
Risk forecasting models and optimal portfolio selection

David Moreno^{a,*}, Paulina Marco^b and Ignacio Olmeda^c

^a*Universidad Carlos III, Dpto. Economía de Empresa, (Madrid) Spain*

^b*Universidad de Valencia, Dpto. de Finanzas Empresariales, (Valencia) Spain*

^c*Universidad de Alcalá, Dpto. de Ciencias de la Computación, (Madrid) Spain*

This study analyses, from an investor's perspective, the performance of several risk forecasting models in obtaining optimal portfolios. The plausibility of the homoscedastic hypothesis implied in the classical Markowitz model is discussed and more general models which take into account asymmetry and time varying risk are analysed. Specifically, it studies whether ARCH-type based models obtain portfolios whose risk-adjusted returns exceed those of the classical Markowitz model. The same analysis is performed with models based on the Lower Partial Moment (LPM) which take into account the asymmetry in the distribution of returns. The results suggest that none of the models achieve a clearly superior average performance. It is also found that models based on semivariance perform as well as those based on the variance, but not better than, even if the evaluation criterion is based on the Reward-to-Semivariance ratio. When attention turns to the analysis of worst case performance, the results are clearly different. Models which employ LPM with a high degree of risk aversion ($n > 2$) as the risk measure are consistently superior to those which employ a symmetric measure, either homoscedastic or heteroscedastic.

I. Introduction

Since the beginning of Modern Finance Theory there has been a constant controversy about the concept of risk, and an increasing interest in ways to measure it. This controversy has been accompanied by a growing investment industry in portfolio models based on sophisticated quantitative methods which require a huge computing power. One should not be surprised by this fact since financial markets are now much more volatile and the use of derivative instruments,

such as options, to hedge risks require 'hi-tech' developments in finance. In particular, the development of new methods of portfolio management is now a paramount issue in the financial community.¹

The roots of the Modern Portfolio Theory (MPT in what follows) can be traced back to Markowitz's (1952) seminal idea that investors should hold mean-variance efficient portfolios. This idea rests on mean-variance investors whose utility depends only on these two factors. The appeal of the mean-variance model is evident. It only requires

*Corresponding author. E mail: jdmoreno@emp.uc3m.es

¹ It is also obvious that particular investors demand more sophisticated financial products as well as risk control systems since the crash of 1987. Concepts such as 'implied volatility' and 'Value at Risk' are now terms commonly used by even the most modest investor.

one to know the first two moments of the probability's distribution in the considered asset's returns to create optimal portfolios which are the solution of a well-defined and easily solvable quadratic programming problem. Obviously, this approach would be correct only if at least one of the underlying assumptions is correct, that is: (1) the utility function of the investor is quadratic and (2) the distribution of returns is conditionally normal.

The first of the assumptions of Markowitz model does not conform to the observed behaviour of investors: it requires that investors increase their absolute risk aversion as their wealth increase, while it is well known that they generally exhibit a higher risk aversion with lower levels of wealth. The second assumption implies that the unconditional mean and variance fully describe the distribution of the asset's returns. Nevertheless, a large number of empirical works document the inappropriateness of the normality hypothesis (see e.g. Taylor, 1995 for a review). In particular, and among other properties, financial returns are skewed and leptokurtic (Mandelbrot, 1963).

Despite of these well known facts, several authors² continue to favour the mean–variance model, not only because of its simplicity but also because it allows one to obtain levels of utility highly correlated with the expected satisfaction of an investor. At this point it seems that one important question is: is it possible to find some alternative measures of risk that outperform the generally employed estimation of unconditional variance?

Stochastic models whose variance evolves with time may provide an explanation of the stylized facts of financial returns. Since Mandelbrot (1963), it is recognized by the financial community that volatility 'clusters' with time. Engle (1982) was probably the first to propose a model of changing volatility. Since the inception of his Autoregressive Conditional Heteroscedastic Model (ARCH, hereafter), an impressive number of papers have employed it, and many extensions and practical applications have been proposed (see Bollerslev *et al.*, 1992, for a review). This study analyses:

- (1) The effect of the modification of the classical Markowitz scheme by introducing heteroscedastic conditional variances in the determination of the optimal portfolios.

- (2) Another possible modification that comes from an interpretation of risk aversion more coherent to rationality: investors should employ risk measures which allow them to reject decisions which would produce rates of return *lower* than a specific target, and accept decisions which would produce rates of return *higher* than the target.³ In this line of reasoning, downside risk measures are employed (see Nawrocki, 1999, for a review of all these measures) which have that property. In fact, even Markowitz (1991) recognizes that from a theoretical, practical, and empirical perspective, these measures, in particular semivariance, are more plausible.

In both modifications, the benefits of portfolio optimization critically depend on how accurately the involved measures can be forecasted. Here the focus will be on the forecasts of risk, rather than of expected returns for two reasons: first, many studies have analysed the forecasts of returns in the context of mean–variance optimization.⁴ The general opinion is that expected returns are very difficult to predict, and that the optimization process is extremely sensitive to these differences. Second, there exists a general impression that volatility, in a wide sense, is easier to estimate than expected returns from historical data.⁵ Moreover, this kind of analyses, to our knowledge, is much scarcer in the framework of portfolio optimization.

The main contributions of our work are the following: First, the approach adopted by other authors (e.g. Nawrocki, 1991) is extended by explicitly considering forecasted instead of realized measures of risk, that is, in the solution of Markowitz's problem the variances and semideviations obtained from ARCH-type or autoregressive models are employed instead of the historical ones. In this sense, the work is closer to the real behaviour of investors who extrapolate, in some pertinent way, the expected optimal portfolio. Second, the performance of dynamic forecasts of symmetric against asymmetric measures of risk is compared. To the present authors' knowledge, none of the studies in the literature have done this type of comparative analysis. Third, a database of country indices is employed whose behaviour is presumably more diverse than if assets of any particular market had been employed, as other authors do. The argument is that country indices may have,

² Levy and Markowitz (1979), Pulley (1981, 1985), Kroll *et al.* (1984), among others.

³ Note that the variance, employed as a risk measures does not have this property, it is symmetric and, consequently, penalizes both gains and losses.

⁴ See Michaud (1989) and Best and Grauer (1991) among others, for a review.

⁵ See, e.g. Merton (1980) and Nelson (1992).

compared to individual stocks, a more predictable structure due to aggregation.

The remainder of the paper is organized as follows. The next section introduces the models employed to measure and forecast risk. Section III describes the dataset as well as the procedure employed in the estimation, prediction and evaluation phases. The procedure is quite computationally intensive. Section IV presents the results obtained with the procedure, as will be seen, no significant differences are found among the models when average performance is evaluated. Finally, it will be seen that when market conditions are adverse, there is a clear superiority of asymmetric models over the symmetric (heteroscedastic or not) ones. The last section summarizes the main results obtained.

II. Risk Models

Time varying models

Engle (1982) was probably the first who provided a simple model of changing volatility, the ARCH model. Though the catalogue of ARCH-type models is huge, probably the most widely used models are Engle's ARCH, and the generalized ARCH process (GARCH) of Bollerslev (1986).

ARCH-type models take into account the excess of kurtosis and volatility clustering. Engle's ARCH assumes that returns are generated by stochastic processes serially correlated, or not, with non-constant variances conditional on the past and constant unconditional variances. For such processes, the recent past gives information about the one-period forecasted variance. The conditional variance of the innovations of an ARCH model of order p or ARCH (p) is parameterized as:

$$\sigma_t^2 = \omega + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 \quad (1)$$

where σ_t^2 is the conditional variance (in this case we assume p lags of the squared residuals), ω and $\alpha_j > 0$, to guarantee a positive variance. The parameters ω , α_j and p are usually estimated by maximum likelihood.

Though the ARCH model allows a parsimonious representation of heteroscedasticity, for practical applications, especially when the volatility is persistent, the usual formulation as an ARCH model

would require a relatively high number of parameters. A more flexible structure is provided by a GARCH model, first suggested by Bollerslev (1986). GARCH stands for Generalized Autoregressive Conditional Heteroscedasticity. GARCH models employ past variances and squared innovations to predict future variances. The conditional variances of the GARCH (p, q) process have the following expression:

$$\sigma_t^2 = \omega + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{k=1}^q \beta_k \sigma_{t-k}^2 \quad (2)$$

where ω , α_j and $\beta_k > 0$, to guarantee a positive variance, and are also estimated by maximum likelihood.

Asymmetric measures of risk

As has been noted earlier, the use of variance as a measure of risk may be inappropriate, since financial returns are asymmetric. Alternatively, one may employ risk measures which penalize asymmetrically both sides of the distribution of returns. By 'downside risk' one refers to a risk measure which is computed over a part of the distribution function, taking into account only returns which are below a particular target.⁶

The first, more obvious measure is the semi-deviation, and a more general measure of downside risk is the Lower Partial Moment (LPM, hereafter) which has been employed by numerous authors (e.g. see Nawrocki, 1999). The LPM measure is defined as follows:

$$LPM(n, i) = \frac{1}{T} \sum_{t=1}^T [\text{Maximum}\{0, (\Phi - R_{it})\}]^n \quad (3)$$

where n is the *degree* of the LPM, T is the total number of observations, Φ is the bound imposed by the investor and R_{it} is the return of asset i in period t . This measure is quite flexible since it is consistent with many utility functions depending on the value of n . Levy and Markowitz (1979) and Kroll *et al.* (1984) showed that some utility functions can be approximated by utility functions which involve only the calculation of mean-variance, but the n -degree LPM defines a much more general class of utility functions. As n increases, one assumes a higher risk aversion by the investor.⁷ This study computes LPM with $n=2, 3, 4$ and 5 to reflect several degrees of risk aversion.

⁶ This view was first proposed by Roy (1952) who shows that investors try to obtain returns over a minimum or catastrophic risk, this is the principle of *Safety First*.

⁷ For n smaller than one 'risk lovers' is considered, $n = 1$ represents 'risk neutrals' and for n higher than one it is considered that investors are 'risk averse' (see Fishburn, 1977). Note that semivariance is equal to LPM with n equal to 2.

Note that since LPM measures below target returns, it can be interpreted as an approximate measure of skewness. Since investors prefer positive skewness and dislike negative skewness, LPM becomes a measure of risk. The higher the LPM value, the greater the degree of negative skewness, and the greater the risk of the investment. As n increases, LPM is still measuring negative skewness. It is simply providing a heavier utility penalty to negative skewness as the degree (n) is increased.

Similarly to the computation of covariances in the classical framework, we can define co-lower partial moment (CLPM, Bawa, 1975; Fishburn, 1977) as⁸:

$$CLPM_{ij,n} = \frac{1}{T} \sum_{t=1}^T \left([\text{Max}\{0, (\phi - R_{it})\}]^{n-1} (\phi - R_{jt}) \right) \quad (4)$$

III. Database and Methodology

Database

The database consists of 35 MSCI (Morgan Stanley Capital International) indices,⁹ expressed in US dollars, covering the period from 31 December 1989 to 1 January 2002 (687 weekly observations). Thirteen emerging markets and 22 developed ones are considered.¹⁰ Returns are calculated as the first difference of log prices.

The main statistics of the raw data are shown in Table 1. As it is common in Finance literature, evidence of leptokurtosis and asymmetry are documented in the returns distribution, which allow the rejection of normality using the Jarque–Bera test. Note that the existence of asymmetric distributions justifies the use of downside risk measures employed in the following section. Strong evidence of ARCH effects are also found by applying Engle’s test for four lags.

Methodology

To deal with heteroscedasticity, for each one of the markets, the study estimated, by maximum likelihood, sixteen different specifications of the AR(z)–ARCH(p) and AR(z)–GARCH(p, q) process ($z = 0, 1, 2, 3$; $p = 1, 2$; $q = 0, 1$). From this range

of models, the best one is selected according to the Akaike Information Criterion (AIC). Then, this model is used to forecast the conditional variance, which is then incorporated into Markowitz’s model to determine the optimal portfolio, that is, the tangency between the Capital Allocation Line (CAL) and the Efficient Frontier. Finally, the performance of these portfolios is compared against that of the portfolio derived from the classical Markowitz framework. To obtain the optimal portfolios constraints are introduced to ensure non-negativity of the weights, this is done because short selling is expensive for individual investors and it is not generally permissible for most institutional investors.

The covariances can be assumed constant or not. They are considered to be constant in time, following Jobson and Korkie (1980) and Jorion (1986) among others. Other authors (e.g. Elton and Gruber, 1973; Eun and Resnik, 1992 or Ledoit and Wolf, 2003) assume that they evolve over time. In this case, more sophisticated models have to be employed, but these extensions are left for future work.

Changing asymmetry is modelled in the following manner: to forecast semivariance a classical AR(p) model is run, with $p = 0, 1, \dots, 5$ on the past values of semivariance. Then, the specification model is chosen which minimizes the AIC, and one proceeds exactly as in the ARCH process, previously described.

A time span divided into two parts is considered: the first one, the in-sample period, is employed for fitting the econometric models. The estimated parameters from the optimal models (in terms of the minimum AIC) are then used to forecast the risk for the next period. These forecasts of risk are then used to compute the optimal portfolio by solving the following quadratic programming formulation:

$$\begin{aligned} & \text{minimize} \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \\ & \text{subject to} \sum_{j=1}^n r_j x_j \geq \rho \\ & \sum_{j=1}^n x_j = 1 \\ & 0 \leq x_j \leq 1, j = 1, 2, \dots, n \end{aligned}$$

⁸ Note, again, that from this expression one can derive the cosemivariance, a particular case of LPM when the degree is equal to two.

⁹ Since this database is well known to the financial researchers a full description is not provided here.

¹⁰ The emerging markets analysed are Indonesia (INO), Korea (KOR), Malaysia (MAL), Philippines (PHI), Taiwan (TAW), Thailand (THA), Argentina (ARG), Brazil (BRA), Chile (CHE), Mexico (MEX), Greece (GRE), Jordan (JOR), and Turkey (TUR). The developed markets are Australia (AUL), Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DEN), Finland (FIN), France (FRA), Germany (GER), Hong Kong (HON), Ireland (IRE), Italy (ITA), Japan (JAP), Luxembourg (LUX), Netherlands (NET), Norway (NOR), Portugal (POR), Singapore (SIP), Spain (SPA), Sweden (SWE), Switzerland (SWI), United Kingdom (UKG) and United States (USA).

Table 1. Descriptive statistics of the weekly series

	Mean	Max	Min	Std. Dev	Skewness	Kurtosis	Jarque Bera	Engle Test
INO	0.248	0.295	0.411	0.065	0.467	11.340	0.000	0.934
KOR	0.091	0.204	0.403	0.053	0.536	9.392	0.000	0.739
MAL	0.038	0.297	0.371	0.049	0.677	14.634	0.000	0.943
PHI	0.039	0.247	0.207	0.046	0.008	6.416	0.000	0.999
TAW	0.011	0.231	0.222	0.051	0.183	5.125	0.000	0.990
THA	0.127	0.248	0.185	0.055	0.085	4.820	0.000	0.983
ARG	0.233	0.805	0.363	0.081	1.253	19.588	0.000	0.993
BRA	0.166	0.208	0.507	0.074	1.157	9.274	0.000	0.989
CHE	0.222	0.140	0.161	0.034	0.104	5.070	0.000	0.992
MEX	0.324	0.180	0.207	0.047	0.316	5.572	0.000	0.992
GRE	0.181	0.185	0.177	0.047	0.064	4.005	0.000	0.993
JOR	0.013	0.085	0.311	0.023	3.247	51.842	0.000	0.990
TUR	0.205	0.290	0.323	0.081	0.164	4.490	0.000	0.992
AUL	0.049	0.072	0.092	0.024	0.213	3.442	0.000	0.986
AUT	0.037	0.165	0.133	0.032	0.130	5.577	0.000	0.994
BEL	0.085	0.075	0.084	0.024	0.325	3.910	0.000	0.988
CAN	0.097	0.098	0.118	0.023	0.358	4.915	0.000	0.997
DEN	0.150	0.079	0.120	0.026	0.263	3.968	0.000	0.990
FIN	0.233	0.233	0.234	0.048	0.357	6.118	0.000	0.994
FRA	0.149	0.091	0.111	0.027	0.304	3.916	0.000	0.992
GER	0.129	0.090	0.131	0.029	0.428	4.389	0.000	0.992
HON	0.156	0.135	0.226	0.038	0.878	6.274	0.000	0.991
IRE	0.128	0.097	0.115	0.029	0.091	3.782	0.000	0.994
ITA	0.049	0.108	0.189	0.034	0.286	4.798	0.000	0.985
JAP	0.128	0.146	0.092	0.032	0.319	4.205	0.000	0.988
LUX	0.083	0.308	0.338	0.040	0.622	18.544	0.000	0.993
NET	0.176	0.067	0.129	0.023	0.481	5.373	0.000	0.987
NOR	0.059	0.137	0.099	0.031	0.025	3.757	0.000	0.995
POR	0.047	0.094	0.104	0.028	0.070	3.914	0.000	0.990
SIP	0.069	0.124	0.140	0.032	0.320	5.369	0.000	0.987
SPA	0.101	0.093	0.137	0.031	0.381	4.439	0.000	0.988
SWE	0.173	0.154	0.190	0.036	0.294	4.707	0.000	0.993
SWI	0.209	0.073	0.098	0.024	0.294	3.685	0.000	0.993
UKG	0.124	0.116	0.082	0.023	0.013	4.056	0.000	0.990
USA	0.213	0.078	0.091	0.021	0.431	4.718	0.000	0.990

Notes: The table shows the descriptive statistics for each of the countries, they are the mean, maximum (max), minimum (min), standard deviation (Std. dev.), skewness and kurtosis. The p value of the Jarque Bera statistic for testing the null hypothesis that the return series are normally distributed is also shown. The last column is the 1 less p value for the Engle's autoregressive conditional heteroscedasticity test. The mean is shown in percentage terms.

where r_j is the expected return, $r_j = E[R_j]$, of security j , σ_{ij} is the expected covariance between returns of security i and of security j , $\sigma_{ij} = E[(R_i - r_i)(R_j - r_j)]$, x_j is the proportion invested in asset S_j . ρ is a parameter representing the minimal rate of return required by an investor.

In cases where the measure of risk employed is the LPM a problem occurs: the CLPM matrix is not symmetric and Markowitz's model requires symmetry, since it is solved through quadratic programming optimization. Following Nawrocki (1992), a symmetric matrix is built which is then employed in the formulation of

Markowitz's model. The alternative formulation of the problem is:

$$\begin{aligned} &\text{minimize } LPM_{n,p} = \sum_{i=1}^n \sum_{j=1}^n x_i x_j LPM_i LPM_j r_{ij} \\ &\text{subject to } \sum_{j=1}^n r_j x_j \geq \rho \\ &\sum_{j=1}^n x_j = 1 \\ &0 \leq x_j \leq 1, j = 1, 2, \dots, n \end{aligned}$$

where r_{ij} is the correlation coefficient between asset S_i and asset S_j .

It must be noted that, in both formulations, the specification of the conditional mean of each of the models is only employed to obtain the residuals, which are used in the specification of the conditional variance, and not to forecast future returns (which are estimated by the historical average). This procedure is used to avoid misspecification in the conditional mean of some of the models, which could lead to suboptimal portfolios. Since the interest lies only in the alternative ways to forecast risk, it is preferred to employ simpler (less accurate) models and to establish the comparisons only in terms of alternative formulations of the risk.

To obtain a number of simulations the above procedure is repeated using non-overlapping and out-of-sample rolling windows so that, say, at iteration 1 the models are built using observations 1–50 and evaluated on observations 51–75, at iteration two models are built using observations 26–75 and evaluated on observations 76–100 and so on. This procedure allows one to obtain 25 different non-overlapping test sets. The table below summarizes the models considered in the work.

	Expected return	Risk measure	Covariance or CoLPM
Model 1	Historical	Historical (variance)	Historical (covariance)
Model 2	Historical	GARCH	Historical (covariance)
Model 3	Historical	Historical (semivariance)	Historical (co LPM)
Model 4	Historical	AR semivariance	Historical (co LPM)
Model 5	Historical	Historical LPM ($n = 3$)	Historical (co LPM)
Model 6	Historical	Historical LPM ($n = 4$)	Historical (co LPM)
Model 7	Historical	Historical LPM ($n = 5$)	Historical (co LPM)

To complete the analysis the worst case is studied, that is, the most pessimistic situation for the investor. From a risk-averse investor's perspective, an unresolved question is whether some of the models perform worse than others in adverse situations, that

is, if it is possible to choose among models so that the performance of one of them is always better than the worst case performance.

In trying to answer the above question the following procedure is proposed. At each time step, from iteration 1 to 25, the performance of each one of the models is registered. Then, an ideal ex-post portfolio from the Markowitz model is considered, which allows one to measure how the stock markets really moved, and take the first five worst results. That is, one considers 20% of the times that the market performed worst. Afterwards, one checks what the performance of the other models was and this set is ordered for the previously detected worst cases. To the extent that a particular model behaves better than the others in these extreme conditions, one can say that it should be, *ceteris paribus*, chosen by risk adverse investors.

IV. Empirical Results

Table 2 shows the annualized returns generated by the seven models in each of the evaluation periods. The results indicate that, on average, the best portfolios are those which employ the LPM and, significantly, those with a degree higher than 2. It seems that the models which incorporate dynamic forecasts of the risk measure (models 2 and 4) produce quite similar results (actually slightly worse) to those which employ the historic measure. On average, the models which employ the LPM ($n > 2$) obtain an annualized return 2% greater than the one obtained with the other models. Note that the results are equally robust along all the testing periods, since all the variances are quite similar.

In terms of annualized risk, all the models also seem to be very similar (see Table 3). Interestingly, the classical Markowitz model achieves the lowest risk level (it is also the second more robust). The worst results correspond to Model 4 which employs forecasts of the semivariance. The less robust models are the ones which employ the LPM ($n > 2$).

Now the possible trade-off between the performance of the models in terms of risk and return will be analysed. The first measure employed is the Sharpe Ratio (Sharpe, 1966).¹¹ The best model seems to be

¹¹ The Sharpe Ratio can be defined as follows:

$$\text{Sharpe Ratio} = \frac{R_i - r_f}{\sigma_i}$$

where R_i is the return of portfolio i and r_f is the risk free rate (3 month Treasury bills are employed) and σ_i is the standard deviation of portfolio i .

Table 2. Mean annualized return in each testing period

Period	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1	21.55	32.14	27.98	27.74	30.17	29.38	22.76
2	30.94	26.53	34.18	31.33	27.00	17.41	13.74
3	70.12	70.49	70.41	70.32	81.37	88.30	94.46
4	60.33	60.85	62.29	63.23	58.56	56.78	55.20
5	66.43	62.97	66.76	65.17	57.44	47.53	38.71
6	23.43	23.62	35.53	39.15	19.49	6.72	3.73
7	25.24	21.58	27.33	26.29	28.67	29.26	28.94
8	31.86	31.63	25.34	24.85	17.38	12.65	10.66
9	13.01	12.45	10.83	11.34	17.91	18.44	16.94
10	28.76	17.49	29.51	27.19	18.97	14.63	14.40
11	19.39	20.82	16.73	15.24	13.47	11.82	8.87
12	13.78	15.23	16.54	13.28	23.04	23.72	22.87
13	24.07	24.11	23.99	24.13	24.03	24.14	24.29
14	1.77	0.94	1.18	0.80	1.67	2.33	2.96
15	34.59	34.43	34.64	34.95	34.16	33.68	33.08
16	21.64	20.52	22.19	21.23	23.09	23.67	24.17
17	10.56	12.70	13.43	6.68	11.57	11.05	10.25
18	28.18	27.52	27.51	27.29	24.16	21.47	19.27
19	7.08	7.49	7.55	8.41	8.01	8.19	8.85
20	3.86	2.89	7.02	7.00	10.74	25.46	33.08
21	17.09	16.30	13.12	11.27	9.43	5.87	2.48
22	1.46	1.42	21.74	20.28	41.04	54.19	63.29
23	56.17	56.57	57.88	57.12	48.21	42.47	37.58
24	21.32	21.32	21.32	21.32	21.32	21.32	21.32
25	37.05	37.05	29.58	30.30	37.05	37.05	37.05
Mean	14.38	14.22	14.93	14.30	16.30	16.69	16.79
Std. dev.	29.92	29.53	30.62	30.47	29.12	28.86	29.13

Notes: The table shows the annualized returns generated by the seven models in each of the evaluation periods, and the mean and the standard deviation (Std. dev.) of these returns for each model. To obtain 25 different non overlapping test sets the next procedure is repeated using non overlapping out of sample rolling windows: at iteration 1 the models are built using observations 1 50 and evaluated on observations 51 75, at iteration 2 models are built using observations 26 75 and evaluated on observations 76 100 and so on.

model 5 (see Table 4), but its ratio is almost equal to the simplest model (the classical Markowitz model), however, model 5 seems to provide more robust results along the testing periods. The worst is, again, model 4. It is also evident that the dynamic versions (models 2 and 4) are worse than the static ones (models 1 and 3). It can also be seen that models which employ a symmetric risk measure (models 1 and 2) provide better results than the ones which employ asymmetric measures (models 3 and 4).

It can be argued, though, that Sharpe's ratio is consistent with variance but not with semivariance. For this reason¹² we also compute the Reward-to-Semivariability (R/SV) Ratio (Roy, 1952).¹³

In the case of the R/SV (see Table 5), it is found that the differences among the models are also not relevant. Unexpectedly, the models which employ the semideviation do not perform better. Again, the worst model is number four and the best the classical Markowitz model.

¹² Another argument to employ this alternative performance measure is that the Sharpe ratio is a biased estimator (see e.g. Ang and Chua, 1979).

¹³ The Reward to Semivariability can be defined as:

$$\frac{R}{SV} = \frac{R_i - R_f}{\text{Semideviation}(R_i)}$$

Table 3. Mean annualized risk in each testing period

Period	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1	10.99	14.61	13.72	13.53	16.23	18.35	18.01
2	25.57	22.15	25.03	23.99	24.75	24.73	25.11
3	19.94	19.86	19.78	19.84	19.14	19.06	19.72
4	20.67	20.61	22.24	22.56	23.01	24.59	26.12
5	16.85	17.06	15.97	16.92	17.12	17.29	16.92
6	14.83	14.94	17.82	18.35	15.61	13.22	10.85
7	8.05	7.30	8.27	9.06	8.96	9.49	9.71
8	10.38	9.78	14.59	14.66	18.83	21.31	22.21
9	15.66	16.05	16.30	16.36	15.67	14.85	15.28
10	10.49	9.01	10.88	10.43	9.18	9.86	10.91
11	10.81	11.71	10.23	11.17	9.14	8.85	8.53
12	8.99	10.79	8.70	10.81	7.47	7.04	6.91
13	8.14	8.09	8.57	8.71	8.53	8.52	8.55
14	9.92	9.92	10.00	9.96	10.27	10.45	10.55
15	7.69	7.65	7.93	8.13	8.20	8.24	8.40
16	12.15	12.19	14.73	15.82	14.75	14.43	14.14
17	19.92	20.03	18.50	20.43	14.89	12.68	10.98
18	22.61	22.50	23.08	23.68	21.01	19.14	17.47
19	25.24	25.72	24.35	24.47	24.09	24.12	24.18
20	23.39	24.37	25.41	24.82	22.54	20.14	18.54
21	20.20	21.65	23.87	25.23	17.61	15.12	15.32
22	24.71	25.20	32.25	30.90	38.44	43.59	43.68
23	20.10	20.31	20.57	19.66	20.62	21.32	21.11
24	19.12	19.12	19.12	19.12	19.12	19.12	19.12
25	14.78	14.78	12.85	13.00	14.78	14.78	14.78
Mean	16.05	16.22	16.99	17.26	16.80	16.81	16.68
Std. dev.	5.99	5.91	6.56	6.27	6.98	7.78	7.94

Notes: The table shows the annualized risk of the seven models in each one of the evaluation periods, and the mean and the standard deviation (Std. dev.) for the entire periods. To obtain 25 different non overlapping test sets we repeat the next procedure using non overlapping out of sample rolling windows: at iteration 1 the models are built using observations 1 50 and evaluated on observations 51 75, at iteration 2 models are built using observations 26 75 and evaluated on observations 76 100 and so on.

Table 6 shows the results obtained with the Wilcoxon test for equal distribution for the values of the previous tables. As can be seen, no statistically significant differences are found among the models.

Tables 7 and 8 provide some insight about the composition of the optimal portfolio for each one of the models. The results of Table 7 show that, on average, the number of assets considered by any model is virtually the same (six or seven different assets). Furthermore, it can be noted that, in general, there is a clear relation among the models: the number of assets increases (decreases) in all the models during the same periods of time. In Table 8, the number of assets which have a weight in the optimal portfolio greater than 3% are considered (which would be, approximately, the percentage in an equally weighed portfolio). This analysis is

performed because the investor could consider not holding an asset with a small weight in order to avoid transaction costs. The difference between emergent and developed stock markets is also shown. The results agree with others studies (e.g. Nawrocki, 1991) that show that the number of assets considered by a LPM approach is lower than from a variance perspective. Moreover, the results also show that the number of assets decreases as the degree of LPM increases. This is due to the fact that this measure of risk takes into account the skewness of the returns distribution and the diversification is achieved more efficiently with a smaller number of assets (see Nawrocki, 1992).

Table 9 shows the correlation coefficients of the returns, generated by each one of the

Table 4. Mean Sharpe ratio in each testing period

Period	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1	1.26	1.67	1.47	1.48	1.38	1.18	0.83
2	1.50	1.53	1.66	1.61	1.39	1.00	0.84
3	3.22	3.25	3.26	3.25	3.94	4.32	4.49
4	2.67	2.70	2.57	2.57	2.32	2.10	1.91
5	3.71	3.46	3.94	3.62	3.13	2.52	2.06
6	1.80	1.80	2.17	2.31	1.45	0.75	0.05
7	2.76	2.54	2.94	2.57	2.86	2.77	2.67
8	2.78	2.92	1.53	1.49	0.76	0.45	0.34
9	0.62	0.57	0.47	0.50	0.94	1.02	0.90
10	2.33	1.46	2.31	2.19	1.59	1.04	0.92
11	2.31	2.25	2.18	1.86	2.08	1.96	1.69
12	0.92	0.90	1.27	0.72	2.35	2.59	2.51
13	2.33	2.35	2.20	2.18	2.22	2.23	2.24
14	0.33	0.42	0.39	0.43	0.33	0.26	0.20
15	3.85	3.85	3.73	3.68	3.56	3.48	3.34
16	1.36	1.27	1.16	1.02	1.22	1.29	1.35
17	0.28	0.38	0.45	0.08	0.44	0.47	0.47
18	1.03	1.00	0.98	0.94	0.91	0.86	0.82
19	0.11	0.12	0.13	0.16	0.15	0.16	0.18
20	0.36	0.30	0.10	0.10	0.67	1.49	2.02
21	0.60	0.52	0.34	0.25	0.25	0.06	0.16
22	0.17	0.17	0.50	0.47	0.92	1.11	1.32
23	3.09	3.08	3.11	3.21	2.63	2.27	2.06
24	1.34	1.34	1.34	1.34	1.34	1.34	1.34
25	2.33	2.33	2.10	2.13	2.33	2.33	2.33
Mean	0.85	0.82	0.82	0.75	0.86	0.84	0.82
Std. dev.	1.92	1.88	1.89	1.84	1.78	1.72	1.67

Notes: The table shows the possible trade off between the performance of the models in terms of risk and return (through the Sharpe's ratio) in each of the evaluation periods, and the mean and the standard deviation (Std. dev.) of those ratios for each model. To obtain 25 different non overlapping test sets the next procedure is repeated using non overlapping out of sample rolling windows: at iteration 1 the models are built using observations 1 50 and evaluated on observations 51 75, at iteration 2 models are built using observations 26 75 and evaluated on observations 76 100 and so on.

models employed. It is found that the correlation between the returns from all models is very high, confirming the previous conclusions that models provide similar results. An inspection of the correlations reveals that one can identify two groups: the one which employs a risk measure which assumes a lower (symmetric or not and dynamic or not) degree of aversion of the decisor (models 1 to 4) and the other which penalizes risk more heavily (models 5, 6 and 7).

In Table 10 one finds interesting differences among the models in terms of the empirical distribution of the returns of the optimal portfolios. The models which employ LPM show less negative, or even, positive skewness. Also, for these models the

kurtosis is more than that of models which employ the variance as the measure of risk; again the kurtosis coefficient increases with the degree of the LPM. The results are in line with those of Nawrocki (1992); note, however that this author performs an ex-ante analysis. From these and the above results it is seen that models which employ the LPM ($n > 2$) as the risk measure constitute a qualitatively different class.

Worst case analysis

Finally, the Worst case analysis has been run. Figs 1 and 2 plot the results obtained in terms of annualized returns and risk, respectively. As can

Table 5. Mean R/SV ratio in each testing period

Period	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1	3.82	5.41	3.59	3.60	2.49	1.94	1.33
2	1.77	1.68	1.96	1.89	1.64	1.20	1.02
3	6.44	6.54	6.60	6.54	8.89	10.66	12.10
4	5.48	5.53	5.39	5.43	4.86	4.36	3.92
5	8.65	7.41	9.79	7.99	6.16	4.51	3.49
6	2.19	2.19	2.61	2.73	1.89	1.06	0.08
7	6.92	5.57	8.26	6.07	7.62	7.01	6.49
8	6.16	6.96	2.86	2.77	1.32	0.76	0.57
9	1.00	0.94	0.73	0.78	1.53	1.61	1.38
10	4.88	2.90	4.60	4.41	2.90	1.72	1.48
11	2.99	2.90	2.86	2.46	2.70	2.56	2.37
12	1.51	1.52	2.22	1.20	5.18	6.08	5.78
13	3.71	3.72	3.54	3.54	3.56	3.60	3.64
14	0.48	0.60	0.56	0.62	0.47	0.37	0.28
15	12.94	12.79	12.31	11.46	10.58	10.02	9.36
16	2.58	2.20	1.98	1.68	2.09	2.27	2.49
17	0.35	0.48	0.58	0.10	0.59	0.68	0.71
18	1.58	1.52	1.53	1.47	1.47	1.41	1.37
19	0.14	0.16	0.18	0.22	0.21	0.22	0.25
20	0.46	0.40	0.14	0.14	0.84	1.74	2.30
21	0.93	0.82	0.55	0.38	0.39	0.09	0.24
22	0.23	0.23	0.75	0.71	1.48	1.84	2.20
23	3.30	3.30	3.31	3.40	2.93	2.60	2.40
24	1.52	1.52	1.52	1.52	1.52	1.52	1.52
25	5.56	5.56	4.50	4.62	5.56	5.56	5.56
Mean	2.39	2.29	2.29	2.02	2.20	2.13	2.08
Std. dev.	4.00	3.89	3.97	3.62	3.62	3.61	3.61

Notes: The table shows the possible trade off between the performance of the models in terms of risk and return in each of the evaluation periods, and the mean and the standard deviation (Std. dev.) of those ratios for each model. To compute the Reward to Semivariability's ratio (R/SV), 3 month Treasury bills are employed as the risk free rate of return. To obtain 25 different non overlapping test sets the next procedure is repeated using non overlapping out of sample rolling windows: at iteration 1 the models are built using observations 1 50 and evaluated on observations 51 75, at iteration 2 models are built using observations 26 75 and evaluated on observations 76 100 and so on.

be seen, here there is a considerable gap between the performances of the different portfolios, obtained by each one of the models. It is quite clear that the losses obtained with models which incorporate the LPM ($n > 2$) as the risk measure (models 5, 6 and 7) are consistently smaller than the ones of the other models (models 1 to 4). It is very relevant how model 7 dominates the rest of the models (on average, it achieves half of the losses obtained from the models that do not use the LPM as the measure of risk). In terms of risk, one can see how the portfolios based on LPM ($n > 2$) attain a reduction in risk levels.

It must be noted that given that: (i) During the studied time periods the mean return was negative, and (ii) It is thought that the investors consider as a risk the obtainment of a loss. It was decided to measure the risk as the total returns under the return zero. So the semideviation is employed with a target equal to zero.

In conclusion, one can affirm that in the case of considering the worst evolution in the market, the portfolios obtained from Models 5, 6 and 7 ($LPM > 2$) outperform all the other portfolios considered in the present study, in terms of return adjusted to risk. Consequently, if an investor is not only interested in the mean performance of his portfolio but also

Table 6. Statistical significance of the measures

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Returns							
Model 1	1	0.93	0.87	0.96	0.81	0.87	0.90
Model 2		1	0.84	0.98	0.75	0.85	0.84
Model 3			1	0.81	0.87	0.96	0.90
Model 4				1	0.85	0.95	0.98
Model 5					1	0.95	0.87
Model 6						1	0.95
Model 7							1
Risk							
Model 1	1	0.98	0.76	0.58	0.99	0.93	0.96
Model 2		1	0.78	0.61	0.95	0.87	0.87
Model 3			1	0.81	0.84	0.72	0.72
Model 4				1	0.61	0.55	0.54
Model 5					1	0.87	0.90
Model 6						1	0.99
Model 7							1
Sharpe Ratio							
Model 1	1	0.98	0.95	0.72	0.96	0.87	0.81
Model 2		1	0.99	0.82	0.99	0.95	0.85
Model 3			1	0.82	0.95	0.93	0.88
Model 4				1	0.79	0.88	0.96
Model 5					1	0.93	0.84
Model 6						1	0.88
Model 7							1
R/SV Ratio							
Model 1	1	0.99	0.88	0.73	0.92	0.92	0.81
Model 2		1	0.96	0.76	0.96	0.96	0.85
Model 3			1	0.85	0.99	0.99	0.85
Model 4				1	0.79	0.90	0.98
Model 5					1	0.98	0.92
Model 6						1	0.93
Model 7							1

Notes: The table shows the results from the Wilcoxon's rank sum test that two populations are identical. A value greater than 0.05 indicates the acceptance of the null.

in the worst case performance, our results clearly indicate that he should employ measures of risk consistent with this fact. Moreover, it has also been shown that if an investor simply tries to maximize his wealth in adverse conditions, he should also employ a risk model that heavily penalizes abnormal (under benchmark) returns, such as those based on the LPM.

V. Conclusions

This study has analysed the performance of risk forecasting models in the construction of optimal

portfolios. Models have been employed that assume homoscedasticity as well as heteroscedasticity. Models have also been considered which assume an asymmetric risk aversion of the decisor. The main conclusion is that, on average, there do not exist any relevant differences between homoscedastic and heteroscedastic models and between symmetric and asymmetric measures of risk. This conclusion is robust along all the criteria employed to evaluate the efficiency of the models.

It was also found that there is a very high and positive correlation between the returns generated by the optimal portfolio, constructed by

Table 7. Number of assets included in the optimal portfolio

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1	13	8	14	14	12	11	10
2	12	10	12	11	11	10	11
3	3	3	4	3	4	4	4
4	3	3	4	3	4	4	4
5	5	5	5	4	6	5	5
6	8	8	7	7	7	7	7
7	9	9	9	8	9	7	7
8	11	11	10	10	9	7	6
9	19	16	17	17	16	15	10
10	11	11	10	10	11	8	5
11	7	7	7	7	6	5	6
12	7	7	7	7	6	6	7
13	4	4	3	3	3	3	3
14	6	7	6	7	6	5	5
15	7	7	7	6	6	6	7
16	8	8	8	9	7	8	7
17	11	11	10	9	10	9	8
18	7	6	5	5	6	6	6
19	6	6	5	4	3	3	3
20	4	4	4	4	3	3	2
21	7	8	6	6	5	4	4
22	10	9	10	10	10	10	7
23	6	5	6	6	5	5	5
24	1	1	1	1	1	1	1
25	1	1	3	3	1	1	1
Mean	7	7	7	7	7	6	6
Std. dev.	4.04	3.43	3.67	3.74	3.61	3.23	2.61

Notes: The table shows the composition of the optimal portfolio generated by the seven models in each of the evaluation periods, and the mean and the standard deviation (Std. dev.) of these number of assets for each model. To obtain 25 different non overlapping test sets the next procedure is repeated using non overlapping out of sample rolling windows: at iteration 1 the models are built using observations 1 50 and evaluated on observations 51 75, at iteration 2 models are built using observations 26 75 and evaluated on observations 76 100 and so on.

Table 8. Number of assets with a weight higher than 3%

	Number of assets	Number of emerging markets assets	Number of developed markets assets
Model 1	5.76	3.04	2.72
Model 2	5.16	2.88	2.28
Model 3	5.16	2.84	2.32
Model 4	5.08	2.80	2.28
Model 5	4.40	2.16	2.24
Model 6	3.64	1.56	2.08
Model 7	3.28	1.24	2.04

Notes: The table shows the number of assets that, on average, compose the optimal portfolio generated by each of the seven models, with a weight higher than 3%.

all the ex-ante models which means that all the models provide quite similar investment decisions. In particular, it was found that ARCH models seem useless for optimal ex-ante portfolio selection and that the models based on semivariance do not show any improvement of efficiency even when this is measured by a consistent statistic such as the Reward-to-Semivariability Ratio. Interestingly, it was found that the returns generated from models which employ LPM, as a measure of risk, have a less negative skewness (which is appreciated by investors) and higher kurtosis.

The most important result is that when one turn one's attention to the analysis of worst case performance, the results are clearly different. Models which employ LPM with a high degree

Table 9. Correlation coefficients of the returns of the optimal portfolios

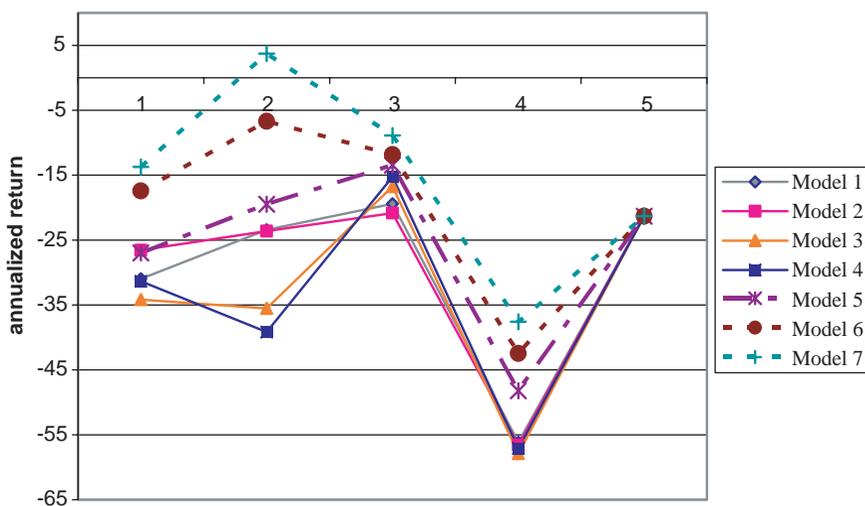
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Model 1	1	0.97	0.98	0.97	0.93	0.87	0.82
Model 2		1	0.94	0.95	0.90	0.84	0.80
Model 3			1	0.99	0.96	0.90	0.86
Model 4				1	0.96	0.90	0.85
Model 5					1	0.98	0.95
Model 6						1	0.99
Model 7							1

Notes: The table shows the correlation coefficients between the portfolios' returns, generated by each one of the models employed.

Table 10. Characteristics of the empirical distribution of portfolios

Portfolios from model i	Mean	Standard deviation	Kurtosis	Skewness
Model 1	14.38	17.27	4.65	0.54
Model 2	14.22	17.39	4.88	0.56
Model 3	14.93	18.33	4.94	0.33
Model 4	14.30	18.47	4.65	0.40
Model 5	16.30	18.25	6.22	0.14
Model 6	16.69	18.55	8.07	0.07
Model 7	16.79	18.51	8.36	0.13

Notes: The table shows the descriptive statistics of the empirical distribution of the optimal portfolio's returns, generated by each one of the seven estimated models. They are the mean, standard deviation (Std. dev.), kurtosis and skewness.

**Fig. 1. Worst cases: annualized returns**

This graph shows the annualized return in percent obtained from each one of the seven models for the five worst cases. To compute the worst case time periods we consider the five iterations in which the ex post evolution of the market really showed the highest losses.

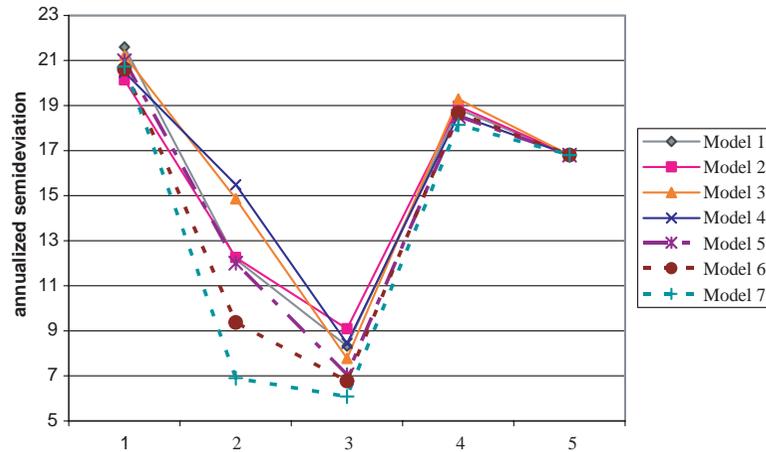


Fig. 2. Worst cases: risk measured as annualized semideviation

This graph shows the risk level measured as the annualized semi deviation (considering a target equal to zero to compute it) obtained from each one of the seven models for the five worst cases. To compute the worst case time periods the five iterations are considered in which the ex post evolution of the market really showed the highest losses.

of risk aversion ($n > 2$) as the risk measure, are consistently superior to those which employ a symmetric measure, either homoscedastic or heteroscedastic. To sum up, the results clearly indicate that the problem of symmetry is much more important than the one of homoscedasticity.

Acknowledgements

We thank David Nawrocki, for helpful comments and suggestions. The contents of this paper are the sole responsibility of the authors.

References

Ang, J. S. and Chua, J. H. (1979) Composite measures for the evaluation of investment performance, *Journal of Financial and Quantitative Analysis*, **14**, 361-84.

Bawa, V. S. (1975) Optimal rules for ordering uncertain prospects, *Journal of Financial Economics*, **2**, 95-121.

Best, M. J. and Grauer, R. R. (1991) On the sensitivity of mean variance efficient portfolios to changes in asset means: some analytical and computational results, *Review of Financial Studies*, **4**, 315-42.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, **31**, 307-27.

Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992) ARCH modeling in finance: a review of the theory and empirical evidence, *Journal of Econometrics*, **52**, 5-59.

Elton, E. J. and Gruber, M. J. (1973) Estimating the dependence structure of share prices implications for portfolio selection, *Journal of Finance*, **28**, 1203-32.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987-1008.

Eun, C. S. and Resnick, B. G. (1992) Forecasting the correlation structure of share prices: a test of new models, *Journal of Banking and Finance*, **16**, 643-56.

Fishburn, P. C. (1977) Mean risk analysis with risk associated with below target returns, *American Economic Review*, **67**, 116-26.

Jobson, J. D. and Korkie, B. (1980) Estimation for Markowitz efficient portfolios, *Journal of American Statistical Association*, **75**, 544-54.

Jorion, P. (1986) Bayes Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis*, **21**, 279-92.

Kroll, Y., Levy, H. and Markowitz, H. M. (1984) Mean variance versus direct utility maximization, *Journal of Finance*, **39**, 47-61.

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**, 603-21.

Levy, H. and Markowitz, H. M. (1979) Approximating expected utility by a function of mean and variance, *American Economic Review*, **69**, 308-17.

Mandelbrot, B. (1963) The variation of certain speculative prices, *Journal of Business*, **36**, 394-419.

Markowitz, H. M. (1952) Portfolio selection, *Journal of Finance*, **7**, 77-91.

Markowitz, H. M. (1991) Foundations of portfolio theory, *Journal of Finance*, **46**, 469-77.

Merton, R. C. (1980) On estimating the expected return on the market: an exploratory investigation, *Journal of Financial Economics*, **8**, 323-61.

Michaud, R. O. (1989) The Markowitz optimization enigma: is 'optimized' optimal?, *Financial Analysts Journal*, **45**, 31-42.

Nawrocki, D. N. (1991) Optimal algorithms and lower partial moment: ex post results, *Applied Economics*, **23**, 465-70.

Nawrocki, D. N. (1992) The characteristics of portfolios selected by N degree lower partial moment, *International Review of Financial Analysis*, **3**, 195-209.

- Nawrocki, D. N. (1999) A brief history of downside risk measures, *Journal of Investing*, **8**, 9-25.
- Nelson, D. B. (1992) Filtering and forecasting with misspecified ARCH Models I: getting the right variance with the wrong model, *Journal of Econometrics*, **52**, 61-90.
- Pulley, L. (1981) A general mean variance approximation to expected utility for short holding periods, *Journal of Financial and Quantitative Analysis*, **16**, 361-73.
- Pulley, L. (1985) Mean variance versus direct utility maximization: a comment, *Journal of Finance*, **40**, 601-2.
- Roy, A. D. (1952) Safety first and the holding of assets, *Econometrica*, **20**, 431-49.
- Sharpe, W. (1966) Mutual fund performance, *Journal of Business*, **39**, 119-38.
- Taylor, M. P. (1995) The economics of exchange rates, *Journal of Economic Literature*, **33**, 13-47.