Modeling and Forecasting Outliers and Level Shifts in Absolute Returns

Philip Hans Franses*
Econometric Institute
Erasmus University Rotterdam

Marco van der Leij
Econometric Institute
Erasmus University Rotterdam

Richard Paap
Rotterdam Institute for Business Economic Studies
Erasmus University Rotterdam

Econometric Institute Report 2001-34

Abstract

Due to high and low volatility periods, time series of absolute returns experience temporary level shifts (that is, periods with outliers) which differ in length and size. In this paper we put forward a new model which can describe and forecast the location and size of such level shifts. Our so-called Switching Regime Censored Latent Effects Autoregression [SR-CLEAR] assumes that technical trading rules may have explanatory value for future volatility. It is assumed that these rules have a time-varying effect on absolute returns, and that this effect appears as an outlier or a level shift. We apply the SR-CLEAR model to nine stock markets and we document its excellent fit and competitive forecasting ability.

Key words: Absolute returns, Outliers, Temporary level shifts, Censored latent effects

---

*This paper has been prepared for the International Conference on Modeling and Forecasting Financial Volatility, September 7-9, 2001, in Perth Australia. All computations in the paper are performed in Ox 2.20 (see Doornik, 1999). The data and programs used in this paper can be obtained from the corresponding author. The third author thanks the Netherlands Organization for Scientific Research (N.W.O.) for its financial support. Address for correspondence: Philip Hans Franses, Econometric Institute H11-34, Erasmus University Rotterdam, P.O.Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: franses@few.eur.nl

1 Introduction

This paper aims at modelling an important feature of absolute stock market returns, where this variable can be viewed as a measure of volatility. A typical feature of absolute returns, at least when a sample of daily data over several years in considered, is that there are periods with large absolute returns, usually associated with high volatility, and periods with small absolute returns and hence low volatility. These periods are not equally long and the level of absolute returns in these periods also varies in size. Hence, the absolute returns seem to experience temporary level shifts which are unequal in length and size. If a temporary level shift lasts for only one observation, it is usually called an outlier. If it lasts for more periods, one can also view a temporary level shift as a sequence of outliers.

It is by now well known that occasional level shifts suggest the possible empirical adequacy of so-called long memory models, see Granger and Hyung (2000), Diebold and Inoue (2001), Bos, Franses and Ooms (1999) and Hyung and Franses (2001), to mention a few references. An example of such a long memory model is proposed in Ding, Granger and Engle (1993). It should be stressed though that a long memory model tries to capture the specific autoregressive dynamics of a time series with occasional level shifts, but that it does not seek to predict the size and relevance of the level shifts themselves. For financial investors, however, it is important to anticipate a period of high level absolute returns, as changes in risk premiums of assets can occur. Hence, investors would benefit from the opportunity to predict a temporary level shift in absolute returns. In this paper we therefore focus on the level shifts instead of on the autoregressive behaviour of absolute returns.

The model we put forward in this paper has the following features. First of all, it aims to describe and forecast outliers and temporary level shifts using predetermined variables. It seems that certain technical trading rules, which usually involve differences between short-run and long-run moving averages of the levels of the stock market, can have predictive power, see Brock, Lakonishok and LeBaron (1992), Gençay (1996) and Franses and van Griensven (1998), to mention just a few studies. When the short-run average rises above the long-run average, one may expect a period of positive returns and the market is going upwards. In contrast, when the short-run average falls below the long-run average, this may be viewed as a prelude to a negative correction, and hence negative returns.

A second feature of our model is that it allows the predetermined trading rules to have a different impact on volatility in times when the stock market goes up then when it goes down. This feature matches with the leverage effect, which entails that volatility increases particularly when stock prices fall, and less in case prices go up, see for example, Braun, Nelson and Sunier (1995), Christie (1982), Glosten, Jagannathan and Runkle (1992), Schwert (1989), among others. For that reason, we allow the model structure and the parameters to differ across two regimes.

A third and rather important feature of our model, which in fact is the main novelty when compared with alternative models, is that it introduces additional uncertainty as to whether the predetermined variable has a relevant impact or not. The idea is that we assume that this variable enters a censored regression for outliers, and only when a stochastic threshold is exceeded, there is a positive contribution to volatility through these outliers. This feature may be relevant in practice as it assumes that the predetermined variable does not always have an effect on absolute returns and that the size of its effect also does not have to be constant over time. To this end, we extend the so-called Censored Latent Effects Autoregression [CLEAR], put forward in Franses and Paap (2001). An important feature of this model is that it allows for time-varying effects of variables without the cost of many additional parameters.

The outline of our paper is as follows. In Section 2, we discuss the autoregressive behaviour of absolute returns, and the ARFIMA model that is frequently used to describe this behaviour. In Section 3, we propose our new model, the Switching Regime CLEAR model [SR-CLEAR]. We discuss its representation and interpretation. We consider maximum likelihood [ML] parameter estimation, and we examine through simulations if data generated from an SR-CLEAR model have features that can be picked up by an ARFIMA model. In Section 4, we consider our new model for nine stock markets, for which we an-

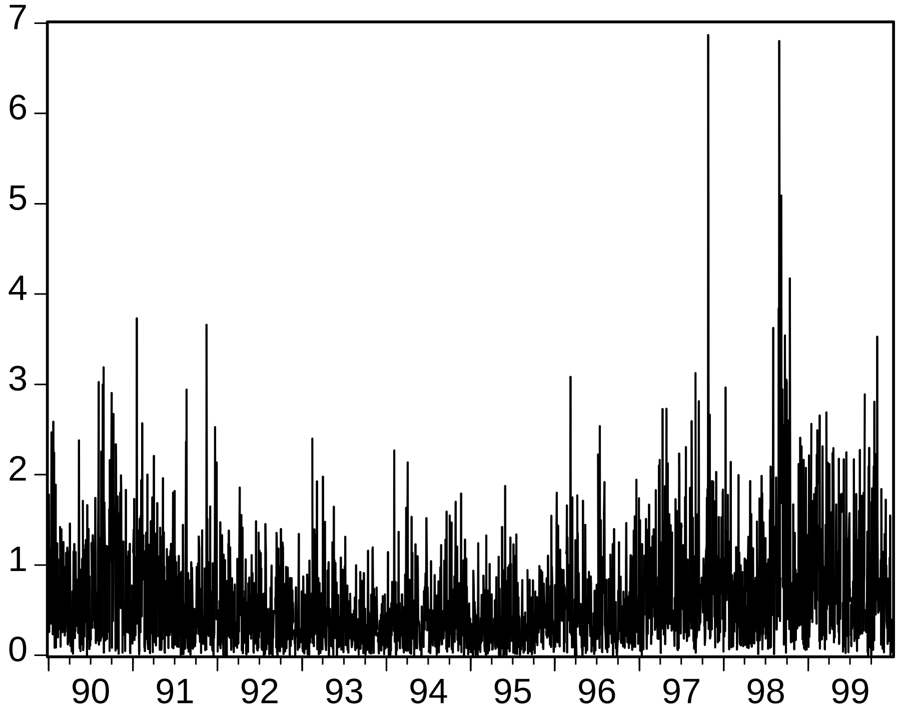

Figure 1: Absolute returns on the S&P 500: 1990-1999

alyze daily data covering 1990-1999. We find that the model fits the data quite well. In Section 5, we use the model for forecasting the daily absolute returns in 2000, and we show that we obtain substantial improvement over alternative models. Finally, in Section 6, we conclude with a discussion of topics for further research.

2 On Modeling Absolute Returns

Consider a daily observed stock market index z_t, and consider the daily absolute returns y_t, that is,

$$y_t = \left| \frac{z_t - z_{t-1}}{z_{t-1}} \right| . \tag{1}$$

A typical graph of y_t is given in Figure 1, which shows the daily absolute returns of the S&P 500 index for 1990-1999. We notice periods where absolute returns temporarily stay at a high level, like during 1997, 1998 and 1999. There are also periods of low level

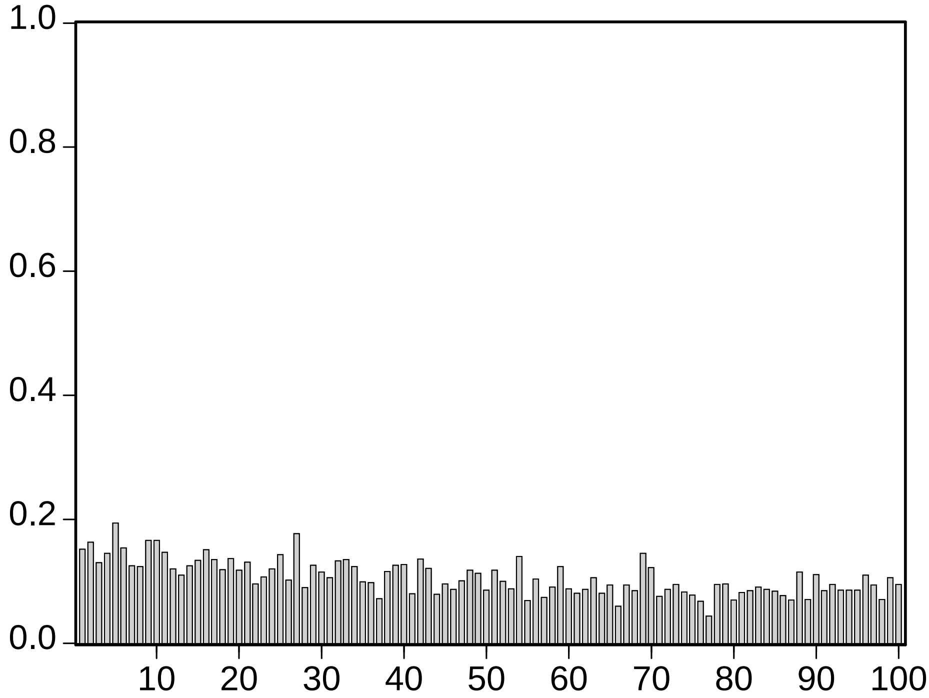

Figure 2: Autocorrelation function of the absolute returns on the S&P 500: 1990-1999

absolute returns, like around 1995. Hence, absolute returns seem to show temporary level shifts.

It has now been well established that such level shifts may suggest long-memory behavior, see Granger and Hyung (2000) and Diebold and Inoue (2001) among others. Indeed, if we look at the estimated autocorrelation function in Figure 2 for the S&P 500 absolute returns, it is evident that the autocorrelation function decays very slowly. This behaviour of the autocorrelation function is present in almost every series of absolute returns, and it is known as long memory. Indeed, long-memory models have frequently been used for measures of financial volatility. As these models may be viewed as a benchmark, we start our discussion with a brief treatment of the ARFIMA model, and next we extend the ARFIMA model to capture temporary level shifts.

2.1 Long-memory model

The ARFIMA(p,d,q) model is widely used to describe series with long memory. The model can be represented by

$$\Phi(L)(1-L)^d(y_t-\mu)=\Theta(L)\varepsilon_t, \tag{2}$$

where ε_t is $N(0,\sigma_\varepsilon^2)$, $t=1,\ldots,T$. $\Phi(L)=(1-\phi_1 L-\cdots-\phi_p L^p)$ and $\Theta(L)=(1+\theta_1 L+\cdots+\theta_q L^q)$ are polynomials of finite orders p and q in the usual lag operator L. For any real $d>-1$, the fractional difference operator $(1-L)^d$ is defined by its Taylor-Maclaurin series, that is,

$$(1-L)^d=\sum_{j=0}^{\infty}\pi_j L^j,\ \ \pi_0=1,\ \ \pi_j=\frac{j-1-d}{j}\pi_{j-1}, \tag{3}$$

which implies that $(1-L)^d$ can be written as an infinite AR polynomial. As $\left|\frac{\pi_j}{\pi_{j-1}}\right|<1$, it is clear that $\lim_{j\to\infty}\pi_j=0$, but that π_j decays only very slowly. Beran (1994) shows that the autocorrelation function implied by (2) decays at a hyperbolic rate. When $|d|<0.5$, and $\Phi(L)$ and $\Theta(L)$ do not contain a unit root, a time series generated by an ARFIMA(p,d,q) model is stationary, and the model parameters can be estimated by maximum likelihood [ML].

Table 1 gives parameter estimates of an ARFIMA$(5,d,0)$ model fitted to the daily S&P 500 absolute returns (1990-1999), where the AR order is chosen such that the estimated residuals are approximately white noise. Maximum likelihood estimation is performed using ARFIMA package 1.01 (Doornik and Ooms, 2001). The maximum likelihood estimate of d is 0.33, and it appears to be different from zero.

2.2 Technical trading rule as an explanatory variable

The ARFIMA(p,d,q) model only uses information from the lagged values of the dependent variable. We might extend this model by including a predetermined explanatory variable x_t. Such a variable might be the ratio between a short-run moving average and a long-run moving average of stock prices. Several studies find that a technical trading rule based

Table 1: Maximum Likelihood Estimation of an ARFIMA$(5, d, 0)$ model, applied to absolute returns on the S&P 500 index[1]

	Parameter	Standard Error
d	0.330	0.033
ϕ_1	-0.287	0.039
ϕ_2	-0.147	0.033
ϕ_3	-0.120	0.029
ϕ_4	-0.039	0.026
ϕ_5	0.019	0.022
μ	0.667	0.111

[1] The model is given in (2). The sample size is 2527 observations.

on such a ratio generate higher profits than a buy-and-hold strategy, see for example Brock, Lakonishok and LeBaron (1992). When the short-run moving average exceeds that of the long-run, the market is going upwards and positive returns are to be expected. When the short-run moving average falls below the long-run moving average, this can be viewed as a prelude to a declining stock market. Hence, a variable constructed from a short-run moving average and a long-run moving average might be a leading indicator of fluctuations of the stock market, and hence of absolute returns.

More specifically, consider the moving average of stock prices over k days, that is,

$$\overline{z}_{k,t} = \frac{1}{k} \sum_{i=t-k+1}^{t} z_i.$$

The ratio of the moving average of stock prices of k_1 days over that of k_2 days is defined as

$$x_t = \frac{\overline{z}_{k_1,t} - \overline{z}_{k_2,t}}{\overline{z}_{k_2,t}}, \tag{4}$$

where usually $0 < k_1 < k_2$. Typically, in practice one takes k_1 to be equal to 1, 2 or 5, and k_2 equal to 50, 100 or 200.

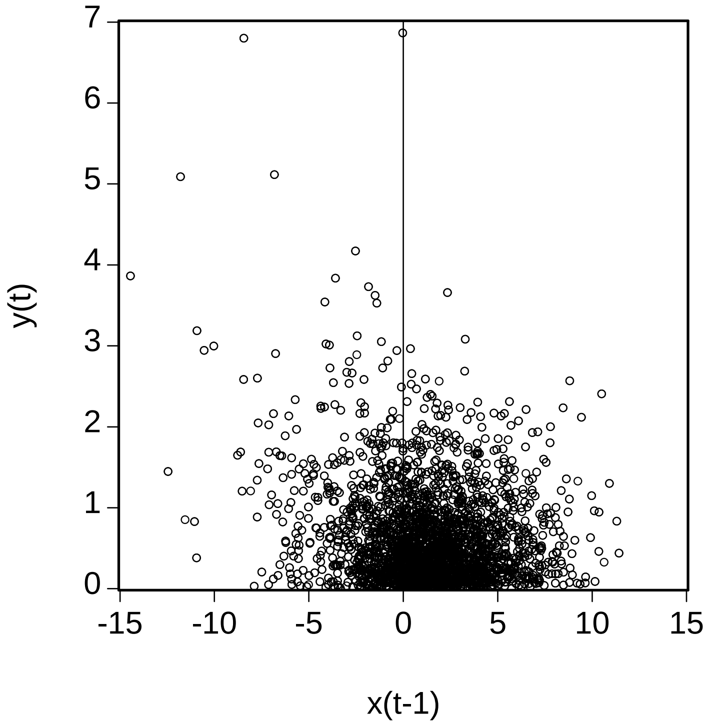

Figure 3: Absolute returns on the S&P 500 index (1990-1999) against the ratio of moving averages of 1 day over 50 days.

Figure 3 shows a scatterplot of the absolute returns on the S&P 500 index (y_t), against the one day lagged value of the ratio between the 1-day moving average and the 50-day moving average (x_{t-1}), for 1990-1999. This graph seems to suggest a relation between absolute returns, and the lagged ratio of moving averages, where higher absolute returns correspond with more negative values of x_{t-1}. This suggests that x_{t-1} has power in explaining part of the y_t. However, the relation between y_t and x_{t-1} does not seem to be linear. There does not seem to be an obvious relation between y_t and x_{t-1} when x_{t-1} is positive. Hence, Figure 3 suggests perhaps an asymmetric and non-linear relation between x_{t-1} and y_t. When stock prices are relatively low, that is when $x_{t-1} < 0$, the corresponding absolute returns are relatively high, and this corresponds with the leverage effect.

One may now want to include a leverage effect in an ARFIMA model. One way to do this is to allow for asymmetric effects of x_{t-1} in an ARFIMAX model, that is, to consider

$$\Phi(L)(1-L)^d(y_t - (\mu_1 + \pi_1 x_{t-1})I_{x_{t-1}>0} - (\mu_2 + \pi_2 x_{t-1})(1 - I_{x_{t-1}>0})) = \Theta(L)\varepsilon_t, \qquad (5)$$

where $I_{x_{t-1}>0}$ is an indicator function, which takes a value 1 when the argument is true and 0 otherwise. Hence, this model captures x_{t-1} in the conditional mean, but allows its effect to depend on its sign. In a sense, this model allows for a time-varying conditional mean, where its value depends on x_{t-1}. In Section 5, we will compare this model with our new model, which is discussed in the next section.

3 Switching-Regime CLEAR Model

In this section we propose a new nonlinear time series model for describing potentially time-varying effects of a predetermined variable on daily absolute returns. Due to these time-varying effects the model should be able to describe and forecast occasional outliers and level shifts. This is in contrast to an ARFIMA model, which focuses only on the autoregressive structure of absolute returns, and an ARFIMAX model which imposes that x_{t-1} always has an effect.

In this section we first discuss its representation. The model does not nest the above long-memory model, but we will show that it can capture long-memory features. For that matter, we generate artificial data for this model, in order to see if it really generates long-memory properties. Next, we discuss estimation by ML and we show through Monte Carlo simulations that this method is quite reliable. We then turn to the construction of residuals, to inference on latent variables and, finally, the construction of out-of-sample forecasts.

3.1 Representation

To allow for time-varying effects of x_{t-1} and also to allow for regime-dependent effects, we propose to consider a Switching Regime CLEAR model. This model builds on the CLEAR model of Franses and Paap (2001), but it additionally introduces two regimes in the model.

Consider first the CLEAR(p) model of Franses and Paap (2001), which for absolute returns could look like

$$y_t = \mu + \rho_1 y_{t-1} + \ldots + \rho_p y_{t-p} + \pi |x_{t-1}| + v_t + \varepsilon_t, \tag{6}$$

where

$$v_t = \begin{cases} \alpha + \beta |x_{t-1}| + u_t & \text{if } \alpha + \beta |x_{t-1}| + u_t > 0 \\ 0 & \text{if } \alpha + \beta |x_{t-1}| + u_t \leq 0, \end{cases} \tag{7}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $u_t \sim N(0, \sigma_u^2)$ for $i = 1, 2$ and where we assume that ε_t and u_t are mutually uncorrelated. Equation (6) consists of a standard ARX(p) model with an additional censored latent variable v_t. From (7) this variable v_t has a conditional censored normal distribution. Hence, the CLEAR(p) model is essentially a combination of an ARX(p) model and a standard Tobit model.

Franses and Paap (2001) apply this model to forecast US unemployment, and find that the model outperforms other non-linear models. A major advantage of the model is that the censored model for v_t introduces the possibility that the effect of $|x_{t-1}|$ varies over time, depending on whether a *stochastic* threshold is exceeded. To see this, notice that according to (7) v_t affects y_t only when $\alpha + \beta |x_{t-1}| + u_t > 0$. If $\beta \neq 0$, and if we define $\tau \equiv -\alpha/\beta$ and $\eta_t \equiv -(1/\beta) u_t$, we can rewrite this condition as:

$$v_t > 0 \text{ if } \begin{cases} |x_{t-1}| > \tau + \eta_t & \text{if } \beta > 0 \\ |x_{t-1}| < \tau + \eta_t & \text{if } \beta < 0. \end{cases} \tag{8}$$

Hence, depending on the sign of β, v_t is only positive when $|x_{t-1}|$ exceeds or falls below a stochastic threshold. If we assume that $\beta > 0$, we can combine (8) with (6) and (7) to obtain

$$y_t = \begin{cases} \mu + \rho_1 y_{t-1} + \ldots + \rho_p y_{t-p} + \pi |x_{t-1}| + \varepsilon_t & \text{if } |x_{t-1}| \leq \tau + \eta_t \\ (\mu + \alpha) + \rho_1 y_{t-1} + \ldots + \rho_p y_{t-p} + (\pi + \beta) |x_{t-1}| + \varepsilon_t & \text{if } |x_{t-1}| > \tau + \eta_t, \end{cases} \tag{9}$$

Equation (9) shows that the CLEAR model is closely related to the Smooth Transition Regression model of Teräsvirta (1994). Hence, the influence of $|x_{t-1}|$ on y_t varies over time. However, due to the stochastic element η_t in the threshold, it might occur that even though $|x_{t-1}| \gg \tau$, $v_t = 0$.

The v_t variable can be interpreted as an innovation outlier at time t, and, depending on the parameters, it may also give a sequence of outliers. If $|x_{t-1}|$ exceeds the stochastic threshold $\tau + \eta_t$, then $v_t > 0$, and this will put y_t at a higher level. Therefore, this model might capture and forecast innovation outliers and temporary level shifts in absolute returns.

Although the CLEAR(p) might perform well, the model does not capture the feature of asymmetry in absolute returns as discussed earlier. We therefore introduce the SR-CLEAR(p) model, which has the following representation,

$$y_t = \begin{cases} \mu_1 + \rho_{1,1} y_{t-1} + \ldots + \rho_{1,p} y_{t-p} + \pi_1 x_{t-1} + v_{1,t} + \varepsilon_{1,t} & \text{if } x_{t-1} > 0 \\ \mu_2 + \rho_{2,1} y_{t-1} + \ldots + \rho_{2,p} y_{t-p} + \pi_2 x_{t-1} + v_{2,t} + \varepsilon_{2,t} & \text{if } x_{t-1} \leq 0, \end{cases} \tag{10}$$

where

$$v_{1,t} = \begin{cases} \alpha_1 + \beta_1 x_{t-1} + u_{1,t} & \text{if } \alpha_1 + \beta_1 x_{t-1} + u_{1,t} > 0 \\ 0 & \text{if } \alpha_1 + \beta_1 x_{t-1} + u_{1,t} \leq 0 \end{cases} \tag{11}$$

and

$$v_{2,t} = \begin{cases} \alpha_2 + \beta_2 x_{t-1} + u_{2,t} & \text{if } \alpha_2 + \beta_2 x_{t-1} + u_{2,t} > 0 \\ 0 & \text{if } \alpha_2 + \beta_2 x_{t-1} + u_{2,t} \leq 0 \end{cases} \tag{12}$$

where $\varepsilon_{i,t} \sim N(0, \sigma^2_{\varepsilon_i})$ and $u_{i,t} \sim N(0, \sigma^2_{u_i})$ for $i = 1, 2$ and where we assume that $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are mutually uncorrelated with $u_{1,t}$ and $u_{2,t}$. One might expect that β_1 is positive and β_2 is negative.

The SR-CLEAR model assumes different parameters in different regimes, depending on the sign of x_{t-1}. Hence, there are different variances in the upward and downward regimes, that is $\sigma^2_{\varepsilon_1}$ and $\sigma^2_{\varepsilon_2}$. It might perhaps be expected that uncertainty of the descriptive quality of the model is higher in cases the stock market goes down, that is, $\sigma^2_{\varepsilon_1} < \sigma^2_{\varepsilon_2}$. The model also allows for different variances of the error terms in the censored regression models for $v_{1,t}$ and $v_{1,t}$, thereby allowing for the possibility that there is perhaps more uncertainty about the size of the innovation outliers in the downward regimes, that is, one might conjecture that $\sigma^2_{u_2} > \sigma^2_{u_1}$. Notice that in both cases the innovation outliers have a positive contribution to volatility. When $|\beta_1|$ is smaller that $|\beta_2|$, one can say that negative returns have more impact on volatility than positive returns have, which can be associated with the leverage effect.

3.2 Data properties

To get an impression of this model, and the data it can describe, we generate 100 times 1000 observations on a simulated stock market index, z_t, absolute returns on that index, y_t, and the ratio between short-run moving average and long-run moving average, x_t, where we use 1 day against 200 days.

The data generating process is not straightforward. Notice that x_t follows from the values of $z_{t-199}, \ldots, z_t$, and that x_t implies y_{t+1} using equations (10), (11) and (12) of the SR-CLEAR model. By assigning either a negative sign or a positive sign to y_{t+1} we are able to compute z_{t+1}, and hence we are able to compute x_t, y_t and z_t recursively.

Based on estimation results reported below, we generate the observations from the following data generating process. We first generate starting values of the stock market index $z_{-199}, \ldots, z_0$, assuming that the percentage returns of these starting values are white noise and standard normally distributed, that is,

$$z_{-199} = 100, \quad z_t = z_{t-1}(1 + \eta_t/100), \quad \forall\, t = -198, \ldots, 0, \tag{13}$$

where $\eta_t \sim N(0,1)$. Next, we generate x_t, y_{t+1}, z_{t+1} for $t = 0, \ldots, 999$, recursively. The explanatory variable x_t is computed as:

$$x_t = \frac{\overline{z}_{1,t} - \overline{z}_{200,t}}{\overline{z}_{200,t}} \tag{14}$$

We compute the dependent variable y_t using

$$y_t = \begin{cases} 0.2 + 0.3y_{t-1} + 0x_{t-1} + v_{1,t} + \varepsilon_{1,t} & \text{if } x_{t-1} > 0 \\ 0.4 + 0.1y_{t-1} + 0.01x_{t-1} + v_{2,t} + \varepsilon_{2,t} & \text{if } x_{t-1} \leq 0, \end{cases} \tag{15}$$

with

$$v_{1,t} = \begin{cases} 0.1 + 0.02x_{t-1} + u_{1,t} & \text{if } 0.1 + 0.02x_{t-1} + u_{1,t} > 0 \\ 0 & \text{if } 0.1 + 0.02x_{t-1} + u_{1,t} \leq 0 \end{cases} \tag{16}$$

and

$$v_{2,t} = \begin{cases} -0.8 - 0.1x_{t-1} + u_{2,t} & \text{if } -0.8 - 0.1x_{t-1} + u_{2,t} > 0 \\ 0 & \text{if } -0.8 - 0.1x_{t-1} + u_{2,t} \leq 0, \end{cases} \tag{17}$$

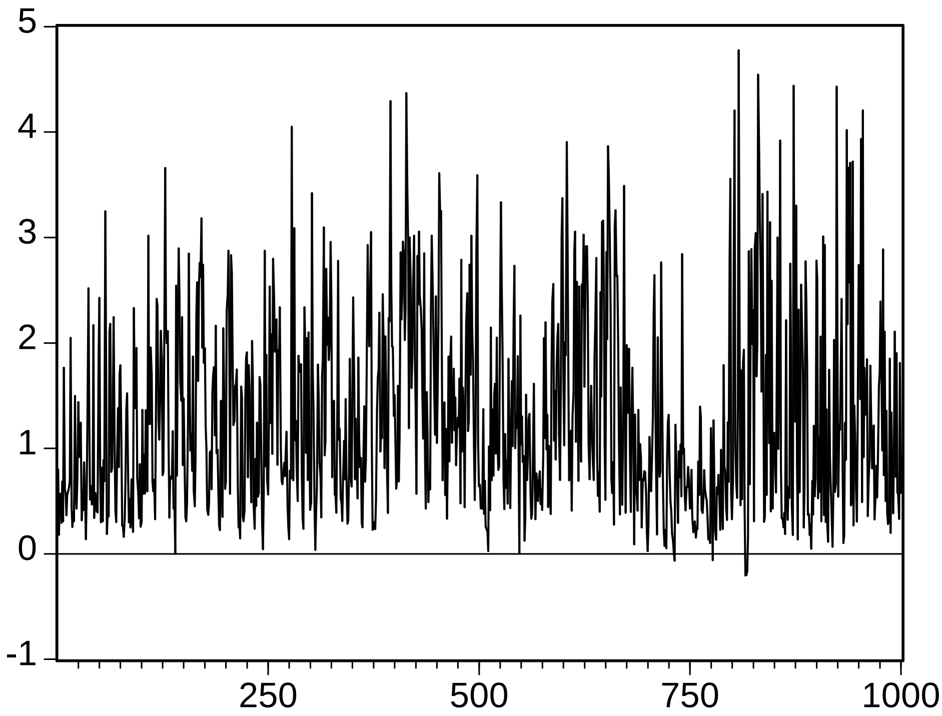

Figure 4: Simulated time series from the SR-CLEAR model (15)-(17)

where $\varepsilon_{1,t} \sim N(0, 0.02)$, $\varepsilon_{2,t} \sim N(0, 0.05)$, $u_{1,t} \sim N(0, 0.8)$ and $u_{2,t} \sim N(0, 2)$. The chosen parameters are similar to parameter estimates that are typically found in practice, see the next section.

Finally, we compute z_t by transforming the absolute return y_t into a positive or negative return, and appending it to the index. Therefore, we consider

$$z_t = z_{t-1}(1 + (2p_t - 1)y_t) \tag{18}$$

where $p_t \sim BIN(1, 0.55)$, with BIN denoting a binomial distribution.

Figure 4 shows a graph of the absolute returns of one of the replications. Notice that the simulated absolute returns are similar to the absolute returns on the S&P 500 in Figure 1, although there are more extreme outliers in the actual S&P 500 time series. The simulation also produces a few 'negative' absolute returns. Although the SR-CLEAR model has a positive censored effect, the generated observation sometimes can be negative due to exceptionally large draws of $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$.

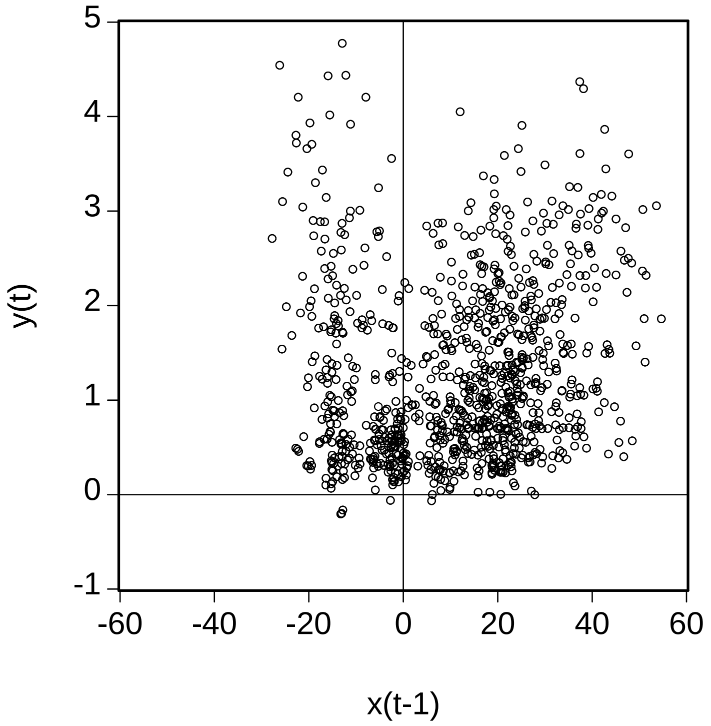

Figure 5: Simulated absolute returns against a ratio of moving averages from the SR-CLEAR model (15)-(17)

However, the simulated time series of absolute returns seems to resemble an actual time series of absolute returns quite well. This can be observed from periods of high values, for example between $t = 400$ and $t = 500$, and periods of low values, between $t = 700$ and $t = 800$. Figure 5 shows the accompanying scatter plot of y_t and x_{t-1}. This graph seems rather similar to that in Figure 3. Hence the simulated time series seems to resemble the relation between x_{t-1} and y_t quite well. Finally, Figure 6 shows the estimated autocorrelation function of the simulation, averaged over 100 replications, together with the theoretical ACF of an ARFIMA$(0, d, 0)$ proces, with $d = 0.3$. The SR-CLEAR model shows autocorrelation at high lags. However, the ACF of the SR-CLEAR model decays at a slightly faster rate than the ACF of this particular ARFIMA$(0, d, 0)$ model.

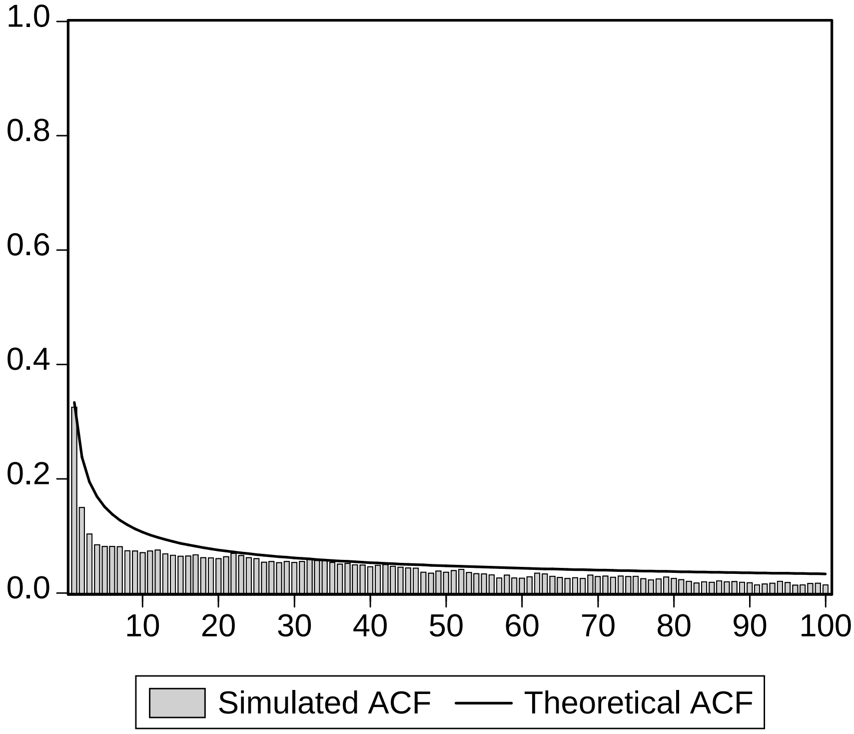

Figure 6: Estimated autocorrelation function of simulated SR-CLEAR data (average over 100 replications), and the theoretical autocorrelation function of an ARFIMA$(0, 0.3, 0)$ process.

3.3 Estimation by Maximum Likelihood

The parameters of the SR-CLEAR model can be estimated by maximizing the log-likelihood function. The log-likelihood function can be written as the sum of the log conditional densities of y_t given its past and given x_{t-1}, that is,

$$\ell(Y_T|X_{T-1};\theta) = \sum_{t=1}^{T} \ln(f(y_t|Y_{t-1}, x_{t-1};\theta)), \tag{19}$$

where $Y_T = (y_T, \dots, y_1)$ and $X_{T-1} = (x_{T-1}, \dots, x_0)$, and where $\theta = (\theta_1, \theta_2)$ collects the parameters in the two regimes. We denote regime 1, when $x_{t-1} > 0$, and regime 2 otherwise.

As we have a switching regime model, the sign of the switching variable x_{t-1} determines whether the log conditional density of y_t depends on θ_1 or θ_2. Hence we can split the log likelihood function into

$$\ell(Y_T|X_T;\theta) = \sum_{t=1}^{T} \ln(f_1(y_t|Y_{t-1}, x_{t-1};\theta_1))I_{x_{t-1}>0} + \sum_{t=1}^{T} \ln(f_2(y_t|Y_{t-1}, x_{t-1};\theta_2))(1 - I_{x_{t-1}>0}), \tag{20}$$

where $f_i(y_t|Y_{t-1}, x_{t-1};\theta_i)$ for $i = 1, 2$ is the density function of the CLEAR model in Franses and Paap (2001). Equation (20) states that we can maximize the log-likelihood of the whole sample, by maximizing the log-likelihood in the separate regimes, and adding them together. Hence, we can simply estimate the SR-CLEAR model by estimating two CLEAR models for the two regimes.

The conditional density function and the log-likelihood function of the CLEAR model are derived in Franses and Paap (2001). In regime i, the (marginal) conditional density function of y_t follows from the joint density function of $v_{i,t}$ and y_t given its past and x_{t-1}, $f_i(y_t, v_{i,t}|Y_{t-1}, x_{t-1};\theta_i)$. As $v_{i,t}$ has a censored normal distribution, the density function of y_t given its past and x_{t-1} in regime i is

$$f_i(y_t|Y_{t-1}, x_{t-1};\theta_i) = \Pr[v_{i,t} = 0|x_{t-1};\theta_i] f_i(y_t|Y_{t-1}, x_{t-1}, v_{i,t};\theta_i)|_{v_{i,t}=0} + \int_{-\alpha_i-\beta_i x_{t-1}}^{\infty} \frac{1}{\sigma_{u_i}} \phi\left(\frac{u_{i,t}}{\sigma_{u_i}}\right) f_i(y_t|Y_{t-1}, x_{t-1}, v_{i,t};\theta_i)|_{v_{i,t}=\alpha_i+\beta_i x_{t-1}+u_{i,t}} du_{i,t}, \tag{21}$$

where $\phi(\cdot)$ is the probability density function of a standard normal distribution. From the properties of the censored normal distribution it follows that

$$\Pr[v_{i,t} = 0|x_{t-1}; \theta_i] = \Phi\left(\frac{-\alpha_i - \beta_i x_{t-1}}{\sigma_{u_i}}\right), \tag{22}$$

for $i = 1, 2$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution. The density function of y_t conditional on its past, x_{t-1} and $v_{i,t}$ in regime i is

$$f_i(y_t|Y_{t-1}, x_{t-1}, v_{i,t}; \theta_i) = \frac{1}{\sigma_{\varepsilon_i}}\phi\left(\frac{y_t - \mu_i - \sum_{i=1}^{p}\rho_{i,p}y_{t-i} - \pi_i x_{t-1} - v_{i,t}}{\sigma_{\varepsilon_i}}\right). \tag{23}$$

Inserting (23) and (22) into (21) gives us the density function of y_t given its past and x_{t-1}, that is,

$$\begin{aligned} f_i(y_t|Y_{t-1}, x_{t-1}; \theta_i) = {} & \Phi\left(\frac{-b_{i,t}}{\sigma_{u_i}}\right)\frac{1}{\sigma_{\varepsilon_i}}\phi\left(\frac{a_{i,t}}{\sigma_{\varepsilon_i}}\right) \\ & + \int_{-b_{i,t}}^{\infty}\frac{1}{\sigma_{u_i}}\phi\left(\frac{u_{i,t}}{\sigma_{u_i}}\right)\frac{1}{\sigma_{\varepsilon_i}}\phi\left(\frac{a_{i,t} - b_{i,t} - u_{i,t}}{\sigma_{\varepsilon_i}}\right)du_{i,t}, \end{aligned} \tag{24}$$

where $a_{i,t} \equiv y_t - \mu_i - \sum_{i=1}^{p}\rho_{i,p}y_{t-i} - \pi_i x_{t-1}$ and $b_{i,t} \equiv \alpha_i + \beta_i x_{t-1}$. Franses and Paap (2001) show that (24) can be simplified to

$$\begin{aligned} f_i(y_t|Y_{t-1}, x_{t-1}; \theta_i) = {} & \Phi\left(\frac{-b_{i,t}}{\sigma_{u_i}}\right)\frac{1}{\sigma_{\varepsilon_i}}\phi\left(\frac{a_{i,t}}{\sigma_{\varepsilon_i}}\right) \\ & + \Phi\left(\frac{(\sigma_{u_i}/\sigma_{\varepsilon_i})a_{i,t} + (\sigma_{\varepsilon_i}/\sigma_{u_i})b_{i,t}}{\sqrt{\sigma_{u_i}^2 + \sigma_{\varepsilon_i}^2}}\right)\frac{1}{\sqrt{\sigma_{u_i}^2 + \sigma_{\varepsilon_i}^2}}\phi\left(\frac{a_{i,t} - b_{i,t}}{\sqrt{\sigma_{u_i}^2 + \sigma_{\varepsilon_i}^2}}\right). \end{aligned} \tag{25}$$

The log-likelihood function (19) can be maximized with respect to θ. The maximum likelihood estimator is asymptotically normally distributed with mean θ and covariance matrix equal to the inverse of the information matrix, $\mathcal{I}(\theta)^{-1}$, where

$$\mathcal{I}(\theta) = -\mathrm{E}\left(\frac{\partial^2\ell(Y_T|X_T; \theta)}{\partial\theta\partial\theta'}\right). \tag{26}$$

To find the maximum likelihood estimator we use the BFGS maximization algorithm, see e.g. Fletcher (1987). This algorithm has the advantage, that it does not require second order derivatives. However, first order derivatives are required, and we use analytical derivatives in our maximization procedure. The information matrix is estimated by evaluating the actual Hessian matrix in (26) at the ML estimates, instead of the expected Hessian matrix.

3.4 Small Sample Properties

To investigate the (small) sample properties of the proposed ML estimator, we perform a small simulation experiment. We generate 1000 times 250 and 1000 observations on the relevant variables, and we estimate the parameters in the models, where we take the true parameters as the starting values. The data generating process (DGP) follows the same procedure as before, but now we use 50 days for the long-run moving average, instead of 200 days, and parameters similar to the parameters found in practice, see the next section. Hence x_t is generated by,

$$x_t = \frac{\overline{z}_{1,t} - \overline{z}_{50,t}}{\overline{z}_{50,t}}, \tag{27}$$

while y_t is generated by

$$y_t = \begin{cases} 0.2 + 0.2y_{t-1} + 0x_{t-1} + v_t + \varepsilon_{1,t} & \text{if } x_{t-1} > 0 \\ 0.4 + 0.1y_{t-1} + 0.01x_{t-1} + w_t + \varepsilon_{2,t} & \text{if } x_{t-1} \leq 0, \end{cases} \tag{28}$$

$$v_{1,t} = \begin{cases} 0.1 + 0.05x_{t-1} + u_{1,t} & \text{if } 0.1 + 0.05x_{t-1} + u_{1,t} > 0 \\ 0 & \text{if } 0.1 + 0.05x_{t-1} + u_{1,t} \leq 0 \end{cases} \tag{29}$$

$$v_{2,t} = \begin{cases} -0.4 - 0.2x_{t-1} + u_{2,t} & \text{if } -0.4 - 0.2x_{t-1} + u_{2,t} > 0 \\ 0 & \text{if } -0.4 - 0.2x_{t-1} + u_{2,t} \leq 0, \end{cases} \tag{30}$$

where $\varepsilon_{1,t} \sim N(0, 0.02)$, $\varepsilon_{2,t} \sim N(0, 0.05)$, $u_{1,t} \sim N(0, 0.8)$ and $u_{2,t} \sim N(0, 2)$.

Next, we estimate the model parameters using ML. The relevant simulation results are given in Table 2. The third column of Table 2 shows the average ML estimator over all replications, while the fourth column shows the root mean squared error [RMSE] of the ML estimator, indicating the precision of the estimator. The ML estimator becomes increasingly more precise with increasing sample size, thereby suggesting consistency. The final six columns in Table 2 deal with the distributional properties of the ML estimator, where we focus on the so-called z-scores. We compute the tail probabilities of these z-scores for different percentiles of the standard normal distribution. We observe that, even for a sample as small as 250 observations, the distribution of the z-scores is rather close to the normal distribution. When we increase the sample size, we see an even closer match and hence the estimator seems asymptotically normally distributed.

Table 2: Properties of the ML estimator for the SR-CLEAR model (28)-(30) for sample sizes of 250 observations and 1000 observations.[1]

parameter	true value	$E[\hat{\theta}]$	$\sqrt{E[(\hat{\theta}-\theta)^2]}$	nominal size z-scores[2]					
				left tail			right tail		
				0.01	0.05	0.10	0.10	0.05	0.01
			250 observations						
$\rho_{1,1}$	0.20	0.20	0.034	0.02	0.07	0.12	0.13	0.07	0.02
μ_1	0.20	0.20	0.050	0.03	0.07	0.13	0.10	0.05	0.02
π_1	0.000	0.000	0.010	0.01	0.05	0.10	0.14	0.08	0.03
$\sigma^2_{\varepsilon,1}$	0.020	0.019	0.006	0.10	0.18	0.25	0.05	0.01	0.00
α_1	0.10	0.11	0.171	0.00	0.03	0.08	0.13	0.08	0.02
β_1	0.050	0.047	0.033	0.02	0.06	0.14	0.08	0.04	0.01
$\sigma^2_{u,1}$	0.80	0.80	0.149	0.04	0.09	0.15	0.06	0.02	0.00
$\rho_{2,1}$	0.10	0.09	0.087	0.04	0.10	0.18	0.13	0.08	0.04
μ_2	0.40	0.40	0.129	0.04	0.09	0.14	0.12	0.08	0.03
π_2	0.010	0.07	0.082	0.07	0.13	0.18	0.14	0.09	0.04
$\sigma^2_{\varepsilon,2}$	0.050	0.042	0.024	0.20	0.31	0.37	0.02	0.01	0.00
α_2	-0.40	-0.40	0.587	0.00	0.01	0.05	0.15	0.10	0.03
β_2	-0.20	-0.20	0.236	0.00	0.05	0.09	0.13	0.07	0.02
$\sigma^2_{u,2}$	2.00	2.07	0.986	0.08	0.13	0.18	0.03	0.01	0.00
			1000 observations						
$\rho_{1,1}$	0.20	0.20	0.015	0.01	0.05	0.09	0.10	0.06	0.01
μ_1	0.20	0.20	0.023	0.02	0.07	0.13	0.09	0.05	0.01
π_1	0.000	-0.000	0.004	0.01	0.05	0.09	0.12	0.05	0.02
$\sigma^2_{\varepsilon,1}$	0.020	0.020	0.003	0.03	0.09	0.16	0.06	0.02	0.01
α_1	0.10	0.11	0.075	0.01	0.04	0.08	0.13	0.07	0.02
β_1	0.050	0.049	0.012	0.02	0.06	0.11	0.09	0.04	0.01
$\sigma^2_{u,1}$	0.80	0.80	0.068	0.02	0.07	0.12	0.08	0.03	0.00
$\rho_{2,1}$	0.10	0.10	0.031	0.02	0.06	0.10	0.11	0.06	0.01
μ_2	0.40	0.40	0.047	0.02	0.06	0.11	0.10	0.05	0.01
π_2	0.010	0.010	0.016	0.02	0.06	0.10	0.11	0.06	0.01
$\sigma^2_{\varepsilon,2}$	0.050	0.049	0.009	0.05	0.13	0.20	0.05	0.02	0.00
α_2	-0.40	-0.40	0.197	0.00	0.03	0.08	0.12	0.07	0.02
β_2	-0.20	-0.20	0.044	0.01	0.05	0.10	0.09	0.05	0.01
$\sigma^2_{u,2}$	2.00	2.00	0.334	0.03	0.09	0.16	0.06	0.02	0.00

[1] The DGP is given by (13), (18), (27), (28), (29) and (30). The number of replications is 1000.

[2] The z-scores are defined as $(\hat{\theta}-\theta)/\hat{\sigma}_{\hat{\theta}}$, where $\hat{\sigma}_{\hat{\theta}}$ denotes the estimated standard error of $\hat{\theta}$. The cells report the empirical size.

In sum, the simulation results suggest that we may safely rely on ML estimation in practice.

3.5 Residuals and Fit

The SR-CLEAR model contains unobserved variables which makes the construction of residuals different from the standard regression case. The fit of the model, or the expectation of y_t given past observations and x_{t-1}, $\mathrm{E}[y_t|Y_{t-1}, x_{t-1}; \hat{\theta}]$, is given by

$$\mathrm{E}[y_t|Y_{t-1}, x_{t-1}; \hat{\theta}] = \mathrm{E}_1[y_t|Y_{t-1}, x_{t-1}; \hat{\theta}_1] I_{x_{t-1}>0} + \mathrm{E}_2[y_t|Y_{t-1}, x_{t-1}; \hat{\theta}_2](1 - I_{x_{t-1}>0}), \quad (31)$$

where

$$\mathrm{E}_i[y_t|Y_{y-1}, x_{t-1}; \hat{\theta}_i] = \hat{\mu}_1 + \hat{\rho}_{i,1} y_{t-1} + \ldots + \hat{\rho}_{i,p} y_{t-p} + \hat{\pi}_i x_{t-1} + \mathrm{E}[v_{i,t}|x_{t-1}; \hat{\theta}_i] \quad (32)$$

for $i = 1, 2$. The expectation of $v_{i,t}$ given x_{t-1} follows from the expectation of a censored normal distribution, that is,

$$\mathrm{E}[v_{i,t}|x_{t-1}; \hat{\theta}_i] = \Phi\left(\frac{\hat{\alpha}_i + \hat{\beta}_i x_{t-1}}{\hat{\sigma}_{u_i}}\right)(\hat{\alpha}_i + \hat{\beta}_i x_{t-1}) + \hat{\sigma}_{u_i} \phi\left(\frac{\hat{\alpha}_i + \hat{\beta}_i x_{t-1}}{\hat{\sigma}_{u_i}}\right), \quad (33)$$

see Franses and Paap (2001). Residuals in each of the two regimes can be defined as the difference between y_t and its conditional expectation, that is

$$\hat{\varepsilon}_t = y_t - \mathrm{E}[y_t|Y_{t-1}, x_{t-1}; \hat{\theta}] \quad (34)$$

The residuals $\hat{\varepsilon}_t$ are heteroscedastic as the conditional variance of y_t equals

$$\begin{aligned}
\mathrm{V}[y_t|Y_{t-1}, x_{t-1}; \theta] &= \mathrm{E}[(y_t - \mathrm{E}[y_t|Y_{t-1}, x_{t-1}; \theta])^2 | Y_{t-1}, x_{t-1}; \theta] \\
&= \mathrm{E}[(\varepsilon_{1,t} + v_{1,t} - \mathrm{E}[v_{1,t}|x_{t-1}; \theta_1])^2 I_{x_{t-1}>0} \\
&\qquad + (\varepsilon_{2,t} + v_{2,t} - \mathrm{E}[v_{2,t}|x_{t-1}; \theta_2])^2 (1 - I_{x_{t-1}>0}) | Y_{t-1}, x_t; \theta] \\
&= (\sigma^2_{\varepsilon_1} + \mathrm{V}[v_{1,t}|x_{t-1}; \theta_1]) I_{x_{t-1}>0} + (\sigma^2_{\varepsilon_2} + \mathrm{V}[v_{2,t}|x_{t-1}; \theta_2])(1 - I_{x_{t-1}>0}).
\end{aligned} \quad (35)$$

The variance of $v_{i,t}$ given x_{t-1} follows from the variance of a censored normal distribution,

$$\mathrm{V}[v_{i,t}|x_{t-1}; \theta_i] = \sigma^2_{u_i} \Phi\left(\frac{\alpha_i + \beta_i x_{t-1}}{\sigma_{u_i}}\right) + (\alpha_i + \beta_i x_{t-1}) \mathrm{E}[v_{i,t}|x_{t-1}; \theta_i] - \mathrm{E}[v_{i,t}|x_{t-1}; \theta_i]^2, \quad (36)$$

see Franses and Paap (2001). The variance of $v_{i,t}$ increases with $\alpha_i + \beta_i x_{t-1}$, and hence the variance of the residuals will be larger when $\alpha_i + \beta_i x_{t-1}$ is large.

3.6 Inference on the censored latent variable

We could interpret the variables $v_{1,t}$ and $v_{2,t}$ as a generator of innovation outliers. Financial analysts and investors are often more interested in the probability of these shocks, than in the point estimate of a variable. With the variables $v_{1,t}$ and $v_{2,t}$, the SR-CLEAR model allow for an intuitive way to draw inference on innovation outliers.

For example, if one wants to predict the probability that a shock (either $v_{1,t}$ or $v_{2,t}$) occurs given x_{t-1}, one uses (22) to compute

$$\begin{aligned}
&\Pr[v_{1,t} > 0 \vee v_{2,t} > 0 | x_{t-1}; \hat{\theta}] = \\
&\quad (1 - \Pr[v_{1,t} = 0 | x_{t-1}; \hat{\theta}_1]) I_{x_{t-1}>0} + (1 - \Pr[v_{2,t} = 0 | x_{t-1}; \hat{\theta}_2])(1 - I_{x_{t-1}>0}) = \\
&\qquad \Phi\left(\frac{\hat{\alpha}_1 + \hat{\beta}_1 x_{t-i}}{\hat{\sigma}_{u_1}}\right) I_{x_{t-1}>0} + \Phi\left(\frac{\hat{\alpha}_2 + \hat{\beta}_2 x_{t-i}}{\hat{\sigma}_{u_2}}\right)(1 - I_{x_{t-1}>0}), \qquad (37)
\end{aligned}$$

where we use the symmetry of the normal distribution.

In other situations, we may want to have ex-post information about innovation outliers. That is, we would like to know when outliers occurred in the past, and how large these outliers were. The variables $v_{i,t}$ for $i = 1, 2$ are unobserved, but we might use the information on Y_t to derive the ex-post probability that $v_{i,t}$ was positive, and the ex-post expectation of the size of $v_{i,t}$. These *conditional* probabilities and expectations could be used to analyze the probability and size of outliers in the estimation sample. For example, Franses and Paap (2001) used the conditional probability of v_t given Y_t from their CLEAR model to identify recession periods in the United States, that is periods where innovation outliers affected US unemployment.

An estimate of the probability that y_t is affected by an outlier (either $v_{1,t}$ or $v_{2,t}$) given Y_t and x_{t-1} is

$$\begin{aligned}
&\Pr[v_{1,t} > 0 \vee v_{2,t} > 0 | Y_t, x_{t-1}; \hat{\theta}] = \\
&\quad (1 - \Pr[v_{1,t} = 0 | Y_t, x_{t-1}; \hat{\theta}_1]) I_{x_{t-1}>0} + (1 - \Pr[v_{2,t} = 0 | Y_t, x_{t-1}; \hat{\theta}_2])(1 - I_{x_{t-1}>0}), \qquad (38)
\end{aligned}$$

where the probability that $v_{i,t} = 0$ given Y_t and x_{t-1} in regime i equals

$$\Pr[v_{i,t} = 0 | Y_t, x_{t-1}; \theta_i] = \frac{\Pr[v_{i,t} = 0 | x_{t-1}; \theta_i] f_i(y_t | Y_{t-1}, x_{t-1}, v_{i,t}; \theta_i)|_{v_{i,t}=0}}{f_i(y_t | Y_{t-1}, x_{t-1}; \theta_i)}. \qquad (39)$$

$\Pr[v_{i,t} = 0|x_{t-1}; \theta_i]$, $f_i(y_t|Y_{t-1}, x_{t-1}, v_t; \theta_i)$ and $f_i(y_t|Y_{t-1}, x_{t-1}; \theta_i)$ are defined in (22), (23) and (21), respectively.

An estimate of the value of the outlier (either $v_{1,t}$ or $v_{2,t}$) that affects y_t is

$$\mathrm{E}[v_{1,t}|Y_t, x_{t-1}; \hat{\theta}]I_{x_{t-1}<0} + \mathrm{E}[v_{2,t}|Y_t, x_{t-1}; \hat{\theta}](1 - I_{x_{t-1}<0}), \tag{40}$$

where the expected value of the shock $v_{i,t}$ given the values of Y_t and x_{t-1} in regime i equals

$$\mathrm{E}[v_{i,t}|Y_t, x_{t-1}; \theta_i] = \frac{\int_{-\alpha_i - \beta_i x_{t-1}}^{\infty} (\alpha_i + \beta_i x_{t-1} + u_{i,t}) \frac{1}{\sigma_{u_i}} \phi\left(\frac{u_{i,t}}{\sigma_{u_i}}\right) f_i(y_t|Y_{t-1}, x_{t-1}, v_{i,t}; \theta_i)|_{v_{i,t}=\alpha_i+\beta_i x_{t-1}+u_{i,t}} du_{i,t}}{f_i(y_t|Y_{t-1}, x_{t-1}; \theta_i)}. \tag{41}$$

for $i = 1, 2$.

3.7 Forecasting

Finally we discuss out-of-sample forecasting using a switching regime CLEAR model. We first consider out-of-sample inference on the probability that a shock will affect the series y_t at time $T+1$. The predictive probability that the series y_t is affected by a positive shock at $T+1$ is given by

$$(1 - \Pr[v_{1,T+1} = 0|x_T; \theta_1)I_{x_T>0} + (1 - \Pr[v_{2,T+1} = 0|x_T; \theta_2)(1 - I_{x_T>0}), \tag{42}$$

where $\Pr[v_{i,T+1} = 0|x_T; \theta_i)$ is defined in (22) for $i = 1, 2$. Additionally, a forecast of the value of this shock is

$$\mathrm{E}[v_{1,T+1}|x_T; \hat{\theta}_1]I_{x_T>0} + \mathrm{E}[v_{2,T+1}|x_T; \hat{\theta}_2](1 - I_{x_T>0}), \tag{43}$$

where $\mathrm{E}[v_{i,T+1}|x_T; \theta_i]$ is defined in (33).

Given forecasts for future shocks, we can construct forecasts for the series y_t. The one-step ahead forecast at time T conditional on Y_T and x_T is given by

$$\begin{aligned}\mathrm{E}[y_{T+1}|Y_T, x_T; \hat{\theta}] &= \int_{-\infty}^{\infty} y_{T+1} f(y_{T+1}|Y_T, x_{T1}; \hat{\theta})\, dy_{T+1} \qquad (44)\\ &= (\hat{\mu}_1 + \hat{\rho}_{1,1} y_{t-1} + \ldots + \hat{\rho}_{1,p} y_{t-p} + \hat{\pi}_1 x_{t-1} + \mathrm{E}[v_{1,t}|x_{t-1}; \hat{\theta}_1])I_{x_T>0} \\ &+ (\hat{\mu}_2 + \hat{\rho}_{2,1} y_{t-1} + \ldots + \hat{\rho}_{2,p} y_{t-p} + \hat{\pi}_2 x_{t-1} + \mathrm{E}[v_{2,t}|x_{t-1}; \hat{\theta}_2])(1 - I_{x_T>0}),\end{aligned}$$

where $f(y_{T+1}|Y_T, x_T; \theta)$ and $\text{E}[v_{i,T+1}|x_T; \theta_i]$ are defined as in (21) and (33), respectively. Note that the one-step ahead forecast corresponds to the fit defined in (32).

4 The SR-CLEAR model in practice

In this section we apply the SR-CLEAR(p) model to several major stock market indices, that is, the Dow Jones (US), the NASDAQ (US), the S&P 500 (US), the Nikkei-225 (Japan), the FTSE-100 (Great Britain), the DAX Xetra (Germany), the CAC-40 (France), the AEX (the Netherlands) and the Hang Seng (Hongkong). All data are obtained from Datastream. We compute absolute returns for these indices, and we consider five ratios of moving averages with different short-run and long-run periods, that is 1 day versus 50 days, 1 day versus 150 days, 5 days versus 150 days, 1 day versus 200 days and 2 days versus 200 days. Brock, Lakonishok and LeBaron (1992) also use these periods for their research on technical trading rules.

Preliminary analysis of the autocorrelation function suggests that a lag order of 5 captures autocorrelation structures quite well. Hence, we estimate a SR-CLEAR(5) model for all nine indices with five different ratios of short-run and long-run moving averages covering the period from 1/1/1990 to 31/12/1999. The size of the sample is over 2500 observations, and the specific number depends on the country-specific non-trading days. Table 3 shows the relevant estimation results for the S&P 500, which turn out to be rather representative, while the other estimation results are shown in the Appendix. In general we find that π_1 and π_2 are not significant, while β_1 is in general significantly positive and β_2 is in general significantly negative. This clearly indicates that the relation between x_{t-1} and y_t is non-linear, and above all, that is seems to get channelled through the censored latent effects variables $v_{1,t}$ and $v_{2,t}$. When the absolute value of x_{t-1} is above a stochastic threshold, this will create an outlier and hence increase absolute returns. When the absolute value of x_{t-1} is very large, that is, when the stock market is extremely bullish or bearish, these outliers will likely come in sequences, and hence there will be a temporary level shift, until $|x_{t-1}|$ comes down again.

The estimation results for the regime where $x_{t-1} < 0$ are rather consistent across

Table 3: Parameter estimates of a Switching Regime CLEAR(5) model for absolute returns on the S&P 500.[a]

y_t	S&P 500				
Period	1/2/1990 – 12/31/1999 (2527 obs.)				
x_t[b]	1-50	1-150	5-150	1-200	2-200
Linear part, when $x_{t-1} > 0$[c]					
μ_1	**0.114**	**0.120**	**0.130**	**0.119**	**0.114**
	(0.014)	**(0.014)**	**(0.015)**	**(0.014)**	**(0.014)**
π_1	-0.004	0.000	-0.000	0.001	0.001
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
$\sigma^2_{\epsilon_1}$	**0.013**	**0.014**	**0.015**	**0.014**	**0.014**
	(0.002)	**(0.002)**	**(0.002)**	**(0.002)**	**(0.002)**
Censored latent effect, when $x_{t-1} > 0$					
α_1	**0.137**	0.089	0.013	0.083	*0.096*
	(0.044)	(0.050)	(0.054)	(0.051)	*(0.049)*
β_1	**0.027**	**0.020**	**0.029**	**0.016**	**0.016**
	(0.010)	**(0.005)**	**(0.006)**	**(0.005)**	**(0.005)**
$\sigma^2_{u_1}$	**0.491**	**0.599**	**0.636**	**0.645**	**0.646**
	(0.031)	**(0.036)**	**(0.039)**	**(0.038)**	**(0.037)**
Linear part, when $x_{t-1} < 0$[c]					
μ_2	**0.194**	**0.196**	**0.178**	**0.262**	**0.274**
	(0.023)	**(0.035)**	**(0.027)**	**(0.046)**	**(0.040)**
π_2	0.010	0.016	0.014	0.014	0.015
	(0.008)	(0.009)	(0.007)	(0.010)	(0.010)
$\sigma^2_{\epsilon_2}$	**0.019**	**0.028**	**0.019**	**0.045**	**0.046**
	(0.004)	**(0.011)**	**(0.006)**	**(0.010)**	**(0.009)**
Censored latent effect, when $x_{t-1} < 0$					
α_2	-0.067	-0.227	-0.041	**-0.555**	**-0.709**
	(0.087)	(0.165)	(0.110)	**(0.213)**	**(0.212)**
β_2	**-0.170**	**-0.132**	**-0.115**	**-0.161**	**-0.182**
	(0.021)	**(0.021)**	**(0.020)**	**(0.031)**	**(0.031)**
$\sigma^2_{u_2}$	**1.120**	**1.260**	**1.098**	**1.489**	**1.545**
	(0.099)	**(0.193)**	**(0.127)**	**(0.235)**	**(0.243)**
Log-Likelihood	-1469.381	-1508.371	-1517.403	-1520.723	-1517.623

[a] The model is given in (10)-(12). Standard errors are given below the parameter estimates in parentheses. Numbers in italics indicate significance at the 5% level, while boldface indicates significance at the 1% level.

[b] 1-50 denotes the explanatory variable is a ratio of 1-day and 50-day moving averages.

[c] Estimates for the AR parameters $\rho_{i,k}$ for $i = 1, 2$ and $k = 1, \ldots, 5$ are not shown in the table to save space.

Table 4: Tests on parameter restrictions in the Switching Regime CLEAR(5) model for absolute returns on the S&P 500.[a]

y_t	S&P 500				
x_t[b]	1-50	1-150	5-150	1-200	2-200
$H_0: \rho_{1,1} = \ldots = \rho_{1,p} = 0$[c]	**38.589**	**32.693**	**33.082**	**33.545**	**33.450**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0: \rho_{2,1} = \ldots = \rho_{2,p} = 0$[c]	8.719	**16.495**	7.782	**26.297**	**27.294**
	(0.121)	**(0.006)**	(0.169)	**(0.000)**	**(0.000)**
$H_0: \pi_1 = \pi_2 = 0$[c]	3.775	3.069	3.787	2.203	3.209
	(0.151)	(0.216)	(0.151)	(0.332)	(0.201)
$H_0: \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0$[d]	**867.212**	**1081.984**	**1118.443**	**1148.274**	**1155.937**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0: \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0$[d]	**451.423**	**279.945**	**262.587**	**243.391**	**254.757**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0: \sigma^2_{u_1} = \sigma^2_{u_2}$[c]	**36.689**	**11.372**	**12.013**	**12.545**	**13.315**
	(0.000)	**(0.001)**	**(0.001)**	**(0.000)**	**(0.000)**
$H_0: \beta_1 = -\beta_2$[c]	**37.852**	**26.109**	**17.246**	**21.578**	**27.121**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

[a] P-values are given below the test statistics in parentheses. Number in italics indicate significance at the 5% level, while boldface indicates significance at the 1% level.
[b] 1-50 denotes the explanatory variable is a ratio of 1-day and 50-day moving averages.
[c] Wald test against the alternative of non-equality.
[d] Likelihood ratio test against $H_A : \alpha_i \neq 0 \vee \beta_i \neq 0 \vee \sigma^2_{u_i} > 0$. The test statistic is asymptotically $\frac{1}{2}\chi^2(2) + \frac{1}{2}\chi^2(3)$ distributed, see Wolak (1989, p. 19-20).

the markets and x_{t-1} variables. The parameter β_2 is always highly significant with values ranging from -0.200 and -0.070. This effect is much stronger than the censored latent effect in the first regime, when x_{t-1} is positive, where we see lower and sometimes insignificant values of β_1 and also unexpected signs. Hence, it looks as if x_{t-1} has a stronger effect on absolute returns when x_{t-1} is negative, and this supports the leverage effect conjecture.

Our choice for a SR-CLEAR(5) model is further substantiated by the test results on parameter restrictions, as shown in Table 4 for the S&P 500. Test results for other stock indices can be found in the Appendix. The first two Wald tests show that lagged values $y_{t-1}, \ldots, y_{t-5}$ generally have a significant effect on y_t, hence a lag order of at least five lags seems justified. However, the significance of the lagged values of y_t tends to be smaller in the regime where $x_{t-1} < 0$. This finding contrasts with the results in Glosten, Jagannathan and Runkle (1992), who find stronger lag effects when $y_{t-1} < 0$ due to the leverage effect. Note however that in our model, the leverage effect is mostly absorbed by the strong censored latent effect in the second regime.

The results for the Wald test in the third row shows that x_{t-1} generally has no significant linear effect on y_t. Hence y_t appears to be mainly explained by the censored latent effects $v_{1,t}$ and $v_{2,t}$. This is also supported by the test results in the fourth and fifth row, which show that the censored latent effects are not redundant. Finally, the seventh and eighth row suggest that β_1 differs significantly from β_2, and similarly that $\sigma^2_{u_1}$ differs significantly from $\sigma^2_{u_2}$. This supports our choice for a switching regime model with two regimes.

To get an impression of the properties of the SR-CLEAR model we show actual values and fitted values for the absolute returns on the S&P 500 in Figure 7. Here x_t is constructed from a 1-day and a 50-day moving average. The fitted values do not show much variation, except in periods when there is a high level of absolute returns or when there are a few large outliers. As a result, the SR-CLEAR seems to pick up the temporary level shift in 1990 fairly well. The SR-CLEAR is also able to filter the extreme outlier of 6.8% on 31/8/1998 during the Asia crisis to a residual of 5.1%.

The graph in Figure 7 also shows, that the SR-CLEAR can miss out on outliers. For example, 1991 showed a period of high level absolute returns, and a few observations of more than 3%, but this is not picked up by the model. The SR-CLEAR model also does not pick up the outlier of 6.9% on 27/10/1997. There are at least two reasons why this might occur. The first reason is that x_{t-1} has an effect on y_t only when it exceeds a *stochastic* threshold. Hence it is possible that an outlier occurs, even if $|x_{t-1}|$ is fairly small. A second reason is that x_{t-1} has relatively little explanatory power in the first regime, as we saw that π_1 and β_1 are often insignificant. This is also illustrated in Figure 8. This figure shows unconditional expectations of $v_{1,t}$ in regime 1 and $v_{2,t}$ in regime 2 as defined in (33), together with x_t. We may consider the unconditional expectation of $v_{i,t}$ as the expectation of a positive outlier. Figure 8 shows that in the first regime, when $x_t > 0$,

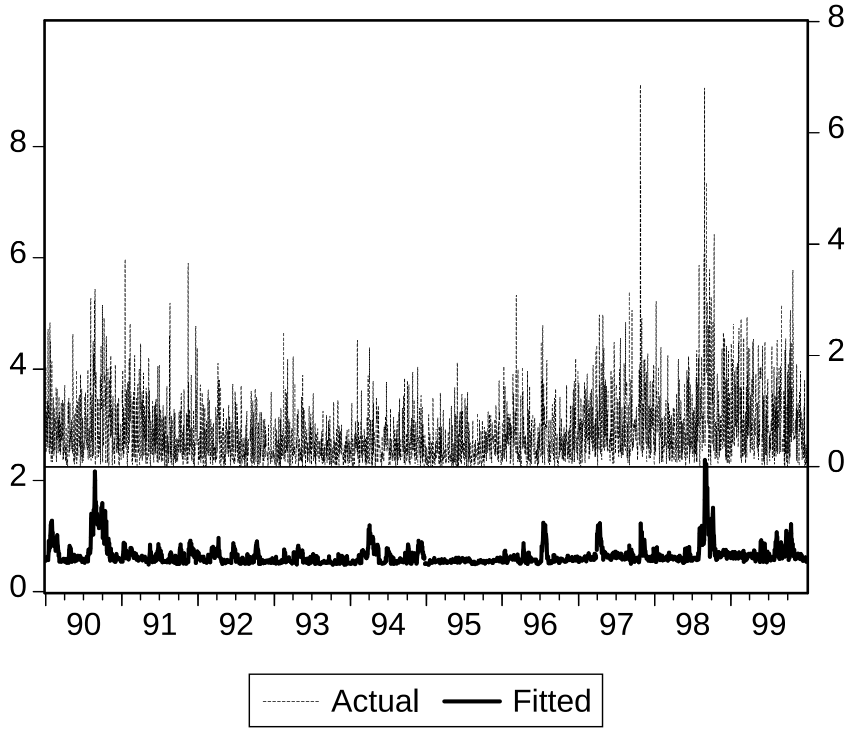

Figure 7: Actual and fitted absolute returns on the S&P 500.

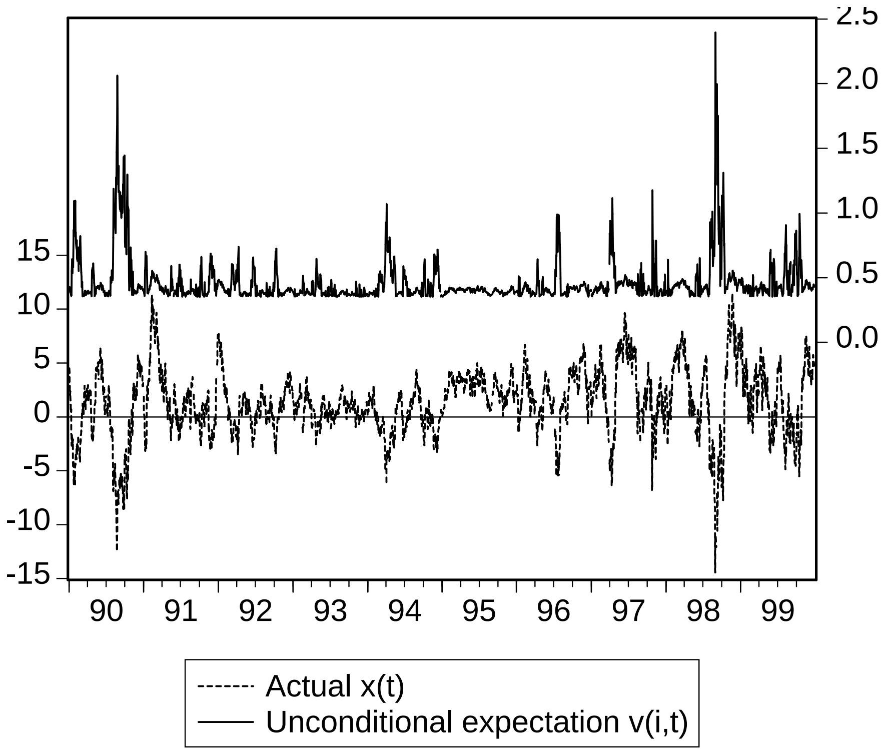

Figure 8: Unconditional expectations of $v_{i,t}$ for $i = 1, 2$, and actual values of x_t.

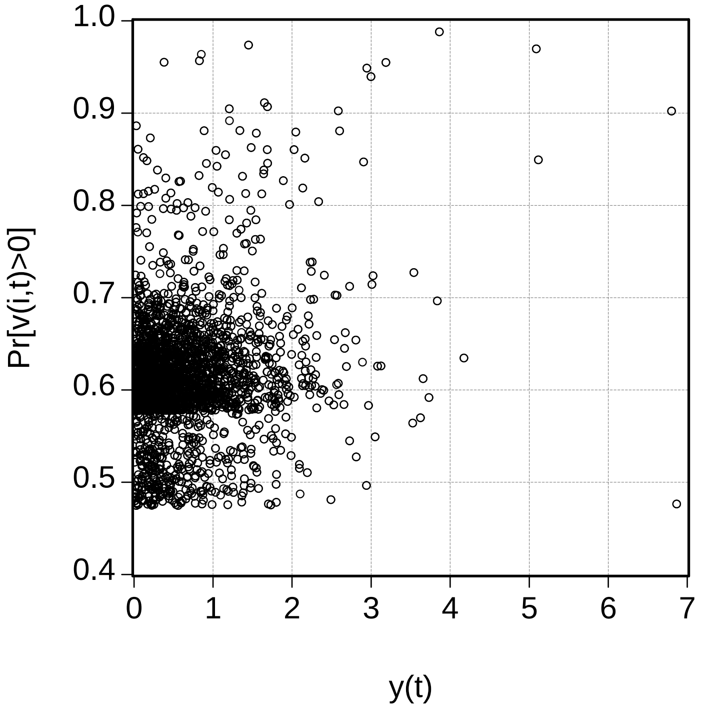

Figure 9: The unconditional probabilities $\Pr[v_{i,t} > 0|x_{t-1}; \theta_i]$ for $i = 1, 2$ against absolute returns on the S&P 500.

there is almost no variation in the unconditional expectation of $v_{1,t}$. Hence, the ratio of moving averages does not seem to be able to explain outliers when $x_{t-1} > 0$. This also suggests that x_{t-1} may not be a very good explanatory variable in the first place.

The SR-CLEAR model could be used for other purposes than point forecasting. One of the nice features of the model is that we have a natural indicator to assess the probability that a shock occurs. We can compute the unconditional probability that $v_{1,t}$ or $v_{2,t}$ is greater than zero, as given in (37). Figure 9 visualizes the relation between absolute returns, and the unconditional probability for both $v_{1,t} > 0$ and $v_{2,t} > 0$. Most probabilities are below 75%, but there are several observations that show high probabilities up to 99%. Some of these high probabilities correspond with high absolute returns of more than 5.0. In Table 5 we show the within-sample predictions and realizations that absolute returns are higher than 3.0, where we use a cut-off point of 75%. The table shows that when we would predict an outlier, absolute returns were larger than 3.0 in 5 of 94 cases. Hence,

Table 5: Within-sample prediction-realization table of high absolute returns for the S&P 500.

		$\Pr[v_{i,t} > 0 \mid x_{t-1}; \hat{\theta}_i] > .75$		
		yes	no	
$y_t > 3.0$	yes	5	13	18
	no	89	2420	2509
		94	2433	2527

we would need 94 signals, to get 5 high absolute returns correct. In 2433 cases we would not have predicted an outlier, and in 2420 cases we would have been correct.

Summarizing, our estimation and evaluation results support our choice for a model with censored latent effects and switching regimes. In the next section, we will see if our SR-CLEAR model also improves forecastibility of absolute returns in comparison with other models.

5 In-sample fit and out-of-sample forecasts

Although the SR-CLEAR model does seem to describe a few of the features of absolute returns quite well, it does not imply that the SR-CLEAR model will be able to accurately forecast absolute returns. To assess the fit and forecasting abilities of the SR-CLEAR model, we compare within-sample and out-of-sample forecasts of the SR-CLEAR model to the forecasts of three other models. These three models are the single regime CLEAR model as discussed in Section 3.1, the ARFIMA model as discussed in Section 2.1 and the ARFIMAX model as discussed in Section 2.2.

We estimate the parameters of the SR-CLEAR(5) model, the CLEAR(5) model, the ARFIMAX$(5, d, 0)$ model, and the ARFIMA$(5, d, 0)$ using ML for nine stock market indices and the five explanatory variables we also used earlier. We create within-sample forecasts of absolute returns and we compute median squared prediction errors [MedSPE] of these forecasts. With our parameter estimates we also compute the log-likelihood and the Akaike Information Criterium [AIC], defined by

$$AIC = -2\frac{\ell(Y_T|X_{T-1};\hat{\theta})}{T} + \frac{2k}{T},$$

where k denotes the total number of parameters. The model with minimum AIC value is to be preferred.

Table 6 shows the results for the within-sample fit, where we use the ratio of 1-day and 50-day moving averages as the explanatory variable. This choice is based on the fact that this x_{t-1} variable generally leads to the largest log-likelihoods across other choices for x_{t-1}. The ARFIMAX model appears better in terms of MedSPE, but the SR-CLEAR model is more adequate in terms of log-likelihood and AIC's. The difference in log-likelihood can be explained by noticing that the ARFIMA and ARFIMAX models assume that the disturbances ε_t on the fitted values have an independent and identical normal distribution. From Figure 1 it becomes clear that the error term is more likely to be heteroskedastic. That is, when the absolute returns are at a temporary high level, they show more variation than when they are at a low level. It is exactly this feature of absolute returns that is captured by the CLEAR and SR-CLEAR model, but that violates the assumptions of the ARFIMA and ARFIMAX model. Hence, the data are in fact 'more likely' to be generated by an SR-CLEAR model than an ARFIMA or ARFIMAX model. On the other hand, the outperformance of the ARFIMA and ARFIMAX models in terms of MedSPE also follows from the assumption of normal disturbances in the ARFIMA model. As a consequence, the ARFIMA and ARFIMAX models focus on minimizing the least squares of the forecast errors, while the CLEAR and SR-CLEAR model also try to capture outliers.

One might now hope that the SR-CLEAR performs better in predicting the probability and the size of outliers. We therefore compare the forecasts of the SR-CLEAR and ARFIMAX model for the 30 largest absolute returns in the S&P 500. Table 7 shows the results. Even for the 30 highest absolute returns, most ARFIMAX forecasts are closer to the actual values than the SR-CLEAR forecasts, hence the ARFIMAX model does perform better in terms of point forecasts. However, the probability density of y_t at the ML estimates of the SR-CLEAR model is higher, and hence it is more likely that an

Table 6: Within-sample performance of several models in forecasting absolute returns.[a]

	SR-CLEAR	CLEAR	ARFIMAX	ARFIMA
	Median Squared One-Step Ahead Prediction Error			
Dow Jones	0.134	0.146	**0.115**	0.124
NASDAQ	0.184	0.190	**0.141**	0.151
S&P 500	0.128	0.146	**0.105**	0.117
Nikkei-225	0.340	0.369	**0.313**	0.328
FTSE-100	0.130	0.135	**0.115**	0.122
DAX Xetra	0.239	0.243	0.194	**0.187**
CAC-40	0.251	0.260	**0.227**	0.243
AEX	0.174	0.180	**0.136**	0.138
Hang Seng	0.415	0.448	**0.359**	0.367
Average Ranking	3.0	4.0	1.1	1.9
	Log-Likelihood			
Dow Jones	**-1514.284**	-1616.535	-2201.445	-2249.254
NASDAQ	**-2077.979**	-2127.996	-2641.322	-2698.868
S&P 500	**-1469.381**	-1563.394	-2174.899	-2228.965
Nikkei-225	**-2749.145**	-2813.716	-3421.967	-3481.714
FTSE-100	**-1667.413**	-1707.856	-2129.806	-2153.979
DAX Xetra	**-2313.140**	-2384.045	-2948.961	-2991.808
CAC-40	**-2373.929**	-2433.837	-2888.600	-2937.442
AEX	**-1935.192**	-1995.684	-2562.337	-2623.060
Hang Seng	**-2958.626**	-3038.195	-3833.843	-3899.531
Average Ranking	1.0	2.0	3.0	4.0
	Akaike Information Criterium			
Dow Jones	**1.240**	1.314	1.786	1.822
NASDAQ	**1.695**	1.726	2.141	2.185
S&P 500	**1.204**	1.271	1.764	1.805
Nikkei-225	**2.237**	2.280	2.771	2.817
FTSE-100	**1.364**	1.387	1.728	1.745
DAX Xetra	**1.885**	1.933	2.389	2.421
CAC-40	**1.934**	1.973	2.340	2.377
AEX	**1.580**	1.620	2.077	2.124
Hang Seng	**2.406**	2.461	3.103	3.154
Average Ranking	1.0	2.0	3.0	4.0

[a] Explanatory variable is the ratio of 1-day and 50-day moving averages. Boldface denotes the best performing model with respect to the variable and statistic.

Table 7: Within-sample forecasting performance of the SR-CLEAR and ARFIMAX model for the top 30 absolute returns on the S&P 500

Date	y_t	x_{t-1}[a]	SR-CLEAR $\hat{y}_t$	SR-CLEAR $f(y_t\|Y_{t-1},x_{t-1};\hat{\theta})$[b]	ARFIMAX $\hat{y}_t$	ARFIMAX $f(y_t\|Y_{t-1},x_{t-1};\hat{\theta})$[c]
27/10/1997	6.866	-0.025	0.679	0.000	0.876	0.000
31/8/1998	6.801	-8.440	1.652	0.000	1.513	0.000
28/10/1997	5.115	-6.819	1.229	0.000	1.475	0.000
8/9/1998	5.090	-12.058	2.293	0.012	2.464	0.000
15/10/1998	4.173	-2.533	0.844	0.001	1.101	0.000
1/9/1998	3.863	-14.445	2.368	0.139	1.967	0.003
27/8/1998	3.837	-3.591	0.931	0.005	1.010	0.000
17/1/1991	3.732	-1.838	0.754	0.003	0.805	0.000
15/11/1991	3.659	2.346	0.532	0.000	0.445	0.000
4/8/1998	3.624	-1.489	0.793	0.005	0.954	0.000
23/9/1998	3.541	-4.152	1.022	0.015	1.302	0.000
28/10/1999	3.527	-1.397	0.728	0.005	0.927	0.000
27/8/1990	3.188	-10.922	2.003	0.198	1.617	0.016
2/9/1997	3.125	-2.393	0.832	0.021	1.012	0.001
8/3/1996	3.083	3.280	0.603	0.000	0.594	0.000
30/9/1998	3.051	-1.163	0.797	0.019	1.152	0.003
6/8/1990	3.024	-4.089	0.978	0.042	0.958	0.001
1/10/1998	3.011	-3.913	0.930	0.039	1.322	0.009
23/8/1990	2.998	-10.033	1.870	0.208	1.600	0.035
9/1/1998	2.966	0.384	0.538	0.001	0.762	0.000
11/9/1998	2.945	-10.547	1.871	0.220	1.793	0.092
21/8/1991	2.941	-0.339	0.676	0.016	0.607	0.000
1/10/1990	2.905	-6.759	1.371	0.118	1.148	0.006
3/9/1999	2.891	-2.458	0.846	0.036	1.126	0.006
16/9/1997	2.813	-0.822	0.705	0.024	0.785	0.001
15/10/1999	2.806	-2.858	0.933	0.052	1.066	0.007
29/4/1997	2.728	-1.094	0.717	0.031	0.757	0.002
11/4/1997	2.728	-3.877	0.989	0.073	0.927	0.005
23/3/1999	2.688	3.251	0.603	0.003	0.744	0.002
9/10/1990	2.673	-2.994	0.862	0.059	0.839	0.004
Average Top 30	3.546	-3.726	1.065	0.045	1.122	0.006
Average sample	0.637	1.343	0.648	0.754	0.633	0.530

[a] Ratio of 1-day and 50-day moving average.
[b] Probability density of y_t given its past and x_{t-1} at the ML estimates of an SR-CLEAR(5) model, as given in (21).
[c] Probability density of y_t given its past and x_{t-1} at the ML estimates of an ARFIMAX$(5, d, 0)$ model.

SR-CLEAR model would have generated the absolute returns as shown in Table 7.

To assess the out-of-sample performance of the four models, we use the parameter estimates to create one-step-ahead forecasts for the year 2000. For the ARFIMA$(5, p, 0)$ model and ARFIMAX$(5, p, 0)$, we use 'naive' forecasts, see Doornik and Ooms (2001). We compare our forecasts in terms of MedSPE and log-likelihood. Table 8 shows the results, where we use the ratio of 1-day and 50-day moving average as the explanatory variable. Interestingly the CLEAR(5) model has the best out-of-sample performance in terms of MedSPE and log-likelihood. While the within-sample performance of the ARFIMA and ARFIMAX models were better in terms of MedSPE, the out-of-sample performance of these models is remarkably worse. The CLEAR(5) model also performs better than the SR-CLEAR(5) model, indicating that the SR-CLEAR may have too many parameters.

To compare the forecasts from the SR-CLEAR(5) model and the ARFIMAX$(5, d, 0)$ model more closely, we show actual values and out-of-sample forecasts from both models in Figure 10. The forecasts from the ARFIMAX model are almost always higher than the forecasts from the SR-CLEAR model. This seems to result in better forecasts for the ARFIMAX model in April and May. However, the absolute returns shift to a lower level from June to September. In this period the forecasts from the SR-CLEAR level immediately jump to a lower level, while the forecasts from the ARFIMAX model only slowly decay. Hence, the SR-CLEAR model is better able to pick up the temporary shift to a lower level of absolute returns.

From this graph it seems that the SR-CLEAR model is not able to pick up large absolute returns in 2000. In Table 9 we examine these large values more closely. The table displays the ten highest absolute returns on the S&P 500 in the year 2000, and the forecasting performance of the SR-CLEAR model and ARFIMAX model for those observations. The out-of-sample performance in forecasting large absolute returns are similar to the within-sample performance. When absolute returns are excessively large, point forecasts from the ARFIMAX model are closer to the actual values than point forecasts from the SR-CLEAR model, but these high absolute returns have a higher probability density in the SR-CLEAR model.

Table 8: Out-of-sample performance of several models in forecasting absolute returns for 2000.[a]

	SR-CLEAR	CLEAR	ARFIMAX	ARFIMA
	Median Squared One-Step Ahead Prediction Error			
Dow Jones	**0.178**	0.180	0.218	0.220
NASDAQ	1.408	1.287	**1.132**	1.215
S&P 500	0.184	**0.177**	0.261	0.289
Nikkei-225	0.351	**0.314**	0.348	0.340
FTSE-100	0.214	**0.208**	0.226	0.254
DAX Xetra	0.364	**0.339**	0.368	0.420
CAC-40	0.327	**0.311**	0.328	0.345
AEX	0.179	**0.156**	0.209	0.218
Hang Seng	0.499	**0.441**	0.523	0.686
Average Ranking	2.3	1.3	2.8	3.6
	Log-Likelihood			
Dow Jones	-315.216	-305.651	**-301.650**	-307.627
NASDAQ	-619.641	-697.211	**-495.890**	-508.120
S&P 500	**-319.758**	-339.870	-320.710	-331.103
Nikkei-225	-293.495	**-280.587**	-328.244	-333.147
FTSE-100	**-261.664**	-267.880	-274.529	-275.627
DAX Xetra	-325.457	**-312.220**	-330.982	-332.414
CAC-40	-313.300	**-305.336**	-322.147	-326.277
AEX	-257.931	**-257.819**	-282.553	-287.084
Hang Seng	-381.928	**-379.765**	-408.647	-412.276
Average Ranking	2.1	1.9	2.4	3.6

[a] Performance of one-step-ahead forecasts for the year 2000. Explanatory variable is the ratio of 1-day and 50-day moving averages. Boldface denotes the best performing model with respect to the variable and statistic.

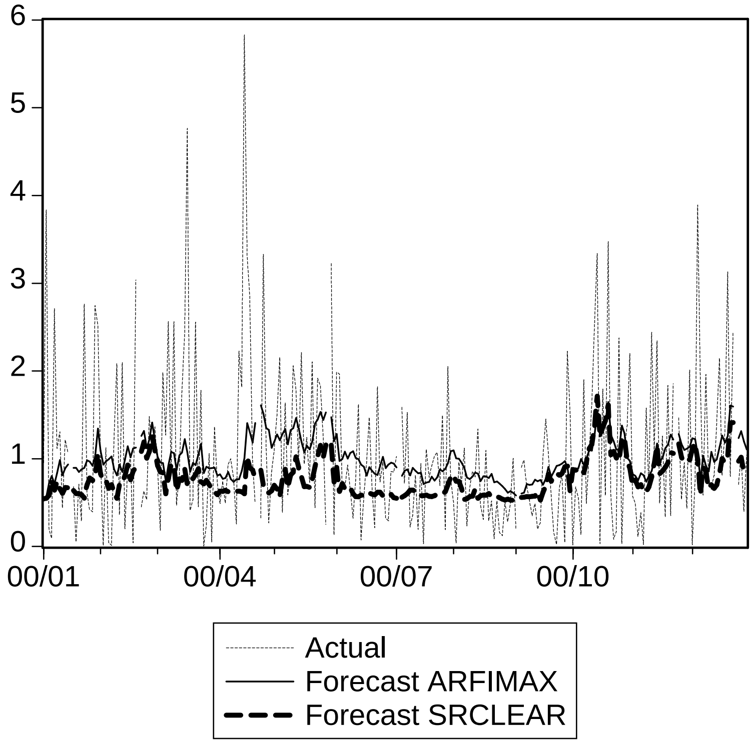

Figure 10: Actuals and one-step ahead forecasts of absolute returns on the S&P 500 for 2000, generated by an SR-CLEAR(5) model and an ARFIMAX$(5, d, 0)$ model.

Table 9: Out-of-sample forecasting performance of the SR-CLEAR and ARFIMAX model for the top 10 absolute returns on the S&P 500 for 2000

			SR-CLEAR		ARFIMAX	
Date	y_t	x_{t-1} [a]	$\hat{y}_t$	$f(y_t \vert Y_{t-1}, x_{t-1}; \hat{\theta})$ [b]	$\hat{y}_t$	$f(y_t \vert Y_{t-1}, x_{t-1}; \hat{\theta})$ [c]
14/4/2000	5.828	0.778	0.602	0.000	1.007	0.000
16/3/2000	4.764	-0.645	0.704	0.000	1.072	0.000
5/12/2000	3.892	-4.224	0.990	0.006	1.050	0.000
4/1/2000	3.834	3.565	0.549	0.000	0.564	0.000
19/10/2000	3.474	-7.930	1.611	0.074	1.647	0.004
13/10/2000	3.338	-9.346	1.713	0.111	1.552	0.005
25/4/2000	3.328	-0.009	0.707	0.006	1.513	0.005
17/4/2000	3.308	-5.005	1.025	0.028	1.404	0.003
30/5/2000	3.225	-5.090	1.154	0.044	1.475	0.006
20/12/2000	3.130	-4.456	1.041	0.040	1.257	0.003
Average Top 10	3.812	-3.236	1.010	0.031	1.254	0.003
Average sample	1.064	-0.501	0.775	0.494	0.999	0.421

[a] Ratio of 1-day and 50-day moving average.

[b] Probability density of y_t given its past and x_{t-1} at the ML estimates of an SR-CLEAR(5) model, as given in (21).

[c] Probability density of y_t given its past and x_{t-1} at the ML estimates of an ARFIMAX$(5, d, 0)$ model.

To get a more complete picture of the out-of-sample forecasting performance of the models, we perform several forecast encompassing tests (see Clements and Hendry, 1993). If we have forecasts from two competing models, the forecasts from the first model are said to encompass the forecasts from the second model, when forecasts from the first model can not be improved by adding features of the second model. Let f_{T+h} be the forecasts from the first model, and $\bar{f}_{T+h}$ the forecasts from the second model. Then f_{T+h} is said to encompass $\bar{f}_{T+h}$ if the coefficient δ in the following regression is zero

$$y_{T+h} - f_{T+h} = \delta(\bar{f}_{T+h} - f_{T+h}) + \eta_{T+h}, \text{ for } h = 1, \dots, H, \tag{45}$$

where y_{T+h} is the true value. In general, out-of-sample forecasts from the SR-CLEAR and CLEAR model do not encompass forecasts from the ARFIMAX model. Hence, the out-of-sample point forecasts of the ARFIMAX model clearly outperform the point forecasts of the SR-CLEAR and CLEAR model.

Although the performance of the SR-CLEAR model in terms of point forecasts is not impressive, the SR-CLEAR model might still be useful for predicting outliers. This is illustrated in Table 11, which shows the out-of-sample predictions and realizations that absolute returns are higher than 3.0, similar to Table 5. The table shows that, if we would signal an outlier in regime i when $\Pr[v_{i,t} > 0 | x_{t-1}; \hat{\theta}_i] > .75$, we would have been correct in 4 of 11 cases. Only on 28 of 252 days there was a signal. Hence, we would only need 28 signals to correctly predict $4/11 = 36\%$ of rather high absolute returns.

6 Concluding Remarks

In this paper we proposed a new model for daily absolute returns, which allows for modeling and forecasting innovation outliers and level shifts, and which picks up part of the long-memory properties in the data. To our knowledge, this has been the first attempt to forecast excessively large absolute returns. We showed that the model fits the data well, at least for the nine stock markets considered. The SR-CLEAR model has a natural indicator to assess the probability of an outlier in absolute returns, and we showed that an application of this indicator performs well in predicting the occurrence of a shock on

Table 10: Encompassing tests for out-of-sample one-step ahead forecasts of three models.[a]

H_0: SR-CLEAR encompasses:	CLEAR		ARFIMAX	
Dow Jones	-0.056	(0.828)	**1.363**	**(0.000)**
NASDAQ	**-0.850**	**(0.000)**	**1.122**	**(0.000)**
S&P 500	**-0.694**	**(0.028)**	**1.119**	**(0.000)**
Nikkei-225	**1.329**	**(0.001)**	0.454	(0.339)
FTSE-100	-0.701	(0.127)	**1.064**	**(0.000)**
DAX Xetra	*0.574*	*(0.084)*	**0.959**	**(0.000)**
CAC-40	**1.286**	**(0.012)**	**1.179**	**(0.000)**
AEX	0.369	(0.500)	**0.823**	**(0.000)**
Hang Seng	*0.550*	*(0.055)*	**0.470**	**(0.050)**
H_0: CLEAR encompasses:	SR-CLEAR		ARFIMAX	
Dow Jones	**1.056**	**(0.000)**	**1.175**	**(0.000)**
NASDAQ	**1.850**	**(0.000)**	**1.098**	**(0.000)**
S&P 500	**1.694**	**(0.000)**	**1.195**	**(0.000)**
Nikkei-225	-0.329	(0.396)	-0.392	(0.320)
FTSE-100	**1.701**	**(0.000)**	**1.089**	**(0.000)**
DAX Xetra	0.426	(0.198)	**0.778**	**(0.000)**
CAC-40	-0.286	(0.576)	**0.941**	**(0.000)**
AEX	0.631	(0.249)	**0.801**	**(0.000)**
Hang Seng	0.450	(0.115)	**0.465**	**(0.008)**
H_0: ARFIMAX encompasses:	SR-CLEAR		CLEAR	
Dow Jones	-0.363	(0.144)	-0.175	(0.304)
NASDAQ	-0.122	(0.385)	-0.098	(0.315)
S&P 500	-0.119	(0.561)	-0.195	(0.214)
Nikkei-225	0.546	(0.250)	**1.392**	**(0.000)**
FTSE-100	-0.064	(0.782)	-0.089	(0.635)
DAX Xetra	0.041	(0.861)	0.222	(0.249)
CAC-40	-0.179	(0.461)	0.059	(0.798)
AEX	0.177	(0.371)	0.199	(0.285)
Hang Seng	**0.530**	**(0.027)**	**0.535**	**(0.002)**

[a] Test equation is given in (45). Explanatory variable is the ratio of 1-day and 50-day moving averages. P-values are given in parentheses. The numbers in italics denote significance at the 10% level, while boldface denotes significance at the 5% level.

Table 11: Out-of-sample prediction-realization table of high absolute returns for the S&P 500.

		$\Pr[v_{i,t} > 0 \vert x_{t-1}; \hat{\theta}_i] > .75$		
		yes	no	
$y_t > 3.0$	yes	4	7	11
	no	24	217	241
		28	224	252

absolute returns, especially out-of-sample.

Although the SR-CLEAR is able to describe features of long memory quite well, the SR-CLEAR model does not seem to outperform long-memory models when we consider point forecasts. This suggests as a further research topic to introduce censored latent effects in long-memory models, thereby building on the nonlinear long-memory model put forward in van Dijk, Franses and Paap (2001). A second research topic that follows from the analysis here, amounts to including censored latent effects components in GARCH models and in models for realized volatility.

Appendix: More estimation results

Table 12: SR-CLEAR(5) model for absolute returns on the Dow Jones.

y_t	Dow Jones				
Period	1/2/1990 – 12/31/1999 (2518 obs.)				
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.153**	**0.155**	**0.159**	**0.143**	**0.145**
	(0.016)	**(0.015)**	**(0.016)**	**(0.015)**	**(0.015)**
π_1	-0.002	-0.000	-0.001	-0.000	-0.001
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
$\sigma^2_{\epsilon_1}$	**0.010**	**0.012**	**0.013**	**0.012**	**0.013**
	(0.002)	**(0.002)**	**(0.002)**	**(0.001)**	**(0.002)**
Censored latent effect when $x_{t-1} > 0$					
α_1	**0.215**	**0.186**	**0.131**	**0.237**	**0.205**
	(0.042)	**(0.045)**	**(0.049)**	**(0.044)**	**(0.046)**
β_1	*0.019*	*0.012*	**0.018**	0.005	0.007
	(0.009)	*(0.005)*	**(0.006)**	(0.005)	(0.005)
$\sigma^2_{u_1}$	**0.444**	**0.557**	**0.603**	**0.568**	**0.588**
	(0.026)	**(0.031)**	**(0.034)**	**(0.030)**	**(0.032)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.159**	**0.237**	**0.199**	**0.253**	**0.255**
	(0.032)	**(0.041)**	**(0.036)**	**(0.038)**	**(0.038)**
π_2	0.002	0.002	-0.001	0.004	-0.006
	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)
$\sigma^2_{\epsilon_2}$	**0.017**	**0.050**	**0.036**	**0.057**	**0.050**
	(0.005)	**(0.015)**	**(0.010)**	**(0.009)**	**(0.010)**
Censored latent effect when $x_{t-1} < 0$					
α_2	0.057	-0.433	-0.245	**-0.775**	**-0.517**
	(0.094)	(0.228)	(0.160)	**(0.216)**	**(0.200)**
β_2	**-0.139**	**-0.125**	**-0.108**	**-0.155**	**-0.125**
	(0.020)	**(0.024)**	**(0.020)**	**(0.030)**	**(0.027)**
$\sigma^2_{u_2}$	**1.038**	**1.406**	**1.198**	**1.901**	**1.688**
	(0.097)	**(0.242)**	**(0.177)**	**(0.309)**	**(0.275)**
Log-Likelihood	-1514.284	-1554.934	-1571.387	-1550.822	-1567.120

See Table 3 for explanatory details.

Table 13: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the Dow Jones.

y_t	Dow Jones				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	**16.895**	**17.196**	**16.597**	**26.833**	**21.095**
	(0.005)	**(0.004)**	**(0.005)**	**(0.000)**	**(0.001)**
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	6.373	6.278	6.682	**18.625**	6.950
	(0.272)	(0.280)	(0.245)	**(0.002)**	(0.224)
$H_0 : \pi_1 = \pi_2 = 0$	0.609	0.180	0.144	0.251	0.620
	(0.738)	(0.914)	(0.931)	(0.882)	(0.733)
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**809.803**	**998.822**	**1025.040**	**1038.225**	**1053.430**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**424.831**	**299.484**	**296.700**	**274.163**	**255.617**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**35.151**	**12.140**	**10.880**	**18.379**	**15.737**
	(0.000)	**(0.000)**	**(0.001)**	**(0.000)**	**(0.000)**
$H_0 : \beta_1 = -\beta_2$	**30.492**	**21.606**	**18.465**	**24.524**	**17.840**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

See Table 4 for explanatory details.

Table 14: SR-CLEAR(5) model for absolute returns on the NASDAQ.

y_t	NASDAQ				
Period	1/2/1990 – 12/31/1999 (2528 obs.)				
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.269**	**0.306**	**0.316**	**0.312**	**0.316**
	(0.026)	**(0.024)**	**(0.024)**	**(0.025)**	**(0.025)**
π_1	-0.005	-0.002	-0.002	-0.002	-0.002
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
$\sigma^2_{\epsilon_1}$	**0.052**	**0.068**	**0.067**	**0.067**	**0.068**
	(0.008)	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**
Censored latent effect when $x_{t-1} > 0$					
α_1	**-0.294**	**-0.507**	**-0.544**	**-0.514**	**-0.540**
	(0.114)	**(0.107)**	**(0.110)**	**(0.113)**	**(0.114)**
β_1	**0.053**	**0.032**	**0.038**	**0.030**	**0.033**
	(0.010)	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**
$\sigma^2_{u_1}$	**1.080**	**1.441**	**1.434**	**1.445**	**1.442**
	(0.125)	**(0.130)**	**(0.129)**	**(0.128)**	**(0.127)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.356**	**0.315**	**0.295**	**0.316**	**0.305**
	(0.038)	**(0.044)**	**(0.040)**	**(0.041)**	**(0.040)**
π_2	0.010	*0.012*	*0.012*	*0.011*	*0.010*
	(0.008)	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*
$\sigma^2_{\epsilon_2}$	**0.084**	**0.085**	**0.081**	**0.089**	**0.088**
	(0.011)	**(0.016)**	**(0.013)**	**(0.015)**	**(0.014)**
Censored latent effect when $x_{t-1} < 0$					
α_2	**-0.732**	**-0.841**	**-0.765**	**-1.108**	**-1.052**
	(0.173)	**(0.281)**	**(0.247)**	**(0.310)**	**(0.296)**
β_2	**-0.166**	**-0.087**	**-0.071**	**-0.088**	**-0.081**
	(0.022)	**(0.019)**	**(0.018)**	**(0.020)**	**(0.020)**
$\sigma^2_{u_2}$	**2.094**	**2.364**	**2.412**	**2.707**	**2.718**
	(0.252)	**(0.397)**	**(0.386)**	**(0.488)**	**(0.483)**
Log-Likelihood	-2077.979	-2118.491	-2118.652	-2120.999	-2119.507

See Table 3 for explanatory details.

Table 15: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the NASDAQ.

y_t	NASDAQ				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	**29.055**	**30.271**	**26.998**	**28.697**	**28.492**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	**16.520**	**34.016**	**40.627**	**35.055**	**38.447**
	(0.006)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \pi_1 = \pi_2 = 0$	3.432	*6.030*	*7.369*	*6.185*	*6.275*
	(0.180)	*(0.049)*	*(0.025)*	*(0.045)*	*(0.043)*
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**740.469**	**851.925**	**854.812**	**890.266**	**892.423**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**406.114**	**311.181**	**316.873**	**287.771**	**286.413**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**13.043**	*4.883*	*5.787*	*6.259*	*6.520*
	(0.000)	*(0.027)*	*(0.016)*	*(0.012)*	*(0.011)*
$H_0 : \beta_1 = -\beta_2$	**22.149**	**7.553**	2.892	**7.678**	*5.610*
	(0.000)	**(0.006)**	(0.089)	**(0.006)**	*(0.018)*

See Table 4 for explanatory details.

Table 16: SR-CLEAR(5) model for absolute returns on the Nikkei-225.

y_t	Nikkei-225				
Period	1/2/1990 – 12/31/1999 (2465 obs.)				
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.236**	**0.216**	**0.205**	**0.214**	**0.200**
	(0.025)	**(0.032)**	**(0.029)**	**(0.036)**	**(0.035)**
π_1	**-0.011**	-0.000	0.001	-0.000	0.000
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
$\sigma^2_{\epsilon_1}$	**0.026**	**0.022**	**0.019**	**0.019**	**0.018**
	(0.004)	**(0.004)**	**(0.003)**	**(0.004)**	**(0.004)**
Censored latent effect when $x_{t-1} > 0$					
α_1	**0.220**	**0.474**	**0.438**	**0.467**	**0.453**
	(0.073)	**(0.075)**	**(0.069)**	**(0.075)**	**(0.072)**
β_1	**0.053**	-0.005	0.001	-0.005	-0.002
	(0.012)	(0.008)	(0.008)	(0.007)	(0.007)
$\sigma^2_{u_1}$	**1.182**	**0.916**	**0.901**	**0.837**	**0.804**
	(0.079)	**(0.066)**	**(0.063)**	**(0.063)**	**(0.059)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.408**	**0.376**	**0.406**	**0.386**	**0.399**
	(0.032)	**(0.035)**	**(0.036)**	**(0.037)**	**(0.038)**
π_2	0.006	0.002	*0.008*	0.003	0.004
	(0.007)	(0.003)	*(0.004)*	(0.003)	(0.003)
$\sigma^2_{\epsilon_2}$	**0.079**	**0.075**	**0.092**	**0.084**	**0.088**
	(0.009)	**(0.010)**	**(0.011)**	**(0.011)**	**(0.011)**
Censored latent effect when $x_{t-1} < 0$					
α_2	**-0.513**	**-0.473**	**-0.611**	**-0.483**	**-0.512**
	(0.117)	**(0.127)**	**(0.141)**	**(0.133)**	**(0.138)**
β_2	**-0.177**	**-0.106**	**-0.112**	**-0.091**	**-0.091**
	(0.015)	**(0.010)**	**(0.011)**	**(0.009)**	**(0.009)**
$\sigma^2_{u_2}$	**2.537**	**2.536**	**2.767**	**2.607**	**2.669**
	(0.184)	**(0.174)**	**(0.200)**	**(0.182)**	**(0.189)**
Log-Likelihood	-2749.145	-2734.542	-2731.466	-2725.423	-2719.553

See Table 3 for explanatory details.

Table 17: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the Nikkei-225.

y_t	Nikkei-225				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	*12.954*	4.750	3.165	4.975	5.573
	(0.024)	(0.447)	(0.675)	(0.419)	(0.350)
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	9.066	7.522	**15.471**	10.751	*11.717*
	(0.106)	(0.185)	**(0.009)**	(0.057)	*(0.039)*
$H_0 : \pi_1 = \pi_2 = 0$	*7.526*	0.381	4.845	0.737	1.889
	(0.023)	(0.827)	(0.089)	(0.692)	(0.389)
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**513.768**	**442.103**	**453.385**	**422.225**	**418.579**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**746.237**	**740.823**	**750.811**	**746.322**	**750.870**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**45.758**	**75.500**	**79.561**	**84.302**	**88.384**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \beta_1 = -\beta_2$	**41.075**	**79.995**	**68.117**	**73.957**	**69.075**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

See Table 4 for explanatory details.

Table 18: SR-CLEAR(5) model for absolute returns on the FTSE-100.

y_t	FTSE-100				
Period	1/2/1990 – 12/31/1999 (2524 obs.)				
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.197**	**0.110**	**0.124**	**0.159**	**0.165**
	(0.028)	**(0.026)**	**(0.037)**	**(0.028)**	**(0.027)**
π_1	-0.003	0.002	0.001	0.001	0.001
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
$\sigma^2_{\epsilon_1}$	**0.024**	**0.011**	*0.014*	**0.021**	**0.022**
	(0.004)	**(0.003)**	*(0.006)*	**(0.004)**	**(0.004)**
Censored latent effect when $x_{t-1} > 0$					
α_1	0.119	**0.349**	**0.312**	**0.265**	**0.246**
	(0.068)	**(0.056)**	**(0.083)**	**(0.064)**	**(0.062)**
β_1	0.021	0.001	0.004	-0.002	-0.003
	(0.012)	(0.005)	(0.006)	(0.006)	(0.006)
$\sigma^2_{u_1}$	**0.587**	**0.429**	**0.447**	**0.507**	**0.529**
	(0.047)	**(0.027)**	**(0.040)**	**(0.038)**	**(0.039)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.244**	**0.401**	**0.429**	**0.459**	**0.470**
	(0.039)	**(0.050)**	**(0.051)**	**(0.053)**	**(0.054)**
π_2	*0.015*	0.006	-0.001	0.004	0.008
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
$\sigma^2_{\epsilon_2}$	**0.034**	**0.087**	**0.089**	**0.105**	**0.106**
	(0.010)	**(0.013)**	**(0.014)**	**(0.015)**	**(0.015)**
Censored latent effect when $x_{t-1} < 0$					
α_2	-0.067	**-0.728**	**-0.851**	**-1.013**	**-1.085**
	(0.115)	**(0.221)**	**(0.249)**	**(0.288)**	**(0.300)**
β_2	**-0.129**	**-0.122**	**-0.138**	**-0.141**	**-0.158**
	(0.017)	**(0.022)**	**(0.025)**	**(0.028)**	**(0.029)**
$\sigma^2_{u_2}$	**0.720**	**1.420**	**1.499**	**1.656**	**1.599**
	(0.076)	**(0.236)**	**(0.263)**	**(0.323)**	**(0.315)**
Log-Likelihood	-1667.413	-1646.912	-1648.767	-1661.233	-1660.300

See Table 3 for explanatory details.

Table 19: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the FTSE-100.

y_t	FTSE-100				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	**23.031**	*11.769*	*11.656*	**22.001**	**23.702**
	(0.000)	*(0.038)*	*(0.040)*	**(0.001)**	**(0.000)**
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	*12.381*	**18.666**	*12.021*	10.286	7.175
	(0.030)	**(0.002)**	*(0.035)*	(0.068)	(0.208)
$H_0 : \pi_1 = \pi_2 = 0$	4.557	1.872	0.184	0.322	1.060
	(0.102)	(0.392)	(0.912)	(0.851)	(0.589)
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**637.718**	**645.646**	**628.676**	**664.212**	**675.386**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**343.323**	**306.789**	**326.053**	**278.747**	**279.255**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	2.195	**17.349**	**15.589**	**12.442**	**11.364**
	(0.138)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.001)**
$H_0 : \beta_1 = -\beta_2$	**26.425**	**28.404**	**27.913**	**25.794**	**29.487**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

See Table 4 for explanatory details.

Table 20: SR-CLEAR(5) model for absolute returns on the DAX Xetra.

y_t	DAX Xetra				
Period	1/2/1990 – 12/31/1999 (2502 obs.)				
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.192**	**0.209**	**0.212**	**0.226**	**0.258**
	(0.021)	**(0.020)**	**(0.023)**	**(0.025)**	**(0.027)**
π_1	0.000	-0.003	*-0.004*	-0.003	-0.003
	(0.003)	(0.002)	*(0.002)*	(0.002)	(0.002)
$\sigma^2_{\epsilon_1}$	**0.023**	**0.027**	**0.032**	**0.033**	**0.041**
	(0.003)	**(0.003)**	**(0.005)**	**(0.007)**	**(0.009)**
Censored latent effect when $x_{t-1} > 0$					
α_1	**0.163**	0.098	0.024	-0.021	-0.135
	(0.058)	(0.063)	(0.081)	(0.097)	(0.125)
β_1	**0.044**	**0.024**	**0.027**	**0.024**	**0.027**
	(0.010)	**(0.005)**	**(0.006)**	**(0.005)**	**(0.005)**
$\sigma^2_{u_1}$	**0.947**	**0.979**	**1.130**	**1.119**	**1.234**
	(0.055)	**(0.058)**	**(0.077)**	**(0.089)**	**(0.116)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.378**	**0.405**	**0.405**	**0.337**	**0.276**
	(0.036)	**(0.043)**	**(0.042)**	**(0.046)**	**(0.045)**
π_2	0.009	**0.019**	*0.010*	*0.008*	0.006
	(0.008)	**(0.005)**	*(0.005)*	*(0.004)*	(0.004)
$\sigma^2_{\epsilon_2}$	**0.074**	**0.097**	**0.068**	**0.053**	**0.047**
	(0.013)	**(0.015)**	**(0.012)**	**(0.011)**	**(0.011)**
Censored latent effect when $x_{t-1} < 0$					
α_2	**-0.603**	**-0.773**	**-0.484**	-0.143	-0.022
	(0.160)	**(0.193)**	**(0.165)**	(0.154)	(0.138)
β_2	**-0.174**	**-0.117**	**-0.108**	**-0.077**	**-0.069**
	(0.018)	**(0.015)**	**(0.014)**	**(0.012)**	**(0.011)**
$\sigma^2_{u_2}$	**2.192**	**2.682**	**2.173**	**2.092**	**1.910**
	(0.230)	**(0.293)**	**(0.217)**	**(0.203)**	**(0.178)**
Log-Likelihood	-2313.140	-2329.426	-2359.105	-2365.418	-2372.957

See Table 3 for explanatory details.

Table 21: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the DAX Xetra.

y_t	DAX Xetra				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	**23.543**	**31.082**	**44.815**	**35.413**	**23.499**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	4.353	**20.442**	3.104	8.011	9.428
	(0.500)	**(0.001)**	(0.684)	(0.156)	(0.093)
$H_0 : \pi_1 = \pi_2 = 0$	1.184	**16.179**	*8.835*	*7.551*	5.010
	(0.553)	**(0.000)**	*(0.012)*	*(0.023)*	(0.082)
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**759.352**	**829.570**	**846.329**	**866.677**	**868.613**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**477.597**	**369.146**	**352.199**	**317.890**	**307.449**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**27.781**	**32.465**	**20.441**	**19.329**	**10.146**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.001)**
$H_0 : \beta_1 = -\beta_2$	**39.952**	**34.780**	**29.598**	**16.072**	**11.282**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.001)**

See Table 4 for explanatory details.

Table 22: SR-CLEAR(5) model for absolute returns on the CAC-40.

y_t	CAC-40				
Period	1/2/1990 – 12/31/1999 (2497 obs.)				
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.208**	**0.187**	**0.205**	**0.185**	**0.186**
	(0.034)	**(0.027)**	**(0.030)**	**(0.030)**	**(0.028)**
π_1	0.003	0.001	0.000	0.001	0.001
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
$\sigma^2_{\epsilon_1}$	**0.021**	**0.016**	**0.021**	**0.020**	**0.017**
	(0.006)	**(0.004)**	**(0.005)**	**(0.005)**	**(0.004)**
Censored latent effect when $x_{t-1} > 0$					
α_1	**0.427**	**0.455**	**0.378**	**0.433**	**0.463**
	(0.069)	**(0.057)**	**(0.067)**	**(0.069)**	**(0.060)**
β_1	0.006	0.005	0.009	0.003	0.003
	(0.009)	(0.005)	(0.005)	(0.004)	(0.004)
$\sigma^2_{u_1}$	**0.803**	**0.783**	**0.842**	**0.871**	**0.836**
	(0.057)	**(0.046)**	**(0.055)**	**(0.057)**	**(0.050)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.336**	**0.383**	**0.357**	**0.389**	**0.325**
	(0.062)	**(0.053)**	**(0.055)**	**(0.046)**	**(0.045)**
π_2	0.004	**0.012**	*0.009*	**0.011**	**0.010**
	(0.007)	**(0.005)**	*(0.004)*	**(0.004)**	**(0.004)**
$\sigma^2_{\epsilon_2}$	**0.041**	**0.044**	**0.041**	**0.035**	**0.033**
	(0.015)	**(0.011)**	**(0.012)**	**(0.008)**	**(0.007)**
Censored latent effect when $x_{t-1} < 0$					
α_2	0.020	0.000	0.088	0.074	0.133
	(0.159)	(0.123)	(0.127)	(0.107)	(0.101)
β_2	**-0.125**	**-0.085**	**-0.076**	**-0.075**	**-0.070**
	(0.016)	**(0.011)**	**(0.011)**	**(0.011)**	**(0.010)**
$\sigma^2_{u_2}$	**1.370**	**1.477**	**1.427**	**1.379**	**1.381**
	(0.143)	**(0.129)**	**(0.129)**	**(0.112)**	**(0.110)**
Log-Likelihood	-2373.929	-2376.619	-2392.231	-2398.406	-2399.014

See Table 3 for explanatory details.

Table 23: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the CAC-40.

y_t	CAC-40				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	7.286	8.598	10.246	*11.724*	10.170
	(0.200)	(0.126)	(0.069)	*(0.039)*	(0.071)
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	3.967	2.385	2.645	4.577	2.241
	(0.554)	(0.794)	(0.755)	(0.470)	(0.815)
$H_0 : \pi_1 = \pi_2 = 0$	1.339	*7.338*	4.087	*7.680*	*8.425*
	(0.512)	*(0.026)*	(0.130)	*(0.021)*	*(0.015)*
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**600.985**	**665.496**	**667.885**	**712.131**	**705.413**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**386.484**	**349.779**	**341.350**	**309.239**	**311.863**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**13.583**	**25.707**	**17.300**	**16.399**	**20.374**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \beta_1 = -\beta_2$	**40.836**	**43.334**	**29.351**	**39.415**	**35.863**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

See Table 4 for explanatory details.

Table 24: SR-CLEAR(5) model for absolute returns on the AEX.

y_t			AEX		
Period		1/2/1990 – 12/31/1999 (2524 obs.)			
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.172**	**0.158**	**0.166**	**0.152**	**0.161**
	(0.020)	**(0.018)**	**(0.018)**	**(0.017)**	**(0.018)**
π_1	**-0.008**	*-0.003*	*-0.003*	*-0.002*	*-0.003*
	(0.003)	*(0.001)*	*(0.002)*	*(0.001)*	*(0.001)*
$\sigma^2_{\epsilon_1}$	**0.021**	**0.021**	**0.024**	**0.022**	**0.023**
	(0.004)	**(0.004)**	**(0.004)**	**(0.003)**	**(0.003)**
Censored latent effect when $x_{t-1} > 0$					
α_1	-0.022	0.024	-0.076	0.037	0.012
	(0.074)	(0.065)	(0.070)	(0.059)	(0.062)
β_1	**0.057**	**0.027**	**0.034**	**0.021**	**0.022**
	(0.010)	**(0.005)**	**(0.005)**	**(0.004)**	**(0.004)**
$\sigma^2_{u_1}$	**0.840**	**0.832**	**0.906**	**0.887**	**0.908**
	(0.061)	**(0.055)**	**(0.062)**	**(0.054)**	**(0.057)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.255**	**0.275**	**0.362**	**0.323**	**0.298**
	(0.027)	**(0.037)**	**(0.036)**	**(0.036)**	**(0.037)**
π_2	0.012	*0.007*	**0.016**	**0.009**	*0.007*
	(0.006)	*(0.003)*	**(0.004)**	**(0.003)**	*(0.003)*
$\sigma^2_{\epsilon_2}$	**0.029**	**0.032**	**0.054**	**0.040**	**0.037**
	(0.005)	**(0.007)**	**(0.008)**	**(0.008)**	**(0.008)**
Censored latent effect when $x_{t-1} < 0$					
α_2	*-0.174*	*-0.282*	**-0.560**	**-0.484**	*-0.414*
	(0.086)	*(0.140)*	**(0.163)**	**(0.169)**	*(0.169)*
β_2	**-0.162**	**-0.090**	**-0.100**	**-0.092**	**-0.087**
	(0.013)	**(0.012)**	**(0.014)**	**(0.014)**	**(0.014)**
$\sigma^2_{u_2}$	**1.201**	**1.639**	**1.879**	**1.815**	**1.764**
	(0.099)	**(0.173)**	**(0.216)**	**(0.211)**	**(0.206)**
Log-Likelihood	-1935.192	-1984.639	-1989.760	-2000.218	-2006.205

See Table 3 for explanatory details.

Table 25: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the AEX.

y_t			AEX		
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	**28.913**	**36.305**	**42.809**	**55.217**	**57.789**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	7.710	6.597	*13.385*	3.736	3.616
	(0.173)	(0.252)	*(0.020)*	(0.588)	(0.606)
$H_0 : \pi_1 = \pi_2 = 0$	**11.845**	**10.313**	**19.840**	**11.045**	**10.043**
	(0.003)	**(0.006)**	**(0.000)**	**(0.004)**	**(0.007)**
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**841.321**	**834.473**	**870.829**	**870.513**	**881.616**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**470.297**	**404.229**	**386.846**	**373.061**	**363.771**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**9.619**	**19.756**	**18.751**	**18.102**	**16.039**
	(0.002)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \beta_1 = -\beta_2$	**40.388**	**23.101**	**18.384**	**24.151**	**20.197**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

See Table 4 for explanatory details.

Table 26: SR-CLEAR(5) model for absolute returns on the Hang Seng.

y_t			Hang Seng		
Period		1/2/1990 – 12/31/1999 (2478 obs.)			
x_t	1-50	1-150	5-150	1-200	2-200
Linear when $x_{t-1} > 0^{1)}$					
μ_1	**0.335**	**0.251**	**0.276**	**0.257**	**0.265**
	(0.030)	**(0.023)**	**(0.033)**	**(0.025)**	**(0.026)**
π_1	-0.004	-0.001	-0.001	-0.001	-0.001
	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)
$\sigma^2_{\epsilon_1}$	**0.062**	**0.036**	**0.047**	**0.042**	**0.043**
	(0.008)	**(0.005)**	**(0.010)**	**(0.006)**	**(0.007)**
Censored latent effect when $x_{t-1} > 0$					
α_1	*-0.280*	-0.054	-0.190	-0.034	-0.085
	(0.124)	(0.087)	(0.138)	(0.093)	(0.098)
β_1	**0.056**	**0.033**	**0.039**	**0.024**	**0.027**
	(0.011)	**(0.005)**	**(0.006)**	**(0.004)**	**(0.005)**
$\sigma^2_{u_1}$	**2.345**	**1.814**	**2.062**	**1.889**	**1.904**
	(0.193)	**(0.117)**	**(0.183)**	**(0.127)**	**(0.132)**
Linear when $x_{t-1} < 0^{1)}$					
μ_2	**0.397**	**0.484**	**0.440**	**0.431**	**0.441**
	(0.041)	**(0.053)**	**(0.054)**	**(0.052)**	**(0.055)**
π_2	0.008	*0.010*	0.008	0.006	0.008
	(0.007)	*(0.005)*	(0.005)	(0.005)	(0.005)
$\sigma^2_{\epsilon_2}$	**0.122**	**0.159**	**0.141**	**0.136**	**0.152**
	(0.023)	**(0.024)**	**(0.025)**	**(0.026)**	**(0.029)**
Censored latent effect when $x_{t-1} < 0$					
α_2	**-1.160**	**-1.516**	**-1.334**	**-1.427**	**-1.465**
	(0.251)	**(0.301)**	**(0.299)**	**(0.314)**	**(0.335)**
β_2	**-0.212**	**-0.148**	**-0.138**	**-0.134**	**-0.132**
	(0.019)	**(0.016)**	**(0.016)**	**(0.015)**	**(0.016)**
$\sigma^2_{u_2}$	**5.194**	**6.641**	**6.438**	**6.831**	**7.051**
	(0.541)	**(0.736)**	**(0.712)**	**(0.775)**	**(0.832)**
Log-Likelihood	-2958.626	-2915.901	-2949.592	-2934.147	-2934.408

See Table 3 for explanatory details.

Table 27: Tests on parameter restrictions in the SR-CLEAR(5) model for absolute returns on the Hang Seng.

y_t	Hang Seng				
x_t	1-50	1-150	5-150	1-200	2-200
$H_0 : \rho_{1,1} = \ldots = \rho_{1,p} = 0$	10.129	*11.735*	10.916	**15.244**	*13.959*
	(0.072)	*(0.039)*	(0.053)	**(0.009)**	*(0.016)*
$H_0 : \rho_{2,1} = \ldots = \rho_{2,p} = 0$	**15.496**	**26.607**	**15.335**	*15.024*	**23.871**
	(0.008)	**(0.000)**	**(0.009)**	*(0.010)*	**(0.000)**
$H_0 : \pi_1 = \pi_2 = 0$	3.609	4.780	3.111	1.944	3.089
	(0.165)	(0.092)	(0.211)	(0.378)	(0.213)
$H_0 : \alpha_1 = \beta_1 = \sigma^2_{u_1} = 0^{1)}$	**962.968**	**1062.256**	**1077.943**	**1056.303**	**1065.254**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \alpha_2 = \beta_2 = \sigma^2_{u_2} = 0^{1)}$	**647.606**	**484.537**	**482.283**	**458.436**	**453.450**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \sigma^2_{u_1} = \sigma^2_{u_2}$	**24.614**	**41.960**	**35.438**	**39.556**	**37.327**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
$H_0 : \beta_1 = -\beta_2$	**48.912**	**47.230**	**34.459**	**47.267**	**39.933**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

See Table 4 for explanatory details.

References

Beran, J. (1994), *Statistics for Long-memory Processes.* London: Chapman and Hall.

Bos, C., P.H. Franses and M. Ooms (1999), Long memory and level shifts: Re-analyzing inflation rates, *Empirical Economics*, 24, 427-450.

Braun, Ph., D. Nelson and A. Sunier (1995), Good news, bad news, volatility and betas, *Journal of Finance*, 50, 1575-1603.

Brock W., J. Lakonishok and B. LeBaron (1992), Simple technical trading rules and the stochastic properties of stock returns, *Journal of Finance*, 47, 1731-1764.

Christie, A. (1982), The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of Financial Economics*, 3, 407-432.

Clements, M. and D. Hendry (1993), On the limitations of comparing mean squared forecast errors, *Journal of Forecasting*, 12, 617-637.

Diebold, F.X. and A. Inoue (2001), Long memory and regime switching, *Journal of Econometrics*, 105, 131-159.

Ding, Z., C.W.J. Granger and R.F. Engle (1993), A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, 1, 83-106.

Doornik, J.A. (1999), *Object-Oriented Matrix Programming Using Ox*, 3rd ed. London: Timberlake Consultants Press

Doornik, J.A. and Ooms, M. (2001), *A Package for Estimating, Forecasting and Simulating Arfima Models: Arfima Package 1.01 for Ox*, Oxford: www.nuff.ox.ac.uk/Users/Doornik.

Fletcher, R. (1987), *Practical Methods of Optimization*, 2nd ed. New York: John Wiley & Sons.

Franses, P.H. and K. van Griensven (1998), Forecasting exchange rates using neural networks for technical trading rules, *Studies in Nonlinear Dynamics and Econometrics*, 2, 109-114.

Franses, P.H. and R. Paap (2001), Censored Latent Effects Autoregression, with an Application to US Unemployment, *Journal of Applied Econometrics*, to appear.

Gençay, R. (1996), Non-linear prediction of security returns with moving average rules, *Journal of Forecasting*, 15, 165-174.

Glosten. L., R. Jagannathan and D. Runkle (1992), On the relation between the expected value and the volatility of the nominal excess return of stocks, *Journal of Finance*, 48, 1779-1801.

Granger, C.W.J., N. Hyung (2000), Occasional Structural Breaks and Long Memory, revised version of *UCSD Working Paper* 99-14.

Hyung, N. and P.H. Franses (2001), Structural breaks and long memory in US inflation rates: Do they matter for forecasting?, *Econometric Institute Research Report*, EI2001-13, Erasmus University Rotterdam.

Schwert, W. (1989), Why does stock market volatility change over time?, *Journal of Finance*, 44, 1115-1153.

Teräsvirta, T. (1994), Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models, *Journal of the American Statistical Association*, 89, 208-218.

Van Dijk, D., P.H. Franses and R. Paap (2001), A nonlinear long memory model for US unemployment, *Journal of Econometrics*, to appear

Wolak, F. (1989), Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models, *Econometric Theory*, 5, 1-35.