GARCH type portfolio selection models with the Markovian approach

Gaetano Iaquinta[1], Sergio Ortobelli Lozza[1], Enrico Angelelli[2]

[1]Department MSIA
University of Bergamo
24127-Via dei Caniana, 2; Bergamo - Italy;
e-mail addresses: sol@unibg.it; gaetano,iaquinta@unibg.it
[2]Dipartimento di Metodi Quantitativi
University of Brescia
25122 -C.da S. Chiara, 50; Brescia - Italy;
e-mail address: angele@eco.unibs.it

***Abstract*—** This paper describes different GARCH type portfolio models using a bivariate Markov process. In particular we approximate the GARCH process with a Markov chain in order to value the price/return distribution at the investor's temporal horizon. Then we discuss the computational complexity of the optimization problem and we implement an heuristic algorithm for the global optimum. Finally we propose an ex-post comparison among portfolio selection strategies based on reward/risk performance ratios.

***Keywords*—** GARCH models, Portfolio selection, Performance strategies, Ex-post analysis , Heuristic, Global optimization, Markov chains.

I. Introduction

IN this paper, we model the return portfolios with a Markov chain that account the GARCH evolution of the returns. In particular, we use the Duan and Simonato's approximation of the returns evolution (Duan, Simonato 2001) in portfolios selection problems. Under this distributional hypothesis we compare the ex-post performance of some portfolio selection strategies.

There is a general consensus on the importance to model the time varying volatility (Engle [1982], Bollerslev [1986]) and the leverage effect (Black [1976]). Several empirical studies have showed that these statistical aspects serve to solve many biases between theoretical and empirical prices (see Bakshi, Cao and Chen [1997], Engle and Mustafa [1992], and Heston and Nandi [2000]). Because there is wide consensus that the variance of the financial asset returns is time variant, a great amount of efforts are directing to realize mathematical models which, by choosing the variance dynamics as the model corner-stone, should be effectively able to model financial prices. Surely the GARCH model is a reference instrument to study the volatility dynamics, and among its advantages there is its high flexibility to be suitable to capture the most important features of the financial variables. In this work we analyze the impact of choices based on the GARCH parametric characterization of financial asset series. It is to note that the passage from the GARCH parametric characterization of financial asset series to the computation of the price/return distribution at some future time is not immediate. In order to build portfolio wealth distribution we use Duan and Simonato's GARCH approximation (Duan, Simonato 2001). Moreover, we extend the Duan and Simonato's ideas to other possible GARCH type models (see Glosten, et al.(1993), Nelson (1991)). As these authors explain many GARCH models and in particular the GARCH(1,1) models can be represented as a bivariate Markovian system (i.e., the state of the process is uniquely represented by price and variance states). This feature allows to approximate GARCH models by a discrete Markov chain. The Markovian and semi-Markovian models has been used in different fields of the financial literature typically in option pricing and credit risk (see, among others, Duan and Simonato (2001); D'Amico and Di Biase, (2009), D'Amico et al. (2009, 2010)), and in portfolio theory (see Angelelli and Ortobelli (2009), Iaquinta et al. (2010)). To build the transition matrixes we use the method discussed by Duan and Simonato (2001), Duan et al. (2003) for parametric Markovian processes.

With parametric portfolio selection models the transition matrix depends on the parameters of the underlying multivariate Markov process and the parameters are functions of the portfolio weights. Therefore we should check for a global optimum for most of the portfolio selection problems. In the paper we implement an optimization heuristic algorithm Angelelli and Ortobelli (2009) that reduces enormously the computational complexity with respect to other global optimization approaches like simulated annealing.

In the following empirical comparison, we present some portfolio selection strategies that use different GARCH

The authors thank for grants COFIN 60% 2010

models. All of them are based on the estimation of the distribution of the returns at future times under the assumption that the residuals of log returns portfolios follow a GARCH process.

The paper is organized as follows.

In Section 2 we show the models implemented. In Section 3 we discuss the Markovian approximation of portfolio value and we formalize the portfolio selection model discussing the computational complexity of the problem. In section 4 we perform an empirical comparison among different portfolio selection models. Finally, we briefly summarize the paper.

II. Portfolio Value With GARCH Volatility Dynamics

Let us consider a discrete-time economy and risky assets with log returns[1] $r_{t+1} = \left[r_{1,t+1},\ldots,r_{n,t+1}\right]'$. If we denote by $x = \left[x_1,\ldots,x_n\right]'$ the vector of the positions taken in the n risky assets, then the portfolio wealth at time $t+1$ is given by

$$W_{t+1}(x) = \sum_{i=1}^{n} x_i \, \mathrm{e}^{r_{i,t+1}} . \tag{1}$$

In particular, we assume that investors want to maximize the performance of their choices at a given future date *T*.

Now we introduce the alternative GARCH volatility dynamics models implemented in this work.

Suppose that under the historical measure P the daily portfolio log-return is described by the following relation:

$$r_{t+1} = \ln\left(\frac{W_{t+1}(x)}{W_t(x)}\right) = \mu + \sigma_{t+1}\varepsilon_{t+1} \tag{2}$$

where $W_t(x)$ is the portfolio value at time *t*, with asset position x, and $\varepsilon_{t+1} \mid \phi_t \sim (0,1)$ under P.

For convention we consider the initial portfolio wealth equal to 1 (i.e. $W_0 = 1$). In this work we use the standard GARCH(1,1) (see Bollerslev, T.(1986)) and some well-known extension of the standard GARCH(1,1). Each model can be represented as:

$$\sigma_{t+1} = f\left(\rho, \sigma_t, \varepsilon_t\right)$$

where the relation expresses that the conditional variance at time $t+1$ is function of the lagged value of the variance (σ_t), the lagged shock (ε_t) and a set of parameters (ρ).

The variance dynamics models we consider are:

Model I: GARCH (1,1) (G11): (see Bollerslev, T.(1986))

$$\sigma_{t+1}^2 = \omega + \beta\sigma_t^2 + \alpha\sigma_t^2\varepsilon_t^2 \tag{3}$$

where $\omega, \alpha > 0$ and $0 < \beta < 1$.

Model II: GJR-GARCH (GJR-G): (see Glosten, et al.(1993))

$$\sigma_{t+1}^2 = \omega + \beta\sigma_t^2 + \alpha\sigma_t^2\varepsilon_t^2 + \gamma\sigma_t^2\varepsilon_t^2 I_t \tag{4}$$

where $\omega, \alpha > 0$ and $0 < \beta < 1$ and $I_t = \begin{cases} 1; \varepsilon_t < 0 \\ 0; \text{otherwise} \end{cases}$

Model III: E-GARCH (E-G): (see Nelson (1991))

$$\ln\left(\sigma_{t+1}^2\right) = \omega + \beta\ln\left(\sigma_t^2\right) + \alpha\left(|\varepsilon_t| - \gamma\varepsilon_t\right) \tag{5}$$

where $\alpha > 0$ and $0 < \beta < 1$.

The parameters of the models are $\theta = (\mu, \rho)$ where μ is the constant drift term and $\rho = (\omega, \alpha, \beta, \gamma)$ is the parameter vector related to the variance dynamics.

In the model II and III the parameter γ allows to model the asymmetric behavior of the variance, sometime called Black's effect. It consists of a greater response of the variance when the news arrived in the market are negative ($\varepsilon_t < 0$) than when the news are positive ($\varepsilon_t > 0$).

All conditions on the parameters ω, α are used to avoid theoretical inconsistence on the value for σ_t^2 (i.e., $\omega < 0$ should mean a potential negative value for the variance, while $\alpha < 0$ should mean that greater shock movements induce a decreasing variance), while the conditions on β allow the variance process to be covariance-stationary. Let us consider the value of the stationary variance level in each model, supposing the weakly stationarity on σ_t^2 then we obtain[2] :

$$h^* = E\left(\sigma_t^2\right) = \frac{\omega}{1 - \beta - \alpha E^P\left(\varepsilon_t^2\right)} \qquad \text{in the G11}$$

$$h^* = E\left(\sigma_t^2\right) = \frac{\omega}{1 - \beta - \alpha E^P\left(\varepsilon_t^2\right) - \gamma E^P\left(\varepsilon_t^2 I_t\right)} \qquad \text{in the GJR}$$

$$q^* = E\left(\ln\sigma_t^2\right) = \frac{\omega - \alpha E^P\left(|\varepsilon_t| - \gamma\varepsilon_t\right)}{1 - \beta} \qquad \text{in the E-G}$$

Since we have to avoid the asymptotic divergence and the negativity of the variance process we need the following additional conditions:

- $\beta + \alpha E^P\left(\varepsilon_t^2\right) < 1$ in the G11
- $\beta + \alpha E^P\left(\varepsilon_t^2\right) + \gamma E^P\left(\varepsilon_t^2 I_t\right) < 1$ in the GJR

Note that in the normal innovation case (i.e., $\varepsilon_{t+1} \mid \phi_t \overset{P}{\sim} N(0,1)$)

$h^* = \dfrac{\omega}{1 - \beta - \alpha}$ in G11, $h^* = \dfrac{\omega}{1 - \beta - \alpha - \gamma/2}$ in GJR,

$h^* \cong \exp\left(q^*\right) = \exp\left(\dfrac{\omega - \alpha\sqrt{2/\pi}}{1 - \beta}\right)$ in E-G.

[1] Generally, we assume the standard definition of log return between time *t* and time *t+1* of asset *i*, as $r_{i,t+1} = \log\frac{S_{i,t+1} + d_{i,[t,t+1]}}{S_{i,t}}$, where $S_{i,t}$ is the price of the *i*-th asset at time *t* and $d_{i,[t,t+1]}$ is the total amount of cash dividends paid by the asset between *t* and *t+1*.

[2] We use the GARCH property that: $E\left(f(\sigma_t)g(\varepsilon_t)\right) = E\left(f(\sigma_t)\right)E\left(g(\varepsilon_t)\right)$ where f and g are some measurable function, since σ_t , and thus each its function, is measurable with respect to the sigma-algebra at time $t-1$

and the conditions are: $\beta+\alpha<1$ in the G11, $\alpha+\beta+\gamma/2<1$ in the GJR-G. In the E-GARCH model the value h^* is approximated with the exponential of the expected log volatility. The GARCH(1,1) model is a benchmark model and it is used for model comparisons. The main characteristics of other GARCH models used can be summarized in:

1. Parsimonious model, only 4 parameters to model the variance dynamics
2. Two state variables: price and variance
3. Time varying variance: GARCH models drive the variance process.
4. Models are potentially able to explain the well-known stylised-facts as the "Leverage effect" (γ parameter) and the "Clustering effect" in the stochastic volatility (β parameter)

III. Portfolio Value Dynamics As Markovian Evolution Process

Duan and Simonato have shown that the GARCH(1,1) model can be represented as a bivariate Markovian system (i.e., the state of the process is uniquely represented by $\left(W_t,\sigma_{t+1}^2\right)$ so the process is markovian of the first order). This feature allows to approximate GARCH models by a discrete Markov chain. Duan and Simonato's analysis can be extended to GJR-GARCH and E-GARCH models as we show here in the following. In particular, we present the Markov chain approximation of a GARCH (1,1) process (Duan, Simonato 2001) adapted to the work models.

Let us consider an underlying portfolio log-return modeled by the equation (2) or equivalently let us consider $\ln W_t(x)=\ln W_{t-1}(x)+\mu+\sigma_t z_t$ where W_t denote the portfolio value at day t. Let Q be some probability measure and σ_t be the variance modeled by G11, GJR or E-G. Let z_t a standardized random variable independently distributed with respect to the information up to time $t-1$, i.e., $\varepsilon_t \mid \phi_{t-1} \overset{Q}{\sim} (0,1)$.

Following Duan and Simonato's suggestions, we form the partitions by using the logarithm of adjusted wealth and log variance for the two state variables considered. The adjusted wealth is used to reduce the dimension of the transition matrix by a wealth conversion. The logarithms of the values used are justified mainly for its better convergence behavior.

The adjusted wealth is computed by $W_t^*=e^{-\tilde{\mu}t}W_t$ where and $\tilde{\mu}=\mu-h^*/2$ is the stationary variance, the pre-adjusted wealth can be easily recover later. Also the unconditional variance can be computed in all the GARCH model mentioned.

Note that in term of log-adjusted wealth the log return dynamics becomes:

$$\ln\left(\frac{W_t^*}{W_{t-1}^*}\right)=\ln\left(\frac{W_t}{W_{t-1}}\right)-\tilde{\mu}=\frac{1}{2}\left(h^*-\sigma_t^2\right)+\sigma_t\varepsilon_t .$$

The unconditional expectation of the continuously compounded return on the adjusted wealth is zero, since $E^Q\left(\sigma_t^2\right)=h^*$ and $E\left(\sigma_t\varepsilon_t\right)=0$.

Let p_t and q_t be the logarithm of the adjusted wealth (let us say log wealth) and the logarithm of the variance respectively (i.e., $p_t=\ln\left(W_t^*\right)$ and $q_t=\ln\left(\sigma_t^2\right)$) then the models can be rewritten with:

$$p_t=p_{t-1}+\frac{1}{2}\left(h^*-e^{q_t}\right)+\sqrt{e^{q_t}}\,\varepsilon_t$$

$$q_{t+1}=\ln\left(\omega+\beta e^{q_t}+\alpha e^{q_t}\varepsilon_t^2\right) \qquad \text{in the G11 case,}$$

$$q_{t+1}=\ln\left(\omega+\beta e^{q_t}+\alpha e^{q_t}\varepsilon_t^2+\gamma e^{q_t}\varepsilon_t^2 I_t\right) \qquad \text{in the GJR case}$$

or

$$q_{t+1}=\omega+\beta q_t+\alpha\left(\left|\varepsilon_t\right|-\gamma\varepsilon_t\right) \qquad \text{in the E-G case.}$$

To find a states partition to approximate the GARCH process we use:

1) A log wealth partition centered on the logarithmic of the initial portfolio wealth: $[p_0-I_p, p_0+I_p]$, where I_p is determined by studying the conditional behavior of the logarithm of the adjusted portfolio wealth over the investor time horizon T:

$$I_p=\delta_p(m)\sqrt{\sum_{t=1}^{T}E^Q\left(\sigma_t^2\mid\phi_0\right)} \tag{6}$$

2) An analytical formula of the conditional variance of the log wealth can be derived for many GARCH processes.

3) Log variance partition: to form the partition we would study the conditional behavior of the logarithm of the variance $q_T=\ln\left(\sigma_T^2\right)$. From the GARCH process features we know that there are two notable values of the variance:

a) the initial variance, which the process starts from,

b) the unconditional variance (h^*) to whom the process asymptotically is attracted.

Both these values have to be considered in the variance partition, but the second has increasing importance as we are far from the begin instant The partition center can be computed as: $q_1^*=\ln\left(\frac{\tau-\min(\tau,T)}{\tau}\sigma_1^2+\frac{\min(\tau,T)}{\tau}h^*\right)$. The value of τ is a temporal index used to form the weights. As it increases as the relative weight of the unconditional variance respect to the initial variance increases. Then in the study of long-term horizon τ has to be small. Anyway it is important to ensure that q_1 belongs to the partition. The log variance partition is $[q_1^*-I_q, q_1^*+I_q]$. In order to compute the width I_q of the partition it should be enough to study $Var^Q\left(q_T\mid\phi_0\right)$, but in-G11 and GJR-G it could result analytically complex. We know by the Jensen inequality that so Duan and Simonato propose to use a width $Var^Q\left(q_T\mid\phi_0\right)\le\ln\left(Var^Q\left(\sigma_T^2\mid\phi_0\right)\right)$:

$$I_q=\ln\left(e^{q_1}+\delta_q(n)\sqrt{Var^Q\left(\sigma_T^2\mid\phi_0\right)}\right)-q_1 \tag{7}$$

Only in the E-G case we have to note that the log variance

partition can be constructed directly by the E-GARCH equation, because it expresses the variance in logarithmic terms: $[q_1^* - I_q, q_1^* + I_q]$ where:

$$q_1^* = \frac{\tau - \min(\tau, T)}{\tau}\ln\left(\sigma_1^2\right) + \frac{\min(\tau, T)}{\tau}\ln\left(h^*\right) \text{ and}$$

$$I_q = \delta_q(n)\sqrt{\sum_{t=1}^{T} Var^Q\left(q_t \mid \phi_0\right)} \tag{8}$$

In the E-G the sum of the conditional variance up to T is given by[3]: $\sum_{t=1}^{T} Var^Q\left(q_T \mid \phi_0\right) = T\alpha^2\left(1-\gamma^2 - \frac{2}{\pi}\right)$

Duan and Simonato showed that $\delta_p(m) \xrightarrow{m\to\infty} \infty$ and $\frac{\delta_p(m)}{m} \xrightarrow{m\to\infty} 0$ are sufficient partition conditions for the approximating Markov chain to converge to its target GARCH process.

The logarithmic adjusted wealth partition and the logarithmic variance partition are equally divided in and odd parts respectively in order to determine the state of the bivariate process:

- $\bar{p}(i) = p_0 + \frac{2i-1-m}{m-1} I_p$ and the corresponding cells are $C(i) = [c(i), c(i+1))$ for $i = 1,...,m$, where $c(1) = -\infty$, $c(i) = \frac{\bar{p}(i-1) + \bar{p}(i)}{2}$ for $i = 2,...,m$ and $c(m+1) = +\infty$
- $\bar{q}(j) = q_1^* + \frac{2j-1-n}{n-1} I_q$ and the corresponding cells are $D(j) = [d(j), d(j+1))$ for $j = 1,...,n$, where $d(1) = -\infty$, $d(j) = \frac{\bar{q}(i-1) + \bar{q}(i)}{2}$ for $j = 2,...,n$ and $d(n+1) = +\infty$.

The Markov transition probability from state (i, j) at time t to state (k, l) at time $t+1$ is defined as

$\pi(i, j; k, l) = \Pr^Q\{p_{t+1} \in C(k), q_{t+2} \in D(l) \mid p_t = \bar{p}(i), q_{t+1} = \bar{q}(j)\}$ for $t = 0,...,T-1$.

It is typical in the GARCH(1,1) models that the variance at time $t+2$ is a deterministic function of the information set at time $t+1$. In particular in the models investigated we can write the variance as function of its lagged value, and two lagged wealth, i.e. $q_{t+2} = \Phi\left(q_{t+1}, p_{t+1}, p_t\right)$.

First we recover ε_{t+1} from the log price equation written one time forward:

$\varepsilon_{t+1} = \frac{p_{t+1} - p_t + \frac{1}{2}\left(e^{q_{t+1}} - h^*\right)}{\sqrt{e^{q_{t+1}}}}$ and substituting in the log variance equation we obtain:

$$\Phi^{G11}\left(q_{t+1}, p_{t+1}, p_t\right) =$$
$$= \ln\left(\omega + \beta e^{q_{t+1}} + \alpha\left(p_{t+1} - p_t + \frac{1}{2}\left(e^{q_{t+1}} - h^*\right) - \gamma\sqrt{e^{q_{t+1}}}\right)^2\right)$$

$$\Phi^{GJR}\left(q_{t+1}, p_{t+1}, p_t\right) =$$
$$\ln\left(\omega + \beta e^{q_{t+1}} + \left(\alpha + \gamma I_t\right)\left(p_{t+1} - p_t + \frac{1}{2}\left(e^{q_{t+1}} - h^*\right)\right)^2\right)$$

$$\Phi^{E-G}\left(q_{t+1}, p_{t+1}, p_t\right) =$$
$$\omega + \beta e^{q_{t+1}} + \alpha\left(\left|\frac{p_{t+1} - p_t + \frac{1}{2}\left(e^{q_{t+1}} - h^*\right)}{\sqrt{e^{q_{t+1}}}}\right| - \gamma\frac{p_{t+1} - p_t + \frac{1}{2}\left(e^{q_{t+1}} - h^*\right)}{\sqrt{e^{q_{t+1}}}}\right)$$

This implies a source of sparsity in the markovian transition matrix: for each combination of (i, j, k) it exists only an index l where the transition probability can be non zero. Thus we can rewrite the Markov transition probability as:

$$\pi(i, j, k, l) = \begin{cases} \Pr^Q\{p_{t+1} \in C(k) \mid p_t = \bar{p}(i), q_{t+1} = \bar{q}(j)\} \\ \qquad if\, \Phi\left(\bar{q}(j), \bar{p}(k), \bar{p}(i)\right) \in D(l) \\ 0, \text{otherwise} \end{cases}$$

The conditional probability can be computed as:

$$\Pr^Q\{p_{t+1} \in C(k) \mid p_t = \bar{p}(i), q_{t+1} = \bar{q}(j)\} =$$
$$= \Pr^Q\left\{c(k) \le \bar{p}(i) + \frac{1}{2}\left(h^* - e^{\bar{q}(j)}\right) + \sqrt{e^{\bar{q}(j)}}\,\varepsilon_{t+1} < c(k+1)\right\} =$$
$$= \Pr^Q\left\{\frac{c(k) - \bar{p}(i) - \frac{1}{2}\left(h^* - e^{\bar{q}(j)}\right)}{\sqrt{e^{\bar{q}(j)}}} \le \varepsilon_{t+1} < \frac{c(k+1) - \bar{p}(i) - \frac{1}{2}\left(h^* - e^{\bar{q}(j)}\right)}{\sqrt{e^{\bar{q}(j)}}}\right\}$$

Clearly these transition probabilities can be easily computed for any classical distributional assumption on the innovations. In this paper we use Gaussian distributed innovations.

Once we have computed the transition matrix M we can obtain the wealth distribution at time T considering the power of the transition matrix M^T. As a matter of fact, given the state (i,j) of the bivariate process (return, variance) corresponding to the k-th raw of the transition matrix, the distribution of the bivariate process at time T conditioned to start by (i,j) state is given by the k-th raw of the matrix M^T. Thus, to get the probability the log wealth is in the state "s" after T steps starting by (i,j) state we have to sum the probabilities for the different variance states, i.e., $\pi_T(i, j, s) = \sum_{i=1}^{n} \pi_T(i, j, s, l)$ where $\pi_T(i, j, s, l)$ is the probability corresponding to the k-th raw of the matrix M^T to go in the state (s,l) after T steps. Doing so we easily obtain the cumulative distribution of the forecasted final wealth for any described GARCH type model.

IV. Portfolio Value Dynamics As Markovian Tree Process

In the portfolio selection problem we assume the initial wealth $W_0 = 1$ and all admissible wealth processes $W(x) = \{W_t(x)\}_{t\ge 0}$ depending on an initial portfolio $x \in S$ are

[3] Note that in E-G case:
$Var^Q\left(q_t \mid \phi_0\right) = \alpha^2 E^Q\left(\left|\varepsilon\right| - \gamma\varepsilon - \sqrt{2/\pi}\right)^2 = \alpha^2\left(1 + \gamma^2 - \frac{2}{\pi}\right)$
where $\varepsilon \sim N(0,1)$

defined on a filtered probability space $\left(\Omega, \Im, \left(\Im_t\right)_{0\le t\le\infty}, \Pr\right)$. The portfolio selection problem when no short sales are allowed, can be represented as the maximization of a functional $f:\left(\Omega, \Im, \Pr\right)\mapsto\mathbb{R}$ applied to the random final wealth W_T(x) obtained with the portfolio weights belonging to the n-dimensional simplex

$$S=\left\{x\in\mathbb{R}^n \mid \sum_{i=1}^n x_i=1; x_i\ge 0\right\},$$

i.e., $\max_{x\in S} f(W_T(x))$

Typical examples are the performance measure type functionals $f(X)=\dfrac{\rho_1(-X)}{\rho_2(X)}$ where $\rho_1(.),\rho_2(.)$ are two positive increasing functions of coherent risk measures (see Rachev et al. (2008) and the reference therein). These functionals are isotone with the monotony order, but $\rho_1(.)$ and $\rho_2(.)$ are consistent with risk averse preferences. Thus, the functional $f(X)=\dfrac{\rho_1(-X)}{\rho_2(X)}$ is not isotone neither with risk lover nor with risk averse preferences. We refer to Rachev et al. (2008) for further examples of the above measures. Here in the following we introduce the two performance type measure used in the choice problem: the Sharpe ratio and the Rachev ratio.

Sharpe ratio (*SR*). The Sharpe ratio (see Sharpe (1994)) serves to value the expected excess return for unity of risk (standard deviation), i.e.,

$$\mathrm{SR}(X)=\frac{E(X-r_b)}{\sigma_{X-r_b}},$$

where r_b is a given benchmark and σ_{X-r_b} is the standard deviation of the random variable $X-r_b$. When the benchmark r_b is the riskfree rate and X is the portfolio return, the Sharpe ratio is isotonic with non-satiable risk averse preferences.

Rachev ratio This performance functional is defined as

$$\mathrm{OA-RR}_{(\alpha,\beta)}(W_T(x))=\frac{ETL_\alpha(W_T(r_b-z_{(x)})-1)}{ETL_\beta(W_T(z_{(x)}-r_b)-1)}$$

When the benchmark r_b is the riskfree rate and $z_{(x)}$ is the chosen portfolio gross return (i.e. $z_{(x)}=\exp(r_{(x)})$) and $W_T(z_{(x)}-r_b)-1$ is the final wealth at time T we obtain investing in the excess return $z_{(x)}-r_b$.

ETL_α is the Expected Tail Loss or Average Value at Risk (AVaR) which is a coherent measure defined as

$$ETL_\alpha(Y)=\frac{-1}{\alpha}\int_0^\alpha F_Y^{-1}(u)du$$

where $F_Y^{-1}(u)=\inf\left\{t\in\mathbb{R}/\Pr\left(Y\le t\right)\ge u\right\}$ is the left inverse of the distribution function. Recall that the classic consistent estimator of expected tail loss is given by

$$ETL_\alpha(Y)\approx\frac{-1}{\alpha T}\sum_{t=1}^T Y_t I_{[Y_t\le F_Y^{-1}(\alpha)]}$$

where $I_{[Y_t\le F_Y^{-1}(\alpha)]}=\begin{cases}1 \text{ if } Y_t\le F_Y^{-1}(\alpha)\\ 0 \text{ otherwise}\end{cases}$.

When the benchmark r_b is the riskfree rate, X is the portfolio return, and the numerator and the denominator are positive (negative), then the Rachev ratio is isotonic (consistent) with non-satiable preferences of investors who are neither risk averse nor risk lover (see Rachev et al. (2008)).

A. Computational Complexity and an heuristic for global optimization

Some recent studies (see Stoyanov et al. (2007), and Rachev et al. (2008)) have classified the computational complexity of reward-risk portfolio selection problems. In particular, Stoyanov et al. (2007) have shown that we can distinguish four cases of reward/risk ratios $\dfrac{f_{(1)}(X)}{f_{(2)}(X)}$ that admit an unique optimum in myopic strategies:

1. The ratio is a quasi-concave function when the risk functional $f_{(2)}(X)$ is convex and the reward functional $f_{(1)}(X)$ is concave.
2. The optimal ratio problem reduces to a convex programming problem when in addition to the conditions of point 1, both functions $f_{(1)}(X)$ and $f_{(2)}(X)$ are positively homogeneous.
3. The optimal portfolio problem reduces to a quadratic programming problem if in addition to the conditions of point 2, the reward function $f_{(1)}(X)$ is linear (or linearizable), and the risk function $f_{(2)}(X)$ is an increasing function of a quadratic form.
4. The optimal ratio problem reduces to a linear programming problem if the reward function $f_{(1)}(X)$ is linear and the risk function $f_{(2)}(X)$ is linearizable .

While the maximization of the Sharpe ratio can be solved as a quadratic type problem (it enters in the third category), the Rachev ratio (that is the ratio between two convex measures) is not included in this classification and it could present more local maxima. Moreover, when we approximate the bivariate process with a Markov chain, the transition matrix change with the portfolio weights x. Thus, the discretization process we adopt when we build up the approximating Markov chain implies that none of the above cases apply and the computational complexity increases. For example, Angelelli and Ortobelli (2009) have shown that non-parametric Markov portfolio models generally admit many local maxima even if it has to give a unique maximum as a consequence of the

monotony of the integral. Thus we can loss the monotony of the utility functional when we adopt our discretization process.

In order to illustrate the situation we consider 2000 historical observations of three components i j and k of the Dow Jones Industrial index and plot the values of different performance measures by varying the portfolio composition x in the 3-dimensional simplex $S = \{(x_i, x_j, x_k) \mid x_i + x_j + x_k = 1,$ $x_i, x_j, x_k \geq 0\}$. In particular, we consider *Rachev ratio* with temporal horizon $T = 20$ days. We consider a Markov chain with 28 states for the portfolio of returns and 6 states for its variance. Figure 1 reports the value of the Rachev ratio when we consider a GARCH(1,1) process. As we could expect the problem present more local maximum.

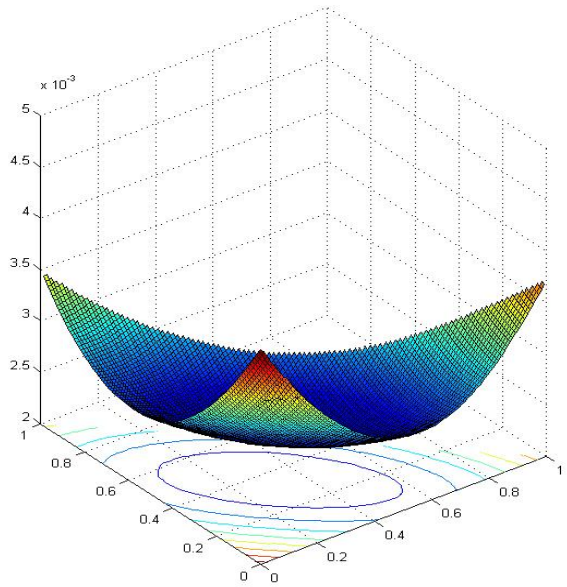

Figure 1. The Rachev ratio performance when we use a GARCH(1,1) model and we vary the composition of a 3 components portfolio.

In order to solve this global optimization problem we implemented a local search algorithm whose required input is an objective function f and an initial feasible solution $\bar{x}$ representing a portfolio from which the search is started. A current solution (portfolio) x is first defined as the initial solution at hand. Then the algorithm tries to iteratively update the current solution by a better one. Improving solutions, if any, are searched on a predefined grid of points fixed on the directions $x - e_i$ for $i = 1,2,...,n$ where x is the current portfolio and e_i is the portfolio where the share of asset i is equal to 1 and all other assets have share equal to 0. If a better solution is found on a search direction the current solution is updated and the search is continued from the new one. If no direction provides an improved solution the search ends. Actually, the search can be performed with two opposite orientations. Indeed the share of an asset i can be either increased or decreased. Accordingly, the algorithm, performs the search in three distinct steps. In the first step the algorithms tries to improve the current solution by increasing the share of assets i (i=1,…,n); in the second step the algorithm tries to improve the current solution by decreasing the share of assets i (i=1,…,n), in the last step the algorithm the first two steps changing iteratively the ordering of the control and the distance between each portfolio. More details are provided below. The general scheme of the algorithm is defined by the following MATLAB-like pseudo code.

```
function x = Optimize(f,xi)
   [x,improved] :=
   improveByIncreasingSingleAssets(f,xi);
   Do:
   [x,improved] :=
   improveByDecreasingSingleAssets(f,x);
   if improved
         [x,improved] :=
         improveByIncreasingSingleAssets(f
         ,x);
   end if
   While improved
   return x;
```

More in details, procedure `improveByIncreasingSingleAssets` tries to improve the current solution x by iteratively increasing the share of single assets in the portfolio. The basic idea is to choose an asset i such that x_i>0, define a finite set of alternative portfolios

$$x'(\beta) = (1-\beta)x + \beta e_i$$

where parameter β is assigned values $\beta_h = \left(\frac{h}{m}\right)^p$ for $h = 1,...,m.$ with $p \geq 1,$ and $m \in \mathbb{N}.$ If there is a h such that $f(x(\beta_h)) > f(x),$ then the current portfolio x is updated by $x(\beta_{h^*})$ where $h^* = \arg\left(\max_h \{f(x(\beta_h))\}\right)$. The integer m defines the number of points in which the objective function will be evaluated, whereas the index $p \geq 1$ defines how the points are distributed on the simplex $(1-\beta)x + \beta e_i$. In particular, for $p = 1$ the points are equidistributed, while if p gets larger the points get more concentrated around x. The directions in the simplex are searched according to the increasing value of i. Without loss of generality we can assume that $f(e_i) > f(e_j)$ for any $i < j$ so that the attempt to increase the share of the asset with highest performance is made first. If a better solution is found on a search direction $x - e_k$ for some k, the current solution x is updated by the new solution and the search is continued on the new directions $x - e_i$ for $i = k+1,...,n,1,2,...,k-1.$ If no direction provides an improved solution the search is stopped. The procedure returns the best solution found, possibly the initial one, and a flag indicating whether an improvement has been achieved during the search.

Similarly, procedure `improveByDecreasingSingleAssets` tries to improve the current solution x by decreasing the share of an asset i such that $0 < x_i < 1$. The scheme is the same as in procedure `improveByIncreasingSingleAssets`, but for any chosen asset i, the set of alternative portfolios $x'(\beta)$ is defined as

$$x'(\beta) = (1-\beta)x + \beta \cdot P_i(x)$$

where $P_i(x)$ is the projection of the portfolio x from portfolio e_i on the hyperplane

$$\begin{cases} \sum_k x_k = 1 \\ x_i = 0 \end{cases}$$

which can be obtained by

$$P_i(x) = \frac{x - (x_i e_i)}{1 - x_i}.$$

The main advantages of this algorithm are:

1. The algorithm permits to approximate the global optimum with an error of $\left(\frac{1}{m}\right)^p$ when the objective function is a non-constant concave function (the optimum is unique) and the lines $x'(\beta)$ are not particular contour lines of the objective function[4].
2. The algorithm checks the m points β_h on the lines $x'(\beta_h)$ of the n-dimensional simplex. So, we can better explore the whole simplex and approximate the global optimum.
3. The computational complexity is much less than that of classic algorithms for global optimum such as Simulated Annealing type algorithms (see Angelelli and Ortobelli 2009).

V. An Ex-Post Empirical Comparison among GARCH type models

In this section, we compare portfolio selection strategies based on the GARCH models introduced in the previous sections. We use 32 assets quoted on the US markets (NYSE and NASDAQ) from 01/02/97 till 06/14/2010 for a total of 3384 daily observations. We compare the performance of:

1) Rachev ratio under the hypothesis the log wealth follows or a *GARCH(1,1)*, or a *GJR-GARCH* or an *E-GARCH*.

2) Sharpe ratio (see Sharpe 1994) under the assumption we consider historical iid returns.

We recalibrate daily the portfolio and for the dynamic strategies we use a temporal horizon T=20 working days.

[4]However, we can still approximate the optimum by updating the solution x with a point $(1-\beta_h)x + \beta_h e_i$ choosing an h among $h = 1,\ldots,m-1$ any time the lines $(1-\beta)x + \beta e_i$ with $\beta \in 0,1]$ are particular level curves of the concave objective function.

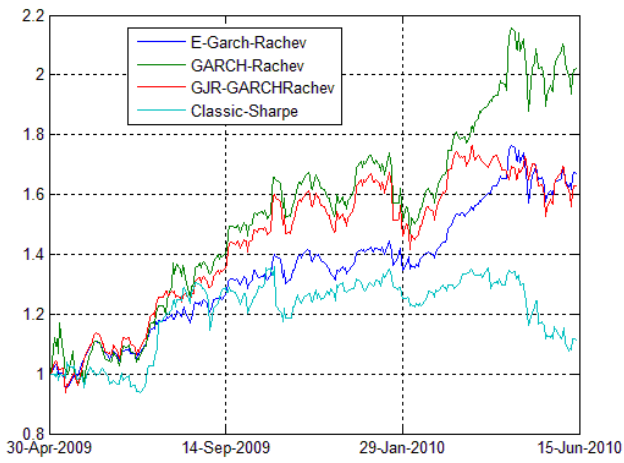

Figure 2. Ex-post final wealth process when European strategies are applied with daily recalibration and temporal horizon T=20 days.

We forecast the future wealth using 28 states for the portfolio of returns and 6 states for its variance. As coefficents of AVaR in the Rachev ratio we use $\alpha = \beta = 0.05$. The comparison consists in the ex post evaluation of the wealth produced by the strategies. For each strategy, we consider an initial wealth $W_0 = 1$ at the date 04/30/2009, and at the k th recalibration ($k = 0,1,2,\ldots$), three main steps are performed to compute the ex-post final wealth:

Step 1 Determine the market portfolio $x_M^{(k)}$ that maximizes the performance ratio $\rho(W(x))$, i.e. the solution of the following optimization problem:

$$\max_{x^{(k)}} \rho(W(x^{(k)}))$$

s.t.

$$\left(x^{(k)}\right)' e = 1,$$

$$x_i^{(k)} \geq 0; \quad i = 1,\ldots,n.$$

As shown by Angelelli and Ortobelli (2009) this type of problems could present more local optimum then we use the heuristic developed from them to approximate the global optimum.

Step 2 Determine the ex-post final wealth given by:

$$W_{t_{k+1}} = W_{t_k}\left(\left(x_M^{(k)}\right)' z^{(ex\ post)}\right),$$

where $z^{(expost)}$ is the vector of observed gross returns between t_k and t_{k+1}.

Step 3 The optimal portfolio $x_M^{(k)}$ is the new starting point for the $(k+1)$-th optimization problem.

Steps 1, 2 and 3 are repeated until the observations are available and for each performance ratio.

The output of this analysis is represented in Figure 2. Figure 2 reports the ex -post wealth process using different GARCH models. In particular these results emphasize the

good performance of the classic GARCH(1,1) model that in the last year present earnings of about the 100%. Instead the other GARCH models are almost never comparable to the classic one. However the comparison with static classic strategy is amazing and it suggests us that we should never use the classic strategies in portfolio choices.

Thus, the empirical results show that volatility GARCH models could be very important in portfolio theory.

VI. Conclusion

This paper examines the impact of GARCH type return evolution in portfolio selection problems. We describe how to approximate GARCH type processes with Markov chains and we deal the portfolio selection problem under these distributional assumptions. Thus we propose algorithms that permit to solve computationally complex problems in acceptable computational times. Finally, we propose an empirical comparison among the myopic portfolio selection models and those based on the GARCH approximation. The ex-post empirical comparison among classic approaches and those based on Markovian trees shows the greater predictable capacity of the latter.

References

[1] Angelelli, E. and Ortobelli, S. (2009) "American and European Portfolio Selection Strategies: the Markovian Approach" Chapter 5 in "Financial Hedging" edited By Patrick N. Catlere, 119-152.

[2] Bakshi, G., C. Cao, and Z. Chen, (1997) "Empirical Performance of Alternative Option Pricing Models," Journal of Finance, 52, 2003-2049.

[3] Black, Fisher, (1976) "Studies of Stock Price Volatility Changes," in: Proceedings of the 1976 Meetings of Business and Economic Statistics Section, American Statistical Association, 177-181

[4] Bollerslev, Tim, (1986) Generalized Autoregressive Conditional Heteroskedasticity., Journal of Econometrics. April, 31:3, pp. 307–27.

[5] D'Amico, G., G. Di Biase, (2009) Dynamic Concentration /Inequality Indices of Economic Systems. In proceedings of the International Conference "Recent Advances in Applied Mathematics" (C.A. Bolucea, V. Mladenov, E. Pop, M. Leba, N. Mastorakis, Eds.), (Tenerife, Spain), pp. 312-316, Wseas Press

[6] 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.

[7] D'Amico G., G. Di Biase, J. Janssen, R. Manca, (2009) Semi-Markov Backward Credit Risk Migration Models Compared with Markov Models. In proceedings of the 3RD International Conference on Applied Mathematics, Simulation, Modelling (P. Pardalos, N. Mastorakis, V. Mladenov and Z. Bojkovic, Eds.), (Athens, Greece), pp. 112-116, Wseas Press

[8] Duan, J. and J. Simonato, (2001) American option pricing under GARCH by a Markov chain approximation., Journal of Economic Dynamics and Control 25, 1689--1718.

[9] Duan, J., E. Dudley, G. Gauthier, and J. Simonato, (2003) Pricing discretely monitored barrier options by a Markov chain, Journal of Derivatives 10, 9--23.

[10] Engle, Robert F., (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation., Econometrica. 50:4, pp. 987–1007.

[11] Engle, R.F. and C. Mustafa, (1992) Implied ARCH Models from Options Prices., Journal of Econometrics, 52, 289-311.

[12] Glosten, Lawrence R., R. Jagannathan and D. E. Runkle, (1993) On the Relation Between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks., Journal of Finance. 48:5, pp. 1779–801.

[13] Heston, S., and S.Nandi, (2000) A closed-form GARCH option valuation model., Review of Financial Studies, 13, 433-451.

[14] Iaquinta G., Ortobelli S., Angelelli E. (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.

[15] Nelson, Daniel B., (1991) Conditional Heteroskedasticity in Asset Returns: A New Approach., Econometrica. 59:2, pp. 347–70.

[16] Rachev, S.,S. Ortobelli, S. Stoyanov, F. Fabozzi, and A. Biglova (2008) Desirable properties of an ideal risk measure in portfolio theory., International Journal of Theoretical and Applied Finance 11(1), 19-54.

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

[18] Stoyanov, S., Rachev, S. and Fabozzi, F. (2007). Optimal financial portfolios, *Applied Mathematical Finance* 14(5), 401-436.