Volume-Return portfolio selection and large scale dimensional problems with bivariate Markov chains

Enrico Angelelli, Sergio Ortobelli Lozza, and Gaetano Iaquinta.

Abstract**—** A portfolio selection problem is proposed under the assumption that financial returns follow homogenous Markovian chains. In this framework we describe two possible applications of bivariate Markov chains in portfolio selection problems. First, we show how to account the joint behavior of future wealth considering a bivariate Markov process and propose a technique to reduce the dimensionality of the large scale portfolio choice problem considering the heavy tails of the returns. Secondly, we describe the (volume, return) portfolio evolution with a bivariate Markov chain and we propose a volume-return portfolio strategy that accounts the investor behavior. Finally, we perform an ex-post analysis to assess the large-scale reduction technique used and the performance of a portfolio strategy that accounts the joint volume-return Markovian behaviour.

Keywords**—**portfolio selection, stable Paretian distributions, bivariate Markov process, reduction of dimensionality, volume return portfolio strategy.

I. Introduction

IN this paper we deal with the large scale portfolio problem using bivariate Markov chains. First we show how to reduce the large scale portfolio selection problem using the asymptotic approximation of future wealth estimated with Markov chains. Secondly, we propose a volume return portfolio strategy and we evaluate its impact on the US stock market.

Traditional financial theory is based on the assumption of financial returns normally distributed. Many empirical studies (fundamental works of Mandelbrot (1963) and Fama (1965) and more recently Rachev and Mittnik (2000), Rachev et al. (2007) and the reference therein) have rejected a normal shape for financial returns distribution. Several research works have been proposed to improve the realism of the underlying financial models. In this paper we assume that asset returns follow Markov processes and thus their evolution is approximated with a Markov chain as suggested in empirical studies (see Cox et al. (1979), Angelelli and Ortobelli (2009), D'Amico et al. (2010)). Moreover we distinguish two possible applications in portfolio theory of bivariate Markov chains.

In the first application, we assume that each couple of assets follows a bivariate Markov process so we consider their joint distribution with a proper Markov chain. Then, we approximate the future wealth taking into account their asymptotic approximation. Next, we use the covariation of stable sub Gaussian distributions (see Samorodnistky and Taqqu (1994)) to reduce the dimensionality of the large scale portfolio problem (see Angelelli et al. 2013a for further discussion). Finally, we propose an empirical comparison to show and evaluate the impact on portfolio problems of the reduction of preselection dimensionality technique. In this empirical analysis we consider all the stocks of the main US stock markets (NASDAQ, and NYSE) during the period of the crisis.

In the second application, we approximate by a bivariate Markov chain the joint distribution of the portfolio of returns and of the portfolio of traded volume increments. Then, we examine a reward-risk portfolio strategy which considers the investor behavior with respect the volume and return evolution. In particular, we consider that investors would like to purchase those assets whose traded volume and wealth are increasing while they want to reduce losses in particular when their traded volume is increasing. Finally, we propose an ex-post empirical comparison to show and evaluate the impact of this portfolio strategy with respect to other performance measures used in dimensional reduction techniques. In this empirical analysis we consider all the most liquid stocks of the main US stock markets (NASDAQ, and NYSE) during the decade 2003-2013.

The paper is organized as follows. In Section 2 we discuss how modeling return series and their asymptotic behavior. Section 3 introduces the technique used to manage the curse of dimensionality and an empirical analysis on the US stock market. In Section 4 we discuss volume-return portfolio selection strategies and an empirical analysis on the most liquid US stocks. Section 5 briefly summarizes the paper.

II. Returns dynamics model

In this section we introduce and discuss the returns dynamics modeling by Markov chains. Then, we analyze some possible consequences of that model. In particular, we show how to determine the distributions of the future wealth, and of a couple of portfolios that follow a bivariate Markov

Enrico Angelelli is with University of Brescia Department Economics and Management, Brescia Italy (e-mail: angele@eco.unibs.it).

Sergio Ortobelli Lozza., is with University of Bergamo Department SAEMQ 24127 Bergamo Italy, and VSB-TU of Ostrava Department of Finance 70121 Ostrava, Czech Republic. (corresponding author to provide e-mail: sol@unibg.it).

Gaetano Iaquinta is a private research analyst, 20017 Rho (Milan) Italy (e-mail: gaetano.iaquinta@unibg.it).

process. Finally, we discuss how to take into account the asymptotic behavior of the future wealth process.

A. *Bivariate Markov processes*

Bivariate processes are used for different objectives in portfolio theory. Typically they are used either to evaluate the association between different financial factors or to reduce the dimensionality of large scale problems (see Iaquinta et al. (2010, 2011) and Angelelli et al. (2011)). In this work we use bivariate Markov processes to value the joint behavior, either, of the wealth and the traded volume of a given portfolio, or, of the future wealth obtained by two different portfolios of returns. Here we analyze how to approximate the joint behavior of two different portfolios by a bivariate Markov chain. Then, we extend this analysis to the return-volume portfolio framework in Section 4.

Similarly to the univariate case (see Angelelli et al. (2013)) assume that an initial wealth $W_0=(W_{0x},W_{0y})'=(1,1)'$ is invested at time $t=0$ in two portfolios of weights $x=[x_1,\cdots,x_n]'$ and $y=[y_1,\cdots,y_m]'$ of n and m risky assets respectively. The vectors x and y represent the percentage of the initial wealth (respectively W_{0x} and W_{0y}) invested in each asset. Let us denote the prices of these assets at time t by $P_t^{(x)}=[P_{1,t}^{(x)},\dots,P_{n,t}^{(x)}]'$ and $P_t^{(y)}=[P_{1,t}^{(y)},\dots,P_{m,t}^{(y)}]'$. The portfolios returns during the period $[t,t+1]$ are given by the vector $Z_{t+1}=(Z_{x,t+1},Z_{y,t+1})'$ with components

$$Z_{x,t+1}=\sum_{i=1}^{n} x_i \frac{P_{i,t+1}^{(x)}}{P_{i,t}^{(x)}} \text{ and } Z_{y,t+1}=\sum_{i=1}^{m} y_i \frac{P_{i,t+1}^{(y)}}{P_{i,t}^{(y)}}.$$

We assume that the portfolios returns $Z_{x,t}$ and $Z_{y,t}$ follow two homogeneous Markov processes.

We introduce the multi-index $i=(i_x,i_y)$ and denote by $z^{(i)}=(z_x^{(i_x)},z_y^{(i_y)})'$, $\quad i\in I:=\{(i_x,i_y):1\le i_x\le N, 1\le i_y\le M\}$ the states of the Markov chain. First we discretize the support of the Markov process $\{Z_t\}$. Given a set of past observations $\{z_{-K},\dots,z_0\}$, we consider the range of the portfolios returns

$$(\min_{k=-K,\cdots 0} z_{x,k}, \max_{k=-K,\cdots 0} z_{x,k})\times(\min_{k=-K,\cdots 0} z_{y,k}, \max_{k=-K,\cdots 0} z_{y,k})$$

and divide it into $N\cdot M$ bi-dimensional intervals $(a_i,a_{i-1})\times(b_j,b_{j-1})$, where $\{a_i\}$ and $\{b_j\}$ are two decreasing sequences given by

$$a_i:=\left(\frac{\min_k z_{x,k}}{\max_k z_{x,k}}\right)^{1/N} \max_k z_{x,k}, \quad i=0,\dots,N,$$

$$b_j:=\left(\frac{\min_k z_{y,k}}{\max_k z_{y,k}}\right)^{1/M} \max_k z_{y,k}, \quad j=0,\dots,M.$$

The idea is to approximate the returns associated to values of the Markov process in $(a_{i_x},a_{i_x-1})\times(b_{i_y},b_{i_y-1})$ by the state $(z_x^{(i_x)},z_y^{(i_y)})$ of the Markov chain defined by

$$z_x^{(i_x)}=\sqrt{a_{i_x}a_{i_{x-1}}}=\max_k z_{x,k}\left(\frac{\max_k z_{x,k}}{\min_k z_{x,k}}\right)^{\frac{1-2i_x}{2N}},$$
$$i_x=1,\dots,N$$

$$z_y^{(i_y)}=\sqrt{b_{i_y}b_{i_{y-1}}}=\max_k z_{y,k}\left(\frac{\max_k z_{y,k}}{\min_k z_{y,k}}\right)^{\frac{1-2i_y}{2M}}$$
$$i_y=1,\dots,M$$

Introducing,

$$u_x:=\left(\frac{\max_k z_{x,k}}{\min_k z_{x,k}}\right)^{1/N} \text{ and } u_y:=\left(\frac{\max_k z_{y,k}}{\min_k z_{y,k}}\right)^{1/M} \qquad (1)$$

we may write $z_x^{(i_x)}=z_x^{(1)}u_x^{1-i_x}$ and $z_y^{(i_y)}=z_y^{(1)}u_y^{1-i_y}$. Assuming the Markov chain $\{Z_t\}$ homogeneous, we denote its transition matrix by $Q=\{q(i,j)\}_{i,j\in I}$, where

$$q(i,j)=P\left(Z_{t+1}=z^{(j)} \mid Z_t=z^{(i)}\right), i,j\in I$$

represents the probability of observing the returns $z^{(j)}$ in $t+1$ being in $z^{(i)}$ at time t. These probabilities are estimated by the maximum likelihood estimates:

$$\hat{q}(i,j)=\frac{\pi_{ij}}{\pi_i}$$

where π_{ij} is the number of observations that transit from $z^{(i)}$ to $z^{(j)}$ and π_i the number of observations in $z^{(i)}$. See Sadek, and Limnios (2002) for the statistical properties of these estimators. Let us now consider the bivariate wealth process generated by the gross returns.

Given wealth W_{t-1} at time $t-1$, the wealth $W_t=(W_{tx},W_{ty})'$ at time t is a bivariate random variable with $N\cdot M$ possible values:

$$W_t=z^{(i)}\otimes W_{t-1}=\left(z_x^{(i_x)}W_{(t-1)x}, z_y^{(i_y)}W_{(t-1)y}\right),\ i\in I.$$

By denoting $i_s=(i_{x,s},i_{y,s})$ the realized state of the Markov chain at time s, the value of W_t is given by

$$W_t=\begin{pmatrix} W_{ox}z_x^{(i_{x,1})}z_x^{(i_{x,2})}\cdots z_x^{(i_{x,t})} \\ W_{oy}z_y^{(i_{y,1})}z_y^{(i_{y,2})}\cdots z_y^{(i_{y,t})} \end{pmatrix}.$$

It is clear that the sequence $\langle i_0,i_1,\dots,i_t\rangle$ identifies uniquely the path followed by the bivariate wealth process up to time t. Thus, using formulas (1), the wealth obtained along the path $\langle i_0,i_1,\dots,i_t\rangle$ is given by

$$W_t = \begin{pmatrix} W_{ox}\left(z_x^{(1)}u_x\right)^t u_x^{-\left(i_{x,1}+i_{x,2}+\cdots+i_{x,t}\right)} \\ W_{oy}\left(z_y^{(1)}u_y\right)^t u_y^{-\left(i_{y,1}+i_{y,2}+\cdots+i_{y,t}\right)} \end{pmatrix}$$

Notice that vectors x and y represent the percentages of the initial wealths. Thus, if we want to evaluate the sample path of the ex-post wealths, we have to recalibrate each portfolio in order to maintain these percentages constant over time.

Interesting enough, describing the gross returns by a general bivariate Markov chain with $N \cdot M$ possible states implies that the number of possible values for W_t grows exponentially with the time. However, in our setting, the final wealth W_t does not depend on the specific path followed by the process, but only on the sums of the indices of the states traversed by the Markov chain in the first t steps. As indices $i_{x,s}$ and $i_{y,s}$ can range in $[1, N]$ and $[1, M]$ respectively, we may have 'only' $[1+t(N-1)]\cdot[1+t(M-1)]$ values as the final wealth W_t. This known as the recombining effect of the Markov chain on the wealth process W_t.

Let us denote the $[1+t(N-1)]\cdot[1+t(M-1)]$ possible values of W_t at time t by

$$w^{(l,t)} = \begin{pmatrix} w_x^{(l_x,t)} \\ w_y^{(l_y,t)} \end{pmatrix} = \begin{pmatrix} (z_x^{(1)})^t u_x^{1-l_x} \\ (z_y^{(1)})^t u_y^{1-l_y} \end{pmatrix},$$

where $l = (l_x, l_y) \in L_t$ and

$L_t := \{(l_x, l_y) : 1 \le l_x \le 1+t(N-1), 1 \le l_y \le 1+t(M-1)\}$.

The possible values of W_t up to time T can be stored in T matrices of dimension $[1+(N-1)T]\times[1+(M-1)T]$ or in a mono-dimensional vector of size:

$$\sum_{t=1}^{T}[1+t(N-1)][1+t(M-1)] = O(NMT^3).$$

The wealth W_t can be represented by a three-dimensional Markovian tree, starting with a single node $w^{((1,1),0)} = (1,1)'$ and presenting at each time instant t the $[1+t(N-1)]\cdot[1+t(M-1)]$ nodes given by $w^{(l,t)}, l \in L_t$.

We are interested in the evolution of such a process $\{W_t\}$, which is clearly connected to the evolution of $\{Z_t\}$. Consider the matrix

$$P_{(W_t,Z_t)} = \{p_{(W_t,Z_t)}(l,i)\}_{l\in L_t, i\in I}$$

with components

$$p_{(W_t,Z_t)}(l,i) = P(W_t = w^{(l,t)} \cap Z_t = z^{(i)}),$$

which represents the probability of obtaining the wealth $w^{(l,t)}$ and to be in state $z^{(i)}$ at time t, and the vector $P_{W_t} = \{p_{W_t}(l)\}_{l\in L_t}$ with components

$$p_{W_t}(l) = P(W_t = w^{(l,t)}), \quad l \in L_t.$$

The probabilities $p_{(W_t,Z_t)}(l,i)$ and $p_{W_t}(l)$ can be computed recursively by the formulas

$$p_{(W_t,Z_t)}(l,i) = \begin{cases} p_i & t=0, l=1 \\ \sum_{h\in I}[p_{(W_{t-1},Z_{t-1})}(l-(i-1)), h)\cdot \\ \quad \cdot q(h,i)] & t>0, l_x-(i_x-1)>0, \\ & l_y-(i_y-1)>0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$p_{(W_t)}(l) = \begin{cases} 1 & t=0, l=1 \\ \sum_{h\in I} p_{(W_t,Z_t)}(l,h) & t>0 \\ 0 & \text{otherwise} \end{cases} \quad \textbf{(2)}$$

where $p_i = P(Z_0 = z^{(i)})$ is the probability that the return at time zero is $z^{(i)}$. We assume these probabilities to be known from past observations.

Even if for computing the distribution of the bivariate process we need more time than the univariate case, the computational complexity of this algorithm is still of polynomial order. This can be easily proved using a similar analysis to the one proposed by Iaquinta and Ortobelli (2006).

B. Modeling the asymptotic behavior of the log returns

The fact that log returns present a distribution with heavier tails than distributions with finite variance is documented in several empirical research works. The empirical investigation (see, among others, Rachev and Mittnik (2000) and the references therein) shows that

$$\Pr(|\ln z_{(x)}| > u) \sim u^{-\alpha} L(u) \qquad \text{as } u \to \infty \quad \textbf{(3)}$$

where $0 < \alpha < 2$ and $L(u)$ is a slowly varying function at infinity, i.e.,

$$\lim_{u\to\infty} \frac{L(cu)}{L(u)} \to 1 \text{ for all } c > 0,$$

Our dataset satisfies the relation (3) for values $1 < \alpha < 2$: This tail condition implies that the log returns $r_{(x)} = \ln z_{(x)}$ distribution admits finite mean and not finite variance and belongs to the domain of attraction of an α-stable law. This asymptotic behavior of data can be modeled assuming that for each portfolio $x \in S$ the forecasted log-wealth $(\tilde{W}_T(x) = \sum_{t=1}^{T} \ln z_{(x),t})$ at a given future time T is in the domain of attraction of an $\alpha(x)$ stable distribution, i.e.:

$$\tilde{W}_T(x) \overset{d}{=} S_{\alpha(x)}(\sigma(x), \beta(x), \mu(x)), \quad \textbf{(4)}$$

where $\alpha(x)\in(0,2]$ is the index of stability, $\sigma(x)$ is the scale parameter, $\mu(x)$ is the location parameter and $\beta(x)$ is the skewness parameter. McCulloch's method (see McCulloch (1986)) provides an efficient technique to derive stable Paretian parameters estimation. In particular, this method requires the knowledge of 5%, 25%, 50%, 75%, 95% quantiles of the log wealth $\tilde{W}_T(x)$ for any portfolio. Optimal portfolio strategies that account the Markovian and asymptotic behavior of the final wealth can be derived by computing reward and risk measures with stable distributions.

Alternatively, we can consider the asymptotic behavior of the future wealth, assuming that the vector of the forecasted log-wealths (obtained investing in each asset) is in the domain of attraction of a particular stable law. Typically, we can assume that the vector of the future log-wealths, denoted by

$$\tilde{W}_T(x)=[\tilde{W}_{T,1},\ldots,\tilde{W}_{T,n}]'$$

where $\tilde{W}_{T,i}$ is the log-wealth at time T obtained investing in the i-th asset, is α-stable sub Gaussian distributed. That is, the characteristic function of $\tilde{W}_T$ has the following form:

$$\begin{aligned}\Phi_{\tilde{W}_T}(u)&=E(\exp(\mathrm{i}u'r))=\\&=\exp(-(u'Vu)^{\frac{\alpha}{2}}+\mathrm{i}u'\mu)\end{aligned} \quad \textbf{(5)}$$

where $V=[v_{ij}]$ is a positive definite dispersion matrix, μ is the mean vector (when $\alpha>1$) and i is the imaginary unit. Since μ_i and $v_{ii}=\sigma_i^2$ are respectively the location parameter and the square scale parameter of the α-stable distributed i-th component $\tilde{W}_{T,i}$, we can estimate the parameters v_{ii} and $\mu_i(i=1,\ldots,n)$ using the McCulloch's quantile estimator, fixing the skewness parameter $\beta=0$ and imposing a common stability parameter α for all the components. Generally, as stability parameter α we use either the empirical mean of the stability parameters of the assets (i.e., $\alpha=\frac{1}{n}\sum_{i=1}^n\alpha_i$) or the stability parameter of the market index, if it exists. As remarked by Kring et al. (2008), the covariation parameter can be seen as the difference of square scale parameters, i.e.:

$$v_{ij}=\sigma^2_{(\tilde{W}_T(e_i)+\tilde{W}_T(e_j))/2}-\sigma^2_{(\tilde{W}_T(e_i)-\tilde{W}_T(e_j))/2} \quad \textbf{(6)}$$

where $\sigma^2_{(\tilde{W}_{T,i}\pm\tilde{W}_{T,j})/2}$ are the square scale parameters of the random variables

$$(\tilde{W}_{T,i}\pm\tilde{W}_{T,j})/2$$

whose distributions can be evaluated using the bivariate Markovian approximation. Thus, to estimate the covariation parameters v_{ij} (with $i\neq j$) of the stable vector we use the estimator

$$\hat{v}_{ij}=\hat{\sigma}^2_{(\tilde{W}_{T,i}+\tilde{W}_{T,j})/2}-\hat{\sigma}^2_{(\tilde{W}_{T,i}-\tilde{W}_{T,j})/2} \quad \textbf{(7)}$$

based on the estimates of the scale parameters of $(\tilde{W}_{T,i}+\tilde{W}_{T,j})/2$ and $(\tilde{W}_{T,i}-\tilde{W}_{T,j})/2$.

III. Dimensional reduction of the large scale portfolio selection problems

In this section we first examine the methodologies to reduce the dimensionality of large scale portfolio problems. Secondly, we propose an ex-post empirical analysis on the preselection of some "optimal" assets based on proper performance measures.

A dynamic portfolio selection problem can be stated in the following terms:

$$\max_{x\in S} f\big(W_T(x)\big) \quad \textbf{(8)}$$

where $f:(\Omega,F,P)\to\Re$ is the investor's performance functional, T is a defined horizon, W_T is the investor's final wealth, S is a $(n-1)$-dimensional simplex of the possible portfolios, x is the vector of weights solution of the choice problem. Recall that for static problems more than one hundred performance measures (see Cogneau and Hübner (2009a-b)) have been proposed that can be easily extended to a dynamic framework.

The typical functionals which are defined under the assumption that the gross return of each portfolio follows a Markov chain with N states are called OA performance (utility) functionals or OA performance measures (Angelelli and Ortobelli (2009)). OA performance measures can be used either to optimize the choices or to reduce the dimensionality of the portfolio problem. In particular we use the following two OA performance functionals in the next empirical analysis:

OA-Sharpe ratio (OA-SR). The classic version of the Sharpe ratio (see Sharpe (1994)) values the expected excess return for unity of risk (standard deviation). With the OA-Sharpe ratio we value the expected excess final wealth at time T for unity of risk, i.e.,

$$OA-SR(W_T(x))=\frac{E(W_T(x)-1)}{\sqrt{x'Q_Tx}} \quad \textbf{(9)}$$

where $Q_T=[q_{ij,T}]$ is the variance covariance matrix of the final wealth obtained with each asset at time T: we consider the Markov joint distribution at time T of i-th and j-th assets and compute their covariance $q_{ij,T}$. However, using Sharpe type measures we generally don't take into account the asymptotic behavior of the wealth (except in the case the optimal portfolios are in the domain of attraction of the Gaussian law).

OA-Stable Sharpe ratio (*OA-SSR*) This performance functional is defined as

$$OA-SSR(W_T(x))=\frac{\mu_{\ln(W_T(x))}}{\sqrt{x'Vx}} \qquad \textbf{(10)}$$

where $V=[v_{ij}]$ is the dispersion matrix computed with formula (6), $\mu_{\ln(W_T(x))}=\mu(x)$ is the mean of the stable distribution that better approximates the log final wealth $\ln(W_T(x))\overset{d}{=}S_a(\sigma(x),\beta(x),\mu_{lm(W_T(x))})$. As stability parameter α we use the empirical mean of the stability parameters of the assets (i.e., $\alpha=\frac{1}{n}\sum_{i=1}^{n}\alpha_i$). As for the Sharpe ratio, this ratio is isotonic with the preferences of non-satiable risk averse investors (see Rachev et al. (2007)). Using this performance measure we also take into account the asymptotic behavior of the future wealth.

A. *The large scale portfolio dimensional problem*

From a statistical point of view the number of observations should increase proportionally with the number of assets (see Papp et al. (2005), Kondor et al. (2007)) to achieve a reasonable statistical approximation of the historical series. Thus when we have a considerable number of assets we need to reduce the dimensionality of the portfolio problem. In the empirical analysis we adopt two methodologies to reduce the dimensionality of the large scale portfolio problems:

1. Preselection of an asset universe subset relevant for some optimality criteria;
2. Approximation of the relevant assets dynamics by some representative common factors.

The following paragraphs describe these methodologies in some details.

1) *Preselection*

Preselection methodology have been proposed and analyzed by Ortobelli et al. (2010, 2011). This methodology is performed to identify some relevant assets involved in the portfolio choice. In our dataset the choice is among more than 1500 US stocks. We suggest to preselect no more than 170 assets following two steps:

Step 1: Order the assets with the two performance measures (9) and (10).

Step 2: Select the "best" 170 assets satisfying the same common criteria (one group of 170 assets obtained by the intersection of the ordered groups of assets).

Once we get the optimal preselected assets we could suggest either to invest in them uniformly (i.e., investing in each asset the same percentage of wealth 1/170) or to further reduce their randomness approximating the series with some factors obtained, for example, with a principal component analysis.

2) *Common factors approximation*

Generally, we use a principal component analysis (PCA) to identify some common factors to approximate the asset returns (see Ross (1978)). The portfolio selection problem dimensionality is reduced by applying a non-Gaussian factor analysis that accounts for the joint Markov evolution of returns and their asymptotic behavior. First we perform a PCA of the preselected gross returns of the stocks to identify the few factors (portfolios) with the highest return variability. Therefore, we replace the original n correlated time series $\{z_i\}_{i=1,\dots,n}$ with n uncorrelated time series $\{R_j\}_{j=1,\dots n}$ assuming that each $\{z_i\}$ is a linear combination of the series $\{R_j\}_{j=1,\dots n}$. Then we implement a dimensionality reduction by choosing only those factors whose variability is significantly different from zero. We call portfolio factors $f_j(j=1,\dots s)$ the s time series in $\{R_j\}_{j=1,\dots n}$ with a significant dispersion measure, while the remaining $n-s$ series with very small dispersion measure are summarized by an error. Thus, each series z_i is a linear combination of the factors plus a small uncorrelated noise:

$$z_i=\sum_{j=1}^{s}a_j f_j+\sum_{j=s+1}^{n}a_j R_j=\sum_{j=1}^{s}a_j f_j+\varepsilon_i \qquad \textbf{(11)}$$

When the series don't present heavy tails the PCA could be applied to the Pearson correlation matrix of wealths obtained by the single gross returns. However, simple tests on these historical series reject the normality assumption and show that the historical series present very heavy tails. Thus, in order to consider the asymptotic behavior of the historical series we apply the PCA to the linear correlation matrix $\widetilde{V}=[\rho_{i,j}]$ (where $\rho_{i,j}=\frac{v_{ij}}{\sqrt{v_{ii}v_{jj}}}$) obtained by the dispersion matrix $V=[v_{ij}]$ of the stable sub-Gaussian hypothesis, that we call sub-Gaussian correlation matrix. We can consider the matrix based only on the historical series of log-returns or the correlation matrix computed on the forecasted log-wealths obtained investing in each asset. Observe that to compute the forecasted correlation matrix we have to use the bivariate Markov process. Once identified the s factors that account for most of variability of the historical gross returns, and other w factors $\widetilde{f}_j(j=1,\dots,w)$ that account for most of variability of the forecasted wealths we further reduce the variability of the error by regressing the series on the factors f_j and $\widetilde{f}_j$ so that we get:

$$z_{i,t+1}=b_{i,0}+\sum_{j=1}^{s}b_{i,j}f_{j,t}+\sum_{h=1}^{w}c_{i,h}\widetilde{f}_{h,t+1}+\varepsilon_{i,t+1}$$

Moreover, since the series present heavy tails we use the median regression on the factors to approximate the returns. Clearly, once we reduce the randomness of the large scale problem, we can use the preselected approximated returns to optimize portfolio selection strategies.

B. *Portfolio preselection in practice*

In this section, we evaluate the impact of the proposed model on the US stock market. In particular, we consider the stocks traded on the NYSE and on the NASDAQ. Since we want to propose as much as possible a realistic empirical analysis, we have developed a dynamic dataset that uses all the useful financial data from DataStream.

Using this dynamic dataset we propose two different ex-post comparisons during a period of about two years (500 daily observations) from 15-Sep-2008 till 31-Aug-2010 . In all the empirical analyses we assume:

a) that investors have a temporal horizon of T=20 working days (thus, for each portfolio strategy we should optimize the portfolio every 20 working days for a total of 25 optimizations);

b) Markov chains have N=9 states;

c) the initial wealth W_0 is equal to 1 at the date 15-Sep-2008.

The first comparison applies the portfolio selection to the approximated preselected returns using all the active assets and the daily observations during the previous ten years (2600 trading days). Thus, with this analysis we use 3100 daily observations overall from 14-May-1998 till 31-Aug-2010. The second comparison applies the portfolio selection to the approximated preselected returns using all the active assets and the daily observations during the previous six months (125 working days). Thus with this analysis we use 625 daily observations overall from 17-Mar-2008 till 31-Aug-2010. For both comparisons we value the empirical evidence from the preselected assets.

1) Empirical evidence from the preselected assets

In order to understand if preselection gives some benefits we compare the ex-post wealth of two portfolio strategies with the behavior of two market indexes: NASDAQ Composite and NYSE Composite. In the first strategy the investor uses a completely diversified portfolio on all active assets either in the last six months or in the last ten years (i.e., he/she invests $1/n$ in each asset where n is the number of available assets at the optimization date). With the second strategy the investor uses a completely diversified portfolio only on the preselected assets. For both strategies the investors recalibrate the portfolio every month (every 20 working days).

Figure 1 reports the sample paths of the ex-post wealth of the two strategies and of the market indexes when we consider all active assets either in the last six months, or in the last ten years. The difference between the strategies based on preselected assets among all those active either in the last ten years or in the last six months suggests that:

1) the set of the preselected assets is completely different in the two cases;

2) preselection works better if it is applied to more assets (as in the case of all assets active during the last six months);

3) the recent entries in the market could have an important impact in the portfolio choices.

Moreover, considering the limited transaction costs we obtain when we use preselection, we also deduce that it makes sense considering a preselected number of assets.

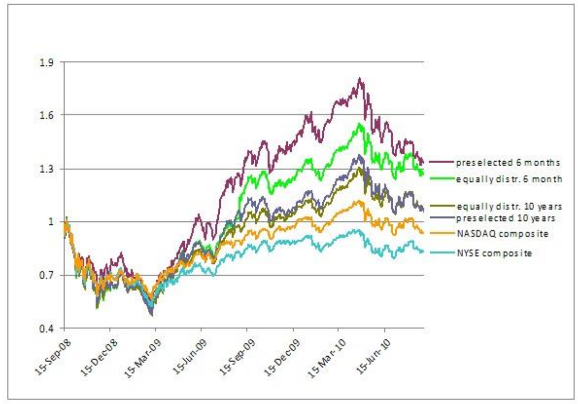


Fig.1: Comparison of the ex-post wealth among the indexes NYSE Composite, NASDAQ Composite and all the equally diversified portfolios among: preselected assets and all active assets either during the last 10 years or during the last 6 months.

IV. Volume-Return portfolio strategies

In this section, we evaluate the impact of the proposed bivariate Markov chain model applied to the volume-return process on the most liquid US stocks. In particular, we consider the stocks traded on the NYSE and on the NASDAQ from 01-Jul-2002 till 01-Jul-2013. Since we want to propose as much as possible a realistic empirical analysis, we have developed a dynamic dataset adopting the same filters adopted by Angelelli et al. (2013). In particular, the dataset uses all the stocks (from DataStream) which, at each recalibration time, are active during the last six months (125 trading days). Moreover, to limit the liquidity risk we use only the stocks whose daily value of volume traded overcome in average the 50 million of USD, where the value of volume traded of i-th asset at time t is given by:

$$Value\ of\ volume\ traded\ of\ i-th\ asset_t = P_{i,t}V_{i,t} \qquad (12)$$

where $P_{i,t}$ and $V_{i,t}$ are respectively the closure price and volume traded of i-th asset at time t.

Doing so, every month (at each recalibration time) we preselect always more than 500 assets in the US stock market which do not present liquidity risk. Since the number of these stocks is still too large, we need to reduce the dimensionality of the large scale portfolio problem. Thus, we preselect the first 100 assets with the greatest values of the performance measures (9) and (10) and, then, we approximate the preselected returns regressing the historical series on few factors obtained applying a proper PCA. In particular, at each recalibration time we consider 14 factors: 7 obtained with the PCA applied to the forecasted sub-Gaussian correlation matrix and the other 7 obtained with the PCA applied to the forecasted Pearson correlation matrix of the historical series.

A. A volume-return performance measure

We assume that the gross returns $Vol_{x,k}$ of the traded volume of a portfolio x evolves exactly as for the wealth in the Markovian case studied in Section II. So we start by an initial volume equal to 1 and we assume that the portfolio of volume with weights x follows a Markov chain with step $v_x := \left(\frac{\max_k Vol_{x,k}}{\min_k Vol_{x,k}}\right)^{1/N}$ and N states

$$q_x^{(i_x)} = \sqrt{b_{i_x} b_{i_{x-1}}} = \max_k Vol_{x,k}\left(\frac{\max_k Vol_{x,k}}{\min_k Vol_{x,k}}\right)^{\frac{1-2i_x}{2N}} =$$

$$= v_x^{\frac{1-2i_x}{2}} \max_k Vol_{x,k}.$$

Doing so we obtain that the joint probability follows formula (2) and the wealth-volume process at time t in the node $l = (l_x, l_{Vol_x}) \in L_t$ is given by

$$\begin{pmatrix} w_x^{(l_x,t)} \\ Vol_x^{(l_{Vol_x},t)} \end{pmatrix} = \begin{pmatrix} (z_x^{(1)})^t u_x^{1-l_x} \\ (q_x^{(1)})^t v_x^{1-l_{Vol_x}} \end{pmatrix}.$$

where $L_t := \{(l_x, l_{Vol_x}) : 1 \le l_x \le 1 + t(N-1), 1 \le l_{Vol_x} \le 1 + t(N-1)\}$.

In this section we consider a reward risk analysis that account of the joint return-volume evolution, but first we have to define some measures of reward and risk according to the investor behavior.

Considering that investors want to maximize the future wealth of the portfolio in particular when the volume traded is increasing we can assume as reward measure at time t the value

$$E(W_t(x) \mid W_t(x) \ge s_1; Vol_t(x) \ge s_2).$$

for some benchmarks (s_1,s_2). In the following empirical analysis we assume (s_1,s_2) the maximum values (for a given l_x) (s_1,s_2)=$(w_x^{(l_x,t)}, Vol_x^{(l_x,t)})$ such that

$$\Pr(W_t(x) \ge w_x^{(l_x,t)}; Vol_t(x) \ge Vol_x^{(l_x,t)}) >= 0.03 \qquad (13)$$

Considering that investors want to minimize the future losses of the portfolio in particular when the volume traded is increasing we can assume as risk measure at time t the value

$$E\big((1 - W_t(x))Vol_t(x) \mid W_t(x) \le s_1; Vol_t(x) \ge s_2\big).$$

for some benchmarks (s_1,s_2). In the following empirical analysis we assume (s_1,s_2) the first values (for increasing values of l_x) (s_1,s_2)=$(w_x^{(1+t(N-1)-l_x,t)}, Vol_x^{(l_x,t)})$ such that

$$\Pr(W_t(x) \le w_x^{(1+t(N-1)-l_x,t)}; Vol_t(x) \ge Vol_x^{(l_x,t)}) >= 0.003. \quad (14)$$

Therefore in the next section we use and apply the following reward-risk performance measure.

OA-ReturnVolume ratio (*OA-RVR*) This performance functional is defined as

$$\rho(W(x)) = \frac{\sum_{t=1}^{T} E(W_t(x) \mid W_t(x) \ge s_{1,t}; Vol_t(x) \ge s_{2,t})}{\sum_{t=1}^{T} E\big((1 - W_t(x))Vol_t(x) \mid W_t(x) \le m_{1,t}; Vol_t(x) \ge m_{2,t}\big)} \quad (15)$$

where $s_{1,t}$ and $s_{2,t}$ are defined by formula (13), while $m_{1,t}$ and $m_{2,t}$ are defined by formula (14) .

B. An empirical comparison

In our empirical analysis we use a dataset of more than ten years from 01-Jan-2003 till 01-July-2013, and assume the following settings:

a) that investors have a temporal horizon of $T = 20$ trading days (thus, for each portfolio strategy we should optimize the portfolio every 20 trading days for a total of 135 optimizations);
b) that investors cannot invest more than 70% in a single asset (i.e.: $x_i \in [0, 0.7]$);
c) Markov chains have $N = 5$ states;
d) the initial wealth W_0 is equal to 1 at the date 01-Jan-2003;
e) we consider 5 basis points of proportional transaction costs, applied at each recalibration time.

We perform a comparisons to evaluate the impact of return-volume approximation by comparing the ex-post performance of different portfolio strategies based on: the OA-Sharpe ratio (9), the OA-Stable Sharpe ratio (10), the OA-ReturnVolume ratio (15).

For each strategy, we have to compute the optimal portfolio composition 135 times and at the k-th optimization ($k = (0,1,2,\ldots,135)$), three main steps are performed to compute the ex-post final wealth:

Step 1 Preselect 100 assets among all those liquid and active in the last six months and then approximate these returns regressing them on few factors obtained by a PCA to reduce the randomness of the problem.

Step 2 Determine the market portfolio $x_M^{(k)}$ that maximizes the performance ratio $\rho(W(x))$ associated to the strategy, i.e. the "ideal" solution of the following optimization problem:

$$\begin{aligned} &\max_{x^{(k)}} \rho(W(x^{(k)})) \\ &s.t. \\ &(x^{(k)})' e = 1, \\ &x_i^{(k)} \le 0.7; x_i^{(k)} \ge 0 \\ &i = 1, \ldots, n \end{aligned}$$

Angelelli and Ortobelli (2009) have observed that the complexity of the portfolio problem is much higher in view of a Markovian evolution of the wealth process. In order to overcome this limit we use the Angelelli and Ortobelli's heuristic algorithm that could be applied to any complex portfolio selection problem that admit more local optima.

Step 3 Differently by Angelelli et al. (2013) we do not recalibrate daily the portfolio maintaining the percentages invested in each asset equal to those of the market portfolio $x^{(k)}$ during the period $[t_k, t_{k+1}]$ (where $t_{k+1} = t_k + T$). Thus, the ex-post final wealth is given by:

$$W_{t_{k+1}} = \big(W_{t_k} - t.c._{t_k}\big)(x_M^{(k)})' z_{(t_k+1)}^{(ex\ post)} \qquad (16)$$

where $t.c._{t_k}$ are the proportional transaction costs we get changing the portfolio, $z_{(t_{k+1})}^{(ex\ post)}$ is the vector of observed gross returns between t_k and t_{k+1}.

Steps 1, 2, and 3 are repeated for all performance ratios until some observations are available.

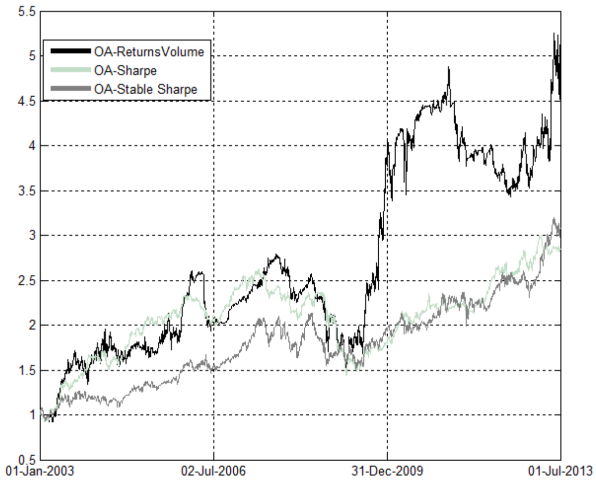


Fig.1: Ex-post comparison of OA-Sharpe ratio, OA-Stable Sharpe ratio, OA ReturnsVolume ratio applied to the preselected assets among all the active stocks in the last 6 months.

The output of this analysis is given in Figure 2 where the results of all strategies applied to the preselected assets among all the active assets in the last six months are reported.
The results obtained from the return-volume strategy (applied to the preselected liquid stocks) present very good results considering the transaction costs and the period of global crisis. In particular, the OA-return-volume ratio (best strategy) gives more than the 18% for year. We also observe that the OA-Stable Sharpe strategy presents the lowest level of risk with few small jumps and an almost linear behaviour. In addition, the OA-Stable Sharpe and the OA-return-volume strategies present higher final wealth than the OA-Sharpe strategy.
The portfolio composition generally changes a lot during the ex-post period. This is confirmed from Figure 3 that describes the portfolio turnover and its diversification. In particular, it examines how the portfolio composition of the OA-return-volume strategy changes during the ex-post period. In the first sub-figure (Fig. 3(a)) we have the percentages invested in each assets at each computation of the optimal portfolio.
The second sub-figure (Fig. 3(b)) points out the percentages Φ_k (k=1,...,25) of the portfolio changed every 20 days obtained by the formula:

$$\Phi_k = \sum_{i=1}^{n} \left| x_{M,i}^{(k)} - x_{M,i}^{(k-1)} \right|$$

In particular Φ_k should belong to the interval [0,2], where the value 0 means that the portfolio composition is not changed during the period $[t_{k-1},t_k]$, while the value 2 corresponds to the case we sell the portfolio and we buy a completely different portfolio. The last sub-figure (Fig. 3(c)) points out the number of:

(1) the quantity of assets used (i.e. those assets whose percentages are greater than zero $x_{M,i}^{(k)}>0$, i=1,...,n);
(2) the quantity of entering assets;
(3) the quantity of exiting assets.

Figure 3b shows a strong turnover in the portfolio composition since we often observe a value near to 2 of function Φ_k. Since the portfolio change a lot we pay high transaction costs every 20 days. Moreover, the portfolio is not very well diversified among all preselected assets because there are always some assets in which the strategy suggests (see Figure 3a) to invest the maximum possible (i.e. 70%). In addition, from Figure 3c we deduce that in all examined period there are no more than 25 assets (among the 100 preselected) with positive weight in the current portfolio.

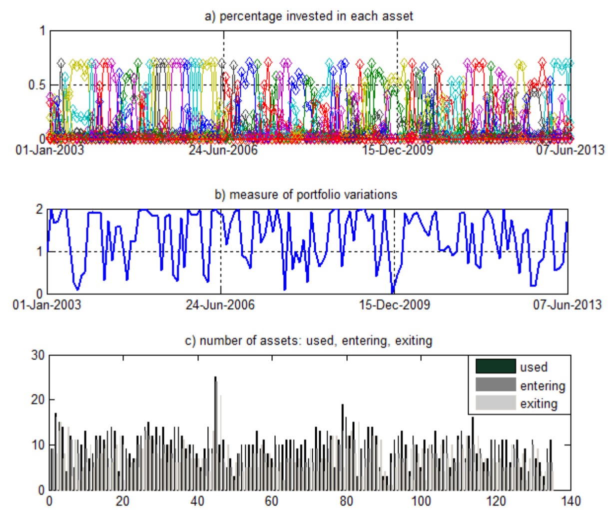


Fig. 3: Portfolio composition and portfolio variations of the OA-ReturnVolume strategy applied to preselected assets among all the active in the last 6 months.

V. Concluding Remarks

The proposed analysis emphasize the importance to use bivariate Markov chains either to reduce the dimensionality of large scale portfolio problems or to consider the joint behavior between the returns and the traded volume.

First we deduce that the preselection could play a crucial rule in portfolio selection, since in all cases we get more wealth than the indexes. This appears much more evident when we compare the final wealth of preselection portfolio strategies with uniform type strategies (strategies where the wealth is invested uniformly among all assets). Moreover, the preselection analysis also suggests that new firms entered in the market could have a strong impact in the portfolio selection.

Secondly, we observe that also the traded volume could play a crucial rule in portfolio selection strategies. In this framework we propose a new performance measure that is able to account properly the investor's return and volume preferences. This empirical analysis shows that it makes sense to account the joint distribution of the returns and of the volume.

ACKNOWLEDGMENT

The paper has been supported by the Italian funds ex MURST 60% 2013 and MIUR PRIN MISURA Project, 2013–2015. The research was also supported through the Czech Science Foundation (GAČR) under project 13-13142S and SP2013/3, an SGS research project of VSB-TU Ostrava, and further- more by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070), including the access to the supercomputing capacity, and the European Social Fund in the framework of CZ.1.07/2.3.00/20.0296 (second author).

REFERENCES

[1] Angelelli, E. and S. Ortobelli (2009), *American and European portfolio selection strategies: the Markovian approach*, in P. N. Catlere. (Ed.), Financial Hedging, Nova Science, New York, pp. 119-152.

[2] Angelelli, E., A. Bianchi and S. Ortobelli Lozza (2011), *Financial Applications of bivariate Markov processes*, Mathematical Problems in Engineering Volume 2011 Article ID 347604, 15 pages, doi:10.1155/2011/347604.

[3] Angelelli E., S. Ortobelli Lozza. and G. Iaquinta (2013), *Portfolio choice: a non parametric Markovian framework* in Mathematical Applications in Science & Mechanics (eds. Trisovic N., Rasteiro D., Tenorio A., Simian D., Minea A., Roushdy M., Salem A.) pp. 237-242, Wseas Press, ISBN: 978-960-474-305-6.

[4] Angelelli E., S. Ortobelli Lozza. and G. Iaquinta (2013a), *Dimensional portfolio reduction with asymptotic Markov processes* in Recent Advanced in Intelligent Control, Modelling & Computational Science(eds. S. Sendra, M. Leba, A. Ionica, K. Kalovrectis, O. Corbi) pp. 95-100, Wseas Press, ISBN: 978-960-474-319-3

[5] Cogneau, P. and G. Hübner (2009a), *The (more than) 100 Ways to Measure Portfolio Performance - Part 1: Standardized Risk-Adjusted Measures*, Journal of Performance Measurement, 13(4), 56-71.

[6] Cogneau, P. and G. Hübner (2009b), *The (more than) 100 Ways to Measure Portfolio Performance - Part 2: Special Measures and Comparison*, Journal of Performance Measurement, 14(1), 56-69.

[7] Cox J.C., S.A. Ross and M. Rubinstein (1979), *Option pricing: a simplified approach*, Journal of Financial Economics 7, 229-263.

[8] D'Amico G., G. Di Biase, J. Janssen and R. Manca (2010), *Semi-Markov Backward Credit Risk Migration. Models: a Case Study*, International Journal of Mathematical Models and Methods in Applied Sciences, 4(1), 82-92.

[9] Fama E. (1965), *The behavior of stock market prices*, Journal of Business 38, 34-105.

[10] Iaquinta, G. and S. Ortobelli (2006), *Distributional approximation of asset returns with nonparametric markovian trees* in International Journal of Computer Science & Network Security 6(11), 69-74.

[11] Iaquinta G., S. Ortobelli Lozza and E. Angelelli (2010), *The Markovian portfolio selection model with GARCH volatility dynamics* in (M. Pirtea, M. Mazilu, J. Strouhal eds.) Selected Topics in Economy & Management Transformation Vol. I, pp.382-387. Wseas Press, ISSN:1792-5983, ISBN: 978-960-474-240-0.

[12] Iaquinta G., S. Ortobelli Lozza and E. Angelelli (2011), *GARCH type portfolio selection models with the Markovian approach*, International Journal of Mathematical Models and Methods in Applied Sciences, 5(2), 308-315.

[13] Kondor I., S. Pafka and G. Nagy (2007), *Noise sensitivity of portfolio selection under various risk measures*, Journal of Banking and Finance 31, 1545-1573.

[14] Kring, S., S. Rachev, M. Höchstötter and F. Fabozzi, (2008), *Estimation of α-stable sub-Gaussian distributions for asset returns*, in G. Bol, S. T. Rachev and R. Würth (Eds.), Risk Assessment: Decisions in Banking and Finance, pp. 111-152.

[15] Mandelbrot B. (1963), *The variation of certain speculative prices*, Journal of Business 26, 394-419.

[16] McCulloch J.H. (1986), *Simple consistent estimators of stable distribution parameters*, Communications in Statistics - Simulation and Computation 15(4), 1109-1136.

[17] Ortobelli S., A. Biglova, S. Rachev and S. Stoyanov (2010), *Portfolio selection based on a simulated copula*, Journal of Applied Functional Analysis 5/2, 177-193.

[18] Ortobelli Lozza S., E. Angelelli and D. Toninelli (2011), *Set-portfolio selection with the use of market stochastic bounds*, Emerging Markets Finance and Trade 47(5), 5-24 .

[19] Papp G., S. Pafka, M.A. Nowak and I. Kondor (2005), *Random matrix filtering in portfolio optimization*. ACTA Physica Polonica B 36, 2757-2765.

[20] Rachev S. and S. Mittnik (2000), *Stable Paretian Models in Finance*, Wiley & Sons, New York.

[21] Rachev S., S. Mittnik, F. Fabozzi, S. Focardi and T. Jasic (2007), *Financial econometrics: from basics to advanced modeling techniques*, Wiley & Sons, Hoboken.

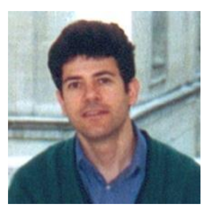
[22] Ross S. (1978), *Mutual fund separation in financial theory-the separating distributions*, Journal of Economic Theory 17, 254-286.

[23] Sadek A. and N. Limnios (2002), *Asymptotic properties for maximum likelihood estimators for reliability and failure rates of Markov chains*, Communication in Statistics-Theory and Methods.

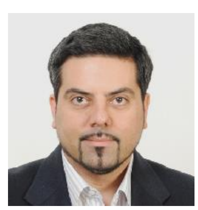
[24] Samorodnitsky G. and M. Taqqu (1994), *Stable non-Gaussian random processes: Stochastic models with infinite variance*, Chapman and Hall, New York.

[25] Sharpe W.F. (1994), *The Sharpe ratio*, Journal of Portfolio Management, 45-58.

Enrico Angelelli is an associate professor in Mathematical Finance at the University of Brescia.
He holds a Ph.D. in Computational Methods for Financial and Economic Forecasting and Decisions from the University of Bergamo.
He taught numerous courses at the University of Brescia, including Computer science for finance, Algorithms and computer programming, Information systems.
His research focuses on the application of operational research techniques to optimization problems ranging from portfolio selection to scheduling and transportation.



Sergio Ortobelli Lozza is an associate professor in Mathematical Finance at the University of Bergamo. He is also visiting Professor at VSB TU Ostrava Department of Finance, Czech Republic. He holds a Ph.D. in Computational Methods for Financial and Economic Forecasting and Decisions from the University of Bergamo. He taught numerous courses at the Universities of Bergamo, Calabria and Milan, including basic and advanced calculus, measure theory, stochastic processes, portfolio theory, and advanced mathematical finance. His research, published in various academic journals in mathematics and finance, focuses on the application of probability theory and operational research to portfolio theory, risk management, and option theory.



Gaetano Iaquinta is a private research analyst. He cooperates with financial institutions in Milan (Italy). He holds a Ph.D. in Computational Methods for Financial and Economic Forecasting and Decisions from the University of Bergamo and visiting Ph.D student in "Finance" at University of Lugano (Switzerland). His working areas and research interests include computational finance, option pricing theory, risk management, multistage stochastic optimization and scenario generation.