Jumps in Rank and Expected Returns Introducing Varying Cross-sectional Risk

Gloria González-Rivera*
Department of Economics

Tae-Hwy Lee
Department of Economics

Santosh Mishra
Department of Economics
University of California, Riverside
Riverside, CA 92521

Abstract

Decision theorists claim that an ordinal measure of risk may be sufficient for an agent to make a rational choice under uncertainty. We propose a measure of financial risk, namely the **V**arying **C**ross-sectional **R**isk (VCR), that is based on a ranking of returns. VCR is defined as the probability of a sharp jump over time in the position of an asset return within the cross-sectional return distribution of the assets that constitute the market, which is represented by the Standard and Poor's 500 Index (SP500). We model the *joint* dynamics of the cross-sectional position and the asset return by analyzing (1) the *marginal* probability distribution of a sharp jump in the cross-sectional position within the context of a duration model, and (2) the probability distribution of the asset return *conditional* on a jump, for which we specify different dynamics in returns depending upon whether or not a jump has taken place. As a result, the marginal probability distribution of returns is a mixture of distributions. The performance of our model is assessed in an out-of-sample exercise. We design a set of trading rules that are evaluated according to their profitability and riskiness. A trading rule based on our VCR model is dominant providing superior mean trading returns and accurate estimation of the Value-at-Risk.

Key words: ARCH, CAPM, Duration, Nonlinearity, Trading rule, VaR.

JEL Classification: C3, C5, G0.

*Corresponding author: gloria.gonzalez@ucr.edu, tel. 1-909-827 1590, fax 787 5685.

1 Introduction

Economists, investors, regulators, decision makers at large face uncertainty in a daily basis. While there is an intuitive notion, probably shared by most of us, on the meaning of uncertainty, which involves the realization of future events within a probabilistic understanding of the world, there is no agreement on how to quantify uncertainty such as becomes an operational measure. The measure of uncertainty depends on who you are and what you do. Financial economists and econometricians for most part equate uncertainty with risk and risk with volatility. We are very familiar with measures such as variance, range, absolute deviation; in general, any measure that can summarize the dispersion of the random variable under analysis. Investors and regulators are not only concerned with measures of volatility but also they monitor the lower tail of the probability distribution of returns. Regulators worry about catastrophic or large losses that can jeopardize the health of the financial system of the economy. On the other hand, decision theorists deal with volatility measures only in particular instances; they prefer to analyze rational choices considering the entire probability distribution of the random variable in question.

Granger (2002) provided very illuminating comments on risk. He reviewed the statistical foundations for the choice of the variance and the mean absolute deviation as appropriate measures of volatility and contrasted these results with the rejection by uncertainty theorists of these measures as appropriate measures of risk. Rational choice under uncertainty requires a connection to an objective function, such as an utility function, for which a variance may not summarize completely the risk faced by an agent. Decision theorists claim that an ordinal measure of risk, i.e. the constructing a ranking of assets, may be sufficient for an agent to make a rational choice under uncertainty. The empirical question is how these theoretical results can help decision makers when they face numerous time series of historical asset prices.

The contribution of this paper is framed in a middle ground between variance advocates and decision theorists. We develop an empirical model of financial returns that widens the conception of risk maintained by financial economists. Our model combines a cardinal measure of risk -conditional volatility- with an ordinal measure -the cross-sectional position of an asset in relation to its peers. That is to say, we *jointly* consider a time-varying coordinate and a cross-sectional coordinate for each asset return. Fixing the time coordinate, we observe the market as a collection of assets returns, and every asset is assigned a cross-sectional position or percentile. From period to period, the position of the asset in relation to the cross-section of assets changes. Our objective is to model the dynamics of the cross-sectional position jointly with the asset return. In Figure 1, we present a stylized description of the problem that we aim to analyze. Let y_{it} be the asset return of firm i at time t, and z_{it} the cross-sectional position of this return, i.e. $z_{it} = \Phi_{cs}(y_{it})$ where $\Phi_{cs}(.)$ is the cumulative distribution function of all asset returns that constitute the market at time t. In Figure 1, for every t, we draw the probability density function of all assets, this is the cross-sectional market distribution, which is time-varying. To illustrate the different dynamics of y_{it} and z_{it}, we choose four points in time. Consider the sequential movements of y_{it} and z_{it} on going from t_1 to t_4. We observe that from t_1 to t_2, the market overall has gone down as well as the return and the cross-sectional position of asset i, $y_{t_1} > y_{t_2}, z_{t_1} > z_{t_2}$. However, from t_2 to t_3, the asset return is lower $y_{t_2} > y_{t_3}$ but its cross-sectional position has improved $z_{t_2} < z_{t_3}$. One may say that this asset has become riskier but, in relation to its peers, is less risky than it used to be. The opposite happens on going from t_3 to t_4. The overall market is going up; for asset i, the return increases $y_{t_3} < y_{t_4}$ but its cross-sectional position is unchanged $z_{t_3} = z_{t_4}$ One may say that now it is less risky but, in relation to its peers, is as risky as it used to be. We are interested in this notion of relative risk. We model the conditional probability of jumping cross-sectional positions, which we call time-**V**arying **C**ross-sectional **R**isk (VCR). It is time-varying because it depends on an information set that changes over time, it is cross-sectional because it depends on the position of the asset in relation to its peers, and it is risk because it is an assessment of the chances of being a winner or a looser within the available set of assets.

This notion of relative risk speaks to the idea that assets returns are related to each other and that risk of an individual asset cannot be understood by solely examining the univariate stochastic process of the asset return. This is the core of the intellectual contribution of the founding fathers of portfolio theory, Markowitz (1959), Sharpe (1964) and Lintner (1965). In the classical Capital Asset Pricing Model (CAPM), the cornerstone of modern portfolio theory, the variance of a portfolio is a sufficient statistic to measure risk, the covariance between any two asset is a sufficient measure of interdependence, risk is fully characterized by the beta of the asset, and expected returns are linear functions of beta. This complete world comes to the expense of very restrictive assumptions, such as quadratic utility functions or normality of asset returns. In this paper, while we do not subscribe to any particular set of assumptions or any particular asset pricing theory, we are faithful, from an empirical perspective, to the idea of interdependence or comovement among assets, which cannot be summarized by a particular moment of the portfolio returns such as the variance. Comovements are brought into our modelling strategy as the time-varying cross-sectional distribution of asset returns. In this general setting, expected returns are nonlinear functions of this notion of relative risk. Linearity is not excluded since it can be viewed as a particular case of a more general model, it just becomes a testable proposition.

When theories as CAPM or APT are empirically tested (Fama and French, 1993, 1996), the most conventional avenue has been to estimate and test in two stages[1]. At the risk to oversimplify an extended body of work, let us summarize the procedure in a couple of bold strokes. First, researchers run a time series regression in cleverly manipulated data sets, out of which, the beta of the asset return or portfolio return is obtained. Secondly, they gather a cross-section of assets or portfolios and run a cross-sectional regression where the betas are the regressors. These regressions -time series and cross-sectional- can be more or less sophisticated in order to take care of a myriad of econometric problems that partly arise from the two-stage procedure. In this paper, the estimation procedure is performed in one-stage. Our analysis is primarily a time series exercise because our ultimate objective is to forecast returns but, concurrently, we also incorporate a cross-sectional dimension. We set our problem as to model the joint distribution of the return and the probability of a (sharp) jump (J_{it}) in the cross-sectional position of the asset, i.e. $f(y_{it}, J_{it}|\Im_{t-1})$ where $\Im_{t-1}$ is an information set up to time $t-1$. Since $f(y_{it}, J_{it}|\Im_{t-1}) = f_1(J_{it}|\Im_{t-1})f_2(y_{it}|J_{it}, \Im_{t-1})$, our task will be accomplished by modelling the conditional distribution of the return and the marginal distribution of the jump. One of the implications of this setting is that the marginal distribution of returns is a mixture of distributions, which may explain the unconditional leptokurtosis that characterize financial asset returns.

On modelling $f_1(J_{it}|\Im_{t-1})$, our paper also connects with the recent literature in microstructure of financial markets and duration analysis (Engle and Russell, 1998). This line of research aims to model events (trades) and waiting times between trades. The question of interest is what is the expected length of time between two trades given some information set. In this paper, the event is the jump in the cross-sectional position of the asset return, however, when we model the expected duration between jumps or its mirror image -the conditional probability of the jump-, our analysis is

[1] For other approaches such as estimation of conditional CAPM, see Jagannathan and Wang (1996).

performed in calendar time as in Hamilton and Jordà (2002). Given some information set, the question of interest is what is the likelihood that tomorrow the return of a given firm is such that there is a sharp change in the position of this firm in relation to the cross-sectional distribution of returns. This calendar time approach is necessary because asset returns are reported in calendar time (days, weeks, etc.) and it has the advantage of incorporating all information that becomes available in each period of time.

The performance of our model is assessed in an out-of-sample exercise within the context of investment decision making. We consider two scenarios. In the first, we deal with an investor whose interest is to maximize profits of a portfolio long in stocks. This scenario may be marginally related to a momentum strategy where the investor is going long on winners and short on losers (Jegadeesh and Titman, 2001, 2002), however the aim of the momentum literature is different from ours in that it searches for macroeconomic factors that can explain abnormal returns obtained with a momentum strategy. The second scenario that we consider is an investor who worries about potential large losses, she wishes to add a Value-at-Risk evaluation to her trading strategy. Profitability and riskiness are the two coordinates in the mind of the investor. We design a set of trading rules that will be compared in the two aforementioned scenarios. The statistical comparison is performed within the framework of White (2000) reality check. A trading rule that exploits the one-step ahead forecast of the ranking of asset returns will be shown to be clearly superior to other rules based on more standard models.

The organization of the paper is as follows. In section 2, we provide our strategy for the joint modelling of asset returns and jumps in the cross-sectional position. We present the estimation results for a sample of weekly returns of those SP500 firms that have survived for the last ten years. In section 3, we assess the out-of-sample performance of our model. We explain the trading rules, loss functions, and the statistical framework to compare different trading rules. Finally, in section 4 we conclude.

2 Cross-sectional position and expected returns

In this section, we propose a bivariate model of *expected returns* and *jumps* in the ranking of a given asset within the cross-sectional distribution of asset returns.

Let y_{it} be the return of the i^{th} firm at time t, and $\{y_{it}\}_{i=1}^{M}$ be the collection of asset returns of the M firms that constitute the *market* at time t. For every time period, we order the asset returns from the smallest to the largest, and we define z_{it}, the *cross-sectional position* of the i^{th} firm within the market, as the percentage of firms that have a return less or equal to the return of the i^{th} firm. We write

$$z_{it} = M^{-1} \sum_{j=1}^{M} \mathbf{1}(y_{it} \geq y_{jt})$$

for $z_{it} \in [M^{-1}, 1]$ and where $\mathbf{1}(.)$ is the indicator function. We say that a sharp jump in the cross-sectional position of the i^{th} firm has occurred when there is a minimum (upwards or downwards) movement of 0.5 in the ranking of the return of the i^{th} firm. We define such a jump as a binary variable $J_{it} = \mathbf{1}(|z_{it} - z_{it-1}| \geq 0.5)$. The choice of the magnitude of the jump is not arbitrary. The sharpest jump that we could consider is 0.5. In every time period, we need to allow for the possibility of jumping, either up or down, in the following period regardless of the present position of the asset. For instance, if we choose a jump greater than 0.5, say 0.7, and $z_{it} = 0.4$, then the probability of jumping up or down in the next time period is zero with probability one. Note that the defined jump does not imply that the return will be above or below the median. As an example, if $z_{it-1} = 0.4$ and $z_{it} = 0.6$, then $J_{it} = 0$ but the return at time t will be above the cross-sectional median of returns.

Our objective is to model the conditional joint probability density function of returns and jumps $f(y_{it}, J_{it}|\Im_{t-1}; \Theta)$, where $\Im_{t-1}$ is the information set up to time $t-1$, which contains past realizations of returns, jumps, and cross-sectional positions. To simplify notation, we drop the subindex i but in the following analysis should be understood that the proposed modelling is performed for every single firm in the market. We factor the joint probability density function as the product of the conditional density of the return and the marginal density of the jump

$$f(y_t, J_t|\Im_{t-1}; \Theta) = f_1(J_t|\Im_{t-1}; \theta_1) f_2(y_t|J_t, \Im_{t-1}; \theta_2)$$

For a sample $\{y_t, J_t\}_{t=1}^{T}$, the joint log-likelihood function is

$$\sum_{t=1}^{T} \log f(y_t, J_t|\Im_{t-1}; \Theta) = \sum_{t=1}^{T} \log f_1(J_t|\Im_{t-1}; \theta_1) + \sum_{t=1}^{T} \log f_2(y_t|J_t, \Im_{t-1}; \theta_2)$$

Let us call $L_1(\theta_1) = \sum_{t=1}^{T} \log f_1(J_t|\Im_{t-1}; \theta_1)$ and $L_2(\theta_2) = \sum_{t=1}^{T} \log f_2(y_t|J_t, \Im_{t-1}; \theta_2)$. The maximization of joint log-likelihood function can be achieved maximizing $L_1(\theta_1)$ and $L_2(\theta_2)$ separately without loss of efficiency because the parameter vectors θ_1 and θ_2 are "variation free" and J_t is weakly exogenous (Engle, 1983).

2.1 Modelling the cross-sectional jump $f_1(J_t|\Im_{t-1};\theta_1)$

In order to model the conditional probability of jumping, we define a counting process $N(t)$ as the cumulative number of jumps up to time t, that is, $N(t) = \sum_{n=1}^{t} J_n$. This is a non-decreasing step function that is discontinuous to the right and to the left and for which $N(0) = 0$. Associated with this counting process, we define a duration variable $D_{N(t)}$ as the number of periods between two jumps. Note that because our interest is to model the jump jointly with returns and these are recorded in a calendar basis (daily, weekly, monthly, etc.), the duration variable needs to be defined in calendar time instead of event time as is customary in duration models. For this reason, it is very likely that the duration between two jumps could remain constant for several time periods. The question of interest is, what is the probability of a sharp jump at time t in the cross-sectional position of the i^{th} firm asset return given all available information up to time $t-1$? This is the conditional hazard rate p_t

$$p_t = \Pr(J_t = 1|\Im_{t-1}) = \Pr(N(t) > N(t-1)|\Im_{t-1}) \tag{1}$$

It is easy to see that the probability of jumping and duration must have an inverse relationship. If the probability of jumping is high, the expected duration must be short, and viceversa. Following Hamilton and Jordà (2002), we specify an autoregressive conditional hazard (ACH)[2] model for (1). Let $\Psi_{N(t)}$ be the expected duration. The expected duration until the next jump in the cross-sectional position is given by $\Psi_{N(t-1)} = \sum_{j=1}^{\infty} j(1-p_t)^{j-1}p_t = p_t^{-1}$. Consequently, to model (1), it suffices to model the expected duration and compute its inverse. A general ACH model is specified as

$$\Psi_{N(t)} = \sum_{j=1}^{m} \alpha_j D_{N(t)-j} + \sum_{j=1}^{r} \beta_j \Psi_{N(t)-j} \tag{2}$$

Since p_t is a probability, it must be bounded between zero and one. This implies that the conditional duration must have a lower bound of one. Furthermore, working in calendar time has the advantage that we can incorporate information that becomes available between jumps and can affect the

[2] The ACH model is a discrete-time version of the autoregressive conditional duration (ACD) of Engle and Russell (1998).

probability of jumping in future periods. We can write a general conditional hazard rate as

$$p_t = [\Psi_{N(t-1)} + \delta^{'} X_{t-1}]^{-1} \tag{3}$$

where X_{t-1} is a vector of relevant calendar time variables such as past cross-sectional positions and past returns.

The log-likelihood function $L_1(\theta_1) = \sum_{t=1}^{T} \log f_1(J_t|\Im_{t-1};\theta_1)$ corresponding to a sample of observed jumps in the cross-sectional position is

$$L_1(\theta_1) = \sum_{t=1}^{T} [J_t \log p_t(\theta_1) + (1 - J_t) \log(1 - p_t(\theta_1))] \tag{4}$$

where $\theta_1 = (\alpha', \beta', \delta')'$ is the parameter vector for which the log-likelihood function is maximized.

2.2 Modelling the conditional return $f_2(y_t|J_t, \Im_{t-1}; \theta_2)$

We assume that the return to the i^{th} firm asset may behave differently depending upon the occurrence of a sharp jump. If a sharp jump has occurred, the return was pushed either towards the lower tail or upper tail of the cross-sectional distribution of returns. In relation to the market, this asset becomes either a looser or a winner. On the other hand, if there is no jump, the return keeps its cross-sectional position in relation to the market. *A priori,* one may expect different dynamics in these two states. A general specification is

$$f_2(y_t|J_t, \Im_{t-1}; \theta_2) = \begin{cases} N(\mu_{1t}, \sigma_{1t}^2) & \text{if } J_t = 1 \\ N(\mu_{0t}, \sigma_{0t}^2) & \text{if } J_t = 0 \end{cases} \tag{5}$$

where μ_t is the conditional mean and σ_t^2 is the conditional variance, potentially different depending upon the existence of a jump. The information set consists of past returns and cross-sectional positions $\Im_{t-1} = \{y_{t-1}, y_{t-2}, ...z_{t-1,} z_{t-2}, ...\}$.

The log-likelihood function $L_2(\theta_2) = \sum_{t=1}^{T} \log f_2(y_t|J_t, \Im_{t-1}; \theta_2)$ is

$$L_2(\theta_2) = \sum_{t=1}^{T} \log \frac{1}{\sqrt{2\pi}} \left[\frac{J_t}{\sqrt{\sigma_{1t}^2}} \exp \frac{-(y_t - \mu_{1t})^2}{\sigma_{1t}^2} + \frac{1 - J_t}{\sqrt{\sigma_{0t}^2}} \exp \frac{-(y_t - \mu_{0t})^2}{\sigma_{0t}^2} \right]$$

where θ_2 includes all parameters in the conditional means and conditional variances.

The implication of (5) together with (1) is that the conditional marginal density function of returns is a mixture of normal density functions where the weights in the mixture are given by the probability of jumping

$$g(y_t|\Im_{t-1}) = p_t \times N(\mu_{1t}, \sigma_{1t}^2) + (1 - p_t) \times N(\mu_{0t}, \sigma_{0t}^2) \tag{6}$$

Our interest is to forecast returns. From (6), it is easy to see that expected returns are nonlinear functions of the information set, even in the simple case where μ_{1t} and μ_{0t} are linear. The expected return is a function of p_t , which we call time-varying cross-sectional risk (VCR). It is time-varying because it depends on the information set, it is cross-sectional because it depends on the positioning of the asset return in relation to the other firms in the market, and it is risk because it assesses the possibility of being a winner or a looser within the full collection of assets. The one-step ahead forecast of the return according to (6) is

$$E(y_{t+1}|\Im_t) = p_{t+1} \times \mu_{1t+1} + (1 - p_{t+1}) \times \mu_{0t+1} \tag{7}$$

2.3 Estimation Results

We collect the weekly returns, from January 1, 1990 to August 31, 2000, for all the firms in the SP500 index that have survived for the last ten years (343 firms). The total number of observations is 599. In Table 1, we summarize the unconditional moments -mean, standard deviation, coefficient of skewness, and coefficient of kurtosis- for the 343 firms. The frequency distribution of the unconditional mean seems to be bimodal with two well defined groups of firms, a cluster with a negative mean return of approximately -0.25%, and another with a positive mean return of 0.25%. For the unconditional standard deviation, the median value is 4.37%. The median coefficient of skewness is 0.01, with most of the firms exhibiting moderate to low asymmetry. All the firms have a large coefficient of kurtosis with a median value is 5.23. We calculate the Box-Pierce statistic to test for up to fourth order autocorrelation in returns and we find mild autocorrelation for about one-third of the firms. However, the Box-Pierce test for up to fourth order autocorrelation in squared returns indicates strong dependence in second moments for all the firms in the SP500 index.

[Table 1 about here]

2.3.1 The duration model

For 343 firms, we fit a conditional duration model as in (2) and (3). The information set consists of past durations, past returns and past cross-sectional positions : $\{D_{N(t)-j}, y_{t-j}, z_{t-j}\}$ for $j = 1, 2, ...$ The duration time series for every firm is characterized by clustering - long (short) durations are followed by long (short) durations- and, consequently the specification of a ACH model may be warranted. We maximize the log-likelihood function (4) with respect to the parameter vector $\theta_1 = (\alpha', \beta', \delta')'$. We estimate different lag structures (linear and nonlinear) of the information set and, based on standard model selection criteria (t-statistics and log-likelihood ratio tests), we obtain the following final specification

$$\begin{aligned} p_t &= [\Psi_{N(t-1)} + \delta' X_{t-1}]^{-1} \\ \Psi_{N(t)} &= \alpha D_{N(t)-1} + \beta \Psi_{N(t)-1} \\ \delta' X_{t-1} &= \delta_1 + \delta_2 y_{t-1} \mathbf{1}(z_{t-1} \leq 0.5) + \delta_3 y_{t-1} \mathbf{1}(z_{t-1} > 0.5) + \delta_4 z_{t-1} \end{aligned}$$

The conditional duration model is an ACH(1,1) with persistence $\alpha + \beta$. There is a nonlinear effect of the predetermined variables on duration. The effect of past returns on duration depends on whether the cross-sectional position of the asset is above or below the median. In Table 2, we report the cross-sectional frequency distributions of the estimates $\hat{\theta}_1 = (\hat{\alpha}', \hat{\beta}', \hat{\delta}')'$ for all the 343 firms. All the parameters are statistically significant at the customary 5% level with the exception of δ_4, for which do not report its frequency distribution.

[Table 2 about here]

For $\hat{\alpha}$, the median is 0.36 with 90% of the firms having an $\hat{\alpha}$ below 0.48. For $\hat{\beta}$, its frequency distribution is highly skewed to the right with a median of 0.06 and with 90% of the firms having a $\hat{\beta}$ below 0.25. The median persistence is 0.45 and for 90% of the firms, the persistence is below 0.63. The estimates $\hat{\delta}_2$ and $\hat{\delta}_3$ have opposite signs, the former is positive and the latter is negative with $\hat{\delta}_2$ being larger in magnitude than $\hat{\delta}_3$. The effect of $\hat{\delta}_2$ and $\hat{\delta}_3$ in expected duration depends on the interaction between the cross-sectional position and the sign of the return. There are four possible scenarios. For instance, when the past asset return is positive and below (above) the median market return, its expected duration is longer (shorter) and the probability of a sharp jump is smaller (larger), other things equal. On the contrary, when the past asset return is negative and below (above) the median market return, its expected duration is shorter (longer) and the probability of a sharp jump is larger (smaller), other things equal. Both $\hat{\delta}_2$ and $\hat{\delta}_3$ have a very skewed cross-sectional frequency distributions. For $\hat{\delta}_2$, the median value is 0.15 with 90% of the firms having a $\hat{\delta}_2$ below 0.55.

For $\hat{\delta}_3$, the median value is -0.11 with 90% of the firms having a $\hat{\delta}_3$ below -0.45. Roughly speaking, for a representative firm with median parameter estimates, the expected duration is between 4 and 5 weeks, and since $E(p_t) \geq [E(\Psi_{N(t-1)} + \delta^{'} X_{t-1})]^{-1}$, a lower bound for the expected probability of a sharp jump is between 20 and 25 %.

2.3.2 The nonlinear model for conditional returns

We proceed to estimate (5). Since this model is already nonlinear, we restrict the specification of the conditional mean and conditional variance in each state ($J_t = 1$ or $J_t = 0$) to parsimonious linear functions of the information set. The preferred specification of the conditional first two moments is

$$\begin{aligned} f_2(y_t|J_t, \Im_{t-1}; \theta_2) &= \begin{cases} N(\mu_{1t}, \sigma^2_{1t}) & \text{if } J_t = 1 \\ N(\mu_{0t}, \sigma^2_{0t}) & \text{if } J_t = 0 \end{cases} \qquad (8) \\ \mu_{1t} &\equiv E(y_t|\Im_{t-1}, J_t = 1) = \nu_1 + \gamma_1 y_{t-1} + \eta_1 z_{t-1} \\ \mu_{0t} &\equiv E(y_t|\Im_{t-1}, J_t = 0) = \nu_0 + \gamma_0 y_{t-1} + \eta_0 z_{t-1} \\ \sigma^2_{1t} &= \sigma^2_{0t} = \sigma^2_t = E(\epsilon^2_t|\Im_{t-1}, J_t) = \omega + \rho\epsilon^2_{t-1} + \tau\sigma^2_{t-1} \end{aligned}$$

where $\epsilon_{t-1} = (y_{t-1} - \mu_{1,t-1})J_{t-1} + (y_{t-1} - \mu_{0,t-1})(1 - J_{t-1})$. We arrive to this specification by implementing a battery of likelihood ratio tests. The descriptive statistics mentioned at the beginning of this section show that the returns are leptokurtic. If the returns are coming from a mixture of distributions, we can explain the unconditional leptokurtosis that we find in the data. Consequently, the tests should aim to gather statistical evidence for the mixture of normals that we propose in (5). In a first instance, we focus on the possibility of different dynamics in the conditional mean in each state with equal constant variances. The first hypothesis of interest is

$$H^1_0 : \nu_1 = v_0, \gamma_1 = \gamma_0, \eta_1 = \eta_0$$

where both the restricted and the unrestricted model have $\sigma^2_{1t} = \sigma^2_{0t} = \sigma^2$. This null is rejected very strongly for all the firms in the SP500 index[3]. In a second instance, we relax the assumption of constant variance across states and write a second null hypothesis as

$$H^2_0 : \nu_1 = v_0, \gamma_1 = \gamma_0, \eta_1 = \eta_0$$

where both the restricted and the unrestricted model have $\sigma^2_{1t} = \sigma^2_{0t} = \sigma^2_t$ with σ^2_t specified as in (8). For all firms, we reject again very strongly this

[3] We do not report all the testing results for the 343 firms but they are available upon request.

null hypothesis and hence, we claim that there are different dynamics in the conditional mean across states as well as dynamics in the conditional variance. In the third instance, we inquire about the possibility of having different dynamics in the conditional variance. Following the rejection of H_0^1 and H_0^2, the third hypothesis that we test is equal conditional variances maintaining the nonlinearity in the conditional mean, i.e. $\mu_{1t} \neq \mu_{0t}$, as in (8),

$$H_0^3 : \sigma_{1t}^2 = \sigma_{0t}^2$$

The unrestricted model has conditional means as in (8) and conditional variances that follow a GARCH(1,1) process with different parameters depending on $J_t = 1$ or $J_t = 0$. We fail to reject the third hypothesis and we settle in a final model as in (8).

The estimation results for the 343 firms are summarized in Table 3. We report the cross-sectional frequency distributions of the parameters estimates of the conditional mean and conditional variance. The majority of the parameters are statistically significant at the 5% level.

[Table 3 about here]

When we consider asset returns for which a sharp jump has taken place, the impact of past returns, $\hat{\gamma}_1$, is predominantly negative (in 75% of the firms $\hat{\gamma}_1 < 0$). The effect of past cross-sectional positions, $\hat{\eta}_1$, is clearly negative for all the firms. These signs are expected. Consider an asset whose past return has been going down and, at the same time, has experienced a move down in its cross-sectional position. A movement down in past returns and cross-sectional positions implies that the likelihood of a sharp jump up is increased. Given that we are considering an asset for which a sharp jump is happening, we should expect that the most likely direction of the jump is up, thus increasing the present expected return. When we consider asset returns for which there is no jump, the effect of past returns, $\hat{\gamma}_0$, on expected returns could be positive or negative across the firms with a median value of 0.15. On the contrary, the marginal effect of past cross-sectional positions, $\hat{\eta}_0$, is clearly positive. This means that asset returns who move up in the cross-sectional ranking of firms, but who have not experienced a sharp jump, tend to have an increase in their expected returns, other things equal. For individual firms we observe that $|\hat{\gamma}_1| > |\hat{\gamma}_0|$ and $|\hat{\eta}_1| > \hat{\eta}_0$, which it is consistent with the notion that, most of the time, sharp jumps in cross-sectional positions must be associated with large movements in expected returns. The median value of $\hat{\gamma}_1$ is -1.11 compared to the median of $\hat{\gamma}_0$ that is 0.15; and the median of $\hat{\eta}_1$ is -0.61 compared to the median of $\hat{\eta}_0$ that is 0.37.

The second part of Table 3 describes the estimates of the parameters of the conditional variance. The model is a standard GARCH(1,1). The persistence is measured by $\rho + \tau$. The median persistence is 0.93, with a median value for $\hat{\rho}$ of 0.06 and a median value for $\hat{\tau}$ of 0.87. A leverage effect in the conditional variance does seem to be warranted, the different specifications of the conditional mean across states take care of potential asymmetries in returns. We run standard diagnostic checks such as the Box-Pierce statistics for autocorrelation in residuals, squared residuals, and standardized squared residuals and we conclude that the residuals, standardized residuals, and standardized squared residuals seem to be white noise. The specification (8) passes standard diagnostic checks for model adequacy. However, a more important aspect of the model is to assess its forecasting performance, which we analyze in the following section.

3 Out-of-sample evaluation of the VCR model

In this section we assess the performance of the proposed VCR model within the context of investment decision making. We consider two major scenarios. First, we deal with an investor whose interest is to maximize profits from trading stocks. We assume that her trading strategy -what to buy, what to sell- depends on the forecast of expected returns based on the VCR model, The superiority of the proposed specification depends on its potential ability to generate larger profits than those obtained with more standard models. In the second scenario, in addition to the return, the investor wishes to assess potential large losses by adding a Value-at-Risk evaluation of her trading strategy. In this case, the modelling of the conditional variance becomes also relevant. Both scenarios provide an out-of-sample evaluation of the VCR model.

3.1 Trading rules

We consider four trading strategies. The first one is called "VCR-position trading rule" and it is based on the one-step ahead forecast of individual asset returns based on the VCR model. We proceed as follows. For each firm in the market (343 firms), we compute the one-step ahead forecast $\hat{y}_{i,t+1}$ of the return as in (7) and predict the ordinal position of the asset in relation to the overall market, that is, $\hat{z}_{i,t+1} = M^{-1} \sum_{j=1}^{M} \mathbf{1}(\hat{y}_{it+1} \geq \hat{y}_{j,t+1})$. The sequence of one-step ahead forecasts is obtained with a rolling sample. For a sample size of T and with the first R observations, we estimate the parameters of the model $\hat{\theta}_R$ and compute the one-step ahead forecast $\hat{y}_{i,R+1}(\hat{\theta}_R)$. Next, using observations 2 to $R+1$, we estimate the model again to obtain $\hat{\theta}_{R+1}$ and calculate the one-step ahead forecast $\hat{y}_{i,R+2}(\hat{\theta}_{R+1})$. We keep rolling the sample one observation at a time until we reach T, to obtain $\hat{\theta}_T$ and the last one-step forecast $\hat{y}_{i,T+1}(\hat{\theta}_T)$. In the first period of the forecasting interval (R,T), the investor observes the predicted ranking and buys the top five performing assets. In every subsequent period, the investor revises her portfolio, selling the assets that fall out of the top performers and buying the ones that raise to the top, and she computes the one-period portfolio return $\pi_t = K^{-1}\sum_{j=1}^{K} y_{jt}$ where $K=5$ is the number of assets in the top performing portfolio. Every asset in the portfolio is weighted equally.

The second trading rule is called "Position trading rule" and, though it takes into account the cross-sectional position of an asset, ignores the non-linearity of the model (mixture of normal densities) for expected returns. The one-step ahead forecast for every asset in the market is obtained from a linear specification of the conditional mean where the regressors are past returns and past cross-sectional positions. As in the previous rule, the ordinal position is predicted and a rolling sample scheme is used to obtain the sequence of one-step ahead forecasts. The investor follows the same strategy as before buying the top five performing assets and revising her portfolio in every period.

The third trading rule is a buy-and-hold strategy of the market portfolio. At the beginning of the forecasting interval, the investor buys the SP500 index and holds it until the end of the interval. At any given t, the one-period portfolio return is $\pi_t = y_{mt}$ where y_{mt} is the return to the SP500 index.

The four trading rule is driven by the market efficiency hypothesis. We call this rule the "Random walk trading rule". If stock prices follow a random walk, the best predictor of price is the previous price, and the best forecast for the return of any given asset is zero. Hence $\pi_t = 0$ for any t and any asset.

In summary, three out of the four trading rules aim to assess the predictability of stock returns: the "VCR-position trading rule" claims that stock returns are non-linearly predictable, the "Position trading rule" claims that stock returns are linearly predictable, and the "Random walk trading rule" claims that stock returns are linearly non-predictable. The "Buy-and-hold the market trading rule" claims that actively managed portfolios have no advantage over passively index investing.

3.2 Loss functions: mean trading return and VaR

The simplest evaluation criterium is to compute the return of each trading strategy over the forecasting interval (R, T). Suppose that there are P periods in this interval. For every trading rule we evaluate the mean trading return $MTR = P^{-1} \sum_{t=R}^{T} \pi_{t+1}$. The rule that provides the largest MTR would be a preferred trading strategy. However, with each trading rule we choose different portfolios that may have different levels of risk. The investor while pursuing a high MTR may also wish to control for catastrophic events maintaining a minimum amount of capital to cushion against excessive losses. Consequently, each trading rule would be evaluated according to their ability to allocate the optimal amount of capital for unlikely events, rendering a Value-at-Risk evaluation criterium necessary.

Consider a portfolio of assets whose return is given by π_t. We are interested in $VaR_{t+1}^{\pi}(\alpha)$, the one-step ahead Value-at-Risk forecast of π_t at a given nominal tail coverage probability α. This is defined as the conditional quantile

$$\Pr(\pi_{t+1} \leq VaR_{t+1}^{\pi}(\alpha)|\Im_t) = \alpha$$

If the density function of π_{t+1} belongs to the location-scale family (eg. Lehmann 1983, p. 20), $VaR_{t+1}^{\pi}(\alpha)$ can be estimated as

$$VaR_{t+1}^{\pi}(\alpha) = \mu_{t+1}(\hat{\theta}_t) + \Phi_{t+1}^{-1}(\alpha)\sigma_{t+1}(\hat{\theta}_t)$$

where $\mu_{t+1}(\hat{\theta}_t)$ is the conditional mean forecast of the return, $\sigma_{t+1}(\hat{\theta}_t)$ is the conditional standard deviation forecast, $\Phi_{t+1}(.)$ is the conditional cumulative distribution function of the standardized portfolio return, and $\hat{\theta}_t$ is the parameter vector estimated with information up to time t.

We evaluate the trading rules according to three VaR based loss functions. The first loss function aims to minimize the required capital to protect the investor against a large negative return, the second loss function assesses which trading rule provides the correct predicted tail coverage probability, and the third loss function is based on quantile estimation and it evaluates which trading rule provides the best estimate of the VaR.

The first loss function V_1, suggested by Bao *et al.* (2002) sets the minimum required capital $MRC_{t+1}(\alpha) = VaR_t^{\pi}(\alpha)$. Over the forecasting period, the trading rule that provides the lowest amount of capital to put aside to protect the investor from a large loss will be preferred

$$V_1 \equiv P^{-1} \sum_{t=R}^{T} MRC_{t+1}(\alpha) = P^{-1} \sum_{t=R}^{T} VaR_t^{\pi}(\alpha)$$

The second loss function V_2 aims to choose the trading rule that minimizes the difference between the nominal and the empirical lower tail probability. It is proposed by Lee and Saltoglou (2001) as an out-of-sample evaluation criterium based on the likelihood ratio statistic of Christofferson (1998). Over the forecasting interval (R, T), consider the following counts $n_0 = \sum_{t=R}^{T} \mathbf{1}(\pi_{t+1} > VaR_t^{\pi}(\alpha))$ and $n_1 = \sum_{t=R}^{T} \mathbf{1}(\pi_{t+1} < VaR_t^{\pi}(\alpha))$. If the VaR has been correctly estimated, n_0 must be $P \times (1-\alpha)$ and n_1 equals to $P \times \alpha$. In fact, the log-likelihood function of a sample $\{\mathbf{1}(\pi_{t+1} < VaR_t^{\pi}(\alpha)\}_{t=R}^{T}$ is $L(\alpha) = (1-\alpha)^{n_0} \alpha^{n_1}$ and the maximum likelihood estimator of α is $\hat{\alpha} = n_1/P$. If we were to test $H_0 : \alpha =$nominal tail probability, the likelihood ratio test $2(\log L(\hat{\alpha}) - \log L(\alpha))$ would be a suitable statistic. The loss function V_2 is based in this statistic. A trading rule that minimizes V_2 will be preferred.

$$\begin{aligned} V_2 &\equiv P^{-1} 2 \left[\log L(\hat{\alpha}) - \log L(\alpha)\right] \\ &= P^{-1} 2 \sum_{t=R}^{T} \left[\mathbf{1}(\pi_{t+1} < VaR_t^{\pi}(\alpha)) \log \frac{\hat{\alpha}}{\alpha} + \mathbf{1}(\pi_{t+1} > VaR_t^{\pi}(\alpha)) \log \frac{1-\hat{\alpha}}{1-\alpha} \right] \end{aligned}$$

The third loss function V_3 chooses the trading rule that minimizes the objective function used in quantile estimation (Koenker and Bassett, 1978)

$$\begin{aligned} V_3 &\equiv P^{-1} \sum_{t=R}^{T} |\pi_{t+1} - VaR_t^{\pi}(\alpha)| \times \\ &\quad \times \left[(1-\alpha) \times \mathbf{1}(\pi_{t+1} < VaR_t^{\pi}(\alpha)) + \alpha \times \mathbf{1}(\pi_{t+1} \geq VaR_t^{\pi}(\alpha))\right] \end{aligned}$$

The trading rule that provides the smallest V_3 is preferred because it indicates a better goodness of fit.

3.3 Comparison of trading rules

The question of interest is, out of the four proposed trading rules, which one is the best? Each rule produces different forecasts that are evaluated according to the four loss functions introduced in the previous section: $-MTR, V_1, V_2$, and V_3 [4]. The best trading rule is the one that provides the

[4] We write a negative sign in the mean trading return ($-MRT$) so as to minimize this function as well as those based on VaR calculations.

minimum loss. However, for every loss function, how can we tell when the difference among the losses produced by each trading rule is statistically significant? Furthermore, is it a pairwise comparison among the four trading rules sufficient? Note that all trading rules are based on the same data, and that their forecasts are not independent. We need a statistical procedure to assess whether the difference among the losses is significant while, at the same time, taking into account any forecast dependence across trading rules and controlling for potential biases due to data snooping. This procedure is the "reality check" proposed by White (2000). Suppose that we choose one of the trading rules as a benchmark. We aim to compare the loss of the remaining trading rules to that of the benchmark. We formulate a null hypothesis where the trading rule with the smallest loss is not any worse than the benchmark rule. If we reject the null hypothesis, there is at least one trading rule that produces less loss than the benchmark. A brief sketch of the formal testing procedure follows.

Let l be the number of competing trading rules ($k = 1, \ldots, l$) to compare with the benchmark rule (indexed as $k = 0$). For each trading rule k, one-step predictions are to be made for P periods from R through T using a rolling sample, as explained in the previous sections. Consider a generic loss function $L(Y, \theta)$ where Y typically will consist of dependent variables and predictor variables. In our case, we have four loss functions: $-MTR, V_1, V_2$, and V_3. The best trading rule is the one that minimizes the expected loss. We test a hypothesis about an $l \times 1$ vector of moments, $E(\mathbf{f}^*)$, where $\mathbf{f}^* \equiv \mathbf{f}(Y, \theta^*)$ is an $l \times 1$ vector with the kth element $f_k^* = L_0(Y, \theta^*) - L_k(Y, \theta^*)$, for $\theta^* = \text{plim}\hat{\theta}_T$ and where $L_0(.)$. is the loss under the benchmark rule and $L_k(.)$ is the loss provided by the k trading rule. A test for a hypothesis on $E(\mathbf{f}^*)$ may be based on the $l \times 1$ statistic $\overline{\mathbf{f}} \equiv P^{-1} \sum_{t=R}^{T} \hat{\mathbf{f}}_{t+1}$,where $\hat{\mathbf{f}}_{t+1} \equiv \mathbf{f}(Y_{t+1}, \hat{\theta}_t)$.

Our interest is to compare all the trading rules with a benchmark. An appropriate null hypothesis is that all the trading rules are no better than a benchmark, i.e., $H_0 : \max_{1 \leq k \leq l} E(f_k^*) \leq 0$. This is a multiple hypothesis, the intersection of the one-sided individual hypotheses $E(f_k^*) \leq 0,\ k = 1, \ldots, l$. The alternative is that H_0 is false, that is, the best trading rule is superior to the benchmark. If the null hypothesis is rejected, there must be at least one trading rule for which $E(f_k^*)$ is positive. Suppose that $\sqrt{P}(\overline{\mathbf{f}} - E(\mathbf{f}^*)) \xrightarrow{d} N(0, \Omega)$ as $P(T) \to \infty$ when $T \to \infty$, for Ω positive semi-definite. White's (2000) test statistic for H_0 is formed as $\bar{V} \equiv \max_{1 \leq k \leq l} \sqrt{P} \bar{f}_k$,which converges in distribution to $\max_{1 \leq k \leq l} G_k$ under H_0, where the limit random vector $G = (G_1, \ldots, G_l)'$ is $N(0, \Omega)$. However, as the null limiting distribution of $\max_{1 \leq k \leq l} G_k$ is unknown, White (2000, Theorem 2.3) shows that the distribution of $\sqrt{P}(\bar{\mathbf{f}}^* - \bar{\mathbf{f}})$ converges to that of $\sqrt{P}(\bar{\mathbf{f}} - E(\mathbf{f}^*))$, where $\bar{\mathbf{f}}^*$ is obtained from the stationary bootstrap of Politis and Romano (1994). By the continuous mapping theorem this result extends to the maximal element of the vector $\sqrt{P}(\bar{\mathbf{f}}^* - \bar{\mathbf{f}})$ so that the empirical distribution of

$$\bar{V}^* = \max_{1\leq k\leq l} \sqrt{P}(\bar{f}_k^* - \bar{f}_k), \tag{9}$$

is used to compute the p-value of $\bar{V}$ (White, 2000, Corollary 2.4). This p-value is called the "Reality Check p-value".

3.4 Evaluation of trading rules

The out-of-sample performance of the aforementioned trading rules is provided in Table 4. In the upper panel, the trading rules are evaluated according to the MTR function, and in the lower two panels according to the VaR loss functions. In both cases, the in-sample horizon is the first 300 periods, $R = 300$, and the out-of-sample horizon is 299 periods, $P = 299$. The stationary bootstrap is implemented with 1000 bootstrap resamples and smoothing parameter $q = 0.25$[5]. In the first column of each panel, we report the benchmark trading rule to which the remaining rules will be compared.

[Table 4 about here]

In the column MTR of the upper panel, we report the value of this function for each trading strategy. The "VCR-position" rule produces a mean trading return of 0.264 that is twice as much as the next most favorable rule, which is the "position" trading rule. This one and the "buy-and-hold the market" rule produce similar results. The "random walk" rule is the least favorable. The statistical difference among the rules is assessed with the White procedure. In the column $White$, we report the reality check p-values. The trading rules "position", "buy-and-hold the market", and "random walk" are clearly dominated with p-values less than 5%, while the "VCR-position" rule is the dominant rule with a p-value of 1.0. We also calculate the Diebold-Mariano-West test statistic that corresponds to a pairwise comparison of the benchmark rule with the best of the alternative rules. This statistic points in the same direction as that of White's: the "VCR-position" is the preferred trading rule when the investor aims to maximize returns.

[5] We experiment with different smoothing parameters $q = 0.50$ and $q = 0.75$ and we obtain similar results.

In the lower two panels of Table 4, we report the out-of-sample performance of the trading rules evaluated according to the loss functions V_1, V_2, and V_3, for two tail nominal probabilities $\alpha = 1\%$ and $\alpha = 5\%$[6]. With respect to V_1, the "position" trading rule seems to dominate statistically the remaining two rules providing the least amount of capital. However, when we consider V_2, the same rule performs very poorly because it estimates a tail probability of 7.2% for a nominal of 1%, and 14% for a nominal of 5%. Thus, a minimum amount of capital comes to the expense of a high failure rate. The "VCR-position" rule delivers the best tail coverage, estimating a tail probability of 1.1% for a nominal of 1% and 4.3% for a nominal of 5%. The White p-value confirms that this is the dominant rule. With respect to the loss function V_3, the three trading rules produce similar losses, however the differences are statistically significant and the White p-value of 1.0 asserts that, once more, the "VCR-position" trading rule dominates the alternative strategies.

4 Conclusions

From an empirical perspective, we have extended the notion of risk. Classical asset pricing theories explain risk as the covariance of individual returns with a set of factors. We have borrowed from the classics the abstract notion of asset interdependence as well as the idea of the market as the main factor that drives returns, but we have chosen a different route on materializing the abstract. Our notion of the market is the cross-sectional distribution of returns at a given moment on time, and our notion of interdependence is the cross-sectional position (ranking) of a given asset return within the full set of assets that constitute the market. Our task has been to investigate the implications of these choices for asset returns.

We have modelled the *joint* dynamics of the cross-sectional position and the asset return by analyzing (1) the *marginal* probability distribution of a sharp jump in the cross-sectional position, and (2) the probability distribution of the asset return *conditional* on a jump. The former is conducted within the context of a duration model, and the latter assumes that there are different dynamics in returns depending upon whether or not a jump has occurred. We have estimated and tested the proposed models with weekly returns of those SP500 corporations that have survived in the index from January 1, 1990 to August 31, 2000. The estimation results for the 343 firms

[6] We have removed from this comparison the "random walk" trading rule because there is indeterminacy on the assets that should form the portfolio.

are diverse but, broadly speaking, we have found that the expected probability of jumping increases when the firm's cross-sectional position is either at the very top or the very bottom of the cross-sectional distribution of returns, hence extreme positions tend to be shorter lived than intermediate ones. For a representative firm such as one with median parameter estimates, we have calculated that the expected duration is between 4 and 5 weeks implying a minimum expected probability of a sharp jump of 20-25%. Furthermore, we found that the expected return is a function of past cross-sectional positions and that there are different dynamics when the return is either at extreme positions (top or bottom of the cross-sectional distribution) or at intermediate positions.

From an investor's point of view, the most relevant question is how useful is this model. We have judged the adequacy of our model in two dimensions: profitability and risk monitoring. Different trading rules are compared within the statistical framework of White's reality check. A trading rule based on the one-step ahead forecast of our model dominates alternative rules based on standard models. It provides superior mean trading returns and at the same time, estimates correctly the Value-at-Risk.

We summarize this research by underlining two main contributions. First, the conditional probability of jumping cross-sectional positions is forecastable. This probability is named **V***arying* ***C****ross sectional* ***R****isk* because provides an assessment of the chances of being a winner or a looser within the available set of assets. Secondly, the marginal probability distribution of returns is a mixture of distributions that can explain the unconditional leptokurtosis found in financial asset returns, even in cases where conditional heteroscedasticity is not present. Based on these two contributions, we can predict the one-step ahead ranking of asset returns.

References

Bao, Y., T.-H. Lee, and B. Saltoglu (2002), "Conditional Value-at-Risk Models", University of California, Riverside.

Christoffersen, P.F. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841-862.

Diebold, F.X. and R.S. Mariano (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 253-263.

Engle, R.F. and J.R. Russell (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica,* 66, 1127-1162.

Fama, E. and K.R. French (1996), "The CAPM is Wanted, Dead or Alive," *Journal of Finance*, 51(5) 1947-58.

Fama, E. and K.R. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33(1) 3-56.

Granger, C. W J (2002), "Some comments on risk", *Journal of Applied Econometrics*, 17(5) 447-456.

Hamilton, J.D, and O. Jordà (2001), "A Model for the Federal Funds Rate Target," forthcoming, *Journal of Political Economy.*

Jegadeesh, Narasimhan, and Sheridan Titman, (2001), "Profitability of momentum strategies: an evaluation of alternative explanations", *Journal of Finance,* 56, 699-720.

Jegadeesh, Narasimhan, and Sheridan Titman, (2002), "Cross-sectional and time-series determinants of momentum returns", *Review of Financial Studies,* 15, 143-157.

Jagannathan, R and Z Wang (1996), " The Conditional CAPM and The Cross-section of Expected Returns", *Journal of Finance,* 51, 3-54.

Jorion, P. (2000) "Value at Risk: The New Benchmark for Managing Financial Risk" McGraw-Hill.

Koenker, R. and G. Bassett (1978), "Regression Quantiles", *Econometrica*, 46(1), 33-50.

Lee, T.-H. and B. Saltoglu (2001), "Evaluating Predictive Performance of Value-at-Risk Models in Emerging Markets: A Reality Check", University of California, Riverside.

McCracken, M.W. (1999), "Robust out-of-sample Inference", *Journal of Econometrics,* 99, 195-223.

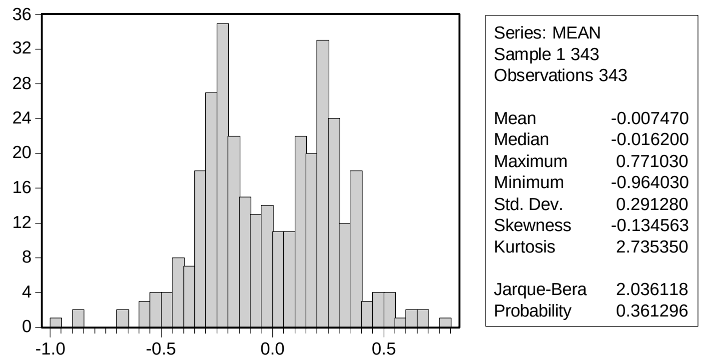
Politis, D.N. and J.P. Romano (1994), "The Stationary Bootstrap", *Journal of American Statistical Association*, 89, 1303-1313.

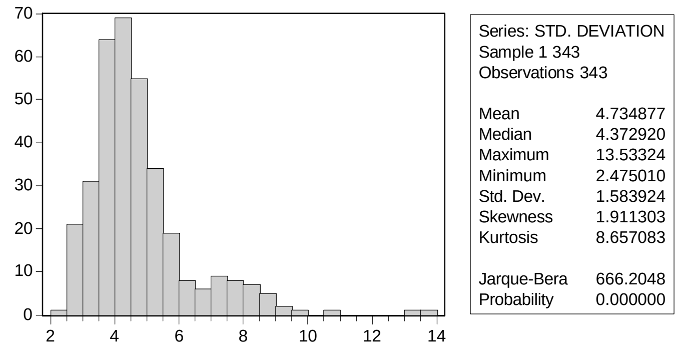
West, K.D. (1996), "Asymptotic Inference about Predictive Ability", *Econometrica*, 64, 1067-1084.

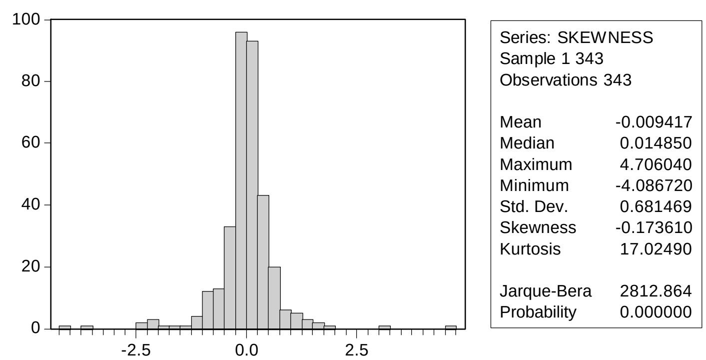
White, H. (2000), "A Reality Check for Data Snooping", *Econometrica*, 68(5), 1097-1126.

Table 1
Descriptive Statistics of Weekly Returns of the SP500 firms
January 1, 1990-August 31, 2000

Cross-sectional frequency distribution of unconditional moments







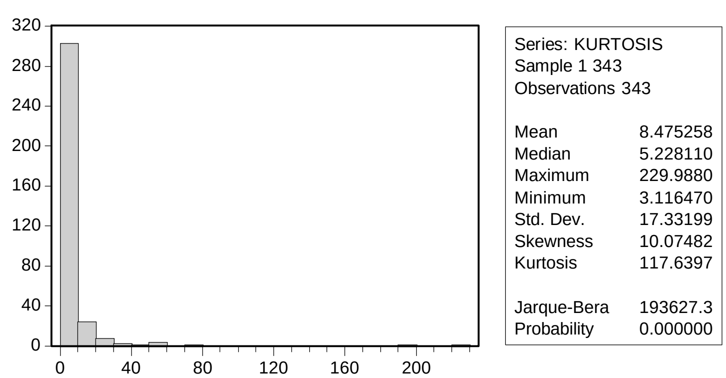


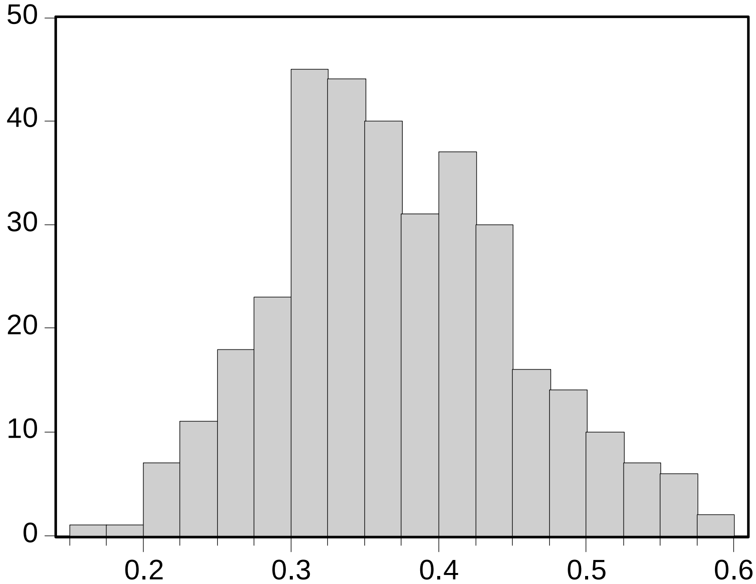
Table 2

Cross-sectional frequency distribution of the estimates of the duration model

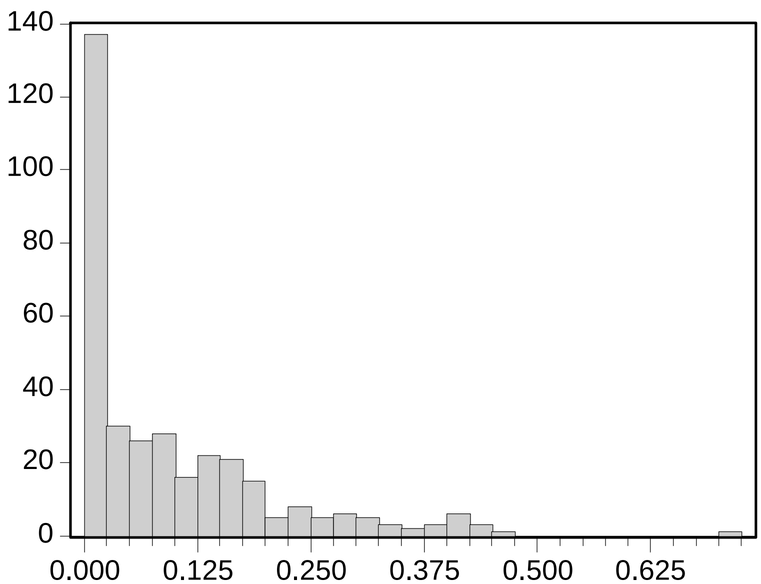
$$p_t = \Pr(J_t = 1 \mid \Im_{t-1}) = [y_{N(t-1)} + d' X_{t-1}]^{-1}$$

$$y_{N(t)} = a\ D_{N(t)-1} + b y_{N(t)-1}$$

$$d' X_{t-1} = d_1 + d_2 y_{t-1} 1(z_{t-1} \leq 0.5) + d_3 y_{t-1} 1(z_{t-1} > 0.5) + d_4 z_{t-1}$$



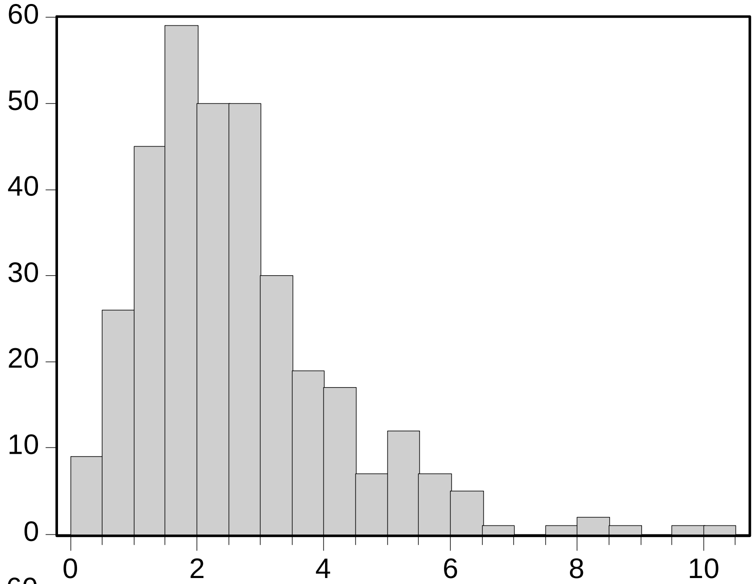
Series: ALPHA	
Sample 1 343	
Observations 343	
Mean	0.371292
Median	0.364000
90% percentile	0.485000
Maximum	0.598000
Minimum	0.167000
Std. Dev.	0.081061
Skewness	0.309177
Kurtosis	2.814036
Jarque-Bera	5.958832
Probability	0.050823



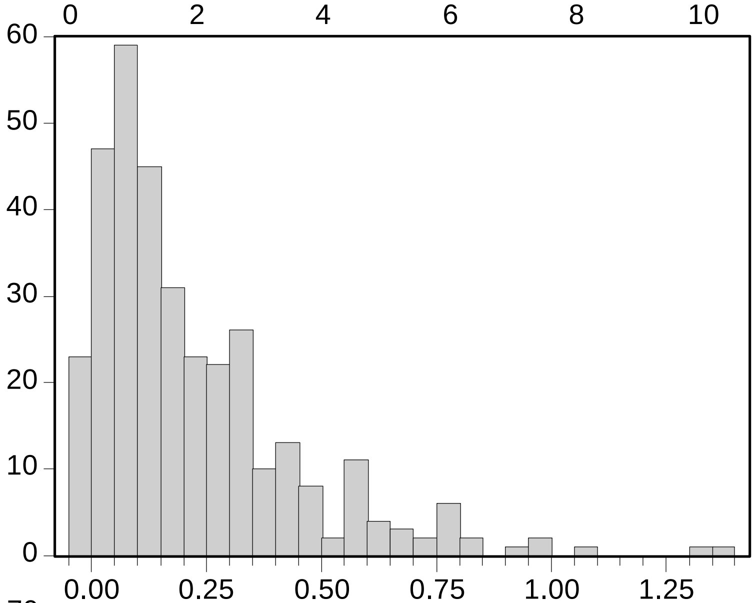
Series: BETA	
Sample 1 343	
Observations 343	
Mean	0.098172
Median	0.056000
90% percentile	0.252000
Maximum	0.724000
Minimum	0.010000
Std. Dev.	0.110893
Skewness	1.780148
Kurtosis	6.773379
Jarque-Bera	384.6472
Probability	0.000000

Table 2 (cont.)

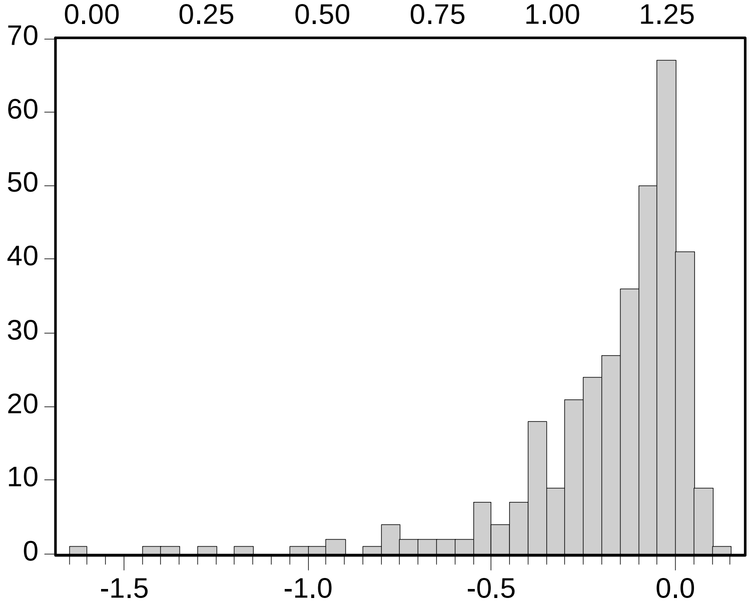
Cross-sectional frequency distribution of the estimates of the duration model



Series: DELTA1	
Sample 1 343	
Observations 343	
Mean	2.621921
Median	2.325000
90%percentile	4.757000
Maximum	10.23000
Minimum	0.008000
Std. Dev.	1.580073
Skewness	1.480903
Kurtosis	6.378431
Jarque-Bera	288.4929
Probability	0.000000



Series: DELTA2	
Sample 1 343	
Observations 343	
Mean	0.218213
Median	0.146000
90% percentile	0.551000
Maximum	1.369000
Minimum	-0.027000
Std. Dev.	0.223472
Skewness	1.804511
Kurtosis	7.277704
Jarque-Bera	447.6692
Probability	0.000000



Series: DELTA3	
Sample 1 343	
Observations 343	
Mean	-0.181584
Median	-0.105000
90% percentile	-0.446000
Maximum	0.135000
Minimum	-1.617000
Std. Dev.	0.244567
Skewness	-2.509015
Kurtosis	11.44026
Jarque-Bera	1377.982
Probability	0.000000

Table 3
Cross-sectional frequency distribution of the estimates of the nonlinear model for expected returns

$$f_2(y_t \mid J_t, \Im_{t-1}; q_2) = \begin{cases} N(m_{1t}, s_{1t}{}^2) & if \quad J_t = 1 \\ N(m_{ot}, s_{ot}{}^2) & if \quad J_t = 0 \end{cases}$$

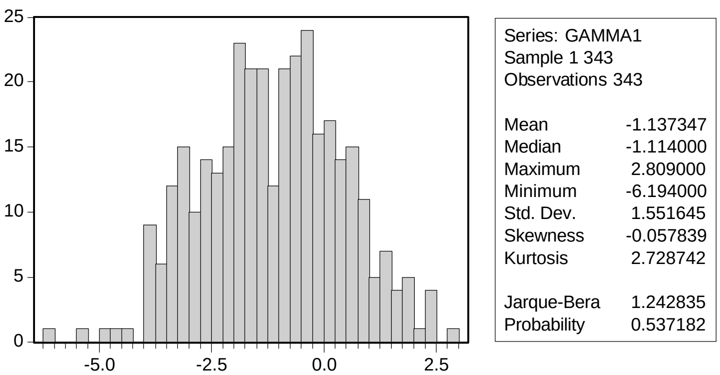
$$m_{1t} = n_1 + g_1 y_{t-1} + h_1 z_{t-1}$$

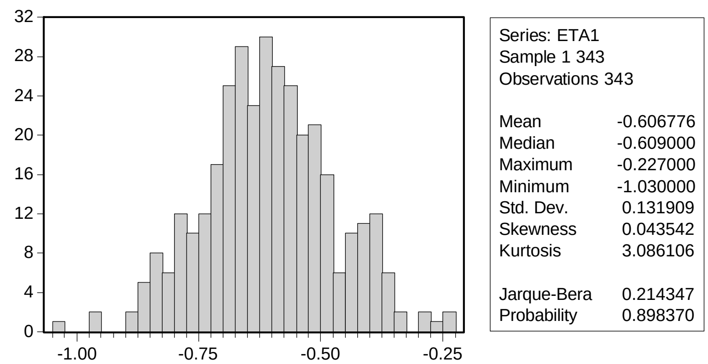
$$m_{0t} = n_0 + g_0 y_{t-1} + h_0 z_{t-1}$$

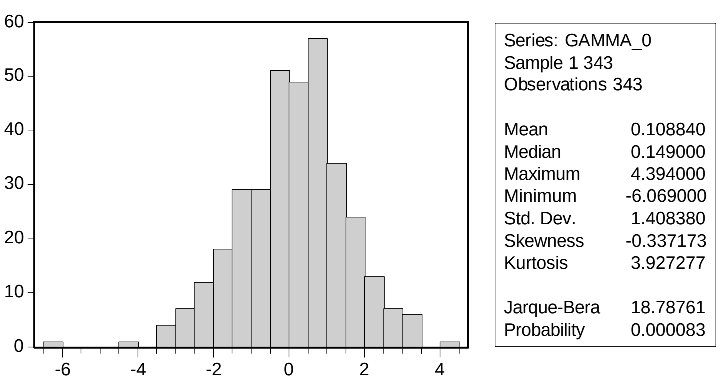
$$s^2{}_t = w + re^2{}_{t-1} + ts^2{}_{t-1}$$

$$with \quad e_{t-1} = (y_{t-1} - m_{1,t-1})J_{t-1} + (y_{t-1} - m_{0,t-1})(1 - J_{t-1})$$

Conditional mean parameter estimates







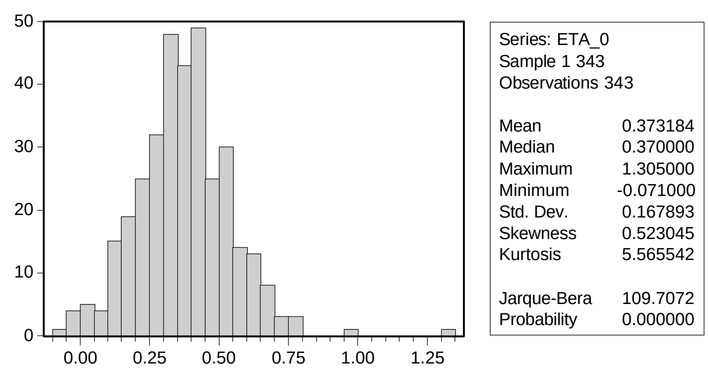
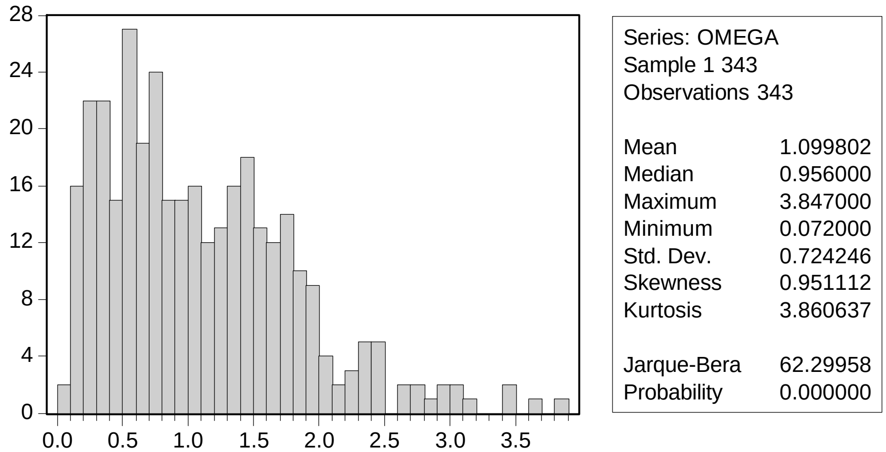
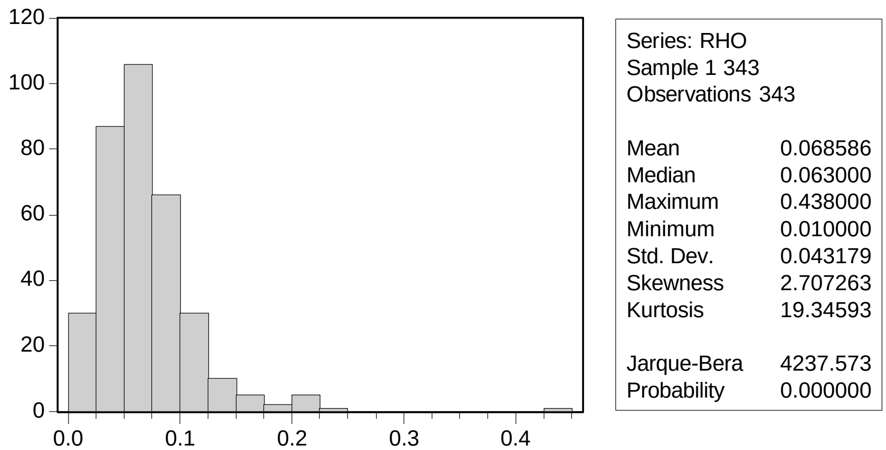


Table 3 (cont.)

Conditional variance parameter estimates





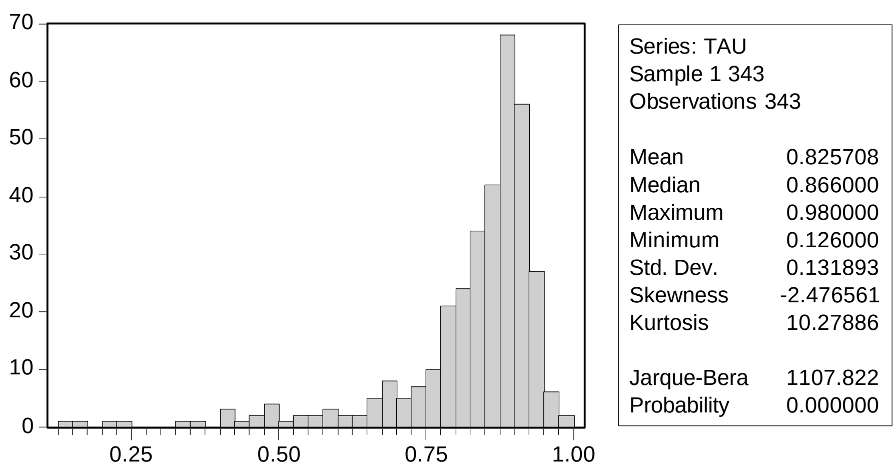


Table 4

Out-of-sample performance of the trading rules

Mean Trading Return

Benchmark trading rule	MTR	White p-value	DMW p-value
VCR-Position	0.264	1.000	0.984
Position	0.131	0.021	0.002
Buy-and-hold market	0.115	0.019	0.001
Random walk	0.000	0.000	0.000

VaR based loss functions

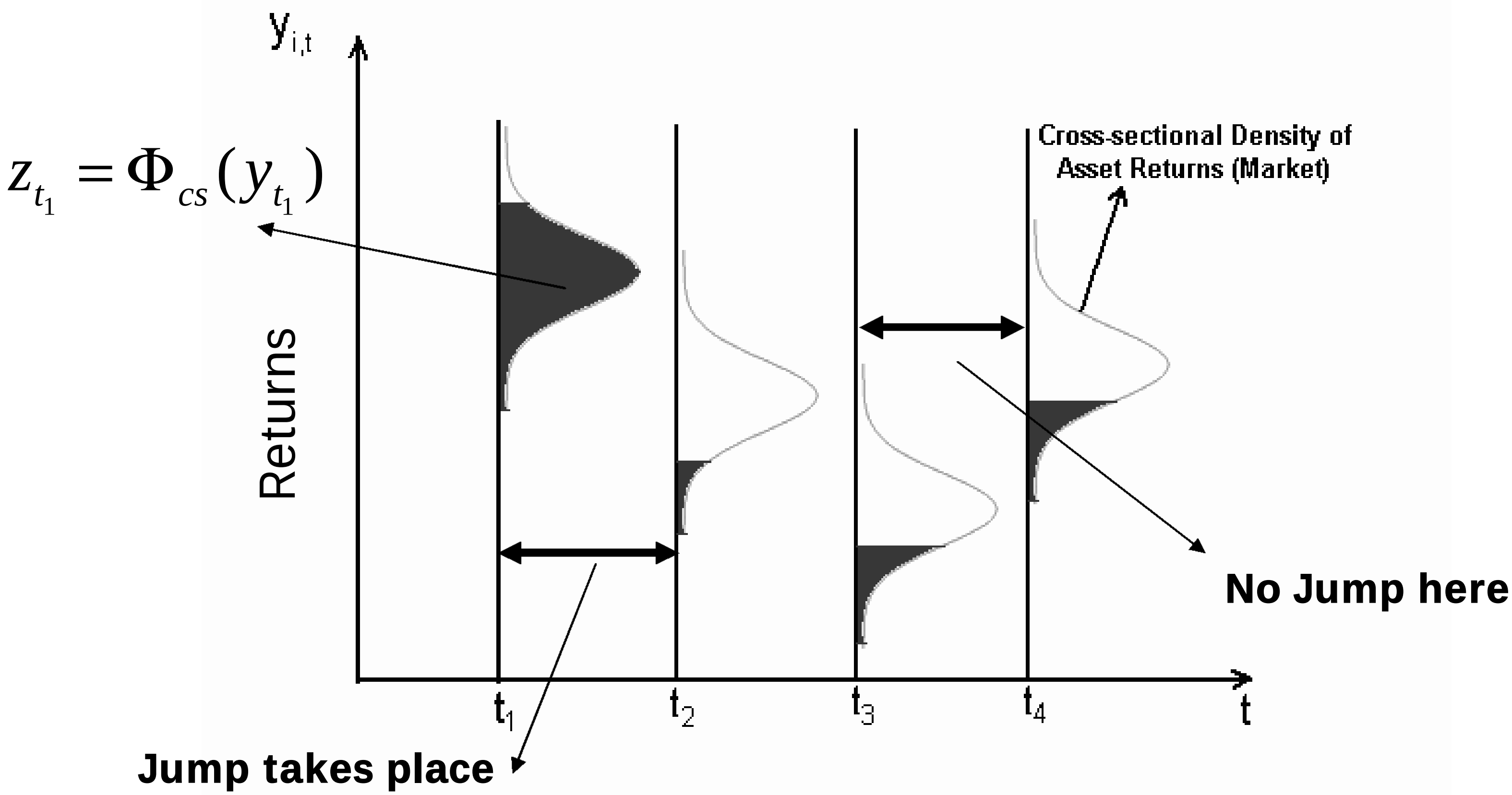
$a = 0.01$

Benchmark trading rule	V_1	White p-value	V_2	$\hat{a}$	White p-value	V_3	White p-value
VCR-Position	1.672	0.000	0.003	0.011	1.000	0.020	1.000
Position	0.980	1.000	0.097	0.072	0.000	0.038	0.010
Buy-and-hold market	2.351	0.000	0.040	0.029	0.000	0.040	0.011

$a = 0.05$

Benchmark trading rule	V_1	White p-value	V_2	$\hat{a}$	White p-value	V_3	White p-value
VCR-Position	1.076	0.000	0.011	0.043	1.000	0.075	1.000
Position	0.565	1.000	0.171	0.140	0.000	0.101	0.034
Buy-and-hold market	1.625	0.000	0.055	0.081	0.000	0.141	0.000

Figure 1
Stylized description of the modelling problem



$y_{t_1} > y_{t_2}$	$y_{t_2} > y_{t_3}$	$y_{t_3} < y_{t_4}$
$z_{t_1} > z_{t_2}$	$z_{t_2} < z_{t_3}$	$z_{t_3} = z_{t_4}$