

Chapter 93 Time Series Modeling and Forecasting of the Volatilities of Asset Returns

By

Tze Leung Lai, Department of Statistics, Stanford University

Haipeng Xing, Department of Applied Mathematics and Statistics, SUNY at Stony Brook

For *The Handbook of Quantitative Finance and Risk Management*

Abstract

Dynamic modeling of asset returns and their volatilities is a central topic in quantitative finance. Herein we review basic statistical models and methods for the analysis and forecasting of volatilities. We also survey several recent developments in regime-switching, change-point and multivariate volatility models.

Keywords: Conditional heteroskedasticity; Stochastic volatility; Regime-switching; Change-point.

93.1 Introduction

Volatility of asset returns or interest rates plays a central role in quantitative finance and appears in basic formulas and theories in derivative pricing, risk management and asset allocation. In derivative pricing, volatility of the underlying asset return plays a key role, and the market convention is to list option prices in terms of volatility units. Nowadays, derivative contracts on volatilities are traded in financial markets, so dynamic models of volatilities are indispensable in pricing such derivatives. In financial risk management, the Basel Accord requires banks and other financial institutions to calculate their minimum capital requirement for covering the risks of their trading positions that are aggregated into a risk measure involving volatility forecasts. In Markowitz's portfolio theory, the optimal weights allocated to the assets in the portfolio involve the volatilities of the asset returns and their correlations. In view of the fundamental importance of volatility modeling and forecasting, there is a large literature on the subject, and surveys of various directions in the literature have been provided by Knight and Satchell (1999), Engle and Patton (2001) and Poon and Granger (2003).

In this chapter, we review statistical methods and models for evaluating and forecasting the volatilities of asset returns. Section 2 describes some stylized facts of asset returns and their volatilities, and gives a brief review of classical volatility models, including the GARCH-type and stochastic volatility models. Section 3 surveys several recent developments, primarily in the area of regime-switching and change-point volatility models and in relating volatilities to exogenous variables. Section 4 summarizes multivariate regression models and describes in this connection some recent work on mean-variance portfolio optimization when the means and covariances of the underlying assets are unknown and have to be estimated from time series data.

93.2 Conditional heteroskedasticity models

93.2.1 Stylized facts on time series of asset returns

Lai and Xing (2008a, pp. 140-144) summarize various stylized facts from the literature on empirical analysis of asset returns and their volatilities. In particular, the empirical analysis has revealed clustering of large changes and returns. Such *volatility clustering* is common in the intra-day, daily and weekly returns in equity, commodity and foreign exchange markets, and results in much stronger autocorrelations of the squared returns than the original returns that are usually weakly autocorrelated. Moreover, there is asymmetry of magnitudes in upward and downward movements of asset returns, and the volatility response to a large positive return is considerably smaller than that to a negative return of the same magnitude. This asymmetry is sometimes referred to as a *leverage effect*. A possible cause of the asymmetry is that a drop of a stock price increases the debt-to-asset ratio of the stock, which in turn increases the volatility of returns to the equity holders. In addition, the news of increasing volatility makes the future of the stock more uncertain, and the asset price and its return therefore become more volatile.

93.2.2 Historic volatility and exponentially weighted moving averages

Let σ_t be the volatility of the asset return r_t at time t . A commonly used estimate of σ_t^2 at time $t - 1$ is the sample variance based on the most recent k observations:

$$\hat{\sigma}_t^2 = \frac{1}{k-1} \sum_{i=1}^k (r_{t-i} - \bar{r})^2, \quad (93.1)$$

where $\bar{r} = \sum_{i=1}^k r_{t-i}/k$. The volatility given by (93.1) is called the *historic volatility*. If k is in days, we can convert (93.1) in the daily basis to the annual volatility by $\sqrt{A}\sigma$, where the annualizing factor A is the number of trading days, usually taken as around 252; see Lai and Xing (2008a, p. 66).

Historic volatility uses equal weights for the observations in the moving window of returns. Since the interest is to estimate the current level of volatility, one may want to put more weight to the most recent observations, yielding an estimate of the form

$$\hat{\sigma}_t^2 = \sum_{i=1}^k \alpha_i u_{t-i}^2 \quad (93.2)$$

where $\sum_{i=1}^k \alpha_i = 1$, and we use u_m to denote the centered log return $r_m - \bar{r}$, or simply the log return r_m (since \bar{r}^2 is typically small in comparison with $k^{-1} \sum_{i=1}^k r_{m-i}^2$). The particular case of (93.2) with $\alpha_i = (1 - \lambda)\lambda^{i-1}$ for $0 < \lambda < 1$, $k = \infty$ and $u_j = 0$ for $j \leq 0$, is called an *exponentially weighted moving average* (EWMA) and has the recursive representation

$$\hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 + (1 - \lambda) u_{t-1}^2. \quad (93.3)$$

RiskMetrics, originally developed by J. P. Morgan and made publicly available in 1994, uses EWMA with $\lambda = 0.94$ for updating daily volatility estimates in its database. Noting that (93.3) says that $\hat{\sigma}_t^2$ is convex combination of $\hat{\sigma}_{t-1}^2$ and u_{t-1}^2 , we can modify (93.3) by including also a long-run variance rate V :

$$\hat{\sigma}_t^2 = (1 - \alpha - \beta)V + \alpha u_{t-1}^2 + \beta \hat{\sigma}_{t-1}^2. \quad (93.4)$$

93.2.3 The GARCH model

How should the weights α_i and γ in the moving average scheme (93.2) be chosen and how should V be estimated? Regarding them as unknown parameters in a model of the time-varying volatilities σ_t^2 , we can estimate them by fitting a model to the observed data. In particular, a model that matches well with (93.2) is Engle's (1982) *autoregressive conditional heteroskedastic* (ARCH) model, denoted by ARCH(k) and defined by

$$u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^k \alpha_i u_{t-i}^2, \quad (93.5)$$

in which ϵ_t are i.i.d. random variables with mean 0 and variance 1, having either the standard normal or the following standardized Student- t distribution. Let x_ν be a Student- t distribution with $\nu > 2$ degrees of freedom. Then $\epsilon_t = x_\nu / \sqrt{\nu/(\nu - 2)}$ has a *standardized Student- t*

distribution with variance 1 and probability density function

$$f(\epsilon) = \frac{\Gamma[(\nu + 1)/2]}{\Gamma[\nu/2]\sqrt{(\nu - 2)\pi}} \left(1 + \frac{\epsilon^2}{\nu - 2}\right)^{-(\nu+1)/2}. \quad (93.6)$$

Since the ARCH model (93.5) is similar to an autoregressive model, weak (or covariance) stationarity requires that the zeros of the characteristic polynomial $1 - \alpha_1 z - \dots - \alpha_k z^k$ lie outside the unit circle. As the α_i are usually assumed to be nonnegative, this requirement is equivalent to

$$\alpha_1 + \dots + \alpha_k < 1. \quad (93.7)$$

Under (93.7), the long-run volatility σ of u_t is given by

$$\sigma^2 = E(u_t^2) = \frac{\omega}{1 - \alpha_1 - \dots - \alpha_k}. \quad (93.8)$$

Other properties of ARCH and estimation of the model will be discussed in Section 3.3.

Bollerslev (1986) subsequently generalized (93.5) to a model of the form

$$u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^h \beta_i \sigma_{t-i}^2 + \sum_{j=1}^k \alpha_j u_{t-j}^2, \quad (93.9)$$

where the random variables ϵ_t are i.i.d. with mean 0 and variance 1 and have either the standard normal or a standardized Student- t distribution. This model is called the *generalized ARCH* (GARCH) model and is denoted by GARCH(h, k). It can be considered as an ARMA model of volatility with martingale difference innovations: Specifically, letting $\beta_i = 0$ if $i > h$, $\alpha_j = 0$ if $j > k$, and $\eta_t = u_t^2 - \sigma_t^2$, we can rewrite (93.9) as

$$u_t^2 = \omega + \sum_{i=1}^{\max(h,k)} (\alpha_i + \beta_i) u_{t-i}^2 + \eta_t - \sum_{j=1}^h \beta_j \eta_{t-j}. \quad (93.10)$$

Moreover, $E(\eta_t | u_s, s \leq t-1) = 0$ because $E(u_t^2 | u_s, s \leq t-1) = \sigma_t^2$. Hence the same invertibility and stationarity assumptions for ARMA models apply to GARCH models; see Lai and Xing (2008a, Section 6.3.2) where maximum likelihood estimation of GARCH parameters, likelihood inference and model-based forecasts of future volatilities are also described.

93.2.4 The exponential GARCH model

As pointed out in Section 2.1, a stylized fact of the volatility of asset returns is the leverage effect that the volatility response to large positive returns is considerably smaller than that to negative returns of the same magnitude. Because the GARCH model is defined by σ_t^2 and u_{t-j}^2 , such leverage effect cannot be incorporated. To address the asymmetry, several modifications of the GARCH model have been proposed in the literature, the most popular being the *exponential GARCH* (EGARCH) model proposed by Nelson (1991). Letting $f_j(\epsilon) = \alpha_j \epsilon + \gamma_j (|\epsilon| - E|\epsilon|)$, we can write the EGARCH(h, k) model in the form

$$u_t = \sigma_t \epsilon_t, \quad \log(\sigma_t^2) = \omega + \sum_{i=1}^h \beta_i \log(\sigma_{t-i}^2) + \sum_{j=1}^k f_j(\epsilon_{t-j}). \quad (93.11)$$

The random variable $f_j(\epsilon_t)$ is a sum of two zero-mean random variables $\alpha_j\epsilon_t$ and $\gamma_j(|\epsilon_t| - E|\epsilon_t|)$, and can be written in the form

$$f_j(\epsilon_t) = \begin{cases} (\alpha_j + \gamma_j)\epsilon_t - \gamma_j E|\epsilon_t|, & \text{if } \epsilon_t \geq 0, \\ (\alpha_j - \gamma_j)\epsilon_t - \gamma_j E|\epsilon_t|, & \text{if } \epsilon_t < 0, \end{cases}$$

which shows the asymmetry of the volatility response to positive and to negative returns. Since (93.11) represents $\log(\sigma_t^2)$ as a linear process, σ_t^2 and therefore u_t also are strictly stationary if the zeros of $1 - \beta_1 z - \dots - \beta_h z^h$ lie outside the unit circle; see Lai and Xing (2008a, Section 6.3.4) where EGARCH forecasts of future volatilities and maximum likelihood estimation of EGARCH parameters are also described.

93.2.5 ARMA-GARCH and ARMA-EGARCH models

We can combine an ARMA(p, q) model for r_t with the GARCH(h, k) model (93.9) for the innovations u_t , leading to the ARMA(p, q)-GARCH(h, k) model

$$\begin{aligned} r_t &= \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + \sum_{j=1}^q \psi_j u_{t-j}, & u_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^h \beta_i \sigma_{t-i}^2 + \sum_{j=1}^k \alpha_j u_{t-j}^2, \end{aligned} \quad (93.12)$$

in which the ϵ_t are i.i.d. standard normal or standardized Student- t random variables. The parameters of the combined ARMA(p, q)-GARCH(h, k) can be estimated by maximum likelihood; see Lai and Xing (2008a, pp. 156-157). The one-step ahead forecasts of the asset return and its volatility at time t are given by

$$\begin{aligned} \hat{r}_{t+1|t} &= \phi_0 + \sum_{i=1}^p \phi_i r_{t+1-i} + \sum_{j=1}^q \psi_j u_{t+1-j}, \\ \hat{\sigma}_{t+1|t}^2 &= \omega + \sum_{j=1}^k \alpha_j u_{t+1-j}^2 + \sum_{j=1}^h \beta_j \sigma_{t+1-j}^2. \end{aligned} \quad (93.13)$$

The k -step ahead forecasts can be obtained by making use of (93.13) and the law of iterated conditional expectations. Replacing the second equation in (93.12) by (93.11) yields the ARMA(p, q)-EGARCH(h, k) model.

93.2.6 Volatility persistence and integrated GARCH models

Fitting GARCH(1,1) models to financial time series often displays high volatility in the sense that the estimated $\alpha + \beta$ values are close to 1. More generally, the *integrated GARCH* (IGARCH) model, introduced by Engle and Bollerslev (1986), has the form (93.9) with $\sum_{j=1}^k \alpha_j + \sum_{i=1}^h \beta_i = 1$. Although IGARCH models have infinite (unconditional) variance, they are in fact strictly stationary. In view of the ARMA representation (10) with martingale difference innovations, the shocks η_t to the variance in an IGARCH model do not decay over time, suggesting integration in variance that is analogous to ARIMA models. Note that an integrated ARMA model, ARIMA(p, d, q), for a unit-root nonstationary time series x_t is a stationary ARMA(p, q) for

$(1 - B)^d x_t$, where B is the backshift operator $Bx_t = x_{t-1}$ and therefore $(1 - B)x_t = x_t - x_{t-1}$. Extending d from positive integers to $-\frac{1}{2} < d < \frac{1}{2}$ yields the fractionally integrated ARMA model, ARFIMA(p, d, q), for which $(1 - B)^d x_t$ is stationary ARMA; see Lai and Xing (2008a, p.219). Noting that (93.10) represents a GARCH model for u_t as an ARMA model for u_t^2 , Baillie et al. (1996) have used fractional integration to extend IGARCH models to FIGARCH (fractionally integrated GARCH) models, with a slow hyperbolic rate of decay for the influence of the past innovations η_{t-j} .

93.2.7 Stochastic volatility models

Although the GARCH(1, 1) model $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha u_{t-1}^2$ resembles an AR(1) model for σ_t^2 , it involves the observed u_{t-1}^2 instead of an unobservable innovation in the usual AR(1) model. Replacing αu_{t-1}^2 by a zero-mean random disturbance has the disadvantage that it no longer ensures σ_t^2 to be nonnegative. A simple way to get around this difficulty is to consider $\log \sigma_t^2$ instead of σ_t^2 , leading to the *stochastic volatility* (SV) model introduced by Taylor (1986):

$$u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = e^{h_t}, \quad h_t = \phi_0 + \phi_1 h_{t-1} + \dots + \phi_p h_{t-p} + \eta_t, \quad (93.14)$$

which has AR(p) dynamics for $\log \sigma_t^2$. The ϵ_t and η_t in (93.14) are assumed to be independent normal random variables, with $\epsilon_t \sim N(0, 1)$ and $\eta_t \sim N(0, \sigma^2)$. There is an extensive literature on the properties of SV models; see Taylor (1994), Capobianco (1996), Ghysels et al. (1996), Shephard (1996), and Barndorff-Nielsen and Shephard (2001).

A complication of the SV model is that unlike usual AR(p) models, the h_t in (93.14) is an unobserved state undergoing AR(p) dynamics, while the observations are u_t such that $u_t | h_t \sim N(0, e^{h_t})$. The likelihood function of $\theta = (\sigma, \phi_0, \dots, \phi_p)^T$, based on a sample of n observations u_1, \dots, u_n , involves n -fold integrals, making it prohibitively difficult to compute the MLE by numerical integration for usual sample sizes. A tractable alternative to the MLE is the following quasi-maximum likelihood (QML) estimator. To fix the ideas, consider the case $p = 1$ and let $\phi_0 = \omega$, $\phi_1 = \phi$. Letting $y_t = \log u_t^2$, note that $y_t = h_t + \log \epsilon_t^2$, where $\log \epsilon_t^2$ is distributed like $\log \chi_1^2$, which has mean -1.27 . Let $\xi_t = \log \epsilon_t^2 - E(\log \chi_1^2)$. Note that this is a linear state-space model with unobserved states h_t and observations y_t satisfying

$$h_t = \omega + \phi h_{t-1} + \eta_t, \quad y_t = h_t + E(\log \chi_1^2) + \xi_t. \quad (93.15)$$

The QML treats ξ_t as if it were normal so that (93.15) is a linear Gaussian state-space model to which the Kalman filter can be applied to give not only the minimum-variance linear predictor $h_{t|t-1}$ of h_t based on observations up to time t but also its variance $\text{Var}(h_{t|t-1})$. Let $e_t = y_t - E(\log \chi_1^2) - h_{t|t-1}$ and $v_t = V_{t|t-1} + \text{Var}(\log \chi_1^2)$. If ξ_t were normal (with mean 0 and variance $\text{Var}(\log \chi_1^2)$), then e_t would be $N(0, v_t)$ and therefore the log-likelihood function would be given by

$$l(\omega, \phi, \sigma^2) = -\frac{1}{2} \sum_{t=1}^n \log v_t - \frac{1}{2} \sum_{t=1}^n e_t^2 / v_t \quad (93.16)$$

up to additive constants that do not depend on (ω, ϕ, σ^2) . The QML estimator that maximizes (93.16) is consistent and asymptotically normal but not efficient because it uses an incorrect density function for the non-normal ξ_t .

Another alternative to the MLE is the Bayes estimator that assumes the following combined normal and inverted chi-square prior distribution of (σ^2, ω, ϕ) :

$$\frac{m\lambda}{\sigma^2} \sim \chi_m^2, \quad (\omega, \phi) | \sigma^2 \sim N((\omega_0, \phi_0), \sigma^2 \mathbf{V}_0) \Big|_{|\phi| < 1}, \quad (93.17)$$

where $|\phi_0| < 1$ and $N(\cdot, \cdot) \Big|_{|\phi| < 1}$ denotes the bivariate normal distribution restricted to the region $\{(\omega, \phi) : |\phi| < 1\}$ so that the corresponding AR(1) model for h_t is stationary. Therefore the conditional distribution of $\boldsymbol{\theta} := (\omega, \phi, \sigma^2)$ given $\mathbf{h} := (h_1, \dots, h_n)$ again has the same form as (93.17) but with $(m, \lambda, \omega_0, \phi_0, \mathbf{V}_0)$ replaced by their posterior counterparts. The difficulty with posterior estimation in SV models is that \mathbf{h} is actually unobservable and the observations are u_1, \dots, u_n . Gibbs sampling can be used to obtain a Monte Carlo approximation to the posterior distribution of $(\boldsymbol{\theta}, \mathbf{h})$. Let $\mathbf{h}_{-t} = (h_1, \dots, h_{t-1}, h_{t+1}, \dots, h_n)$, which removes h_t from \mathbf{h} . Using $f(\cdot | \cdot)$ to denote conditional densities, we can use the AR(1) dynamics for h_t to obtain

$$f(h_t | \mathbf{u}, \boldsymbol{\theta}, \mathbf{h}_{-t}) \propto \frac{1}{\sigma_t} \exp\left(-\frac{u_t^2}{2\sigma_t^2}\right) \frac{1}{\sigma_t^2} \exp\left(-\frac{(h_t - u_t)^2}{2\nu^2}\right), \quad (93.18)$$

where $\nu^2 = \sigma^2 / (1 + \phi^2)$ and $\mu_t = [\omega(1 - \phi) + \phi(h_{t-1} + h_{t+1})] / (1 + \phi^2)$. Moreover, the conditional distribution of $\boldsymbol{\theta}$ given \mathbf{u} and \mathbf{h} is a combined normal and inverted χ^2 :

$$\frac{m\lambda + \sum_{t=2}^n v_t^2}{\sigma^2} \sim \chi_{m+n-1}^2, \quad (\omega, \phi) | \sigma^2 \sim N((\omega_*, \phi_*), \sigma^2 \mathbf{V}_*) \Big|_{|\phi| < 1}. \quad (93.19)$$

Therefore, given \mathbf{h} and (ω, ϕ) , let $v_t = h_t - \omega - \phi \log h_{t-1}$ for $2 \leq t \leq n$ and generate $1/\sigma^2$ from the χ^2 -distribution in (93.19). With σ^2 thus generated, let $\mathbf{z}_t = (1, \log h_{t-1})^T$,

$$\mathbf{V}_*^{-1} = \left(\sum_{t=2}^n \mathbf{z}_t \mathbf{z}_t^T \right) / \sigma^2 + \mathbf{V}_0^{-1}, \quad (\omega_*, \phi_*)^T = \mathbf{V}_* \left\{ \mathbf{V}_0^{-1} (\omega_0, \phi_0)^T + \sum_{t=2}^n h_t \mathbf{z}_t / \sigma^2 \right\},$$

and generate (ω, ϕ) from the truncated bivariate normal distribution in (93.19). The Gibbs sampler iterates these two steps in generating $h_1, \dots, h_n, \sigma^2, \omega$, and ϕ until convergence; see Jacquier, Polson and Rossi (1994) for details. Implementation and software for Gibbs sampling can be found on the WinBUGS Web site (<http://www.mrc-bsu.cam.ac.uk/bugs/>). Empirical evidence has suggested that the u_t in (93.14) have fatter tails than normal; see Gallant, Hsich and Tauchen (1997). To accommodate leptokurtic distributions for ϵ_t in SV models, it is convenient to use $\epsilon_t = \sqrt{\lambda_t} z_t$, in which z_t are i.i.d. standard normal variables that are independent of λ_t , and λ_t has an inverted χ^2 -distribution: $\nu / \lambda_t \sim \chi_\nu^2$. This framework implies that $\epsilon_t \sim \sqrt{\lambda_t} z_t$ has Student's t -distribution with ν degrees of freedom. Harvey and Shephard (1996) propose another extension of the SV model to describe the leverage effect of asset return volatilities by allowing correlations between ϵ_t and η_t . These extended SV models can still be estimated by Gibbs sampling when certain prior distributions are imposed on the unknown parameters; see Jacquier, Polson and Rossi (2004). A survey of various estimation methods for SV models is given by Broto and Ruiz (2004).

93.3 Regime-switching, change-point and spline-GARCH models of volatility

As mentioned in Section 2.6, empirical studies of exchange rates, interest rates and stock returns have found high volatility persistence in the estimation of the GARCH(1, 1) parameters ω , α and β , giving values of the MLE $\hat{\lambda}$ of $\lambda := \alpha + \beta$ close to 1. In his comments on Engle and Bollerslev (1986), Diebold (1986) noted that in fitting GARCH models to interest rate data the choice of a constant term ω not accommodating to shifts in monetary policy regimes might have led to an apparently integrated GARCH model. Subsequent work in the literature has demonstrated that if parameter changes are ignored in fitting time series models whose parameters in fact undergo occasional changes, then the fitted models tend to exhibit long memory; see Perron (1989), Lamoureux and Lastrapes (1990), Hillebrand (2005), Lai and Xing (2006) and the references therein. In this section we give a brief survey of the development in incorporating possible parameter changes over time in the volatility models of Section 2.

93.3.1 Regime-switching volatility models

Regime-switching volatility models allow parameter jumps over a small number of regimes. The seminal paper by Hamilton and Susmel (1994) on regime-switching ARCH models provides the basic ideas and paves the way for subsequent developments. They modify the usual ARCH model (93.5) by introducing a scale factor κ_t that remains constant within the same unobservable regime:

$$u_t = \kappa_t \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^k \alpha_i (u_{t-i} / \kappa_{t-i})^2 + \gamma (u_{t-1} / \kappa_{t-1})^2 1_{\{u_{t-1} \geq 0\}}, \quad (93.20)$$

where $1_{\{u_{t-1} \geq 0\}}$ is an indicator variable which assumes the value 1 or 0 and which is used to incorporate the leverage effect. Suppose that there are K possible values of κ_t , which is assumed to be a Markov chain with a given $K \times K$ transition probability matrix. Then one has a hidden Markov model with K states and its likelihood function can be maximized numerically. Hamilton and Susmel have found that (93.20) provides a better fit and better forecasts for the weekly returns of the value-weighted portfolio of stocks traded on NYSE from the week ending on July 3, 1962 to the week ending on December 29, 1987. They attribute most of the persistence in stock price volatility to Markovian switching among low, moderate and high volatility regimes.

Cai (1994) discusses the possibility of extending regime switching to GARCH models while considering a similar regime-switching ARCH model. He comments that “combining the Markov-switching model with GARCH includes tremendous complications in actual estimation.” Gray (1996) points out that this difficulty arises because the hidden states of the parameters in regime-switching GARCH models are path-dependent. To circumvent this difficulty, he suggests modeling σ_t^2 by a finite mixture of conditional variances $\sigma_{t,k}^2$ ($k = 1, \dots, K$), with each $\sigma_{t,k}$ following a GARCH process, instead of using directly a regime-switching model for σ_t^2 . The weights in the mixture are determined by the transition probabilities of an underlying Markov chain.

Regime-switching stochastic volatility (RSSV) models have been introduced by So, Lam and Li (1998), who modify (93.14) in the case $p = 1$ as

$$u_t = \sigma_t \epsilon_t, \quad \log(\sigma_t^2) = \omega_t + \phi \log(\sigma_{t-1}^2) + \eta_t, \quad (93.21)$$

in which ω_t is an unobservable Markov chain assuming K possible values. They use Gibbs sampling to estimate ω_t and ϕ .

93.3.2 Spline-GARCH models

Engle and Rangel (2008) propose to modify the GARCH model (93.9) in the following way to allow structural changes over a long time-scale while maintaining GARCH-type changes over a short time-scale:

$$y_t = \mu + \nu_t \sqrt{h_t} \epsilon_t, \quad (93.22)$$

in which ϵ_t is an i.i.d. sequence and h_t is generated by the GARCH model

$$h_t = \left(1 - \sum_{i=1}^I a_i - \sum_{j=1}^J b_j\right) + \sum_{i=1}^I a_i w_{t-i}^2 + \sum_{j=1}^J b_j h_{t-j}, \quad \text{with } w_s = \sqrt{h_s} \epsilon_s. \quad (93.23)$$

Whereas the vector of unknown GARCH parameters, $\boldsymbol{\eta} = (a_1, \dots, a_I, b_1, \dots, b_J)^T$, is assumed to be time-invariant, the additional factor ν_t is used to model long-term changes in volatilities. Engle and Rangel propose to use an exponential spline function with prespecified knots to represent ν_t by

$$\log \nu_t = \boldsymbol{\beta}^T \mathbf{x}_t + \psi_0 t + \sum_{k=1}^K \psi_k (t - t_k)_+^2, \quad (93.24)$$

where \mathbf{x}_t are exogenous variables (including the constant 1 to incorporate the intercept) and $z_+ = \max(z, 0)$. The number K and the locations of the knots are determined exogenously so that the parameters $\mu, \psi_0, \dots, \psi_K, \boldsymbol{\eta}$ and $\boldsymbol{\beta}$ of the spline-GARCH model can be estimated by maximum likelihood.

93.3.3 Stochastic change-point ARX-GARCH models

Besides the inherent computational complexity, another issue with regime-switching models is the interpretation of the regimes and the choice of its number. An alternative approach to regime switching is to incorporate possible parameter changes of unknown magnitude and at unknown times. McCulloch and Tsay (1993) consider autoregressive models with random shifts in mean and error variances, and Wang and Zivot (2000) assume the number of change-points to be known and use conjugate priors to model changes in regression parameters and error variance. Markov chain Monte Carlo methods are used to estimate the time-varying parameters. Lai, Liu and Xing (2005) have recently introduced a much more tractable model that has explicit recursive formulas for estimating the time-varying parameters. The model has been further extended to include GARCH dynamics (93.9) in the short time-scale. Here we summarize both models and the associated estimators.

Consider a stochastic regression model with time-varying regression coefficients and error variances of the form

$$y_t = \boldsymbol{\beta}_t^T \mathbf{x}_t + \nu_t \sqrt{h_t} \epsilon_t, \quad (93.25)$$

where the stochastic regressor \mathbf{x}_t consists of lagged variables y_{t-1}, y_{t-2}, \dots and exogenous variables, the regression coefficient $\boldsymbol{\beta}_t$ and the long-term volatility level ν_t are piecewise constant, and h_t is generated by the GARCH model (93.23) that relates short-term volatility fluctuations to conditional heteroskedasticity. We use the following Bayesian model to describe the stochastic dynamics of the piecewise constant $\boldsymbol{\theta}_t = (\boldsymbol{\beta}_t^T, \tau_t)^T$, where $\tau_t = (2\nu_t^2)^{-1}$. Letting $I_t = \mathbf{1}_{\{\boldsymbol{\theta}_t \neq \boldsymbol{\theta}_{t-1}\}}$, which is the indicator variable (assuming the values 0 and 1) that shows whether a parameter change occurs at time t , the prior distribution assumes that $I_t, t \geq t_0$, are independent with $P(I_t = 1) = p$, and that

$$\boldsymbol{\theta}_t = (1 - I_t)\boldsymbol{\theta}_{t-1} + I_t(\mathbf{z}_t^T, \gamma_t),$$

where $(\mathbf{z}_1^T, \gamma_1), (\mathbf{z}_2^T, \gamma_2), \dots$ are i.i.d. random vectors such that

$$\gamma_t \sim \chi_m^2/\lambda, \quad \mathbf{z}_t | \gamma_t \sim N(\mathbf{z}, \mathbf{V}/(2\gamma_t)). \quad (93.26)$$

To begin with, suppose that instead of (93.23), h_t is given by $h_t \equiv 1$, which has been considered by Lai, Liu and Xing (2005) who make use of $J_t = \max\{j \leq t : I_j = 1\}$ to derive explicit recursive formulas for the Bayes estimates $E(\boldsymbol{\beta}_t | \mathcal{Y}_t)$ and $E(\sigma_t^2 | \mathcal{Y}_t)$, where $\mathcal{Y}_t = (\mathbf{x}_1, y_1, \dots, \mathbf{x}_t, y_t)$. Note that conditional on \mathcal{Y}_t and $J_t = j$, the distribution of $\boldsymbol{\theta}_t (= \boldsymbol{\theta}_{t-1} = \dots = \boldsymbol{\theta}_j)$ is given by

$$2\tau_t \sim \chi_{m+t-j+1}^2/\lambda_{j,t}, \quad \boldsymbol{\beta}_t | \tau_t \sim N(\mathbf{z}_{j,t}, \mathbf{V}_{j,t}/(2\tau_t)), \quad (93.27)$$

where

$$\begin{aligned} \mathbf{V}_{i,j} &= (\mathbf{V}^{-1} + \sum_{t=i}^j \mathbf{x}_t^T \mathbf{x}_t)^{-1}, & \mathbf{z}_{i,j} &= \mathbf{V}_{i,j}(\mathbf{V}^{-1} \mathbf{z} + \sum_{t=i}^j \mathbf{x}_t y_t), \\ \lambda_{i,j} &= \frac{1}{2} \lambda + \mathbf{z}^T \mathbf{V}^{-1} \mathbf{z} + \sum_{t=i}^j y_t^2 - \mathbf{z}_{i,j}^T \mathbf{V}_{i,j}^{-1} \mathbf{z}_{i,j}. \end{aligned} \quad (93.28)$$

Lai, Liu and Xing (2005) make use of the following recursion for $p_{j,t} = P(J_t = j | \mathcal{Y}_t)$:

$$p_{j,t} \propto p_{j,t}^* := \begin{cases} pf(y_t | I_t = 1) & \text{if } j = t, \\ (1-p)p_{j,t-1}f(y_t | \mathcal{Y}_{j,t-1}, J_t = j) & \text{if } j \leq t-1, \end{cases} \quad (93.29)$$

where we use $f(\cdot)$ to denote conditional densities as in (93.18). Let $g_{i,j} = (m + j - i + 1)/2$ and

$$f_{ij} = (\det(\mathbf{V}_{i,j}))^{1/2} \Gamma(g_{i,j}) \lambda_{i,j}^{-g_{i,j}}, \quad f_{00} = (\det(\mathbf{V}))^{1/2} \Gamma(m) \lambda^{-m/2}. \quad (93.30)$$

Then the conditional densities in (93.29) can be expressed in terms of the quantities in (93.30), yielding

$$p_{j,t}^* = \begin{cases} pf_{tt}/f_{00} & \text{if } j = t, \\ (1-p)p_{j,t-1}f_{jt}/f_{j,t-1} & \text{if } j \leq t-1. \end{cases} \quad (93.31)$$

Hence $p_{j,t} = p_{j,t}^* / \sum_{i=k+1}^t p_{i,t}^*$ and

$$E(\boldsymbol{\beta}_t | \mathcal{Y}_t) = \sum_{j=1}^t p_{j,t} \mathbf{z}_{j,t}, \quad E(\nu_t^2 | \mathcal{Y}_t) = \sum_{j=1}^t p_{j,t} \frac{\lambda_{j,t}}{2(g_{j,t} - 1)}. \quad (93.32)$$

Lai, Liu and Xing (2005) also develop explicit formulas for $E(\boldsymbol{\beta}_t|\mathcal{Y}_n)$ and $E(\nu_t^2|\mathcal{Y}_n)$ when $1 \leq t \leq n$. Let $\tilde{J}_t = \min\{j > t : I_{j+1} = 1\}$ be the future change-time closest to t . Conditional on \mathcal{Y}_n and $(J_t, \tilde{J}_t) = (i, j)$ with $i \leq t < j$, the posterior distribution of $\boldsymbol{\theta}_t (= \boldsymbol{\theta}_i = \dots = \boldsymbol{\theta}_j)$ is

$$2\tau_t \sim \chi_{m+j-i+1}^2/\lambda_{ij}, \quad \boldsymbol{\beta}_t|\tau_t \sim N(\mathbf{z}_{ij}, \mathbf{V}_{ij}/(2\tau_t)). \quad (93.33)$$

Using Bayes' theorem to combine the forward and backward filters, they have shown that

$$E(\boldsymbol{\beta}_t|\mathcal{Y}_n) = \sum_{1 \leq i \leq t \leq j \leq n} \xi_{ijt} \mathbf{z}_{i,j}, \quad E(\nu_t^2|\mathcal{Y}_n) = \sum_{1 \leq i \leq t \leq j \leq n} \xi_{ijt} \frac{\lambda_{i,j}}{2(g_{i,j} - 1)}, \quad (93.34)$$

where $\xi_{ijt} = \xi_{ijt}^*/\{p + \sum_{1 \leq i' \leq t < j' \leq n} \xi_{ijt}^*\}$,

$$\xi_{ijt}^* = \begin{cases} pp_{it} & i \leq t = j, \\ (1-p)p_{it}q_{t+1,j}f_{ij}f_{00}/(f_{it}f_{t+1,j}) & i \leq t \leq j \leq n, \end{cases} \quad (93.35)$$

and analogous to $p_{i,t}$ for the forward filter, the weights $q_{t+1,j}$ for the backward filter are given by

$$q_{t+1,j} \propto q_{t+1,j}^* = \begin{cases} pf_{t+1,t+1}/f_{00} & \text{if } j = t+1, \\ (1-p)q_{t+2,j}f_{t+1,j}/f_{t+2,j} & \text{if } j > t+1. \end{cases}$$

To extend these recursions to the case where h_t is given by (93.23), Lai and Xing (2008b) first assume that the parameter vector $\boldsymbol{\eta} = (a_1, \dots, a_I, b_1, \dots, b_J)^T$ in (93.23) is known. For $n \geq j+1$, they use the recursions

$$\begin{aligned} \mathbf{P}_{n,j} &= \mathbf{P}_{n-1,j} - \left\{ \mathbf{P}_{n-1,j} \mathbf{x}_n \mathbf{x}_n^T \mathbf{P}_{n-1,j} / (\hat{h}_{n,j} + \mathbf{x}_n^T \mathbf{P}_{n-1,j} \mathbf{x}_n) \right\}, \\ \lambda_{n,j} &= \lambda_{n-1,j} + \left\{ (y_n - \hat{\boldsymbol{\beta}}_{n-1,j}^T \mathbf{x}_n)^2 / (\hat{h}_{n,j} + \mathbf{x}_n^T \mathbf{P}_{n-1,j} \mathbf{x}_n) \right\}, \\ \hat{\boldsymbol{\beta}}_{n,j} &= \hat{\boldsymbol{\beta}}_{n-1,j} + \left\{ \mathbf{P}_{n-1,j} \mathbf{x}_n (y_n - \hat{\boldsymbol{\beta}}_{n-1,j}^T \mathbf{x}_n) / (\hat{h}_{n,j} + \mathbf{x}_n^T \mathbf{P}_{n-1,j} \mathbf{x}_n) \right\}, \\ \hat{\nu}_{n,j}^2 &= \frac{m+n-j-2}{m+n-j-1} \hat{\nu}_{n-1,j}^2 + \frac{1}{m+n-j-1} \left\{ (y_n - \hat{\boldsymbol{\beta}}_{n-1,j}^T \mathbf{x}_n)^2 / (\hat{h}_{n,j} + \mathbf{x}_n^T \mathbf{P}_{n-1,j} \mathbf{x}_n) \right\}, \\ \hat{h}_{n,j} &= \left(1 - \sum_{i=1}^I a_i - \sum_{l=1}^J b_l\right) + \sum_{l=1}^J b_l \hat{h}_{n-l,j} + \sum_{i=1}^I a_i (y_{n-i} - \hat{\boldsymbol{\beta}}_{n-1,j}^T \mathbf{x}_{n-i})^2 / \hat{\nu}_{n-1,j}^2. \end{aligned} \quad (93.36)$$

The weight $p_{j,n}$ are given by (93.29) with $\mathbf{V}_{j,n}$ and $\mathbf{z}_{j,n}$ replaced by

$$\begin{aligned} \mathbf{V}_{j,n} &= (\mathbf{V}^{-1} + \sum_{t=j}^n \mathbf{x}_t^T \mathbf{x}_t / \hat{h}_{t,j})^{-1}, \quad z_{j,n} = \mathbf{V}_{j,n} (\mathbf{V}^{-1} \mathbf{z} + \sum_{t=j}^n \mathbf{x}_t y_t / \hat{h}_{t,j}). \\ \lambda_{j,n} &= \frac{1}{2} \lambda + \mathbf{z}^T \mathbf{V}^{-1} \mathbf{z} - \mathbf{z}_{j,n}^T \mathbf{V}_{j,n}^{-1} \mathbf{z}_{j,n} + \sum_{t=j}^n y_t^2 / \hat{h}_{t,j}. \end{aligned} \quad (93.37)$$

With the modifications given by (93.36) and (93.37), they extend the formulas (93.32) for the Bayes estimates $E(\boldsymbol{\beta}_t|\mathcal{Y}_t)$ and $E(\nu_t^2|\mathcal{Y}_t)$. Similarly they modify the formulas (93.34) for the Bayes estimates $E(\boldsymbol{\beta}_t|\mathcal{Y}_n)$ and $E(\nu_t^2|\mathcal{Y}_n)$ when the h_t are generated by the GARCH model (93.23). They also describe how the hyperparameters $p, \mathbf{z}, \mathbf{V}, \boldsymbol{\eta}$ can be estimated by maximum likelihood or variants thereof. Approximations that involve only a bounded number of components in the mixtures in (93.32) and (93.34) and associated approximations to the likelihood function are also given in Lai and Xing (2008b).

93.4 Multivariate volatility models and applications to mean-variance portfolio optimization

For a multivariate time series $\mathbf{y}_t \in \mathbb{R}^p$, let $\boldsymbol{\Sigma}_t = \text{Cov}(\mathbf{u}_t | \mathcal{F}_{t-1})$ be the conditional covariance matrix of \mathbf{u}_t given the information set \mathcal{F}_{t-1} of events up to time $t-1$. There is an extensive literature on dynamic models of $\boldsymbol{\Sigma}_t$. In Section 4.1 we give a brief review of this literature and point out the difficulties with these multivariate volatility models due to the “curse of dimensionality”. In Section 4.2 we describe some recent work to circumvent these difficulties in connection with Markowitz’s mean-variance portfolio optimization when the means and covariances of the asset returns in the future period are unknown and have to be estimated from historical time series.

93.4.1 Multivariate GARCH models

A straightforward extension of the univariate GARCH(1, 1) model $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha u_{t-1}^2$ to the multivariate setting is to use the *vech* (half-vectorization, or half-stacking) operator that transforms a $p \times p$ symmetric matrix \mathbf{M} into a vector, denoted by $\text{vech}(\mathbf{M})$, which consists of the $p(p+1)/2$ lower diagonal (including diagonal) elements of \mathbf{M} . This leads to a multivariate GARCH model of the form

$$\text{vech}(\boldsymbol{\Sigma}_t) = \boldsymbol{\omega} + \mathbf{B}\text{vech}(\boldsymbol{\Sigma}_{t-1}) + \mathbf{A}\text{vech}(\mathbf{u}_t \mathbf{u}_t'), \quad (93.38)$$

where \mathbf{A} and \mathbf{B} are $\frac{p(p+1)}{2} \times \frac{p(p+1)}{2}$ matrices. Note that (93.38) involves $2\left(\frac{p(p+1)}{2}\right)^2 + \frac{p(p+1)}{2}$ parameters. Moreover, \mathbf{A} and \mathbf{B} have to satisfy certain constraints for $\boldsymbol{\Sigma}_t$ to be positive definite, making it difficult to fit such models.

Instead of using the *vech* operator, an alternative approach is to replace σ_s^2 and u_s^2 in (93.9) by $\boldsymbol{\Sigma}_s$ and $\mathbf{u}_s \mathbf{u}_s^T$ and the scalar parameters $\omega, \alpha_j, \beta_i$ by matrices, leading to

$$\boldsymbol{\Sigma}_t = \mathbf{A}\mathbf{A}^T + \sum_{i=1}^h \mathbf{B}_i \boldsymbol{\Sigma}_{t-i} \mathbf{B}_i^T + \sum_{j=1}^k \mathbf{A}_j (\mathbf{u}_{t-j} \mathbf{u}_{t-j}^T) \mathbf{A}_j^T; \quad (93.39)$$

see Engle and Kroner (1995) who refer to their earlier work with Baba and Kraft. The model, commonly called BEKK, involves a lower triangular matrix \mathbf{A} and always gives nonnegative definite $\boldsymbol{\Sigma}_t$. The number of parameters is $p^2(k+h) + p(p+1)/2$, which increases with p^2 .

Many other multivariate generalizations of the GARCH(h, k) model and SV model have been proposed in the literature; see the recent survey by Bauwens, Laurent and Rombouts (2006). An important statistical issue is the curse of dimensionality in parameter estimation unless p is small. Another issue that arises if one models the GARCH dynamics separately for the volatilities of the components of \mathbf{u}_t and for their correlations as in (93.38) is that the resultant matrix may not be nonnegative definite.

93.4.2 A new approach and its application to mean-variance portfolio optimization

Markowitz’s celebrated mean-variance portfolio optimization theory assumes that the means and covariances of the underlying asset returns are known. In practice, they are unknown and

have to be estimated from historical data. Plugging the estimates into Markowitz’s efficient frontier that assumes known parameters has led to portfolios that may perform poorly and have counter-intuitive asset allocation weights; this has been referred to as the “Markowitz optimization enigma.” Lai, Xing and Chen (2008) recently introduced a new approach to resolve the enigma. Not only does the new approach provide substantial improvements over previous methods, but it also allows flexible modeling to incorporate dynamic features and structural changes in the training sample of historical data, as illustrated by simulation and empirical studies. They propose to use a stochastic regression model of the form

$$r_{it} = \boldsymbol{\beta}_i^T \mathbf{x}_{it} + \epsilon_{it}, \quad (93.40)$$

for the time series of asset returns of the i th asset, where the components of \mathbf{x}_{it} include 1, factor variables such as the return of a market portfolio like the S&P500 index at time t , and lagged variables $r_{i,t-1}, r_{i,t-2}, \dots$. They also include conditional heteroskedasticity in (93.40) by assuming the ϵ_{it} to follow the GARCH(1,1) model

$$\epsilon_{it} = s_{it}z_{it}, \quad s_{it}^2 = \omega_i + a_i s_{i,t-1}^2 + b_i r_{i,t-1}^2, \quad (93.41)$$

in which z_{i1}, z_{i2}, \dots are i.i.d. with mean 0 and variance 1. Their basic idea is to introduce covariates (including lagged variables to account for time series effects) and possibly also conditional heteroskedasticity so that the unmodeled errors ϵ_{it} or z_{it} can be regarded as i.i.d. Therefore, instead of i.i.d. returns that are commonly assumed in mean-variance optimization theory, their new approach combines domain knowledge of the m assets with time series modeling to obtain better predictors of the returns and their volatilities for the future period.

The model (93.40) is intended to produce i.i.d. $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{mt})^T$ or i.i.d. $\mathbf{z}_t = (z_{1t}, \dots, z_{mt})^T$ after adjusting for conditional heteroskedasticity as in (93.41). Note that (93.40)-(93.41) models the the asset returns separately, instead of jointly in a multivariate regression or multivariate GARCH model which has the difficulty of having too many parameters to estimate. While the vectors $\boldsymbol{\epsilon}_t$ (or \mathbf{z}_t) are assumed to be i.i.d., (93.40) (or (93.41)) does not assume their components to be uncorrelated since it treats the components separately rather than jointly. This avoids the curse of dimensionality of vector ARX or multivariate GARCH models. Another innovation of the approach of Lai, Xing and Chen (2008) is their new formulation of mean-variance portfolio optimization, when the means and covariances of the asset returns are unknown and have to be estimated from data, as a stochastic optimization problem to which stochastic control techniques can be applied.

REFERENCES

- Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74, 3-30.
- Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21, 79-109.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.

- Barndorff-Nielsen, O.E. and Shephard, N.G. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial econometrics. *Journal of the Royal Statistical Society, Series B*, 63, 167-241.
- Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys*, 18, 613-649.
- Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics*, 12, 309-316.
- Capobianco, E. (1996). State-space stochastic volatility models: a review of estimation algorithms. *Applied Stochastic Models and Data Analysis*, 12, 265-279.
- Diebold, F.X. (1986). Modeling the persistence of conditional variances: A comment. *Econometric Reviews*, 5, 51-56.
- Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflations, *Econometrica*, 50, 987-1007.
- Engle, R.F. and Bollerslev, T. (1986). Modeling the persistence of conditional variances. *Econometric Reviews*, 5, 1-50.
- Engle, R.F., and Patton, A.J. (2001). What good is a volatility model? *Quantitative finance*, 1, 237-245.
- Engle, R.F. and Rangel, J.G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies*, 21, 1187-1222.
- Gallant, A.R., Hsieh, D. and Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics*, 81, 159-192.
- Ghysels, E., Harvey, A.C. and Renault, E. (1996). Stochastic volatility. In G.S. Maddala and C.R. Rao (eds.), *Statistical Methods in Finance*, pp. 119-191. North Holland, Amsterdam.
- Gray, S.F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42, 27-62.
- Hamilton, J.D. and Susmel, R. (1994). A conditional heteroskedasticity and change in regimes. *Journal of Econometrics*, 64, 307-333.
- Harvey, A.C. and Shephard, N. (1996). Estimation of asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics*, 14, 429-434.
- Hillebrand, E. (2005). Neglecting parameter changes in GARCH models. *Journal of Econometrics*, 129, 121-138.
- Jacquier, E., Polson, N.G. and Rossi, P.E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics*, 12, 371-417.
- (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122, 185-212.
- Knight, J. and Satchell, S. (1998). ed. *Forecasting volatility in the Financial Markets*. Butterworth Heinemann.
- Lai, T.L., Liu, H. and Xing, H. (2005). Autoregressive models with piecewise constant volatility and regression parameters. *Statistica Sinica*, 15, 279-301.
- Lai, T.L., and Xing, H. (2006). Structural change as an alternative to long memory in financial time series. In *Advances in Econometrics*, Vol. 20 (H. Carter and T. Fomby, eds.), pp. 209-228.

- Amsterdam: Elsevier.
- (2008a). *Statistical Models and Methods for Financial Markets*. New York: Springer.
- (2008b). Hidden Markov models of structural changes in econometric time series. Technical Report, Department of Statistics, Stanford University.
- Lai, T.L., Xing, H. and Chen, Z. (2008). Mean-variance portfolio optimization when means and covariances are unknown. Technical Report, Department of Statistics, Stanford University.
- Lamoureux, C.G. and Lastrapes, W.D. (1990). Persistence in variance, structural change and the GARCH model. *Journal of Business and Economic Statistics*, 8, 225-234.
- McCulloch, R.E. and Tsay, R.S. (1993). Bayesian inference and prediction for mean and variance shifts in autoregressive time series. *Journal of the American Statistical Association*, 88, 968-978.
- Nelson, D.B. (1990). Stationarity and persistence in GARCH(1,1) model. *Econometric Theory*, 6, 318-334.
- Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, 57, 1361-1401.
- Poon, S., and Grandner, C.W.J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41, 478-539.
- Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In *Statistical Models in Econometric, Finance and Other Fields* (O.E. Barndorff-Nielsen, D. R. Cox and D.V. Hinkley, eds.), pp. 1-67. London: Chapman & Hall.
- So, M.K.P., Lam, K. and Li, W.K. (1998) A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics*, 16, 244-253.
- Taylor, S. (1986). *Modeling Financial Time Series*. Chichester: Wiley.
- (1994). Modeling stochastic volatility: a review and comparative study. *Mathematical Finance*, 4, 183-204.
- Wang, J. and Zivot, E. (2000). A Bayesian time series model of multiple structural changes in level, trend and variance. *Journal of Business and Economic Statistics*, 18, 374-386.