Mixture Gaussian Time Series Modelling of Long-Term Market Returns

Albert C.S. Wong, FSA, PhD
Department of Finance, The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
Telephone: 2609 7648
Fax: 2603 6586
E-mail: albertw@baf.msmail.cuhk.edu.hk

Wai Sum Chan, FSA, PhD
Department of Statistics and Actuarial Science, The University of Hong Kong
Pokfulam Road, Hong Kong
Telephone: 2857 8318
Fax: 2585 9041
E-mail: chanws@hku.hk

Mixture Gaussian Time Series Modelling of Long-Term Market Returns

Abstract

Stochastic modelling of investment returns is an important topic for actuaries who deal with variable annuity and segregated fund investment guarantees. The traditional lognormal stock return model is simple, but it is generally less appealing for longer-term problems. In recent years, the use of regime switching lognormal (RSLN) processes for modelling maturity guarantees has been gaining popularity. However, it is unclear whether this class of models is able to fit the tails of the stock return distributions well. In this paper, we introduce the class of mixture Gaussian time series processes for modelling long-term stock market returns. The mixture time series process is extremely flexible in modelling tails of return distributions. Monthly data from the Toronto Stock Exchange 300 and the Standard and Poor's 500 indices are used to illustrate the mixture time series modelling procedures. The fits of the mixture models to the data are compared to the lognormal and RSLN models. Efficient method of generating large numbers of return scenarios using the fitted mixture models is also discussed.

1. Introduction

Stochastic equity return modelling plays an important role in measuring the obligations created by segregated fund investment guarantees in Canada. The March 2002 final report of the CIA Task Force on Segregated Fund Investment Guarantees (TFSFIG) provides useful guidance for appointed actuaries applying stochastic techniques to value segregated fund guarantees in a Canadian GAAP valuation environment. The full report of the Task Force (TFSFIG, 2002) is published by the Canadian Institute of Actuaries and is available from the CIA website address given in the reference section of this paper.

In the United States, the Life Capital Adequacy Subcommittee (LCAS) of the American Academy of Actuaries issued the C-3 Phase II Risk-Based Capital (RBC) proposal in December 2002. The full report of the Subcommittee (LCAS, 2002) is published by the American Academy of Actuaries and is available from the Academy's website. In addition to the interest rate risk for interest-sensitive products, the C-3 Phase II report also addresses the equity risk exposure inherent in variable products with guarantees, such as (1) guaranteed minimum death benefits (GMDBs); (2) guaranteed minimum income benefits (GMIBs); and guaranteed minimum accumulation benefits (GMABs). Stochastic scenario analysis is recommended to determine minimum capital requirements for these variable products. The actuary performing the analysis has the responsibility of developing and validating scenario models. Again, stochastic equity return modelling plays an important role in these analyses. The chosen equity return generator must be calibrated to specified distribution percentiles to ensure sufficiently fat tails. These percentiles, in the United States, are based on historical experience of the S&P 500 total return index as a proxy for returns on broadly diversified U.S. equity fund.

There are a large number of potential stochastic models for equity returns. Regulators, both in the U.S. and Canada, do not restrict the use of any model that reasonably fits the historical data. We shall first describe the traditional lognormal stock return model. The lognormal model has a long and illustrious history, and has become "the workhorse of the financial asset pricing literature" (Campbell *et al.*, 1997, p.16). Most of the empirical studies use monthly data. Define the one-period *simple net return* as

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1}$$

where P_t is the end-of-month stock price (or market index), with dividends reinvested. The *continuously compounded return* or *log return* Y_t is defined to be the natural logarithm of its gross return:

$$Y_t = \log(1 + R_t) = \log(P_t) - \log(P_{t-1}). \tag{2}$$

The lognormal model assumes that log returns Y_t are independently and identically distributed (IID) normal with mean μ and variance σ^2.

The independent lognormal (ILN) model is simple, scalable and tractable. But as attractive as the lognormal model is, it is not consistent with all properties of historical equity returns. At short horizons, observed returns have negative skewness and strong evidence of excess kurtosis with time varying volatility (Panjer *et al.*, 1998, p.438; and Campbell *et al.*, 1997, p.379).

The class of Markov regime switching models (Hamilton, 1990 and 1994) has been widely employed in the literature to explain various empirical phenomena in an observed economic time series. Hardy (2001) proposes using regime switching lognormal (RSLN) processes for modelling monthly equity returns. She shows that the class of RSLN models fits the TSE 300 and S&P 500 monthly total returns much better than the ILN does. The RSLN model is defined as

$$Y_t = \mu_{S_t} + \sigma_{S_t}\, \varepsilon_t \tag{3}$$

where $S_t = 1, 2, \ldots, k$ denotes the unobservable state indicator which follows an ergodic k-state Markov process and ε_t is a standard normal random variable which is IID over time. In most situations, $k = 2$ or 3 (that is two- or three-regime models) is sufficient for modelling monthly equity returns (Hardy, 2001). The stochastic transition probabilities that determine the evolution in S_t is given by

$$Pr\{S_{t+1} = j \quad | \quad S_t = i\} = p_{ij},$$
$$0 \leq p_{ij} \leq 1, \qquad \sum_{j=1}^{k} p_{ij} = 1 \quad \text{for all} \quad i,$$

so that the states follow a homogenous Markov chain.

The use of RSLN processes for modelling maturity guarantees has been gaining popularity. However, it is unclear whether this class of models is able to fit the tails of the stock return distributions well. In this paper, we introduce the class of mixture Gaussian time series processes for modelling long-term stock market returns. The mixture time series process is extremely flexible in modelling tails of return distributions. The remaining of this paper is organised as follows. Section 2 introduces the basic mixture Gaussian autoregressive (MAR) model. Extension of the basic model to include a component for autoregressive conditional heteroscedasticity (ARCH) is also discussed. Section 3 illustrates the MAR modelling procedures using monthly data from the Toronto Stock Exchange 300 and the Standard and Poor's 500 total return indices. The fits of the mixture models to the data are compared to the ILN and RSLN models. Finally, Section 4 provides some concluding remarks.

2. The Models

2.1 The MAR Model

Wong and Li (2000) introduces the class of mixture autoregressive (MAR) models. The model is basically a mixture of K Gaussian autoregressive (AR) components, defined by

$$F(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \Phi \left(\frac{y_t - \phi_{k0} - \phi_{k1} y_{t-1} - \ldots - \phi_{kp_k} y_{t-p_k}}{\sigma_k} \right). \tag{4}$$

We denote this model by MAR$(K; p_1, p_2, \ldots, p_K)$. Here $F(y_t|\mathcal{F}_{t-1})$ is the conditional cumulative distribution function of Y_t given the past information, evaluated at y_t; $\mathcal{F}_t$

is the information set up to time t; $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution; $\alpha_1 + \ldots + \alpha_K = 1$, $\alpha_k > 0, k = 1, \ldots, K$. Note that the mixture of independent normal models is a special case of the MAR model with $p_1 = p_2 = \ldots = p_K = 0$.

Several interesting properties make the MAR model a promising candidate for financial time series modelling. As the conditional means of the components depend on past values of the time series, the shape of the conditional distributions of the series changes over time. The conditional distributions can be changed from short-tailed to fat-tailed, or from unimodal to multimodal.

Another important feature of the MAR model is the ability to model changing conditional variance. The conditional variance of y_t, which is dependent on the conditional means of the components, is given by

$$\text{Var}(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \sigma_k^2 + \sum_{k=1}^{K} \alpha_k \mu_{k,t}^2 - \left(\sum_{k=1}^{K} \alpha_k \mu_{k,t} \right)^2 .$$

The term $\sum \alpha_k \mu_{k,t}^2 - (\sum \alpha_k \mu_{k,t})^2$ is non-negative and is zero only if $\mu_{1,t} = \mu_{2,t} = \ldots = \mu_{K,t}$. The conditional variance is large when the $\mu_{k,t}$'s differ greatly. The conditional variance is smallest when the $\mu_{k,t}$'s are all the same. We call this smallest possible conditional variance, $\sum \alpha_k \sigma_k^2$, the baseline conditional variance or volatility.

As the MAR model is a mixture of autoregressive models, it is expected that the range of autocorrelations of the time series generated by the MAR model should be similar to those of an AR model. It is easy to show that the autocorrelation at lag j, ρ_j, is given by

$$\rho_j = \sum_{i=1}^{p} \left(\sum_{k=1}^{K} \alpha_k \phi_{ki} \right) \rho_{|j-i|} \qquad j = 1, \ldots, p,$$

where $p = \max(p_1, \ldots, p_K)$. Note that these equations are similar to the Yule-Walker equations for the ordinary AR(p) process, where the coefficient $\sum_{k=1}^{K} \alpha_k \phi_{ki}$ replaces the lag i coefficient of the AR(p) process.

2.2 The MARCH model

Despite the above interesting properties, the MAR model suffers from the following limitation in the modelling of nonlinear time series. The squared autocorrelation structure of the MAR model is quite simple and is analogous to that of the AR model (Granger and Newbold, 1986, p.309 and Wong, 1998). The absence of a squared autocorrelation structure of the MAR models is a shortfall which limits its application to financial data especially.

Wong and Li (2001) introduces the class of mixture autoregressive conditional heteroscedastic (MARCH) models. The MARCH model is a generalisation of the MAR process. This model consists of a mixture of K autoregressive components with autoregressive conditional heteroscedasticity, i.e., the conditional mean of y_t follows an AR process while the conditional variance of y_t follows an ARCH process (Engle, 1982). The MARCH models are defined by

$$\begin{aligned} F\left(y_t|\mathcal{F}_{t-1}\right) &= \sum_{k=1}^{K} \alpha_k \Phi\left(\frac{e_{k,t}}{\sqrt{h_{k,t}}}\right), \\ e_{k,t} &= y_t - \phi_{k0} - \phi_{k1} y_{t-1} - \ldots - \phi_{kp_k} y_{t-p_k}, \\ h_{k,t} &= \beta_{k0} + \beta_{k1} e^2_{k,t-1} + \ldots + \beta_{kq_k} e^2_{k,t-q_k}. \end{aligned} \tag{5}$$

We denote this model by MARCH$(K; p_1, p_2, \ldots, p_K; q_1, q_2, \ldots, q_K)$. To avoid the possibility of zero or negative conditional variance, the following conditions for β_{ki}'s must be imposed: $\beta_{k0} > 0$ $(k = 1, \ldots, K), \beta_{ki} \geq 0$ $(i = 1, \ldots, q_k; k = 1, \ldots, K)$. Note that the MAR model is a special case of the MARCH model with $q_1 = q_2 = \ldots = q_K = 0$.

One additional feature of the MARCH model is its greater flexibility in the modelling of changing conditional variance. The conditional variance of y_t is given by

$$\mathrm{Var}(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k h_{k,t} + \sum_{k=1}^{K} \alpha_k \mu^2_{k,t} - \left(\sum_{k=1}^{K} \alpha_k \mu_{k,t}\right)^2.$$

The first term allows the modelling of the dependence of the conditional variance on the past "errors". The second and third terms model the change of the conditional variance due to the difference in the conditional means of the components.

The squared autocorrelations of the time series generated by a MARCH model are similar to those of an ARCH model. As an example, for a MARCH(K;0,...,0;1,...,1) model with $\phi_{k0} = 0$ ($k = 1, \ldots, K$), the autocorrelations of the squared time series are given by

$$\mathrm{corr}(Y_t^2, Y_{t-l}^2) \;=\; \left(\sum \alpha_k \beta_{k1}\right)^l .$$

Note that the squared autocorrelation function is similar to those of an ARCH(1) model with the lag 1 coefficient replaced by the coefficient $\sum \alpha_k \beta_{k1}$. As a generalisation of the ARCH model, the range of possible squared autocorrelations should be as great as that of the corresponding standard ARCH process. In the case of a Markov switching generalised ARCH process, Francq and Zakoian (1999) showed that the squared process behaves like a special ARMA model.

2.3 Model Estimation

The estimation of the parameters of the MARCH model and the MAR model can be done by maximum (conditional) likelihood method. The EM algorithm (Dempster, Laird and Rubin, 1977) for maximising the log-likelihood, which is the most readily available procedure in estimating mixture type models, is employed. One advantage of the EM algorithm is that it ensures that the likelihood values increase monotonically. See McLachlan and Basford (1988), and McLachlan and Krishnan (1997) for a discussion of the EM algorithm and other alternatives. The standard errors of the parameter estimates can be computed by the Louis' method (1982), after the EM estimation. The details of the EM estimation methods for the MAR and MARCH models are discussed in Wong and Li (2000, 2001).

2.4 Model Selection

There are two aspects of model selection in the MARCH models, namely, the number of components and the order of each AR-ARCH component. Here, we do not discuss the selection problem for the number of components, K, as it is difficult to handle even in the special case of the MAR model (Wong and Li, 2000). The use of Bayesian information criterion (BIC) (Schwarz, 1978) to choose K is somewhat non-standard as it corresponds to testing problems with nuisance parameters that do not exist under the null hypothesis (Davis, 1977, 1987). However, a two-component MARCH model should be sufficient in most applications.

After the number of components K has been decided, the BIC can be used for the selection of the orders, p_K and q_K, of each AR-ARCH components. Wong and Li (2000, 2001) illustrate the performance of the minimum BIC procedure with simulation studies. They find that the minimum BIC procedure has a good performance. Moreover, they also find that the minimum AIC procedure (Akaike, 1973) is not appropriate for the model selection problem of the MARCH models.

3. Empirical Results

3.1 TSE 300 Total Return Series

We consider monthly returns on the TSE 300 index, with dividends reinvested, from January 1956 to December 1999, giving a total of 527 log return observations. The dataset has been analysed by Hardy (2001) and it is listed as Appendix B in the March 2002 final report of the CIA Task Force on Segregated Fund Investment Guarantees.

Descriptive statistics of the data series are first computed. Table 1 (the second column) gives the sample mean, standard deviation, coefficient of skewness, coefficient of excess kurtosis, minimum, lower 97.5, 95 and 90 percentiles, first 8 sample autocorrelations of the series, and first 8 sample autocorrelations of the squared series. The TSE 300 data are negatively skewed with a large coefficient of excess kurtosis (3.91). The minimum value observed in October 1987, when the TSE 300 log return was -0.2552. The series has a marginally significant first-order autocorrelation and a significant lag-2 serial correlation for the squared observations.

The independent lognormal (ILN) model, the two-regime RSLN model (RSLN2) and the two-point mixture of independent normal distributions (MIND2) that are considered in Hardy (2001) and Klein (2002) are fitted to the TSE data. The maximum likelihood estimated model parameters are given in Table 2. In order to capture the first-order autocorrelation and the second-order squared autocorrelation in the data, both the MAR(2;1,0) model and the MARCH(2;1,0;2,0) model are also considered. These two models are also preferred by the BIC within the class of MAR-type processes. The conditional maximum likelihood estimates of these models are summarised in Table 2.

Comparison of the "goodness-of-fit" of different types of fitted models to the data is not a trivial task. Little is known about the power of the model selection criteria, such as AIC and BIC, for discriminating classes of stochastic models that are not embedded. The likelihood ratio test could be problematic for testing regime-switching-type of models, because of the possible nuisance parameter situation (Chan and Tong, 1986 and 1990). Furthermore, simulations show that these standard model selection criteria could be sensitive to the effective number of observations (Ng and Perron, 2003). The log-likelihoods, AIC and BIC values of the various models are reported in Table 1. Their normalized (by the sample size) values are also given.

In this paper, we compare the fitted models by examining their model characteristics, such as moments; percentiles; autocorrelations; and squared autocorrelations. These characteristics will be compared to those of the data. The characteristics of the fitted models are obtained through simulations. For each fitted model, one million scenarios are generated. Each scenario is a simulated log return time series of 527 observations. For models that require starting values, long-term sample averages are used for initialisation. For example, the starting regime in the RSLN2 model is chosen randomly according to the unconditional probabilities (π_1 and π_2) of being in each regime. The averages of sample characteristics for the one million generated scenarios are calculated for each fitted model. The probability of the fitted model would generate a value at least as small as the October 1987 TSE Crash, i.e.,

$$Pr(Crash) = Pr\Big\{ \min_{1 \le t \le 527} Y_t \le -0.2552 \Big\}, \tag{6}$$

is also computed empirically. The results are recorded in Table 1.

The ILN model generally does not provide a good fit to the historic monthly TSE 300 total return series. The model skewness and excess kurtosis are 0.00044 and −0.01228, respectively. They deviate enormously from the observed sample skewness (−0.90987) and excess kurtosis (3.91362). Furthermore, under the ILN model, it would be almost impossible to produce an observation as small as the October 1987 Crash value.

The RSLN2 model has a coefficient of skewness of −0.55946 and a coefficient of excess kurtosis of 2.48449. Both values are 36-38% less than the observed figures. On the average, the RSLN2 model produces the minimum value of −0.2030 out of the 527 observations. It is somewhat 20% less than the observed minimum (−0.2552) in the data. Under the RSLN2 model, there is a 7.84% chance of producing a value as small as the observation in the 1987 Crash. This result is consistent with the analysis in Hardy (2001, p.47). Finally, we observe that the first two lags of the squared autocorrelations of the RSLN2 model are significant. It implies that the model is able to produce volatility clustering (i.e., serial correlations of conditional variances).

The skewness and excess kurtosis of the MIND2 model are −0.7200 and 3.01036, respectively. They are much closer to the observed values as compared to the RSLN2 model. The average of the minimum values generated by the MIND2 model is −0.2238 and its corresponding probability of producing the Crash is around 21%.

The characteristics produced by the MAR(2;1,0) model are reasonably close to the data as compared to other models. The only shortfall of the model would be its inability to explain the second-order volatility clustering observed in the TSE data. On the other hand, the MARCH(2;1,0;2,0) process models explain well the lag-2 volatility clustering at the expense of suffering a 10% downward bias in mean.

Our limited experience of stochastic modelling of the TSE data shows that there is no single model can be identified as superior to all others. The MAR-type models introduced in this paper could explain better the skewness, excess kurtosis and outliers observed in the data.

3.2 S&P 500 Total Return Series

Monthly S&P 500 total return data from 1956 to 1999 ($n = 527$) are considered. We choose the same data period as the TSE series for ease of comparison. Table 3 (the second column) reports the sample statistics computed from the data series. Analogous to the TSE series, the S&P data are negatively skewed with a large coefficient of excess kurtosis (3.01). The minimum value observed in October 1987, when the S&P 500 log return was -0.2425. The first three lags of the squared autocorrelations are fairly large. They indicate that some low-order volatility clustering effects exist in the data.

As in the previous section, the ILN, RSLN2 and MIND2 models are fitted. In addition, two MARCH-type models are selected by the BIC. They are the MARCH(2;0,2;0,0) model and the MARCH(2;0,3;0,0) model. The estimation results of these models are given in Table 4.

We compare the characteristics of the fitted models and those of the data. Again, there is no single model found to be superior to all others in describing the S&P 500 data. Compared to the RSLN2 model, the MARCH(2;0,3;0,0) process could explain better the kurtosis and the farther end of the left tail (i.e., crashes) of the data.

4. Concluding Remarks

The main objective of this paper is not to reject the use of any class of stochastic processes for modelling the dynamics in the TSE and S&P data. Instead, we suggest widening the set of candidate models. We have shown that the class of MARCH models provides a better capture of the kurtosis (thickness of the tails) and extreme observations (such as October 1987) in the TSE and S&P data. The MARCH processes can also model volatility clustering in the data flexibly.

The presence of outliers (crashes) is a general phenomenon across stock markets (Sornette, 2003). In the standard time series outlier literature (e.g., Tsay 1986 and 1988), a "contaminated" time series is modelled as an autoregressive moving-average (ARMA) process plus outlier components. The time series can be "cleaned" (i.e., adjusted) by iterating the detection-estimation-adjustment cycles (Chen and Liu, 1993).

In actuarial applications, whether or not it is appropriate to adjust the data for the outliers depends on the purpose to which the model so derived will be used. If the model will be used in an application for which extreme stochastic fluctuations are less important (e.g. to ensure that premiums are adequate in most but not extreme scenarios) then it may be preferable to use a model based on outlier-adjusted data. If however the model will be used in an application for which extreme stochastic fluctuations are important (such as ensuring that investment guarantee reserves are sufficient to keep an insurance company solvent in all but the most extreme scenarios), then a model which is sympathetic to outliers in the data ought to be used. The class of MARCH models, which is extremely flexible in modelling tails of return distributions, might be useful to actuaries in this situation.

Finally, efficient algorithm for generating large number of scenarios using the MARCH model has been developed. A sample S-plus (Krause and Olson, 2002) programme is given in the Appendix.

5. References

Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle," in *The Second International Symposium on Information Theory* (eds B. N. Petrov and F. Csake), pp. 267-281. Hungary: Akademiai Kiado.

Campbell, J. Y., Lo., A. W., and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

Chan, K. S., and Tong, H. (1990). "On Likelihood Ratio Tests for Threshold Autoregression," *Journal of the Royal Statistical Society*, B52, 496-476.

Chan, W. S., and Tong, H. (1986). "On Tests for Nonlinearity in Time Series Analysis," *Journal of Forecasting*, 5, 217-228.

Chen, C., and Liu, L. M. (1993), "Joint Estimation of Model Parameters and Outlier Effects in Time Series," *Journal of the American Statistical Association*, 88, 284-297.

Davies, R. B. (1977), "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64, 247-254.

Davies, R. B. (1987), "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74, 33-43.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm" (with discussion), *Journal of Royal Statistical Society*, Ser. B, 39, 1-38.

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

Francq, C., and Zakoian, J. M. (1999), "Linear-representations Based Estimation of Switching-regime GARCH Models," in the *Proceedings of the 1999 NBER/NSF Time Series Conference*, Taipei.

Granger, C. W. J., and Newbold, P. (1986), *Forecasting Economic Time Series*, San Diego: Academic Press.

Hamilton, J. (1990), "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics*, 45, 39-70.

Hamilton, J. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

Hardy, M. (2001), "A Regime-Switching Model of Long-Term Stock Returns," *North American Actuarial Journal*, 5, 41-53.

Klein, G. E. (2002), Discussion on "A Regime-Switching Model of Long-Term Stock Returns," *North American Actuarial Journal*, 6, 171-176.

Krause, A., and Olson, M. (2002), *The Basics of S-Plus*, New York: Springer.

Life Capital Adequacy Subcommittee (2002), *Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Products with Guarantees (Excluding Index Guarantees)*, American Academy of Actuaries.

`http://www.actuary.org/pdf/life/rbc_16dec02.pdf`

Louis, T. A. (1982), "Finding the Observed Information Matrix when Using the EM Algorithm," *Journal of Royal Statistical Society*, Ser. B, 44, 226-233.

McLachlan, G. J., and Basford, K. E. (1988), *Mixture Models: Inference and Applications to Clustering*, New York: Marcel Dekker.

McLachlan, G. J., and Krishnan, T. (1997), *The EM Algorithm and Extensions*, New York: Wiley.

Ng, S., and Perron, P. (2003), "A Note on the Selection of Time Series Models", Working Paper, Department of Economics, Johns Hopkins University, USA.

`http://server.econ.jhu.edu/people/ng//papers/aic.pdf`

Panjer, H. (1998), *Financial Economics: With Applications to Investments, Insurance, and Pensions*, Schaumburg: Actuarial Foundation.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461-464.

Sornette, D. (2003), *Why Stock Markets Crash: Critical Events in Complex Financial Systems*, Princeton: Princeton University Press.

Task Force on Segregated Fund Investment Guarantees (2002), *The Final Report*, Canadian Institute of Actuaries.

`http://www.actuaries.ca/publications/2002/202012e.pdf`

Tsay, R. S. (1986), "Time Series Model Specification in the Presence of Outliers," *Journal of the American Statistical Association*, 81, 132141.

Tsay, R. S. (1988), Outliers, Level Shifts and Variance Changes in Time Series," *Journal of Forecasting*, 7, 1-20.

Wong, C. S., and Li, W. K. (2000), "On a Mixture Autoregressive Model," *Journal of the Royal Statistical Society*, Ser. B, 62, 95-115.

Wong, C. S., and Li, W. K. (2001), "On a Mixture Autoregressive Conditional Heteroscedastic Model," *Journal of American Statistical Association*, 96, 982-995.

Wong, C. S. (1998), *Statistical Inference for Some Nonlinear Time Series Models*, PhD Thesis, The University of Hong Kong.

Appendix

```
######### Simulation of the MARCH(2;1,0;2,0) Model for the
######### TSE 300 Total Return Series

tse300.sim <- rep(0,527)
tse300.sim[1] <- log(tse300[2]/tse300[1])
tse300.sim[2] <- log(tse300[3]/tse300[2])
tse300.sim[3] <- log(tse300[4]/tse300[3])
for (i in 4:527) {
    p <- runif(1)
    if (p <= 0.9508) {
        error1 <- tse300.sim[i-1] - 0.0107 - 0.0612*tse300.sim[i-2]
        error2 <- tse300.sim[i-2] - 0.0107 - 0.0612*tse300.sim[i-3]
        h <- 0.0011 + 0.0740*error1^2 + 0.1348*error2^2
        tse300.sim[i] <- 0.0107 + 0.0612*tse300.sim[i-1] + sqrt(h)*rnorm(1)
    }
    else
        tse300.sim[i] <- -0.0671 + sqrt(0.0065)*rnorm(1)
}
tse300.sim
```

Table 1 Comparison of the TSE 300 Data and the Fitted Models

		Data		*ILN*		*RSLN2*		*MIND2*		*MAR(2;1,0)*		*MARCH(2;1,0;2,0)*	
					diff(%)		diff(%)		diff(%)		diff(%)		diff(%)
	Mean	0.00814		0.00813	**-0.1%**	0.00809	**-0.6%**	0.00813	**-0.1%**	0.00807	**-0.8%**	0.00731	**-10.2%**
	Std Dev	0.04511		0.04510	**0.0%**	0.04496	**-0.3%**	0.04502	**-0.2%**	0.04502	**-0.2%**	0.04527	**0.3%**
	Skewness	-0.90987		0.00044	**-100%**	-0.55946	**-38.5%**	-0.72000	**-20.9%**	-0.72751	**-20.0%**	-0.77711	**-14.6%**
	Kurtosis	3.91362		-0.01228	**-100%**	2.48449	**-36.5%**	3.01036	**-23.1%**	3.26681	**-16.5%**	3.04879	**-22.1%**
Y(1)	**Minimum**	-0.25519		-0.12954	**-49.2%**	-0.20299	**-20.5%**	-0.22378	**-12.3%**	-0.23205	**-9.1%**	-0.22608	**-11.4%**
Y(13)	**L97.5%**	-0.09359		-0.08118	**-13.3%**	-0.09494	**1.4%**	-0.09036	**-3.4%**	-0.09567	**2.2%**	-0.09401	**0.5%**
Y(26)	**L95.0%**	-0.06530		-0.06674	**2.2%**	-0.06700	**2.6%**	-0.06415	**-1.8%**	-0.06683	**2.3%**	-0.06617	**1.3%**
Y(52)	**L90.0%**	-0.04533		-0.05024	**10.8%**	-0.04438	**-2.1%**	-0.04416	**-2.6%**	-0.04493	**-0.9%**	-0.04504	**-0.6%**
	Pr(Crash)			0.00000		0.07840		0.21010		0.28522		0.22450	
	log L			885.6		922.7		913.4		912.5		918.4	
	AIC			883.6		917.7		908.4		906.5		910.4	
	SBC			879.4		903.9		897.8		893.7		893.4	
	Effective # Obs			527		527		527		526		524	
	Normalized log L			885.6		922.7		913.4		914.2		923.7	
	Normalized AIC			883.6		917.7		908.4		908.2		915.6	
	Normalized SBC			879.4		903.9		897.8		895.4		898.5	
Autocorrelations													
(a) Raw Series													
	Lag	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**
	1	0.0813	1.97	-0.0019	-0.04	0.0336	0.74	-0.0019	-0.04	0.0574	1.32	0.0560	1.29
	2	-0.0628	-1.45	-0.0020	-0.05	0.0246	0.55	-0.0017	-0.04	0.0012	0.03	0.0012	0.03
	3	0.0425	0.94	-0.0021	-0.05	0.0178	0.40	-0.0020	-0.04	-0.0018	-0.04	-0.0018	-0.04
	4	-0.0307	-0.74	-0.0018	-0.04	0.0127	0.29	-0.0020	-0.05	-0.0024	-0.05	-0.0019	-0.04
	5	0.0470	1.06	-0.0018	-0.04	0.0089	0.20	-0.0020	-0.04	-0.0023	-0.05	-0.0021	-0.05
	6	0.0131	0.29	-0.0020	-0.05	0.0061	0.14	-0.0020	-0.05	-0.0022	-0.05	-0.0022	-0.05
	7	-0.0627	-1.42	-0.0018	-0.04	0.0040	0.09	-0.0020	-0.05	-0.0020	-0.05	-0.0020	-0.05
	8	-0.0689	-1.56	-0.0019	-0.04	0.0023	0.05	-0.0017	-0.04	-0.0020	-0.05	-0.0021	-0.05
(b) Squared Series													
	Lag	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**
	1	0.0248	0.51	-0.0019	-0.04	0.1392	2.93	-0.0021	-0.04	0.0006	0.02	0.0785	1.62
	2	0.1054	2.35	-0.0017	-0.04	0.1029	2.19	-0.0019	-0.04	-0.0019	-0.04	0.1270	2.98
	3	0.0045	0.10	-0.0019	-0.04	0.0752	1.61	-0.0018	-0.04	-0.0018	-0.04	0.0163	0.36
	4	0.0411	0.93	-0.0019	-0.04	0.0551	1.19	-0.0018	-0.04	-0.0020	-0.04	0.0146	0.32
	5	0.0494	1.12	-0.0019	-0.04	0.0400	0.87	-0.0019	-0.04	-0.0017	-0.04	0.0008	0.02
	6	0.0011	0.02	-0.0019	-0.04	0.0289	0.63	-0.0017	-0.04	-0.0019	-0.04	-0.0005	-0.01
	7	0.0010	0.02	-0.0017	-0.04	0.0204	0.44	-0.0019	-0.04	-0.0021	-0.04	-0.0021	-0.04
	8	0.0008	0.02	-0.0017	-0.04	0.0143	0.31	-0.0019	-0.04	-0.0020	-0.04	-0.0027	-0.06

Note: diff(%) = relative difference to the observed value from the data

Table 2 Fitted Model Parameters for TSE 300 TR Data

Model	***Parameters***			
ILN	$\mu = 0.00814$	$\sigma = 0.04511$		
RSLN2	$\mu_1 = 0.0123$	$\sigma_1 = 0.0347$	$p_{12} = 0.0371$	$\pi_1 = 0.8500$
	$\mu_2 = -0.0157$	$\sigma_2 = 0.0778$	$p_{21} = 0.2101$	$\pi_2 = 0.1500$
MIND2	$\mu_1 = 0.0118$	$\sigma_1 = 0.0374$		$\alpha_1 = 0.9237$
	$\mu_2 = -0.0357$	$\sigma_2 = 0.0872$		$\alpha_2 = 0.0763$
MAR(2;1,0)	$\phi_{10} = 0.0106$	$\phi_{11} = 0.0636$		
	$\sigma_1 = 0.0382$			$\alpha_1 = 0.9427$
	$\phi_{20} = -0.0425$			
	$\sigma_2 = 0.0931$			$\alpha_2 = 0.0573$
MARCH(2;1,0;2,0)	$\phi_{10} = 0.0107$	$\phi_{11} = 0.0612$		
	$\beta_{10} = 0.0011$	$\beta_{11} = 0.0740$	$\beta_{12} = 0.1348$	$\alpha_1 = 0.9508$
	$\phi_{20} = -0.0671$			
	$\beta_{20} = 0.0065$			$\alpha_2 = 0.0492$

Table 3 Comparison of the S&P 500 Data and the Fitted Models

		Data		*ILN*		*RSLN2*		*MIND2*		*MARCH(2;0,2;0,0)*		*MARCH(2;0,3;0,0)*	
					diff(%)		**diff(%)**		**diff(%)**		**diff(%)**		
	Mean	0.00963		0.00962	**-0.1%**	0.00965	**0.2%**	0.00962	**-0.1%**	0.00973	**1.0%**	0.00943	**-2.1%**
	Std Dev	0.04156		0.04155	**0.0%**	0.04144	**-0.3%**	0.04149	**-0.2%**	0.04126	**-0.7%**	0.04177	**0.5%**
	Skewness	-0.63995		0.00044	**-100%**	-0.49307	**-23.0%**	-0.42111	**-34.2%**	-0.49126	**-23.2%**	-0.48958	**-23.5%**
	Kurtosis	3.01395		-0.01228	**-100%**	2.08630	**-30.8%**	1.88026	**-37.6%**	2.00858	**-33.4%**	2.34149	**-22.3%**
Y(1)	**Minimum**	-0.24253		-0.11721	**-51.7%**	-0.18541	**-23.6%**	-0.17507	**-27.8%**	-0.19033	**-21.5%**	-0.19926	**-17.8%**
Y(13)	**L97.5%**	-0.08207		-0.07266	**-11.5%**	-0.07977	**-2.8%**	-0.08210	**0.0%**	-0.07616	**-7.2%**	-0.07673	**-6.5%**
Y(26)	**L95.0%**	-0.05921		-0.05935	**0.2%**	-0.05810	**-1.9%**	-0.05930	**0.1%**	-0.05714	**-3.5%**	-0.05771	**-2.5%**
Y(52)	**L90.0%**	-0.04196		-0.04416	**5.3%**	-0.03984	**-5.0%**	-0.03997	**-4.7%**	-0.04007	**-4.5%**	-0.04053	**-3.4%**
	Pr(Crash)			0.00000		0.06600		0.02550		0.10234		0.16780	
	log L			929.4		952.5		946.7		947.6		950.3	
	AIC			927.4		946.5		941.7		940.6		942.3	
	SBC			923.1		933.6		931.0		925.7		925.3	
	Effective # Obs			527		527		527		525		524	
	Normalized log L			929.4		952.5		946.7		951.2		955.7	
	Normalized AIC			927.4		946.5		941.7		944.2		947.7	
	Normalized SBC			923.1		933.6		931.0		929.2		930.6	
Autocorrelations													
(a) Raw Series													
	Lag	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**
	1	0.0254	0.60	-0.0019	-0.04	0.0250	0.56	-0.0020	-0.05	-0.0019	-0.04	-0.0019	-0.04
	2	-0.0352	-0.80	-0.0020	-0.05	0.0135	0.31	-0.0018	-0.04	-0.0020	-0.05	-0.0020	-0.05
	3	0.0005	0.01	-0.0021	-0.05	0.0069	0.16	-0.0021	-0.05	-0.0019	-0.04	-0.0015	-0.03
	4	0.0003	0.01	-0.0018	-0.04	0.0031	0.07	-0.0019	-0.04	-0.0017	-0.04	-0.0017	-0.04
	5	0.0768	1.76	-0.0018	-0.04	0.0008	0.02	-0.0020	-0.05	-0.0020	-0.05	-0.0020	-0.05
	6	-0.0543	-1.24	-0.0020	-0.05	-0.0004	-0.01	-0.0020	-0.05	-0.0019	-0.04	-0.0020	-0.05
	7	-0.0423	-0.96	-0.0018	-0.04	-0.0011	-0.03	-0.0021	-0.05	-0.0020	-0.05	-0.0018	-0.04
	8	-0.0400	-0.91	-0.0019	-0.04	-0.0016	-0.04	-0.0018	-0.04	-0.0018	-0.04	-0.0020	-0.05
(b) Squared Series													
	Lag	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**	**acf**	**t-ratio**
	1	0.1204	2.63	-0.0019	-0.04	0.0975	2.16	-0.0017	-0.04	0.0671	1.49	0.0713	1.49
	2	0.0603	1.32	-0.0017	-0.04	0.0546	1.22	-0.0021	-0.04	0.0510	1.17	0.0515	1.10
	3	0.0595	1.32	-0.0019	-0.04	0.0299	0.67	-0.0019	-0.04	0.0046	0.10	0.1571	3.72
	4	0.0292	0.66	-0.0019	-0.04	0.0160	0.36	-0.0019	-0.04	0.0005	0.01	0.0207	0.45
	5	0.0079	0.18	-0.0019	-0.04	0.0082	0.18	-0.0020	-0.04	-0.0019	-0.04	0.0129	0.28
	6	0.0256	0.57	-0.0019	-0.04	0.0032	0.07	-0.0017	-0.04	-0.0022	-0.05	0.0237	0.52
	7	-0.0502	-1.13	-0.0017	-0.04	0.0007	0.02	-0.0017	-0.04	-0.0021	-0.05	0.0030	0.06
	8	-0.0072	-0.16	-0.0017	-0.04	-0.0007	-0.01	-0.0021	-0.04	-0.0025	-0.05	0.0005	0.01

Note: diff(%) = relative difference to the observed value from the data

Table 4 Fitted Model Parameters for S&P 500 TR Data

Model	***Parameters***				
ILN	$\mu = 0.00963$	$\sigma = 0.04156$			
RSLN2	$\mu_1 = 0.0126$	$\sigma_1 = 0.0350$	$p_{12} = 0.0398$		$\pi_1 = 0.9050$
	$\mu_2 = -0.0185$	$\sigma_2 = 0.0748$	$p_{21} = 0.3798$		$\pi_2 = 0.0950$
MIND2	$\mu_1 = 0.0129$	$\sigma_1 = 0.0335$			$\alpha_1 = 0.8485$
	$\mu_2 = -0.0088$	$\sigma_2 = 0.0686$			$\alpha_2 = 0.1515$
MARCH(2;0,2;0,0)	$\phi_{10} = 0.0121$				
	$\beta_{10} = 0.0012$	$\beta_{11} = 0.0709$	$\beta_{12} = 0.0520$		$\alpha_1 = 0.9519$
	$\phi_{20} = -0.0363$				
	$\beta_{20} = 0.0062$				$\alpha_2 = 0.0481$
MARCH(2;0,3;0,0)	$\phi_{10} = 0.0114$				
	$\beta_{10} = 0.0010$	$\beta_{11} = 0.0670$	$\beta_{12} = 0.0406$	$\beta_{13} = 0.1690$	$\alpha_1 = 0.9745$
	$\phi_{20} = -0.0642$				
	$\beta_{20} = 0.0065$				$\alpha_2 = 0.0255$