LUISS

Dipartimento
di Impresa e Management

Cattedra Asset Pricing

An empirical application of a parametric approach to portfolio choices on Italian stock market

Prof. Paolo Porchia

RELATORE

Prof. Marco Pirra

CORRELATORE

Lidia Buldrini Matr. 711411

CANDIDATO

Anno Accademico 2019/2020

Index

Introduction

Since ancient times, the decision of how to distribute available wealth among different investment opportunities has been an important issue for the humankind. In fact, we can already find examples of investment advices in the literature of the first centuries of the first millennium, which suggested to invest wealth in different areas of economy, such as land as well as merchandise.

In recent times, however, the problem of allocating available funds among different assets has become relevant not only for individual investors who are willing to guarantee themselves an income for their retirement time, but also for institutional investors which want to design portfolios in order to finance their long-term stream of spending. Naturally, the way in which decision are taken differs between the two of them. In fact, optimal portfolio decisions depend on the available financial assets in the investable universe, the risk and return profile, the preferences and circumstances of the investor.

In financial literature, the pioneers in optimal portfolio choice framework have been Markowitz (1952), Merton (1969-1971), Samuelson (1969) and Fama (1970). Their formulations have been revolutionary for that time but now well understood. Consequently, nowadays academic research focuses on the identification of critical issues in real world portfolio choices and their influence on investors.

This paper has the purpose of testing the applicability on Italian stock market of one of the new econometric approaches formulated in literature which uses parametric portfolio weights in order to find the optimal ones that maximizes investor's utility.

This is a new area of econometrics which detaches from traditional approaches. Instead of basing its activity on a two step procedure which first draws inference on parameters or tries to identify the return distribution and then solves for the parameters which maximize utility, the econometrician now specifies portfolio weights as a function of observable quantities and then solves for the maximum level of utility.

In this stream of literature, I will focus on the parametric model to portfolio choice formulated by Brandt, Santa-Clara and Valkanov in 2009. In their approach, the authors model portfolio weights as a linear function of three firm characteristics: market capitalization, book-to-market ratio and one-year lagged return. After this phase, they calculate optimal portfolio weights that maximize investor's utility.

Brandt et al.’s approach has four main advantages. First, it avoids the returns’ modelling phase, which is central in traditional econometrics, by directly focusing on drawing inference on portfolio weights’ parameters. Second, it includes implicitly the connection among returns, covariances, variances and firm characteristic. Third, it helps estimation problems like overfitting or imprecise coefficient estimates. At last, it allows the econometrician to easily test hypotheses, whether they are joint or single.

This work is articulated as follow. The first chapter gives a brief overview of the portfolio choice problem in literature, starting with Markowitz paradigm across econometric approaches, from traditional ones to the more recent. This is propaedeutic for a better understanding of the exposition of the subsequent research.

The second chapter focuses on new econometric approaches and particularly on the portfolio choice model formulated by Brandt, Santa-Clara and Valkanov in 2009. There, its strengths and weaknesses will be highlighted.

The third and last chapter presents the results of my research. I will expose and comment finding for the simple linear policy case and for the model extensions considered. Then, I will also draw the main conclusions about the applicability of the model on the Italian stock market.

Chapter 1

Portfolio choice problem in literature

1.1 The theoretical problem: modelling issues

In financial literature, after the formulations of well-known authors such as Markowitz, Merton and Fama that are considered as pillars in the portfolio choice problem framework, econometricians tried to incorporate in their models an increasing number of features from real-life cases, combining theory and practical evidence. Among them, there is no representation that can be considered superior to the other, because each has its own basic assumptions, strengths and weaknesses. However, as every portfolio choice problem, they share the same modelling issues.

The first feature that arises when defining a portfolio policy is the choice of the utility function that better represents investor's preferences. The most commonly used in practice is the family of hyperbolic absolute risk aversion (HARA) because its value function is homogeneous in wealth. This class includes constant relative risk aversion function (CRRA), logarithmic utility, quadratic utility and constant absolute risk aversion (CARA). Although, toward this category of functions are moved many critics. Among them, there is the fact that they create an odd connection between risk aversion and the elasticity of intertemporal substitution, which are in practice one the reciprocal of the other. Another one is the consequentiality created between the time horizon considered for the investment opportunity and the asset allocations. Finally, they do not take into account investor's behaviour anomalies. Despite this short explanation, I will go deeply into this topic in the next chapter, when I will analyse a portfolio choice problem.

The second one regards the possibility of inserting intermediate consumption into a portfolio choice problem. We know from economic theory that the investor could be willing to choose to consume part of his wealth instead of wating for terminal utility. In this way the investor for each period has to determine both asset allocation and his optimal level of consumption. This feature can be inserted both in discrete time portfolio choice problem and in continuous time constructions.

Another important element to assess or assume when stating a portfolio choice problem is whether the market in which we operate is complete or not. A financial market can be defined as complete when "the state variables governing investment opportunities are driven by the same stochastic process that drive asset returns” (Campbell (2002)). In turn, we can distinguish two types of market completeness:

- Statically complete markets
- Dynamically complete markets.

Markets are statically complete when there is a number of traded assets at least equal to the number of future possible outcomes. On the other side, markets are said to be dynamically complete when investors, by trading dynamically available assets, can create a variety of securities that give as payoff one unit of consumption in a particular state and zero units in all other ones.

This classification has an influence on the complexity of the resolution of a portfolio choice problem.

The definition of the time horizon of the investment opportunities is another important element to consider. In practice, the investor has rarely a clear view on which will be the final date of his investment opportunity. However, in literature we can distinguish between an infinite and a finite horizon which lead to different computational procedures. Differently from what one could expect, the infinite case is easier to solve because we only need to find a steady state solution. On the contrary, finite-horizon cases can only be solved numerically and involve finding a solution for each time period considered. In the limit, a sequence of finite solutions can converge to an infinite one as the time horizon increases.

The last important feature to consider when modelling a portfolio choice problem is the presence of frictions on the market and background risk. Frictions are all problematic elements directly connected to the risky security. An example are transaction costs. On the contrary, background risk incorporates all risks that an investor has to face other than the ones directly connected to the risky security. Examples are investment and consumption of housing and entrepreneurial income.

After this brief overview, I now will analyse more in detail the issues of portfolio choice problem, across the most important econometric approaches in existing literature. I will start with Markowitz paradigm, which is probably the most famous portfolio choice formulation.

1.1.1 Markowitz paradigm

During its lifetime, an investor has often to face a portfolio choice problem. For example, in some countries, people have to choose in which pension fund invest their money in order to get a reimbursement when they will retire from work. These resources can be invested in many assets, but the main problem is how to allocate them when making the portfolio choice.

The most famous and commonly used method to assign weights to the assets in a portfolio is the mean-variance framework, formulated by Markowitz in 1952. Before going through computation, I want to start from a graph that best synthetises this framework.

MEAN-STANDARD DEVIATION DIAGRAM

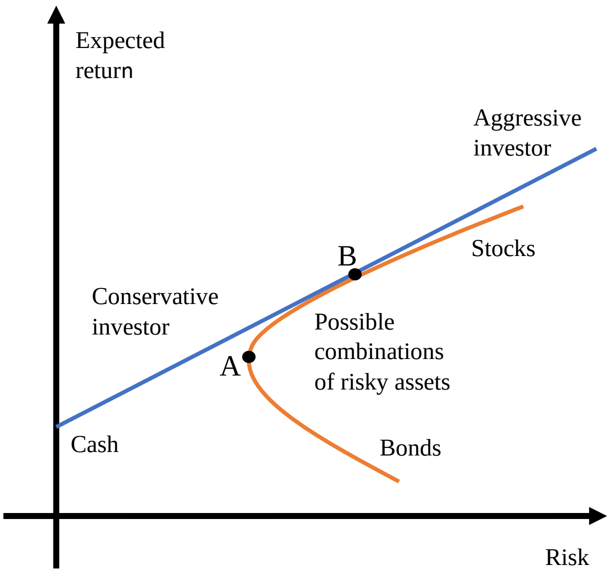

Figure 1.1 Source: Campbell J. Y. and L. M. Viceira (2002), “Strategic Asset Allocation”

Figure 1.1 summarizes all possible investment combinations when we consider only three assets in the investable universe: cash, stocks and bonds. The horizontal axis represents risk assumed by the investor, while the vertical axis is the expected return that he receives as a function of the risk taken on. When the investor has available only bonds and stocks, he will choose the weights for each asset that would minimize his variance. The curved line represents all possible portfolio allocations. Point A denotes the minimum variance portfolio. On the other side, the straight line exemplifies all possible combinations of assets when we include also the risk-free security in the portfolio. Besides, point B represents the extreme case in which the investor employs all its resources in the risk-free asset, like, for example, cash. As we continue on the line, the strategy of the investor becomes more and more aggressive, because he takes on more risk but he receives a higher compensation. The tangency point between the two lines, the straight and the curved one, is the optimal portfolio allocation among the three assets considered in the mean-variance framework.

All this explanation can be generalized for including more assets by using matrix notation.

Consider R_t, a K dimensional vector representing assets returns. It has mean ρ (K×1) and variance Σ (K×K). We then construct portfolio return by multiplying returns vector by the weights assigned to each of them. In matrix notation:

$$w'R_t$$

Where w (K×1) is the vector of portfolio weights. In this case we want to determine optimal portfolio weights in order to minimize portfolio variance:

$$\min_w w'\Sigma w$$

Under the following conditions:

$$w'\rho = \mu$$

$$w'1_K = 1$$

Where μ is the desired level of returns. We then set the Lagrangian:

$$w'\Sigma w - \lambda_1(\rho' w - \mu) - \lambda_2(1_K' w - 1)$$

$$w'\Sigma w - (w'B - a')\lambda$$

Where:

$$B = [\rho, 1_K]$$

$$a = \begin{bmatrix} \mu \\ 1 \end{bmatrix}$$

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

The first order conditions are equal to:

$$2\Sigma w - B\lambda = 0_K$$

$$(w'B - a')' = 0_2$$

Then, making computations, optimal portfolio weights are equal to:

$$w = \Sigma^{-1}B[B'\Sigma^{-1}B]^{-1}a$$

$$= \Sigma^{-1}[\rho, 1_K]\left[\begin{bmatrix}\rho' \\ 1_K'\end{bmatrix}\Sigma^{-1}[\rho, 1_K]\right]^{-1}\begin{bmatrix}\mu \\ 1\end{bmatrix}$$

However, if we consider the simpler case in which $B = 1_K$ and $a = 1$, we obtain the following result:

$$w = \Sigma^{-1}1_K[1_K'\Sigma^{-1}1_K]^{-1}$$

$$= \frac{\Sigma^{-1}1_K}{[1_K'\Sigma^{-1}1_K]}$$

During the years, many critics has been moved toward mean-variance paradigm. These can be synthetized in three main point. First, Markowitz formulation regards only single-period portfolio choice problem. Second, it embodies a utility maximization only for investors whose preferences are represented by a quadratic utility function. Consequently, preferences toward higher-order return moments are not considered.

1.1.2 Intertemporal utility maximization in discrete and continuous time

In the previous paragraph I have illustrated Markowitz paradigm and how it is focused only on a single-period portfolio choice problem. Now, I will go through a more generical formulation of a portfolio choice problem, which considers a multiperiod time horizon.

In the dynamic asset allocation framework, we can distinguish between discrete and continuous time. I will start with the case of a discrete time formulation. We are in a discrete time setting when the investor has a well defined timeline for its investment opportunity: it begins at time zero and ends at time T; there are some intermediate dates in which the investor has the possibility to rebalance his portfolio. Now, consider an investor who wants to maximise its expected utility over wealth by investing in a risk-free asset an in risky assets. The time horizon is $t + \tau$. Then, we can state this problem as:

$$V(\tau, W_t, z_t) = \max_{\{w_s\}_{s=t}^{t+\tau-1}} E_t[u(W_{t+\tau})]$$

under the condition of the budget constraint:

$$W_{s+1} = W_s\left(w_s' r_{s+1} + R_s^f\right)$$

$V(\tau, W_t, z_t)$ is called value function and represents the expectations of the investor at time t (the starting point of the investment) regarding the utility that he will get at the end of the time horizon. These expectations are influenced by the initial level of wealth W_t, and by z_t which denotes the information set available at the time. The value function can be interpreted as an assessment of the quality of the investment: if it is high, the opportunity is good; otherwise it is not worth it. $u(W_{t+\tau})$ is the general notation used for the utility function. Depending on the function we choose to use, we substitute it with the corresponding specific formulation (e.g. CRRA, logarithmic etc.). The budget constraint must be positive every period and indicates the amount of accessible resources for each period. The time subscript s denotes intermediate dates in which is possible to rebalance the portfolio. w_s' are portfolio weights and r_{s+1} and R_s^f are returns of risky and risk-free assets respectively.

We can then rewrite the multiperiod portfolio choice problem above as a single-period one:

$$V(\tau, W_t, z_t) = \max_{\{w_s\}_{s=t}^{t+\tau-1}} E_t[u(W_{t+\tau})]$$

$$= \max_{w_t} E_t\left[\max_{\{w_s\}_{s=t+1}^{t+\tau-1}} E_{t+1}[u(W_{t+\tau})]\right]$$

$$= \max_{w_s} E_t\left[V\left(\tau - 1, W_t\left(w_t' r_{t+1} + R_t^f\right), z_{t+1}\right)\right]$$

Such that:

$$V(0, W_{t+\tau}, z_{t+\tau}) = u(W_{t+\tau})$$

The function obtained is said Bellman equation and is fundamental for the solution of any dynamic portfolio choice problem. The FOCs for its solution are derived from the following expression:

$$E_t\left[V_2\left(\tau - 1, W_t\left(w_t' r_{t+1} + R_t^f\right) z_{t+1}\right) r_{t+1}\right] = 0$$

V_2 denotes the partial derivative with respect to the second argument of the value function.

When the FOCs of the multiperiod case are different form the ones of the single period because excess returns are not independent from innovations hedging demands arise. This happens because deviating from the single-period portfolio choice, the investors tries to hedge against changes in the investment opportunities.

On the other side, we have portfolio choices in continuous time. In this case the investor must take investment decisions regarding portfolio choice constantly because portfolio is rebalanced every instant.

The starting point is, as before, the following statement:

$$V(\tau, W_t, z_t) = \max_{\{w_s\}_{s=t}^{t+\tau-1}} E_t[u(W_{t+\tau})]$$

In this case we must assume that risky asset prices p_t and the vector representing the information set z_t evolve together as correlated Itô vector process:

$$\frac{dp_t}{p_t} - rdt = \mu^p(z_t, t)dt + D^p(z_t, t)dB_t^p$$

$$dz_t = \mu^z(z_t, t)dt + D^z(z_t, t)dB_t^z$$

r is instantaneous risk free rate, while μ^p and μ^z denote conditional mean vectors, B_t^p and B_t^z Brownian motion processes and D^p and D^z conditional diffusion matrices. The budget constraint takes the form:

$$\frac{dW_t}{W_t} = \left(w_t'\mu_t^p + r\right)dt + w_t'D_t^p dB_t^p$$

Consequently, the Bellman equation is equal to:

$$0 = \max_{w_t}\left[V_1(\cdot) + W_t\left(w_t'\mu_t^p + r\right)V_2(\cdot) + \mu_t^{z\prime}V_3(\cdot) + \frac{1}{2}W_t^2 w_t'\Sigma_t^p w_t V_{22}(\cdot) + W_t w_t' D_t^p \rho_t' D_t^{z\prime} V_{23}(\cdot) + \frac{1}{2}tr[\Sigma_t^z V_{33}(\cdot)]\right]$$

Where ρ_t' is the correlation matrix of B_t^p and B_t^z, and Σ_t^p and Σ_t^z denote covariance matrices:

$$\Sigma_t^p = D_t^p D_t^{p\prime}$$

$$\Sigma_t^z = D_t^z D_t^{z\prime}$$

FOCs are derived from the following equation:

$$\mu_t^p V_2(\cdot) + W_t w_t' \Sigma_t^p V_{22}(\cdot) + D_t^p \rho_t' D_t^{z\prime} V_{23}(\cdot) = 0$$

Optimal portfolio weights are:

$$w_t^* = -\frac{V_2(\cdot)}{W_t V_{22}(\cdot)}\left(\Sigma_t^p\right)^{-1}\mu_t^p - \frac{V_2(\cdot)}{W_t V_{22}(\cdot)}\frac{V_{23}(\cdot)}{V_2(\cdot)}\left(\Sigma_t^p\right)^{-1} D_t^p \rho_t' D_t^{z\prime}$$

This solution can be split in two components, myopic demand and hedging demand respectively. The first part of the expression is the one that summarizes optimal portfolio weights for a myopic portfolio choice. A portfolio choice is said to be myopic when the investment horizon is irrelevant, and the investor doesn't know what will happen in subsequent future periods. On the contrary, the second portion is given by the difference between a dynamic problem and a myopic one.

If we take the limit to zero of the continuous time Bellman equation, we obtain its discrete time counterpart. However, the two results do not share the same properties because continuous time policies cannot always assure that the wealth is positive.

I have so far analysed the formalization of a portfolio choice problem in both discrete and continuous time cases. Although, there could be some cases in which for the investor is more convenient to invest myopically. We can recognize four cases that fall into this category:

- Single-period horizon
- Logarithmic utilities
- Constant investment opportunities
- Stochastic investment opportunities that cannot be hedged

The first instance is an obvious one, and it does not require further explanation.

When an investor has preferences that can be expressed through a logarithmic utility function, the utility that he will get from the terminal wealth will be exactly equal to the one derived from the sum of the utilities of returns for each period. So, this case can be in this way reconducted to the one of a myopic portfolio choice.

As I said before explaining discrete-time portfolio choice problem, hedging demands arise when the investor tries to hedge against changes in investment opportunities. When investment opportunities are the same for each time period, the value function does not depend anymore on the changes in these opportunities and there are no hedging demands. Consequently, we go back to the case of a myopic portfolio choice. In the investable universe there are constant investment opportunities in two cases:

- Assets returns are i.i.d.
- Conditional moments of returns are stochastic.

When we have excess returns that are independent from the innovations the investor cannot use the available assets to hedge, so there are no hedging demands and the portfolio choice is again myopic.

1.2 Traditional econometric approaches

In portfolio choice framework, econometrics has the role of defining the data generating process. This can be done in two ways: using plug-in estimation or through decision theory. In plug-in estimation, also called calibration, the econometrician estimates portfolio weights and then he inserts them into an analytical solution in the investor's problem. The purpose is to formulate explanatory statements. On the contrary, in decision theory approach, the econometrician tries to identify the return distribution in order to find optimal portfolio weights.

In the next sections, I will examine with more detail both approaches, starting with plug-in estimation.

1.2.1 Plug-in estimation

Indeed, in case of plug-in estimation the econometrician indicates parameters for the data generating process and then insert them in the investor's problem. The aim is to find, again, optimal portfolio weights.

As done before, we can make a distinction between an investor which has a short-term horizon and another one that looks more forward. To be clearer, I will start with the simpler case of a single-period portfolio choice problem, and then extend to the more general case of a multiperiod choice.

Optimal portfolio weights are a function of risk preferences, of the information set available and of the parameters specified and estimated by the econometrician in the data generating process:

$$w^* = w\left(\phi, z_t, \hat{\theta}\right)$$

We need to assume that the estimator $\hat{\theta}$ is consistent and that its asymptotic distribution is:

$$\sqrt{T}\left(\hat{\theta} - \theta\right) \xrightarrow{T\to\infty} N[0, V_\theta]$$

Consequently, the asymptotic distribution of the optimal portfolio weights estimator can be computed in the following way using the delta method:

$$\sqrt{T}(\hat{w}_t^* - w_t^*) \xrightarrow{T\to\infty} N[0, w_3(\cdot)V_\theta w_3(\cdot)']$$

where $w(\cdot)$ maps the evolution of portfolio weights.

To get better the sense of the former passages, we use the mean-variance framework as exemplification. Considering the excess returns r_{t+1} to be i.i.d., the first step involves the estimation of the first and second moments:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_{t+1}$$

$$\hat{\Sigma} = \frac{1}{T-N-2}\sum_{t=1}^{T}(r_{t+1}-\hat{\mu})(r_{t+1}-\hat{\mu})'$$

Then we insert these estimates in the numerical solution. Thus, optimal portfolio weights would be equal to:

$$\hat{w}^* = \left(\frac{1}{\gamma}\right)\hat{\Sigma}^{-1}\hat{\mu}$$

As we can see the optimal portfolio weights depend much from asymptotic approximations. They also inherit estimation errors of parameter estimates.

This problem becomes more relevant in case of a multiperiod portfolio choice. In fact, when the investor has a longer-term investment horizon, the estimation error of the first time period is passed from one period to the subsequent ones and it accumulates until the last period considered.

We can easily see this outcome from the following example. We consider optimal portfolio weights as mapped by a set of recursive function, as before dependent from risk preferences, information set, parameter estimates and, in this case, also by time:

$$w^*_{t+\tau-1} = w(1, \phi, z_{t+\tau-1}, \theta)$$

$$w^*_{t+\tau-2} = w(2, \phi, z_{t+\tau-2}, \theta, w^*_{t+\tau-1})$$

$$w^*_{t+\tau-3} = w(3, \phi, z_{t+\tau-3}, \theta, \{w^*_{t+\tau-1}, w^*_{t+\tau-2}\})$$

$$\vdots$$

$$w^*_{t} = w(\tau, \phi, z_{t}, \theta, \{w^*_{t+\tau-1}, \dots, w^*_{t+1}\})$$

Then, asymptotic standard errors of portfolio weight estimates are equal to:

$$\sum_{s=t+1}^{T-1} \frac{\partial w(t, \phi, z_t, \theta, \{w_s^*\}_{s=t+1}^{T-1})}{\partial w_s^*} \frac{\partial w_s^*}{\partial \theta}$$

This expression confirms the fact that, due to the recursive nature of the process, estimation errors cumulate through time periods.

This loss in precision derived form pug-in estimation can be measured by using certainty equivalent loss:

$$CE\ loss = CE - E\left[\widehat{CE}\right]$$

Where

$$CE = w^{*\prime}\mu - \frac{\gamma}{2} w^{*\prime}\Sigma w^*$$

$$\widehat{CE} = \widehat{w}^{*\prime}\mu - \frac{\gamma}{2} \widehat{w}^{*\prime}\Sigma \widehat{w}^*$$

So, the loss is given by the difference of certainty equivalent calculated with true portfolio weights and with estimated portfolio weights. It is easy to understand that these values strongly depend on risk aversion, covariance matrices of estimates and returns.

During the years financial literature has underlined the fact that the use of asymptotic results in plug-in estimation cannot lead always to reliable outcomes because samples considered for the estimation are finite. Thus, three main solution have been proposed in order to augment plug-in estimates: portfolio constraint, factor models and shrinkage estimation.

The imposition of constraint on portfolio weights improves sample estimates by limiting their range of variation and eliminating more extreme ones. There are plenty of constraint that can be set, however the most common are limits to short selling activity, maximum positions in securities of a determined sector or industry or boundaries in the level of borrowing investable in risky assets.

The second way of enhancing results of plug-in estimation is through the use of a factor structure for the assets covariation. This action has the result of reducing the number of free parameters of the matrix. In order to capture all elements of the covariation matrix, a multifactor model is considered:

$$r_{i,t} = \alpha_{i,t} + \beta_i' f_t + \varepsilon_{i,t}$$

β_i' represents the vector of factor loading, f_t denotes the generic expression for factor realization and $\varepsilon_{i,t}$ uncorrelated residuals.

Consequently, the return covariance matrix is obtained:

$$\Sigma = B\Sigma_f B' + \Sigma_\varepsilon$$

Where B is the matrix of stacked factor loadings.

The main problem is this case is represented by the choice of factors. Financial literature suggests three possible solutions to this issue. The first answer is to get them directly from returns by using a proper statistical procedure. The other ones suggest making a choice based on an empirical work or economic theory.

The last solution identified in order to reduce the estimation error in plug-in estimation is shrinkage estimation. In this case, by using proper estimators, estimates are induced to converge toward a common value thus reducing extreme estimation errors. The first estimator identified was the one for sample means, however later estimators were calculated also for the sample covariance matrix and directly for optimal portfolio weights. The estimator for the sample mean is equal to:

$$\mu_s = \delta\mu_0 + (1-\delta)\bar{\mu}$$

Where δ must be included between zero and one. The optimal shrinkage factor δ results from the following expression:

$$\delta^* = min\left[1, \frac{(N-2)/T}{(\bar{\mu}-\mu_0)'\Sigma^{-1}(\bar{\mu}-\mu_0)}\right]$$

The estimator for the covariance matrix is instead:

$$\hat{\Sigma}_s = \delta\hat{S} + (1-\delta)\hat{\Sigma}$$

Where $\hat{S}$ is the optimal shrinkage target. In this case the optimal shrinkage factor corresponds to:

$$\delta^* \simeq \frac{1}{T}\frac{A-B}{C}$$

Where:

$$A = \sum_{i=1}^{N} \sum_{j=1}^{N} asy\ var\left[\sqrt{T}\hat{\sigma}_{i,j}\right]$$

$$B = \sum_{i=1}^{N} \sum_{j=1}^{N} asy\ cov\left[\sqrt{T}\hat{\sigma}_{i,j}, \sqrt{T}\hat{s}_{i,t}\right]$$

$$C = \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\hat{\sigma}_{i,j} - \hat{s}_{i,t}\right)$$

Appling shrinkage estimation directly on optimal portfolio weights could be an advantageous solution because it is easier to specify particular objectives of the shrinkage procedure. Furthermore, a direct link between mean and variance is created. In this case the estimator is:

$$\hat{x}_s^* = \delta x_0 + (1 - \delta)\hat{x}^*$$

Where x_0 is the shrinkage target.

1.2.2 Decision theory

The alternative to plug-in estimation, where the econometrician specifies some parameters that will be estimated through the process, is to bypass the step of finding optimal parameters and try to determine return distribution by exploiting personal beliefs and data observed. This is what is done with decision theory approach.

In this framework we need to rewrite the investor's utility maximization problem as:

$$\max_{w_t} \int u(w_t' r_{t+1} + R^f) p(r_{t+1}|Y_T) dr_{t+1}$$

where Y_T denotes data available to the investor. $p(r_{t+1}|Y_T)$ is the return distribution identified by the investor. It is subjective because it is determined on the base of his prior beliefs. This is equal to write:

$$p(r_{t+1}|Y_T) = \int p(r_{t+1}|\theta) p(\theta|Y_T) d\theta$$

However, in order to solve it, we need to know posterior distribution of the parameters as well. They are obtained through the Bayes' theorem:

$$p(\theta|Y_T) = \frac{p(Y_T|\theta)p_0(\theta)}{p(Y_T)} \propto p(Y_T|\theta)p_0(\theta)$$

Thus, the maximization problem can be rewritten as:

$$\max_{w_t} \int \left[\int u(w_t' r_{t+1} + R^f) p(r_{t+1}|\theta) dr_{t+1} \right] p(\theta|Y_T) d\theta$$

As we can see, just having knowledge of data and prior beliefs of the investor we can find optimal portfolio weights based on the determined return distribution.

The choice of the type of priors ends up being fundamental in the Bayesian approach. In financial literature priors can be of two kinds: uninformative and informative. Uninformative priors are the ones which incorporate very little information about parameters. On the contrary, informative priors include subjective information. Although both of them have their area of application, uninformative priors are most widely used in empirical applications because they are more comparable to traditional statistical approaches. In fact, the principal problem when using informative priors is to keep analytical tractability of posterior distribution. This problem is solved by econometricians using conjugate priors, which are helpful for updating posteriors with new information.

An alternative way of specifying priors is to determine them on the base of theoretical outcomes of an economic model. This can be done through the mixed estimation method, which more generally allows to combine two sets of data into a single posterior distribution.

Supposing i.i.d returns and starting beliefs expressed as:

$$r_{t+1} \sim MNV[\mu, \Sigma]$$

$$p(\mu) = MNV[\bar{\mu}, \Lambda]$$

Then, the new information set is included generating the following distribution:

$$p(v|\mu) = MNV[P\mu, \Omega]$$

Where v are new information, P is a matrix which combines returns and Ω denotes error covariance matrix.

The three most common application which include economic view in Bayesian framework are Black-Litterman model, the return forecasting with belief in no predictability formulated by Kandel, Stambaugh and Connor, and belief in asset pricing model formulated by Pastor. The most famous is probably the Black-Litterman model, in which benchmark beliefs are obtained by inferring risk premia that would induce an investor to hold assets in proportion of market capitalization in the mean variance framework.

So far, I have tried to explain how parameters uncertainty is tried to be ruled out. However, the same path could be followed in order to solve model uncertainty on which portfolio choice model to choose in order to get the best result. In this case we would have:

$$p(M_j|Y_T) = \frac{p(Y_T|M_j)p(M_j)}{\sum_{j=1}^{J} p(Y_T|M_j)p(M_j)}$$

And

$$p(Y_T|M_j) = \int p(Y_T|M_j,\theta_j)p(\theta_j|M_k)d\theta_j$$

This finds its best application when model the uncertainty about the true model is very high.

1.3 Alternative econometric approach

As we have seen, traditional econometric approaches generally involve a path composed by two main steps. The first one in focused on the estimation of parameters of the data generating process or of the return distribution, while the second one is dedicated to the calculation of optimal portfolio weights on the base of results derived from the former step. Thus, there is a high dependence of the applied model outcome on inference made in the first part of the process.

The most recent stream of literature proposes new approaches in order to avoid or at least reduce this reliance. Inside this flow we can identify two principal groups of methodologies: the ones which uses parametric portfolio weights and the others which exploit nonparametric portfolio weights.

In the following sections I will go into details of each category, presenting to the reader, as examples, the principal models formulated by Brandt et al.

1.3.1 Parametric portfolio weights

The technique of parametric portfolio weights allows the econometrician to skip the first step of traditional approaches by modelling optimal portfolio weights as a function of observable quantities (Brandt (2010)) and then solve this function in order to maximize utility.

One of the models that follows this path is the one formulated by Brandt and Santa-Clara in 2006. This approach is aimed at a dynamic portfolio selection by extending the asset space through the inclusion of mechanically managed portfolios. In this augmented asset space are included two types of portfolios: conditional portfolios, which invest a sum proportional to the state variable considered, and timing portfolios which invest one period in a risky asst and the other ones in a risk-free security. They demonstrate that this is equal to solving a conditional problem with parametrized weights.

This is just one of the recent econometric approaches which involve parametric portfolio weights. Thus, the following chapter of this work will be completely dedicated to one of those, the approach formulated by Brandt et al. in 2009.

1.3.2 Nonparametric portfolio weights

The stream of nonparametric portfolio weights starts form the fact that when using parametric portfolio weights there could be made some mistakes when modelling the weight function.

Brandt's nonparametric approach determines optimal portfolio weights as the ones which satisfy sample analogues of Euler equation or FOCs, in which the conditional expectations are replaced by nonparametric regression.

Considering a single period portfolio choice, we start from the following Euler equation:

$$E_t\left[u'\left(w_t' r_{t+1} + R_t^f\right) r_{t+1}\right] = 0$$

Supposing i.i.d. returns and constant weights through different states, and replacing conditional expectations, we obtain the following estimator for portfolio weights:

$$\widehat{w} = \left\{ w : \frac{1}{T} \sum_{t=1}^{T} u'\left(w_t' r_{t+1} + R_t^f\right) r_{t+1} = 0 \right\}$$

Removing the previous assumptions on weights and returns, the estimator is equal to:

$$\widehat{w}(z) = \left\{ w : \frac{1}{T h_T^K} \sum_{t=1}^{T} \omega\left(\frac{z_t - z}{h_T}\right) u'\left(w_t' r_{t+1} + R_t^f\right) r_{t+1} = 0 \right\}$$

As we can see, in this case there is a dependence of the final result on the macroeconomic state variable. ω denotes the kernel function, which weights marginal utility maximizations and h_T is the sequence of kernel bandwidths.

1.4 Which model is the best one? Empirical considerations

At the beginning of this chapter I have stated that we cannot identify, from a theoretical point of view, a model that is preferable to the others because each of them is based on different assumption and has different implications. However, when required to make a choice, we can base on empirical findings present in literature.

Thus, in this last section I move from the theoretical analysis made so far toward more empirical considerations, in order to assess strengths and weaknesses of each model.

Table 1.1 reports the main empirical evidences emerged in literature in recent years, in particular in the research carried out by DeMiguel et al. (2009).

Among all theoretical formulations previously exposed, I will now take into consideration ten asset allocation models: an equally weighted portfolio, the classical mean-variance approach formulated by Markowitz, Bayesian approaches to estimation error (with the use of diffuse priors, Bayes-Stein shrinkage portfolio, and beliefs determined by asset pricing models), moment restrictions (minimum variance portfolio, and value weighted portfolio), imposition of constraints on the portfolio policy, the combination of portfolio strategies and, in the end, the parametrization of portfolio weights formulated by Brandt. However, this last one will be further analysed in the subsequent chapters as already anticipated.

The equally weighted portfolio strategy is the easiest to implement because it assigns the same weight to every asset present in the portfolio. However, beyond its apparent simplicity, it is generally the best performing in comparison to other optimizing strategies considered by DeMiguel et al. when trading on the stock market. Indeed, it has the highest Sharpe ratio, the best CER and the lowest turnover. This is mainly a consequence of the fact that it does not rely on estimation of parameters such as mean and variance as other alternative strategies. Definitely, the largest errors are usually made when estimating the mean. This superiority remains also when both mean and variance are unknown.

As stated before, there are many critical issues linked to the correct application of the mean-variance framework formulated by Markowitz. One of those regards the estimation of the mean and variance matrices, in which a large estimation error can be made. If no corrections are made in order to deal with this estimation error, the portfolio strategy is not able to beat the naïve 1/N. However, when the mean needs to be estimated and the variance is known, Markowitz's strategy is more likely to perform

	Asset allocation mod.	Advantages	Disadvantages
	1/N	It has the lowest turnover. It is the best performing strategy when mean and variance are both unknown. It has the best CEQ. It beats dynamic strategies.	
Classical approach	Sample-based mean-variance (Markowitz)	If mean is unknown/known and variance is known/unknown, it is more likely to beat 1/N when the no. of periods is high and the no. of assets is low. Its performance improves if the idiosyncratic asset volatility is higher than 20%.	It is no able to outperform the 1/N if no corrections are made for dealing with the estimation error. It has the highest level of turnover. It requires long estimation windows.
Bayesian approach to estimation error	Diffuse priors	It reduces turnover compared to Markowitz approach.	Not effective at dealing with estimation error.
	Bayes-Stein shrinkage portfolio	It has a higher Sharpe ratio compared to Markowitz model. It reduces turnover compared to Markowitz approach.	Not effective at dealing with estimation error.
	Bayes beliefs determined by asset pricing models	It reduces turnover compared to Markowitz approach.	Not effective at dealing with estimation error.
Moments restrictions	Minimum variance portfolio	It performs better than Markowitz approach. It has a lower turnover compared to the Bayesian approach and Markowitz model. It exploits info about correlations leading to a reduction in extreme portfolio weights.	It doesn't beat the 1/N strategy. It ignores estimates of expected returns
	Value weighted portfolio	It has a null turnover because the investor holds the market portfolio.	
	Portfolio constraints	No short sale constraint reduces turnover consistently.	They improve performance only for small estimation windows.
	Combination of portfolio strategies	They are more effective in reducing the estimation error. They exploit the correlation structure to reduce risk.	They ignore data on expected returns.
	Parametric portfolio weights (Brandt et al.)	It reduces statistical estimation problem. It has a higher performance and a higher turnover which grows with the number of assets. Using info on stock characteristics improves performance.	

Table 1.1 Source: Self-elaboration on De Miguel V., L. Garlappi and R. Uppal (2009) data

better than the 1/N strategy if the number of periods is high and the number of assets is low. This happens because the number of parameter estimates grow with the number of assets considered, and so does the estimation error; while a wider estimation window reduces it because the econometrician has more elements for making computations. Other than the longer estimation period required, another weakness of the mean-variance framework is the high level of turnover, which in the presence of transaction costs can erode an optimal performance. Furthermore, the presence of an idiosyncratic volatility of assets higher than 20% can improve the performance of this strategy.

The Bayesian approaches have the advantage to lead to a reduction in turnover and in an increase in the performance in comparison to the Markowitz approach. However, they are not yet effective in dealing with the estimation error.

Among asset allocation models with moments restrictions, I consider a value weighted portfolio and a minimum variance portfolio. The first assigns weights in proportion to market capitalization, while the other takes the weights associated with the portfolio which has the lowest variance. The value weighted portfolio has the strength of being the one with the lowest turnover, because the investor holds the market portfolio, and thus he does not trade at all. On the other side, the minimum variance strategy has a higher turnover compared to the value weighted but this is lower than the one of the Markowitz approach. It has also a better performance among other strategies, but not enough to outperform the naïve one. Although, the minimum variance strategy has also the great advantage of leading to a reduction in extreme portfolio weights through the exploitation of information about correlations.

Generally, the imposition of a constraint on a portfolio strategy has the effect of improving the performance of the portfolio only if short estimation windows are considered. Otherwise, limitations in trading, such as the no short sale constraint, has the only consequence of reducing turnover consistently. Although, this could be an advantage in the presence of a high level of transaction costs.

On the other side, combinations of different asset allocation methods are the most effective in dealing with the estimation error even if they generally ignore data on expected returns. Anyway, these approaches are able to exploit the correlation structures in order to reduce the overall risk of the portfolio.

Nevertheless, among all approaches considered, the one with parametric portfolio weights formulated by Brandt et al. is the only one which can outperform the naïve 1/N strategy. In fact, by using information on stock characteristics, it improves its performance and reduces the estimation error. The trade-off is the fact that it leads to an increase in asset turnover.

These evidences, as the rest of this work, are referred to the stock market, because, as highlighted by Ackermann et al. (2017), in the currency markets other portfolio optimization strategies lead to a better performance compared to the one of the naïve 1/N strategy.

Chapter 2

Brandt's model: a parametric approach for portfolio weights estimation

In this chapter I will illustrate in a more detailed way Brandt et al.'s model formulated in 2009, beginning from its fundamental through its extensions and highlighting its strengths and weaknesses. As we will see, this is a perfect example of an alternative econometric approach based on parametric portfolio weights.

The conclusion drawn so far is that Markowitz's model leaves an open point. In fact, even if the mean-variance approach is considered as traditional and it is the most well-known, it sometimes involves computational issues that can hardly be solved by a simple investor; it would rather be implemented by a professional. Indeed, it requires the asset manager to model expected returns, variances and covariances to be applied. Furthermore, it is focused on a single period horizon where the investor cannot do anything but waiting for the investment to end, without having the possibility of revising his investment decision at intermediate dates.

In order to solve these issues Brandt, Santa-Clara and Valkanov have formulated in 2009 an alternative parametric approach which involves modelling optimal portfolio weights as a function of firm characteristics. These characteristics can be various, such as one-year lagged return, market capitalization, book-to-market ratio, asset growth, leverage and so on. The choice of course depends mainly on the type of market and the aim of the empirical application. Then, the coefficients are estimated by maximizing the utility function of the potential investor.

At the beginning of their paper, the authors have underlined four main advantages of their alternative parametric approach. First, it gives a direct focus on portfolio weights escaping the long and tedious step of modelling returns and characteristics, which is central in traditional econometric approaches. Second, it encapsulates in an implicit way the connection between returns, characteristics, variances and covariances. Third, it avoids the frequent problem in statistical computations of overfitting or not precise coefficient estimates, because it requires only the estimation of N weights, where N is the number of assets considered in the portfolio every period. Finally, it allows to easily test hypotheses, whether they are individual or joint.

2.1 Methodology: a simple linear policy

In an investment universe in which there is the possibility to allocate funds in N stocks in each period, the investor has to choose weights in order to solve the following problem of utility maximization:

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} E_t\left[u(r_{p,t+1})\right] = E_t\left[u\left(\sum_{i=1}^{N_t} w_{i,t}\, r_{i,t+1}\right)\right]$$

$$w_{i,t} = f(x_{i,t};\theta)$$

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t}\theta^T \hat{x}_{i,t}$$

This is the case in which portfolio weights are a linear function of θ and $\hat{x}_{i,t}$, but this application can be extended to more general functions. $r_{p,t+1}$ represents portfolio returns and $r_{i,t+1}$ are stock returns. $\bar{w}_{i,t}$ is the matrix of portfolio weights assumed as a benchmark. Usually it is an equally weighted portfolio. $1/N_t$ is the term that allows for the number of stocks to vary in time. θ ($F_t \times 1$) is the vector of parameter estimates, which is representative of the weight assigned to each firm characteristic selected. These θs are constant in time because otherwise they would be influenced also by past return history and not only by firm-specific characteristics. $\hat{x}_{i,t}$ ($N_t \times F_t$) is the matrix of normalized firm characteristics. As stated before, these can be various, from book-to-market ratio or market capitalization to annual asset growth. N_t represents the number of asset and F_t is the number of characteristics chosen. For every period, the following inequality must be valid: $N_t > F_t$. The multiplication of the two matrices, $\theta^T \hat{x}_{i,t}$, embodies the deviations from the benchmark of optimal portfolio weights. This product can be interpreted also as the adjustment made to the benchmark portfolio in order to improve its performance.

In Brandt et al.'s opinion, the standardization process of the characteristics brings two main advantages. First, the cross-sectional average of the deviation from the benchmark portfolio is zero, making the sum of optimal weights equal to one. Second, data of characteristics could be non-stationary, while standardized characteristics are granted to be stationary.

It is worth noting that with the application of this model the complexity in making computations doesn't grow with the number of stocks included in the optimal portfolio, but with the amount of characteristics considered.

2.2 Objective function

Brandt et al.'s model can be applied to a wide range of utility functions, even the ones affected by investor's behaviours. The only prerequisite that it is imposed is that the utility function must have a unique solution.

As we have seen, in financial literature, when modelling a portfolio choice, the assumptions made regarding the utility function over wealth are very important. We can choose among three main categories of functions, depending on the prerequisites set:

- Exponential utility family
- Power utility family
- Quadratic utility

The exponential utility function takes the form:

$$U(W_{t+1}) = -\exp\left(-\theta * W_{t+1}\right)$$

where W_{t+1} represents wealth and θ is the coefficient of absolute risk aversion, which embodies the absolute monetary amount that an investor is willing to pay in order to avoid taking on risk of a gamble of an absolute dimension. In this case the coefficient θ is a constant and returns are assumed to be normally distributed. Relative risk aversion, which is the proportion of wealth that an investor is willing to pay in order to avoid a bet of a given size relative to risk, is increasing in wealth.

Power utility functions take the form:

$$U(W_{t+1}) = \frac{W_{t+1}^{1-\gamma} - 1}{(1-\gamma)}$$

Where W_{t+1} is again wealth and γ is the constant coefficient of relative risk aversion. Here asset returns are assumed to be lognormally distributed and absolute risk aversion is declining in wealth. In the limit case of $\gamma = 1$, the power utility function turns into a log utility one. The most commonly used power utility function is Constant Relative Risk Aversion (CRRA) function.

Finally, we have quadratic utility, which is represented by the following expression:

$$U(W_{t+1}) = a + b * W_{t+1}$$

Both absolute and relative risk aversion grow as wealth goes up. No distributional assumptions are made.

In any case, the curvature of the utility function is an indicator of the level of risk aversion of the investor.

Each of these categories has its advantages and disadvantages. For example, CRRA function has the benefit of incorporating preferences toward higher moments of the distribution, without the addition of new parameters on preferences, it is optimal under a partial myopic behaviour of the investor and it is twice continuously differentiable, which is a plus in an optimization process. On the other side, there are several weaknesses as well. First, some CRRA could not be able to find the global optimum, for example when the problem is not well stated. Second, in some cases these functions, especially with characteristics-based portfolios (as pointed out by Ammann et al. (2016)), lead the investor to take on extreme positive or negative position in portfolio weights. Third, there are cases in which expected utility does not take finite values.

On the other side, quadratic utility functions have the strength of allowing for closed-form solution that are able to guarantee the identification of global maxima.

In my empirical application, I decided to use CRRA utility function, following the path set by Brandt et al. in most of their applications. So, standard CRRA utility over wealth through this paper would take the form:

$$u(r_{p,t+1}) = \frac{(1 + r_{p,t+1})^{1-\gamma}}{1-\gamma}$$

However, this is a discretionary choice. For example, Ammann et al. (2016) in their application of the model decide to use a quadratic utility function.

2.3 Statistical inference

As already said before, one of the most important advantages of Brandt et al.'s model is the possibility to test individual or joint hypotheses on optimal portfolio weights in a simple way. This is due to the fact that the optimal thetas estimated can be interpreted as a generalized method of moment estimator. In fact, the portfolio policy can be stated also as:

$$\max_{\theta} \frac{1}{T} \sum_{t=0}^{T-1} u(r_{p,t+1}) = \frac{1}{T} \sum_{t=0}^{T-1} u\left(\sum_{i=1}^{N_t} (\bar{w}_{i,t} + \frac{1}{N_t} \theta^T \hat{x}_{i,t}) * r_{i,t+1}\right)$$

which is subject to the first order condition:

$$\frac{1}{T} \sum_{t=0}^{T-1} h(r_{t+1}, x_t; \theta) \equiv \frac{1}{T} \sum_{t=0}^{T-1} u'(r_{p,t+1}) * \left(\frac{1}{N_t} * \hat{x}_t^T r_{t+1}\right) = 0$$

Consequently, the covariance matrix of optimal estimates can be computed in two ways:

- As asymptotic covariance matrix
- With bootstrap technique.

The first method exploits the property of the method of moments estimator:

$$\Sigma_\theta \equiv AsyVar[\hat{\theta}] = \frac{1}{T}[G^T W^{-1} G]^{-1}$$

where:

$$G \equiv \frac{1}{T}\sum_{t=0}^{T-1} \frac{\partial h(r_{t+1}, x_t; \theta)}{\partial \theta} =$$

$$= \frac{1}{T}\sum_{t=0}^{T-1} u''(r_{p,t+1})\left(\frac{1}{T}\hat{x}_t^T r_{t+1}\right)\left(\frac{1}{T}\hat{x}_t^T r_{t+1}\right)^T$$

The consistent estimator of the covariance matrix of $h(r, x; \theta)$, W, can be estimated in two ways. In the first case we have:

$$W = \frac{1}{T}\sum_{t=0}^{T-1} h(r_{t+1}, x_t; \hat{\theta})\, h(r_{t+1}, x_t; \hat{\theta})^T$$

This equation can be used if marginal utilities are uncorrelated by assumptions.

The alternative method implies using the Newey-West procedure, that produces an estimator which is consistent even in the presence of autocorrelation and heteroscedasticity. It involves the selection of a lag length in order to identify the number of lagged residuals used to evaluate autocorrelation. According to the authors, the procedure is asymptotically equal to the one that is optimal under a mean-squared error loss function; however, it could result is size distortions. It can be computed as:

$$\widehat{W}_t^{NW}(\theta) = \Gamma_{lT}(\theta) + \sum_{l=1}^{L}\left(1 - \frac{l}{L+1}\right)\Gamma_{lT}(\theta) + \Gamma_{lT}(\theta)'$$

where:

$$\Gamma_{lT} = \frac{1}{T}\sum_{i=1}^{T-l} h(r_t, x_t; \hat{\theta})\, h(r_{t+1}, x_t; \hat{\theta})$$

On the other side, the bootstrap technique implies the creation of many random samples of returns and characteristics from the original sample. After the estimation of optimal thetas, the covariance matrix of coefficients must be estimated from any bootstrapped subsample created. This method has the advantage to capture features of data that are non-normal, and it does not rely on asymptotic results as the former case. This can be a useful property for example in countries such as emerging ones where samples available to make analysis are not very big.

I have decided to use in my application Newey-West estimator in the asymptotic covariance framework, because it seemed to me the most suitable one for the aim of my research. Although, Brandt et al. use the bootstrap technique across their paper.

2.4 Extensions to the simple linear policy

Until now I have exposed the basic application of the model, the so called simple linear policy, and its fundamental assumptions. However, the authors have included in their paper some possible extensions, refinements and restriction that can be imposed to the parametric portfolio choice framework. Now I will focus on the two that I have implemented in my empirical application. However, in order to provide the reader with a better understanding of the model, the last subsection is dedicated to the other one included in the original paper which exposes the improvement made to remove the assumption of constant coefficient through time.

2.4.1 Transaction costs

Transaction costs are all those expenses related to the stock exchange, fees, or more broadly to the trading activity. These can be of various entity and are not constant from country to country.

However, transaction costs can be a relevant part when computing portfolio returns in a determined portfolio policy. As such, they are inserted in the theoretical computation in the following way:

$$r_{p,t+1} = \sum_{i=1}^{N_t} w_i r_{i,t+1} - c_{i,t}\left|w_{i,t} - w_{i,t-1}\right|$$

and

$$T_t = \sum_{i=1}^{N_t}\left|w_{i,t} - w_{i,t-1}\right|$$

T_t denotes asset turnover. It represents the absolute amount of exchanges made from one time period to the other for each asset. This value is multiplied by $c_{i,t}$, which is a function that synthetises transaction costs. This product is then subtracted for each period from portfolio returns.

As we can see, asset turnover and transaction costs are strictly connected. In fact, the effect of transaction cost on portfolio return can be reduced with a reduction in the level of turnover. This consideration is independent on the way in which these costs are modelled.

As I said, $c_{i,t}$is the generic term that includes transaction costs. Though, they can be modelled in different ways. For example, they can be a constant proportion through time, or approximated directly from liquidity measures of the market. Another possibility, the one that I decided to implement in my empirical application following the authors' path, is to determine them as a function of asset characteristics. I choose to use market capitalization; however, other measures can easily be used. Anyway, transaction costs are usually higher for smaller firms and thinner for bigger ones.

Nevertheless, transaction costs can have a relevant impact on characteristics-base portfolios. In fact, in the presence of trading costs the simple linear policy presented above is far from being optimal.

Furthermore, in the past have other authors like Davis and Norman (1990) have proven that the optimal portfolio policy is characterized by a boundary around the target weight for the risky asset inside which it is not convenient to trade.

Taking these considerations into account, Brandt et al. have decided to find their solution in order to identify the boundary that makes for the investor convenient whether to trade or not. The starting point is the initial portfolio given by the equation of the portfolio policy:

$$w_{i,0} = \overline{w}_{i,0} + \theta^T x_{i,0}$$

After that, they define each period a target portfolio and a hold portfolio:

$$w_{i,t}^t = \overline{w}_{i,t} + \theta^T x_{i,t}$$

$$w_{i,t}^h = w_{i,t-1} \frac{1 + r_{i,t}}{1 + r_{p,t}}$$

At this point, they consider two possible outcomes:

- The hold portfolio is close to the target portfolio
- The hold portfolio is far from the target one.

In the first case, for the investor is better not to trade, and the no-trade region, the distance between the two portfolios, is defined as:

$$w_{i,t} = w_{i,t}^h$$

if

$$\frac{1}{N_t} \sum_{i=1}^{N_t} \left(w_{i,t}^t - w_{i,t}^h \right)^2 \leq\ k^2$$

On the other side, the investor should trade on the edge of the no trade region. The optimal portfolio in this case turns to:

$$w_{i,t} = \alpha_t w_{i,t}^h + (1-\alpha_t) w_{i,t}^t$$

if

$$\frac{1}{N_t}\sum_{i=1}^{N_t}\left(w_{i,t}^t - w_{i,t}^h\right)^2 > k^2$$

where

$$\alpha_t = \frac{k\sqrt{N_t}}{\sqrt{\sum_{i=1}^{N_t}\left(w_{i,t}^t - w_{i,t}^h\right)^2}}$$

The authors underline anyway that this is only an approximation to the optimal solution.

In the following chapter, when I will apply transaction costs to the portfolio strategy, I would just limit their use to the demonstration of how their weight is relevant for the final portfolio return.

Although, transaction costs can, in some cases, even bring some benefits. In fact, recent studies conducted by DeMiguel et al. (2019) have shown that they can influence the number of characteristics that are jointly significant for an investor's optimal portfolio. Indeed, they increase this number. This finding has two main consequences. First, the joint use of characteristic produces a reduction in asset turnover by eliminating superfluous transactions, thus helping in trading diversification. The second, deducted directly from the former point, there is an increase in investor's utility.

2.4.2 Weight constraints

In financial literature, as we have seen for the improvement of plug-in estimates, there are many ways in which constraint can be imposed to portfolio weights. The most common ones are setting a limit in the maximum exposure in a sector, industry or in a single security, a ceiling in the amount of borrowings to invest in a risky asset and restrictions on short positions that an investor can take on. The reason underlying these limitations can be various, from market imperfections to investor's preference in asset management. For example, Brazil, as other emerging markets, forbids naked short selling in order to prevent speculations.

The restraint imposed most frequently is the no short sale constraint, which limits the amount of short positions that an investor can take on by completely cutting negative weights. In practice this restriction is an extreme condition that can be applied for example when there is a high risk of speculation. This is the case when the regulator wants to protect the market.

In Brandt et al.'s paper, the no short sale constraint is imposed in order to try to find a solution for the elevated level of negative weights and, consequently, leverage that the model induces the investor to take on. However, this is of course not the optimal solution because it is the extreme case in which borrowings are completely cancelled out. Indeed, investors generally are most likely to take on intermediate levels of leverage.

Furthermore, the imposition of this restriction requires a renormalization of portfolio weights, because without taking any action and with the cancellation of negative positions, optimal portfolio weights would sum more than one. This can be done in the following way:

$$w_{i,t}^{+} = \frac{max[0, w_{i,t}]}{\sum_{j=1}^{N_t} max[0, w_{j,t}]}$$

A problem arises when imposing this constraint. In fact, the renormalization makes the function of the optimal portfolio policy not differentiable in $w_{i,t} = 0$. As a consequence, we can't use the asymptotic estimator derived form the method of moments. However, this challenge can be easily overcome by using the bootstrap technique.

2.4.3 Coefficient variation through time

As already stated, in the simple linear policy case, the optimal coefficient estimates are supposed to be time invariant. This assumption is essential in order to guarantee that optimal portfolio weights are not affected by the history of past returns.

Although, this hypothesis can be easily surmounted with the introduction of a vector of predictor that models the factors which optimal coefficient are supposed to vary with in the linear portfolio policy equation:

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t}\theta^T\left(z_t \otimes x_{i,t}\right)$$

where z_t denotes the vector of predictor and $\otimes$ the Kronecker product.

2.5 Other alternative econometric approaches

Brandt et al.'s portfolio model is just one of the possible developments of the recent econometric approaches formulated in order to find a solution to the portfolio problem. In this paragraph I will give the reader a brief overview over two alternative approaches formulated in recent years, other than the ones already exposed in the previous chapter.

The first one is a simulation approach framed by Brandt, Goyal, Santa Clara and Stroud. This model aims at solving a discrete-time portfolio choice problem and applies when there is a broad number of assets with arbitrary return distribution and a wide number of state variables with non-stationary dynamics or path dependency. This simulation-based approach is articulated on three main steps. The first one involves the simulation of large sample paths of asset returns or state variables (e.g. by using posterior beliefs). The second implies finding the optimal portfolio and consumption policies in a recursive way using standard dynamic programming. In this part the algorithm proceeds until it arrives at zero. In the third and last stage, portfolio weights that maximise investor’s utility are obtained.

The principal scope of this model is to find a solution to limitations emerged from other models of the contemporaneous literature. In fact, the other approaches have in common simplistic assumptions on portfolio return distribution (e.g. Epstein-Zin and CRRA utility functions) and the incapacity of handling practical issues like the management of a large amount of state variables.

The second econometric approach embodies a completely new area of research that exploits machine learning techniques in order to solve portfolio selection problems. In fact, when making a regression, in the absence of linear functions, the artificial intelligence technique of artificial neural networks can be used because it is considered a good statistical representation of non-linear functions. For example, feed-forward and recurrent neural networks can be trained using market and macroeconomic data with the purpose of predicting returns of some asset, and then the results can be applied in a portfolio strategy involving the same assets. This is for sure a revolutionary approach that will be widely used in future years.

Chapter 3

An Empirical Application: the Italian stock market

In chapter one and two, I have presented the main streams of literature on portfolio choice problem, giving a wide space to the solution proposed by Brandt et al. Then, in this chapter I will describe the empirical application I choose in order to verify the effectiveness of Brandt's model also on the Italian market for the time period between 2000 and 2019. First, the reader will be provided with details about data used, preliminary computation undertaken on raw data and performance indicators used to draw conclusions on the model outcomes. Next, I will analyse results for each case considered.

For every output presented, the investor could only invest in stocks and he was assumed to have preferences described by a CRRA utility function.

3.1 Data

For the empirical application I decided to use Italian data. I considered firm-level monthly returns of Italian companies listed on Milan Stock Exchange for the period between January 2000 and December 2019. All data were taken from Eikon Thompson Datastream. For each period and firm considered, I constructed three main characteristics: market capitalization (Mkt cap), book to market ratio (BtM) and momentum (MoM). The first is defined as the log of price per share multiplied by the number of shares outstanding; the second is defined as the log of one plus book equity divided by market equity; the third was calculated each month as the compounded return between prior months t-12 and t-2. These three characteristics were chosen because, considering that are the most widely used in financial literature, I wanted to verify their relevance also on Italian market. For every computation, I took into account the standard timing convention for which we leave a six months lag after the end of a fiscal year to be sure that financial information would have been publicly available and so there is no insider trading activity.

Each characteristic has been standardized in order to have mean zero and standard deviation equal to one as in Brandt et al. (2009).

I started with an initial sample of 216 companies. From this I eliminated the ones with too many missing values and the 30% with the lower market capitalization. By doing so, I ended up with a sample of 69 listed firms.

The following figures (from 3.1 to 3.6) show the evolution in time of the mean and standard deviation of the non-standardized characteristics. These could be useful to go backward from a specific assessment of a standardized characteristic to the corresponding original value.

AVERAGE MARKET CAPITALIZATION

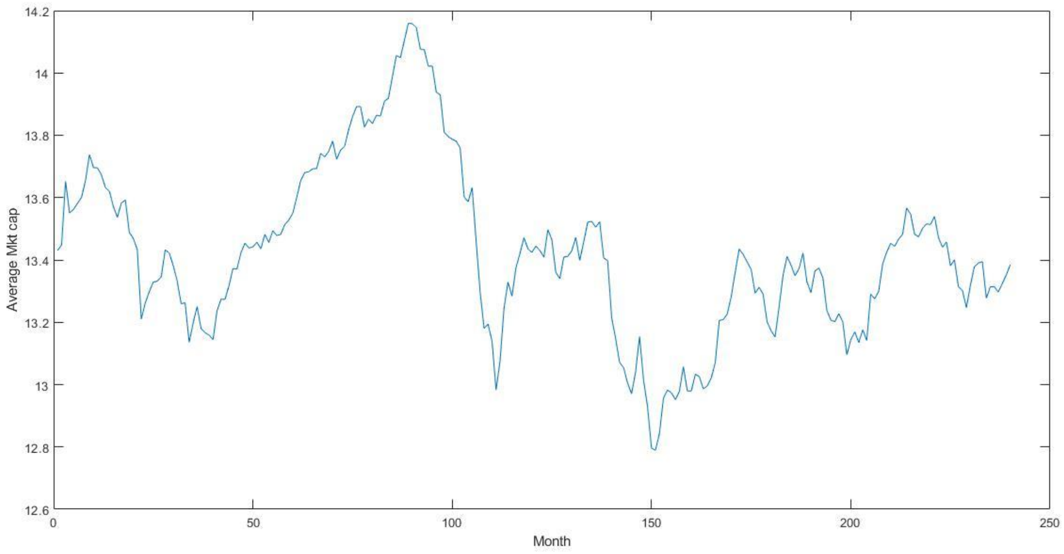

Figure 3.1 Source: self-elaboration of Eikon Thompson Datastream data

MARKET CAPITALIZATION STANDARD DEVIATION

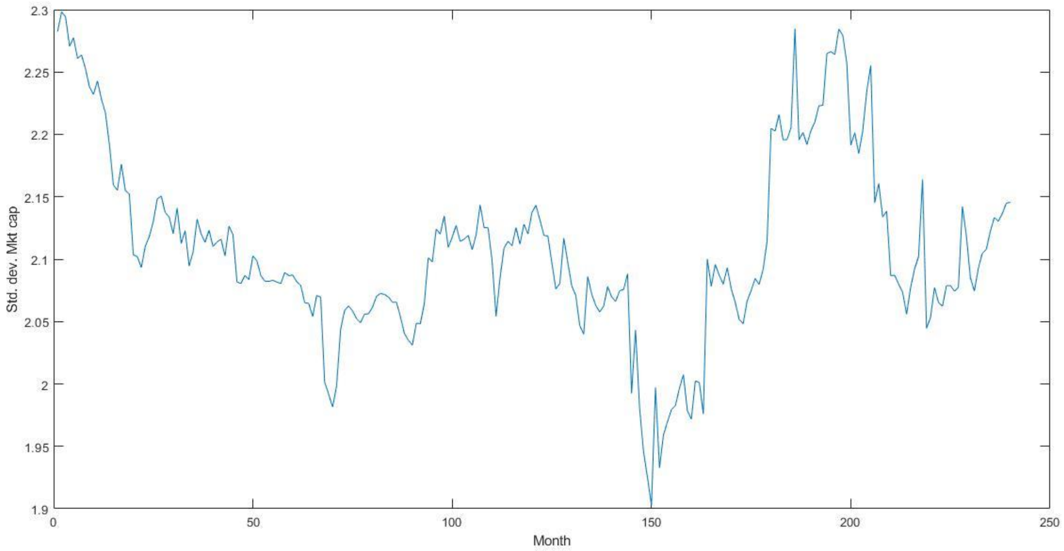

Figure 3.2 Source: self-elaboration of Eikon Thompson Datastream data

AVERAGE BOOK-TO-MARKET RATIO

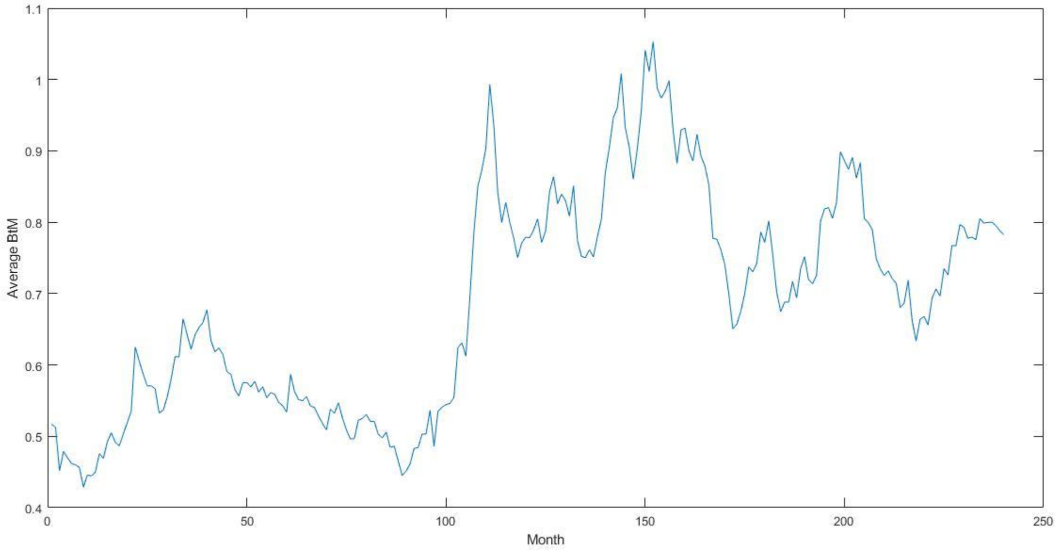

Figure 3.3 Source: self-elaboration of Eikon Thompson Datastream data

BOOK-TO-MARKET RATIO STANDARD DEVIATION

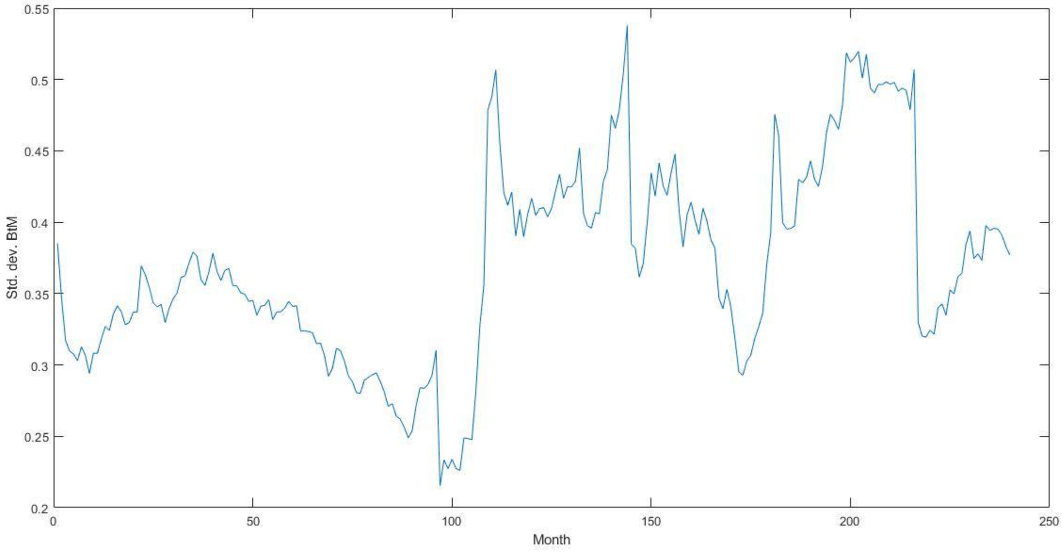

Figure 3.4 Source: self-elaboration of Eikon Thompson Datastream data

AVERAGE MOMENTUM

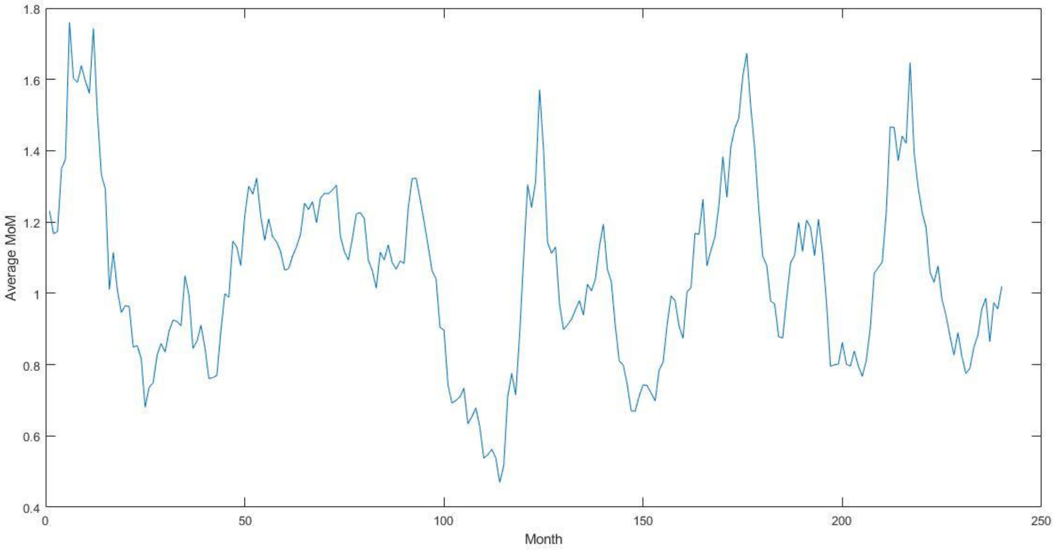

Figure 3.5 Source: self-elaboration of Eikon Thompson Datastream data

MOMENTUM STANDARD DEVIATION

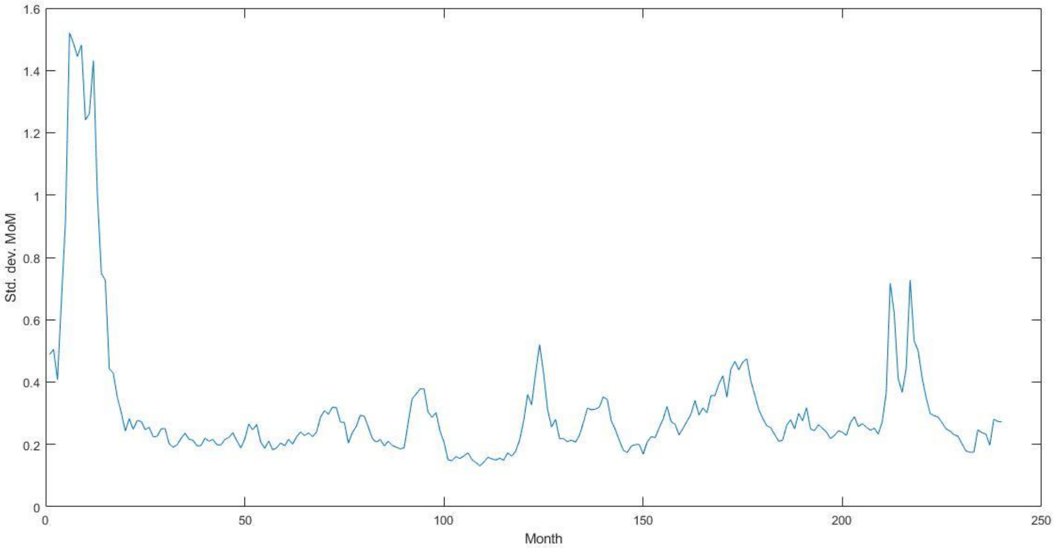

Figure 3.6 Source: self-elaboration of Eikon Thompson Datastream data

3.2 Performance indicators

In order to have a better understanding of the outcomes of the empirical application and for easiness of comparison I decided to use the following performance indicators.

- Sharpe ratio (SR)
- Certainty Equivalent Return (CER)
- Asset Turnover (T)
- Herfindal Hirschman Index (HHI)
- Proportion of Negative Weight (PNW)
- Mean and Standard Deviation of portfolio return
- Maximum and minimum weight.

The Sharpe Ratio is a risk-adjusted method to evaluate the performance of an investment. It is given by the ratio between the expected return of the investment and its standard deviation:

$$SR = \frac{E[R]}{\sqrt{Var(R)}}$$

It provides us with a measure of the excess return per unit of deviation in a certain investment opportunity. For a better interpretation, it should be used in comparison with other values for other investments. The higher it is, the better.

The Certainty Equivalent Return is calculated as:

$$CER = r_p - \left(\frac{\gamma}{2}\right) * \sigma_p^2$$

where r_p represents the average portfolio return and σ_p^2 optimal portfolio variance. It measures the impact of the entire distribution of returns according to risk preferences of the investor.

The Asset Turnover is computed as:

$$T_t = \sum_{i=1}^{N_t} \left| w_{i,t} - w_{i,t-1} \right|$$

It is given for each period by the sum of all absolute changes in weights that occur from one period to the subsequent, net of the ones that occur in a mechanical way. It shows the entity of the variation of the composition of the portfolio from time to time.

The Herfindal Hirschman Index is determined in the following way:

$$HHI = \frac{1}{T}\sum_{t=1}^{T} w_t' w_t$$

where w_t are again portfolio weights. It is used in order to evaluate the level of diversification of a portfolio.

The Proportion of Negative Weights is given by:

$$PNW = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_t}\frac{1_{\{w_{i,t}<0\}}}{N_t}$$

It measures the percentage amount of short position that are taken in a portfolio. It assesses the quantity of leverage of the portfolio.

The mean portfolio return is the average value of the portfolio returns, while the standard deviation of portfolio returns is given by the square root of the average square deviation of each data point from its mean value. It gives us an idea of the volatility of the portfolio.

Computing the maximum and minimum weight let us know the range of variation of weights. This information, used jointly with the other ones, would illustrate the entity of long and short positions that the model leads to take on.

3.3 Results

In the following paragraphs I will describe the results for the simple linear policy case and for the two extensions considered: the presence of transaction costs and the restriction of the no short sale constraint. Then, for the first two cases, I further extended them making the risk aversion coefficient vary and reducing the number of assets included in the portfolio.

The benchmark portfolio is assumed to be equally weighted.

3.3.1 Simple linear policy: the base case

In the following table (Table 3.1) are represented the results for the base case of the model, computed for a coefficient of risk aversion equal to five.

SIMPLE LINEAR PORTFOLIO POLICY WITH GAMMA EQUAL TO FIVE

	$\gamma = 5$
$\Theta_{Mkt\ cap}$	0.4268
Std. err.	0.0064
Θ_{BtM}	-4.2743
Std. err.	0.0035
Θ_{MoM}	0.7112
Std. err.	0.0035
Max w	0.1963
Min w	-0.3793
No. of negative weights	5657
PNW	0.343
T	1.5768
CER	-0.0269
Mean portf. Return	-0.0098
Std. portf. Return	0.0828
SR	-0.1184
HHI	0.3035
Mean Mkt cap	13.4237
Mean BtM	0.6883
Mean MoM	1.0512

Table 3.1 Source: Self-elaboration on Eikon Thompson Datastream data

The table, as the other ones included in this work, is articulated in four main sections. The first one describes the estimates obtained for the parameter and their standard errors. The second reports the descriptive statistics relative to the distribution of portfolio weights. The third part displays the performance indicators used to assess the portfolio. At last, they are exhibited values of average characteristics before the standardization process.

All data are not expressed in percentage and they are calculated on a monthly basis.

After the characteristics have been standardized, we can make in an easier way a comparison among them.

From the table we can see that the estimates are positive for market capitalization and momentum, while is negative for the book-to market capitalization. This means that the deviation of the optimal weights from the benchmark portfolio considered are growing with market capitalization and momentum, and they are slowing down with book to market ratio. This evidence is not consistent with literature. In fact, usually investors underweight companies that have a low book-to market ratio, the so-called growth firms, bad past returns and large firms (the ones with negative alphas). At the same time, stockholders overweight value firms (the ones with a high book-to-market ratio), past winners and the ones with a low market capitalization. These theoretical statements cannot be verified here. The deviation from the benchmark portfolio is positive for market capitalization, and so for big firms, and negative when considering book-to-market ratio. On the other side, the evidence found in literature on the lagged one-year return is confirmed.

The weights of the optimal portfolio fluctuate within a range with a maximum of 0.1963 and a minimum of -0.3793. The total amount of negative position is 5 657 out of 16 491, with a proportion of negative weight equal to 34.3% of the total.

The mean return of the portfolio is negative and equal to -0.98% on a monthly basis, while the standard deviation is equal to 0.0828. The turnover is 1.5768 and it is relatively high. With a negative Sharpe ratio and a negative Certainty Equivalent Return, -0.1184 and -0.0269 respectively, this portfolio has not one of the best performances, even if the lost is not so high. The Herfindal Hirschman Index is close to 0.3035. This bad performance may be due to lack of significance of the characteristics and the period selected. If fact, the performance of the model is closely related with how explanatory are certain variables for a specific market. Perhaps Italy would require other characteristics.

Figure 3.7 exhibits portfolio monthly gross returns across time. As can be easily seen, almost half of the optimal portfolio returns are above one.

GROSS PORTFOLIO RETURNS

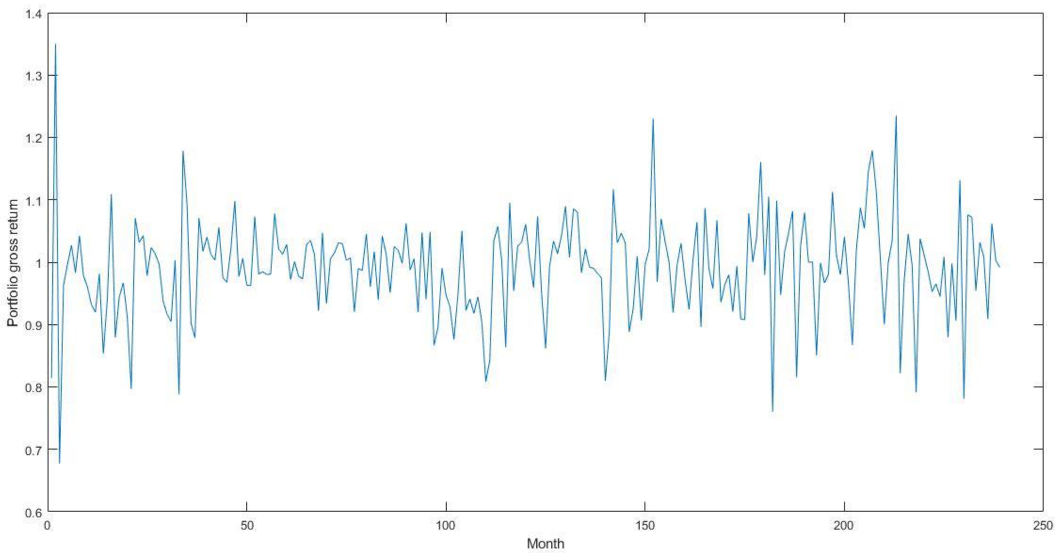

Figure 3.7 Source: self-elaboration of Eikon Thompson Datastream data

The average value of the three characteristics considered are 13.4237 for market capitalization, 0.6883 for book-to-market ratio and 1.0512 for momentum.

3.3.1.1 A comparison with other portfolio strategies

In the following table (Table 3.2) I will make a comparison with two alternative portfolio strategies in order to better understand the former results. I have considered an equally weighted portfolio and a value weighted portfolio. For the first one I have assigned an equal weight (0.0145) to each firm considered. In the second one the weights are assigned in proportion to market capitalization.

The coefficient of risk aversion, represented by the gamma in the CRRA utility function, is set as in the former case equal to 5.

SIMPLE LINEAR PORTFOLIO POLICY COMPARED TO EW AND VW PORTFOLIOS

	Brandt et al.	**EW**	**VW**
Max w	0.1963	0.0145	0.0227
Min w	-0.3793	0.0145	0.0065
No. of negative weights	5657	0	0
PNW	0.343	0	0
T	1.5768	0	0.018
CER	-0.0269	-0.0062	-0.0066
Mean portf. Return	-0.0098	0.003	0.0027
Std. portf. Return	0.0828	0.0608	0.0611
SR	-0.1184	0.0493	0.0442
HHI	0.3035	0.0144	0.0148

Table 3.2 Self-elaboration on Eikon Thompson Datastream data

The boundary of variation of weights distribution for the value weighted portfolio is from a lower bound of 0.0065 and a higher one of 0.0227. The equally weighted portfolio has no change as all weights are set to 0.0145. The portfolio strategy implemented by Brandt et al.'s model has a much wider range. Also the amount of short positions assumed is higher, while it set to zero for other strategies. This high level of negative position has been already underlined by Ammann et al. (2016), who have considered the fact that portfolios based on characteristics, in their optimal solution are very leveraged and many stocks should be shorted. However, this case is not always tolerable by an investor. The elevated level of the turnover compared to the value weighted strategy, equal to 0.018, shows a large trading activity for the optimal portfolio policy. The optimal portfolio is implementable anyway, although it is more likely to be affected by transaction costs. The low turnover of the other strategies can be due to variables that are stable in time. Though, the optimal portfolio does not push the investor to take too extreme bets.

The mean portfolio return is higher in both value and equally weighted strategies when compared to the optimal portfolio policy. We have a 0.3% portfolio return and a 0.27% respectively. So, it is convenient for the investor to invest in these alternative strategies. These values translate into a negative SR for the optimal portfolio policy and positive ones for the VW case (0.0442) and the EW (0.0493), which has the best performance. As expected, the optimal portfolio policy has a standard deviation that is considerably larger than the value weighted and equally weighted portfolios, which have a volatility almost equal. The CER is one of the best statistics to summarise performance because, as said before, it has the ability to catch the impact of the distribution of returns as a function of the risk preference of the investor. In this case the negative values obtained are highly influence by the relatively high level of volatility, the negative performance of the first portfolio, for which it is consistently lower, and by the chosen value for the coefficient of risk aversion.

However, consistently with other findings, the optimal portfolio is more diversified (the HHI is equal to 30.35%) in comparison with the other strategies (HHI of the EW is 1.44% and the VW is 1.48%). As expected, the EW portfolio has the worst level of diversification.

3.3.1.2 Gamma variation

Table 3.3 and table 3.4 display results for the base case and the two alternative strategies, value weighted and equally weighted, when we make the risk aversion coefficient vary inside the CRRA utility function. The values of gamma considered are 2, 5, 20 and 90. 90 is close to the extreme case of γ =100, for which the investor is highly sensible to risk; however, this level is not tolerable in practice.

By doing so, I want to highlight how influent are investors’ preferences in the estimation of the parameters for the optimal portfolio policy.

Data are always expressed on a monthly basis and not in percentage terms.

DIFFERENT VALUES FOR RISK AVERSION COEFFICIENT IN BASE CASE

	γ = 2	**γ = 5**	**γ = 20**	**γ = 90**
$\Theta_{Mkt\ cap}$	1.7709	0.4268	-0.5146	-0.5829
Std. err.	0.0147	0.0064	3.19E-04	2.91E-12
Θ_{BtM}	-9.2297	-4.2743	-1.3858	-0.7727
Std. err.	0.007	0.0035	2.05E-04	1.54E-12
Θ_{MoM}	0.2204	0.7112	1.0181	1.1691
Std. err.	0.008	0.0035	1.73E-04	1.45E-12
Max w	0.3198	0.1963	0.1447	0.1523
Min w	-0.8364	-0.3793	-0.145	-0.1053
No. of negative weights	6432	5657	4030	3634
PNW	0.39	0.343	0.2444	0.2204
T	2.8541	1.5768	1.2115	1.3093
CER	-0.0501	-0.0269	-0.0351	-0.1536
Mean portf. return	-0.0283	-0.0098	0.0013	0.0036
Std. portf. return	0.1475	0.0828	0.0603	0.0591
SR	-0.1919	-0.1184	0.0216	0.0609
HHI	1.3349	0.3035	0.0632	0.0486

Table 3.3 Source: self-elaboration on Eikon Thompson Datastream data

DIFFERENT VALUES FOR RISK AVERSION IN EW AND VW STRATEGIES

	EW				**VW**			
	γ = 2	**γ = 5**	**γ = 20**	**γ = 90**	**γ = 2**	**γ = 5**	**γ = 20**	**γ = 90**
Max w	0.0145	0.0145	0.0145	0.0145	0.0227	0.0227	0.0227	0.0227
Min w	0.0145	0.0145	0.0145	0.0145	0.0065	0.0065	0.0065	0.0065
No. of negative weights	0	0	0	0	0	0	0	0
PNW	0	0	0	0	0	0	0	0
T	0	0	0	0	0.018	0.018	0.018	0.018
CER	-0.0007	-0.0062	-0.0340	-0.1726	-0.0010	-0.0066	-0.0346	-0.1746
Mean portf. Return	0.003	0.003	0.003	0.003	0.0027	0.0027	0.0027	0.0027
Std. portf. return	0.0608	0.0608	0.0608	0.0608	0.0611	0.0611	0.0611	0.0611
SR	0.0493	0.0493	0.0493	0.0493	0.0442	0.0442	0.0442	0.0442
HHI	0.0144	0.0144	0.0144	0.0144	0.0148	0.0148	0.0148	0.0148

Table 3.4 Source: self-elaboration on Eikon Thompson Datastream data

As the value of gamma decreases, the parameter estimate for market capitalization characteristic increases. On the contrary, as the risk aversion of the investor increases, both estimates for book-to-market and momentum increase. While this variation is of relevant entity for book-to-market characteristic, it becomes less prominent for market capitalization and one-year lagged return. This means that book-to-market characteristic is much more sensible to variation in risk aversion and expected returns. Market capitalization and momentum are instead related mainly with expected returns.

Even the distribution of weights is significantly affected by the variation of gamma. For lower levels of gamma, the investor takes on a rising number of short positions, making more extreme bets. In fact, the minimum weight of the optimal portfolio goes from -0.1053 for $\gamma = 90$, to -0.8364 for $\gamma = 2$. This consequently makes the level of leverage grow. Also the maximum weight taken is higher for lower values of risk aversion (0.3198 for $\gamma = 2$ and 0.1523 for $\gamma = 90$). The variability of weights is much higher for less risk-averse investor. For gamma equal to 2 the turnover is high. This would make transaction costs relevant in proportion. As risk aversion increases, turnover reduces until the value of 1.2115 for $\gamma = 20$. It partially grows again for $\gamma = 90$. The level of leverage taken by the investor is considerably reduced for higher levels of gamma.

Consequently, this evidence affects the distribution of portfolio returns. The mean portfolio return has it minimum, -2.83%, for $\gamma = 2$ and its maximum value, 0.36%, for $\gamma = 90$. It turns positive when gamma is equal to 20. As expected from previous results, portfolio volatility is lower for higher values of risk aversion. Positive returns have an impact for the computation of the Sharpe ratio, which becomes increasingly positive for $\gamma = 20$ and $\gamma = 90$. The impact of return distribution according to risk preferences of the investor is of course higher for higher values of gamma; however, they cannot be compared with each other being computed for different values of gamma. Not surprisingly, we have a better level of diversification for lower levels of risk aversion.

3.3.1.3 Variation of number of stocks

In this paragraph I will analyse the effect of a variation in the number of stocks in the portfolio on performance and other statistics. Tables 3.5, 3.6 and 3.7 exhibit the results for a selected random subset of 35 assets from the original portfolio.

VARIATION OF NUMBER OF ASSET IN THE OPTIMAL PORTFOLIO

	69 firms				**35 firms**			
	$\gamma = 2$	**$\gamma = 5$**	**$\gamma = 20$**	**$\gamma = 90$**	**$\gamma = 2$**	**$\gamma = 5$**	**$\gamma = 20$**	**$\gamma = 90$**
$\Theta_{Mkt\ cap}$	1.7709	0.4268	-0.5146	-0.5829	3.4008	1.1642	-0.2678	-1.1951
Std. err.	0.0147	0.0064	3.19E-04	2.91E-12	0.0432	0.0429	0.0014	5.71E-13
Θ_{BtM}	-9.2297	-4.2743	-1.3858	-0.7727	-11.9304	-5.2935	-1.965	-1.403
Std. err.	0.007	0.0035	2.05E-04	1.54E-12	0.0246	0.0255	0.0008	3.25E-13
Θ_{MoM}	0.2204	0.7112	1.0181	1.1691	-0.5853	0.2879	0.7957	1.0288
Std. err.	0.008	0.0035	1.73E-04	1.45E-12	0.0335	0.0148	0.0003	2.17E-13
Max w	0.3198	0.1963	0.1447	0.1523	0.7649	0.3502	0.2164	0.258
Min w	-0.8364	-0.3793	-0.145	-0.1053	-1.5373	-0.6814	-0.2597	-0.2266
No. of negative weights	6432	5657	4030	3634	3695	3337	2311	2358
PNW	0.39	0.343	0.2444	0.2204	0.4417	0.3989	0.2763	0.2819
T	2.8541	1.5768	1.2115	1.3093	8.4746	3.6985	2.2289	2.4308
CER	-0.0501	-0.0269	-0.0351	-0.1536	-0.0812	-0.0381	-0.0404	-0.1944
Mean portf. return	-0.0283	-0.0098	0.0013	0.0036	-0.0386	-0.0137	-0.0007	0.0026
Std. portf. return	0.1475	0.0828	0.0603	0.0591	0.2064	0.0987	0.063	0.0644
SR	-0.1919	-0.1184	0.0216	0.0609	-0.1870	-0.1388	-0.0111	0.0404
HHI	1.3349	0.3035	0.0632	0.0486	4.5448	0.9121	0.1663	0.1477
Mean Mkt cap	13.4237	13.4237	13.4237	13.4237	13.4479	13.4479	13.4479	13.4479
Mean BtM	0.6883	0.6883	0.6883	0.6883	0.6807	0.6807	0.6807	0.6807
Mean MoM	1.0512	1.0512	1.0512	1.0512	1.059	1.059	1.059	1.059

Table 3.5 Source: self-elaboration on Eikon Thompson Datastream data

VARIATION OF NUMBER OF ASSET IN EW STRATEGY

	69 firms				35 firms			
	γ = 2	**γ = 5**	**γ = 20**	**γ = 90**	**γ = 2**	**γ = 5**	**γ = 20**	**γ = 90**
Max w	0.0145	0.0145	0.0145	0.0145	0.0286	0.0286	0.0286	0.0286
Min w	0.0145	0.0145	0.0145	0.0145	0.0286	0.0286	0.0286	0.0286
No. of negative weights	0	0	0	0	0	0	0	0
PNW	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0
CER	-0.0007	-0.0062	-0.0340	-0.1726	-0.0001	-0.0059	-0.0349	-0.1703
Mean portf. return	0.003	0.003	0.003	0.003	0.0038	0.0038	0.0038	0.0038
Std. portf. return	0.0608	0.0608	0.0608	0.0608	0.0622	0.0622	0.0622	0.0622
SR	0.0493	0.0493	0.0493	0.0493	0.0611	0.0611	0.0611	0.0611
HHI	0.0144	0.0144	0.0144	0.0144	0.0285	0.0285	0.0285	0.0285

Table 3.6 Self-elaboration on Eikon Thompson Datastream data

VARIATION OF NUMBER OF ASSET IN VW STRATEGY

	69 firms				35 firms			
	γ = 2	**γ = 5**	**γ = 20**	**γ = 90**	**γ = 2**	**γ = 5**	**γ = 20**	**γ = 90**
Max w	0.0227	0.0227	0.0227	0.0227	0.0398	0.0398	0.0398	0.0398
Min w	0.0065	0.0065	0.0065	0.0065	0.0156	0.0156	0.0156	0.0156
No. of negative weights	0	0	0	0	0	0	0	0
PNW	0	0	0	0	0	0	0	0
T	0.018	0.018	0.018	0.018	0.032	0.032	0.032	0.032
CER	-0.0010	-0.0066	-0.0346	-0.1746	-0.0002	-0.0060	-0.0347	-0.1784
Mean portf. return	0.0027	0.0027	0.0027	0.0027	0.0036	0.0036	0.0036	0.0036
Std. portf. return	0.0611	0.0611	0.0611	0.0611	0.0619	0.0619	0.0619	0.0619
SR	0.0442	0.0442	0.0442	0.0442	0.0582	0.0582	0.0582	0.0582
HHI	0.0148	0.0148	0.0148	0.0148	0.029	0.029	0.029	0.029

Table 3.7 Source: self-elaboration on Eikon Thompson Datastream data

When observing parameter estimates for the reduced sample, we can note that they follow the same path as the full sample: the estimates for market capitalization characteristic increase when risk aversion diminish, while both book-to-market and momentum estimates increase for higher values of gamma. However, although the absolute values of parameter evaluations grow in the case of the reduced sample for market capitalization and book-to-market ratio, for momentum estimate this absolute value decreases.

The reduction in the number of assets causes an increase in the variability of weights distribution. In fact, the boundaries created by maximum and minimum weights widen for every gamma considered. Also the proportion of negative weights rises consistently. This means that investors are push toward more extreme bets, taking on a higher level of leverage. Even the turnover is more than doubled in the reduced case, enhancing the entity of transaction costs.

When 35 firms are considered, there is a significant reduction in mean portfolio return and, as a consequence, in Sharpe ratio. At the same time, there is an upturn in portfolio volatility. This evidence confirms what already found by Fletcher (2017). In fact, according to his findings, there is a “positive relation between the number of securities in the optimal portfolios and the performance when using randomly selected subset of the investment universe”. Not surprisingly, with less assets the optimal portfolio is less diversified. Even certainty equivalent return is higher in reduced-sample case.

Average values of characteristics rise with the reduction of asset considered.

3.3.2 Transaction costs

In this section I will analyse the influence of transaction costs on the model’s estimates. We know that transaction costs can change much from stock to stock and they consistently reduced over time, although they remain still higher for smaller firms. Taking into account these considerations, I have decided to model transaction costs as in Brandt et al. (2009):

$$z_{i,T} = A_T(0.006 - 0.0025 * Mkt\ cap_{i,T})$$

where A_T is a factor used to model the linear decrease of transaction costs in time. The Market capitalization term in this case has been computed as the market capitalization at time T divided by the maximum value of it across all assets.

This function has been multiplied by turnover and then subtracted from portfolio return.

Figure 3.7 shows the evolution downwards of transaction costs in time. Maximum, minimum, median and average values are represented.

TRANSACTION COSTS VARIATION THROUGH TIME

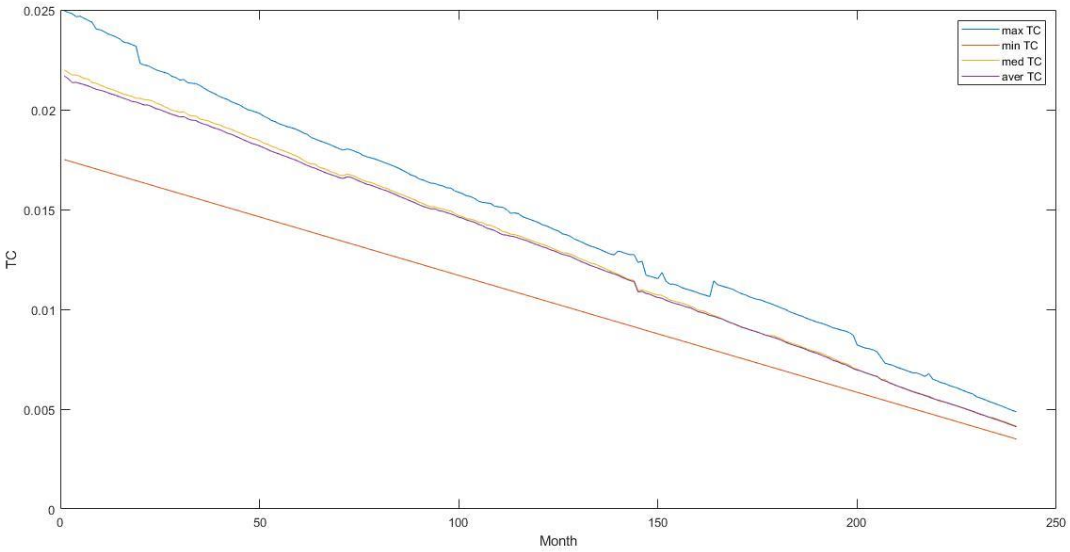

Figure 3.7 Source: self-elaboration of Eikon Thompson Datastream data

In table 3.8 are represented fluctuations of results. All computations are made for a coefficient of risk aversion equal to five. For easiness of comparison, in the table there have been inserted also data referring to the base case and the alternative strategies previously considered.

Data are expressed, as always, not in percentage terms and on a monthly basis.

TRANSACTION COSTS IN THE SIMPLE LINEAR POLICY

	BC	TC	EW	VW
$\Theta_{\text{Mkt cap}}$	0.4268	0.0823	0	0
Std. err.	0.0064	0.0073	0	0
Θ_{BtM}	-4.2743	-3.7938	0	0
Std. err.	0.0035	0.0038	0	0
Θ_{MoM}	0.7112	0.2232	0	0
Std. err.	0.0035	0.0038	0	0
Max w	0.1963	0.1391	0.0145	0.0227
Min w	-0.3793	-0.3331	0.0145	0.0065
No. of negative weights	5657	5353	0	0
PNW	0.343	0.3246	0	0
T	1.5768	1.1532	0	0.018
CER	-0.0269	-0.0213	-0.0062	-0.0066
Mean portf. return	-0.0098	-0.0126	0.003	0.0027
Std. portf. return	0.0828	0.0762	0.0608	0.0611
SR	-0.1184	-0.1654	0.0493	0.0442
HHI	0.3035	0.2246	0.0144	0.0148

Table 3.8 Source: Self-elaboration on Eikon Thompson Datastream data

For the simple linear policy, the inclusion of transaction costs brings a reduction in absolute values of parameter estimates. This is a direct consequence of the cost of trading. The optimal theta for market capitalization approaches to zero.

The distribution of weights becomes less broad, considering the fact that the maximum value goes from 0.1963 to 0.1391 in presence of transaction costs and the minimum from -0.3793 to -0.3331. Reflecting the fact that trading is much more expensive, turnover decreases as well as short position taken.

The already negative mean portfolio return drops, even if volatility declines. As shown by the Sharpe ratio, performance worsen with the burden of transaction costs. There is a lower level of portfolio diversification. However, certainty equivalent return reduces: the impact on return distribution of risk preference of the investor becomes less relevant.

3.4.2.1 Gamma variation

Table 3.9 exhibits what happens when I make gamma change. As in the instance of the simple linear policy, I want to analyse how relevant risk aversion is when estimating optimal thetas, even in the presence of transaction costs. As before, gamma is set equal to 2, 5, 20 and 90.

All data are on a monthly basis and not in percentage terms.

TRANSACTION COSTS WITH GAMMA VARIATION

	BC				TC			
	$\gamma = 2$	**$\gamma = 5$**	**$\gamma = 20$**	**$\gamma = 90$**	**$\gamma = 2$**	**$\gamma = 5$**	**$\gamma = 20$**	**$\gamma = 90$**
$\Theta_{Mkt\ cap}$	1.7709	0.4268	-0.5146	-0.5829	1.2464	0.0823	-0.6725	-0.6586
Std. err.	0.0147	0.0064	3.19E-04	2.91E-12	0.017	0.0073	2.81E-04	3.12E-13
Θ_{BtM}	-9.2297	-4.2743	-1.3858	-0.7727	-7.9305	-3.7938	-1.4919	-0.7863
Std. err.	0.007	0.0035	2.05E-04	1.54E-12	0.0085	0.0038	1.54E-04	1.61E-13
Θ_{MoM}	0.2204	0.7112	1.0181	1.1691	0.1607	0.2232	0.6448	1.1119
Std. err.	0.008	0.0035	1.73E-04	1.45E-12	0.0098	0.0038	9.85E-05	8.12E-14
Max w	0.3198	0.1963	0.1447	0.1523	0.2725	0.1391	0.1103	0.1477
Min w	-0.8364	-0.3793	-0.145	-0.1053	-0.7136	-0.3331	-0.133	-0.1031
No. of negative weights	6432	5657	4030	3634	6276	5353	3687	3613
PNW	0.39	0.343	0.2444	0.2204	0.3806	0.3246	0.2236	0.2191
T	2.8541	1.5768	1.2115	1.3093	2.4234	1.1532	0.8499	1.2482
CER	-0.0501	-0.0269	-0.0351	-0.1536	-0.0490	-0.0213	-0.0385	-0.1675
Mean portf. return	-0.0283	-0.0098	0.0013	0.0036	-0.0324	-0.0126	-0.0023	-0.001
Std. portf. return	0.1475	0.0828	0.0603	0.0591	0.1289	0.0762	0.0602	0.0592
SR	-0.1919	-0.1184	0.0216	0.0609	-0.2514	-0.1654	-0.0382	-0.0169
HHI	1.3349	0.3035	0.0632	0.0486	0.97	0.2246	0.0586	0.048

Table 3.9 Source: Self-elaboration on Eikon Thompson Datastream data

Giving a close look at the result, I can confirm findings for the simple linear policy case. Both the optimal thetas for book-to-market and momentum characteristics go up with as I take higher values of gamma. Estimate for market capitalization characteristic goes down as gamma increases. As before this change is stronger for the theta estimated for book-to-market characteristic, reflecting the fact that it is influenced by expected returns and risk aversion.

The introduction of transaction costs and the change in risk aversion parameter lead again to a reduction in turnover and in the number of negative and extreme positions taken. This makes the position more sustainable for the investor.

Mean portfolio returns are negative for every risk aversion value considered. As already stated, transaction costs worsen the situation for the investor. Again, portfolio volatility goes down with the two variations made, but not in a significant way. The optimal portfolio is better diversified in the absence of transaction costs; however, the difference becomes negligible for $\gamma = 90$.

3.4.2.2 Variation of number of stocks

In this section, I want to test once more whether the evidence found by Fletcher (2017) is true even in the presence of transaction costs. I newly select a random subsample of 35 firms. Table 3.10 presents the results.

All data are on a monthly basis and not in percentage.

TRANSACTION COSTS WITH ASSET VARIATION

	69 firms				**35 firms**			
	$\gamma = 2$	**$\gamma = 5$**	**$\gamma = 20$**	**$\gamma = 90$**	**$\gamma = 2$**	**$\gamma = 5$**	**$\gamma = 20$**	**$\gamma = 90$**
$\Theta_{Mkt\ cap}$	1.2464	0.0823	-0.6725	-0.6586	2.5792	0.8708	-0.3464	-1.3434
Std. err.	0.017	0.0073	2.81E-04	3.12E-13	0.0922	0.0523	0.0013	5.08E-13
Θ_{BtM}	-7.9305	-3.7938	-1.4919	-0.7863	-9.4562	-4.3971	-1.8127	-1.3265
Std. err.	0.0085	0.0038	1.54E-04	1.61E-13	0.0592	0.0325	0.0009	3.14E-13
Θ_{MoM}	0.1607	0.2232	0.6448	1.1119	0.0835	0.1949	0.6048	1.0263
Std. err.	0.0098	0.0038	9.85E-05	8.12E-14	0.0523	0.0165	0.0002	2.07E-13
Max w	0.2725	0.1391	0.1103	0.1477	0.6135	0.2891	0.1871	0.2592
Min w	-0.7136	-0.3331	-0.133	-0.1031	-1.2266	-0.5606	-0.2307	-0.2198
No. of negative weights	6276	5353	3687	3613	3640	3208	2135	2440
PNW	0.3806	0.3246	0.2236	0.2191	0.4351	0.3835	0.2552	0.2917
T	2.4234	1.1532	0.8499	1.2482	6.5651	3.032	1.816	2.402
CER	-0.0490	-0.0213	-0.0385	-0.1675	-0.0011	-0.0348	-0.0423	-0.1920
Mean portf. return	-0.0324	-0.0126	-0.0023	-0.001	-0.0407	-0.0162	-0.0037	-0.0013
Std. portf. return	0.1289	0.0762	0.0602	0.0592	0.1665	0.0862	0.0621	0.0651
SR	-0.2514	-0.1654	-0.0382	-0.0169	-0.2444	-0.1879	-0.0596	-0.0200
HHI	0.97	0.2246	0.0586	0.048	2.8967	0.6268	0.1384	0.1498

Table 3.10 Source: Self-elaboration on Eikon Thompson Datastream

Even in the presence of the burden of transaction costs, we can see the positive relation between the number of securities included in the optimal portfolio and its performance. The optimal portfolio is much more volatile in the reduced portfolio case, but better diversified.

Analysing the distribution of portfolio weights, we can see that once more maximum and minimum position take are higher in absolute value. This fact has been already underlined in the simple linear policy case. The investor makes more extreme bets.

Optimal thetas increase in absolute value, exception made for parameter estimate of momentum characteristic.

3.4.3 No short sale constraint

In this section I will discuss results referring to the imposition of the no short sales constraint. As exposed before, the imposition of that limitation implies setting negative positions equal to zero. Although this operation would make the sum of weights higher than one, so I had renormalized them.

As stated before, already Ammann et al. (2016) had highlighted the fact that characteristics-based portfolios push investors toward high level of leverage, letting them short a higher amount of stocks. However, for investors who are really risk averse this level of leverage is barely tolerable. This can be translated in an elevated turnover and, consequently, in increased transaction costs.

The imposition of the no short sale constraint is the solution proposed by Brandt et al. (2009) in their paper for the high level of leverage. Although, this limit would only completely eliminate leveraged positions which would be an acceptable solution only for extremely risk adverse investor. Indeed, the average investor would be more likely to take a position with a medium-small level of leverage.

There are anyway some cases in which this constraint is reasonable, for example when there is a country with incomplete market, as in Brazil, or when there is a temporary restriction imposed by the regulator for special situations. The principal aim is trying to avoid speculations.

Table 3.11 exhibits results obtained for the simple linear policy case and for the presence of no short selling constraint when gamma is equal to five.

NO SHORT SALE CONSTRAINT

	No short sale	BC
$\Theta_{Mkt\ cap}$	0.4263	0.4268
Std. err.	0.0064	0.0064
Θ_{BtM}	-4.2740	-4.2743
Std. err.	0.0035	0.0035
Θ_{MoM}	0.7117	0.7112
Std. err.	0.0035	0.0035
Max w	0.0806	0.1963
Min w	0	-0.3793
No. of negative weights	0	5657
PNW	0	0.343
T	0.4023	1.5768
CER	-0.0269	-0.0269
Mean portf. return	-0.0098	-0.0098
Std. portf. return	0.0828	0.0828
SR	-0.1184	-0.1184
HHI	0.0305	0.3035

Table 3.11 Source: Self-elaboration on Eikon Thompson Datastream data

There is a very small difference when comparing the optimal thetas for the two cases. Estimates for market capitalization and book-to-market ratio go down, while, on the contrary, the one for momentum grows.

Not surprisingly, there are no negative positions and turnover is consistently reduced.

There are no relevant variations for mean portfolio return, standard deviation and certainty equivalent return. Usually the average portfolio return of a portfolio with the no short sale constraint has a lower return for two main reasons. First, short positions can be used to increase the exposure in long positions by taking more leverage. Second, the investor with this limitation cannot take advantage of both positive and negative forecasts. The level of diversification is consistently reduced when the constraint is imposed.

3.4 Analysis of results: hypotheses testing

So far, I have assumed that the bad performance of the model derives primarily from the choice of firm characteristics. In fact, market capitalization, book-to-market ratio and momentum might not be explanatory variables for Italian stock market.

In order to assess the robustness of this assumption, I have decided to use the approach formulated by Fama and MacBeth (1973). Through this test I want to evaluate whether these characteristics are able to explain the cross-section of returns of stocks for the time period selected.

We know that in Arbitrage Pricing Theory (APT) and more generally in multifactor models, the expected return of a marketable security can be written as the sum of each factor multiplied by the sensitivity of the security to that particular factor (factor beta). A factor is an element which synthetizes a different component of the overall systematic risk which affects the asset. On the other side, a factor beta embodies the percentage variation in the expected return of the selected security for 1% change in the excess return of the factor portfolio. Anything could be a factor, from accounting measures to macroeconomic variables, thus we have to carefully select them in order to have an accurate estimate and not end up with irrelevant regressors.

Therefore, in my case firm characteristics are the factors which are supposed to explain the exposure of estimated expected returns to that particular factor.

The Fama-MacBeth methodology consists in two different steps. The first one involves running a classical times series regression aimed at identifying beta coefficients for the explanatory variables considered. The second requires running a cross-sectional regression of stock returns against the betas estimated in the previous step in order to find the risk premia λ. If my hypothesis is correct, the resulting risk premia should not be statistically significant.

Table 3.12 exhibits the results deriving from the execution of this procedure.

ESTIMATED RISK PREMIA

	Mean	Variance	t-stat
λ Mkt cap	-0.0284	0.5633	-0.5850
λ BtM	0.0255	0.1714	0.9522
λ MoM	0.0798	0.5745	1.6276

Table 3.12 Source: Self-elaboration on Eikon Thompson Datastream data

As we can see from the table above, risk premia are indeed not statistically significant, because we cannot reject the null hypothesis of λ equal to zero at a 5% confidence level. Though, the risk premium relative to the one-year lagged return is close to the critical value, meaning that it is only weakly significant.

Thus, I can definitely affirm that the suboptimal performance of the model for the time period considered is mainly due to the fact that the characteristics chosen are not able to explain the cross-section of stocks returns on Italian stock market.

Although, this is just one of the possible determinants of the suboptimal performance of the model. In fact, as highlighted by De Miguel et al. (2009), alternative portfolio strategies are more likely to beat the performance of an equally weighted portfolio if the estimation period considered is very large (at least 6 000 months are required for a portfolio of 50 assets), and the number of assets considered is small. This is a direct consequence of the fact that for longer periods we have more data available for computations and if there are less assets there is consequently a smaller number of parameters to estimate, thus reducing the possibility of making an error in the process. Even if I considered only 69 assets, the estimation window might not be enough broad.

Other evidences found in my work can find a confirmation in literature exposed at the end of chapter one. For example, when I impose the no short sale constraint on a broad time period there is a consistent reduction in turnover, while the performance does not improve.

Conclusion

Across this paper I have analysed portfolio choice problem from different points of view.

In the first chapter we have seen as Markowitz's mean-variance paradigm has been the most widely used since its formulation in 1952. Although, I have also underlined the fact that the mean-variance framework leaves two open points. The first is that it requires complex and tedious calculations regarding the step of modelling expected returns, variances and covariances. The second issue is that the paradigm refers only to a single period investment, so the investor cannot change its investment decision at intermediates dates.

Subsequently, I have focused the attention on more general dynamic approaches to the portfolio choice problem, first considering discrete time formulation and later the one in continuous time.

Then, I have exposed two streams of literature regarding the econometric methodologies for solving the portfolio choice problem: traditional econometric approaches and recent econometric approaches. Regarding the first category, I have stressed the attention on pug-in estimation and decision theory. The first involves the initial step of drawing inference on parameters of the data generating process and then substituting the values found in the analytical expression of the investor's problem, while the second implicates the preliminary step of modelling return distribution and then tries to find optimal portfolio weights. As emerged, traditional approaches have the relevant disadvantages of transmitting estimation errors made in the first phase to the second one. On the other side, new approaches start from this evidence and try, through the use of parametric or nonparametric weights, to avoid the first step and directly estimate optimal portfolio weights.

In the second chapter I have presented the recent econometric approach formulated by Brandt, Santa-Clara and Valkanov in 2009. This model parametrises portfolio weights by expressing them as a function of firm characteristics. It shows to have many practical advantages and exhibits good results for the application made by the authors on the US stock market.

However, I cannot affirm the same thing when considering the Italian stock market. In the third chapter I have presented results of my empirical application which focuses on a subset of 69 Italian firms listed on Milan stock exchange for the period between 2000 and 2019. The model produces negative average portfolio returns in the simple linear policy case as well as when portfolio constraints such as no short sale constraint or transaction costs are included. The only exception is represented by the instance in which we consider in the base case a coefficient of relative risk aversion higher than 20. Thus, this portfolio policy is beaten even by much simpler policies such as an equally weighted portfolio and a value weighted portfolio.

This bad performance could be due to a wrong choice of firm characteristics. In fact, the results for the optimal parameters estimated do not confirm evidence found in literature about the characteristics chosen: market capitalization, book-to-market ratio and momentum. Consequently, these characteristics appear to be not optimal for the application of Brandt et al.'s model on Italian stock market. To test this hypothesis, I have performed a two-stage regression following Fama-MacBeth methodology, which has confirmed the assumption made.

For future applications I advise to use firm characteristics that represent better the Italian stock market. Another way for improvement in performance is the extension of the estimation window considered. This recommendation is influenced by a study conducted by DeMiguel et al., which has shown that alternative optimizing portfolio strategies require to consider data of very large time period for a small number of assets in order to perform better than an equally weighted portfolio.

References

Ackermann F., P. Walt and K. Schmedders (2017), "Optimal and Naïve Diversification in Currency Markets", *Management Science*, Vol. 63, No. 10, 3347-3360

Ammann M., G. Coqueret and J.P. Schade (2016), "Characteristics-Based Portfolio Choice with Leverage Constraints", *University of St. Gallen Working Paper*, No. 6, February (http://ssrn.com/abstract=2736324)

Berk J. and P. DeMarzo (2017), *Corporate Finance*, Pearson

Brandt M. W. (2010), "Portfolio Choice Problem", of Ait-Sahalia Y. and Hansen L. P., *Handbook of Financial Econometrics, Volume 1: Tools and Techniques*, 269-336

Brandt M. W., A. Goyal , P. Santa-Clara and J. R. Stroud (2005), "A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability", *The Review of Financial Studies*, Vol. 18, No. 3, 831-873 (https://doi.org/10.1093/rfs/hhi019)

Brandt M. W. and P. Santa-Clara (2006), "Dynamic Portfolio Selection by Augmenting the Asset Space", *The Journal of Finance*, Vol. 61, No. 5, 2187-2217

Brandt M. W., P. Santa-Clara and R. Valkanov (2009), "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns", *Review of Financial Studies*, Vol. 22, No. 9, 3411-3447

Brooks (2002), *Introductory Econometrics for Finance*, Cambridge University Press

Campbell J. Y., Y. L. Chan and L. M. Viceira (2003), "A Multivariate Model of Strategic Asset Allocation", *Journal of Financial Economics*, No. 67, 41-80

Campbell J. Y. and L. M. Viceira (2002), "Strategic Asset Allocation", *Oxford University Press*

Chonghui J., D. Jiangze Y. and An and A. Yunbi (2019), "Combining the Minimum-Variance and Equally Weighted Portfolios: Can Portfolio Performance Be Improved?", *Economic Modelling*, Vol. 80, No. 260-274

Davis M. and A. Norman (1990), "Portfolio Selection with Transaction Costs", *Mathematics of Operations Research*, Vol. 15, 676-713

De Miguel V., L. Garlappi and R. Uppal (2009), "Optimal Versus Naïve Diversification: How Inefficient is the 1/N Portfolio Strategy?", *The Review of Financial Studies*, Vol. 22, No. 5, 1915-1953

DeMiguel V., A. Martin-Utrera, F. J. Nogales and R. Uppal (2019), "A Transaction-Cost Perspective on the Multitude of Firm Characteristics" (https://ssrn.com/abstract=2912819)

Denault M., E. Delage and J. G. Simonato (2017), "Dynamic Portfolio Choice: A Simulation-and-Regression Approach", Springer Science, Vol. 18, 369-406 (https://doi.org/10.1007/s11081-017-9347-4)

Djehiche Y. (2018), "A Neural Network Approach to Portfolio Choice", KTH Royal Institute of Technology

Fama E. F. and J. MacBeth (1973), “Risk, Return and Equilibrium: and Empirical Test”, The Journal of Political Economy, Vol. 81, No. 3, 607-636

Fama E. F. (1970), “Multiperiod Consumption-Investment Decisions”, *American Economic Review*, Vol. 60, 163-174

Fletcher J. (2017), “Exploring the Benefit of using Stock Characteristics in Optimal Portfolio Strategies”, *The European Journal of Finance*, Vol. 23, No. 3, 192-210 (https://doi.org/10.1080/1351847X.2015.1062036).

Gu S., B. Kelly and D. Xiu (2020), “Empirical Asset Pricing via Machine Learning”, The Review of Financial Studies, Vol. 33, No. 5, 2223-2273

Markowitz H. (1952), “Portfolio Selection”, *The Journal of Finance*, Vol. 7, 77-91

Medeiros M. C., A. M. Passos and G. F. R. Vasconcelos (2014), “Parametric Portfolio Selection: Evaluating and Comparing to Markowitz Portfolios”, *Rev. Bras. Finanças*, Vol. 12, No. 2, 257-284

Merton R. C. (1969), “Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case”, *Review of Economics and Statistics*, Vol. 51, 247-257

Merton R. C. (1971), “Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory*, Vol. 3, 373-413

Newey W. K. and K. D. West (1994) “Automatic Lag Selection in Covaraince Matrix Estimation”, *Review of Economic Studies*, Vol. 61, No. 4, 631-653

Samuelson P. A. (1969), “Lifetime Portfolio Selection by Dynamic Stochastic Programming”, *Review of Economic and Statistics*, Vol. 51, 239-246

Wakker P. P. (2008), “Explaining the Characteristics of the Power (CRRA) Utility Family”, *Health Economics*, No.17, 13329-1344 (https://doi.org/10.1002/hec.1331)

Zhang Y., X. Li and S. Guo (2017), “Portfolio Selection Problems with Markowitz’s Mean-Variance Framework: A Review of Literature”, *Springer Science*, No. 17, 125-158 (https://doi.org/10.1007/s10700-017-9266-z)

Summary

This paper has the purpose of testing the successful applicability on Italian stock market of one of the new econometric approaches formulated in literature in recent years. This methodology exploits parametric portfolio weights in order to find the ones which maximizes the investor's expected utility. Thus, by avoiding the preliminary step of estimating return distribution or drawing inference on parameters, the econometrician reduces the estimation error usually made and which can influence portfolio performance.

Before going deeply into the exposition of this approach, I will start analysing the theoretical framework which has led to its construction.

According to financial literature, every portfolio choice problem shares five main modelling issues: the choice of the utility function, the possibility of inserting intermediate consumption, the assessment of whether the market in which we operate is complete or not, the time horizon of the investor and the presence of frictions or background risk on the market.

The utility function is the function specifying investor's preferences. Among these, the class most widely used is the one of the hyperbolic absolute risk aversion (HARA).

If the econometrician decides to insert intermediate consumption, the investor will have to choose each period how much wealth to invest and the quantity to use for consumption.

In literature, financial markets are said to be completed when "the state variables governing investment opportunities are driven by the same stochastic process that drives asset returns" (Campbell (2002)). Moreover, complete markets can be further labelled as statically complete and dynamically complete. This classification has an influence on the complexity of the resolution of a portfolio choice problem.

The definition of the time horizon is another important element to consider. In fact, in literature we can distinguish among finite and infinite horizons. However, in practice, the investor has rarely a clear view on which will be the final date of his investment opportunity.

Market frictions are all elements strictly connected to the risky security, which can create an issue. An example are transaction costs. On the contrary, background risk incorporates all risks not directly linked to the security that an investor has to face.

Since its formulation in 1952, the most widely used methodology used for solving the portfolio choice problem has been Markowitz's mean-variance framework. However, it is not so easy to implement because it requires the preliminary step of modelling returns, variances and covariances. Thus, for this reason, during the years many critics have been moved toward its use. Among them, for example, there is the fact that it refers only to a single period investment, so that the investor cannot change his investment decision at intermediates dates.

If we move to a more generical formulation of the portfolio choice problem, we can consider multiperiod time horizons. This is the case of the dynamic asset allocation framework. In this context, we can make an ulterior distinction between discrete time and continuous time. We are in a discrete time setting when the investor has a well-defined timeline for its investment opportunity and at some intermediate dates there is the possibility to rebalance his portfolio. On the other side, we have portfolio choices in continuous time. In this case the investor must take investment decisions regarding portfolio choices constantly because portfolio is rebalanced every instant.

However, there are some cases in which for the investor it is more convenient to invest myopically. This happens mainly in four circumstances: the investor has a single-period horizon, investor's preferences are represented by logarithmic utilities, there are constant investment opportunities and when there are stochastic investment opportunities that cannot be hedged.

In financial literature we can distinguish between two streams of economic approaches: the traditional and the alternative ones. In traditional econometric approaches the econometrician has the purpose of defining the data generating process. This can be done in two different ways: with plug-in estimation and through decision theory. In plug-in estimation, also called calibration, the econometrician estimates portfolio weights and then he inserts them into an analytical solution in the investor's problem. On the contrary, in decision theory approach, the econometrician tries to identify the return distribution in order to find optimal portfolio weights.

The main weakness of the plug-in estimation methodology is the fact that the outcomes depend mainly from asymptotic results and the estimation error made, especially when we have multiperiod portfolio choices, because the error committed in one time period cumulates through time.

In order to solve this issue, three main solution have been proposed during the years: portfolio constraint, factor models and shrinkage estimation. The imposition of a constraint on portfolio weights improves sample estimates by limiting their range of variation and eliminating more extreme ones. On the contrary, the use of a factor structure for the assets covariation result in a reduction in the number of free parameters of the matrix. With shrinkage estimation, through the use of a proper estimator, estimates are induced to converge toward a common value in order to reduce extreme estimation errors.

As anticipated before, decision theory approach tries to determine the return distribution by exploiting personal beliefs and data observed. In financial literature these prior beliefs are divided in two categories: informative and uninformative priors. Uninformative priors are the ones which incorporate very little information about parameters. On the contrary, informative priors include subjective information. Priors can be also specified on the base of theoretical outcomes of an economic model, through the mixed estimation method. This allows more generally to combine two sets of data into a single posterior distribution.

The most recent stream of literature proposes new approaches in order to avoid or at least reduce the estimation error generated by traditional ones. Inside this flow we can identify two principal groups of methodologies: the ones which uses parametric portfolio weights and the others which exploit nonparametric portfolio weights.

The technique of parametric portfolio weights models optimal portfolio weights as a function of observable quantities (Brandt (2010)) and then solves this function in order to maximize utility.

On the other side, the stream of nonparametric portfolio weights starts form the fact that when using parametric portfolio weights there could be made some mistakes when modelling the weight function.

In practice, however, the best performing model on the stock market proves to be the naïve equally weighted portfolio strategy, especially when considering an elevated number of assets and short estimation window.

Among the recent econometric approaches which uses parametric portfolio weights, I will now discuss the one formulated by Brandt, Santa-Clara and Valkanov in 2009. This formulation involves modelling optimal portfolio weights as a function of firm characteristics. These characteristics can be various, such as one-year lagged return, market capitalization, book-to-market ratio, asset growth, leverage and so on. The choice of course depends mainly on the type of market and the aim of the empirical application. Then, the coefficients are estimated by maximizing the utility function of the potential investor.

Brandt et al., at the beginning of their work, outline the main advantages of their model. First, it gives a direct focus on portfolio weights escaping the long and tedious step of modelling returns and characteristics. Second, it encapsulates in an implicit way the connection between returns, characteristics, variances and covariances. Third, it avoids the frequent problem in statistical computations of overfitting or not precise coefficient estimates. Finally, it allows to easily test hypotheses, whether they are individual or joint.

In an investment universe in which there is the possibility to invest in N stocks in each period, the investor has to choose weights in order to solve the following problem of utility maximization:

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} E_t\left[u(r_{p,t+1})\right] = E_t\left[u\left(\sum_{i=1}^{N_t} w_{i,t}\, r_{i,t+1}\right)\right]$$

$$w_{i,t} = f(x_{i,t};\theta)$$

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t}\theta^T \hat{x}_{i,t}$$

Where the utility function is a CRRA utility function which takes the form:

$$u(r_{p,t+1}) = \frac{(1 + r_{p,t+1})^{1-\gamma}}{1-\gamma}$$

However, other utility functions can be used.

This is the case in which portfolio weights are a linear function of θ and $\hat{x}_{i,t}$. $r_{p,t+1}$represents portfolio returns and $r_{i,t+1}$ are stock returns. $\bar{w}_{i,t}$ is the matrix of portfolio weights assumed as a benchmark. θ is the vector of constant parameter estimates, which is representative of the weight assigned to each firm characteristic selected. $\hat{x}_{i,t}$ is the matrix of normalized firm characteristics. This standardization brings with two advantages: it makes the sum of optimal weights equal to one and it grants that characteristics considered are stationary. $\theta^T \hat{x}_{i,t}$, embodies the deviations from the benchmark of optimal portfolio weights.

The computational complexity grows with the number of firm characteristics considered.

As, stated before, Brandt et al.'s model allows to easily test individual and joint hypotheses. This is because optimal thetas estimated can be interpreted as a generalized method of moment estimator. Thus, by exploiting this property, the covariance matrix of optimal estimates can be computed as asymptotic covariance matrix or alternatively with the bootstrap technique. In my empirical application I decide to use the first technique, although the authors use the second one across their paper.

So far, I have exposed the basic application of the model, the so called simple linear policy, and its fundamental assumptions. However, the authors have included in their paper some possible extensions, refinements and restriction that can be imposed to the parametric portfolio choice framework. Thus, I will now focus on their analysis.

The first model extension is the inclusion of transaction costs. Transaction costs are all those expenses related to the stock exchange, fees, or more broadly to the trading activity which can vary consistently from country to country. From a practical point of view, they are calculated as the total asset turnover from one period to the subsequent one multiplied by the function which synthetises costs. Asset turnover is a measure of the absolute amount of exchanges made from one time period to the other for each asset. This amount is then subtracted from portfolio return. The cost function can be anyway modelled in different ways.

Transaction costs can have a relevant impact on characteristics-base portfolios. In fact, in the presence of trading costs the simple linear policy presented above is far from being optimal.

Although, transaction costs can, in some cases, even bring some benefits. In fact, recent studies conducted by DeMiguel et al. (2019) have shown that they can influence the number of characteristics that are jointly significant for an investor's optimal portfolio.

Another model extension is the introduction of a portfolio constraint. In financial literature there are many ways in which these restrictions can be imposed to portfolio weights. However, the one imposed most frequently is the no short sale constraint, which limits the amount of short positions that an investor can take on by completely cutting negative weights. In practice, this limitation is imposed for example when the regulator wants to protect the market from speculations.

Nevertheless, in Brandt et al.'s paper, the no short sale constraint is aimed at trying to find a solution for the elevated level of negative weights and, consequently, leverage that the model induces the investor to take on. However, this is a drastic solution because leverage is completely cancelled out.

The last model extension considered is about time-varying coefficients. In the simple linear policy case, the optimal coefficient estimates are supposed to be time invariant. This assumption is essential in order to guarantee that optimal portfolio weights are not affected by the history of past returns. Although, this hypothesis can be easily surmounted with the introduction of a vector of predictor that models the factors which optimal coefficient are supposed to vary with in the linear portfolio policy equation.

Brandt et al.'s portfolio model is just one of the possible developments of the recent econometric approaches formulated in order to find a solution to the portfolio problem. I will now report two other examples which fall into this category.

The first one is a simulation approach framed by Brandt, Goyal, Santa Clara and Stroud. This model aims at solving a discrete-time portfolio choice problem and applies when there is a broad number of assets with arbitrary return distribution and a wide number of state variables with non-stationary dynamics or path dependency.

The second econometric approach embodies a completely new area of research that exploits machine learning techniques in order to solve portfolio selection problems.

Hitherto, I have presented the main streams of literature on portfolio choice problem, giving a wide space to the solution proposed by Brandt et al. Then, I will now describe the empirical application I choose in order to verify the effectiveness of Brandt's model also on the Italian stock market for the time period between 2000 and 2019.

As experimental application I decided to use Italian data. I considered firm-level monthly returns of Italian companies listed on Milan Stock Exchange for the period between January 2000 and December 2019. All data were taken from Eikon Thompson Datastream. For each period and firm considered, I constructed three main characteristics: market capitalization (Mkt cap), book to market ratio (BtM) and momentum (MoM).

For every computation I considered the standard timing convention for which we leave a six months lag after the end of a fiscal year to be sure that financial information would have been publicly available. Each characteristic has been standardized as in Brandt et al. (2009).

I started with an initial sample of 216 companies. From this I eliminated the ones with too many missing values and the 30% with the lower market capitalization. By doing so, I ended up with a sample of 69 listed firms.

In order to better analyse the performance of the model, I have chosen nine performance metrics: Sharpe ratio (SR), Certainty equivalent return (CER), Asset turnover (T), Herfindal Hirschman Index (HHI), Proportion of negative weights (PNW), mean and standard deviation of portfolio return and maximum and minimum weights. The Sharpe ratio provides us with a measure of the excess return per unit of deviation in a certain investment opportunity. The CER measures the impact of the entire distribution of returns according to risk preferences of the investor. On the contrary, the asset turnover shows the entity of the variation of the composition of the portfolio from time to time. Then, HHI is used in order to evaluate the level of diversification of a portfolio. Finally, the proportion of negative weights assesses the quantity of leverage in the portfolio.

Next, I will go through the outcomes of this application. I will start with the results for the simple linear policy case and for the two extensions considered: the presence of transaction costs and the restriction of the no short sale constraint. Then, for the first two cases, I further extended them making the risk aversion coefficient vary and reducing the number of assets included in the portfolio.

The benchmark portfolio is assumed to be equally weighted and the investor is supposed to have preferences represented by a CRRA utility function.

In the simple linear policy case, with risk aversion equal to five, the deviation of the optimal weights from the benchmark portfolio considered are growing with market capitalization and momentum, and they are slowing down with book to market ratio. However, this evidence is not consistent with literature, because usually investors underweight companies that have a low book-to market ratio bad past returns and large firms, and they overweight value firms, past winners and the ones with a low market capitalization. On the other side, the evidence found in literature on the lagged one-year return is confirmed.

The performance of the portfolio is negative and the turnover is high. This bad performance may be due to lack of significance of the characteristics and the period selected. If fact, the performance of the model is closely related with how explanatory are the chosen variables for a specific market.

When comparing the performance of Brandt et al.'s model with the one of an equally weighted and a value weighted portfolio, we can see that they perform much better even if they have a lower level of diversification and are easier to implement.

Then, I make the value of the coefficient of risk aversion vary, starting from the initial value of five, and then considering also the figures of two, twenty and ninety. By doing so I want to highlight how influent are investors' preferences in the estimation of the parameters for the optimal portfolio policy.

From this analysis emerges that book-to-market characteristic is much more sensible to variation in risk aversion and expected returns, while market capitalization and momentum are related mainly with expected returns.

The distribution of weights is significantly affected by the variation of gamma: for lower levels of gamma, the investor takes on a rising number of short positions, making more extreme bets.

On the other side, the mean portfolio return turns positive when gamma is equal to 20.

Afterwards, I analyse the effect of a variation in the number of stocks in the portfolio on performance and other statistics.

When observing parameter estimates for the reduced sample, we can note that they follow the same path as the full sample. However, although the absolute values of parameter evaluations grow in the case of the reduced sample for market capitalization and book-to-market ratio, for momentum estimate this absolute value decreases.

The reduction in the number of assets causes an increase in the variability of weights distribution and in the proportion of negative weights.

Furthermore, there is a significant drop in mean portfolio return, which is accompanied by an upturn in portfolio volatility. This evidence confirms what already found by Fletcher (2017) about the positive relation between portfolio performance and number of assets included in it.

After varying the number of assets included in the portfolio, I have decided to also include transaction costs to the original simple linear policy. These have been expressed as a function of market capitalization and they have been modelled to also comprise their linear decrease in time. The outcome is that with the insertion of transaction costs, there is a reduction in absolute values of parameter estimates and the already negative mean portfolio return drops.

Even in the presence of transaction costs, if I make gamma and the number of stocks vary, the outcomes are the same as for the simple linear policy case.

Now, I will discuss results referring to the imposition of the no short sales constraint. As exposed before, the imposition of the no short sale constraint is the solution proposed by Brandt et al. (2009) in their paper for the high level of leverage that results with the application of their model.

Although, this limit would only completely eliminate leveraged positions which would be an acceptable solution only for extremely risk adverse investor.

There is a very small difference when comparing the optimal thetas for the simple linear policy case and the one with the imposition of the no short selling constraint. Not surprisingly, there are no negative positions and turnover is consistently reduced.

Furthermore, there are no relevant variations for mean portfolio return, standard deviation and certainty equivalent return, even if usually the average portfolio return of a portfolio with the no short sale constraint has a lower return.

So far, I have assumed that the bad performance of the model derives primarily from the choice of firm characteristics, which might not be explanatory for the Italian stock market.

With the purpose of evaluating whether these characteristics are able to explain the cross-section of returns of stocks for the time period selected, I have decided to apply the approach formulated by Fama and MacBeth (1973). If my hypothesis is correct, the resulting risk premia should not be statistically significant.

Observing the t-statistics, I can definitely affirm that the suboptimal performance of the model for the time period considered is mainly due to the fact that the characteristics chosen are not explanatory for stock returns on Italian market. In fact, only the one-year lagged return is weakly significant.

Although, this is just one of the possible determinants of the suboptimal performance of the model. In fact, as highlighted by De Miguel et al. (2009), the estimation window considered, and the number of assets included in the portfolio can have an influence on the possibility for alternative optimizing models to outperform an equally weighted portfolio.