"A Threshold Model for Italian Stock Market Volatility. Potential of These Models: Comparing Forecasting Performances and Evaluation of Portfolio-Risk with Other Models."

di

Asmara Jamaleh

GRETA, Venezia

"A Threshold Model for Italian Stock Market Volatility.

Potential of These Models: Comparing Forecasting Performances

and Evaluation of Portfolio-Risk with Other Models."

1 - Introduction: What’s this Work About.

This work investigates the potential of Threshold Models in interpreting some characteristics of Financial Markets through an application to Italian Stock Market Volatility.

A SETAR (Self-Exciting Threshold AutoRegressive) model is fitted to a volatility series built using daily returns of the Italian Stock Index MIB30.

The model is then used to obtain volatility forecasts from 1 to 30 step-ahead and Value-at-Risk (VaR) estimates. VaR is a methodology recently introduced by the Basle Committee for measuring the market risk of a portfolio with regard to financial institutions, especially banks.

The performances of the SETAR model are evaluated in comparison with other competitive models, such as a linear AR, a GARCH, a GARCH-L model and, for VaR purposes, also the methodology used by Riskmetrics. This enables us to assess not only if this class of models is worth further investigation since at least comparable with models considered appreciable, but also if it is capable of capturing aspects of empirical data that the other ones can’t explain.

Specific attention is given to the results obtained in days of particular market turbulence, such as the days around the slump of Tuesday October 28, 1997, due to the financial crisis of Asian markets.

The results obtained along this investigation show a generalized better performance of the SETAR model over the other ones. In fact it isn’t only capable of capturing the various characteristics of volatility, but is in particular the only one which can distinguish a persistent shock of the market from an extraordinary shock. Its superior performance is even more evident in particularly turbulent moments of the market. Both aspects ar appreciable for operational purposes too.
As regards VaR for instance, it is basic for a model to provide accurate VaR estimates, even in extreme market conditions. It’s also important to understand if a model tends to over-estimate or under-estimate VaR, since it will conduct to

higher or lower covering costs respectively, which aren' t negligible with respect to the financial management of a bank. The SETAR model, for instance, in the days of the slump provides not only more accurate VaR estimates, but doesn't either conduct, differently from the other models, to higher covering costs, costs that aren't absolutely necessary looking at the real market conditions.

From a modelling point of view the innovative procedure introduced by Tsay is employed here. He pointed out this methodology in order to make Threshold Models tractable in application, since previous techniques were too expensive.

Therefore the present investigation is also a sort of test for the real implementability of these models.

It wil be shown throughout next sections that both the effectiveness of the modelling procedure proposed by Tsay and the appreciable performance exhibited over the other models encourage further investigation of Threshold Models from a methodological and an applied point of view.

1.1 - Motivations. Economic, Financial and Econometric Foundations.

"It seems to be generally accepted that economics is non-linear" [Granger and Teräsvirta, 1993], from the specification of investment functions to the business cycle itself. Non-linearity is also present in the specific context of Financial Markets from investors' attitudes towards risk to the generating processes of financial variables such as stock returns, interest rates, exchange rates, and so on.

Therefore, modelling such variables demands us to draw on from non-linear literature, but given the extent of this last one, it is suitable before to try and isolate the reasons of these non-linear behaviours. A rather unanimous conclusion reached in the econometric literature finds out the *Volatility* phenomenon, in the specific form of Conditional Heteroskedasticity. Together with the search for an adequate non-linear model, we therefore need to search for an adequate functional specification for conditional heteroskedasticity.

One of the main results in this field is represented by GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models, but in spite of the interesting characteristics of these models, there are still unexplained and unsolved non-linearities. Research is therefore open, and it will begin by applying to those non-linear models whose characteristics seem compatible with the evidence of the analysed cases, verifying this compatibility on real data.

A class of non-linear models that has recently began arousing the attentions of theorists and applied analysts in the analysis of financial markets is that of Threshold Models. They have been introduced by Tong in 1980 [Tong and Lim, 1980], but the complexity of the modelling procedure has been a factor that has somewhat restrained their utilization, especially in application. This is why Tsay [1989] proposed a "relatively simple" procedure " as compared with that outlined by Tong and Lim", hoping it could "help exploit the potential" of these models "in application".

In the present work we will therefore investigate the interpretative potentialities of these models, building a Threshold Model for the Italian Stock Market Volatility.

Modelling volatility is important both because it sheds further light on the generating process of the returns, and it is variously assumed as a risk measure: in investment decisions, based for example on CAPM or APT approaches, in the option pricing formulas, in the problems of financial risk management (definition of hedge ratios, carried out through derivative instruments, with high leverage effect), in the evaluation of portfolio risk through the new VaR (Value-at-Risk) methodology. In a context then, such the present-day one, we can't any longer keep separate the financial from the macroeconomic frame: the barriers between the credit sector and the securities (*"valori mobiliari"*) sector of the Financial Market are progressively falling down, with an increasing presence, in the macroeconomic field, not only of the monetary and credit policy transmission effects, but also of the whole sector of the Financial Market. It becomes so more well-founded the idea of not insignificant relations of Stock Market volatility with (real and nominal) Macroeconomic variables: " changes in the level of stock market volatility can have important effects on capital investment, consumption, and other business cycle variables" [Schwert, 1989].

The Threshold Model presented here, specified through the procedure proposed by Tsay (which will result effective and encouraging), after the usual evaluation of conformability to the data, is employed to obtain 1) volatilty forecasts, and 2) VaR estimates (this last application hasn't been tried yet with Threshold Models). If forecasting is one of the main purposes of modelling, portfolio risk evaluation, representing a more operative utilization, gives us a further and more interesting measure of prediction accuracy than usual indicators: models with analogous forecasting performances could in fact reveal themselves to be not equivalent from an operative point of view, since for VaR purposes it is important not only the precision of a model, but if it tends to underestimate or overestimate too.

The performances of the Threshold Model built here are then compared with those of other models of interest in this kind of analysis: in fact it is sensible to make a careful study of newly explored models if they show themselves reasonably comparable with those already well-tested for the same purposes, and even more if they are capable of filling the gaps left by the latter ones. Among the models employed for comparisons we have: 1) two GARCH-type models, since they have been created just for conditional heteroskedasticity; 2) one linear model, as it is recommended in every non-linearity analysis [Granger and Teräsvirta, 1993], in order to judge if the performances of a non-linear model justify the greater modelling and/or computational costs compared to those of the linear one, usually simpler; 3) for the VaR section also the model utilized by RiskMetrics, that publishes forecasts for the VaR estimates.

A specific attention is devoted to the results we have obtained in the last days of October 1997, a rather critical period for the consequences even on the Italian Stock Exchange of the crisis in the Far East markets (the press has written that October 28, Tuesday, will enter the history of Italian stock exchange). In fact it is important to analyse the performances of a model in periods of particular market turbulence, since in that case it becomes even more necessary to have reliable evaluation instruments available, being the situation more unstable.

This analysis has been conducted on the Italian case since it is both of immediate interest for us, and most of the studies applied to Threshold Models examine other Countries. In particular, Italian Stock Market, given its peculiarity (not big dimensions, and continuous alternations between following other major markets and going against the general trend) needs renovated attention: both in consideration of the financial innovation process set up in these last years, and in view of the prospects of European Monetary Union. As recent facts (runnings to stock investments and excessive increases in the Italian stock exchange, especially in April '98) show, putting in evidence that stock investment could or would become a more and more alternative option to the one in public securities[1], not only for the institutional investor, but also for the private one or the small individual saver.

[1]Even for the necessity of an alignment of Italian interest rates (and, in a certain sense, of the market) with the European ones.

1.2 - Considerations on Stock Returns: towards Threshold Models.

In a rather representative article (since other previous and following authors' thesis, both for foreign markets and the Italian ones, converge on it), Hsieh [1991] deals with the subject of non-linearities in (USA) stock returns. The presence of non-linear dynamics is mainly due to the rejection of the i.i.d. (identical and independent distribution) hypothesis for the returns, which coincides with the rejection of the Random Walk hypothesis. The simplest version of a Random Walk is: $P_t = P_{t-1} + \varepsilon_t$, where P_t are the stock prices and ε_t is a disturbance term such that $\varepsilon_t \sim$ i.i.d.$(0, \sigma^2)$. Having a Random Walk structure, price changes evolve in a rather unpredictable way, and the prediction (the conditional expected value) for the price in t+1, made in t, will be the price happened in time t: in this way, forecasts for tomorrow price, based on today price, cannot be improved by using also the information incorporated in the past prices. As a result, the only significant information for market operators is that provided by the most recent price, which coincides with the definition of (informative) Efficiency[2] of the Market. Searching for plausible sources of the i.i.d. hypothesis rejection, that is 1) non-stationarity (as synonymous with structural changes), 2) presence of low complexity chaotic dynamics, 3) conditional heteroskedasticity (time-varying disturbances variance, conditionally on some information set), Hsieh (and other authors with him) finds in this last one the origin of the non-linear dynamics.

For modelling purposes it is therefore suitable to make a careful study of this aspect. But since, Hsieh concludes, models for these purposes created, fundamentally GARCH-type models, "do not fully capture the nonlinearity in stock returns", we must explore the potentialities of other non-linear models. So we will now introduce Threshold Models, given their interesting characteristics in this respect (see section 3).

[2] An important notion in a non-linearity analysis, as we shall see in the applicative part of this work (section 5).

2 - Introduction to Threshold Models.

Threshold Models [Tong and Lim, 1980; Tong, 1983 and 1995] are a class of non-linear models, for which the non-linearity of a process reduces itself to a local linear approximation across states: the total non-linear behaviour of the process is decomposed into regimes, inside which it is linear, and transitions across regimes are controlled by so-called "threshold" variables, variously definable.

The most widely investigated class, both at theoretic and applied level, is that of SETAR (Self-Exciting Threshold Autoregressive) models.

2.1 - An Interesting Threshold Model: the SETAR Model.

The univariate series y_t is a **SETAR(*l*; k, ..., k)** model, where k is repeated *l* times, if:

$$y_t = a_0^{(j)} + \Sigma_{i = 1, \dots, k}\, a_i^{(j)} y_{t-i} + \sigma^{(j)} \varepsilon_t,$$

conditional on $y_{t-d} \in R_j$, j = 1, 2, ..., *l*.

The terms "Self-Exciting" indicate that the dynamic of the process is generated by the process itself "d" times lagged (*endogenous* "threshold" dynamic).

· In this model,

(-) *l* is the number of regimes and R_j, j = 1, ..., *l*, are the regimes;

(-) k is the AR order in each regime;

(-) $r_1, \dots, r_{l-1}$ are the threshold parameters, or threshold values, or simply the thresholds ($r_0 = -\infty$, $r_l = +\infty$);

(-) d is the delay parameter or threshold lag;

(-) $\{\varepsilon_t^{(j)}\}$ is a sequence of (heterogeneous) strict white noises.

· A SETAR model is a piecewise linear model in the space of y_{t-d}, and is capable of providing accurate "local approximations" in this space. It is not, however, piecewise linear in time. Alternatively, in a certain sense, one can interpret a SETAR model as a *switching* linear regression model, with the difference that the switching mechanism is controlled by the threshold variable y_{t-d}, not by the time index t. In fact the dynamic of the model depends on the dimension of y_{t-d}, that is the value taken by y_{t-d}. This is why, at each time t, the process falls in one or another regime, according to the value taken by y_{t-d}. That is, if, at time t, $y_{t-d} \in R_j$, (i.e. $r_{j-1} < y_{t-d} \leq r_j$), j = 1, ..., *l*, the AR coefficients are $a_i^{(j)}$, with i = 0, 1, ..., k, and the disturbance variance is σ_j^2.

· There are many variants available of the SETAR model here presented: the AR orders can vary across different regimes, giving rise to a SETAR(l; k_1, ..., k_l), and so can the delay parameter; other lags can be included in the conditioning expression y_{t-d}; only the disturbance variance, or the constant term could vary across regimes.

· The SETAR model shows sudden transitions from one regime to another. Smooth (in time) regime switchings are offered by STAR (Smooth Threshold Autoregressive, see for instance [Guégan, 1994]) models, which, among other things, have been employed in business cycle modelling (see for example Teräsvirta and Anderson, 1992]).

2.2 - Peculiarities of the SETAR Model.

The dynamic mechanism, which generates the "threshold" structure of the SETAR model, is endogenous (the conditional expression depends on the process itself y_t, but "d" periods lagged, therefore on y_{t-d}, the *delay* variable) and of a dimensional kind (the regime transition happens when y_{t-d} reaches a given numeric value, the threshold). So the model is capable of keeping simultaneously into account many kinds of information[3], temporal and dimensional, quantitative and qualitative information, exhibiting its own dynamic not only in the temporal space (through the linear autoregressive structure in each regime[4]), but also in the dimensional one of the delay variable. It is capable of contemporaneously recording 1) the magnitude of past events (we consider significant), through the AR formulation, but also 2) their "quality", through the "threshold structure. We can better understand the importance of this characteristic through a comparison with a linear AR model. In fact this last one, at each time t, ponders the events happened in times t-1, t-2, ..., always by the same coefficients. While a SETAR *at each time t* first looks at what happened in

[3]Without expensive modelling and/or computational efforts, which could make their use not convenient.

[4]At each time only one regime is activated.

t-d, and consequently ponders the facts by the coefficients of one regime or another. Therefore it's capable of discriminating events that, even being of the same magnitude, are qualitatively different.

3 - Threshold Models, Financial Markets and Volatility.

Threshold Models are able to reproduce behaviours often observed on real data from Financial Markets: cyclical and asymmetric behaviours, occasional bursts in presence of outliers, jump phenomenons, time irreversibility.

They have been employed[5] in modelling stock returns, interest rates, exchange rates, both independently and inserted in the so-called second generation non-linear models, in which the part in-mean and the one in-variance of a process can be formulated as different models. If, say, $y_t = g(x_{t-1}; a) + \varepsilon_t$ is the part in-mean of a process, and $Var(\varepsilon_t | x_{t-1}) = \sigma_t^2 = f(x_{t-1}; q)$ is the part in-variance, where a and q are vectors of parameters, and x_{t-1} is the information set available until time t-1, g(·) could be a Threshold Model and f(·) a GARCH, ..., or g(·) a Threshold Model and f(·) a Threshold-GARCH.

Searching for a functional specification for conditional heteroskedasticity, a model must be capable of capturing at least some of the empirical stylized facts of volatility (of the binomial returns-volatility): leptokurtic non-conditional distribution of asset returns; clustering (periods of high-low volatility are followed by periods of high-low volatility); leverage effect (prices movements are negatively correlated with volatility); persistence of the shocks on volatility; *smile* effect (biases in evaluation of option prices, arising from the use of implied volatilities); influence on volatility of information arrivals (since frequencies of information arrivals and prices recording are different); volatility comovements of speculative markets of different Countries.

The behaviours reproducible by Threshold Models lend themselves to describe, variously combined, some aspects of volatility too: one can associate asymmetric behaviours with leverage effect, time irreversibility and jumps with clustering, cyclical patterns with information arrivals and persistence, bursts with fat tails in the returns distribution.

If these characteristics suggest to try and use a Threshold Model for the Italian

[5] See Chappell et al.[1996], Zakoian [1994], Li [1996].

Stock Market Volatility, the specific model we are going to employ here is, according to the considerations of section 2.2, a SETAR model.

4 - Building SETAR Models: *Key* Instruments.

The main idea of the modelling procedure proposed by Tsay [1989; 1991; Cao and Tsay 1993], is "to transform a SETAR model into a switching regression problem, that is into a regular change-point problem in linear regression analysis, for which statistics can be derived to test for model changes and to explore the dynamic structure of the process. This is achieved by using the concepts of *local estimation* and *arranged autoregression*".

Local estimation, exploiting *arranged autoregression*, aims to obtain a sequence of estimates that, once put on a graph, allow us to locate the position of the thresholds. While *arranged autoregression* is employed also after the identification of the thresholds, in order to proceed with the final estimation of the other parameters of the model.

Local estimation is a sequential estimation which recursively provides a sequence of "local" estimates, since it employs a fixed-length window of data: beginning with some initial data, at every step it adds a new observation deleting the corresponding oldest one. It can be done efficiently by a recursive least-squares algorithm or, in case of missing values in the data, by a Kalman filter.

Arranged autoregression is simply an autoregression, but it uses the magnitude of the *threshold variable* y_{t-d} to arrange the data, not the time index t. Consider, for instance, a SETAR(2), where the first regime is given by $L_1 =]-\infty, 0[$, and the second one is $L_2 = [0, \infty[$:

$$y_t = \begin{cases} 1.3\, y_{t-1} - 0.4\, y_{t-2} + \varepsilon_t^{(1)} & \text{if } y_{t-2} < 0 \\ 0.3\, y_{t-1} + 0.4\, y_{t-2} + \varepsilon_t^{(2)} & \text{if } y_{t-2} \geq 0, \end{cases}$$

Such a set-up cannot yield consistent parameter estimates by an ordinary autoregression. The problem can be solved using an *arranged autoregression*. To understand better its mechanism consider for example 10 consecutive observations of the proposed model (see table [4]). The table clearly shows the set-up of the components inside vectors $\mathbf{y}_t$, $\mathbf{y}_{t-1}$, $\mathbf{y}_{t-2}$. In an ordinary autoregression it coincides with the time order, in an *arranged autoregression* with the dimensional one (given by the magnitude of the threshold variable y_{t-2}): the elements of $\mathbf{y}_{t-2}$ follow an increasing order, and the elements of $\mathbf{y}_t$ and $\mathbf{y}_{t-1}$ are consequenly rearranged, preserving the original "intra-time" relation.

For final estimation purposes, when the delay variable and the thresold value (-0.41, say) are known, the first four rows of the arranged autoregression are the data we employ to estimate (by OLS, *ordinary least squares* method) the parameters in the first regime, the last four ones to estimate the parameters in the second regime: for each regime we therefore *separately* run an OLS estimation.

The joint utilization of Local estimation and Arranged autoregression (that is the so-called *Recursive Local Fitting*: see Appendix [I]) to locate thresholds provides the so-called *scatterplots*: they plot the local estimates of some sample statistics (such as AR-coefficients and their own *t-values*) versus rearranged (in increasing order) values of the threshold variable y_{t-d}. The local estimates should be stable before the first change point enters the estimation window: so, ideally, the scatterplots should look like a step function, with jumps indicating the values of the threshold variable at which the regime changes. In practice, since the windows overlap sequentially, the plots tend to show certain smooth transition from one regime to another. But it isn't a serious problem (this point requires further research [Tsay, 1991]) because the main objective of these scatterplots, which aren't formal tests, is to provide information on the possible partitions of the space of the threshold variable. Once the threshold values are located, we can partition the space into several regimes, and estimate (separately) an AR model with an appropriate order in each regime.

4.1 - The Modelling Procedure for SETAR Models.

Estimating a SETAR model[6] requires previously the identification of some parameters: the AR order[7] k, the delay parameter d, the number of regimes *l*, and the thresholds $r_1, ..., r_{l-1}$. So, the modelling procedure proposed by Tsay consists

---

[6]Which consists in estimating the AR coefficients and the disturbance variances of the various regimes.

[7]Maximum order, since (see step 10) the AR order may be different in different regimes, and at the beginning we have to set at least an upper bound.

of several steps (described in the following):

1) select the maximum AR order, according to the PACF and ACF functions (and Ljung-Box statistic), and/or AIC [Tong and Lim, 1980] and/or SC information criteria;

2) perform nonlinearity tests[8] against an unspecified alternative (the *F-test* of Tsay [Cao and Tsay, 1993] and the *Augmented F-test* of Luukkonen, Saikkonen and Teräsvirta[9] [Guégan, 1994]);

3) select a set of possible values for the delay parameter d;

4) for each value of d entartained, perform linearity tests against the hypothesis of threshold nonlinearity (*Threshold test* and *General Nonlinearity test* of Tsay, illustrated in Appendix [2]);

5) choose d according to the outputs of the tests at step 4;

6) perform a *Recursive Local Fitting* to locate, looking at the *scatterplots*, the possible values of the thresholds;

7) estimate[10] a SETAR for each possible threshold value entartained at step 6;

8) select the threshold value which allow you to minimize the global AIC (we are using the principle of the selection criteria, associated with a loss function to be minimized);

9) evaluate the adequacy of the adopted specification, analysing the residuals, by, say, the BDS test [Brock et al., 1991] and the ACF and PACF functions, to check, respectively, the i.i.d. and the uncorrelation assumption of the residuals;

10) refine, if necessary, the estimated model, by using the AIC (and/or SC) criterion, and other model evaluation techniques, to provide a proper specification for the AR orders, the delay parameter, and the threshold values.

---

[8]Before fitting a non-linear model [Granger and Teräsvirta, 1993], it is recommended to test the non-linearity of the data, using general linearity tests and/or tests against a specific alternative.

[9]These tests and the ones at step 4 are among the most suitable for Threshold Models, but we can perform other tests too.

[10]With known thresholds (as in the Tsay's procedure) the OLS estimates of the AR coefficients and of the residual variances are strongly consistent (see Guégan [1994]) if the process y_t, following a SETAR model, is ergodic, and the disturbance variances in all regimes are finite.

Application to the Italian Stock Market Volatility

5 - MIB30 and Volatility: Data Investigation.

Daily closing prices (from January 4, 1994 to December 30, 1997) of MIB30 stock index are employed here to model the Italian Stock Market Volatility. Building the volatility series requires the following steps:

1) calculate daily returns R_t: $R_t = \ln(P_t) - \ln(P_{t-1})$ using daily stock index prices P_t;

2) filter returns from possible forms of linear dependence[11] due to calendar[12] anomalies (or regularities) and nonsynchronous trading[13] effect, regressing returns on calendar (impulse) dummies and on a certain number of lagged returns themselves R_{t-1}, ...;

3) inspect if there is a GARCH effect for the conditional variance of the residuals $\hat{u}_t$ obtained by filtering the returns in the way illustrated at step 2. If it is present adopt a GARCH specification for them;

4) build the volatility series[14]: $\sigma_t = |\hat{u}_t| \sqrt{(\pi/2)}$;

5) since this volatility measure is skewed, use its Box-Cox transformation, in order to increase the efficiency of parameter estimation and to aid model interpretation [15]: $y_t = (\sigma_t^{\lambda} - 1)/\lambda$.

The Threshold Model will be fitted to this series y_t.

---

[11]It is common practice in any non-linearity analysis: in fact, getting rid of every linear dependence, possible residual dependencies should be non-linear.

[12]They are stock prices changes sistematically recorded in specific calendar dates, say, initial day of trading week, or of trading month, ...: checking their existence is a way to evaluate (informative) efficiency of the market.

[13]This effect arises when we record some data (which are generated at irregular intervals) at constant time intervals. It can cause biases in the autocorrelations: we can therefore detect its presence analysing these last ones.

[14]Many authors build the volatility series in this way, since it provides an unbiased estimator for the conditional standard deviation (volatility) of returns.

[15]λ is chosen in order to allow the distribution of y_t to be as near as possible to the Gaussian one: λ=0.4.

5.1 - Italian Stock Returns: Empirical Evidence.

Here are the empirical results from data analysis:

1) regressing returns on calendar dummies, that is days of the week (Monday, ..., Friday), Post-Holiday days (open market days following holidays different from Saturday and Sunday) and initial days of trading month[16] (following the last day of trading month, that is the contangoes-day, or "*giorno dei riporti*": InTradMo), Monday falling tendency and the settlement (end of trading month) effect have resulted significant;

2) from the inspection of autocorrelation functions (ACF e PACF) of the residuals obtained at step 1), we found significative the autocorrelation at lag 7. Running a rather long[17] (until lag 30) residual autoregression the presence of a nonsynchronous trading effect ha been confirmed;

3) the ARCH-test on the residuals obtained regressing returns simultaneously on calendar dummies and their own lags has pointed out the presence of a GARCH effect.

Concluding this analysis, the stock index returns, filtered from linear dependencies and with a GARCH(1,1) specification for the conditional variance, are described by equations [1]-[2]. Table [1] shows estimation outputs.

$$[1] \quad R_t = a_1Mon + a_2Tue + a_3Wed + a_4Thu + a_5Fri + \\ + a_6PostHol + a_7InTradMo + a_8R_{t-1} + a_9R_{t-2} + \\ + a_{10}R_{t-3} + a_{11}R_{t-4} + a_{12}R_{t-5} + a_{13}R_{t-6} + \\ + a_{14}R_{t-7} + a_{15}R_{t-8} + a_{16}R_{t-9} + a_{17}R_{t-10} + \hat{u}_t$$

$$[2] \quad \sigma_t^2 = b_0 + b_1\, \hat{u}_{t-1}^2 + c\, \sigma_{t-1}^2$$

We have now available the proper estimates $\hat{u}_t$ to build the volatility series y_t.

---

[16] After the shift (February 16, 1996) from *"Negoziazione a Termine"* (meaning "term trading") to *"Liquidazione a Contante"* (meaning "cash settlement"), the dummy for the beginning of trading month has been set equal to 0.

[17] Since (Pagan, [1990]) nonsynchronous trading could be compatible with a MA(1) residual structure, MA(1) dei residui (really pointed out on some sub-samples).

5.2 - Volatility Models for Comparisons.

We present here the models used in the comparisons with the SETAR model (tables [2]-[3] show estimation outputs).

The GARCH models proposed here are a GARCH(1,1) and a GARCH-L(1, 1)[18]. This last one, compared to the GARCH model, introduces an asymmetric term in order to take into account the leverage effect:

[3] $\sigma_t^2 = b_0 + b_1\, \hat{u}_{t-1}^2 + c\, \sigma_{t-1}^2 + d \times I(\hat{u}_{t-1} \leq 0) \times \hat{u}_{t-1}^2$.

As regards the linear model, the best linear model fitting the data has resulted to be an AR(5), estimated on the same series y_t employed for the SETAR:

[4] $y_t = a_0 + a_1\, y_{t-1} + a_2\, y_{t-2} + a_3\, y_{t-3} + a_4\, y_{t-4} + a_5\, y_{t-5} + \varepsilon_t$.

6 - A SETAR Model for Italian Stock Market Volatility.

We show here the most significant results of the analyses conducted to specify the SETAR model for the Volatility of the Italian Stock Market.

1) Choosing the AR order: we have found k = 5 as maximum plausible AR order, which is compatible with our data, since there are five days in the trading week.

2) Linearity tests against an unspecified alternative: the BDS test points out some non-linear dependences in the data (tables [5a] and [5b]). So do the *F-test* and the *Augmented F-test* [19] (table [6]): the first one, in particular, heavily rejects the null of linearity for all AR orders, including k = 5, the selected one.

3) Possible values for the delay parameter: following several authors (Tong, Tsay, ...) we choose those values for d (integer numbers) such that $1 \leq d \leq k$, here $1 \leq d \leq 5$.

4) Detecting "threshold" non-linearity: the *Threshold test* and the *General Nonlinearity test* test the null of linearity against the "threshold" non-linearity. Tables [7] and [8] show rejection of the null. It is important to notice that we have the most significant results, when changing d, at k = 5, which is the AR order we have chosen before.

---

[18] The order (1, 1) has been selected among several possible combinations by using the Schwarz (SC) and the Akaike (AIC) information criteria.

[19] We need to have known or previously fixed the AR order to perform these tests.

5) Identification of the delay parameter: we need to have known, or previously fixed, both the AR order and the delay parameter to perform the *Threshold test* and the *General Nonlinearity test*. Since d, on the contrary, is unknown, Tsay [1989] proposes to perform these tests for several values of d (chosen at point 3) and select that value of d which provides the most significant outputs of the tests (that is the minimum *p-value*, or tail-probability). He therefore suggests such tests also as a technique for the identification of d. For such purpose it is useful to look at the outputs of the tests even for values of k different from the pre-selected one: tables [7] and [8] show that, when changing k, d = 1 seems the most plausible choice.

6) Looking for possible threshold values: once we have selected d = 1 as the delay parameter, we have to detect at which value/s of y_{t-1} there is/are the transition/s from one regime to another. Almost all the scatterplots (see for instance figures [1a,b] and [2a,b])[20] show a burst, at which the relatively stable[21] plot of the estimates begins to be- come unstable, displaying a break, or discontinuity, ap-proximatively around -2.055. So we have two regimes: the threshold will be a point belonging to a sufficiently large interval around -2.055, and detected by a grid-search.

7) Estimation to select the threshold value: we estimate, by OLS, an AR(5) model using the data corresponding to the values of y_{t-1} smaller than -2.055 (first regime), and an AR(5) using the data corresponding to the values larger than -2.055 (second regime). Then we calculate the overall AIC (adding the AIC of the first regim to the AIC of the second regime). We apply this procedure to every point (by a grid search) in the interval chosen at step 6.

8) Identifying the threshold value: among the points inspected at step 7, the overall AIC is minimized when y_{t-1} = -2.0682226, which we select as the threshold value.

In this way we have specified and estimated a SETAR(2; 5, 5; 1) model:

---

[20]As it is common practice, the last 13 points of the scatterplots have been omitted, since extreme observations tend to become less frequent. We have reported only two graphs for spatial requirements, but the other graphs give the same indications as the two reported graphs.

[21]Tsay uses the terms "relatively stable", since the scatterplots of the different (or all of the) coefficients mightn't be clear (as it happens in the cases examined by the author). But we emphasize that their object is to give some indication about the presence of a break, identifying, just approximately, the location of the threshold.

$$y_t = -1.549 - 0.081y_{t-1} + 0.124y_{t-2} + 0.117y_{t-3}$$

$$+0.054y_{t-4} + 0.055y_{t-5} + \sigma_1\varepsilon_t \qquad \text{if } y_t \leq -2.068$$

$$y_t = -1.416 + 0.217y_{t-1} + 0.028y_{t-2} - 0.011y_{t-3}$$

$$+0.014y_{t-4} + 0.084y_{t-5} + \sigma_1\varepsilon_t \qquad \text{if } y_t > -2.068.$$

Following Cao and Tsay [1993] we delete only the more non-significant variables (t-value criterion, table [9]): second, third and fourth lag in the second regime.

9) Evaluating the adopted specification: analysing the standardized residuals (that is $e_t^{(j)}/s_t^{(j)}$), since the disturbance variance too is allowed to be different in different regi- mes) of the estimated model, the adopted specification seems adequate. There are no longer residual autocorre- lations looking at the PACF and ACF functions, and the BDS test fail to reject the i.i.d. hypothesis of the residuals.

The *fitting* analysis too confirms the adequacy of the adopted specification: the estimated values over the whole sample period reproduce rather well the features of the real values. Particularly, the estimated model has been capable of recognizing, for a good part, periods of high and of low volatility. In fact 1) it locates the most part of 1994 in the regime of high volatility, with increasing frequency month after month (the expectations of growing interest rates, which is a factor influencing the stock market volatility, occured in the middle of August); 2) more than 60% of the estimated volatilities of the years 1995 and 1996 fall inside the regime of normal-low volatility (after the investors' pessimism to Italian markets, causing falling stock prices in '95, we had a stagnation in '96, producing rather low average volatilities); 3) the two thirds of the '97 estimates falling in the regime of high volatility belong to the second six months of the year (the crisis from the Far East markets had appeared on other markets in summer) and the last week of October 1997 is located in the uppermost band of the regime of high volatility.

6.1 - Interpreting the Model: Remarks on Volatility.

Given the interpretative potential of the SETAR model (section 2.2), the estimation results allow us to explain better some points, which will have important consequences in terms of forecasting and portfolio risk evaluation.

· The identification of two regimes clearly identifies a regime of high and a regime of ordinary-low volatility.

· This fact, together with the identification of y_{t-1} as the threshold variable, easily describes clustering: at each time t the SETAR model fits the volatility using the parameters of the same regime (identified by the level, high or ordinary-low, of the volatility which has occurred at time t-1) in which the real volatility has fallen the previous time t-1. This is the analytical version of the fact that low - high volatility tend to be followed by low- high volatility.

· The finding d = 1 as the delay parameter, allows a timely updating of the information. In fact, the conditioning event (which is the interpretative key of the facts, permitting the model to decide if it has to ponder them by the coefficients of one or the other regime) is the value of y_{t-1}: this means that the information contained in it can be incorporated from the immediately following instant.

· The mechanism controlling the transition from one regime to another allows the model to notice (quite timely, since it is based on y_{t-1}) changes in the trend, even though unexpected, provided they are sufficiently[22] large. This is the case of *shocks*, a frequent phenomenon in financial markets, and typical of the volatility: if, say, a shock[23] occurs in t-1, there will be in t a regime transition.

· The same mechanism, pondering the events by parameters which are different according to the activated regime (section 2.2), allows the model to recognize the extraordinary or persistent nature of a *shock*. In fact a *shock*, while having the same magnitude, might be immediately absorbed by the market or last for a longer time.

7 - Forecastings for the Italian Stock Market Volatility.

Forecasting is one of the main objectives of modelling, and volatility forecasting is fundamental, as we have seen, in the operational context of Financial Markets.

We judge here the forecasting performances of the employed Threshold Model not only in absolute terms, but also comparing them with other alternative models,

[22]We have a regime change overstepping the threshold.

[23]Since it means (by definition) recording a particularly small or large value, with consequent overstepping of the threshold (in either direction, even though in this context the events inducing high volatility are more important).

particularly a GARCH(1,1), a GARCH-L(1,1), a linear AR(5).

We make short, medium and long-term predictions (that is 1, 2, ..., 5, ... 10, ... 30 days ahead). The forecasting horizon is October 16, 1997 - December 30, 1997: therefore we have re-specified and re-estimated all the models on the reduced sample period, obtained eliminating those days, and we have used them to make predictions. We have started from October 16 in order to include the days in which the crisis of the Far East markets has affected the Italian stock exchange: in fact it is very important to evaluate the performances of a model also in particularly turbulent periods for the markets, during which it becomes even more necessary (being more difficult) to have reliable forecasts available.

For each model[24] the one-step ahead predictor at time t for time y+1, $\hat{y}_{t,\,t+1}$, is given by the conditional expected value[25], $\hat{y}_{t,\,t+1} = E(y_{t+1} | I_t)$, where I_t is an information set including past information available until time t. While multi-step ahead forecasts are obtained by Monte Carlo method for the SETAR model (given its appreciable performances compared to other methods, as regards Threshold Models[26] too), and recursively (that is employing iteratively the formulas of the respective models) for the GARCH-type models (in the GARCH-L formula the conditional expected value of the indicator function is set equal to 0.5) and the linear model.

We adopt as indicators of the forecasting performances several measures based on the difference between predicted and real values (AAD, *absolute average deviation*; MSE, *mean square error*; *Theil Index*; MEDSE, *median square error*), and the correct-signs percentage, both 1-step and 30-step (according to this last one, we forecast for t + 1, t + 2, ..., t + 30, having time t as the time origin, and calculate on the values so obtained the percentage of increasing or decreasing changes correctly forecasted). Better forecasts are provided by models having lower values of the indicators AAD, MSE, Theil Index, MEDSE (we calculate the ratio of the indicator of an alternative model on the indicator of the SETAR model: if the ratio is greater than 1 the SETAR model provides more accurate forecasts), and larger values of the correct-signs percentages.

We use several indicators in order to give a more objective judge of the

---

[24]We obtain predictions also using a Random Walk model, only for theoretic "curiosity", since it is usually the benchmark for the part-in-mean of returns or rates.

[25]In the Random Walk case, the forecast made in t for t+1, as for t+2, ..., is the value occurred in t.

[26]See for instance Clements and Smith [1997].

obtained results. Since a model could be preferable according to one indicator, and worse according to another one, while if it is preferable according to many indicators, its "supremacy" is more reliable.

For the same reason we evaluate the predictive performances on several forecasting horizons, in order to investigate if a model is sistematically preferable or just in some specific period. Table [10] and figures [3a, b, c, d] show the results obtained for the period October 16 - December 30, 1997.

7.1 - Importance and Implications of Forecasting Performances of the SETAR Model.

· Table [10] show that the SETAR model displays the best performance according to all the indicators, and for all the prediction horizons, that is for 1-step, 2-step, ..., 30-step forecasts[27].

· This result is reliable since analogous results have been obtained for other several forecasting periods analysed.

· The reduction in 1-step MSE and 1-step Theil Index, compared to the corresponding AAD, is due to the fact that the first two measures employ squared prediction errors, so they are affected by larger errors, as in the case of outliers. In fact we have here the shock of Tuesday October 28, which reveals itself to be unexpected to all the models, making them to record a prediction error approximately of the same magnitude. In order to reduce the outlier effect we therefore calculate the MEDSE (in place of MSE and/or Theil Index): the priority of the SETAR model turns then fully evident.

· The graphics show that GARCH[28] forecasts tend to reproduce the features of true values, but with a certain delay, exactly equal to the number of steps ahead employed to obtain the predictions[29]. That is, forecasted points look like a (ahead) shift of the true ones. If this can be interpreted as a form of volatility persistence, which, for fitting purposes, makes the model capable of reproducing this empirical stylized fact of the volatility, for forecasting purposes it may become a

[27]The SETAR model displays a better performance even than the linear one in the majority of the cases, and always in the 1-step case.

[28]The GARCH-L graphics are quite similar.

[29]We report only (for spatial requirements) the 1-step and 5-step forecasts graphs, but intermediate and following graphs clearly, and progressively, show this behaviour.

drawback. Because of the alternating (almost alternatively increasing or decreasing) behaviour of volatility, this characteristic may affect the forecasting performance, especially in presence of transitory shocks. We can see better this effect pushing forecasts ahead, that is beyond 1-step. See, for example, the graphic of 5-step forecasts: the shock of Tuesday October 28 persists until the following Tuesday, considerably increasing the prediction error. In fact that shock was absolutely extraordinary, and was immediately absorbed by the market: the following week volatility was not only already gone down, but had continued falling until Friday. But the GARCH model isn't capable of recognizing the nature of the shock, treating it as if it was persistent. The consequences of this fact depend of course on the estimation results, since the memory of the persistence effect is related to the values of estimated coefficients, but it is in any case a structural characteristic of the GARCH model[30].

· While the SETAR model doesn't present this problem: its endogenous "threshold" mechanism, based on the delay variable, here y_{t-1}, (see sections 2.2 and 6.1), allows it not only to notice (timely) a shock, but also to distinguish its extraordinary or persistent nature. So the model is able to follow, even in perspective, that is forecasting, terms, the market.

Obviously also here the sensitivity of the model depends on estimation results (number of regimes, values of coefficients and significance of variables): its characteristic structure anyway provides the model with a potential advantage, which explains the better performances of the SETAR model compared to the other models employed.

· We then point out that the SETAR model displays the best performance among all the other models as regards the 1-step correct-signs percentage. But it is the best, even in absolute terms, as regards the 1-to-30-step percentage, recording the value of 52%. This fact is of immediate interest when we want to know if in the short, medium and long term volatility will tend to increase or diminish. This is important in every *scenery* analysis, both in the context of Financial Markets, both in the macroeconomic one (if we accept the idea [Schwert, 1989] of relations between stock volatility and the volatility of macroeconomic variables).

[30]In fact they are able to reproduce the persistence effect.

8 - VaR: a New Measure for Portfolio-Risk.

In the last ten years the trading activity of financial institutions has been considerably growing (given the acceleration of the process of financial innovation) with a consequent increase of their risk esposure. So they have increased their effort to develop risk management systems, which could provide suitable measures for risk exposure. Financial regulators have also begun to focus their attention on the use of such systems by regulated institutions.

As a result, the Basle Committee on Banking Supervision, in the 1996 Basle Proposal, suggests to financial institutions the procedures by which they have to fix some criteria for capital requirements assessments, in relation to their own market risk exposure. The 1993 EU Capital Adequacy Directive (CAD) regulates such criteria: a financial institution with significant trading activity has the choice between a regulation model, laid down by the CAD, and "internal models", created by the institution itself, which are consistent with the Basle Proposal.
If the CAD model quantifies market risk solely in relation to the institution's trading book, the instructions given by the Basle Proposal take into account important market conditions too, being based on the VaR estimates, which depend on market volatility. It is indeed important to have a model providing good volatility forecasts also in the context of this new method [Aussenegg and Pichler, 1997] for Portfolio Risk evaluation, based on VaR estimates.

In an expected probability distribution of returns, VaR is definable as that negative return (that can be interpreted as a loss) lying on the left tail of a distribution (the Gaussian distribution is usually assumed), at which we have a critical probability level (α), previously chosen, say 5%, 1%, or whatever, provided it is rather low. VaR for time t is defined by: $\text{Prob}(R_t \leq \text{VaR}) = \alpha$, R_t being the returns.

Such low probability values are consistent with larger losses, since this approach aims to provide forecastings of larger losses, at a pre-selected significance level: portfolio investment decisions aren't based only on the expected return-risk relation, but also on hedging costs. So we need to measure, in probabilistic terms, this kind of losses, using VaR as a risk measure: with a 1-day holding period (from t to t+1), the capital requirement for general market risk is: $CR_t = \sqrt{10} \times \max\{\text{VaR}_t; f/60 \times \Sigma_{j=0, ..., 59} \text{VaR}_{t-j}\}$, where f is a multiplication factor, depending on the *zone* in which the model is placed according to Backtesting responses.

Backtesting is a test methodology to evaluate the overall reliability of a model for VaR purposes, according to the indications given by the 1996 Basle Proposal. Having a sample of T data it requires estimating the model on a certain number of initial values, t for instance, and forecasting VaR for time t+1. We then update the sample with the observation of time t+1, re-estimate the model and forecast VaR for time t+2, and so on, recursively, for all the remaining T - t data[31]. For each j = t + 1, ..., T - 1, being R(j) the return observed at time j, and VaR(j) the VaR forecasted in time j-1 for time j, we count the number of times (N) we have obtained that R(j) $\leq$ VaR(j). The nearest this number is, in percentage (that is N/[T-t]), to the pre-selected probability level, the most reliable is the model. More precisely the Basle Proposal fixes some "zones": for $\alpha = 1\%$, retaining the last 250 data to perform the test, if the above-mentioned relation occurs less than 5 times (out of 250) the model is regarded as reasonable, from 5 to 9 times, some questions arise about the quality of the model, 10 or more times the model is considered unreliable, and the supervisor will require the bank to find a better model.

Distribution forecastings are usually 1-day predictions. Forecasting a probability distribution requires us to have a model both for the part-in-mean of returns and their volatility, to forecast its mean and variance.

Even though a financial institution can create its own models for VaR estimation, there are in any case some schemes available, already predisposed. The one provided by Riskmetrics, for instance, which publishes the related values, follows the EWMA (Exponentially Weighted Mo- ving Average) approach. Assuming a normal distribution, it calculates volatility as the square root of:

$$\sigma_t = (1 - \lambda) \times \Sigma_{t=1, \ldots, T} \lambda^{t - 1} \times (R_t - \mu)^2$$

where l ($0 < \lambda < 1$) is a constant (decay factor), set by RiskMetrics equal to 0.94; T is the number of observations in the estimation window[32] we chose; R_t are the real returns; m is the expected value of returns (usually set by RiskMetrics equal to 0); t is the time index, but in inverse time order (the most recent observation has t =1).

[31]One can otherwise perform a *rolling* estimation.

[32]It's the same thing as the sample period for a model.

8.1 - Estimating VaR: Comparing SETAR with Other Models.

Estimating VaR requires us to obtain previously the expected probability distribution of returns, so we need the forecasting equations of mean return and related volatility: the mean-forecasting equation given by the SETAR model is the same as the GARCH and linear AR models (equation[33] [1]), while the volatility-forecasting equation is specific for each model. So, different performances in estimating VaR will be attributable to the capability of the models of providing more or less reliable volatility prediction, since they have a common mean-equation, while it is volatility modelling the discriminating factor.

Before estimating VaR for specific days, we perform the Backtesting procedure: if a model passes the test, the specific predictions it provides will be regarded as reliable. The Backtesting procedure has been performed on the last 550 data of the sample (having in total 989 data), so, for each model, the estimation has been initialized on the first 989-550 = 439 data, and then updated (in the parameters) recursively[34] every 5 data, increasing progressively the sample dimension.

Choosing a significance level of 5%, the most reliable model is the one that will provide the nearest value to 5%.

· As we can see (table [14]) the SETAR model displays the best performance; RiskMetrics, the GARCH and GARCH-L models impose not negligible hedging costs, while the linear model exposes the investor to a risk against which he isn't covered. Which becomes more dangerous in periods of greater turbulence, as it has been during the days of the crisis in Far East markets (last days of October and first days of November 1997).

Then, in table [15] are reported the 1-day VaR estimates for the last week of October 1997, the most critical, because of the big slump of Tuesday 28.

· The goodness of the estimates provided by a model is judged according to the

[33]Since this specification had appeared adequate (section 5.1). As regards GARCH-L, its mean-equation isn't identical with GARCH equation, but differences are after all negligible. RiskMetrics assumes zero-mean.

[34]Possible non-significant variables haven't been deleted, since it is a recursive estimation. As regards RiskMetrics, calculating volatility as an exponential weighted moving average, using the term estimation is improper: the procedure is anyway recursive, with one observation at a time.

closeness to the value provided by the Historical Simulation Approach. This approach employs as the VaR(t+1) estimate, the α-percentile (of the left-tail) in the set of returns (rearranged in increasing dimension order[35]) from time 1 to time t. We can see that once again the SETAR model has the best performance, while all the other models give very misleading results.

· It is very important to have such performance in such problematic days. In fact, from a regulator's point of view, a model is required not only to provide the most accurate VaR estimates as possible, but also to keep its own good performance even in particularly turbulent periods. In fact in these situations the need of having adequate risk management systems available becomes greater, since risk exposure, already generally increased during these last ten years, considerably grows in similar days.

· In estimating VaR, it isn't important only to provide the most accurate forecasting as possible. Especially for a bank it is important to know if a model tends to underestimate or overestimate VaR. In fact, in the first case it displays a prudential attitude, since it attributes to some loss levels, larger than the pre-selected critical threshold, a greater probability for them to occur than the real one. In the investment consequent to this kind of estimate, one will then tend to make bigger hedges than is necessary. As a result, one is less exposed to risk, but bears additional hedging costs (not negligible) that aren't necessary at all. While, in the opposite case, hedging costs are smaller than those required against the losses which one is interested in, and with respect to which VaR has been calculated, but one becomes more exposed to the related risk[36].

As we can see, during the analysed week the SETAR model, compared to the other models, tends to overestimate VaR, which is particularly interesting. In fact, if from a regulator's point of view a model should provide the most accurate VaR estimates as possible, from a bank's point of view they should 1) be as accurate as necessary and 2) lead to low capital requirements. From a bank's point of view predictive accuracy means that the model should be reasonable according to the backtesting results, but it simultaneously should satisfy an economic criterion of minimization of costs of capital. This is why between two models accettable from

---

[35]In this case the distribution isn't assumed, as models usually do, but it is used the historical one.

[36]If, for a model, *Backtesting* gives a probability level smaller than the critical one, it tends to underestime VaR, in the opposite case it tends to overestimate.

a regulator's point of view, a bank will choose the model that tends to provide VaR overestimates or smaller underestimates.

Therefore, even from a bank's point of view the SETAR model displays here an appreciable performance: in fact, the Backtesting shows that it tends to underestimates less than the other models (it is the nearest, by diminution, to 5%), and during the analysed week, it provides also some overestimates, versus the excessive underestimates of all the other alternative models.

· We see that the big slump in price of Tuesday 28 is unexpected for all the models, since their prediction errors have almost the same dimension.

· But after this slump, while the GARCH and GARCH-L models, which are capable of noticing a shock immediately, incorporate the volatility burst, displaying some persistence (as do, even though less evidently, RiskMetrics and the linear model), the SETAR model incorporates this impulse, but regards an unexpected and extraordinary burst exactly as it is: from the next day on, as the slump is absorbed by the market, so it is absorbed by the SETAR model, and only by itself among all the other models.

This capability of distinguishing the extraordinary or persistent nature of a shock is due to the peculiar nonlinear structure of the SETAR model: recalling that it follows an AR(5) in each regime, while the linear model used is an AR(5), it becomes even clearer the advantage offered by the "threshold" nonlinearity (section 2.2). The AR(5) incorporates the burst of Tuesday 28, and, given the autoregressive linear structure, it retains *linearly* memory of it: *linearly* means that, for every y_t, it always ponders what happened in t-1, ..., t-5 by the same weights. So, in the days following Tuesday 28, the burst will be memorized as any other volatility display, even a normal one. But in this way it prevails the dimensional impact of the shock over its transitory nature.

· We can make analogous considerations about the GARCH and GARCH-L models, and Riskmetrics. More specifically GARCH-type models are capable of capturing volatility persistence. But in similar situations, in which the shock is going to be absorbed immediately, that is not to persist, this characteristic becomes disadvantageous. In fact, in such days, in the general nervousness of the market, there is even more need of having reliable forecasts available. Whilst the excessively prudential attitude of the GARCH and GARCH-L models and RiskMetrics leads the investor towards an equally prudential attitude, but with the burden of unjustified costs.

· The situation here is particularly critical, since we have verified that in the following week the GARCH and GARCH-L models (but also RiskMetrics) persist in underestimating considerably VaR, while in those days the crisis isn't only already absorbed, but there is also a progressive decrease of volatility, lasting until Friday. On the contrary, in the SETAR model already since October 31, it hasn't been persisting the burst effect, but it has been prevailing the adjustment effect.

· The SETAR model has revealed itself the only model capable of (timely) recognizing the extraordinary or persistent nature of a shock.

9 - Concluding Remarks.

The results obtained in this work have fulfilled the intentions with which it was been undertaken, that is to investigate the interpretative potentialities of other non-linear models (in this case Threshold Models) compared to the ones usually employed, but not entirely satisfactory (we can just recall Hsieh [1991]: "ARCH-type models do not fully capture the nonlinearity in stock returns"), in the analysis of financial markets. In fact the Threshold Model built for the Volatility of the Italian Stock Market has revealed itself capable of capturing the clustering effect and the asymmetric reaction to negative or positive shock of a certain dimension. But, above all, it is the only one, among all the models considered here, which has been able to recognize, quite timely, the extraordinary or persistent nature, of a *shock*. This characteristic, among the others, has allowed it to display a better performance than alternative models (such as GARCH, GARCH-L, linear model, RiskMetrics), both in terms of *volatility forecasting* and of *portfolio risk evaluation* by the new VaR methodology. Not only in the sample period here reported, but also in other periods investigated for check purposes, displaying even better performances during a particularly turbulent period for the market, that is the last days of October 1997, because of the crisis in the Far East markets. This preminence is attributable to the capability of the model of interpreting the events both from a *quantitative* and a *qualitative* point of view, since it is allowed at any instant to change regime according to some continuously updatable information (incorporated by the magnitude of the delay variable). It is worth pointing out the efficacy here displayed by Tsay's new modelling procedure: thanks to it the significant performances of the Threshold Model haven't required expensive modelling and/or computational costs, which would have otherwise discouraged from employing these models. In fact he thought to have proposed a "relatively simple" methodology, hoping it could "help exploit the potential of TAR models in application" [Tsay, 1989]: so, in the course of this work, we haven't only been allowed to appreciate the effectiveness of this procedure, but also the theoretically-founded expectations on the potential of these models (based on the peculiarity of their dynamic mechanism) have been supported by the empirical evidence here investigated. Therefore all the results encourage further research and investigation of Threshold Models, both from a methodologic and an applied (to other problems concerning Financial Markets) point of view.

REFERENCES

Aussenegg, W., Pichler, S., 1997, "Empirical Evaluation of Simple Models to Calculate Value-at-Risk of Fixed Income Instruments", *Working Paper*.

Brock, W.A., Hsieh, D.A., LeBaron, B., 1991, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, MIT Press, Cambridge.

Cao, C.Q., Tsay, R.S., 1993, "Nonlinear Time-Series Analysis of Stock Volatilities", in M.Hashem Pesaran and Simon M. Potter (eds), *Nonlinear Dynamics, Chaos and Econometrics*, John Wiley & Sons, 157-177.

Chappell, D., Padmore, J., Mistry, P., Ellis, C., 1996, "A Threshold Model for the French Franc/Deutschmark Exchange Rate", *Journal of Forecasting*, 15, 155-164.

Clements, M. P., Smith, J., 1997, "A Monte Carlo study of the forecasting performance of empirical SETAR models", University of Warwick, *Working Paper*.

Granger, C.W.J., Teräsvirta, T., 1993, *Modelling Nonlinear Economic Relationships*, Oxford University Press, New York.

Guégan, D., 1994, *Séries Chronologiques Non Linéaires à Temps Discret*, Economica, Paris.

Hsieh, D. A., 1991, "Chaos and Nonlinear Dynamics: Application to Financial Markets", *The Journal of Finance*, 46(5),1839-1877.

Li, C.W., Li, W.K., 1996, "On a Double-Threshold Autoregressive Heteroschedastic Time Series Model", Journal of Applied Econometrics, 11, 253-274.

Pagan, A. R., Schwert, G. W., 1990, "Alternative Models for Conditional Stock Volatility", *Journal of Econometrics*, 45, 267-290.

Schwert, G. W., 1989, "Why Does Stock Market Volatility Change Over Time?", *The Journal of Finance*, 44(5), 1115-1147.

Teräsvirta T., Anderson, H.M., 1992, "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models", *Journal of Applied Econometrics*, 7, S119-136.

Tong, H., 1995, *Non-linear Time Series. A Dynamical System Approach*, Clarendon Press Oxford, New York.

Tong, H., Lim, K.S., 1980, "Threshold Autoregression, Limit Cycles and Cyclical Data (with Discussion of the Paper)", *Journal of the Royal Statistical Society*, B-42(3), 245-292.

Tong, H., 1983, *Threshold Models in Non-linear Time Series Analysis*, Springer-Verlag, New York.

Tsay, R.S., 1991, "Detecting and Modeling Nonlinearity in Univariate Time Series Analysis", *Statistica Sinica*, 1, 431-451.

Tsay, R.S., 1989, "Testing and Modeling Threshold Autoregressive Processes", *Journal of the American Statistical Association*, 84(405), 231-240.

Zakoian, J. M., 1994, "Threshold Heteroskedastic Models", *Journal of Economic Dynamics and Control*, 18, 931-955.

Table [1]

Variables	Coefficients	t-values
	Returns: Part-in-Mean	
Monday	-0.00159	-1.70232
Tuesday	0.00204	2.27428
Wednesday	0.00029	0.31945
Thursday	0.00081	0.80987
Friday	8.56×10^{-5}	0.09238
post-holiday	0.00141	0.56824
initial day trading month	0.00573	1.95797
$\mathbf{R_{t-1}}$	0.06064	3.53059
R_{t-2}	-0.00721	-0.14528
$\mathbf{R_{t-3}}$	0.04482	1.82728
R_{t-4}	-0.00419	-0.09738
$\mathbf{R_{t-5}}$	-0.04146	-2.5018
R_{t-6}	-0.00849	-0.11191
$\mathbf{R_{t-7}}$	-0.068	-3.33722
R_{t-8}	0.01876	1.06724
R_{t-9}	-4.87×10^{-5}	-0.002
R_{t-10}	0.01744	1.00375
	GARCH specification for Volatility	
b_0 (GARCH)	3.25×10^{-5}	51.12256
b_1 (GARCH)	0.14471	50.6594
c (GARCH)	0.6852	69.81542

Table [2]

GARCH-L (1, 1)		
Variables	Coefficients	t-values*
b_0	3.3052×10^{-5}	2.906712
b_1	0.10939995	2.979126
c	0.68490815	8.6635115
d	0.065093	1.344735

* We point out that the estimated coefficients of the forecasting-equation (omitting the last 50 data) are much more significant.

Table [3]

Variables	Coefficients	t-values
constant	-1.4609	-11.377
AR(1)-coeff.	0.051363	1.611
AR(2)-coeff.	0.080359	2.516
AR(3)-coeff.	0.064345	2.011
AR(4)-coeff.	0.037569	1.176
AR(5)-coeff.	0.069308	2.172

Table [4]

Ordinary autoregression					Arranged autoregression				
time	Y_t	y_{t-1}	y_{t-2}	regime	time	y_t	y_{t-1}	y_{t-2}	regime
3	-0.41	1.21	1.31	L_2	8	0.12	-1.85	-3.08	L_1
4	0.21	-0.41	1.21	L_2	9	0.58	0.12	-1.85	L_1
5	-1.12	0.21	-0.41	L_1	7	-1.85	-3.08	-1.12	L_1
6	-3.08	-1.12	0.21	L_2	5	-1.12	0.21	-0.41	L_1
7	-1.85	-3.08	-1.12	L_1	10	1.28	0.58	0.12	L_2
8	0.12	-1.85	-3.08	L_1	6	-3.08	-1.12	0.21	L_2
9	0.58	0.12	-1.85	L_1	4	0.21	-0.41	1.21	L_2
10	1.28	0.58	0.12	L_2	3	-0.41	1.21	1.31	L_2

Table [5a]

BDS test of the Volatility series				
m \ ε (*)	50 %	100 %	150 %	200 %
2	1.864169516	2.124598036	2.189606136	2.891728508
3	2.393786122	2.504387321	2.443056243	2.993312524
4	2.478191824	2.472415786	2.468097391	3.02584115

Table [5b]

BDS test of the Volatility series filtered by linear dependences [AR(5) filter]				
m \ ε (*)	50 %	100 %	150 %	200 %
2	1.555387886	1.745148729	1.553221853	2.24134242
3	1.959060198	2.091375126	1.897707305	2.501698944
4	1.648204867	1.824062538	1.822494832	2.419873909

(*) m denotes the embedding dimension, ε the percentage applied to standard deviation of the data

Table [6]

	F-test of Tsay		**Augmented F-test of Luukkonen et al.**	
AR order	test-statistic	Tail-probab.	test-statistic	Tail-probab.
1	9.2387857	0.002431613	9.8553754	0.007243233
2	4.5970825	0.003338916	15.333898	0.009026974
3	2.8228693	0.00993989	18.241817	0.032469449
4	3.128174	0.000605365	34.226972	0.001910575
5	2.1700888	0.005986367	36.139457	0.014810618

Table [7]

Threshold test					
	AR order: 5				
Delay parameter: d	k = 1	k = 2	k = 3	k = 4	**k = 5**
d = 1	4.5592765 [0.01071562]	3.9486001 [0.00819845]	18.520833 [$\mathbf{1.61\times10^{-13}}$]	37.081638 [**0.0000000**]	30.648253 [**0.0000000**]
d = 2	23.854260 [8.13×10^{-11}]	16.304724 [2.54×10^{-10}]	10.819412 [1.45×10^{-8}]	13.805061 [7.15×10^{-13}]	31.841277 [7.44×10^{-14}]
d = 3	20.165926 [2.73×10^{-9}]	14.116861 [5.39×10^{-9}]	10.471003 [2.729×10^{-8}]	9.8926031 [3.22×10^{-9}]	10.669977 [1.81×10^{-11}]
d = 4	0.18006411 [0.83524733]	0.34420029 [0.79336884]	0.24143918 [0.91485895]	3.3648421 [0.00511846]	3.0755128 [0.00551722]
d = 5	2.8094993 [0.06077433]	14.719508 [2.33×10^{-9}]	16.746221 [4.86×10^{-13}]	14.141385 [4.87×10^{-13}]	12.713210 [1.33×10^{-13}]

Table [8]

General Nonlinearity test					
AR order: 5					
Delay parameter: d	k = 1	k = 2	k = 3	k = 4	**k = 5**
d = 1	2.9725879 [0.00703167]	2.5550840 [0.00668233]	7.1918190 [**1.13×10^{-12}**]	13.420037 [**2.32×10^{-13}**]	11.380308 [**0.0000000**]
d = 2	8.7329507 [2.93×10^{-9}]	6.4501720 [6.45×10^{-9}]	4.4631743 [5.52×10^{-7}]	5.3634019 [1.97×10^{-10}]	11.377824 [0.0000000]
d = 3	8.9125960 [1.83×10^{-9}]	6.5341252 [4.72×10^{-9}]	4.6172930 [2.69×10^{-7}]	4.1939526 [1.57×10^{-7}]	4.4499370 [2.51×10^{-9}]
d = 4	3.9172126 [0.00071727]	2.6241948 [0.00535925]	1.8645330 [0.03513263]	1.9986620 [0.01306701]	1.9120363 [0.01237347]
d = 5	3.9812383 [0.00061210]	7.1575368 [4.56×10^{-10}]	7.1661346 [1.25×10^{-12}]	6.0542752 [3.52×10^{-12}]	5.3035160 [8.08×10^{-12}]

Table [9]

Estimation outputs for the SETAR model				
	Regime 1		**Regime 2**	
Variable	*Coefficient*	*t-Student*	*Coefficient*	*t-Student*
constant	-1.5487329	**-7.281555**	-1.4163591	**-6.397839**
y_{t-1}	-0.0812831	-1.2435016	0.21719318	**2.5273943**
y_{t-2}	0.1241825	**2.8092206**	0.028371614	0.62547741
y_{t-3}	0.11679665	**2.6563919**	-0.010904324	-0.23652465
y_{t-4}	0.05375207	1.2294935	0.014428353	0.31452342
y_{t-5}	0.054710488	1.2201495	0.083818575	**1.8787135**

Comparing forecasting performances of the SETAR model versus other models

	SETAR			**GARCH / SETAR**			**GARCH-L / SETAR**			**Random Walk / SETAR**		
Steps n.	AAD	MSE	THEIL	AAD	MSE	THEIL	AAD	MSE	THEIL	AAD	MSE	THEIL
1	0.00882	0.00015	0.70998	1.20710	1.07652	1.03755	1.20638	1.07098	1.03488	1.34506	1.33810	1.15676
2	0.00889	0.00017	0.73588	1.18774	1.37855	1.17411	1.19790	1.37654	1.17326	1.29356	1.63872	1.28012
3	0.00894	0.00017	0.74220	1.18277	1.35714	1.16496	1.17468	1.35161	1.16259	1.26632	1.68950	1.29981
4	0.00915	0.00017	0.74441	1.19078	1.33579	1.15576	1.20416	1.34553	1.15997	1.32299	1.82578	1.35121
5	0.00929	0.00018	0.75127	1.24088	1.41062	1.18769	1.23996	1.41013	1.18749	1.48043	2.04788	1.43104
6	0.00930	0.00018	0.74258	1.17897	1.34662	1.16044	1.17370	1.33631	1.15599	1.30288	1.96561	1.40200
7	0.00936	0.00018	0.74714	1.19003	1.23662	1.11203	1.18805	1.22616	1.10732	1.30970	1.84097	1.35682
8	0.00945	0.00018	0.74508	1.23268	1.23930	1.11324	1.22928	1.23171	1.10982	1.53764	2.06992	1.43872
9	0.00922	0.00018	0.74929	1.18105	1.27379	1.12862	1.17988	1.26782	1.12597	1.39179	2.13285	1.46042
10	0.00815	0.00011	0.70780	1.21981	1.43688	1.19870	1.22406	1.44278	1.20115	1.42697	2.76459	1.66270
11	0.00719	7.2E-05	0.65804	1.29031	1.70837	1.30704	1.28750	1.70334	1.30512	1.63218	3.99570	1.99892
12	0.00696	6.6E-05	0.66696	1.29720	1.69789	1.30303	1.29613	1.68838	1.29937	1.64605	4.18949	2.04682
13	0.00703	6.7E-05	0.66573	1.19796	1.41922	1.19131	1.20156	1.42469	1.19360	1.44597	3.25976	1.80548
14	0.00692	6.6E-05	0.66957	1.22793	1.46411	1.21000	1.22746	1.46363	1.20980	1.65593	3.79080	1.94699
15	0.00720	6.9E-05	0.68890	1.22772	1.51806	1.23209	1.22459	1.50721	1.22768	1.46138	3.83999	1.95959
16	0.00713	6.9E-05	0.67557	1.25986	1.51475	1.23075	1.25701	1.50936	1.22856	1.61855	4.14558	2.03607
17	0.00699	6.7E-05	0.65853	1.25250	1.44163	1.20068	1.25403	1.44519	1.20216	1.66585	4.15241	2.03774
18	0.00708	6.9E-05	0.68051	1.24520	1.43452	1.19771	1.24486	1.43176	1.19656	1.76243	4.40459	2.09871
19	0.00706	6.9E-05	0.67222	1.22789	1.34143	1.15820	1.22713	1.33829	1.15684	1.50418	3.68474	1.91956
20	0.00694	6.7E-05	0.65348	1.21685	1.37035	1.17062	1.21691	1.37068	1.17076	1.62114	4.25602	2.06301
21	0.00710	7.1E-05	0.66105	1.21891	1.32344	1.15041	1.21651	1.31857	1.14829	1.79902	4.53114	2.12864
22	0.00722	7.3E-05	0.66327	1.20056	1.31958	1.14873	1.20032	1.31478	1.14664	1.70372	4.59974	2.14470
23	0.00726	7.3E-05	0.65388	1.14358	1.23605	1.11177	1.14483	1.23580	1.11166	1.68007	4.00268	2.00067
24	0.00708	7.1E-05	0.67072	1.15090	1.26622	1.12526	1.14968	1.26408	1.12431	1.78612	4.57012	2.13778
25	0.00716	7.2E-05	0.66881	1.19870	1.34383	1.15924	1.19642	1.33732	1.15642	1.94207	5.21704	2.28408
26	0.00704	7.0E-05	0.64840	1.18102	1.27514	1.12922	1.17770	1.26861	1.12633	1.60288	4.60756	2.14652
27	0.00722	7.3E-05	0.66323	1.20871	1.30914	1.14418	1.20741	1.30621	1.14289	2.18171	6.11831	2.47352
28	0.00724	7.5E-05	0.65720	1.18420	1.27298	1.12826	1.18230	1.26805	1.12607	2.04957	6.04234	2.45811
29	0.00731	7.6E-05	0.66904	1.19853	1.26197	1.12337	1.19533	1.25659	1.12097	1.70995	5.17234	2.27427
30	0.00731	7.7E-05	0.65853	1.18955	1.20773	1.09896	1.18715	1.20513	1.09778	1.86152	5.08202	2.25433

Correct-signs percentage		SETAR	Garch	Garch-l	Rand. Walk
steps n.	1	35%	31%	33%	31%
	from 1 to 30	52%	45%	45%	

(1-step) GARCH Median square error/SETAR Median square error RATIO =1.74334587
(1-step) GARCH-L Median square error/SETAR Median square error RATIO =1.83104304

Table [14]

Back-Testing (tail-probability = 5 %)

SETAR	RiskMetrics	GARCH	GARCH-L	AR
4.9091 %	3.8182 %	3.8182 %	3.6364 %	6 %

Table [15]

		VaR					
date	Price	Historical Simulation Approach	**SETAR (2; 5, 5; 1)**	RiskMetrics	GARCH (1, 1)	GARCH-L (1, 1)	AR(5)
Mo-Oct.27	22625	-0.02131	-0.02174	**-0.02588**	**-0.026**	**-0.026**	**-0.01892**
Tu-Oct.28	**21217**	-0.02148	-0.0151	-0.02774	-0.02611	-0.02611	-0.01788
We-Oct.29	22286	-0.0215	-0.01981	**-0.03733**	**-0.04714**	**-0.04713**	**-0.02574**
Th-Oct.30	21724	-0.0215	-0.01964	**-0.04126**	**-0.04516**	**-0.04515**	**-0.02535**
Fr-Oct.31	21737	-0.02165	-0.02339	**-0.0413**	**-0.04994**	**-0.04993**	**-0.03193**

Appendix [1]

We precede that the outputs obtained in this work have required us to prepare a specific software, building the programmes in the programming language GAUSS. So, also the algorithms illustrated in the Appendices have been implemented using this language.

*** Recursive Local Fitting ***

The Recursive local fitting [Tsay, 1991] can be carried out as follows:

(1) order the observations as in an arranged autoregression and denote by "s" the new order index for arranging the data (originally it was "t", the time index);

(2) initialize the estimation procedure by fitting, via the OLS method, an AR(k) model to the first m (which is the dimension of the estimation window: Tsay usually [1991, 1993] employs about 1/6 of all the data) data cases corresponding to y_s, with s = 1, 2, ..., m;

(3) proceed with the estimation by

3a) adding the next available data case, that is y_{m+1} and

3b) deleting the first data case, the "oldest" in the rectangular window, that is y_1;

(4) repeat step 3) until all the data cases have been processed.

The recursive algorithm here employed is called RWP (Rectangularly-Weighted-Past): it consists of two steps at each iteration, the one to include the new data case, the other to delete the "oldest "one. Denote

- the vector of OLS estimates of AR coefficients by Φ_v, when the last data case in the rectangular window corresponds to y_v;
- the corresponding $(X'X)^{-1}$ matrix by P_v;
- the vector of regressors corresponding to the data case y_j by X_j.

Then, the addition of the data case corresponding to y_{v+1} can be done recursively by calculating

$\Phi^*_{v+1} = \Phi_v + P_v X_{v+1} [1 + X'_{v+1} P_v X_{v+1}]^{-1} [y_{v+1} - X'_{v+1}\Phi_v]$,

$P^*_{v+1} = P_v - P_v X_{v+1} [1 + X'_{v+1} P_v X_{v+1}]^{-1} X'_{v+1} P_v$,

where Φ^*_{v+1} e P^*_{v+1} are respectively the OLS estimates and the $(X'X)^{-1}$ matrix obtained with the m + 1 data cases corresponding to y_j, with j = v + 1 - m, ..., v + 1.

The deletion of the first data case in the rectangular window, that is the case corresponding to y_{v+1-m}, can be done by calculating

$\Phi_{v+1} = \Phi^*_{v+1} + P^*_{v+1} X_{v+1-m} [X'_{v+1-m} P^*_{v+1} X_{v+1-m} - 1]^{-1} [y_{v+1-m} - X'_{v+1-m}\Phi^*_{v+1}]$,

$P_{v+1} = P^*_{v+1} + P^*_{v+1} X_{v+1-m} [X'_{v+1-m} P^*_{v+1} X_{v+1-m} - 1]^{-1} X'_{v+1-m} P^*_{v+1}$.

Appendix [2]

*** Threshold test and General Nonlinearity test of Tsay ***

The <u>Threshold test</u> [Tsay, 1989] is based on an *arranged autoregression*, which is carried out recursively (adding one observation at each iteration, then increasing progressively the sample dimension) to calculate the so-called normalized predictive errors. It requires k (AR order) and d (delay parameter) to be known, and is carried out as follows:

(1) order the data for an *arranged autoregression*, which is in fact performed on the first rearranged m data (Tsay suggests m = (T/10)+k, where T is the sample dimension). Denote the vector of the OLS estimates of the parameters, obtained on the first m data, by b_m, the associated $(X'X)^{-1}$ matrix by P_m, and the vector of regressors corresponding to the first next observation (y_{m+1}) entering the recursive estimation at the second iteration by x_{m+1};

(2) proceed with the recursive autoregression. The recursive least squares estimates can be calculated efficiently by using the following formulas:
$b_{m+1} = b_m + k_{m+1}(y_{m+1} - x'_{m+1} b_m)$, $D_{m+1} = 1 + x'_{m+1} P_m x_{m+1}$, $k_{m+1} = P_m x_{m+1}/ D_{m+1}$, and
$P_{m+1} = \{I - P_m[(x_{m+1}\ x'_{m+1})/ D_{m+1}]\}P_m$, while the predictive errors are given by $A_{m+1} = y_{m+1} - x'_{m+1} b_m$, and the normalized predictive errors by $e_{m+1} = a_{m+1} / \sqrt{D_{m+1}}$;

(3) regress, via OLS, the normalized predictive errors so obtained e_v on the corresponding y_{v-i} and the constant (v = m + 1, ..., T - d - h + 1, h = max{1, k-d+1) and i = 1, ..., k), and save the residuals ε_v;

(4) calculate the test statistic $F = (\Sigma\ e^2 - \Sigma\ \varepsilon^2)\times (T\text{-}d\text{-}m\text{-}k\text{-}h)/ (\Sigma\ \varepsilon^2)\times (k+1)$, where the summation index is v = m+1, ..., T-d-h+1. F follows approximatively an F-distribution, with (k+1) and (T-d- m - k - h) degrees of freedom.

The <u>General Nonlinearity test</u> of Tsay [1991] is identical with the Threshold test as concerns the first 2 steps, but then requires the following further steps:

(3) regress the normalized predictive errors e_v on

- the corresponding y_{v-i} and the constant (v = m + 1, ..., T - d - h + 1 and i = 1, ..., k);
- $(y_{v-i}\, e_{v-i}, e_{v-i}\, e_{v-i-1})$ for i = 1, ..., k;
- $\{y_{v-1} \exp(-y_{v-1}/\gamma), G(z_{v-d}), y_{v-1} G_{t-d}\}$, where γ is a normalization constant, say $\gamma = \max\{|y_{t-1}|\}$, y_t being the series of examined data, $z_{t-d} = (y_{t-d} - My_d)/ S_d$, with My_d and S_d sample mean and sample standard deviation of y_{t-d}, respectively, and $G(\cdot)$ the cumulative distribution function of the standardized normal random variable;

(4) save the estimated residuals ε_v obtained at step 3, and calculate the test statistic, which is identical with the Threshold test, but (k+1) and (T-d-m-k- h) are substituted by 3(k+1) and [T-m-3(k+1)].

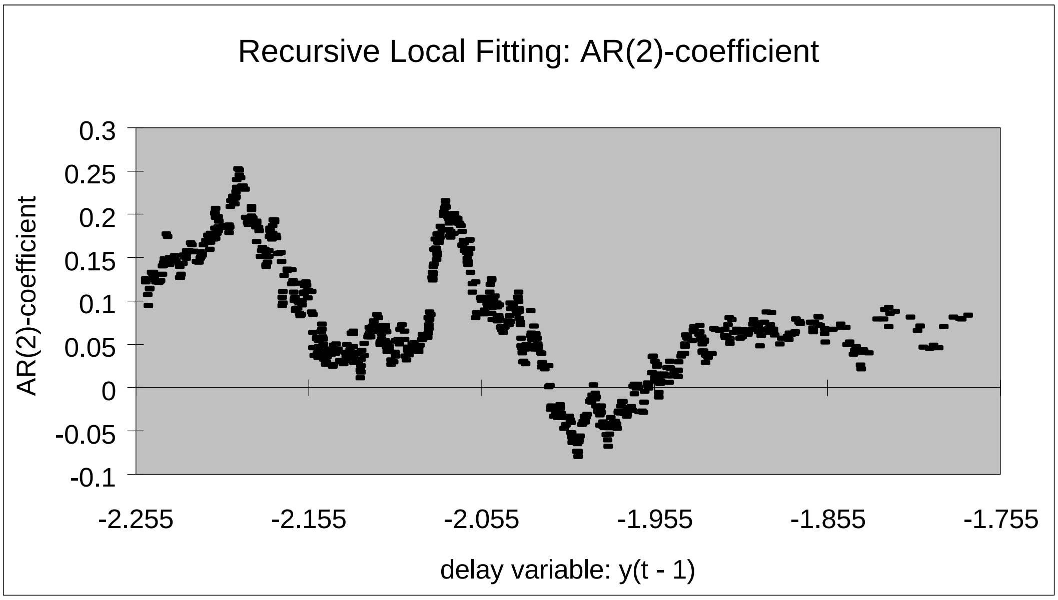


FIGURE [1a]

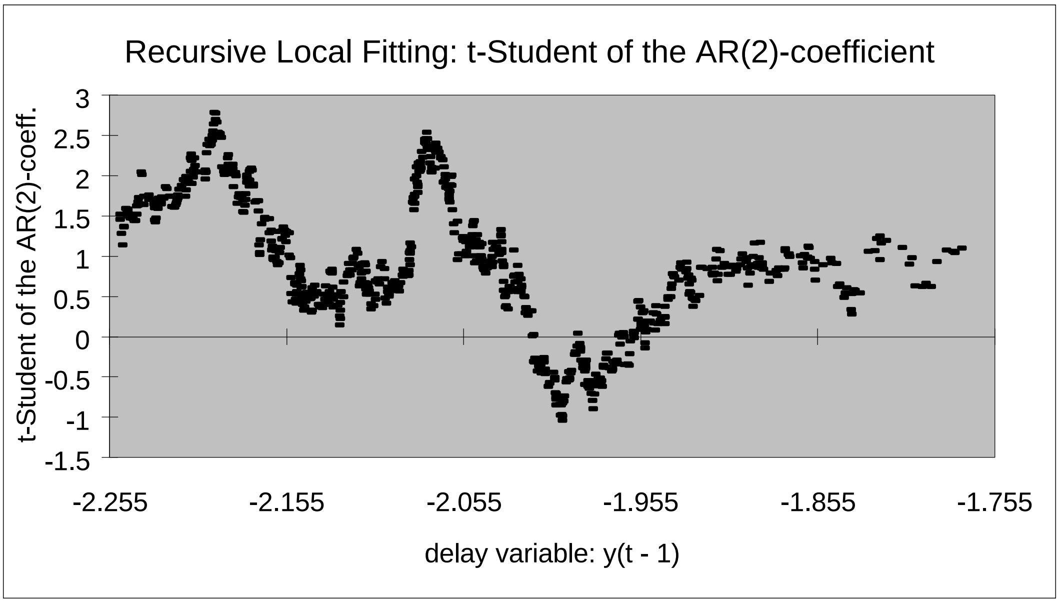


FIGURE [1b]

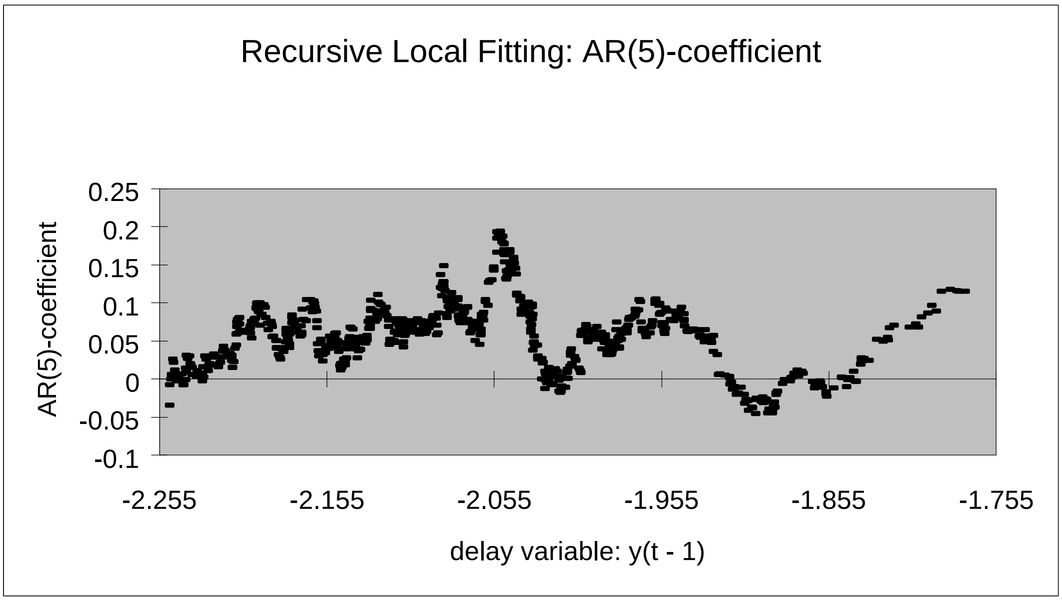


FIGURE [2a]

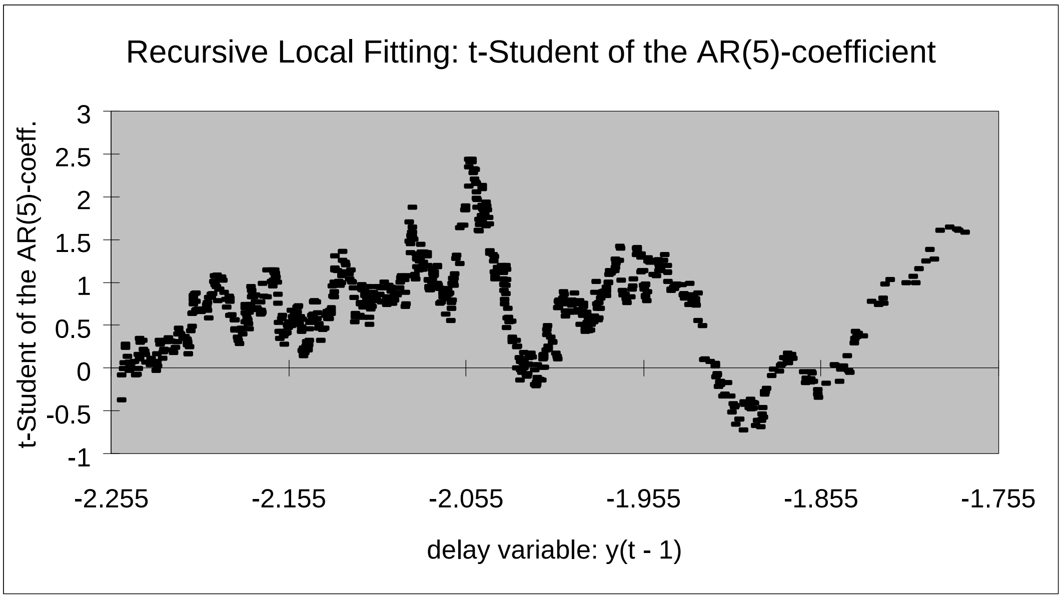
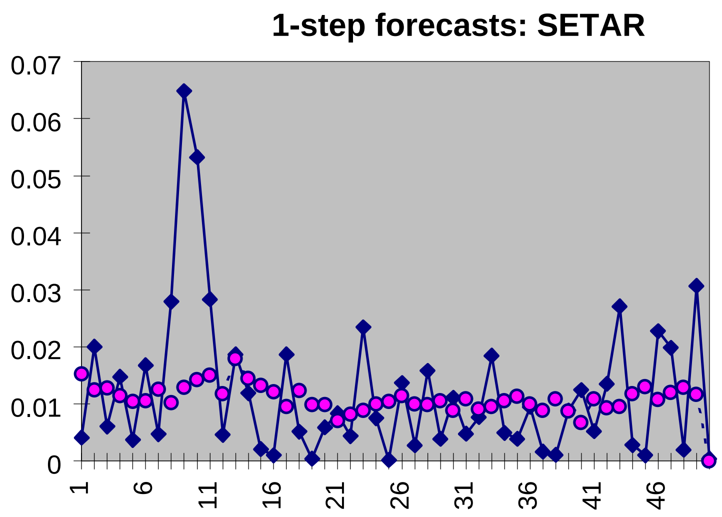
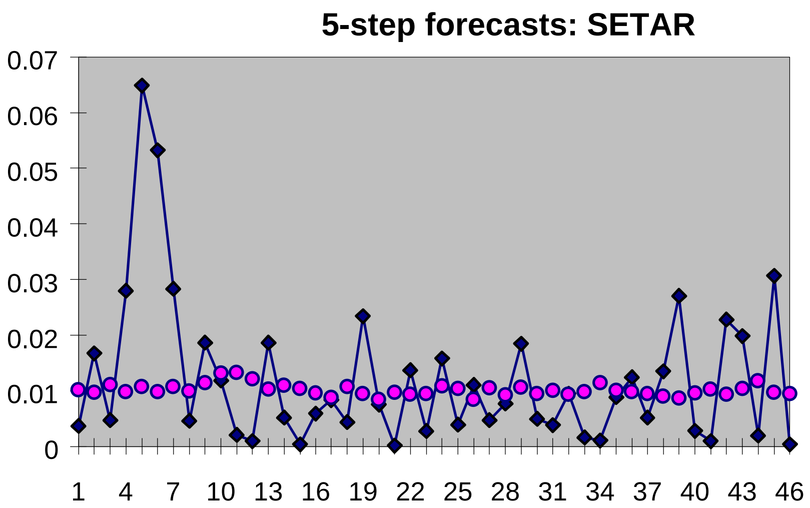


FIGURE [2b]



Rhombus points denote true values, round points denote forecasted values.

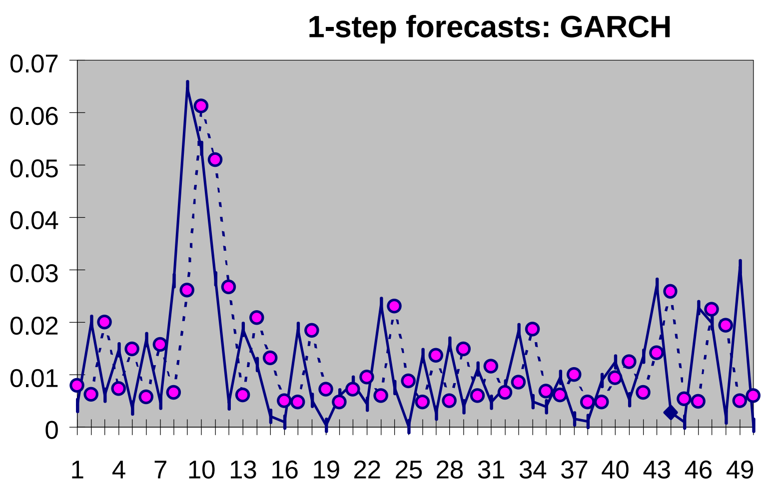
Figure [3a]



Rhombus points denote true values, round points denote forecasted values.

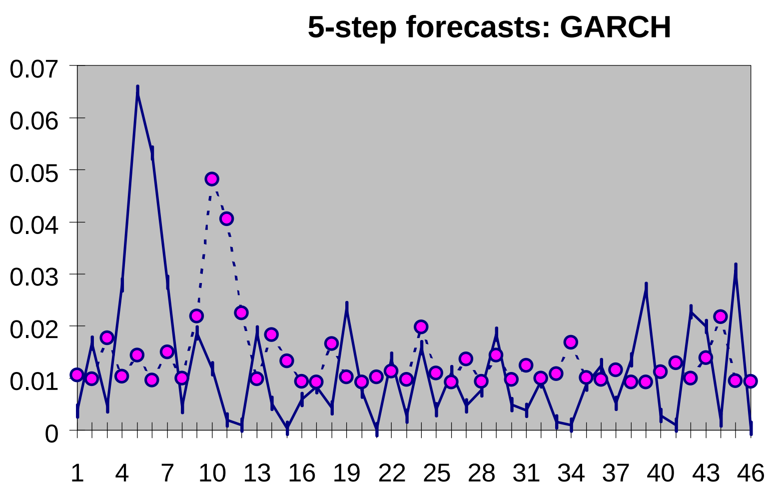
Figure [3b]

The pick denotes the slump of Tuesday October 28, 1997



Rhombus points denote true values, round points denote forecasted values.

Figure [3c]



Rhombus points denote true values, round points denote forecasted values.

Figure [3d]

The pick denotes the slump of Tuesday October 28, 1997.