FINANCIAL TIME SERIES

ASSET PRICE AND RETURN

Let $P_t$ be the price of an asset at time $t$ and $F_h$ be the public information available at time $h$. Suppose that the asset pays dividend $D_t$ in the time interval $(t - 1, t]$. Given $F_h$, financial time series is concerned with time evolution of $P_t$ and $D_t$, or equivalently, the corresponding return series for $t > h$. Two types of return are commonly used in practice. The first return is the simple return $R_t$ defined as

$$ R_t = \frac{(P_t + D_t) - P_{t-1}}{P_{t-1}} \quad \text{or} \quad P_t + D_t = (1 + R_t)P_{t-1}. $$

It is the percentage change in value of the asset from date $t - 1$ to date $t$. The second return is the continuously compounded return or log return defined as

$$ r_t = \ln \left( \frac{P_t + D_t}{P_{t-1}} \right) = \ln(P_t + D_t) - \ln(P_{t-1}). $$

The two returns are related by $r_t = \ln(1 + R_t)$, but each enjoys its own convenient properties in applications. For example, a multiperiod log return is simply the temporal aggregation of the single period log returns in the time interval, and the simple return of a portfolio is a weighted average of individual simple returns with weight on each asset being the percentage of the portfolio’s value invested in that asset.

Another important concept in finance is risk. A sound financial investment should result in a payoff that exceeds the risk-free interest rate. Consequently, some financial studies use risk-adjusted returns. The excess return of an asset is the difference between the asset’s return and the return of some riskless reference asset, e.g., a short-term U.S. treasury bill return. We use log returns in this article, but the ideas and methods discussed apply to other returns.

Table 1(a) gives some summary statistics of daily log returns, in percentages, of selected U.S. stocks and the Standard and Poor 500 index. The table shows clearly some special characteristics of daily stock returns. First, sample mean of a daily log returns is small, but sample variance is substantial. Second, there exists some (negative) skewness. Third, the return series have high excess kurtosis. Fourth, the serial correlations, if any, are weak for daily stock returns. Table 1(b) shows the same statistics for monthly log returns of selected exchange rates. Similar characteristics are observed, but the exchange rate series have significant lag-1 serial correlation.

Figure 1 shows the time plot of daily log returns of the S&P 500 index. The plot shows periods of high and low variabilities. This phenomenon is termed volatility clustering in finance. Figure 2 shows the sample autocorrelation function (ACF) of the daily log returns of S&P 500 index and the absolute series of log returns. The ACF of the log returns is small, but that of the absolute series is substantial and decays slowly. Thus, the log returns are basically serially uncorrelated, but highly dependent.

Based on the empirical features of the series, the exact conditional distribution of $r_t$ given $F_h$ with $t > h$ is hard to specify in practice. Much of the research therefore focuses on the time evolution of the first two conditional moments of $r_t$ given $F_h$. Let

$$ \mu_t = E(r_t|F_{t-1}) \quad \text{and} \quad \sigma_t^2 = \text{var}(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}] $$

be the 1-step ahead conditional mean and variance of $r_t$. The aim of financial time series analysis is to study the properties of and the relationship between $\mu_t$ and $\sigma_t$. The quantity $\sigma_t$ is commonly referred to as the volatility of the series. It represents a measure of financial risk and plays an important role in derivative pricing. The fact that $\sigma_t$ is time-varying is called conditional heteroscedasticity.

FUNDAMENTAL AND TECHNICAL ANALYSES

A key feature that distinguishes financial time series analysis from other statistical analyses is that the important volatility variable $\sigma_t$ is not directly observable. As such, special models and methods are devised to analyze financial time series, and the
analysis can be classified into two general approaches. The first approach studies the relationship between \( \mu_t \) and other economic and market variables such as interest rates, growth rate of the gross domestic product, inflation, index of consumer confidence, unemployment rate, trade imbalance, corporate earnings, book-to-market ratio, etc. This is called the fundamental analysis and it often employs regression models with time series errors, including conditional heteroscedasticity, and factor analysis. Denote the vector of explanatory variables by \( x_t \), which may include unity as its first element. The regression model is

\[
 r_t = \beta x_t + \epsilon_t, \quad t = 1, \ldots, n,
\]

where \( n \) is the sample size, \( \epsilon_t \) is the error term satisfying \( E(\epsilon_t|x_t) = 0, E(\epsilon_t^2|x_t) = \sigma^2(\epsilon_t|x_t) \), and \( \epsilon_t \) and \( \epsilon_j \) are uncorrelated if \( i \neq j \). The ordinary least squares method is used to estimate the unknown parameter \( \beta \), but adjustment must be made to account for the heteroscedasticity in variables. Write the model in matrix form as \( Y = X\beta + E \). Then, \( \hat{\beta} = (X'X)^{-1}X'Y \) and \( \text{cov}(\hat{\beta}) = (X'X)^{-1} \), where \( \sigma^2 = \sum_{t=1}^{n} e_t^2 x_t' \) with \( e_t = y_t - \beta x_t \). See [7], [16] and [20] for further details. When the dimension of \( x_t \) is large, dimension reduction methods such as factor and principal component analyses are used to simplify the model.

The second approach to analysis of financial time series explores the dynamic dependence of \( r_t \) given \( F_{t-1} \) and is referred to as the technical analysis. Time series methods and stochastic processes are the main tools used in this approach. Dynamic linear models are postulated for the dependence structure of \( \mu_t \) and \( \sigma_t^2 \). Because the serial correlation in \( r_t \) is weak, as shown in Table 1 and Figure 2(a), the dynamic structure for \( \mu_t \) is relatively simple. It is often a constant or a simple function of the elements of \( F_{t-1} \), e.g., a simple autoregressive model. On the contrary, the dynamic dependence in \( \sigma_t^2 \) is more complex. In particular, it is well-known that both \( \{r_t^2\} \) and \( \{|r_t|\} \) series have strong serial correlations. See Figure 2(b). As another example, consider the daily log returns of Alcoa stock in Table 1. The Box-Ljung statistics of \( \{r_t^2\} \) and \( \{|r_t|\} \) give \( Q(10) = 574 \) and \( Q(10) = 519 \), respectively. These statistics are highly significant compared with their asymptotic chi-squared distribution with 10 degrees of freedom under the assumption that \( \{r_t\} \) are independent and identically distributed (iid) random variables. Modeling the series \( \{\sigma_t^2\} \) is referred to as volatility modeling. For ease in discussion, we rewrite the return as

\[
 r_t = \mu_t + \sigma_t = \mu_t + \sigma_t \epsilon_t,
\]

where \( \sigma_t^2 = E(\sigma_t^2|F_{t-1}) \) and \( \epsilon_t \) is an iid sequence of random variables with mean zero and variance 1. Volatility modeling focuses on the \( \{\mu_t\} \) or \( \{\sigma_t\} \) series. In practice, \( \epsilon_t \) is often assumed to be standard normal or a standardized Student-\( t \) distribution with \( v \) degrees of freedom or a generalized error distribution. In many applications, it is essential

### Table 1. Summary Statistics of Selected Log Return Series, where Returns are in Percentages, \( \rho_1 \) is the lag-1 Autocorrelation and \( Q(10) \) is the Box-Ljung Statistic of the First 10 lags of Autocorrelation

<table>
<thead>
<tr>
<th>Series</th>
<th>Mean</th>
<th>Stand. Error</th>
<th>Skew.</th>
<th>Excess Kurt.</th>
<th>Min</th>
<th>Max</th>
<th>( \rho_1 )</th>
<th>Q(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Daily stock returns from 1980.1 to 2001.12 with sample size 5556</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>0.043</td>
<td>1.039</td>
<td>-2.147</td>
<td>47.11</td>
<td>-22.90</td>
<td>8.71</td>
<td>0.03</td>
<td>26.4</td>
</tr>
<tr>
<td>IBM</td>
<td>0.048</td>
<td>1.331</td>
<td>-0.457</td>
<td>13.89</td>
<td>-26.09</td>
<td>12.37</td>
<td>-0.04</td>
<td>18.8</td>
</tr>
<tr>
<td>Intel</td>
<td>0.082</td>
<td>2.824</td>
<td>-0.260</td>
<td>4.90</td>
<td>-24.89</td>
<td>23.41</td>
<td>0.05</td>
<td>40.6</td>
</tr>
<tr>
<td>GM</td>
<td>0.035</td>
<td>1.859</td>
<td>-0.290</td>
<td>7.06</td>
<td>-23.60</td>
<td>13.65</td>
<td>-0.00</td>
<td>12.9</td>
</tr>
<tr>
<td>Alcoa</td>
<td>0.054</td>
<td>1.995</td>
<td>-0.273</td>
<td>9.29</td>
<td>-27.58</td>
<td>13.15</td>
<td>0.06</td>
<td>40.4</td>
</tr>
<tr>
<td>(b) Monthly returns of exchange rate: 1971.1 to 2002.9 with sample size 380</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP-US</td>
<td>-0.285</td>
<td>2.802</td>
<td>-0.556</td>
<td>0.94</td>
<td>-10.52</td>
<td>8.06</td>
<td>0.35</td>
<td>59.2</td>
</tr>
<tr>
<td>US-UK</td>
<td>-0.115</td>
<td>2.431</td>
<td>-0.174</td>
<td>1.67</td>
<td>-11.08</td>
<td>9.52</td>
<td>0.35</td>
<td>51.3</td>
</tr>
</tbody>
</table>
to combine the fundamental and technical analyses.

VOLATILITY MODEL

GARCH Family of Models

There are three approaches to volatility modeling. The first approach originated from [9] specifies a fixed function for $\sigma_t^2$ and is called the generalized autoregressive conditional heteroscedastic (GARCH) modeling. A GARCH($p, q$) model for $\sigma_t^2$ is defined as

$$
\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \eta_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,
$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j > 0$. Let $\eta_t = \alpha_t^2 - \sigma_t^2$. It is easy to show that $E(\eta_t) = 0$ and $E(\eta_t \eta_{t+j}) = 0$. 

Figure 1. Daily log returns of the Standard and Poor 500 index from January 1980 to December 2001.

Figure 2. Autocorrelations of daily returns of S&P 500 from 1980 to 2001: (a) log returns (b) absolute series. The horizontal lines are two standard-error limits.
for \( i \neq j \). Thus, \( \{ \eta_i \} \) is a serially uncorrelated series. However, they are not identically distributed. We can rewrite the GARCH model as

\[
a_t^2 = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \sigma_{t-i}^2 + \eta_t - \sum_{j=1}^{q} \beta_j \eta_{t-j},
\]

which is in the form of an autoregressive moving-average (ARMA) model for the \( \{ \sigma_t^2 \} \) series, where \( \alpha_0 = 0 \) if \( i > p \) and \( \beta_i = 0 \) if \( i > q \). See [4]. Many properties of GARCH models can be deduced from those of ARMA models. Two of the properties are of particular relevance in financial time series analysis. First, a given model large elements in \( \{ a_t^2 \} \) or \( \{ \sigma_t^2 \} \) produce a large \( \sigma_t^2 \), which in turn implies a higher probability for obtaining a shock \( \alpha_t \) with large size. GARCH models thus are capable of describing volatility clustering mentioned before. Second, the model may possess high excess kurtosis. Indeed, the excess kurtosis of \( \alpha_t \) of a GARCH(p,q) model is

\[
K_\alpha = \frac{K_0 + K_\alpha^{(g)} + 5 K_0 K_\alpha^{(e)}}{1 - \frac{5}{2} K_0 K_\alpha^{(e)}},
\]

where \( K_\alpha \) is the excess kurtosis of \( \epsilon_t \) and \( K_\alpha^{(e)} \) is the excess kurtosis of the GARCH model when \( \epsilon_t \) is standard Gaussian. See [3] and [19]. Therefore, even if \( \epsilon_t \) is standard Gaussian for which \( K_\alpha = 0 \), the GARCH model can have positive excess kurtosis provided that the fourth moment exists. For instance, for a GARCH(1,1) model with standard Gaussian innovations \( \epsilon_t \), we have

\[
K_\alpha^{(e)} = \frac{6 \alpha_1^2}{1 - 2 \alpha_1^2 - (\alpha_1 + \beta_1)^2} > 0,
\]

which is positive if \( \alpha_1 > 0 \) and \( 1 - 2 \alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0 \). Therefore, \( K_\alpha = K_\alpha^{(e)} > 0 \) for a Gaussian GARCH(1,1) series.

A major weakness of GARCH models is that they assume symmetric impacts of positive and negative shocks \( \alpha_{t-i} \) on the volatility \( \sigma_t^2 \). On the other hand, empirical evidence shows that large positive and negative shocks have rather different effects on asset returns. To overcome this weakness, many variants of GARCH models have been proposed in the literature, e.g. the exponential GARCH model in [18]. For further discussion on this approach to volatility modeling, see Chapter 3 of [19].

Conditional maximum likelihood method is often used in GARCH estimation. Let \( A_m = (\alpha_1, \cdots, \alpha_m)' \) be the vector of shocks \( \alpha_t \) to a return series \( \{ r_t \}_{t=1}^n \), \( \theta \) be the parameter vector of a GARCH(p,q) model in Eq. (1), \( m = \max(p,q) \), and \( H_m = (\sigma_1, \cdots, \sigma_m)' \). Conditioned on \( A_m \) and \( H_m \), the likelihood function of \( A_m \) can be obtained by recursive conditioning

\[
f(A_n | A_m, H_m, \theta) = f(a_n | A_{n-1}, H_m, \theta) \times f(a_{n-1} | A_{n-2}, H_m, \theta) \cdots f(a_1 | A_0, H_m, \theta),
\]

where \( a_t = r_t - \mu_t = \sigma_t \epsilon_t \). If \( \epsilon_t \) is Gaussian, the estimate can be obtained by minimizing

\[
L(\theta) = \sum_{i=m+1}^{n} \left[ \ln(\sigma_i^2) + \frac{a_i^2}{\sigma_i^2} \right].
\]

(2)

Even if \( \epsilon_t \) is not Gaussian, \( L(\theta) \) of Eq. (2) is often minimized to obtain a quasi maximum likelihood estimate (QMLE) of \( \theta \). Denote the true parameter by \( \theta_0 \) and let \( \ell = 1 + p + q \) be the number of parameters of a GARCH(p,q) model. Rewrite \( \sigma_i^2 \) as a function of the past error terms \( \{ a_{t-i}^2 \}_{i=1}^{\infty} \), i.e.

\[
\sigma_i^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{\ell} \beta_i} + \sum_{i=1}^{\ell} \alpha_i a_{t-i}^2
\]

\[
+ \sum_{i=1}^{\ell} \sum_{k=1}^{q} \cdots \sum_{j_k}^{q} \beta_{j_1} \cdots \beta_{j_k} a_{t-i-j_1-\cdots-j_k}^2,
\]

where the infinite sum disappears for ARCH models. Let \( U \) be the \( \ell \)-dimensional vector of the first derivatives of \( \sigma_i^2 \) with respect to \( \theta \) and \( M = E_\theta(\sigma_i^2 UU') \) be an \( \ell \times \ell \) matrix, where \( E_\theta \) denotes expectation when \( \theta \) takes its true value \( \theta_0 \). Under the conditions: (a) \( p > 0 \) and \( \alpha_i > 0 \) for \( i = 0, \cdots, p \) and (b) if \( q > 0, \beta_i > 0 \) for \( i = 1, \cdots, q \), Hall and Yao [14] show that elements of \( M \) exist and are finite. Let \( \hat{\theta} \) be any local minimum of \( L(\theta) \) in Eq. (2) that occurs within radius \( \zeta \) of \( \theta_0 \) for sufficiently small but fixed \( \zeta > 0 \). Hall and Yao show that the limiting distribution of \( \hat{\theta} \)
depends on the distribution of $\epsilon_t^2$ and can be summarized as below.

**Theorem 1.** Assume $M$ is nonsingular and the conditions (a) and (b) mentioned above hold. Also, assume that the small radius $\xi$ of $\theta_0$ is strictly positive and sufficiently small. (i) If $E(\epsilon_0^4) < \infty$ then $n^{1/2}(\hat{\theta} - \theta_0)$ is asymptotically normally distributed with mean zero and variance $\tau^2M^{-1}$, where $\tau^2 = E(\epsilon_0^4) - 1$. (ii) If $E(\epsilon_0^4) = \infty$ but the distribution of $\epsilon_t^2$ is in the domain of attraction of the normal law, then $n^{\alpha - 1}(\hat{\theta} - \theta_0)$ is asymptotically normally distributed with mean zero and variance $M^{-1}$, where $\lambda_n$ is defined as

$$\lambda_n = \inf\{\lambda > 0 : nH(\lambda) \leq \lambda^2\} \quad \text{and} \quad H(\lambda) = E[\epsilon_0^4I(\epsilon_0^2 \leq \lambda)],$$

where $I(.)$ is the usual indicator function. (iii) If the distribution of $\epsilon_t^2$ is in the domain of attraction of a stable law with exponent $\gamma$, then the limiting distribution of $\hat{\theta}$ follows a multivariate stable law determined by the value of $\gamma$.

This theorem is particularly relevant because financial time series tend to have heavy tails. Bootstrap methods are suggested by Hall and Yao [14] to obtain critical values of $\hat{\theta}$.

**Stochastic Volatility Model**

An alternative approach to volatility modeling is to postulate a stochastic model for the latent volatility. Simple autoregressions are often used resulting in the model

$$\ln(\sigma_t) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln(\sigma_{t-i}) + \eta_t, \quad (3)$$

where $\{\eta_t\}$ is an iid Gaussian series with mean zero and variance $\sigma^2$ and all of the zeros of the polynomial $1 - \sum_{i=1}^{p} a_ix^i$ are outside the unit circle. Log transformation is used in Eq. (3) to ensure the positiveness of the volatility.

The additional innovations $\{\eta_t\}$ substantially increase the flexibility of stochastic volatility (SV) models in modeling the continually changing market conditions. However, this advantage comes with the price of complexity in estimation. Let $\theta = (\alpha_0, \alpha_1, \ldots, \alpha_p, \sigma, \gamma)\tau$ be the parameter vector of the SV model and $H_n = (\sigma_1, \ldots, \sigma_n)\tau$ be the volatility vector. The likelihood function of $A_n$ for a SV model is a mixture over the $n$-dimensional $H_n$ as

$$f(\hat{A}_n|\theta) = \int f(A_n|H_n, \theta)f(H_n|\theta)dH_n$$

$$= \int f(A_n|H_n)f(H_n|\theta)dH_n.$$

For a large or moderate sample size $n$, an effective way to evaluate this likelihood function is to treat $H_n$ as augmented data and apply Monte Carlo methods such as the Markov Chain Monte Carlo (MCMC) method with Gibbs sampling. See Chapter 10 of [19] and the references therein.

**Realized Volatility**

The advent of high-frequency financial data, e.g. the transaction-by-transaction data of stock markets, makes quadratic variations of intraday returns a viable alternative to volatility modeling. The basic idea of this approach has been around for many years. For instance, daily log returns were used in [12] to estimate the volatility of monthly log returns. Consider the problem of estimating daily volatility of a log return series. Using the additive property of log returns, the daily return is a temporal aggregation of intraday returns, i.e.

$$r_t = \sum_{i=1}^{n} r_{m,i \cdot},$$

where $m$ denotes the intraday time interval measured in minutes, $r_{m,i \cdot}$ is the $i$th $m$-minute log return in date $t$, and $n$ is the total number of $m$-minute returns in date $t$. The volatility of $r_t$ is

$$\text{var}(r_t|F_{t-1}) = \text{var}(\sum_{i=1}^{n} r_{m,i \cdot}|F_{t-1}) = \sum_{i=1}^{n} \text{var}(r_{m,i \cdot}|F_{t-1})$$

$$+ 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{cov}(r_{m,i \cdot}, r_{m,j \cdot}|F_{t-1}). \quad (4)$$

Under some general conditions, the volatility in Eq. (4) can be estimated directly from
the intraday log returns. For instance, if the intraday returns \( \{r_{m,i,t}\}_{i=1}^{n} \) are serially uncorrelated, then

\[
\text{var}(r_{t}|F_{t-1}) = \hat{\sigma}_{t}^{2} = \sum_{i=1}^{n} r_{m,i,t}^{2}
\]

is a consistent estimate of the square of daily volatility. This estimate is called realized volatility in [2]. If \( r_{m,i,t} \) is serially correlated with lag-1 correlation, i.e. it follows a moving-average model of order 1, then

\[
\hat{\sigma}_{t}^{2} = \sum_{i=1}^{n} r_{m,i,t}^{2} + 2 \sum_{i=1}^{n-1} r_{m,i,t} r_{m,i+1,t}
\]

is a consistent estimate of the square of daily volatility. For intraday index returns, e.g. S&P 500 index, the lag-1 serial correlation appears to be important, especially when the time interval is short. For exchange rate and individual stock series, serial correlations of intraday log returns are relatively small. The efficacy of realized volatility as an estimate of volatility depends heavily on the excess kurtosis of the intraday log returns; the higher the excess kurtosis the less efficient the estimate.

Experience shows that the marginal distribution of logarithms of realized volatility is approximately normal. Also the log volatility series exhibits characteristics of long-range dependence, e.g. sample autocorrelations of the log volatility series decay slowly as lag increases. Linear time series models, including fractionally differenced models, have been used to fit the \( \{\ln(\hat{\sigma}_{t}^{2})\} \) series and the fitted model is used to produce volatility forecasts.

**Remark 1.** Our discussion of volatility is based on asset returns. In options markets, the observed prices of traded options can be used to deduce the corresponding volatility using the well-known Black-Scholes formulas. The resulting volatility is called implied volatility. Implied volatility depends on the specific options used, however. At-the-money short-term call options are often used in practice.

**Illustration**

Figure 3(a) shows the monthly log returns, in percentage, of the S&P 500 index from 1926 to 2001 for 912 observations. This series has been extensively studied in the literature.
because it is the underlying process of one of the most actively traded options in the world. If Gaussian GARCH models are entertained, we obtain

\[ r_t = 0.695(0.144) + \sigma_t \epsilon_t \]
\[ \sigma_t^2 = 0.693(0.182) + 0.115(0.016)\sigma_{t-1}^2 + 0.866(0.015)\epsilon_{t-1}^2, \]

where the number in parentheses denotes the asymptotic standard error. Let \( \epsilon_t = (r_t - 0.695)/\hat{\sigma}_t \) be the standardized residuals. The Box-Ljung statistics of \( \{ \epsilon_t \} \) and \( \{ \epsilon_t^2 \} \) give \( Q(10) = 11.68(0.31) \) and \( Q(10) = 5.44(0.86) \), where the number in parentheses denotes p-value. Thus, the simple GARCH(1,1) model is adequate in describing the first two conditional moments of the monthly log returns of S&P 500 index. However, the skewness and excess kurtosis of the standardized residuals \( \{ \epsilon_t \} \) are \(-0.717 \) and \( 2.057 \), respectively, indicating that \( \epsilon_t \) is not normally distributed.

Figure 3(b) shows the annualized volatility series, i.e. \( \hat{\sigma}_t \sqrt{2T} \) of the GARCH(1,1) model. The GARCH(1,1) model in Eq. (5) exhibits several features of GARCH models commonly seen in practice. First, \( \hat{\alpha}_1 + \hat{\beta}_1 = 0.9811 \), which is close to but less than 1, signifying that the volatility is highly persistent. It also suggests that jumps may exist in volatility. Second, \( \hat{\alpha}_1 \) is small (close to 0.1) in magnitude, but \( 1 - 2(\hat{\alpha}_1 + \hat{\beta}_1)^2 = 0.08 > 0 \). Thus, the log returns have high excess kurtosis. Indeed, the fitted GARCH(1,1) model gives an excess kurtosis of 7.58, which is very close to the excess kurtosis 7.91 of the data; see the formula mentioned before. Third, the unconditional mean of \( \{ \sigma_t^2 \} \) series is \( \bar{\sigma}_0/(1 - \hat{\alpha}_1 - \hat{\beta}_1) = 36.98 \) which is larger than 31.99 of the sample mean of \( \sigma_T^2 \), indicating that the model might be misspecified.

For stochastic volatility modeling, we apply MCMC methods with Gibbs sampling to the data; see Chapter 10 of [19] for further details. Using 4000 iterations with the first 1000 iterations as burn-in, we obtain the model

\[ r_t = 0.864(0.139) + \sigma_t \epsilon_t \]
\[ \ln(\sigma_t^2) = 0.158(0.048) + 0.946(0.016)\eta_t, \]
\[ \hat{\sigma}_t = 0.081(0.019), \]

where the coefficients are posterior means of the last 3000 Gibbs iterations and the numbers in parentheses are posterior standard errors. Figure 3(c) shows the annualized volatility, i.e. posterior mean, of the data. The plot shows a similar pattern as that of the GARCH(1,1) model, but the volatility peaks of the 1980s appear to be sharper than those of the GARCH(1,1) model. The large AR(1) coefficient also supports the persistence in volatility.

Forecasts of GARCH models can be obtained in the same way as those of ARMA models. Here estimated parameters are often treated as given. On the other hand, one uses simulation to produce forecasts for SV models. The resulting forecasts account for parameter uncertainty, but require intensive computation.

HIGH-FREQUENCY DATA

The availability of transaction-by-transaction market data also opens a new avenue for empirical study of market microstructure. For example, intraday 5-minute returns of 345 stocks traded in the Taiwan Stock Exchange were used in [5] to establish statistically and economically significant magnet effects caused by daily price limits. By magnet effect, we mean that the price accelerates toward the limits when it gets closer to the limits. High-frequency data have many characteristics that do not occur when the data are observed at a low frequency. First, tradings do not occur at equally spaced time intervals. Second, trading intensity varies across trading hours; typically higher intensities occur at the open and close of the market in a trading day. Third, transaction prices assume discrete values. Fourth, each transaction produces several variables including time of trade, bid and ask prices, transaction price, trading volume, depths of bid and ask, etc. Fifth, the sample size is large. For example, there are more than 133,000 trades for the IBM stock during the normal trading hours in December 1999.

These characteristics give rise to special features of high-frequency data that become new challenges to statistical analysis. It is well-known that nonsynchronous trading can
lead to negative serial correlations in portfolio as well as individual stock returns. The bid-ask spread, i.e. the difference between ask and bid prices, introduces strong negative lag-1 serial correlation in stock price changes when the data frequency is high. Consequently, existence of serial correlations in observed asset returns does not necessarily contradict with the efficient market hypothesis in finance.

The characteristics also provide new research opportunities for statisticians. For example, the second characteristic leads to marked diurnal patterns in intraday series, i.e. daily seasonabilities. How to model effectively the diurnal patterns and what are the effects of the pattern on statistical properties of the data remain open. As another example, the time interval between consecutive trades not only is necessary in analyzing the unequally spaced time series but also may contain valuable information about the arrival of new events; trading tends to be heavier when there is new information available. Consideration of the information content embedded in trading intervals leads to the development of new statistical models such as the conditional autoregressive duration model in [10]. See [6] and Chapter 5 in [19] for further discussion on high-frequency financial data. In general, the impacts of institutional arrangements of security markets on financial time series become more evident when the observational frequency is high. One must handle such institutional effects properly in order to draw sensible inference about the series under study.

CONTINUOUS-TIME MODEL

Applications of financial time series are necessarily based on discrete-time observations. Much of the theory in asset pricing, however, is based on continuous-time models. An active research area in financial engineering literature is to estimate a continuous-time model using discretely observed time series data. A simple diffusion equation for asset price $P_t$ is

$$dP_t = \mu(P_t)dt + \sigma(P_t)dW_t,$$

where $W_t$ is a standard Brownian motion (or Wiener process), and $\mu(P_t)$ and $\sigma(P_t)$ are smooth functions satisfying certain regularity conditions so that $P_t$ exists via stochastic integration. For instance, the Black-Scholes pricing formulas for European options are based on the Geometric Brownian motion,

$$dP_t = \mu P_t dt + \sigma P_t dW_t,$$

where $\mu$ and $\sigma$ are constant. In this simple case, one can apply the Itô’s Lemma to obtain the distribution for the log return of the asset from time $t$ to $t + \Delta$. The distribution is Gaussian as

$$r_\Delta = \ln \left( \frac{P_{t+\Delta}}{P_t} \right) \sim N \left( \mu \Delta, \sigma^2 \Delta \right).$$

For daily log returns, $\Delta = 1/252$. Therefore, we can easily estimate $\mu$ and $\sigma$ from the daily log returns as

$$\hat{\sigma} = \frac{s}{\sqrt{\Delta}} \quad \text{and} \quad \hat{\mu} = \frac{\bar{r}}{\Delta} + \frac{s^2}{2\Delta},$$

where $\bar{r}$ and $s^2$ are the sample mean and sample variance of the returns, respectively.

In most applications the functions $\mu(P_t)$ and $\sigma(P_t)$ are unknown and there is no closed-form solution for $P_t$ available. Furthermore, to better describe the empirical characteristics of implied volatility such as volatility smile, $\sigma(P_t)$ is extended to be stochastic driven by another Wiener process. The smile effect denotes the observation that implied volatilities of option prices are not constant across the range of options but vary with strike price and the time to maturity of the option. Write $\sigma(P_t) = f(Y_t)$, where $f$ is a positive function. A general diffusion equation for $Y_t$ is

$$dY_t = \alpha(v_0 - Y_t) dt + \beta(Y_t) dz_t^*,$$

where $v_0$ is a constant denoting the long-term mean of $Y_t$, $0 < \alpha < 1$ is the rate of mean reversion of the $Y_t$ process, and $z_t^*$ is a standard Brownian process correlated with $W_t$. A particular process for $Y_t$ that attracts much attention is the Ornstein-Uhlenbeck process

$$dY_t = \alpha(v_0 - Y_t) dt + \beta dz_t^*.$$

For a given starting value $Y_0$, $Y_t$ is a Gaussian process with mean $v_0 + (Y_0 - v_0)e^{-\alpha t}$ and variance $\beta^2(1 - e^{-2\alpha t})/(2\alpha)$, which has an invariant distribution $N(v_0, \beta^2/(2\alpha))$ as $t \to \infty$. Another commonly used model for $Y_t$ is the Cox-Ingersoll-Ross (CIR) model

$$dY_t = \alpha(v_0 - Y_t) dt + \beta \sqrt{Y_t} dz_t^*.$$
Figure 4. Weekly (Wednesday) U.S. three-month treasury bill yields from 1/1/1954 to 10/5/1997.

Table 2. Statistics and Estimation Results for Weekly U.S. 3-Month Treasury Bill Rates From 1/1/1954 to 10/5/1997

<table>
<thead>
<tr>
<th>(a) Descriptive statistics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Article</td>
<td>Mean</td>
<td>Median</td>
<td>St. Dev.</td>
<td>Skewness</td>
</tr>
<tr>
<td>Eraker</td>
<td>0.0555</td>
<td>0.0154</td>
<td>0.0286</td>
<td>1.0865</td>
</tr>
<tr>
<td>Tsay</td>
<td>0.0555</td>
<td>0.0513</td>
<td>0.0286</td>
<td>1.0866</td>
</tr>
</tbody>
</table>

| (b) Estimates of CEV model with $\Delta t = 1$ |
|---------------------------|------------------|------------------|------------------|------------------|
| Article | Stat. | $\theta_r \times 1000$ | $\kappa_r$ | $\sigma_r$ | $\beta$ |
| Eraker | Mean | 0.1589 | -0.0025 | 0.0189 | 0.7608 |
| St.Dev. | 0.0551 | 0.0014 | 0.0010 | 0.0175 |
| Tsay | Mean | 0.1589 | -0.0025 | 0.0188 | 0.7605 |
| St.Dev. | 0.0547 | 0.0014 | 0.0011 | 0.0182 |

| (c) Estimate of SV model with $\Delta t = 1$, rates are in percentages. |
|---------------------------|------------------|------------------|------------------|------------------|
| Article | Stat. | $\theta_r \times 1000$ | $\kappa_r$ | $\sigma_r$ | $\kappa_z$ | $\sigma_z$ | $\beta$ |
| Eraker | Mean | 0.0123 | -0.0020 | 0.0392 | -0.0201 | 0.2457 | 0.7813 |
| St.Dev. | 0.0020 | 0.0005 | 0.0079 | 0.0056 | 0.0267 | 0.1257 |
| Tsay | Mean | 0.0176 | -0.0027 | 0.0257 | -0.0139 | 0.1578 | 1.2636 |
| St.Dev. | 0.0006 | 0.0016 | 0.0054 | 0.0057 | 0.0189 | 0.0915 |

Combining the two diffusion equations, we have a general continuous-time model with stochastic volatility:

$$dP_t = \mu(P_t)dt + \sigma(P_t)dw_t$$  \hspace{1cm} (6)  
$$\sigma(P_t) = f(Y_t)$$  
$$dY_t = \mu_y(Y_t)dt + \sigma_y(Y_t)dz_t^*.$$  \hspace{1cm} (7)

In practice, we can rewrite $z_t^*$ as $z_t^* = \rho w_t + \sqrt{1 - \rho^2} z_t$, where $z_t$ is a standard Brownian motion independent of $w_t$ and $\rho$ is the instantaneous correlation coefficient between $w_t$ and $z_t^*$. The correlation $\rho$ is often found to be negative and is referred to as the leverage effect between stock price and its volatility.

Estimation of the diffusion equations in Eqs. (6) and (7) based on discretely observed...
data is an important issue in derivative pricing. For a time interval $\Delta$, a discrete-time approximation of the model is

$$P_{t+\Delta} - P_t = \mu(P_t)\Delta + \sigma(P_t)\sqrt{\Delta}\epsilon_t$$

$$Y_{t+\Delta} - Y_t = \mu_y(Y_t)\Delta + \sigma_y(Y_t)\sqrt{\Delta}\epsilon'_t$$

where $\epsilon_t$ and $\epsilon'_t$ are bivariate normal random variables with mean zero, variance 1, and correlation $\rho$. This naive discretization of the diffusion equations may lead to discretization bias in parameter estimates, especially when $\Delta$ is large. Clever statistical methods have been devised to treat values of the processes between two consecutive observations as missing values to reduce the approximation bias. This is often done via the MCMC methods as in [11] and [8]. On the other hand, if $\Delta$ is too small, then the effect of market microstructure on $P_{t+\Delta} - P_t$ becomes evident. A trade-off between approximation accuracy and observational noise caused by market microstructure is needed. Other methods of estimating diffusion equations include the efficient method of moments in [13], nonparametric methods in [1], and the generalized method of moments in [15], among others. All of the available methods require intensive computation. They are used mainly for daily or weekly data. Finally, random jumps can be added to the diffusion equations in Eqs. (6) and (7) to better describe the empirical characteristics of implied volatility.

Illustration

Figure 4 shows the weekly (Wednesday) U.S. treasury bill yields from 1/1/1954 to 10/5/1997 for 2288 observations. This series was used in [11] to illustrate the estimation of diffusion equations. We use the same data here with our own programs to reproduce the results. Table 2(a) gives the descriptive statistics of the data that match nicely with those given in [11]. First, consider the simple constant elasticity of variance (CEV) model in the form

$$dY_t = (\theta + \kappa Y_t)dt + \sigma Y_t^\beta dw_t,$$

where $Y_t$ is the interest rate. Except for $\beta$, the conditional posterior distributions of the parameters are standard and can be drawn easily in MCMC iterations. A hybrid accept/reject Metropolis-Hastings algorithm is used to draw $\beta$. Table 2(b) gives the estimation results based on 20,000 Gibbs iterations. It is reassuring to see that results of two independent estimations are in good agreement. Next, consider the stochastic volatility (SV) model in the form

$$dY_t = (\theta + \kappa Y_t)dt + \sigma Y_t^\beta dw_{1t},$$

$$dV_t = \kappa V_t dt + \sigma dw_{2t},$$

where for simplicity $w_{1t}$ and $w_{2t}$ are two independent standard Brownian motions. Here the computational intensity increases substantially because the volatility series is a latent process. Table 2(c) shows the result of SV model for the U.S. interest rate series. The interest rates are in percentages. The estimates for the interest rate diffusion equation are similar to those of the CEV model, but they show some differences with those of Eraker (2001).

Acknowledgments

The author’s work was supported in part by NSF under the grant DMS-9971903 and by the Graduate School of Business, University of Chicago. The author thanks Mr. J.C. Artigas for developing the estimation programs.

REFERENCES


See also ARCH AND GARCH MODELS; FINANCE, STATISTICS IN; and TIME SERIES.

Ruey S. Tsay