 0.2$ .

In actual fact, the value at risk calculation put forward by J.P.Morgan (Longerstaey *et al.*, 1995, Duffie *et al.*, 1997). could be used in order to include the information deriving from the presence of correlation between the series considered and hence to assess the overall risk rather than the risk of the individual share.

Further confirmation of the adequacy of the CVAR-BEKK model with respect to the series observed was sought before selecting the shares to be included in the portfolio. Specifically, the *confidence interval* at the level of significance of 95% contains the *actual* value  $\tau_{+1}$  in 94 % of the series. The CVAR-BEKK model can therefore be considered reliable for most of the series for the purposes of prediction.

The next step after verification of the suitability of the model was prediction of the prices of the shares as well as their intrinsic values. Table 3 shows the values predicted on the basis of model (7), once again restricted to the ten shares selected for the sake of brevity.

Table 3: Forecast return, volatility, intrinsic value, potential value  
– stocks included in portfolio

Series	pf	varf	vinf	indm	vin_p
4	0,008	0,082	3,622	10	3,614
63	0,001	0,044	0,732	8	0,731
130	0,011	0,063	0,611	8	0,6
49	0,009	0,047	0,515	4	0,506
109	0,012	0,044	0,229	3	0,217
159	0,019	0,039	0,193	6	0,174
85	0,031	0,09	0,204	3	0,173
54	0,007	0,046	0,109	6	0,102
60	0,006	0,079	0,05	10	0,044
34	0,009	0,081	0,037	2	0,028

*pf*= forecasted return ;

*varf*= forecasted volatility

*vinf*= forecasted intrinsic value;

*indm*= sector index;

*vin\_p*= forecasted intrinsic value – price.

On the basis of potentials, understood as the difference between share price and intrinsic value, the ten shares with the highest potential returns were then selected.

Two criteria of ranking were used, namely the partial criterion and the total criterion.

The former involves arranging the values of potential of all the shares considered in decreasing order for every industrial sector (Goodman and Peavy III, 1983) and selecting the first share in each.

In the latter, the values of potential of all the shares considered are arranged in decreasing order regardless of industrial sector and the first ten are chosen.

The use of the partial criterion is connected with the relationship between P/E and share performance manifested most strongly in each industrial sector. As Goodman and Peavy III write, “*firms in the same industry tend to cluster in the same relative P/E ranking, detected return*

*differences between P/E groups may be attributable to industry performances rather than P/E level. This bias is eliminated by using P/E relative to its industry.”*

The total criterion has the advantage that the selection is unconnected with the industrial sector of the share in question and therefore not necessarily influenced by the possibly negative trend of individual sectors. Its application thus means selection of the ten best shares in absolute terms (Nicholson, 1960).

The choice between these two criteria of ranking is obviously subjective in that it depends on the opinions of investors. This paper considers both, which evidently lead to the selection of different portfolios.

Table 4 lists the optimal allocations, i.e. the solutions of the problems of optimization (12) and (13); column 4 shows the results for the best portfolio. It has a *monthly average return* of 1.9%, a *monthly standard deviation* of 0.655, and a *Sharpe index* of 0.029.

*Table 4: portfolios weights, volatility, return, Sharpe*

Total Ranking			Partial Ranking		
Series	w(i) prop.	w(i) opt	w(i) opt	w(i) prop.	Series
4	0,462	0	0	0,299	142
63	0	0,581	0	0	34
130	0	0	0	0,225	109
49	0,003	0	0	0	49
109	0,216	0	1	0,344	159
159	0,204	0	0	0,009	65
85	0	0	0	0	63
54	0	0	0	0,123	147
60	0,004	0,419	0	0	4
34	0,112	0	0	0	28
Return	0,011	0,003	0,019	0,011	Return
St. Dev.	1,185	0,77	0,655	0,913	St. Dev.
Sharpe	0,009	0,004	0,029	0,012	Sharpe

## **5. Conclusions**

The selection of a share portfolio has historically constituted a complex problem that has no single solution but depends both on market conditions and on the information available to investors. In other words, the choice of shares to invest in must be based on objective criteria making it possible to assess risk and return without ignoring investors' opinions. To this end, the paper suggests the use of a model for the analysis of multiple historical series with a view to the prediction of share return and associated risk but also taking the indications of the market into account at the same time in the specification of the model itself. Variables obtained as functions of P/E and EPS have thus been used together with the market index as regressors of the combined model (1) and (8).

The innovative choices in the construction of a portfolio selection model put forward here regard two distinct aspects. The first concerns transition from a model of the VAR (1) type for the prediction of a multiple historical series (Campbell, 2003) to one of the CVAR (p) type, which makes it possible to take into consideration any cointegration of the series considered and therefore constitutes an improvement of the information available for estimation purposes. The subsequent use of a combination of CVAR and BEKK models, which extends the results of Bollerslev, Engle and Wooldrige (2), makes it possible to consider also the temporally variable correlation between the volatility of the series and the volatility of the market index within the estimation procedure.

The second concerns the choice of criterion for the selection of shares, which is addressed here by seeking to insert a typical concept of finance such as intrinsic value into the primarily statistical context of the prediction of a multiple historical series. An intrinsic value estimated by means of the combined CVAR-BEKK model is used to obtain a "potential value" serving as a basis to rank the different shares and then select the top ten.

The method put forward was applied to the series of 150 shares with highest capitalization quoted on the Italian stock exchange and led to the selection of 10 shares constituting a portfolio with an average monthly return of 1.9% and a risk of 0.655.

Comparison of the results of the CVAR-BEKK model put forward here and those obtained by means of models of the VAR (1) and DVEC type found in the literature was carried out on the basis of the values of log-likelihood of the models themselves. In other words, since one model could produce a higher value of return than another but nevertheless prove less reliable, it was decided to assess the models' performance in terms of correspondence to the series observed. The log-likelihood of the CVAR-BEKK model always proves greater than that of the other models, thus indicating more accurate representation of the series observed and hence better predictions.

Further developments of the work will regard the study of "value at risk", understood as assessment of the greatest loss possible, as well as identification of possible structural breaks of the individual series of share returns with a view to making the model more adaptable.

## References

Bollerslev T., Engle R.F., Nelson D.B. (1994), ARCH Model, *Handbook of Econometrics*, Elsevier Science, Amsterdam, vol. IV, pp. 2959–3038.

Bollerslev T., Engle R.F., Wooldridge J.M. (1988), A Capital Asset Pricing Model with Time-Varying Covariance, *Journal of Political Economy*, vol. 96, n. 1, pp. 116–131.

Brandt M.W., Goyal A., Santa-Clara P., Stroud H. R., (2005), A simulation approach to dynamic portfolio choice with an application to learning about return predictability, *The Review of Financial Studies*, vol.18 , n. 3, pp. 831–873 .

Brown K., Reily F. (2008), *Investment Analysis and Portfolio Management*, South-Western College Publishing.

Campbell, J.Y., Chan Y.L., Viceira L.M. (2003), A Multivariate Model of Strategic Asset Allocation, *Journal of Financial Economics*, vol. 67, n.1, pp. 41–80.

Datastream Global Equity Indices (2008), user guide, issue 5, Thomson Reuters Ltd.

Duffie D., Pan J. (1997), An overview of value at risk, *Journal of Derivatives*, pp. 7–48.

Engle R.F., Kroner, K.F. (1995), Multivariate Simultaneous Generalized ARCH, *Econometric Theory*, vol. 11, n.1, pp. 122–50.

Evans J. L., Archer S. H., (1968), Diversification and the Reduction of Dispersion: an Empirical Analysis, *The Journal of Finance*, vol. 23, pp. 761–767.

Goldfarb D., Idnani A. (1983), A numerically stable dual method for solving strictly convex quadratic programs, *Mathematical Programming*, vol. 27, pp. 1–33.

Goodman D.A., Peavy III J.W. (1983), Industry Relative Price-Earnings Ratios as Indicators of Investment Returns, *Financial Analysts Journal*, vol. 39, n. 4, 60–66 .

Hamilton J.D., (1994), *Time Series Analysis*, Princeton University Press.

Higham N., (2002), Computing the nearest correlation matrix - a problem from finance, *IMA Journal of Numerical Analysis*, vol. 22, pp. 329–343.

Hlouskova J., Schmidheiny K., Wagner M., (2002), Multistep Predictions from Multivariate ARMA-GARCH Models and their Value for Portfolio Management, *Diskussionsschriften*, Universität Bern.

Hosking J. R. M., (1980), The multivariate portmanteau statistic, *Journal of American Statistical Association*, vol. 75, pp. 602–608.

Johansen S., Juselius, K. (1990), Maximum likelihood estimation and inference on cointegration with applications to the demand for money, *Oxford Bulletin of Economics and Statistics*, vol. 52, pp. 169–210.

Johansen S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive models*, Oxford University Press, Oxford.

Longerstaey J., More L. (1995), *Introduction to RiskMetrics 4th ed.*, Morgan Guaranty Trust Company, New York.

Lutkepohl H., (2007), *New Introduction to Multiple Time Series Analysis*, Springer.

Markowitz H. (1952), Portfolio Selection, *Journal of Finance*, n.7, pp. 77–91.

Nicholson S. F., (1960), Price-earnings ratios, *Financial Analysts Journal*, vol. 16, n. 4, pp. 43-50.