https://doi.org/10.26219/heal.aueb.6306

ATHENS UNIVERSITY
OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

ARCH and GARCH Modeling in the Process of Portfolio Optimization

By

Andreas G. Golfis

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
June 2008

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Η χρήση των μοντέλων ARCH και GARCH στην διαδικασία κατασκευής βέλτιστων χαρτοφυλακίων

Ανδρέας Γ. Γκόλφης

ΔΙΑΤΡΙΒΗ
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Ιούνιος 2008

https://doi.org/10.26219/heal.aueb.6306

DEDICATION

Στην Οικογένειά μου
Βασιλική, Γεώργιος και Χαράλαμπος

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

ACKNOWLEDGEMENTS

I am grateful to my supervisor I. Vrontos for his continuous advice and support throughout this research. In addition I wish to thank the staff of the Library of Athens University of Economics and Business and the support team of the Laboratory of Department of Statistics for their precious help.

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

VITA

I was born in Athens in 1978. I finished my High – School Studies in 1996 (3rd Lyceum of Vironas) and after taking examinations, I attended Athens University of Economics and Business, Department of International and European Economic Studies (1997).

As an undergraduate student I took courses in the following areas: Economic Theory (Macro & Micro Economics), Finance – Portfolio Management, Econometrics, Integration Policies, Public Sector & Fiscal Policies, E. U. – Law of Enterprises and Competitiveness etc.

In 2002, I finished my undergraduate studies and I took up my postgraduate studies in Applied Statistics, at the Department of Statistics (A.U.E.B.). During my postgraduate studies, I took courses in the following scientific areas: Computational Statistics, Experimental Design – Linear & Generalized Linear Models, Stochastic Methods in Finance, Time Series, Multivariate Analysis, Probability theory, Sampling techniques, Environmental Statistics etc.

For the last two years I have been occupied as an Investment Consultant in several private sector companies and public sector organizations. I mainly focus on marketing and viability plans related to subsidies for small and medium enterprises coming through the European Developing Programes.

I am interested in environmental issues, ancient culture and Greek civilization (literature, philoshophy, history etc).

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

ABSTRACT

Andreas Golfis

ARCH and GARCH Modeling in the Process of Portfolio Optimization

June 2008

In this dissertation we rely on two optimization techniques expounded by Markowitz (1952) and Tobin (1958) in order to construct optimal portfolios, out of a pool of 36 stocks, observed over a total of 2.780 days. We apply these two portfolio optimization techniques using daily discrete returns for an estimation period of 2.415 daily observations and allowing for short sales. The performance of the optimal portfolios, during the evaluation period, shows no remarkable difference when it comes to the way of modeling a stock's return volatility - either using the historical method or GARCH modeling.

In total we construct 9 optimal portfolios. We put together the first three of them following the methodology of Markowitz (1952) and by setting portfolio's expected return level (on a daily basis) at 0.10%, 0.15% and 0.20%. Then, drawing on Tobin's (1958) approach, we construct other three optimal portfolios by taking into account three yearly risk-free rates 3.0%, 3.5% and 4.0%. The remaining three portfolios are built by estimating a $GARCH(1,1)$ model and then using the Tobin's approach with three yearly risk-free rates of 3.0%, 3.5% and 4.0%.

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

ΠΕΡΙΛΗΨΗ

Ανδρέας Γκόλφης

Η χρήση των μοντέλων ARCH και GARCH στην διαδικασία κατασκευής βέλτιστων χαρτοφυλακίων

Ιούνιος 2008

Στην διατριβή αυτή βασιζόμαστε σε δύο τεχνικές βελτιστοποίησης, οι οποίες αναπτύχθηκαν από τον Markowitz (1952) και Tobin (1958), με σκοπό να κατασκευάσουμε βέλτιστα χαρτοφυλάκια, από μία ομάδα 36 μετοχών για τις οποίες υπάρχουν παρατηρήσεις για συνολικά 2.780 ημέρες. Εφαρμόζουμε τις δύο αυτές τεχνικές βελτιστοποίησης χαρτοφυλακίου, χρησιμοποιώντας ημερήσιες διακριτές αποδόσεις και λαμβάνοντας υπ’ όψη την δυνατότητα διενέργειας «ανοιχτών πωλήσεων» (short sales). Η πορεία των βέλτιστων χαρτοφυλακίων κατά την περίοδο αξιολόγησης δεν έδειξε σημαντικές διαφορές αναφορικά με τον τρόπο υποδειγματοποίησης της μεταβλητότητας των αποδόσεων των μετοχών – είτε χρησιμοποιώντας ιστορική μέθοδο υποδειγματοποίησης, είτε το υπόδειγμα GARCH.

Συνολικά κατασκευάζουμε εννέα (9) βέλτιστα χαρτοφυλάκια. Συγκροτούμε τα πρώτα τρία (3) βασιζόμενοι στην μεθοδολογία του Markowitz και θέτοντας την αναμενόμενη απόδοση του χαρτοφυλακίου (σε ημερήσια βάση) ίση με 0.10%, 0.15% και 0.20%. Κατόπιν χρησιμοποιώντας την μέθοδο του Tobin (1958), κατασκευάζουμε τρία άλλα βέλτιστα χαρτοφυλάκια λαμβάνοντας υπ’ όψη τρία επιτόκια χωρίς κίνδυνο ίσα με 3.0%, 3.5% και 4.0% αντίστοιχα. Τα υπόλοιπα τρία χαρτοφυλάκια κατασκευάζονται από την εκτίμηση ενός $GARCH(1,1)$ μοντέλου και κατόπιν χρησιμοποιώντας την μέθοδο του Tobin, με τρία επιτόκια χωρίς κίνδυνο ίσα με 3.0%, 3.5% και 4.0% αντίστοιχα.

https://doi.org/10.26219/heal.aueb.6306

TABLE OF CONTENTS

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

LIST OF TABLES

https://doi.org/10.26219/heal.aueb.6306

https://doi.org/10.26219/heal.aueb.6306

LIST OF FIGURES

https://doi.org/10.26219/heal.aueb.6306

1. INTRODUCTION

The objective of this dissertation is to examine whether it is feasible by drawing on different *portfolio optimization techniques*, each of which may model a stock's volatility either through the simple historical variance measure or through a GARCH (or an ARCH) model, to construct optimal portfolios that can bring in substantial profits to the investor. Our interest in conditional variance stems for the fact that conditional variance, in some cases, may be important than the unconditional variance. For instance, a stock holder would be interested in forecasts of the rate of return and its variance over the holding period. The unconditional variance (i.e., the long-run forecast of the variance) would be unimportant if he plans to buy the asset at time t and sell at time $t+1$.

We hope that the results of our study will help us in two aspects. At first, we will see whether a particular method of volatility modelling can bring in higher profits to the investor. Then, we will decide whether it is possible by using ex-post data in order to work with different optimization models (which nevertheless require ex-ante data,), and of course different measures of modelling stock volatility, to earn descent returns on our optimally constructed portfolios. The two optimization techniques we shall employ in the construction of optimal portfolios have been expounded by Markowitz (1952) and Tobin (1958).

After this introductory first chapter, we expound in detail the methodological framework of the dissertation in Chapter 2, while in Chapter 3 we present and discuss the empirical findings of our analysis.

1.1 THE THEORETICAL FRAMEWORK

The theoretical framework of the study is made up of the process of portfolio optimization and volatility modelling.

1.1.1 The Investment Management Process and Portfolio Optimization

The topic of our dissertation is related with the fourth step of the investment management process. According to Fabozzi (1999) the **investment management process** consists of five steps

1. setting investment objectives
2. establishing the investment policy
3. selecting a portfolio strategy
4. selecting the assets
5. measuring and evaluating performance

Let us say a few things on each of the above steps.

Setting Investment Objectives

Setting investment objectives constitutes a critical part of the process of portfolio construction. There are four main portfolio objectives: (a) stability of principle, (b) income, (c) growth of income, and (d) capital appreciation (Strong, 2000).

Establishing the Investment Policy

The **investment policy** has to do with the way an investor, or an institution, must distribute its funds among the different classes of assets.

There are several things to consider in this step of the investment management process. First, an investor, or an institution, may wish to maintain a minimum level of liquidity and safety in its investments, hence this attitude trims the amount of funds available for investment in different risky assets. Second, there may be other considerations that an investor needs to take into consideration before he decides which assets to long.

Selecting the Portfolio Strategy

There are two types of portfolio strategies: (a) active portfolio strategy, and (b) passive portfolio strategy (Fabozzi, 1999).

An **active portfolio strategy** relies on forecasting techniques in order to attain better results (i.e. excess returns) than the returns of a portfolio, which is simply broadly diversified and tracks a market index. Apparently, an investor who selects an active management does not believe in the efficient market hypothesis. Instead, the investor hopes that he can profit from the stock market through any number of strategies, including optimization strategies.

On the other hand, a **passive portfolio strategy** relies on diversification to match the performance of some stock market index. Apparently, the followers of passive management believe in the efficient market hypothesis, according to which stock markets incorporate and reflect all information (past, present, and inside information), so searching for “winning” stocks is futile.

Selecting the Assets

In this stage of the investment process the investment manager evaluates individual securities and constructs an **efficient,** or **optimal portfolio**, that is, a portfolio that provides the greatest *expected return* for a given level of *risk*. The construction of an efficient portfolio requires that we make some assumptions about the investors’ stance toward risk (Fabozzi, 1999:39). The most widely used assumption concerning the investors’ stance is that investors are risk averse. This implies that when investors are faced with two investments having the same expected return but different risk then they will prefer the one with the lower risk.

Measuring and Evaluating Performance

The last step in the investment management process involves the measurement and the evaluation of the portfolio’s performance. It is a very important step, since if the evaluation outcome turns out to be not as expected the investor may change his investment strategy.

1.1.2 Modeling the Return of an Asset as a Stochastic Process

Estimates of expected returns and volatilities and their effects on asset and derivative prices are essential in financial decision making. According to Engle (2001), the return r on an asset or portfolio is equal to the mean value of it (that is, the expected value of r based on past information) plus the standard deviation of r (that is, the square root of the variance) times the error term for the present period.

We know that stock price fluctuations have several important characteristics. First, over the long run, stock prices go up. They are said to "drift." This represents the return from bearing risk. Second, stock prices are random. We know that Wiener processes are random, though we cannot use the basic Wiener process for every stock because different stocks have different volatilities. We can, however, transform the basic Wiener process to give it a different volatility. Third, it should be harder to forecast stock prices further into the future than nearby. That does not mean that stock prices are very predictable but that the margin of error, which is related to the variance of the future stock price, should be greater when predicting far into the future than when predicting into the near future. The final property is that a stock price should never be allowed to become negative. Corporate shareholders have limited liability so the minimum value of their shares is zero.

So given the above characteristics of stock price fluctuations, we can write the return of a stock, r_t, over a holding period (say a month) as follows (Chance, 1994)

$$r_t = \mu dt + \sigma Z_t \sqrt{dt} \qquad (1.1)$$

where, Z_t is a standardized normal variable, σ is the stock's returns standard deviation, and μ the mean return of the stock over a given holding period.

The econometric challenge is to specify how the information is used to forecast the mean (μ) and variance (σ^2) of the return, conditional on past information. While many specifications have been considered for the mean return and have been used in efforts to forecast future returns, virtually no methods were available for the variance before the introduction of ARCH and GARCH models.

If S_t is the stock price at time t, then using equation (1.1) we can find the stock price after time period dt as follows

https://doi.org/10.26219/heal.aueb.6306

$$\begin{aligned} S_{t+dt} &= S_t\left(1+r_t\right) = \\ &= S_t\left[1+\mu dt+\sigma Z_t\sqrt{dt}\right] \end{aligned}$$

Thus, it follows that the capital gains over the holding period dt must be

$$\frac{dS_{t+dt}}{S_t}=\frac{S_{t+dt}-S_t}{S_t}=\mu dt+\sigma Z_t\sqrt{dt} \qquad (1.2)$$

Equation (1.2) describes the most widely used model for stock price behaviour over a very short-period of time; this model is known as the model of **geometric Brownian motion** (Hull, 1997: 217).

For example, if a stock pays no dividends, provides an expected return of 10% per annum with continuous compounding, and it has a volatility of 15% per annum, then the process for this stock price must be

$$\frac{dS_{t+dt}}{S_t}=0.10dt+0.15Z_t\sqrt{dt}$$

Suppose that we are considering weekly changes in the stock price, i.e. $dt=7/365=0.0192$. If the initial stock price is €20 then the change in the price of the stock over the week is given by the following equation

$$\begin{aligned} dS_{t+dt} &= 20\left[0.10\left(0.0192\right)+0.15Z_t\sqrt{0.0192}\right] = \\ &= 0.0385+0.4160Z_t \end{aligned}$$

1.1.3 Modeling the Volatility of a Stock’s Returns with ARCH and GARCH Models

As we pointed out previously, our goal is using certain portfolio optimization models to construct an *optimal* portfolio. These models make use of the stock volatility measure, which is typically the standard deviation of its returns.

Hence a main focal point of this dissertation is the *modelling of a stock’s returns volatility*; this modelling can be either static or dynamic in nature. Despite the fact that a stock’s volatility is measured by the standard deviation of its returns, volatility modelling is cast in terms of formulating models for the variance.

Many financial time series (such as a stock’s returns) exhibit *volatility clustering*, which means that the series have periods where volatility is low and other periods where volatility is high. For the purposes of our dissertation we interpret “volatility" to mean “conditional variance". The conditioning set will always be the behaviour of relevant variables up to time t, i.e. the time when we observe the series. There have recently been developed many different models for data that shows volatility clustering. The main difference between competing models will often be the choice of which variables to condition on, and (as usual) the choice of functional forms. One of the main ideas of asset-pricing is that the variability of an asset should be reflected in its price. An investor would expect an asset with high variance to give a higher return, in order for him to want to hold it. This is the main point of the CAPM and APT models that are common in finance.

There are various reasons that account for the interest in volatility modelling. To begin with, one of the main ideas of the asset-pricing models is investors must be rewarded only for taking on systematic risk, which is represented by the firm's beta coefficient. This type of risk is related to the covariance between the stock's returns and the returns in the market portfolio. Then, option-pricing formulas give the prices of options (and other derivative instruments) in terms of volatility of the underlying asset. Finally, banks and other financial institutions apply so-called value-at-risk models to assess risks in their marketable assets.

For all these reasons, we need to model volatility (risk). However, according to standard economic reasoning risk should not be measured as the *unconditional variance* but rather as the conditional variance. The ARCH and GARCH models are supposed to model a variable's conditional volatility

Measuring a Stock's Returns Volatility

One way to measure and forecast future stock volatility is to assume that a past risk measure (such as the standard deviation) constitutes the best estimate the future risk measure,

Historical volatility reflects how far an instruments price has deviated from its average price (mean) in the past. On a yearly basis, this number represents the one standard deviation % price change expected in the year ahead. For example, if a stock is trading at $100 and has a volatility of 20% then there is a 68% probability (1 standard dev = 68% probability) that the price will be in the range 80 to 120 a year from now. Similarly there is a 95% probability that the price will be between 60 and 140 a year from now (2 standard deviations 1.96). The higher the volatility number the higher the volatility.

When a stock's volatility is captured by its standard deviation or variance (σ^2), we compute it by modelling the stock's excess return, denoted by $r_{e,t}$, in period t, as follows

$$\hat{\sigma}_t^2 = \frac{1}{T}\sum_{j=1}^{T}\left(r_{t-j} - \overline{r}\right)^2 = \frac{1}{T}\sum_{j=1}^{T} r_{e,t-j}^2 \tag{1.3}$$

Where

$$\overline{r} = \frac{1}{T}\sum_{j=1}^{T} r_{t-j} \tag{1.4}$$

In the above equation T represents the returns' time-horizon, over which their variance is computedNotice, the sample variance of stock returns defined as in (1.3) is a biased estimator of the population variance. This moves us from an unbiased estimate of the variance to a maximum likelihood (ML) estimate, which is the standard estimation approach in volatility models. An unbiased estimator (s_t) of the population variance will be

$$s_t^2 = \frac{1}{T-1}\sum_{j=1}^{T}\left(r_{t-j} - \overline{r}\right)^2 = \frac{1}{T-1}\sum_{j=1}^{T} r_{e,t-j}^2 \tag{1.5}$$

Another way to model a stock's volatility is to condition the future forecast of it on past information using either ARCH or GARCH modelling.

To illustrate the notion of GARCH modelling of a stock's volatility, notice the model of variance in (1.3) resembles a moving average model. All observations, from the more recent at time period $t-1$ to the most distant at time period $t-T$, are given equal weight (in four-period example 1/4), while all observations before $t-T$ are given no weight; of course the choice of T is left to the analyst's discretion.

It is often, however, more reasonable to have more recent values of $r_{e,t}$ play a greater role than earlier values. In that case recent excess returns of the security should be weighted heavily in the moving average model of volatility.

A model that accomplished the above type of weighting involves the **exponential smoothing**, according to which the model of a stock's volatility becomes

$$\sigma_t^2 = \alpha r_{e,t-1}^2 + \alpha\left(1-\alpha\right) r_{e,t-2}^2 + \alpha\left(1-\alpha\right)^2 r_{e,t-3}^2 + \ldots \tag{1.6}$$

The summation above extends all the way back through the length of the series. Multiplying both sides of model (1.2) by $\left(1-\alpha\right)$ and lagging it by one period we get

$$\left(1-\alpha\right)\sigma_{t-1}^{2}=\alpha\left(1-\alpha\right)r_{e,t-2}^{2}+\alpha\left(1-\alpha\right)^{2}r_{e,t-3}^{2}+\alpha\left(1-\alpha\right)^{3}r_{e,t-4}^{2}+\ldots \qquad (1.7)$$

And subtracting model (1.7) from model (1.6) yields the following

$$\sigma_{t}^{2}=\alpha r_{e,t-1}^{2}+\left(1-\alpha\right)\sigma_{t-1}^{2} \qquad (1.8)$$

The above equation shows that an estimate of the volatility for period t (made at period $t-1$) is calculated from σ_{t-1} , i.e. the estimate that was made one day ago of the volatility for period $t-1$, and $r_{e,t-1}^{2}$, that is, the most recent observation on changes in the market variable. The above equation method emphasizes that the volatility (σ_{t-1}) on a given day $t-1$ is actually used as a predictor for the volatility (σ_{t}) of the next day t . The parameter α governs how responsive the estimate of the daily volatility is to the most recent observations on the $r_{e,t-1}^{2}$. Once the value of parameter α is known and an initial variance (σ_{t-1}^{2}) is given such as from the first few observations, then all the variances can simply be computed.

For example, suppose that $\alpha=0.10$, the volatility for day $t-1$ is 1% per day, and the excess return of a stock during day $t-1$ was 2%. This means that $\sigma_{t-1}^{2}=0.01^{2}$ and $r_{e,t-1}^{2}=0.02^{2}$. Hence, based on equation (1.8), the estimate for the volatility for day t, σ_{t} , must be 1.14% per day

$$\begin{aligned}\sqrt{\sigma_{t}^{2}}&=\sqrt{0.10\times0.0004+\left(1-0.10\right)0.0001}=\\&=\sqrt{0.00013}=0.0114\end{aligned}$$

In GARCH modeling, we slightly modify equation (1.8) as follows .

$$\sigma_{t}^{2}=\gamma V+\alpha r_{e,t-1}^{2}+\beta\sigma_{t-1}^{2} \qquad (1.9)$$

With $\gamma+\alpha+\beta=1$. Setting $\omega=\gamma V$ in equation (1.9) we derive the GARCH(1,1) model

$$\sigma_{t}^{2}=\omega+\alpha r_{e,t-1}^{2}+\beta\sigma_{t-1}^{2}$$

https://doi.org/10.26219/heal.aueb.6306

And the long-run variance rate V is given by the following relation

$$V = \frac{\omega}{1-\alpha-\beta}$$

Suppose for example that a GARCH-(1.1) model estimated from daily data turns out to be

$$\sigma_t^2 = 0.000002 + 0.13 r_{e,t-1}^2 + 0.86\sigma_{t-1}^2$$

Form the above estimated model it follows that that the long-run variance rate (V) and the long-run volatility per day ($\sqrt{V}$) have as follows

$$V = \frac{\omega}{1-\alpha-\beta} = \frac{0.000002}{1-0.13-0.86} = 0.0002$$

$$\sqrt{V} = 0.0141$$

Suppose now that the current estimate of the volatility on day $t-1$ is $\sigma_{t-1} = 1.6\%$ per day and that the most recent proportional change in the market variable on day $t-1$ is $r_{e,t-1} = 0.01$. The new variance rate then is

$$\sigma_t^2 = 0.000002 + 0.13\left(0.01\right)^2 + 0.86\left(1.6\right)^2 = 0.00024$$

And the new volatility is 1.53% (=$\sqrt{0.00024}$) per day

https://doi.org/10.26219/heal.aueb.6306

1.2 THE METHODOLOGICAL FRAMEWORK

The methodological framework of the dissertation has the following structure. First we rely on two methods used in structuring optimal portfolios: (a) the method introduced by Markowitz (1952), and (b) the method used Tobin (1958). Both portfolio optimization techniques rely on the computation of the variance-covariance matrix $\mathbf{\Sigma}$ of stock returns.

Then, the computation of matrix $\mathbf{\Sigma}$, based on the data of stock returns of the *estimation period*, can be carried out in two ways. The first way involves historical data and second way the use of a GARCH model. Finally, given the two different approaches used in determining matrix $\mathbf{\Sigma}$ we derive different optimal portfolios (i.e. different optimal weights), according to the two different optimization methods; these optimal portfolios are tested during the *evaluation period*.

https://doi.org/10.26219/heal.aueb.6306

2. METHODOLOGY

In this chapter we shall analyze two methods used in structuring optimal portfolios (Sections 2.1 and 2.2). One optimization technique was expounded by Markowitz (1952) and another by Tobin (1958). Both techniques make use of the asset's volatility as a "primary" input in the optimization process, and they measure this volatility based on historical data on stock returns.

Many analysts look at the volatility over the last week or last month or even the last five years to give them an estimate of what the volatility is today and what it is likely to be in the future. The thing that's difficult is that they don't know which window to use. If they use something like a five-year window, they will include a lot of information that may not be very relevant for today. And if then include just a one-week window they will get a very noisy estimate.

An alternative way of modeling a stock's return volatility is through the use of an ARCH or GARCH model (Sections 2.3 and 2.4). These models give us the best of both of these worlds by using the weighted average of past volatility, where we give high weights to recent volatility evidence and small weights to the distant past.

ARCH models, introduced by Engle (1982), and GARCH models, which are generalizations of the ARCH models and were introduced by Bollerslev (1986), aim at modeling the conditional volatility of a time series variable. This volatility is always taken to mean *conditional volatility*, where the conditioning is set to be the behavior of relevant variable up to time $t-1$ -the time when we observe the series.

Many time series variables (particularly financial ones such as stock returns) exhibit changes in volatility over time; these changes tend to be serially correlated, with groups of highly volatile observations occurring together. This is known as **volatility clustering,** that is, a time-series variable may have periods where volatility is low and other periods where volatility is high. Indeed, in considering time series of price changes, Mandelbrot (1963) observed that large changes tend to be followed by large changes − of either sign − and small changes tend to be followed by small changes.

https://doi.org/10.26219/heal.aueb.6306

2.1 OPTIMAL PORTFOLIOS BASED ON THE MARKOWITZ APPROACH

Prior to Markowitz's work, investors focused on assessing the risks and rewards of individual securities in constructing their portfolios. Standard investment advice was to identify those securities that offered the best opportunities for gain with the least risk and then construct a portfolio from these. Acting on this investment perspective an investor might conclude that if all the stocks of a particular sector offered good risk-reward characteristics, he would have to compile a portfolio consisting entirely from these stocks; intuitively, this would be foolish.

Markowitz (1952) formalized this intuition in his path-breaking work by proposing that that investors should focus on selecting portfolios based on their overall risk-reward characteristics instead of merely compiling portfolios from securities that each individually have attractive risk-reward characteristics. In a nutshell, inventors should select portfolios not individual securities.

In what follows, we show how we compute a portfolio’s expected return and risk

2.1.1 A Portfolio’s Expected Return and its Risk

Modern portfolio optimization theory makes use of **ex ante data** (i.e. before the event), since it requires as inputs expected security returns and risk measures (that will capture future volatility). In other words, in all optimization models there is uncertainty that is resolved during the course of events with the use of the ex antes values, i.e. of expected returns and risk.

Every stock i in the market has a certain return r_{it}, which is stochastic (i.e. its future value is not known with certainty), or random variable. Every random variable takes values based on different scenarios, each with a probability less than or equal to 1, and it is broken down into a fixed component and a pure random component (Dougherty, 1992:14). So in the case of a stock’s return we have:

$$r_{it} = \mu_i + u_{it}$$

where

μ_i: the expected return of stock i

u_i: the random component in the return of stock i

Theoretically speaking, the expected return on stock i is given by the following formula:

$$\mu_i = E(r_i) = \sum_{j=1}^{M} p_{ij} r_{ij} \tag{2.1}$$

where

p_{ij} : the probability that state j will occur

r_{ij} : the return of asset i at state j

M : the number of states in the economy

However a problem arises when all we can observe are the past stock prices and not the probability distribution of each stock. Further, the expected return on a stock may not be determined according to any asset pricing model, such as the CAPM. In that case, the simplest approach to estimate the expected return on a stock (or any other financial asset) is to rely on historical **ex post data**. In other words, from the past returns observed over a given time window of size T (estimation period), we approximate the expected return on a stock i (μ_i) with the arithmentic mean return, $\overline{r_i}$, defined as follows:

$$\overline{r_i} = \frac{1}{T} \sum_{j=1}^{T} r_{i,t-j}$$

Unfortunately, this approach of *ex post* measures of return (and risk) has proven disappointing in some case, like the one of real estate portfolio diversification (Mueller and Laposa 1995). Similarly, Sivtanides and Southard (2000) have showed that historical risk measures (i.e. a measure of standard deviation based on historical stock returns) are poor indicators of future volatility.

Much of the failure of the *ex ante* portfolio performance is due to the estimation of risk in the portfolio selection procedure (Lee and Stevenson 2000).

The Expected Return on a Portfolio of Stocks

Suppose now that we consider constructing a portfolio $P = \sum_{i=1}^{N} w_i r_i$, which made up of N stocks (assets) with expected returns $\overline{r_1}, \ldots, \overline{r_N}$. In that case, the **expected return on a portfolio** is simply the weighted average of the N expected returns

https://doi.org/10.26219/heal.aueb.6306

$$\mu_p = E(P) = \sum_{i=1}^{N} w_i E(r_i) = \sum_{i=1}^{N} w_i \mu_i$$

where

w_i : the proportion of funds committed to asset i

N = the number of assets

Or using matrix notation and switching to sample estimates, we can write

$$\overline{r}_p = \mathbf{w}'\overline{\mathbf{r}} \qquad (2.2)$$

Where

$\overline{\mathbf{r}}$: the $N \times 1$ column-vector of average returns, i.e. $\overline{\mathbf{r}}' = (\overline{r}_1, \dots, \overline{r}_N)$

$\mathbf{w}$: the $N \times 1$ column-vector of weights

Since the average return $\overline{r}_p$ is a weighted average of N random variables, it follows that it is a random variable in its own.

The Risk on a Portfolio of Stocks

The variance on a portfolio consisting of N assets can be computed applying variance rules on $P = \sum_{i=1}^{N} w_i r_i$, in which case we get

$$\sigma_p^2 = \operatorname{var}(P) = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{k=1}^{N} w_i w_k \sigma_{ik} \text{, for } k \neq i \qquad (2.3)$$

where

σ_i^2 : the variance of asset's i returns

σ_{ik} : the covariance between the returns of asset i and asset k

Theoretically, the variance of a stock's returns is a weighted average of the squares of the deviations of the asset's returns from its expected value

$$\sigma_i^2 = \sum_{j=1}^{M} p_{ij} \left(r_{ij} - \overline{r}_i\right)^2 \qquad (2.4)$$

https://doi.org/10.26219/heal.aueb.6306

If the return for each security follows a *normal distribution*, then we can apply the **reproducible property** of the normal distribution. That is, if we let $\{R_i\}_{i\geq 1}$ be sequence of N independent continuous random variables such that

$$R_i \sim N\left(\overline{r}_i,\ \sigma_i^2\right) \text{ for } i = 1, 2, .., N$$

Then, the random variable

$$P = \sum_{i=1}^{N} w_i R_i \sim N\left(\overline{r}_p,\ \sigma_p^2\right) \qquad (2.5)$$

this means that we can completely describe its distribution by two terms: the mean (expected return) and the variance of returns.

2.1.2 An Analytic Approach to Portfolio Optimization According to Markowitz's Approach

Suppose that we are choosing a portfolio from N securities with an expected return vector $\overline{\mathbf{r}}' = \left(\overline{r}_1, \ldots, \overline{r}_N\right)$ and a covariance matrix $\mathbf{\Sigma}$.

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \ldots & \sigma_N^2 \end{bmatrix} \qquad (2.6)$$

The question is which of these available N stocks (assets) constitute the best choice for an investor. That is, we need to know out of these N securities which to select and what proportions of the available money capital to devote to each selected security. As we shall shortly show, in this context of analysis, there is not a single portfolio that is best; instead there are many efficient (optimal) portfolios, each of which offers the smallest possible risk for a given level of return.

More formally, in Markowitz framework of risk aversion, consumers prefer more return to less return, and like less variance to more variance. Further, investors must be compensated with higher returns in order to accept higher risk.

In other words, **risk-aversion** dictates that:

1. Any investor who chooses to tolerate a risk level of σ_p^2 will construct a portfolio with *maximum* expected return possible among those portfolios with the same variance σ_p^2.
2. Any investor who opts for an expected return $E\left(r_p\right)$ on the portfolio will want the portfolio with the *minimum variance* possible among those with the same expected return $E\left(r_p\right)$.

Put in mathematical terms, the **optimal investment decision** involves finding a set of optimal weights $\mathbf{w}'$ that minimize the portfolio's risk σ_p, subject to the constraint that a given level of the portfolio's expected return $\mu = \overline{\mathbf{r}}\mathbf{w}$ is attained (Chou *et al.* 2004)

where

$\sigma_p = \sqrt{\mathbf{w}'\mathbf{\Sigma}\mathbf{w}}$

$\mathbf{w}$: the $N \times 1$ column-vector of weights

$\mathbf{\Sigma}$: the variance-covariance matrix; it is a symmetric matrix

$\overline{\mathbf{r}}$: the $N \times 1$ column-vector of expected returns, where $\overline{\mathbf{r}}' = \left(\overline{r}_1, \ldots, \overline{r}_N\right)$;

The solution to the above maximization problem is given by the following expression

$$\mathbf{w}^* = \frac{C - \mu B}{\Delta}\mathbf{\Sigma}^{-1}\mathbf{1} + \frac{\mu A - B}{\Delta}\mathbf{\Sigma}^{-1}\overline{\mathbf{r}} \qquad (2.7)$$

Where

$$A = \mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1} \qquad (2.8)$$

$$B = \mathbf{1}'\mathbf{\Sigma}^{-1}\overline{\mathbf{r}} \qquad (2.9)$$

$$C = \overline{\mathbf{r}}'\mathbf{\Sigma}\overline{\mathbf{r}} \qquad (2.10)$$

$$\Delta = AC - B^2 \qquad (2.11)$$

https://doi.org/10.26219/heal.aueb.6306

Varying the expected return on the portfolio we derive another set of optimal weights. All the portfolios that meet the aforementioned criteria (i.e. for a given level of return they have the minimum variance) make up the **efficient frontier**; put differently, portfolios along the efficient frontier are optimal portfolios in the sense that they have the minimum risk for a given level of return.

Now let us consider the case of individual investors. Still we don't know which portfolio along the efficient frontier the individual investor will select. In order to deal with this problem we need to introduce investors' preferences. In the following figure we draw the efficient frontier along with a set of indifference curves for an investor.

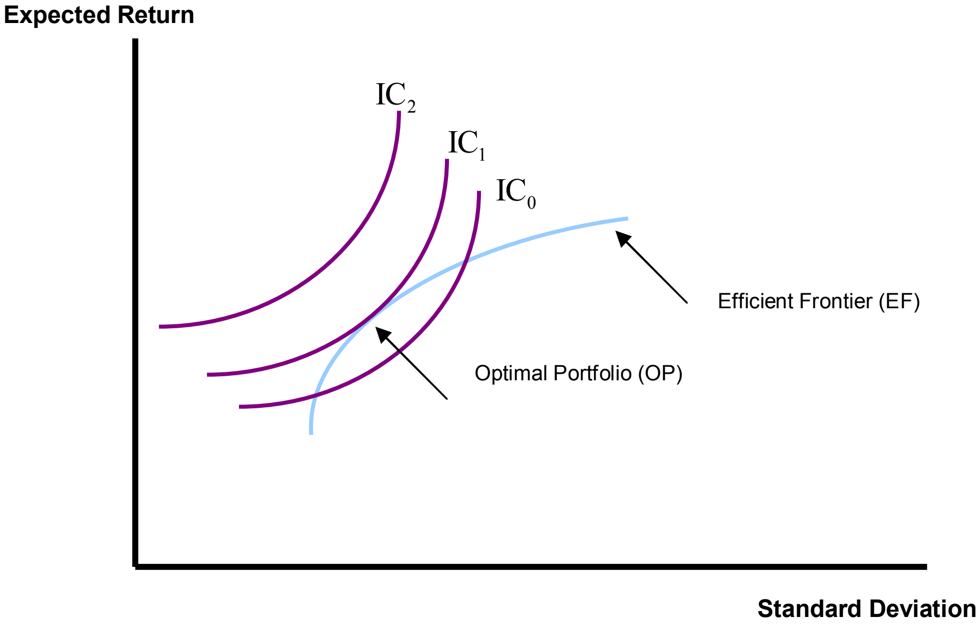

Figure 2.1-1: The Optimal Portfolio for an Investor- No Riskless Lending And Borrowing

The optimal portfolio for this particular investor is at the point of tangency between the indifference curve and the efficient frontier. Notice, the portfolio (OP) is the optimal for one particular individual with the specific form of indifference curves and not for all investors. Another investor who is very risk averse, namely the slope of his indifference curves is much steeper than in the previous case, will end up with a different optimal portfolio along the same efficient frontier; this is shown in the following figure.

https://doi.org/10.26219/heal.aueb.6306

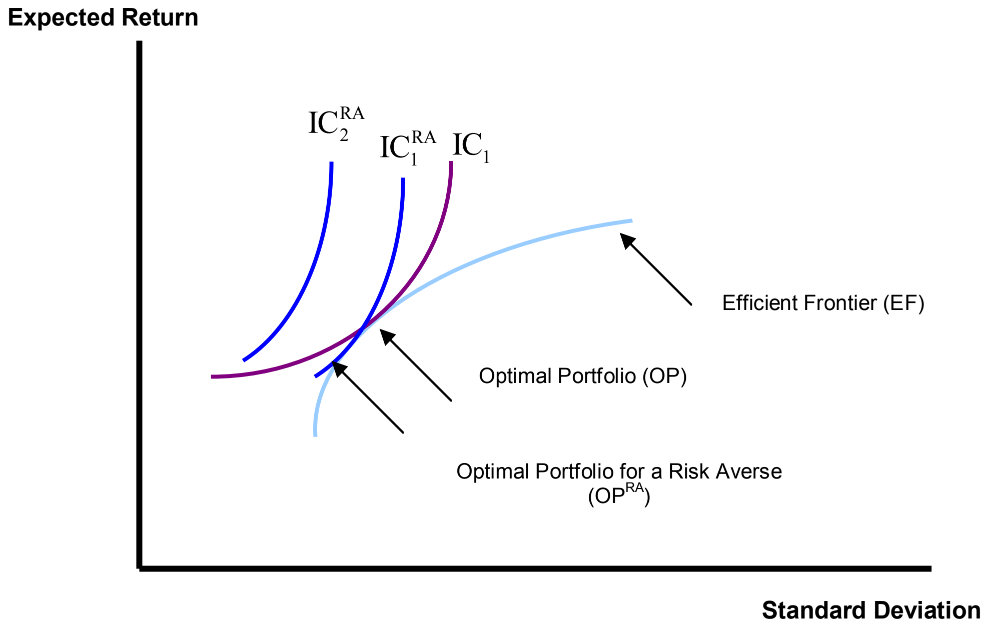

Figure 2.1-2: The Optimal Portfolio for a Risk Averse Investor- No Riskless Lending And Borrowing

Notice the chosen optimal portfolio OP^{RA} of the risk averse has smaller risk and expected return than the optimal portfolio of an investor with an average tolerance towards risk. Finally, the optimal portfolio for a risk lover will involve higher risk but also higher expected return.

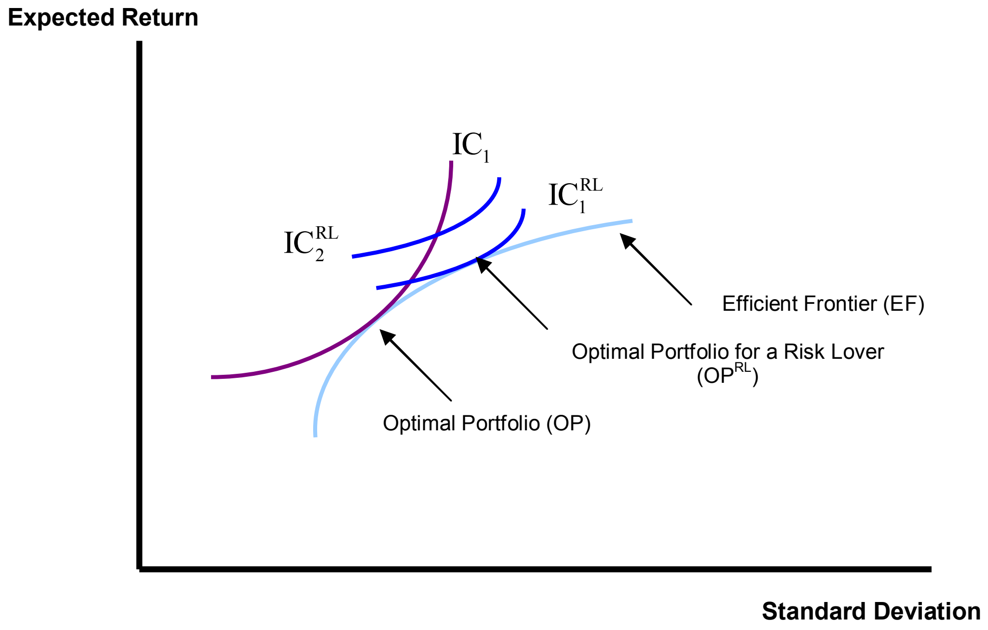

Figure 2.1-3: The Optimal Portfolio for a Risk Loving Investor - No Riskless Lending And Borrowing

https://doi.org/10.26219/heal.aueb.6306

2.2 OPTIMAL PORTFOLIOS BASED ON THE TOBIN APPROACH

In the previous section we argued that, on the one hand, there is no efficient portfolio that stands out, and, on the other hand, a number of efficient portfolios exist depending on the amount of risk an investor is willing to tolerate. Put differently, an investor can adjust his preferred level of risk by moving along the efficient frontier. However the interesting point is to see whether an optimal portfolio exists regardless of the individual investors' tolerance towards risk.

Tobin (1958) resolved this problem by showing that if investors could invest in a riskless asset (such as a government security) in addition to a risky stock portfolio, then the efficient frontier consists of combinations of the risk-free asset (such as a bank account) and an optimal portfolio of risky assets. That is, as we shall shortly present graphically, when riskless lending and borrowing is allowed every investor, irrespective of their level of risk tolerance, would like to hold a single a portfolio of risky assets (the tangency portfolio).

Hence, our objective in this section is twofold. First we shall show that, under the capital asset pricing theory, there is a single optimal portfolio of risky assets that is preferred to all other efficient portfolios of risky assets, and that risk is adjusted by borrowing or lending against a single optimal risky portfolio (Cohen et *al.*1982). Second, we shall present the way to compute this optimal portfolio of risky assets.

2.2.1 The Maximization Problem: A Graphical Approach

The efficient frontier developed in the previous section assumes that all securities (or portfolios) on the efficient set are risky. Alternatively, an investor could combine a risky investment with an investment in a riskless security, such as an investment in government debt. The introduction of a risk-free asset significantly alters the properties of the efficient frontier, and instead of a curve we obtain a straight line in the expected return standard deviation space (this is the same space with the Markowitz analysis).

Consider the situation where the investor can place a portion w_A of his funds in the risky portfolio A, and either lend or borrow at the risk free rate. In this case the expected return on his portfolio is given by the following equation

$$\overline{r}_p = w_A \overline{r}_A + \left(1 - w_A\right) r_F \qquad (2.12)$$

where

$\overline{r}_A$: the expected return on the risky asset A, and r_F is the return on the risk-free asset

The risk on his portfolio should accommodate the fact that the riskless asset has not risk. Since one asset has no risk it follows that the risk of the entire portfolio consists of the risk of the risky asset A, that is, it must be the case that $\sigma_p^2 = w_A^2 \sigma_A^2$, or, after taking the square root $\sigma_p = w_A \sigma_A$. Solving the last equation with respect to the portion of funds placed on the risky portfolio, we obtain the following

$$w_A = \frac{\sigma_p}{\sigma_A}$$

Substituting the above expression for w_A into the expected return formula for the portfolio, we can obtain the following relationship.

$$\overline{r}_P = \frac{\sigma_P}{\sigma_A}\overline{r}_A + \left(1 - \frac{\sigma_P}{\sigma_A}\right) r_F$$

A small rearrangement of terms yields the following formula

$$\overline{r}_P = r_F + \frac{\overline{r}_A - r_F}{\sigma_A}\sigma_P \qquad (2.13)$$

The above line describes the relation between the *expected return* and *risk* for one risky portfolio (i.e. portfolio A) and the risk-free asset. This equation, which is known as the **opportunity line,** is characterized by its slope $\left(\overline{r}_{\mathrm{A}} - r_F\right)/\sigma_{\mathrm{A}}$. Inevitably, there will be other risky portfolios, such as portfolio B, which can be combined with the riskless asset; this combination will result in a set of new opportunity line with a slope $\left(\overline{r}_{\mathrm{B}} - r_F\right)/\sigma_{\mathrm{B}}$. So, we would like to know which optimal opportunity line, created by different combinations of a risky assets (or portfolios) and the risk-free rate, has the highest slope.

The optimal opportunity line, known as the **capital market line** (Ross *et al*. 1999), is the opportunity line with the highest slope, and it can be found at the point of tangency between any opportunity line and the Markowitz efficient frontier, as shown in the following figure.

https://doi.org/10.26219/heal.aueb.6306

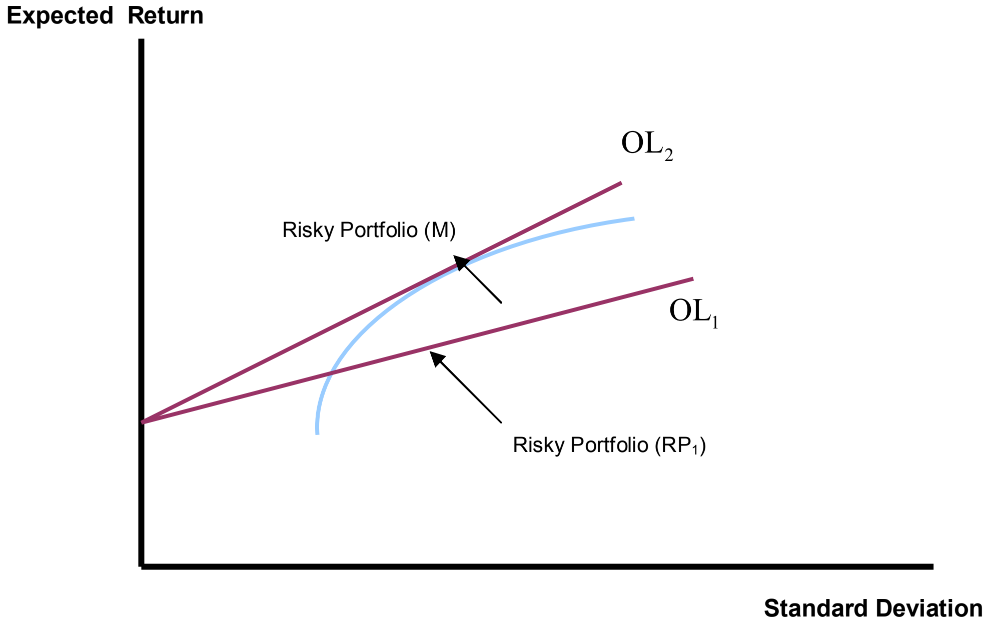

Figure 2.2-1: The Capital Market Line

The above point of tangency, between the opportunity line OL_2 - which, in this case, represents the capital market line - and the efficient frontier, determines the optimal risky portfolio M (i.e. the weights invested in each risky asset that comprises portfolio M), which all investors, regardless of their risk preferences, must hold. Notice, the slope of the capital market line ,i.e. $(\bar{r}_{\mathrm{M}} - r_F)/\sigma_{\mathrm{M}}$, is the highest of all slopes that could result by combining the riskless asset with a risky portfolio, situated either along the Markowitz efficient frontier or within it.

Next, each investor, depending on his degree of risk aversion, will decide how much of his funds will be invested in the optimal risky portfolio M and how much in the riskless asset (Figure 2.2-1).

https://doi.org/10.26219/heal.aueb.6306

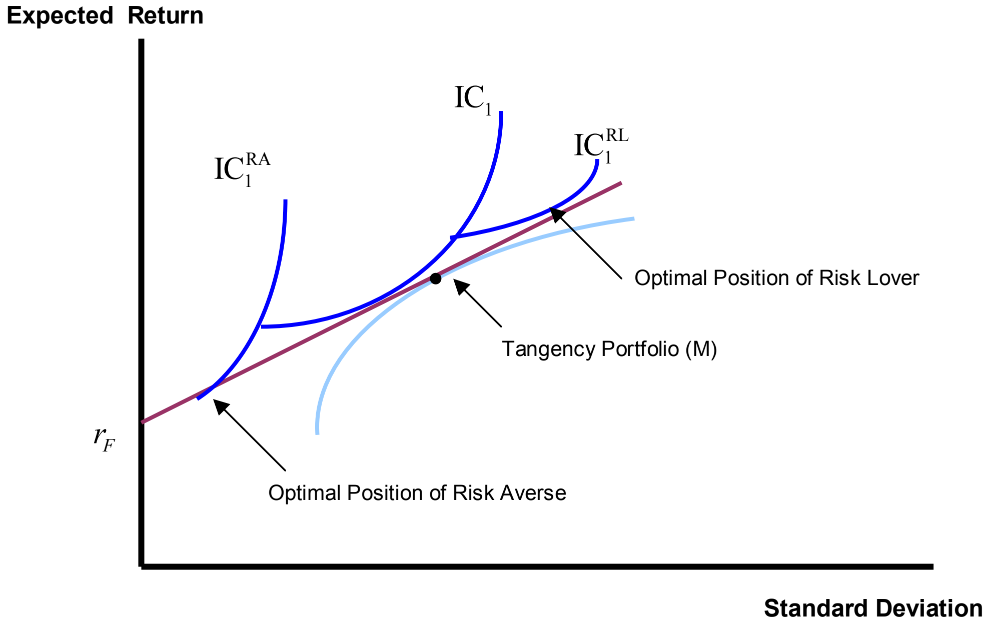

Figure 2.2-2: The Determination of The Optimal Positions When Riskless Lending And Borrowing Is Allowed

It is evident from the above figure that the more tolerant an investor is toward risk the less he will invest in the riskless asset. For example, a risk-averse investor, whose indifference curve is $\mathrm{IC}_1^{\mathrm{RA}}$, would end up at a point on the straight line between r_F and M. Alternatively, a risk-loving investor, whose indifference curve is $\mathrm{IC}_1^{\mathrm{RL}}$, would borrow at the risk-free rate r_F and place the loaned capital and its own funds at M.

2.2.2 The Maximization Problem: An Analytical Approach

Let us now see how we can derive analytically the composition of the tangency portfolio M. The details of the process can be found in Elton and Gruber (1995: 98-101). To begin with, in order to form the optimal portfolio we have to maximize the following objective function, or **theta coefficient** of the portfolio P

$$\theta = \frac{\overline{r}_p - r_F}{\sigma_p} \qquad (2.14)$$

Subject to the constraint $\sum_{i=1}^{N} w_i = 1$

where $\overline{r}_p = \sum_{i=1}^{N} w_i \overline{r}$ is the expected return on the portfolio, r_F is the risk-free rate, and $\sigma_p = \sqrt{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}}$ is the portfolio's standard deviation (risk).

https://doi.org/10.26219/heal.aueb.6306

Given that $r_F = 1 \times r_F$ then according to the above constraint we can write the objective function as follows

$$\begin{aligned}\theta &= \frac{\sum_{i=1}^{N} w_i \left(\overline{r}_i - r_F\right)}{\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}}} = \\ &= \sum_{i=1}^{N} w_i \left(\overline{r}_i - r_F\right)\left[\left(\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}\right)^{-1/2}\right]\end{aligned} \qquad (2.15)$$

The solution to the above maximization problem involves finding the solution to the following system of simultaneous equations

$$\begin{aligned}&\frac{\partial\theta}{\partial w_1} = 0\\ &\frac{\partial\theta}{\partial w_2} = 0\\ &\vdots\\ &\frac{\partial\theta}{\partial w_N} = 0\end{aligned} \qquad (2.16)$$

According to the product rule, the optimal weight for stock k, for $k = 1,\ldots,N$, must be as follows:

$$\begin{aligned}\frac{\partial\theta}{\partial w_k} = &-\frac{1}{2}\sum_{i=1}^{N} w_i \left(\overline{r}_i - r_F\right)\left[\left(\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}\right)^{-3/2}\right]\left(2 w_k \sigma_k^2 + 2\sum_{j=1}^{N} w_j \sigma_{kj}\right)\\ &+\left[\left(\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}\right)^{-1/2}\right]\left(\overline{r}_k - r_F\right)\end{aligned}$$

For $j \neq k$. Multiplying the derivative by $\left(\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}\right)^{1/2}$ we get

https://doi.org/10.26219/heal.aueb.6306

$$\left(\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}\right)^{1/2} \frac{\partial\theta}{\partial w_k} = -\frac{\sum_{i=1}^{N} w_i \left(\overline{r}_i - r_F\right)}{\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}} \left(w_k\sigma_k^2 + \sum_{j=1}^{N} w_j \sigma_{kj}\right) + \\ + \left(\overline{r}_k - r_F\right) = 0$$

Finally, defining by λ the following quantity

$$\lambda = \frac{\sum_{i=1}^{N} w_i \left(\overline{r}_i - r_F\right)}{\sum_{j=1}^{N}\sum_{i=1}^{N} w_i w_j \operatorname{cov}_{ij}}$$

We can then write

$$\frac{\partial\theta}{\partial w_k} = -\lambda\left(w_k\sigma_k^2 + \sum_{j=1}^{N} w_j \sigma_{kj}\right) + \left(\overline{r}_k - r_F\right) = 0$$

Suppose for example that we are considering three assets; then we can write the following

$$\frac{\partial\theta}{\partial w_1} = -\lambda\left(w_1\sigma_1^2 + \sum_{j=1}^{3} w_j \sigma_{1j}\right) + \left(\overline{r}_1 - r_F\right) = 0$$

$$\frac{\partial\theta}{\partial w_2} = -\lambda\left(w_2\sigma_2^2 + \sum_{j=1}^{3} w_j \sigma_{2j}\right) + \left(\overline{r}_2 - r_F\right) = 0$$

$$\frac{\partial\theta}{\partial w_3} = -\lambda\left(w_3\sigma_3^2 + \sum_{j=1}^{3} w_j \sigma_{3j}\right) + \left(\overline{r}_3 - r_F\right) = 0$$

Or in terms of a system

$$\begin{aligned}
\left(\overline{r}_1 - r_F\right) &= \lambda w_1\sigma_1^2 + \lambda w_2\sigma_{12} + \lambda w_3\sigma_{13} \\
\left(\overline{r}_2 - r_F\right) &= \lambda w_2\sigma_2^2 + \lambda w_1\sigma_{21} + \lambda w_3\sigma_{23} \\
\left(\overline{r}_3 - r_F\right) &= \lambda w_3\sigma_3^2 + \lambda w_1\sigma_{31} + \lambda w_2\sigma_{32}
\end{aligned}$$

Defining a new variable $Z_k = \lambda w_k$, for $k = 1,\ldots,3$, we can write the above as follows

$$\begin{aligned}
\left(\overline{r}_1 - r_F\right) &= Z_1\sigma_1^2 + Z_2\sigma_{12} + Z_3\sigma_{13} \\
\left(\overline{r}_2 - r_F\right) &= Z_1\sigma_{21} + Z_2\sigma_2^2 + Z_3\sigma_{23} \\
\left(\overline{r}_3 - r_F\right) &= Z_1\sigma_{31} + Z_2\sigma_{32} + Z_3\sigma_3^2
\end{aligned}$$

The above system of equations can be written in matrix notation as follows

$$\overline{\mathbf{r}} - \mathbf{i}r_F = \mathbf{\Sigma}\mathbf{Z} \tag{2.17}$$

where $\mathbf{r}$ is a 3×1 column-vector of expected returns , $\mathbf{i}$ is a 3×1 column-vector of 1s , $\mathbf{Z}$ is a 3×1 column-vector of Z s, and $\mathbf{\Sigma}$ = a 3×3 symmetric variance-covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$$

Solving the system (2.16) for $\mathbf{Z}$ we get

$$\mathbf{Z}^* = \mathbf{\Sigma}^{-1}\left(\overline{\mathbf{r}} - \mathbf{i}r_F\right) \tag{2.18}$$

Finally, in order to find the optimal weights we calculate the following

$$w_i = Z_i / \mathbf{Z}^{*\prime}\mathbf{i} \tag{2.19}$$

If short-sales are not allowed then we need to add the following additional constraint $w_i \geq 0$ in our calculations (Elton and Gruber, 1995 p.104).

https://doi.org/10.26219/heal.aueb.6306

2.3 ARCH MODELS OF STOCK VOLATILITY

The ARCH models constitute an effort to forecast the future variance of a random variable. Before the introduction of these models, the primary descriptive tool was the **rolling standard deviation**, namely the standard deviation was calculated using a fixed number of the most recent observations (Engle, 2001). For example, the risk of a stock, i.e. the standard deviation of its returns, could be calculated every day using the most recent month (22 business days) of data. This model of volatility forecasting, which assumes that the variance of tomorrow's return is an equally weighted average of the squared residuals from the last 22 days, is the first formulation of the ARCH model (Engle, 2001).

The assumption of equal weights seems unattractive, as one would think that the more recent events would be more relevant and therefore should have higher weights. Furthermore the assumption of zero weights for observations more than one month old is also unattractive. The ARCH model proposed by Engle (1982) let these weights be parameters to be estimated. Thus, the model allowed the data to determine the best weights to use in forecasting the variance.

2.3.1 The Intuition behind the ARCH Models

To illustrate the presence of ARCH effects in financial data, we make use of the **Athens Composite Share Price Index** (or more commonly called the **General Index**), which represents the bulk of the value in the Greek equity market.

In the following figure we look at the weekly levels and returns of General Index from 2002 through early February 2007.

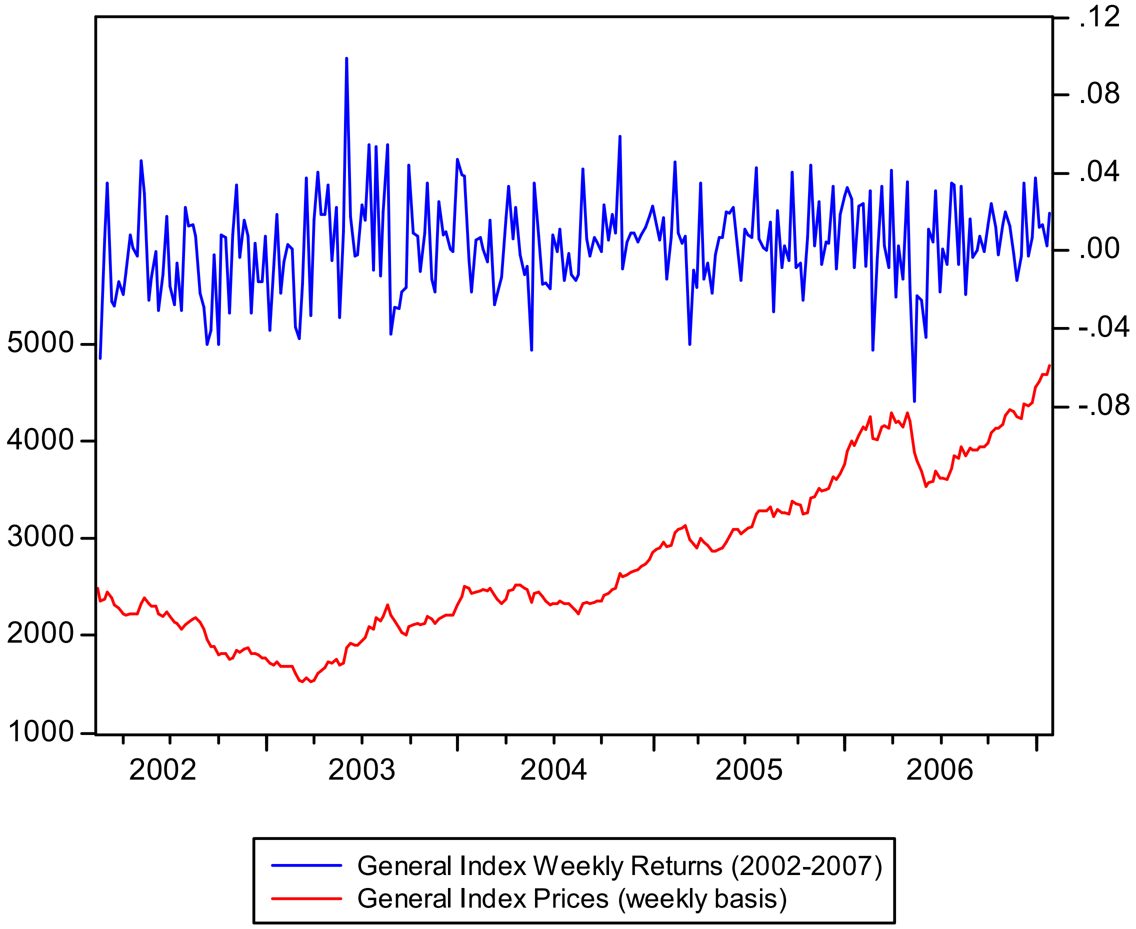

Figure 2.3-1: General Index Weekly Prices and Returns, 2002:02 -2007:02

The raw data (i.e. the index's weakly closing prices) are shown on the left axis. At the beginning of 2002 the index was priced at 2,488.78 and at beginning of February 2007 at 4,779.71. Based on the average monthly growth rate of the index, we conclude that for every euro invested in the index, in 2002, the investor would receive $1.92 by February 2007 (plus the stream of dividends that would have been received by this index). The returns on the General Index, computed as $(P_t / P_{t-1}) - 1$, are shown at the top of Figure 2.3-1. This return series is centered around zero throughout the sample period even though prices are sometimes decreasing (for example from 2002 until early 2003) and sometimes increases (like from early 2003 onward). Further, we observe that the most dramatic weakly return decline in the General Index, which dwarfs all other returns in the size of the decline, occurred at the end of the first quarter of 2006.

https://doi.org/10.26219/heal.aueb.6306

Next, we decompose the volatility of the above return series by looking at portions of the whole history. For example, Figures 2.3-2 and 2.3-3 show the same graph (i.e. the levels and the returns of the General Index) before December 21st 2003 (this is a period of declining prices for the General Index) and after December 21st 2003 (this is a period of rising prices for the General Index)

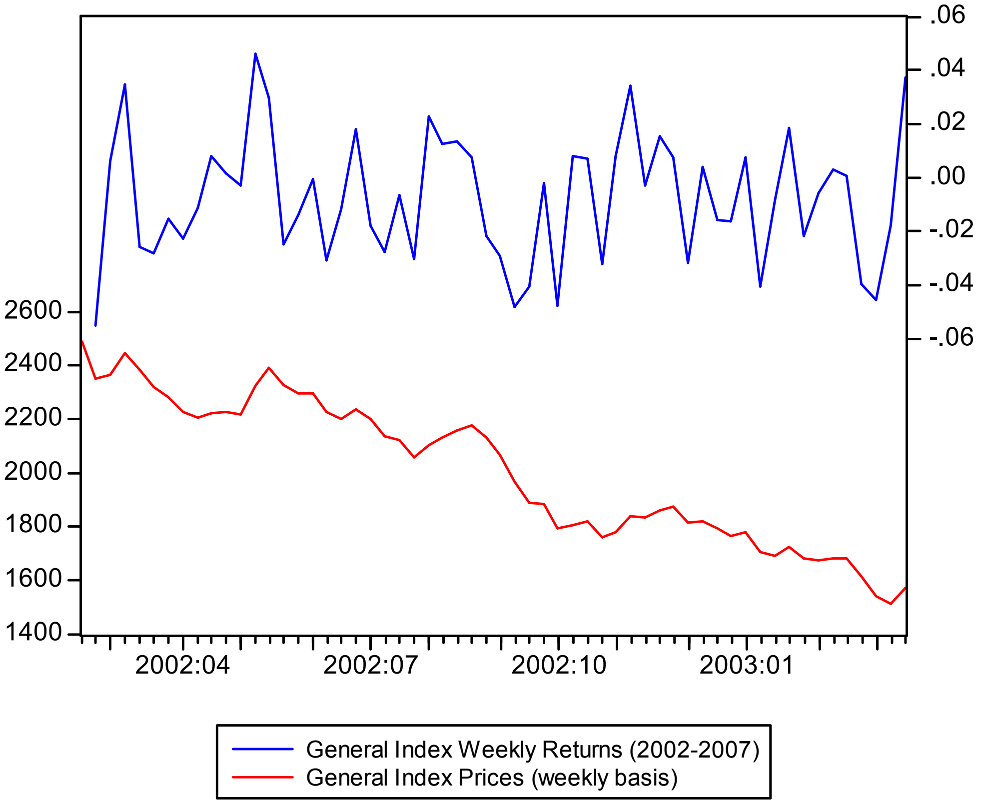

Figure 2.3-2: General Index Weekly Prices and Returns, 2002:02 -2003:03

https://doi.org/10.26219/heal.aueb.6306

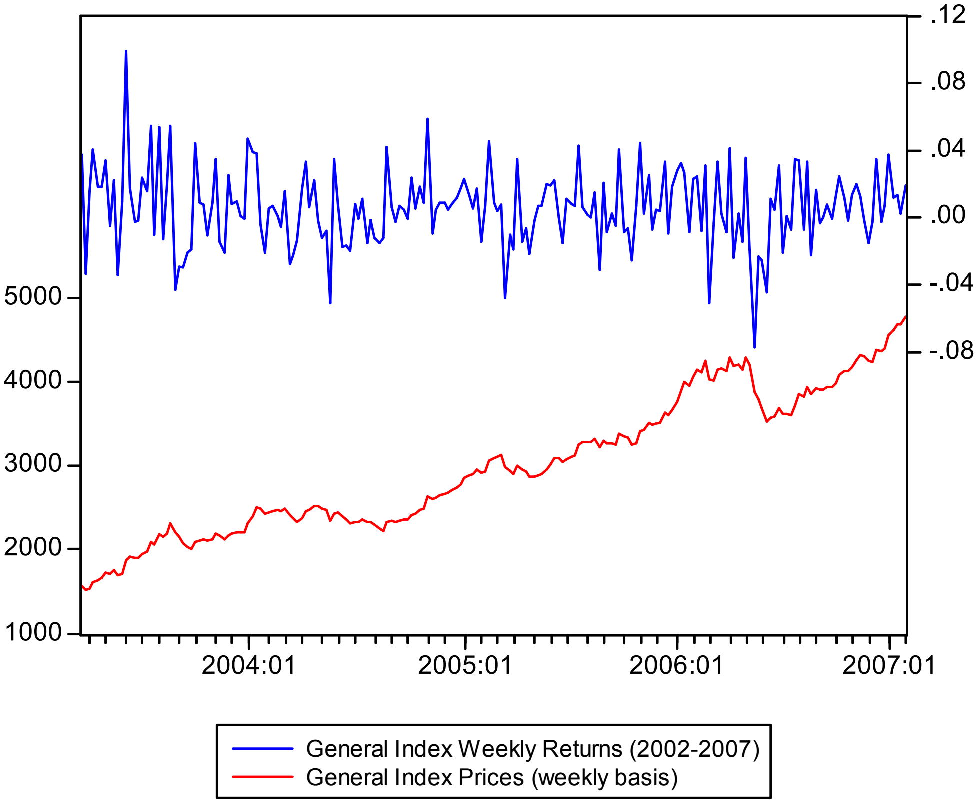

Figure 2.3-3: General Index Weekly Prices and Returns, 2003:03 -2007:02

Looking at the above two figures, it is obvious that the magnitude of the changes is sometimes large and sometimes small, that is, the volatility of the index's returns is changing depending on the historical period considered; this is the effect that ARCH is designed to measure, and it is known as volatility clustering. Specifically, we observe that volatility is higher when prices are falling (historical period 2002:02- 2003:03) and lower when prices are rising. In other words, volatility tends to be higher in bear markets and lower in bull markets.

In order to be more accurate concerning the aforementioned findings, we consider some descriptive statistics about the General Index (Table 2.3-1).

https://doi.org/10.26219/heal.aueb.6306

Table 2.3-1: Illustration of The Presence Of Arch Effects In The Greek Equity Market (2002-2007)

	General Index Returns		**General Index Price**	
	Bearish Market 2002:02 – 2003:02	**Bullish Market 2003:02 – 2007:02**	**Bearish Market 2002:02 – 2003:02**	**Bullish Market 2003:02 – 2007:02**
Mean	-0.0077	0.0059	2,988.75	2,007.62
Median	-0.0064	0.0060	2,904.14	2,062.89
Maximum	0.0461	0.0990	4,779.71	2,488.78
Minimum	-0.0550	-0.0777	1,517.94	1,513.88
Standard Deviation	**0.0236**	**0.0229**	819.05	269.534

The mean return is close to zero relative to the standard deviation for both periods. Specifically, it is -0.77% per week or about -33.10% per year for the brearish period and 0.59% per week or about 35.78% per year for the bullish period (given a compounding considering 52 trading weeks during the year). However, the (weakly) standard deviation of returns is higher, slightly though, in the bearish period. Specifically, in the bearish period its 2.36% per week or about 17.02% per year for the bearish period, and 2.29% per week or about 16.51% per year for the bullish period.

2.3.2 The ARCH (1) Model

To illustrate the ARCH(1) model, we shall consider two specifications for the **mean equation**, which is the econometric model from which we can derive the conditional mean of the process y_t. ARCH processes are used to model the conditional variance of the process y_t.

Consider first the simple AR(1) model as the mean equation for the time-series process y_t.

$$y_t = \mu_t + \varepsilon_t \tag{2.20}$$

where

$\mu_t = \phi_0 + \phi_1 y_{t-1}$

$\{\varepsilon_t\}$ is a sequence of iid white noise variables; this implies that $E(\varepsilon_t) = 0$ and $\operatorname{var}(\varepsilon_t) = \sigma^2$ for $t = 1, \ldots, T$.

https://doi.org/10.26219/heal.aueb.6306

To find the conditional variance $\operatorname{var}\left(y_t | \Phi_{t-1}\right)$ of the AR(1) model, we have to hold y_{t-1} constant in which case the only source of variation at y_t is ε_t. We can show this as follows:

$$\begin{aligned} \operatorname{var}\left(y_t | \Phi_{t-1}\right) &= \operatorname{var}\left(\phi_0 + \phi_1 y_{t-1} + \varepsilon_t | \Phi_{t-1}\right) = \\ &= \operatorname{var}\left(\phi_0 | \Phi_{t-1}\right) + \operatorname{var}\left(\phi_1 y_{t-1} | \Phi_{t-1}\right) + \operatorname{var}\left(\varepsilon_t | \Phi_{t-1}\right) = \\ &= \operatorname{var}\left(\varepsilon_t | \Phi_{t-1}\right) = \operatorname{var}\left(\varepsilon_t\right) \end{aligned}$$

So the conditional variance $\operatorname{var}\left(y_t | \Phi_{t-1}\right) = \sigma^2_{y_t|\Phi_{t-1}}$ is constant since it is equal to $\operatorname{var}\left(\varepsilon_t\right)$,which, in turn, is equal to σ^2 for $t = 1,\ldots,T$. If however the conditional variance $\operatorname{var}\left(y_t | \Phi_{t-1}\right) = \sigma^2_{y_t|\Phi_{t-1}}$ is not constant then we have an ARCH process. For example, it could be the case that

$$\sigma^2_{y_t|\Phi_{t-1}} = \sigma^2_{\varepsilon_t|\Phi_{t-1}} = \alpha_0 + \alpha_1 \varepsilon^2_{t-1} \tag{2.21}$$

An alternative specification of an $ARCH\left(p\right)$ process has as follows

$$\varepsilon_t = u_t \sigma_t \tag{2.22}$$

The conditional variance in the above scheme is

$$\begin{aligned} \operatorname{var}\left(\varepsilon_t | \Phi_{t-1}\right) &= \operatorname{var}\left(u_t \sigma_t | \Phi_{t-1}\right) = \\ &= \sigma_t^2 \operatorname{var}\left(u_t | \Phi_{t-1}\right) = \\ &= \sigma_t^2 \operatorname{var}\left(u_t\right) = \sigma_t^2 \end{aligned}$$

where $\sigma^2_{y_t|y_{t-1}} = \alpha_0 + \alpha_1 y^2_{t-1}$ and $\operatorname{var}\left(u_t\right) = 1$

Of course, we can extend the analysis to the classical regression, so that the ARCH process applies to the disturbance term. In that case, the ARCH (1) model specifies that the conditional variance of the disturbance term at time *t* depends on the squared error term at time $t-1$.

Consider next the k-variable regression model as the *mean equation*[1]

$$y_t = \boldsymbol{\beta}' \mathbf{x}_t + \varepsilon_t \tag{2.23}$$

where

$\mathbf{x}_t$: a $\left(K+1\right) \times 1$ column vector of K independent variables plus the value 1

$\boldsymbol{\beta}$: a $\left(K+1\right) \times 1$ column vector of K population parameters plus the constant term

---

[1] In the financial the mean equation is the single-index model (see Section 3.4)

https://doi.org/10.26219/heal.aueb.6306

If the above regression model contains ARCH errors, then the disturbance term ε_t of the above k-variable regression model must adhere to the following process

$$\varepsilon_t = u_t\left[\alpha_0 + \alpha_1\varepsilon_{t-1}^2\right]^{1/2} \tag{2.24}$$

where

$\{u_t\}$ is a sequence of iid random variables with zero mean and unit variance; this implies that $E(u_t)=0$ and $\operatorname{var}(u_t)=1$ for $t=1,\dots,T$

The obvious generalization for p lags has as follows

$$\varepsilon_t = u_t\left[\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \dots + \alpha_p\varepsilon_{t-p}^2\right]^{1/2} \tag{2.25}$$

In order to estimate the regression model described by equations (2.23) and (2.24), we need to estimate both $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}' = [\alpha_0 \quad \alpha_1]$. It can be easily shown that the conditional mean of y_t with respect to $\mathbf{x}_t$ is $E(y_t|\mathbf{x}_t)=\boldsymbol{\beta}'\mathbf{x}_t$. The conditional variance of the disturbance term ε_t, in equation (2.23), with respect to ε_{t-1} depends on the information set Φ_{t-1} available at time $t-1$.

$$\begin{aligned}
\operatorname{var}\left(\varepsilon_t|\Phi_{t-1t-1}\right) &= \operatorname{var}\left[u_t\left(\alpha_0+\alpha_1\varepsilon_{t-1}^2\right)^{1/2}\middle|\Phi_{t-1t-1}\right] \\
&= \left(\alpha_0+\alpha_1\varepsilon_{t-1}^2\right)\operatorname{var}\left(u_t|\Phi_{t-1t-1}\right) = \\
&= \left(\alpha_0+\alpha_1\varepsilon_{t-1}^2\right)\operatorname{var}\left(u_t\right) = \\
&= \alpha_0+\alpha_1\varepsilon_{t-1}^2
\end{aligned}$$

And in the case of p lags the conditional variance becomes

$$\operatorname{var}\left(\varepsilon_t|\Phi_{t-1}\right) = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \dots + \alpha_p\varepsilon_{t-p}^2 \tag{2.26}$$

Equations (2.23) and (2.24) define the ARCH(1) model, according to which the conditional variance $\operatorname{var}(\varepsilon_t|\Phi_{t-1})=\sigma_t^2$ of the disturbance term ε_t is a function of the previous' period variability, which is expressed by the square of ε_{t-1} (squared innovations). In other words, the ARCH model parameterizes the conditional variance of a random variable in a linear fashion.

To find the unconditional variance of ε_t in equation (2.23) we apply the variance decomposition rule on the conditional variance $\operatorname{var}\left(\varepsilon_t \middle| \Phi_{t-1}\right) = \sigma_t^2$

$$\operatorname{var}\left(\varepsilon_t\right) = \operatorname{var}_{t-1}\left[E\left(\varepsilon_t \middle| \varepsilon_{t-1}\right)\right] + E_{t-1}\left[\operatorname{var}\left(\varepsilon_t \middle| \varepsilon_{t-1}\right)\right] \tag{2.27}$$

The conditional expectation $E\left(\varepsilon_t \middle| \varepsilon_{t-1}\right)$ can be found as follows

$$E\left(\varepsilon_t \middle| \varepsilon_{t-1}\right) = \left[\alpha_0 + \alpha_1 \varepsilon_{t-1}^2\right]^{1/2} E\left(u_t \middle| \varepsilon_{t-1}\right) = 0 \tag{2.28}$$

Thus, we have

$$\begin{aligned} \operatorname{var}\left(\varepsilon_t\right) &= E_{t-1}\left[\operatorname{var}\left(\varepsilon_t \middle| \varepsilon_{t-1}\right)\right] = \\ &= E\left(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2\right) = \alpha_0 + \alpha_1 E\left(\varepsilon_{t-1}^2\right) \end{aligned} \tag{2.29}$$

If the process is weakly stationary, i.e. $\operatorname{var}\left(\varepsilon_t\right) = \operatorname{var}\left(\varepsilon_{t-1}\right) = \sigma^2$ then the unconditional variance of u_t becomes

$$\operatorname{var}\left(\varepsilon_t\right) = \frac{\alpha_0}{1-\alpha_1} \tag{2.30}$$

Thus, unconditionally, the disturbance term ε_t in the regression model (2.23)is distributed normally with mean zero and variance $\alpha_0 / \left(1-\alpha_1\right)$, which is constant and does not depend on time

2.3.3 Testing for ARCH

To test for the ARCH (p) process, in the disturbance term of the classical regression model, we test the following null hypothesis

$$H_0 : \alpha_1 = \ldots = \alpha_p$$

Johnston and Dinardo (1997*,* p.196) suggested the following testing procedure for the above null hypothesis.

1. Given the mean equation $y_t = \boldsymbol{\beta}'\mathbf{x}_t + u_t$ we obtain, using observations, $\left\{y_t, \mathbf{x}_t\right\}$ for $t = 1, \ldots, T$, the residuals $\left\{e_t = y_t - \hat{\boldsymbol{\beta}}'\mathbf{x}_t\right\}$ of it. Notice $\hat{\boldsymbol{\beta}}'$ has been estimated ignoring any ARCH effects and as such it is consistent and asymptotically normally distributed, but inefficient (Grene 2000, p.798).

https://doi.org/10.26219/heal.aueb.6306

2. Using these estimated residuals, and in particular the observations $\{e_t\}$ for $t = 2,\dots,T$, we regress e_t^2 on a constant and a number of explanatory variables $\{e_{t-j}^2\}_{j=1}^{p}$, as the following econometric model shows

$$e_t^2 = b_0 + b_1 e_{t-1}^2 + \dots + b_p e_{t-p}^2 + \text{error} \tag{2.31}$$

3. The OLS regression of the above model will give us the vector of estimated parameters $\hat{\mathbf{b}} = \left[\hat{b}_0, \hat{b}_1, \dots, \hat{b}_p\right]'$. Under the null hypothesis of conditional homoscedasticity, the *Lagrange Multiplier* $((T-1)R^2)^2$ statistic has a limiting distribution in chi-square with p degrees of freedom. If the test statistic is greater than the critical value, we shall reject the null hypothesis in favor of the alternative hypothesis of conditional heteroscedasticity.

2.3.4 The ARCH in Mean Model

Up to this point, we have assumed that the expected return $\bar{r}$ is simply a constant. However, according to the asset pricing models expected return should be positively related to risk, since higher risk needs to be compensated for by higher expected returns. Indeed, according to modern finance theory volatility may be related to risk premia on assets. One way to model this trade-off between expected return and risk (conditional volatility) is to make the conditional mean depend on the conditional variance, the latter being a measure of how risky an asset is.

This idea of the ARCH in Mean model, which was introduced by Engle *et al* (1987), is mathematically expressed by incorporating the conditional variance into an equation for the conditional mean. The net return to asset i can be modelled as the sum of its expected risk premium μ_t plus an asset specific shock ε_t

$$r_{it} = \mu_t + \varepsilon_t \tag{2.32}$$

---

[2] T stands for the number of included observations, and R^2 for the unadjusted R-squared of the regression model (2.30)

According to Sharpe (1964) the expected risk premium on a stock is related to the expected risk premium on the market and the stock's beta coefficient (a measure of the stock's volatility) as follows

$$\mu_i = \overline{r}_i - r_F = \beta_i\left(\overline{r}_m - r_F\right) \qquad (2.33)$$

For $i = 1,\ldots,N$ assets ; $\overline{r}_i$ is the expected return on asset i, r_F the risk-free rate, r_m the rate of return on the market index (it is taken to be a random variable), and β_i the stock's beta coefficient; this measures the expected change in the security return given a change in the market return.

In equation (2.32) the stock's expected risk premium is related to its beta coefficient, while in the ARCH in Mean model the stock's risk premium is related to its conditional variance (σ_i^2)

$$\mu_{it} = a + b_i\sigma_i^2 \qquad (2.34)$$

Combining we obtain the mean equation

$$r_{it} = a + b_i\sigma_i^2 + \varepsilon_t \qquad (2.35)$$

Notice, the firm's specific shock s_t follows an ARCH (or a GARCH) variant. Also, we see in equation (2.35) that an increase in risk, given by the conditional variance, leads to a rise in the mean return ($\mu_i = a + b_i\sigma_i^2$). The value of b gives the increase in returns needed to compensate for a give increase in risk; so it is a measure of risk aversion.

2.4 GARCH MODELS

In empirical applications it is often difficult to estimate models with large number of parameters in the ARCH models. To circumvent this problem Bollerslev (1986) proposed the Generalized ARCH model.

2.4.1 Univariate GARCH Models

To estimate a $GARCH(p,q)$-model, we have to consider two distinct specifications, one for the conditional mean and one for the conditional variance. The $GARCH(p,q)$ model, for the mean equation (2.36), expresses the conditional variance of the excess return (defined from the mean equation) as a linear function of p squared disturbances and q lagged conditional variances as follows (Bollerslev, 1986)

$$y_t = \mu + \varepsilon_t \tag{2.36}$$

$$\operatorname{var}\left(\varepsilon_t \middle| \Phi_t\right) = \sigma_t^2 = \omega + \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 \tag{2.37}$$

where

ε_t : the excess return at time period t

Φ_t denotes the information set available at time t; this information set is assumed to include all values of ε_t up to and including ε_{t-1}

In the above model we have to notice the following. First, we see that the conditional variance $\operatorname{var}\left(\varepsilon_t \middle| \Phi_t\right) = \sigma_t^2$ depends on the squared values ε_{t-j}^2 in the previous p periods, and the conditional variance σ_{t-i}^2 in the previous q periods. Second, although financial markets may experience excessive volatility from time to time, it appears that volatility will eventually settle down to a long run level; this is know as **volatility mean reversion**. Put technically, the unconditional variance, or **long-term variance,** V **,**can be derived from the $GARCH(p,q)$ as follows (Bollerslev, 1986)

$$\operatorname{var}\left(\varepsilon_t\right) = V = \frac{\omega}{1 - \left(\sum_{j=1}^{p} \alpha_j + \sum_{i=1}^{q} \beta_i\right)} \tag{2.38}$$

Based on the above unconditional variance we conclude that a $GARCH\left(p,q\right)$ process is correctly estimated when (a) the mean of the process is non-zero, i.e. $\alpha_0 > 0$ and (b) the stationarity conditions are met, that is, $\sum_{j=1}^{p} \alpha_j + \sum_{i=1}^{q} \gamma_i < 1$.

In practice, the most frequent model to fit financial time series is the $GARCH(1,1)$ model (Johnston and Dinardo 1997, p.197), which has the following form

$$\operatorname{var}\left(\varepsilon_t \middle| \Phi_t\right) = \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{2.39}$$

https://doi.org/10.26219/heal.aueb.6306

The conditional variance equation specified in (2.39) is a function of three terms:

- The constant: ω
- News about volatility from the previous period, measured as the lag of the squared residual from the mean equation: ε_{t-1}^2 (the ARCH term).
- Last period's forecast variance: σ_{t-1}^2 (the GARCH term).

Usually the GARCH coefficient β_1 in the $GARCH(1,1)$ model is found to be around 0.9 for many daily (or weakly) financial time series. Thus, given a large value of β_1, it is obvious that large values of the conditional variance in the previous period, σ_{t-1}^2 will be followed by large values of the conditional variance this period, σ_t^2

By repeated substitution in equation (2.39) we derive

$$\operatorname{var}\left(\varepsilon_t \middle| \Phi_t\right) = \frac{\omega}{1-\beta_1} + \alpha_1\left(\varepsilon_{t-1}^2 + \beta_1\varepsilon_{t-2}^2 + \beta_1^2\varepsilon_{t-3}^2 + \ldots\right) \tag{2.40}$$

That is, the current variance is a function of the square of the disturbance terms in all previous periods. Also notice that if $0 < \beta_1 < 1$, the coefficient (or weights) of the lagged values of ε_t^2 decline geometrically.

Now suppose that the mean equation is described by the **single-index model,** as follows (Elton and Gruber 1995, pp.130-131)

$$r_{it} = a_{it} + \beta_i r_{mt} + \varepsilon_{it} \tag{2.41}$$

for $i = 1, \ldots, N$ assets over $t = 1, 2, \ldots, T$ time periods, where a_i is the return on asset i that is independent of the market performance, r_m is the rate of return on the market index (it is take to be a random variable), β_i is a constant that measures the expected change in the security return given a change in the market return., and ε_{it} is a random variable with an expected value of zero. This random variable accounts for the unexpected returns in asset i.

A GARCH(1,1) model for the disturbance term of the above mean relation is given by the following equation.

$$\operatorname{var}\left(\varepsilon_t \middle| \Phi_t\right) = \sigma_t^2 = \alpha_0 + \alpha_1\left(r_{it} - a_{it} - \beta_i r_{mt}\right)^2 + \gamma_1\sigma_{t-1}^2 \tag{2.42}$$

https://doi.org/10.26219/heal.aueb.6306

According to equation (2.42), interpreted in a financial context, an investor predicts this period's variance (volatility) by forming a weighted average of a long term average (α_0), the forecasted variance from last period (the GARCH term, σ_{t-1}^2), and information about volatility observed in the previous period (the ARCH term,). If the asset return r_i was unexpectedly large in either the upward or the downward direction (i.e large disturbance term ε_t), then the investor will increase the estimate of the variance for the next period. This model is also consistent with the volatility clustering, where large changes in returns are likely to be followed by further large changes.

2.4.2 Multivariate GARCH Models

It is often of interest to model the joint evolution of two or more series of returns, or two or more asset prices. If interest centers on the first moment of returns, and we are willing to assume homoskedasticity, then it would be natural to use a VAR. However, for many assets, it is generally the second moment that is primarily of interest, not the first. We need to model the evolution of volatility for each of the series, allowing the volatilities to be correlated across series.

Suppose that we observe the following returns vector $\mathbf{y}_t$

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \tag{2.43}$$

$$\operatorname{var}\left(\mathbf{y}_t \middle| \Phi_t\right) = \operatorname{var}\left(\boldsymbol{\varepsilon}_t \middle| \Phi_t\right) = \mathbf{H}_t \tag{2.44}$$

Where

$\mathbf{y}_t' = \left(y_{1t,} y_{2t,} \ldots, y_{Mt,}\right)$

$\boldsymbol{\mu}$: an $M \times 1$ vector of sample means

$\mathbf{H}_t$: the time-varying conditional variance-covariance matrix

$\boldsymbol{\varepsilon}_t$: a vector of random variables (in the financial context it represents the excess returns)

Φ_t denotes the information set available at time t; this information set is assumed to include all values of ε_t up to and including ε_{t-1}

Each random variable in vector $\mathbf{y}_t$ has a conditional mean equation $E\left(\mathbf{y}_t|\Phi_t\right)$ and for all of them we have a conditional variance-covariance matrix $\mathbf{H}_t$

$$\operatorname{var}\left(\mathbf{y}_t|\Phi_t\right)=\begin{bmatrix} h_{11,t} & h_{12,t} & \cdots & h_{1M,t} \\ h_{21,t} & h_{22,t} & & h_{2M,t} \\ \vdots & \vdots & \vdots & \vdots \\ h_{M1,t} & h_{M2,t} & \cdots & h_{MM,t} \end{bmatrix}=\mathbf{H}_t \tag{2.45}$$

where h_{ii} is the time-varying conditional variance of variable i, for $i=1,\ldots,M$ and h_{ij} is the time-varying conditional covariance between variable i and j.

For example, the vector $\mathbf{y}_t$ of random variables contains the random variables (y_{1t}, y_{2t}), observed for two assets. In that case

$$\mathbf{H}_t=\begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$$

Full multivariate generalization of a GARCH model involves the specification of a system of dynamic equations for the elements of a conditional varianceco-variance matrix $\mathbf{H}_t$ subject to positive definiteness constraints (Engle and Kroner, 1995). Fully general models model involve many parameters even when the number of variables that are modelled jointly is only moderately large, and the computational difficulties and uncertainty caused by estimating too many parameters often outweighs the benefits of multivariate modelling.

This has led researchers to consider various restricted versions of the general model, such as the leading multivariate models are the Diagonal VEC model (Bollerslev *et al.*, 1988), the Constant Correlation Model (Bollerslev, 1990), and the Dynamic Conditional Correlations (Engle and Sheppard, 2001).

The VEC Specification and the Diagonal Model

Bollerslev *et al.*(1988) have proposed the **VEC specification** of conditional variance-covariance matrix $\mathbf{H}_t$, according to which each element of the conditional variance-covariance matrix $\mathbf{H}_t$ is a function of every element of the lagged conditional variance-covariance matrices and outer products of lagged realizations.

For a $\mathrm{VEC}(p,q)$ we have the following model.

$$\mathbf{h}_t = \boldsymbol{\mu} + \sum_{i=i}^{p} \mathbf{A}_i \mathbf{h}_{t-i} + \sum_{j=1}^{q} \mathbf{B}_j \boldsymbol{\varepsilon}_{t-j} \tag{2.46}$$

where $\boldsymbol{\mu}$ is a $\frac{M(M+1)}{2} \times 1$ vector of constants, $\mathbf{h}_t$ stands for $vech(\mathbf{H}_t)$ which is a $\frac{M(M+1)}{2} \times 1$ column-vector of variances and covariances. The $vech(\cdot)$ operator stacks the lower triangular part of an $M \times M$ matrix into an $\frac{M(M+1)}{2} \times 1$ vector . Finally, $\boldsymbol{\varepsilon}_t$ is $vech(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$, which is a $\frac{M(M+1)}{2} \times 1$ column-vector of disturbances

$\mathbf{B}, \mathbf{A}$ are coefficient matrices of dimension $\frac{M(M+1)}{2} \times \frac{M(M+1)}{2}$,

The intuition of the VEC approach has as follows. We can define the sample time-varying variance-covariance matrix $\boldsymbol{\Sigma}$ as follows

$$\boldsymbol{\Sigma}_T = \frac{1}{T-1} \sum_{t=1}^{T} (\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})' \tag{2.47}$$

For example, suppose that the random vector $\mathbf{y}_t$ consists of random variables $(y_{1t}, y_{2t})'$ observed over $t = 1, \ldots, 3$ periods; the time-varying variance-covariance matrix $\boldsymbol{\Sigma}$ becomes as follows

$$\begin{aligned}\boldsymbol{\Sigma}_3 &= \frac{1}{3-1} \sum_{t=1}^{3} (\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})' = \\ &= \frac{1}{2}(\mathbf{y}_1 - \boldsymbol{\mu})(\mathbf{y}_1 - \boldsymbol{\mu})' + \frac{1}{2}(\mathbf{y}_2 - \boldsymbol{\mu})(\mathbf{y}_2 - \boldsymbol{\mu})' + \frac{1}{2}(\mathbf{y}_3 - \boldsymbol{\mu})(\mathbf{y}_3 - \boldsymbol{\mu})'\end{aligned}$$

We can use exponentially decreasing weights to allow for the following time-varying covariance matrix

$$\begin{aligned}\boldsymbol{\Sigma}_t &= \lambda(\mathbf{y}_{t-1} - \boldsymbol{\mu})(\mathbf{y}_{t-1} - \boldsymbol{\mu})' + \lambda^2(\mathbf{y}_{t-2} - \boldsymbol{\mu})(\mathbf{y}_{t-2} - \boldsymbol{\mu})' + \ldots \\ &= \sum_{m=1}^{\infty} \lambda^m (\mathbf{y}_{t-m} - \boldsymbol{\mu})(\mathbf{y}_{t-m} - \boldsymbol{\mu})'\end{aligned} \tag{2.48}$$

https://doi.org/10.26219/heal.aueb.6306

And after some algebraic manipulation, we obtain the following model for the time varying – variance-covariance matrix (see Section 1.1.2)

$$\begin{aligned}\mathbf{\Sigma}_t &= (1-\lambda)(\mathbf{y}_{t-1}-\boldsymbol{\mu})(\mathbf{y}_{t-1}-\boldsymbol{\mu})' + \lambda\mathbf{\Sigma}_{t-1} = \\ &= (1-\lambda)\boldsymbol{\varepsilon}_{t-1} + \lambda\mathbf{\Sigma}_{t-1}\end{aligned} \tag{2.49}$$

From the above equation given λ and an initial estimate $\mathbf{\Sigma}_1$, the time-varying exponentially weighted covariance matrices can be easily computed.

Let us consider the $\mathrm{VEC}(1,1)$ model, which has the following form

$$\mathbf{h}_t = \boldsymbol{\mu} + \mathbf{A}_1\mathbf{h}_{t-1} + \mathbf{B}_1\boldsymbol{\varepsilon}_{t-1} \tag{2.50}$$

In case where we consider two assets, we have

$$\mathbf{H}_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix} \text{ and } \boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' = \begin{bmatrix} \varepsilon_{1,t}^2 & \varepsilon_{12,t} \\ \varepsilon_{21,t} & \varepsilon_{2,t}^2 \end{bmatrix}$$

$$\mathbf{A}_1 = \begin{bmatrix} \alpha_{11}^1 & \alpha_{12}^1 & \alpha_{13}^1 \\ \alpha_{21}^1 & \alpha_{22}^1 & \alpha_{23}^1 \\ \alpha_{31}^1 & \alpha_{32}^1 & \alpha_{33}^1 \end{bmatrix} \text{and } \mathbf{B}_1 = \begin{bmatrix} \beta_{11}^1 & \beta_{12}^1 & \beta_{13}^1 \\ \beta_{21}^1 & \beta_{22}^1 & \beta_{23}^1 \\ \beta_{31}^1 & \beta_{32}^1 & \beta_{33}^1 \end{bmatrix}$$

Given the above matrices, the $vech(\cdot)$ operator stacks the lower triangular part of the 2×2 time-varying conditional variance-covariance matrix $\mathbf{H}_t$, into an $\frac{2(2+1)}{2}\times 1$ vector

$$\mathbf{h}_t = \begin{bmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{bmatrix}$$

Also the $vech(\cdot)$ operator stacks the lower triangular part of the 2×2 variance-covariance matrix $\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'$, into an $\frac{2(2+1)}{2}\times 1$ vector

$$\boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{2,t}\varepsilon_{1,t} \\ \varepsilon_{2t}^2 \end{bmatrix}$$

So, the three conditional variance equations have as follows (for simplicity we disregard the superscript from the elemets of matrices $\mathbf{A}_1$ and $\mathbf{B}_1$)

$$\begin{aligned} h_{11,t} &= \mu_{11} + \alpha_{11} h_{11,t-1} + \alpha_{12} h_{12,t-1} + \alpha_{13} h_{22,t-1} + \\ &+ \beta_{11} \varepsilon_{1,t-1}^2 + \beta_{12} \varepsilon_{1,t-1} \varepsilon_{2,t-1} + \beta_{13} \varepsilon_{2,t-1}^2 \end{aligned} \tag{2.51}$$

$$\begin{aligned} h_{21,t} &= \mu_{21} + \alpha_{21} h_{11,t-1} + \alpha_{22} h_{12,t-1} + \alpha_{23} h_{22,t-1} + \\ &+ \beta_{21} \varepsilon_{1,t-1}^2 + \beta_{22} \varepsilon_{1,t-1} \varepsilon_{2,t-1} + \beta_{23} \varepsilon_{2,t-1}^2 \end{aligned} \tag{2.52}$$

$$\begin{aligned} h_{22} &= \mu_{22} + \alpha_{31} h_{11,t-1} + \alpha_{32} h_{12,t-1} + \alpha_{33} h_{22,t-1} + \\ &+ \beta_{31} \varepsilon_{1,t-1}^2 + \beta_{32} \varepsilon_{1,t-1} \varepsilon_{2,t-1} + \beta_{33} \varepsilon_{2,t-1}^2 \end{aligned} \tag{2.53}$$

In order to obtain a feasible model for empirical work, we must impose restrictions on the parameter matrices of the $\mathrm{VEC}(1,1)$ model. Bollerslev *et al.*(1988) proposed to use **diagonal** parameter matrices such that the conditional variance of one variable only depends on lagged squared values of the same variable, and the conditional covariances between two variables only depend on lagged values of cross-products of these variables. This model reduces substantially the number of parameters, but potentially important causalities are excluded.

Thus, if we set, in the above equations, $\alpha_{ij} = 0$ and $\beta_{ij} = 0$, for $i \neq j$ we get the parameterization of the bivariate GARCH (1,1)

$$h_{11,t} = \mu_{11} + \alpha_{11} h_{11,t-1} + \beta_{11} \varepsilon_{1,t-1}^2 \tag{2.54}$$

$$h_{21,t} = \mu_{21} + \alpha_{22} h_{21,t-1} + \beta_{22} \varepsilon_{2,t-1} \varepsilon_{1,t-1} \tag{2.55}$$

$$h_{22,t} = \mu_{22} + \alpha_{33} h_{22,t-1} + \beta_{33} \varepsilon_{2,t-1}^2 \tag{2.56}$$

In general, an M -variate model will comprise $M(M+1)/2$ equations each one of which will contain a set of parameters describing the evolution of the dependent variable. The number of parameters of a fairly rich volatility model soon becomes big enough, in some cases even for M < 6, to make the estimation infeasible.

https://doi.org/10.26219/heal.aueb.6306

The Constant Correlation Model

Bollerslev (1990) introduced a new class of MGARCH models, the Constant Conditional Correlations or CCC, for which variances and covariances can be modeled separately.

Given the time-varying conditional variance-covariance matrix $\mathbf{H}_t$, the conditional correlation ρ_{ijt} between any two random variables i and j has as follows

$$\rho_{ijt} = \frac{h_{ijt}}{\sqrt{h_{iit}}\sqrt{h_{jjt}}} \tag{2.57}$$

In the CCC model, Bollerslev (1990) assumes that for all random variables i and j the following is true

$$h_{ijt} = \rho_{ij}\sqrt{h_{iit}h_{jjt}} \tag{2.58}$$

Notice ρ_{ij} does not depend on time t, so the conditional correlation between any two random variables is constant through time. In case of two random variables (y_{1t}, y_{2t}), the time-varying conditional variance-covariance matrix $\mathbf{H}_t$ has as follows

$$\mathbf{H}_t = \begin{bmatrix} h_{11,t} & \rho_{12}\sqrt{h_{11t}h_{22t}} \\ \rho_{21}\sqrt{h_{11t}h_{22t}} & h_{22,t} \end{bmatrix}$$

Hence we have to model the evolution of the 2 diagonal elements ($h_{11,t}$ and $h_{22,t}$) and estimate 1 constant correlation. Next, we denote by and $\mathbf{R}$ the constant correlation matrix. In the case of the two random variables we have

$$\mathbf{R} = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{21} & 1 \end{bmatrix}$$

Then, we consider the diagonal matrix $\mathbf{D}$, whose non-zero elements, which appear in the main diagonal are the conditional standard deviations $\sqrt{h_{11,t}}$ and $\sqrt{h_{22,t}}$

$$\mathbf{D}_t = \begin{bmatrix} \sqrt{h_{11,t}} & 0 \\ 0 & \sqrt{h_{22,t}} \end{bmatrix}$$

Each of the conditional standard deviations $\sqrt{h_{11,t}}$ and $\sqrt{h_{22,t}}$ can be modeled with a univariate GARCH model allowing for different specifications and lags.

https://doi.org/10.26219/heal.aueb.6306

Finally, the whole conditional variance-covariance matrix $\mathbf{H}_t$ is then constructed from these univariate processes, the constant correlation matrix $\mathbf{R}$, and the time-varying diagonal matrix $\mathbf{D}$ as follows

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t \tag{2.59}$$

The Dynamic Conditional Correlation Model

In the dynamic conditional correlation model, Engle and Sheppard (2001) allow for time-varying conditional correlations. Specifically, instead of equation (2.58), in this model the conditional variance-covariance matrix $\mathbf{H}_t$ is partitioned as follows

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \tag{2.60}$$

where $\mathbf{R}_t$ is a time-varying correlation matrix.

2.4.3 Criticism of the GARCH Models

In the recent past there some authors, like Cumby (et al. 1993) and Figlewski (1997) have debated on whether simple averages of historical volatility can predict future volatility better than GARCH models. In response, Andersen and Bollerslev (1998) provided integrated volatility as a significantly more accurate measure than simple squared or absolute realized returns. Through the latter they showed how volatility models provide strikingly accurate volatility forecasts.

Then, the attention was turned to the fact that volatilities of assets and markets tend to move together in time. This widely accepted feature of the data has led to the multivariate modelling of volatilities

https://doi.org/10.26219/heal.aueb.6306

3 EMPIRICAL FINDINGS AND CONCLUSIONS

As we pointed out in the second chapter, we rely on two optimization techniques expounded by Markowitz (1952) and Tobin (1958) in order to construct optimal portfolios, out of a pool of 36 stocks. We apply these two portfolio optimization techniques using daily discrete returns for the estimation period consisting of 2.415 observations; we also allow for short sales.

The whole procedure consists of three steps. First, using the two different optimization techniques we derive the “optimal weights” of the stocks making up the optimal portfolios. Second, based on these estimated weights, we evaluate the performance of 9 optimal portfolios during the *evaluation period*, by computing the portfolios’ average return and risk.

The first three portfolios are constructed following the methodology of Markowitz (1952). That is, we set the portfolio’s expected return level (on a daily basis) at 0.10%, 0.15% and 0.20% and we come up with optimal portfolios 1, 2 and 3 whose, daily risk equals 0.72%, 0.99% and 1.35% respectively. Then, drawing on Tobin’s (1958) methodology, we construct optimal portfolios 4, 5, and 6 by taking into account three yearly risk-free rates 3.0%, 3.5% and 4.0%; it turns out that these portfolios have daily expected return equal to 0.1589%, 0.1610% and 0.1632%, respectively, and a daily portfolio risk equal to 1.05%, 1.06% and 1.08% respectively.

Next, we introduce the GARCH modeling in our analysis by fitting a $GARCH(1,1)$ model to describe the variance of the stocks’ returns during the estimation period. From the estimation of the $GARCH(1,1)$ we derive 36 estimates of the stocks’ **long-term, variance rate** V. As a way to incorporate multivariate GARCH modeling in our analysis we follow Bollerslev (1990) in constructing the conditional variance-covariance matrix $\mathbf{H} = \mathbf{D}_t\mathbf{R}\mathbf{D}_t$, where $\mathbf{D}$ is a diagonal matrix, whose elements are the estimates of long-term standard deviation rates, $\mathbf{R}$ is the constant correlation matrix. Having derived $\mathbf{H}$ we rely on Tobin’s approach to derive optimal portfolios based on three yearly risk-free rates 3.0%, 3.5% and 4.0%; the daily expected return for optimal portfolios 7, 8 and 9, are 0.1720%, 0.1745% and 0.1772% respectively, while the daily portfolio’s risk 1.27%, 1.29% and 1.32% respectively.

Finally in the third step, we evaluate the performance of 7 optimal portfolios during the **ex-post testing period,** which extends from day 2.416 to day 2.780. Based on stocks’ returns

https://doi.org/10.26219/heal.aueb.6306

during the testing period, for every euro invested in them, at the beginning of the testing period (i.e. day 4.416), portfolios 1, 2, 3, 4, 5, 6, and 8 yielded, at the end of the testing period (i.e. day 2.780) €1.032, €1.18, and €1.33, €1.21, €1.22, and €1.22, €1.34 respectively.

3.1 DATA

In this study, we shall consider allocations across the 36 stocks listed on the Athens Stock Exchange. The following table shows the means and standard deviation of these stocks (on a daily basis) based on the entire period, 1-2.780.

Table 3.1-1: Mean and Standard Deviation of Stocks (Daily Basis)

STOCK	MEAN DAILY RETURN	DAILY STANDARD DEVIATION
1	0.00099	0.0152
2	0.00156	0.0267
3	0.00185	0.0269
4	0.00204	0.0373
5	0.0008	0.0225
6	0.0023	0.04208
7	0.00075	0.0204
8	0.0007	0.01750
9	0.0008	0.01861
10	0.000	0.02025
11	0.0007	0.01524
12	0.00071	0.01501
13	0.0022	0.03979
14	0.0014	0.02491
15	0.00069	0.01789
16	0.00139	0.0414
17	0.0015	0.0402
18	0.00260	0.0533
19	0.0005	0.01798
20	0.0006	0.02069
21	0.00067	0.0173
22	0.00055	0.0224
23	0.00051	0.0109
24	0.00077	0.01817
25	0.0011	0.0329
26	0.00246	0.0500
27	0.000690	0.017
28	0.00034	0.0346
29	0.00034	0.0346
30	0.00083	0.0349

https://doi.org/10.26219/heal.aueb.6306

31	0.00313	0.0553
32	0.00087	0.01934
33	0.00084	0.02116
34	0.00163	0.03199
35	0.00049	0.021759
36	-0.00034	0.0250

The combinations of risk and return offered by the 36 securities in the index can also be presented graphically (Figure 3.1-1)

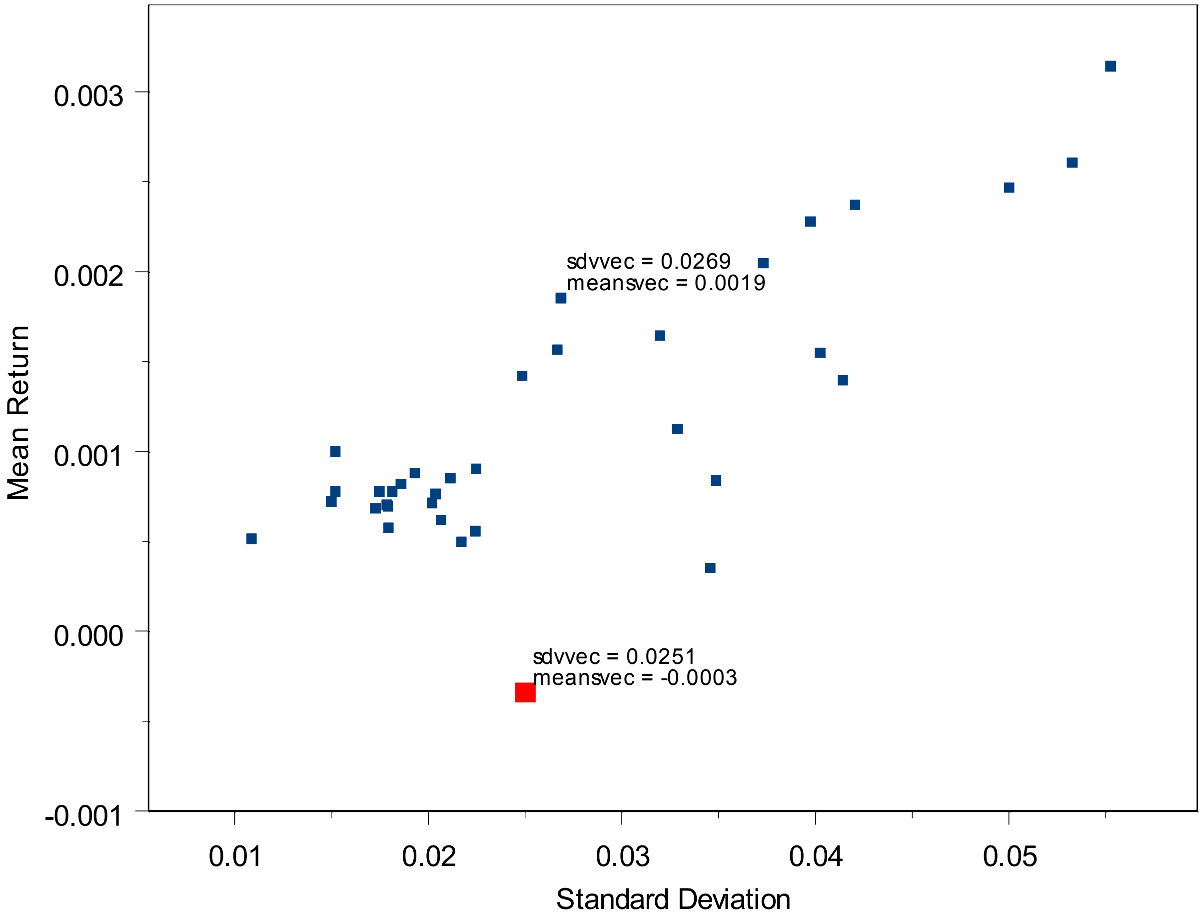

Figure 3.1-1: Risk and Return Characteristics Of The 36 Stocks

From the above figure we observe that some stocks are clearly dominated by other stocks. For example, stock 36 has approximately the same risk level (2,51%) as stock 3, however the average daily return of the former (-0.03%), is far below the average daily return of the latter (0.19%); thus stock 36 is clearly dominated by stock 3.

https://doi.org/10.26219/heal.aueb.6306

Notice the portfolio optimization procedures will be applied using daily discrete returns on the common stocks for the time period 1-2.415. This time period is technically known as the **estimation period.**

3.1.1 Definition of Data

There are two basic types of returns: discrete and continuous. Ignoring dividends, if we represent by $P_{i,t}$ the purchase price of a share of stock i at time-period t (which in our case apparently involves a specific day) and by $P_{i,t-1}$ the stock price in the previous time period, then the *holding period* **discrete return** , r_{di}, for stock i is calculated as follows:

$$r_{di} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} = \frac{P_{i,t}}{P_{i,t-1}} - 1 \qquad (3.1)$$

Also given the stock prices $P_{i,t}$ and $P_{i,t-1}$ we can also define the *holding period* **continuous compounding return**, r_{ci} ,for stock i as follows as:

$$r_{ci} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \qquad (3.2)$$

Then, having a series of returns (whether discrete or continuous) we can compute the stock's **average return**. There are two commonly quoted measures of average return, which rarely agree with each other: (a) the **arithmetic average return** and (b) the **geometric average return.** For a series of n discrete returns, $r_{d1},\ldots,r_{dn}$, the arithmetic mean ($\overline{r}_a$) and the geometric mean ($\overline{r}_g$) are calculated by the following formulas

$$\overline{r}_{ai} = \frac{1}{n}\sum_{j=1}^{n} r_{dij} \qquad (3.3)$$

$$\overline{r}_g = \left[\prod_{i=1}^{n}\left(1 + r_{dj}\right)\right]^{1/n} - 1 \qquad (3.4)$$

To illustrate the difference between the above two measures of average return let us consider a two-period example with the following characteristics: $P_0 = \$100$, $r_{d1} = -0.50$ and $r_{d2} = 1.00$. The question is whether the investor made money over the two periods or not.

Since arithmetic average return for the two periods is 25% - $\overline{r}_d = (1.00 - 0.50)/2 = 0.25$ - we must conclude that at the end of the two-year period the wealth of the investor should have risen to $156.25 [=100 (1+0.25)].However the story is different when we take into consideration the geometric average return, which is calculated as follows

$$\begin{aligned} P_1 &= P_0(1+r_1) = \$100(1-0.50) = \$50 \\ P_2 &= P_1(1+r_2) = \\ &= P_0(1+r_1)(1+r_2) = \$50(1+1.00) = \$100 \end{aligned}$$

The truth is that we didn't make any money since, based on the above returns, the terminal wealth remained at $100. So the geometric average is closer to investment experience .

In closing this section we need to underscore an important point. Our previous discussion on the calculation of average returns depends on what has happened in the past. However, the formulas for the return on a portfolio make use of the expected return. So in order to be able to use average return instead of expected return the implicit assumption we are making is that the daily distribution of returns of the past will be similar to the daily distribution of returns for the coming days. For instance if the security i has an arithmetic average return over the past 100 trading days of $\overline{r}_{ai}$ then we assume that this is the mean of the probability distribution of returns and that this mean is not going to change when an additional observation is added to the existing sample.

3.1.2 Estimation and Testing Periods

The **estimation period** will be set at 2.415 days, that is, 2.415 days on stock returns will be used to calculate the portfolio inputs (such as the means and the variance-covariance matrix), for the portfolio optimization. In other words, we shall use a 2.415-day rolling window to calculate certain stock characteristics like mean return, and stock risk; these inputs will then be used in order to form the optimal portfolios according to certain optimization methods (the major differentiation among them is the way they model stock return volatility).

Then, the **ex-post testing period** extends from day 2.416 to day 2.780; this will be the period over which the portfolios' buy-and-hold returns will be compared. The ex-post forecasts are performed to test the strength of each optimization procedure.

https://doi.org/10.26219/heal.aueb.6306

3.2 OPTIMAL PORTFOLIOS BASED ON THE HISTORICAL MODELLING OF STOCK RETURN VOLATILITY

3.2.1 Optimal Portfolios According to the Markowitz Approach

Based on the set up of the maximization problem described by equations (2.7)-(2.10) we see that in order to find the tangency portfolio, we calculate the variance-covariance matrix (2.6) using stock returns for the estimation period.

We remind the reader that the mean of all mean daily returns of the 36 stocks, during the estimation period was 0.11%. Thus we set the expected daily portfolio's return level at 0.10%, 0.15% and 0.20% Once we have the optimal weights, we can then compute the two important characteristics of the optimal portfolio, namely the expected return and risk.

Substituting the relevant numbers in the case where short sales are allowed yields a **daily expected return** for portfolios 1, 2 and 3 equal to **0.10%**, **0.15%** and **0.20%** respectively. These portfolios had a **monthly daily risk equal to 0.72%, 0.99%** and **1.35%** respectively**.**

3.2.2 Optimal Portfolios According to Tobin's Approach

In order to find optimal portfolios using Tobin's approach, we rely on equations (2.17) and (2.18) and the same variance-covariance matrix (2.6). The output of this optimization process takes into consideration three yearly risk-free rates (3.0%, 3.5% and 4.0%). Substituting the relevant numbers in the case where short sales are not allowed yields a **daily expected return** for portfolios 4, 5 and 6 equal to **0.1589%**, **0.1610%** and **0.1632%** respectively. These portfolios had a **daily portfolio risk equal to 1.05%, 1.06%** and **1.08%** respectively**.**

https://doi.org/10.26219/heal.aueb.6306

3.3 OPTIMAL PORTFOLIOS WITH GARCH EFFECTS IN STOCK RETURNS

3.3.1 Testing for ARCH and GARCH Effects in Stock Returns and the Long-Term Volatility Rates

At first we shall test whether there are significant ARCH and GARCH effects in the returns of the stocks under consideration.

So, we estimate a simple $GARCH(1,1)$ model with no regreessors in the mean equation. The mean equation and the equation for the conditional variance of excess returns have as follows:

$$r_{i,t} = \beta_0 + \varepsilon_{i,t} \tag{3.5}$$

$$\operatorname{var}\left(u_t \middle| \Phi_t\right) = \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \sigma_{t-1}^2 \tag{3.6}$$

where $r_{i,t}$ is the daily return for stock i at day t, for $t = 1, \ldots, 2.415$

The sum of regression coefficients ($\alpha_1 + \gamma_1$) expresses the influence of the variability of variables from the previous period on the current value of the variability; this value is usually close to 1.0, which is a sign of increased inertia in the effects of shocks on the variability of returns on financial assets.

The following table contains estimates for the population parameters β_0, α_0, α_1, and γ_1 of equations (3.5) and (3.6)

Table 3.3-1: GARCH (1,1) Estimation Of The 36 Stocks. Estimation Period, 1-2.415

EST. COEF.	**SR1**	**SR2**	**SR3**	**SR4**	**SR5**	**SR6**	**SR7**	**SR8**	**SR9**	**SR10**
$\hat{\beta}_0$	0.00111	0.00203	0.00185	0.00272	0.00091	0.00311	0.00111	0.0010	0.00096	0.00096
$\hat{\alpha}_0$	1.50E-06	6.26E-06	3.14E-05	3.55E-05	5.13E-06	1.37E-05	5.40E-06	1.10E-06	1.53E-05	2.28E-06
$\hat{\alpha}_1$	0.0497	0.0336	0.0846	0.0452	0.0454	0.0239	0.0525	0.040	0.0867	0.0177
$\hat{\gamma}_1$	0.9446	0.9585	0.8718	0.9299	0.9447	0.9688	0.9379	0.9583	0.8709	0.9770
$\hat{\alpha}_1+\hat{\gamma}_1$	0.9944	0.9922	0.9564	0.9752	0.9902	0.9928	0.9904	0.9987	0.9576	0.9947
	SR11	**SR12**	**SR13**	**SR14**	**SR15**	**SR16**	**SR17**	**SR18**	**SR19**	**SR20**
$\hat{\beta}_0$	0.00074	0.00079	0.00267	0.00163	0.00072	0.00300	0.00228	0.0027	0.00098	0.00095
$\hat{\alpha}_0$	1.52E-06	6.03E-07	8.48E-05	9.11E-05	2.10E-05	3.37E-05	4.84E-05	4.94E-05	1.16E-05	3.04E-06
$\hat{\alpha}_1$	0.0406	0.0268	0.0467	0.0932	0.1416	0.0502	0.0581	0.038	0.0848	0.0253
$\hat{\gamma}_1$	0.9533	0.9709	0.9001	0.7617	0.7980	0.9323	0.9153	0.9450	0.8794	0.9679
$\hat{\alpha}_1+\hat{\gamma}_1$	0.9939	0.9977	0.9468	0.8550	0.9397	0.9825	0.9735	0.9839	0.9642	0.9932
	SR21	**SR22**	**SR23**	**SR24**	**SR25**	**SR26**	**SR27**	**SR28**	**SR29**	**SR30**
$\hat{\beta}_0$	0.00082	0.00107	0.00069	0.00103	0.00150	0.00273	0.00073	0.0012	0.00070	0.00078
$\hat{\alpha}_0$	3.43E-06	1.69E-05	3.39E-06	3.05E-06	1.00E-05	3.07E-05	1.90E-05	5.10E-06	1.25E-05	8.65E-05
$\hat{\alpha}_1$	0.0601	0.0762	0.1010	0.0426	0.0273	0.0372	0.1167	0.0633	0.0305	0.1048
$\hat{\gamma}_1$	0.9302	0.8916	0.8760	0.9481	0.9638	0.9523	0.8287	0.9292	0.9596	0.8287
$\hat{\alpha}_1+\hat{\gamma}_1$	0.9903	0.9679	0.9771	0.9908	0.9911	0.9895	0.9454	0.9926	0.9901	0.9336
	SR31	**SR32**	**SR33**	**SR34**	**SR35**	**SR36**				
$\hat{\beta}_0$	0.00412	0.00087	0.00092	0.00148	0.00066	0.00045				
$\hat{\alpha}_0$	4.10E-05	6.14E-06	4.09E-05	0.000214	5.57E-06	1.48E-05				
$\hat{\alpha}_1$	0.0295	0.0300	0.1415	0.1094	0.0367	0.1183				
$\hat{\gamma}_1$	0.9572	0.9533	0.7717	0.6847	0.9535	0.8706				
$\hat{\alpha}_1+\hat{\gamma}_1$	0.9868	0.9833	0.9132	0.7941	0.9903	0.9889				

In the following table we present the t-statistics related to the above estimated coefficients

Table 3.3-2: t-Statistics of GARCH (1,1) Estimation Of The 36 Stocks

Stock	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\gamma}_1$
SR1	3.732	9.076	158.90
SR2	3.393	6.157	124.68
SR3	5.888	8.505	23.54
SR4	4.301	5.510	91.70
SR5	6.188	9.147	131.96
SR6	6.302	6.732	5.97
SR7	5.305	9.013	64.62
SR8	3.700	6.785	79.33
SR9	6.701	11.271	48.88
SR10	6.224	9.500	348.74
SR11	3.412	7.203.	122.75
SR12	2.440	5.803	150.64
SR13	7.896	6.679	38.31
SR14	6.276	6.848.	19.06
SR15	3.977	8.702	121.24
SR16	7.614	13.57	225.66
SR17	7.266	10.34	69.01
SR18	7.037	10.212	113.383
SR19	6.623	9.825	35.10
SR20	3.585	6.411	183.38
SR21	5.493	9.167	92.30
SR22	5.007	6.107	92.97
SR23	4.839	7.759	45.92
SR24	7.429	11.589	227.42
SR25	5.125	8.425	255.45
SR26	6.030	7.268	173.82
SR27	6.844	11.264	72.41
SR28	4.889	9.901	127.89
SR29	4.294	5.920	76.47
SR30	1.077	10.305	8.26
SR31	3.886	8.123	200.50
SR32	7.605	9.874	37.64

https://doi.org/10.26219/heal.aueb.6306

SR33	7.659	12.65	141.68
SR34	9.578	8.994	17.60
SR35	1.255	12.85	42.66
SR36	9.125	15.062	110.86

In the above table we see that there are significant ARCH and GARCH effects in the stocks' returns.

Then, using equation (2.39) and the estimated coefficients $\hat{\alpha}_1$, $\hat{\gamma}_1$, we can compute the unconditional variance of the $GARCH(1,1)$ model.

$$V = \frac{\alpha_0}{1-\alpha_1-\gamma_1} \qquad (3.7)$$

The above the unconditional variance is known as the **long-term, variance rate** V; this is also the unconditional variance of u_{t+1}. In order to have a finite unconditional variance, we require (a) the constant in the $GARCH(1,1)$ process to be non-zero, i.e. $\alpha_0 > 0$ and (b) the stationarity conditions to be met, that is $\alpha_1 + \gamma_1 < 1$. From the above table, we see that the conditions for the correct estimation of the $GARCH(1,1)$ model are met.

In the following table, we present our estimates for V for the 36 stocks.

https://doi.org/10.26219/heal.aueb.6306

Table 3.3-3: Estimates of the Stocks' Long-Run Volatility Rate

EST. COEF.	**SR1**	**SR2**	**SR3**	**SR4**	**SR5**	**SR6**	**SR7**	**SR8**	**SR9**	**SR10**
$V=\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}_1}$	0.00026	0.00080	0.00072	0.00143	0.00052	0.00191	0.00056	0.0008	0.00036	0.00043
$\sqrt{V}=\sqrt{\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}}}$	0.0164	0.0283	0.0268	0.0378	0.0229	0.0437	0.0237	0.0295	0.0190	0.0208
	SR11	**SR12**	**SR13**	**SR14**	**SR15**	**SR16**	**SR17**	**SR18**	**SR19**	**SR20**
$V=\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}_1}$	0.00025	0.00026	0.00159	0.00062	0.00034	0.00193	0.00182	0.0030	0.00032	0.00045
$\sqrt{V}=\sqrt{\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}}}$	0.0158	0.0162	0.0399	0.0250	0.0186	0.0439	0.0427	0.0554	0.0180	0.0212
	SR21	**SR22**	**SR23**	**SR24**	**SR25**	**SR26**	**SR27**	**SR28**	**SR29**	**SR30**
$V=\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}_1}$	0.00036	0.00052	0.00014	0.00033	0.00112	0.00294	0.00034	0.0006	0.00126	0.00130
$\sqrt{V}=\sqrt{\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}}}$	0.0188	0.0229	0.0121	0.0182	0.0335	0.0542	0.0186	0.0264	0.0355	0.0361
	SR31	**SR32**	**SR33**	**SR34**	**SR35**	**SR36**				
$V=\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}_1}$	0.00310	0.00037	0.00047	0.00104	0.00058	0.00133				
$\sqrt{V}=\sqrt{\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\gamma}}}$	0.0557	0.0192	0.0217	0.0322	0.0240	0.0365				

We can then, following Bollerslev (1990), create the matrix $\mathbf{D}$, a diagonal matrix, whose elements are the estimates of long-term standard deviation rates derived from the $GARCH(1,1)$ model, and the constant correlation matrix $\mathbf{R}$. Thus, the variance-covariance matrix $\mathbf{H}$ is computed as $\mathbf{H}=\mathbf{D}_t\mathbf{R}\mathbf{D}_t$

https://doi.org/10.26219/heal.aueb.6306

3.3.2 Optimal Portfolios with GARCH Effects in Stocks

Having derived the variance-covariance matrix $\mathbf{H}$, as previously discussed, we can then rely on Tobin's approach to derive optimal portfolios based on three yearly risk-free rates - 3.0%, 3.5% and 4.0%. The results are shown in the following table

Table 3.3-4: The Optimal Weights According To Tobin's Optimization Framework – Short Sales Allowed And GARCH Effects In Stock Returns

Weights	$r_F = 3.0\%$	$r_F = 3.5\%$	$r_F = 4.0\%$
	Optimal Weights Portfolio 7	**Optimal Weights Portfolio 8**	**Optimal Weights Portfolio 9**
w_1	0.07048	0.07216	0.0739
w_2	0.0298	0.03030	0.03084
w_3	0.14302	0.14618	0.14950
w_4	0.02558	0.02642	0.02729
w_5	0.04478	0.04439	0.04399
w_6	0.02069	0.02153	0.02242
w_7	-0.06625	-0.06847	-0.070797
w_8	-0.03893	-0.03900	-0.03907
w_9	0.06561	0.06618	0.06677
w_{10}	-0.0073	-0.007737	-0.00813
w_{11}	0.14174	0.14234	0.14298
w_{12}	-0.0024	-0.00330	-0.00416
w_{13}	0.07257	0.074448	0.076418
w_{14}	0.08590	0.08809	0.09039
w_{15}	0.03180	0.031448	0.031072
w_{16}	-0.0083	-0.00843	-0.008526
w_{17}	0.00246	0.00255	0.00264
w_{18}	0.04910	0.050236	0.05142
W_{19}	0.00403	0.002790	0.001484
W_{20}	-0.0111	-0.011895	-0.01263
W_{21}	-0.01225	-0.013304	-0.01440
W_{22}	-0.02377	-0.024934	-0.026154
W_{23}	0.1986	0.196333	0.193904
W_{24}	0.03244	0.0322816	0.032115
W_{25}	0.01349	0.0137599	0.014043
W_{26}	0.0356	0.036435	0.037262
W_{27}	-0.02056	-0.021723	-0.022939
W_{28}	-0.06399	-0.0660464	-0.068194
W_{29}	-0.02954	-0.0303	-0.0312352

https://doi.org/10.26219/heal.aueb.6306

W_{30}	0.00707	0.00681	0.0065381
W_{31}	0.06250	0.06391	0.0654081
W_{32}	0.10349	0.1042210	0.104985
W_{33}	0.09300	0.0938154	0.09466
W_{34}	0.07715	0.0788821	0.08069
W_{35}	-0.00684	-0.008085	-0.0093
W_{36}	-0.11953	-0.1222734	-0.12514

Substituting the relevant numbers in the case where short sales are not allowed yields a **daily expected return** for portfolios 4, 5 and 6 equal to **0.1720%**, **0.1745%** and **0.1772%** respectively. These portfolios had a **daily portfolio risk equal to 1.27%, 1.29%** and **1.32%** respectively**.**

3.4 COMPARING THE PERFORMANCE OF THE OPTIMAL PORTFOLIOS

So far we have derived the optimal weights of various optimal portfolios, along with their expected return and risk characteristics. This was useful when it comes to examining and evaluating the portfolios' past performance.

We know turn to the evaluation of portfolios' performance, by examining these optimal portfolios during the evaluation period 2.416-2.780. Having constructed the 36 vectors of daily returns during the evaluation period, one for each of the 36 stock, subsequently we can easily compute the terminal value (at day 2.780) per euro of investment in a particular stock.

3.4.1 The Performance of Optimal Portfolios with Historical Modeling of Stock Return Volatility

Using the Markowitz's approach in portfolio optimization, we constructed three optimal portfolios, using stock returns from the estimation period. These optimal portfolios had average daily returns of 0.10%, 0.15% and 0.20%, respectively, and daily risk (standard deviation) of 0.72%, 0.99% and 1.35%, respectively. We know want to see how these optimal portfolios fared during the evaluation period.

When the weights of these optimal portfolios are multiplied by the vector of cumulative returns for the ex-post estimation period, we find that for each euro invested in portfolio 1, 2, and 3, the investor got back at the end of the year €1.032, €1.18, and €1.33, respectively.

In order to find optimal portfolios using Tobin's approach, we rely on equations (2.17), (2.18). These equations for the estimation period 1-2.415 have as follows

$$\bar{\mathbf{r}}_{1:2.415} - \mathbf{i}r_F = \mathbf{\Sigma}_{1:2.415}\mathbf{Z}_{1:2.415} \tag{3.8}$$

Where

$\mathbf{\Sigma}_{1:2.415}$ = a 36×36 symmetric variance-covariance matrix

Solving the above for $\mathbf{Z}_{1:2.415}$ we get

$$\mathbf{Z}^{*}_{1:2.415} = \mathbf{\Sigma}^{-1}_{1:2.415}\left(\bar{\mathbf{r}}_{1:2.415} - \mathbf{i}r_F\right) \tag{3.9}$$

Finally, in order to find the optimal weights we calculate the following

$$w_i = Z_i / \mathbf{Z}^{*\prime}\mathbf{i} \tag{3.10}$$

In constructing the optimal portfolio using this approach we take into consideration a yearly risk-free rate of 3.0%, 3.5%, and 4%. In these optimal portfolio, for each euro invested the investor got back at the end of the evaluation period €1.21, €1.22, and €1.22 respectively

3.4.2 Optimal Portfolios with GARCH Effects in Stocks

Using the estimation period 1 – 2.415, we fit a $GARCH(1,1)$ model, from which we extract the long-term volatilities (expressed as standard deviations) of the 36 stocks. Then, using these estimates of volatility we form the matrix $\mathbf{D}_{1:2.415}$, a diagonal matrix, whose elements are the aforementioned estimates of standard deviation rates, and the constant correlation matrix $\mathbf{R}_{1:2.415}$. So the variance-covariance matrix $\mathbf{H}_{1:2.415}$, for the estimation period, s computed as $\mathbf{H}_{1:2.415} = \mathbf{D}_{1:2.415}\mathbf{R}_{1:2.415}\mathbf{D}_{1:2.415}$ .Having derived the variance-covariance matrix $\mathbf{H}_{1:2.415}$ we can then rely on Tobin's approach to derive optimal portfolios based on a risk-free rate of 3.5%.

In this optimal portfolio, for each euro invested the investor got back at the end of the year €1.34

https://doi.org/10.26219/heal.aueb.6306

3.5 CONCLUSIONS

In this dissertation using the theory on portfolio optimization, we constructed and calculated the characteristics (i.e. the expected return and risk) of 9 different optimal portfolios; this was done over an estimation period covering 2.415 days. The construction of these portfolios (which essentially involves selecting the stocks that shall be included and determining their weights) differed in two respects. First, different portfolio optimization methods – the method based on Markowitz (1952) and the method based on Tobin (1958)- were used in order to select the stocks and derive the optimal weights of each stock included in the optimal portfolio. Second, in the construction of some optimal portfolios the GARCH (1,1) modelling was used in order to compute standard deviations of stock returns.

The employment of different optimization techniques and the different ways of modelling a stock's volatility resulted in no significant differences cropping up, as far the performance of the optimal portfolios, in the testing period, is concerned.

Generally, we can spot three basic shortcomings in this dissertation. To begin with, the study's major shortcoming is the limited amount of companies examined, since it considers only limited amount of stock (only 36). As a result, we cannot assert that study's conclusions are indisputable.

Then, another shortcoming of the dissertation is related to the type of ex post data used as a proxy for expected returns (i.e. ex ante data). Specifically, in computing the optimal portfolios the study uses ex post data derived from only arithmetic returns and average arithmetic returns. An alternative set of ex post data could be obtained with continuous monthly returns and geometric average returns.

Finally, the future testing period is restricted to only to a 6-month period, so any conclusions derived are relevant only for the short term.

For further research, we could include more variables in our analysis. For example, one very important variable is the size of portfolio or the market value of the stocks (market like Greek stock market are characterized sometimes for "thin" trading in some stocks), another is the use of derivatives and if that is related to the observed risk/return profile.

https://doi.org/10.26219/heal.aueb.6306

REFERENCES

1. Andersen, T.G. and Bollerslev, T. (1998). “Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts.” *International Economic Review* 39: 885-905.
2. Bollerslev, T. (1990). “Modeling the Coherence in Short-run Nominal Eevchange Rates: A Multivariate Generalized ARCH model.” *Review of Economics and Statistics,* 72: 498-595
3. Bollerslev, T., Engle, R. And Wooldridge, J. (1988). “A Capital Asset Pricing Model with Time-Varying Covariances.” *Journal of Political Economy,* 96: 116-131
4. Bollerslev, T. (1986). “Generalized Autoregressive Conditional Heteroskedasticity.” *The Journal of Econometrics,* 31: 307-327
5. Chance, D. (1994) “The ABCs of Geometric Brownian Motion.” *Derivatives Quarterly,* 1 (Winter): 41-47.
6. Chou, P., LI, W. And Zhou, G. (2004). “Portfolio Optimization under Asset Pricing Anomalies”. *Japan and the World Economy* 18: 121-142.
7. Cohen, J., Zinbarg, E. And Zeikel, A. (1982). *Investment Analysis and Portfolio Management.* 4th ed. Richard D. Irwin, Inc.
8. Cumby, R., Figlewski, S., and Hasbrouck, J. (1993). “Forecasting Volatility and Correlations with EGARCH Models.” *Journal of Derivatives*, Winter: 51-63.
9. Dougherty, C. (1992). *Introduction to Econometrics*. Oxford University Press
10. Elton, E. and Gruber, M. (1995). *Modern Portfolio Theory and Investment Analysis.* John Wiley & Sons.
11. Engle, R. (2001). "GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives,* 15(4):157-168
12. Engle, R. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation." *Econometrica,* 50:987-1008
13. Engle, R. and Kroner, K. (1995). “Multivariate simultaneous GARCH.” *Econometric Theory*, 11:122-150.
14. Engle, R. and Sheppard, K. (2001). “Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH.” *Working Paper 8554, National Bureau of Economic Research.*
15. Engle, R. Lilien, D, and Robins, R. (1987). “Estimation of Time Varying Risk Premia in The Term Structure: The Arch-M Model.” *Econometrica*, 55: 391–407.

16. Fabozzi, F. (1999). *Investment Management.* 2nd ed. Prentice-Hall, Inc.
17. Fama, E. (1970). "Efficient Capital Markets: A Review Of Theory And Empirical Work," *Journal of Finance* 25, 383-417.
18. Figlewski, S. (1997). "Forecasting Volatility." *Financial Markets, Institutions and Instruments,* 6: 1-88.
19. Greene, W. (2000). *Econometric Analysis.* 4thed. Prentice Hall International, Inc.
20. Hull, J. (1997). *Options, Futures, and Other Derivatives.* 3rd ed. Prentice Hall International, Inc.
21. Johnston, J and Dinardo, J. (1997). *Econometric Methods.* 4th ed. McGraw-Hill Inc.
22. Kolb, R. (1996). *Investments.* Blacwell Publishers.
23. Lee, S. and Stevenson, S. (2000). "Real Estate Portfolio Construction and Estimation Risk". Paper presented at the 7th European Real Estate Society Conference, Bordeaux, June: 14-16.
24. Mandelbrot, B. (1963). "The Variation of Certain Speculative Prices." *Journal of Business* 36: 394-419.
25. Markowitz, H. (1952). "Portfolio selection". *Journal of Finance*, 7 (1): 77-91.
26. Merton, R. (1973). "An Intertemporal Asset Pricing Model". *Econometrica.*
27. Merton, R. (1971). "Optimum Consumption and Portfolio Rules in a Continuous-Time Model". *Journal of Economic Theory.*
28. McCulloch, J. (1985). "On Heteroscedasticity." *Econometrica,* 53(2)
29. Mueller, G, and Laposa, P. (1995). "Property-Type Diversification in Real Estate Portfolios: Size and Return Perspective". *The Journal of Real Estate Portfolio Management,* 1(1): 39-50
30. Pagan, A. and Schwert, W. (1990). "Alternative Methods for Conditional Stock Volatility". *Journal of Econometrics*, 45: 267-290.
31. Pindyck, R. and Rubinfeld, D. (1991). *Econometric Models and Economic Forecasts.* 3rd ed. McGraw-Hill Inc.
32. Ross, S., Westerfield, R. and Jaffe, J. (1999). *Corporate Finance.* 5th ed.. Irwin/McGraw-Hill
33. Saunders, M., Lewis, P. and Thornhill, A. (2003). *Research Method for Business Students*. 3rd ed. FT Prentice Hall.
34. Sharpe, W. (1964). "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance,* 19: 425-442

https://doi.org/10.26219/heal.aueb.6306

35. Sivitanides, P. and Southard, J. (2000). “Does Historical Volatility Capture True Risk in Real Estate?” Paper presented at the ARES Conference, Santa Barbara, CA.
36. Strong, R. (2000). *Portfolio Construction, Management & Protection.*.2nd ed..
37. Tobin, J. (1958). “Liquidity Preference as Behavior Toward Risk.” *The Review of Economic Studies*.
38. Weiss, N. (2005). *Course in Probability*. Addison-Wesley.

https://doi.org/10.26219/heal.aueb.6306