

Realistic Statistical Modelling of Financial Data

Tina Hviid Rydberg

BNP Paribas, London, UK. E-mail: tina.rydberg@bnpparibas.com

Summary

The aim of this paper is to present some of the stylized features of financial data which have received a lot of attention both from practitioners and those with more theoretical backgrounds. Some of the models resulting from these efforts are reviewed and discussed. To facilitate the discussion two data sets are used: one of these contains all US trades in IBM stocks in 1995 at NYSE.

Key words: Aggregational Gaussianity; Asymmetry–symmetry; Continuous time models; Discrete time models; Fat tails; High frequency data; Pricing derivatives; Seasonality; Serial correlation; Volatility clustering; Quasi long range dependence.

1 Introduction

The evaluation of risk and the pricing of derivative assets are based on statistical models describing underlying asset prices. The quality of the risk measures and the validity of prices are therefore strongly dependent on how well the statistical model captures the behavior of the underlying asset. Shortcomings of the statistical model can be disastrous, and not only for financial speculators. For example anyone who has an equity linked pension scheme will be seriously affected by mispriced derivatives as most pension funds and insurance companies generate a large part of their surplus by holding a large portfolio of financial assets. The market for mortgages is another example of an industry that benefits from products created based on statistical models. How else could building societies or banks offer mortgage deals with fixed interest rates, when the real market interest rate set by, e.g. the European Central Bank or the Bank of England is variable.

Popular models such as the Black & Scholes (1973) model, based on the geometric Brownian motion, have very nice mathematical properties which have been extensively used to price options. These models, however, lack most of the empirically found features of financial data. More elaborate models can therefore be found in the literature which try to encompass these empirical findings. Statistically these newer models have nice properties in terms of both fitting the observed data and telling economically meaningful stories. But from a mathematical finance point of view, where the focus is on pricing derivatives, they have serious drawbacks. This is because models such as the stochastic volatility models generate markets where derivative prices are not uniquely determined: how to solve this problem is largely unresolved.

Pricing of options and other derivative assets are not the only reasons for building statistical models. One may also wish to gain insight into the market microstructure to answer questions such as: how do different types of trades influence the market? or, how do news releases influence the trading pattern and price movements? Answering these types of questions also requires a statistical model.

The focus of this paper is the statistical modelling of financial data, i.e. the econometrics of

financial data. In order to keep an open mind on how to resolve the modelling problem, the features of financial data, rather than the models, are accentuated.

Two data sets, described in Section 2, are used throughout the paper to facilitate the discussion. Section 3 describes some of the stylized features of financial data, which need to be taken into account when developing models. The stylized facts discussed in Section 3 are not only valid for the two data sets used here, but have shown to be extremely general and present across different types of financial assets and sampling frequencies. These stylized features demonstrate why this is an intellectually challenging area for statisticians to participate in.

The paper aims to provide a basis of knowledge about how far financial econometrics research has developed. Therefore, despite our wish not to promote any particular type of modelling approach, Section 4 of the paper includes brief summaries of discrete and continuous time models. More extensive survey articles of these particular areas may be found in Bollerslev, Engle & Nelson (1994), Ghysels, Harvey & Renault (1996) and Shephard (1996). In Section 4.2.4 very recent research of high frequency financial data is discussed. Section 5 considers problems relating to pricing and hedging in some of the presented models. Section 6 concludes.

2 Data

A financial data set usually consists of a record of trades in stocks or bonds traded at some stock exchange, or, by quoted interest rates or foreign exchange rates. Other examples would be commodities traded at the Chicago Board of Trade (CBOT).

The data sets can be *intra-daily* “tick-by-tick” trades data (e.g. a tick was 1/8 of a dollar in 1995 at the New York Stock Exchange). This means that every single trade in a specific stock or bond is recorded, with the time the trade took place, the amount traded, settling price (the price at which the trade took place), settling date and sometimes what type of traders took part in the trade. These different types of traders include market makers, banks and brokers (these are not disjoint groups).

An *intra-daily* data set could also consist of the quotes. This would include the time the quote is posted, the bid price (the price at which the market maker is willing to buy), ask price (the price at which the market maker is willing to sell) and the depth (how many shares the market maker is willing to buy/sell at the best bid/ask).

If the data set is constituted of only *daily data*, the most you can hope for is opening price, closing price, highest price, lowest price, traded amount (volume) and number of trades. More commonly the daily data sets analyzed in the literature contain only closing prices.

The data used to facilitate the discussion in this paper are:

1. daily closing prices of IBM stocks from DataStream. They cover the period January 2, 1973 to January 25, 1996. The data set consists of dates and prices. The data are shown in Figure 1.
2. Intra-daily prices of IBM stocks from NYSE (New York Stock Exchange), recorded in the TAQ (trades and quotes) data base. They cover the period January 3, 1995 to December 29, 1995. The data set consists of dates, exchange (New York, Cincinnati, Pacific, . . .), time, price and amount. The length of the total data set when all exchanges in the US are considered is 413,906. Figure 2 shows all the prices of trades on four individual days.

For notational convenience we make the following definitions. Let $\{S_t\}$ denote the stock price process and let

$$P_t = \log S_t - \log S_{t-1},$$

denote the logarithmic increase of S over the interval $(t-1, t]$. P_t is usually referred to as the “log returns” or simply the “returns”. The returns for the daily prices of IBM stocks are shown in Figure 3.

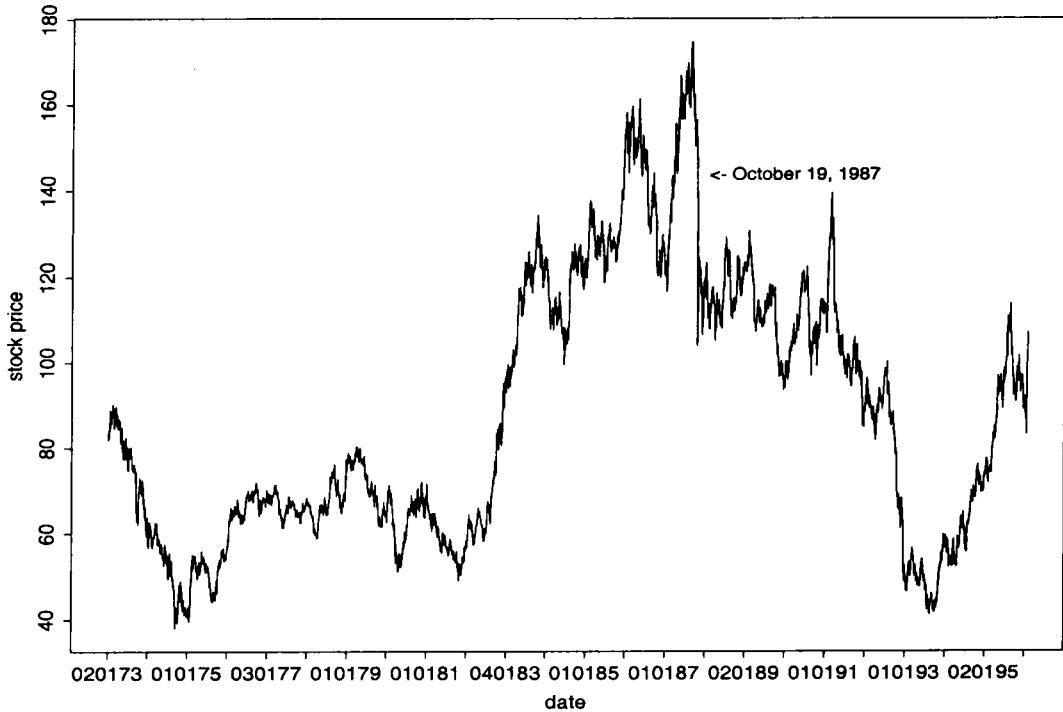


Figure 1. Plot of the IBM prices in the period January 2, 1973 to January 25, 1996. Note for instance the large price changes around October 19, 1987, also known as “Black Monday”.

The estimates of the mean and the empirical variance are the usual

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \text{var}(x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2,$$

and the skewness and kurtosis are given by

$$s(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3 / \text{var}(x)^{3/2}, \quad k(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4 / \text{var}(x)^2$$

For the data sets used here these numbers can be found in Table 1 (See also Campbell, Lo & MacKinlay (1997) who give other versions of the formulae for s and k). Note how the kurtosis decreases when the data are aggregated. Unfortunately, the kurtosis for the half hourly returns is not directly comparable since only data from 1995 were available at this frequency. The data set for the period January 2, 1973 to January 25, 1996 contains the crash in 1987 which completely changes the level of the kurtosis, see Figure 4.

3 Stylized Features of Financial Data

The “stylized features” of financial data will be discussed below. When building new models to describe financial data it is a great help to have some clear-cut guidelines that describe the most important features which the models must encompass. These guidelines are referred to as “stylized

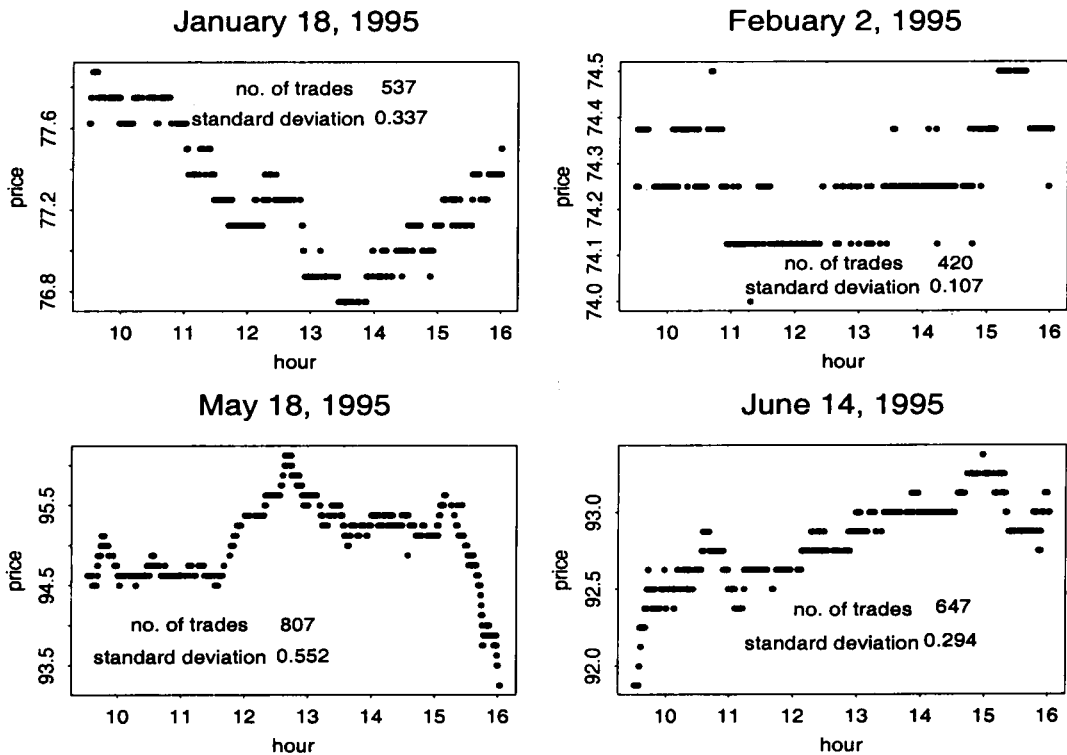


Figure 2. Plot of all traded IBM prices at the New York stock exchange on four different days in 1995. A trade is represented as a dot.

features" or "stylized facts". For daily and lower frequencies of sampling the stylized features are quite well known and some of them are described in for example, Taylor (1986, Ch. 2). For the higher frequencies they are not well established, because the availability of intra-day data is rather new. However, for foreign exchange rates Guillaume, Dacorogna, Davé, Müller, Olsen & Pictet (1997) have reported what they believe to be stylized facts of intra-daily foreign exchange rates. Another source of information on features of high frequency data is Ghysels, Gouriéroux & Jasiak (1998).

Table 1

Estimated parameters for IBM log returns in the period January 2, 1973 to January 25, 1996. Note that the half hourly returns are only from 1995 so these numbers are not really comparable. Also the low frequencies contain 1987 which really contaminates the picture, see Figure 4. The corresponding empirical densities are shown in Figure 7.

resolution	mean	variance	skewness	kurtosis
half hour	-4.6×10^{-5}	1.4×10^{-5}	-0.691	13.27
daily	4.01×10^{-5}	0.0002	-0.638	22.36
weekly	0.0002	0.0010	-0.107	6.26
monthly	0.0004	0.0043	-0.354	4.44

Fat tails

Fat tails imply that the distribution of the returns has tails which are heavier than the tails of a normal distribution, in other words the tails decrease slower than $\exp(-x^2)$. The most celebrated

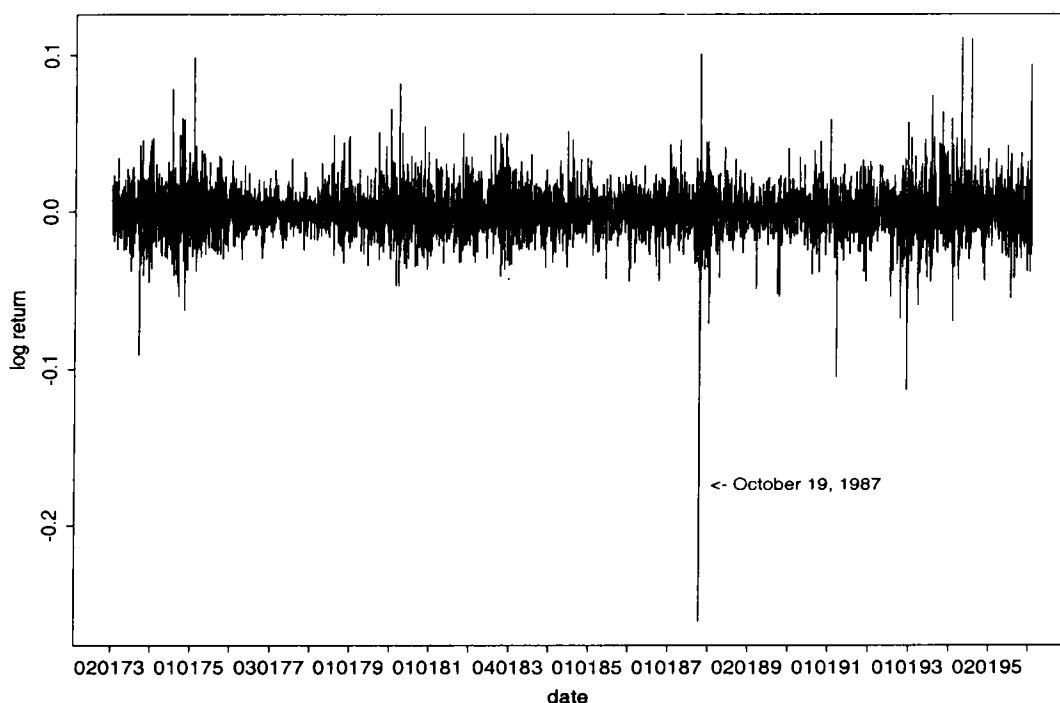


Figure 3. Plot of the IBM log returns in the period January 2, 1973 to January 25, 1996. The log returns exhibits bulks of large price changes.

model in mathematical finance is the previously mentioned Black & Scholes (1973) model which assumes normally distributed log returns. The assumption of normally distributed log returns has, however, long been a disputed topic.

Mandelbrot (1963) argued that the price changes of cotton futures showed much heavier tails than normal; in fact, he claimed that the distribution was a stable law without a second moment. However, because the serial correlation was not taken into account, Mandelbrot's conclusion has been disputed. It is still an ongoing debate how many moments actually exist. There does seem, however, to be general agreement that daily data do have finite variance and thereby the class of stable laws are not suitable models. For a discussion of this see Eberlein & Keller (1995).

For the data considered here the log densities are shown for different resolutions in Figure 7. In Figure 4 we also see how dramatic an effect 1987 has on the statistics of later years ("Black Monday" was on October 19th, 1987. Interestingly, in October 1997, exactly 10 years later, the stock market faced another major crash.). This shows that one should be very careful using statistics based on long time series, i.e. low frequencies, without checking if they are consistent with statistics of higher frequencies. Of course, 1987 is very extreme and one could probably screen some of the observations around October 19th, but, which ones?

When it comes to very high frequencies of intra-daily data the question is much more difficult. In two different markets, the foreign exchange and the inter money bank, Müller, Dacorogna & Pictet (1996) find evidence that the second moment exists but the fourth moment might not. This is not only of relevance when searching for the best fitting distribution but also when studying the serial correlations of the squared or absolute returns, which will be discussed later. See also Davis & Mikosch (2000) who discuss the problems related to estimating the sample autocorrelation in models with heavy tails.

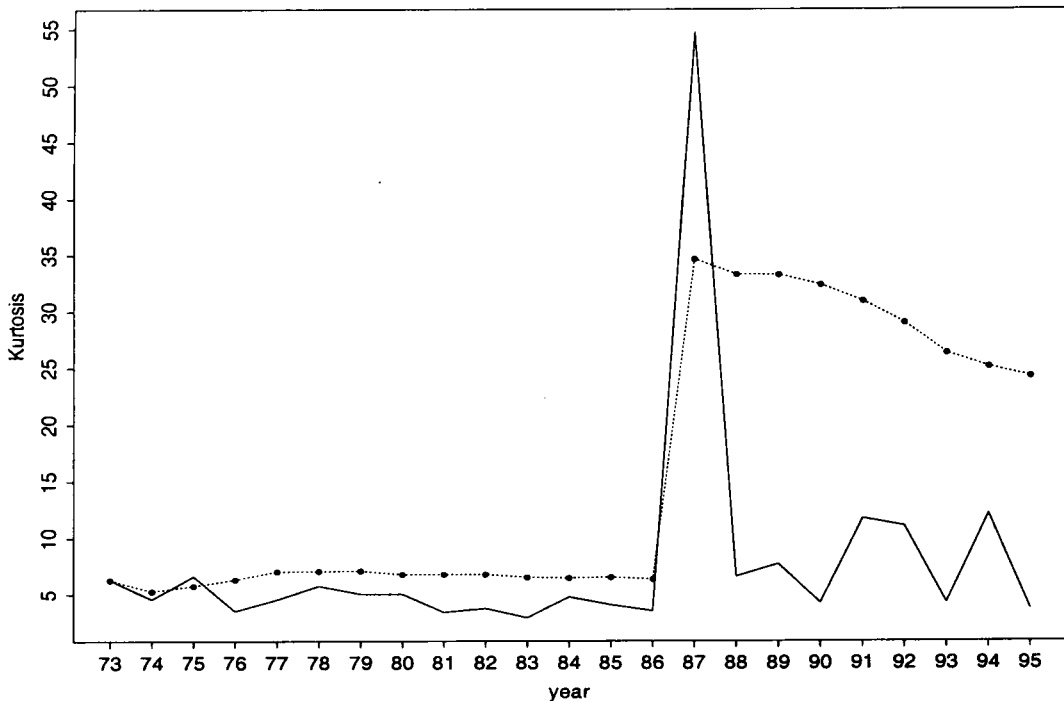


Figure 4. Plot of the empirical kurtosis of the IBM log returns in the period January 2, 1973 to January 25, 1996. The solid line connects the kurtosis of individual years and the dashed line is the kurtosis of the whole period from 1973 to that year.

If the model for the returns does not have enough probability mass in the tails the price and hedge strategies based on the model will not take the large returns into account (see Section 5 for a definition). This will result in seriously mispriced derivatives and hedge strategies which do not cover all of the risk concerned with holding a specific portfolio.

Just one word of caution here. As we shall see, most financial time series have quite persistent correlation structures. If this is not modelled correctly we may get an incorrect picture of the tail behavior. Specifically, this could lead to a model where the tails are in fact too heavy.

Asymmetry–Symmetry

There is evidence that the distribution of stock returns are slightly negatively skewed. One explanation could be that traders react more strongly to negative information than positive information. As an example see Figure 1 where the drop on October 19th, 1987 is very big and afterwards it takes many much smaller increases to get back to the previous level. This type of behavior, in daily returns, is specifically modelled by the EGARCH model (see Section 4.1.3 for further discussion of this phenomena). For high frequency data we also find evidence of negatively skewed tails (see Rydberg & Shephard (1998a)).

Another type of asymmetry is the asymmetry in the lagged correlation. In Figure 5 and 6 it is seen that weekly “coarse volatility”, as defined in formula (1) predicts weekly “fine volatility” better than the opposite way around. This fact was first noticed for exchange rates in Müller, Dacorogna, Davé, Olsen, Pictet & Weizsäcker (1996). In other words this means that

$$\varrho(F_w, C_{w-1}) \geq \varrho(C_w, F_{w-1}),$$

where ϱ is the correlation function and C and F are given by:

$$C_w = |P_{w_5} + P_{w_4} + P_{w_3} + P_{w_2} + P_{w_1}| = |\log S_{w_5} - \log S_{w_0}|, \quad (1)$$

and

$$F_w = (|P_{w_5}| + |P_{w_4}| + |P_{w_3}| + |P_{w_2}| + |P_{w_1}|) / 5,$$

where w denotes the week and w_i the days of the week (see Müller, Dacorogna, Davé, Olsen, Pictet & Weizsäcker (1996)). In words C can be expressed as the absolute price change over a full week. For any one stock exchange a week is a 5 day period, since weekends are not counted. The weekly “fine volatility” F is the mean of the absolute daily price changes over a full week.

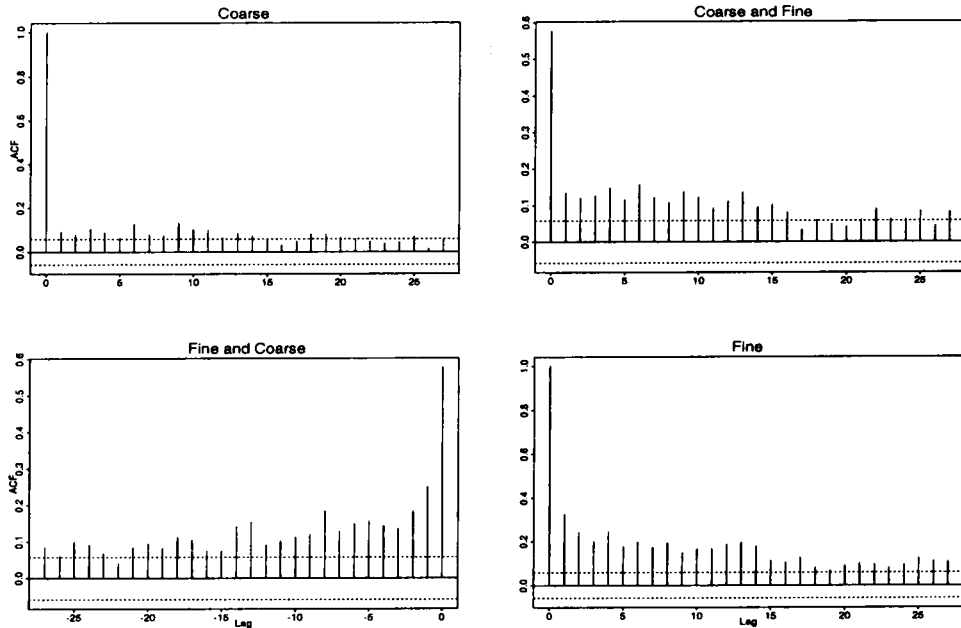


Figure 5. Plot of the autocorrelation between the fine and the coarse volatilities of daily IBM prices in the period January 2, 1973 to January 25, 1996.

Volatility Clustering

The term volatility clustering refers to the fact that large price changes come in clusters, again see Figure 1. See also Figure 3 which shows the log returns, here the clustering is more visible. Creating mathematical/statistical models that describe this feature has been the aim of a lot of research activity. For proposed models in the econometrics literature; see, for instance the review papers on ARCH by Bollerslev, Chou & Kroner (1992), Bollerslev, Engle & Nelson (1994) and Shephard (1996) (see Section 4 of this paper for a very short treatment).

A very different approach is taken in the mathematical finance literature see for instance, the review paper on diffusion stochastic volatility (SV) models by Taylor (1994). In this paper Taylor compares ARCH and SV models. However, presumably neither of these models would have been able to model the price changes around October 19th, 1987, since the price change of that Monday was extreme, in fact it was approximately 23%. The crash of the market started on the preceding

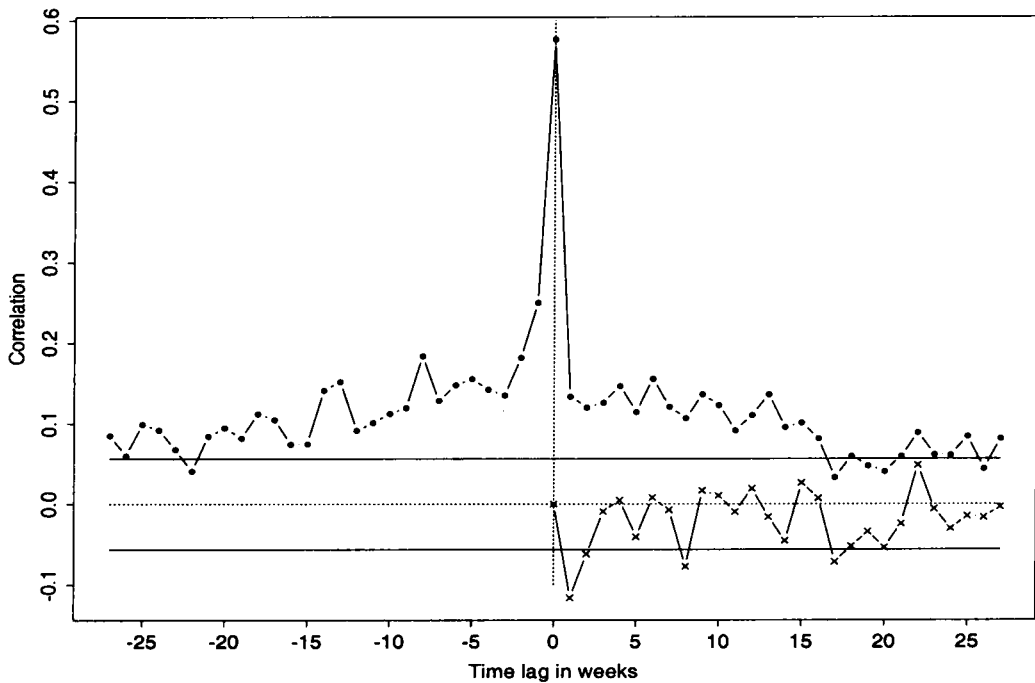


Figure 6. Plot of the lagged correlation q_t of the fine and coarse volatilities of daily IBM prices in the period January 2, 1973 to January 25, 1996. The curve with \times as marks is the difference $q_t - q_{t-1}$. The solid lines indicates $\pm 1.96/n^{1/2}$.

Friday. One often used explanation is that the market was over valued. If such extreme events should be taken into account it would be necessary to add extra components to the models with features derived from risk theory. However, it is an open question whether such extremes should be modelled directly or separately.

When we are only dealing with daily data it is not possible to give much further description of why the price changes the way it does and the reasons for the volatility clusters. If we had more information available such as the number of trades, the trading volume and the variation within a specific day, it would help us gain insight. We will return to this subject in Section 4.

Economic information and news play an important role in the formation of the price process and the traded volume. Andersen & Bollerslev (1998) report on volatility increases around releases of macroeconomic reports. Also the information content of the observed trades series influences the prices process. Without exaggeration, one could say that, the market is a bit hysterical, because some types of information like big sell orders might trigger a big group of people to sell out of a specific asset. This, of course, causes the price to fall. In this sense the volatility is driven by information.

Aggregational Gaussianity

For decreasing sampling frequencies the central limit law sets in and the distribution of the log returns tends towards a Gaussian law. Decreasing sampling frequencies is equivalent to aggregation since,

$$\log S_5 - \log S_0 = (\log S_5 - \log S_4) + \dots + (\log S_1 - \log S_0) = P_5 + \dots + P_1.$$

An example of this is shown in Figure 7 where the empirical densities for four different sampling frequencies are shown together with the estimated normal density.

This stylized fact also rules out the Stable laws as suitable models for describing the marginal log

returns, as they do not aggregate to a normal distribution. Together with the fact that log returns have fat tails, this suggests that natural models for describing the marginal distributions of the log returns should be families of laws which are fat tailed, closed under convolution (this would be desirable when working with portfolios, as the value of a portfolio is the sum of the individual assets) and contain the normal law as a limit. One class of distributions which allows for this is the normal inverse Gaussian. In fact this particular class of distributions has been shown to be capable of fitting log returns on many kinds of assets, see Barndorff-Nielsen (1995), Barndorff-Nielsen (1997), Rydberg (1997a) and Rydberg (1999).

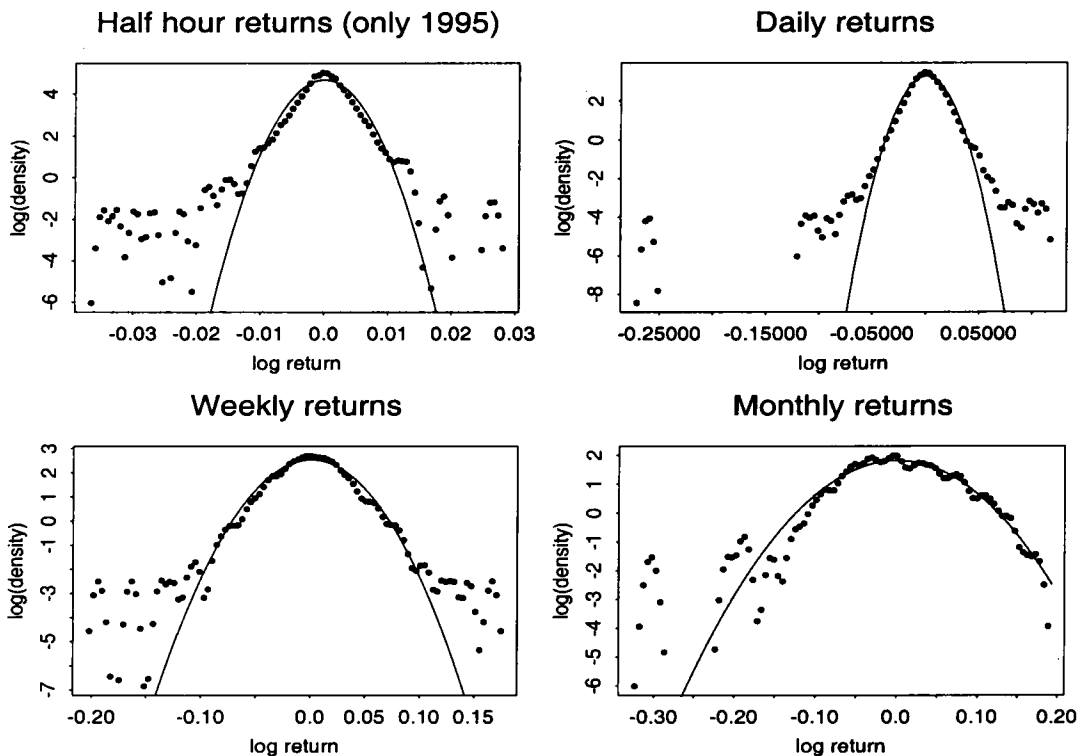


Figure 7. Plot of the empirical densities (rescaled histograms) of the IBM log returns in the period January 2, 1973 to January 25, 1996 and the estimated normal density. Again the half hourly data are only from 1995. The ordinate axis is on a log scale in order to exhibit the tails clearly.

Quasi Long Range Dependence

The log returns themselves, of all kinds of assets (e.g. see Figure 9) hardly show any serial correlation but this does not mean that they are independent. In order to study the dependence and correlation structure of the log returns it is necessary not only to study the log returns themselves, but also an appropriate function of the log returns. In this context “appropriate” is related to the variance i.e. an appropriate function is a function that reveals more information about the variance of the asset prices.

The most widely used function is probably the square, which does reveal information about the serial correlation of the variance, since it enlarges large returns and diminishes small ones. Using the squared log returns might pose a problem, however, since in some cases, for instance for foreign exchange rates, the 4th moment might not exist (see Guillaume, Dacorogna, Davé, Müller, Olsen & Pictet (1997)). For the data studied here it should be possible to use the 4th moment, see Figure 4.

Another possible function is the absolute value of the log returns. The absolute value will compare the log returns size-wise, and a strong correlation between first lags will then, like the use of the squared log returns, reveal volatility clustering, i.e. clusters of big movements of the price.

For a comparison of squares and absolute values see Figure 9. Note that, in the interval of interest, the absolute values of the log returns span a much larger interval than the squared log returns. How much that actually is can be seen from Figure 8. It is also striking that the absolute returns show a much stronger persistence than the squared returns. This is in accordance with Ding, Granger & Engle (1993), who in a study of serial correlation look at functions of the form $|x_t|^d$ and find as an empirical fact that the values of d close to 1 are the ones for which the property of high autocorrelations for long lags is most significant. However, the absolute values are much more difficult to handle analytically than the squares, so perhaps this is the reason why the square is the preferred function.

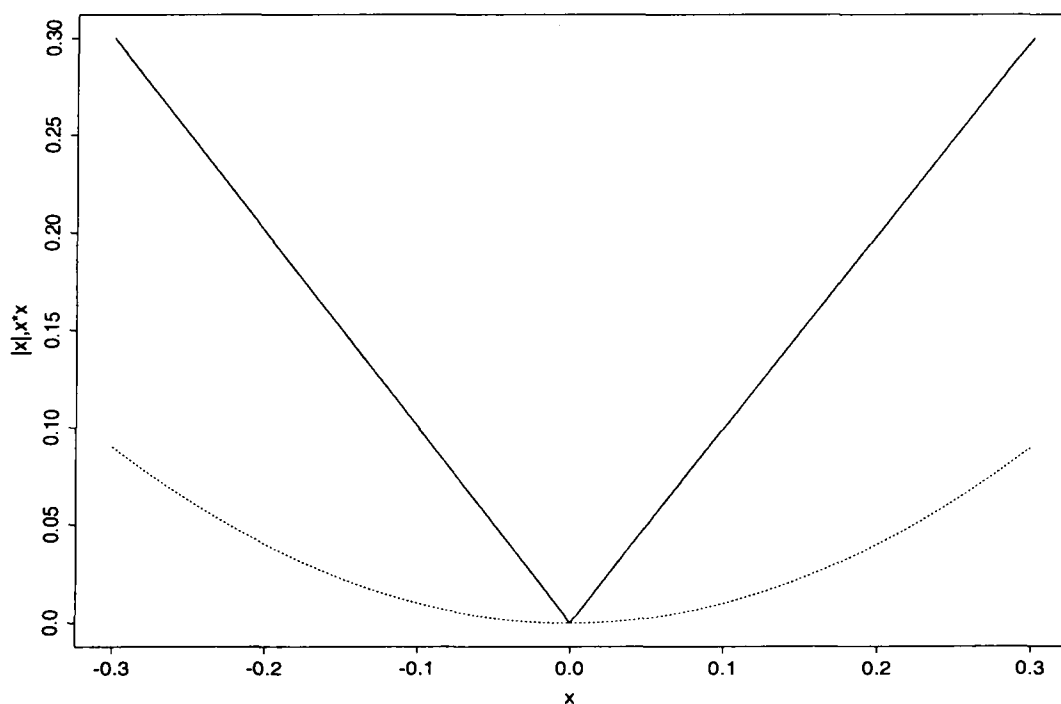


Figure 8. Plot of $|x|$ and x^2 in the range where the log returns lie. x^2 is represented by the dashed line. It is worth noticing how much x^2 smooths out the differences in this area.

The result of using a model where the correlation is misspecified is similar to the result of using a model where the distribution is misspecified. This is because large returns can be a result of either, uncorrelated events which have a distribution with a thick tail, or, negatively correlated events. However, it is often difficult to distinguish the effects of large negative correlations in the first few lags and the tail behavior of the distribution of the returns, especially if one is not searching for it.

Seasonality

For intra-daily frequencies the above stylized features still hold but some extra complexity is added. For a thorough study of high frequency foreign exchange rates see Guillaume, Dacorogna, Davé, Müller, Olsen & Pictet (1997). Exchange rates and stock prices are not directly comparable. One big difference is that the exchange rates are something world wide, e.g. the US Dollar is traded against the Deutsch Mark or the Euro in all of the four major markets: the Pacific, covering Australia

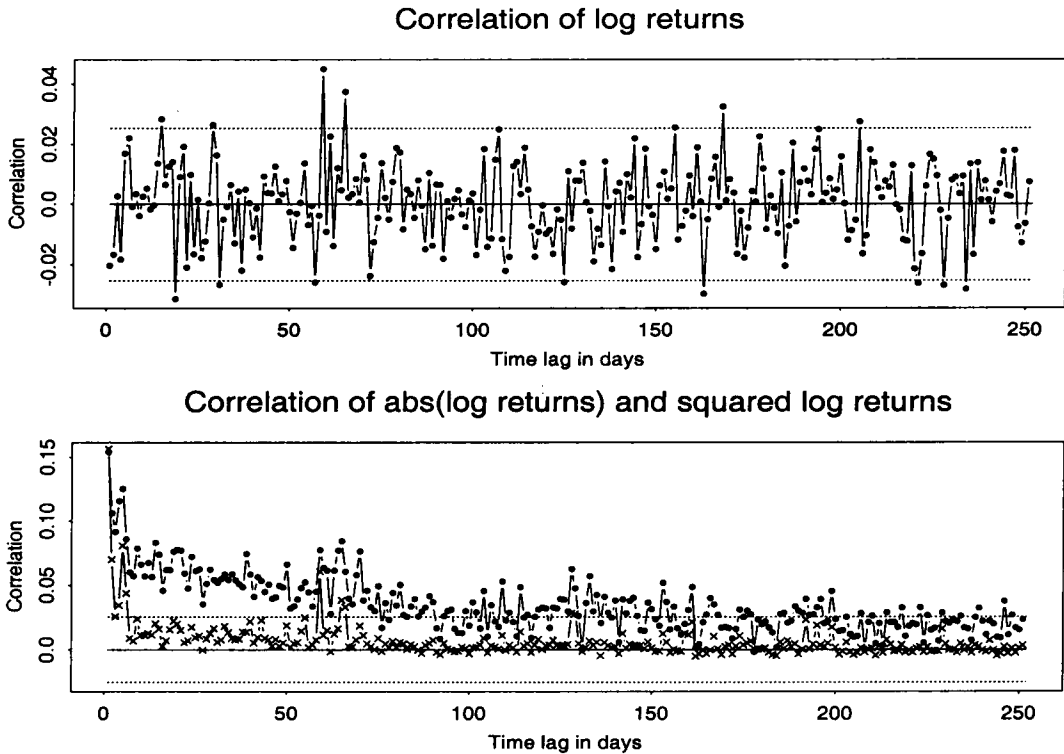


Figure 9. Plot of the auto correlation function for the IBM log returns in the period January 2, 1973 to January 25, 1996. In the bottom picture the \times represents the squared log returns. The dashed lines indicates $\pm 1.96/n^{1/2}$. 250 lags are shown since this corresponds to the number of days in a business year.

and New Zealand; the Far East, covering Tokyo, Singapore and Hong Kong; the European market with the major exchanges being London and Frankfurt am Main; and the North American with New York as the major exchange. The opening hours of these markets span almost all 24 hours of the day, except for a gap between the closure of the American markets and the opening in the Pacific and the Far Eastern, thereby the world market does not really close, see Guillaume, Dacorogna, Davé, Müller, Olsen & Pictet (1997).

Stocks are mainly traded at the national stock exchanges. Therefore, they are of a different nature to foreign exchange rates. The prices of stock trades are quoted within the business day which is usually somewhere between the hours of 8.30 and 16.30. The fact that the market closes overnight does in some instances create a strong overnight effect, perhaps due to the larger amount of news revealed to the traders while the market is closed.

In any case, the market is very busy in the mornings, slows down around lunch, and picks up again in the afternoon, see Figure 11. It is very clear from Figure 11 that the New York stock exchange is only open from 9.30–16.00. On a weekly basis one sees that the trading is slower on Mondays than the rest of the week, and the strength of this phenomena for the IBM stocks can be seen from Figure 10.

4 Models

Many different types of models have been suggested for the modelling of financial data, some more complicated than others. These quite different types of models are the topic of this section.

The least ambitious goal, which has received a lot of attention since the beginning of this century,

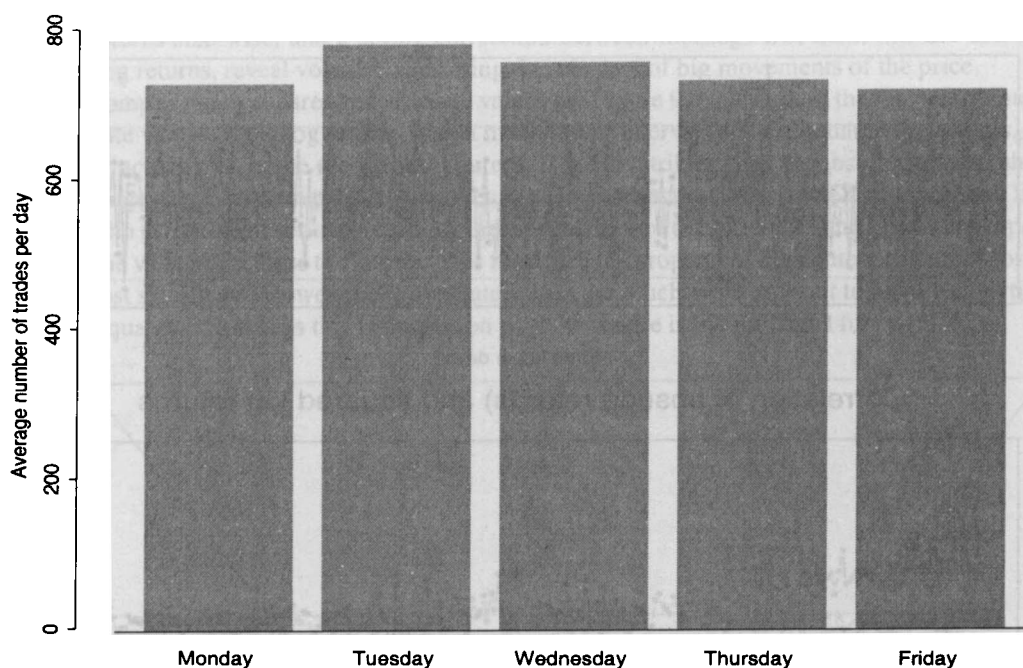


Figure 10. *The histogram depicts the average number of trades in the IBM stock per day in a given day of the week in 1995.*

starting with Bachelier (1900), is to give a statistical/mathematical model which is capable of describing the behavior of a single stock via one stochastic process. Even this simple scenario is not easy and several very different types of models have been suggested.

Furthermore, the type of proposed models are very much influenced by specific goals. The goal being, for instance, pricing of options; in which case one strives for models that allow for an equivalent martingale measure, since this makes pricing of derivative assets easy. This branch of finance tends to be highly loaded with difficult probability theory, especially martingale theory. This branch of finance is usually called “mathematical finance”. Another goal could be forecasting, in which case it is the correlation structure that is put under the microscope. It is this area of finance that uses the most statistical methods and is referred to as “econometrics”.

In the following two sections several models proposed for modelling financial data are described and discussed. The models are divided into two main groups: discrete time models and continuous time models. In principle the models could also have been divided according to whether the model belongs to the econometric or the mathematical finance literature. However, division according to a mathematical concept, rather an apparent disciplinary boundary, is more intellectually justified.

4.1 Discrete Time Models

The discrete time models have been quite popular, one reason being that they (usually) allow for maximum likelihood analysis. This is a much more complicated matter for continuous time models, such as diffusion processes, see Pedersen (1995), Bibby & Sørensen (1995), Gallant & Tauchen (1996), Gallant & Long (1997) and Elerian, Chib & Shephard (1998). Even though it is easier to write down the likelihood for the discrete time models estimation can still be rather difficult when innovations are non Gaussian. See Shephard (1996) who discusses the use of Markov chain Monte

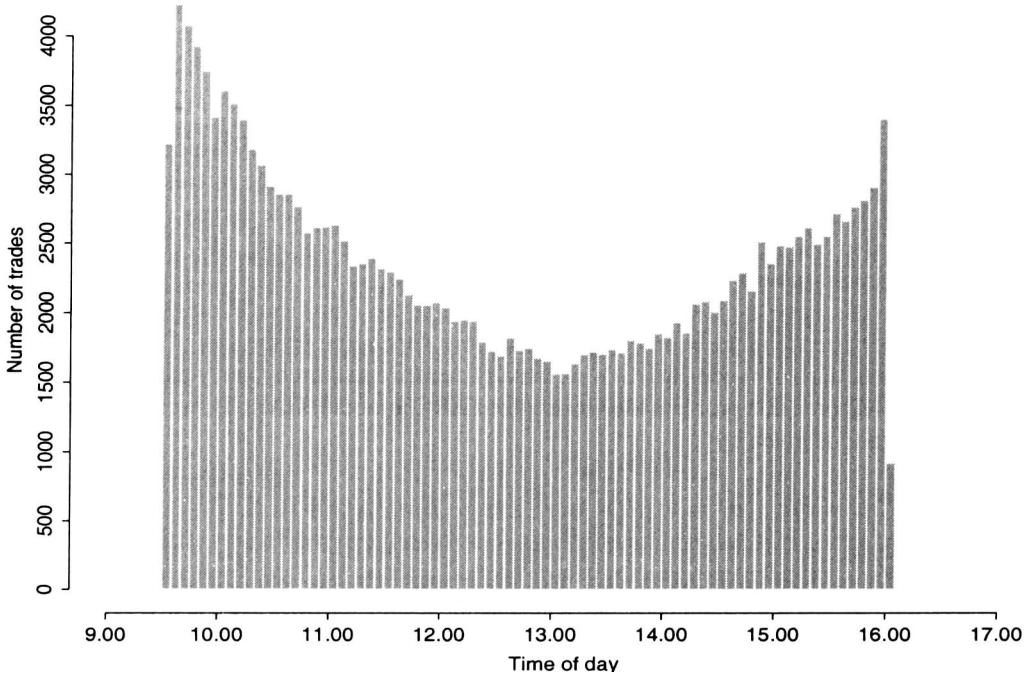


Figure 11. Number of trades as a function of 5 minute periods.

Carlo methods and the Kalman filter. Many financial time series are only observed at fixed discrete time points, as is the case if the data set consists of daily closing prices. The discrete nature of such types of data are also one of the motivations behind discrete time models.

4.1.1 ARCH type models

The most widely used discrete time model is probably the ARCH model (Auto Regressive Conditional Heteroscedasticity) developed by Engle (1982). The simplest ARCH model, the ARCH(1) is given by

$$y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = w + \alpha y_{t-1}^2, \quad t = 1, \dots, T,$$

where the ε_t 's are independent identically distributed and $\varepsilon_t \sim N(0, 1)$. In order to ensure that $\sigma_t^2 \geq 0$ for all t we need $\alpha \geq 0$. For the ARCH(1) model

$$\frac{E(y_t^4)}{(E(y_t^2))^2} = 3 \frac{1 - \alpha^2}{1 - 3\alpha^2},$$

if $3\alpha^2 < 1$. If this condition is fulfilled then the process for the squares, y_t^2 , is covariance stationary and the autocorrelation function is given by

$$\rho(s) = \alpha^s.$$

A general ARCH(q) model is given by

$$\sigma_t^2 = w + \sum_{i=1}^q \alpha_i y_{t-i}^2.$$

Now, if we rewrite the innovation part, we see that

$$\varepsilon_t^2 = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + m_t$$

where $m_t = \varepsilon_t^2 - \sigma_t^2$ which is a martingale difference. In addition, the model corresponds to an AR(q) for the squared innovations. Certain regularity conditions are needed, and the interested reader may consult Bollerslev, Engle & Nelson (1994), which is a very comprehensive survey study of ARCH type models. See also Shephard (1996), where the statistical aspects of both ARCH and stochastic volatility models are discussed.

The ARCH models have excess kurtosis accommodating the empirical findings. Furthermore, they show some persistence in the squared autocorrelations depending on the size of q . In empirical applications of the ARCH(q) model large q 's are usually needed and thereby many parameters (α_i 's). In an attempt to overcome this problem the Generalized ARCH model, GARCH, was introduced by Bollerslev (1986) and is given by

$$y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = w + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad t = 1, \dots, T.$$

Neither the ARCH nor the GARCH models take the asymmetry into account. Also, ARCH and GARCH do not model what is called *leverage*, which is the fact that volatility is negatively correlated with changes in stock returns in the sense that "bad news" tends to give an increase in volatility and "good news" a decrease in volatility. A model which is developed in Nelson (1991) with these aspects in mind is the Exponential GARCH, EGARCH, model

$$\log \sigma_t^2 = \alpha_t + \sum_{k=1}^{\infty} \beta_k g(y_{t-k}), \quad \beta_1 \equiv 1$$

where

$$g(y_t) = \theta y_t + \gamma (|y_t| - E|y_t|).$$

The above-mentioned are only a few of the ARCH type models which have been considered in the econometrics literature. For more on these models see Bollerslev, Engle & Nelson (1994) and Olsen & Associates (1995).

Let us just briefly look at one of these more exotic ARCH models, the HARCH model developed in Müller, Dacorogna, Davé, Olsen, Pictet & Weizsäcker (1996). This model is a result of an extremely thorough study of the correlation structure of the foreign exchange market. In Müller, Dacorogna, Davé, Olsen, Pictet & Weizsäcker (1996) it is observed that coarse volatility predicts fine volatility better than the other way around. Together with other findings about the behavior of volatility the HARCH model is proposed, where a HARCH(n) is defined as follows:

$$y_t = \varepsilon_t \sigma_t,$$

$$\sigma_t^2 = c_0 + \sum_{j=1}^n c_j \left(\sum_{i=1}^j y_{t-i} \right)^2,$$

$$c_0 > 0, \quad c_n > 0, \quad c_j \geq 0 \quad \text{for } j = 1, \dots, n-1.$$

Here the ε_t 's are independent identically normal or t -distributed. The HARCH models differ from any other ARCH formulation by the way past returns are taken into account. Past squared returns are entered aggregated over different interval sizes. In Müller, Dacorogna, Davé, Olsen, Pictet &

Weizsäcker (1996) a HARCH(1024) is estimated. To estimate 1024 parameters would be impossible. In order to reduce the number of parameters and get a more parsimonious representation, ranges of j where the c_j 's are the same are represented as components and thereby the number of parameters is reduced to 6. The 6 new components represent different perceptions of volatility, as seen by different types of traders. The six new intervals correspond to $j = 1, 2-4, 5-16, 17-64, 65-256$ and $257-1024$. These j 's are not directly transferable to real time since they are chosen under ϑ -time (see Dacorogna, Gauvreau, Müller, Olsen & Pictet (1994)). They can be roughly translated to a minimum of 10 minutes, 15 minutes, 1 hour, 4 hours, 1 day and $3\frac{1}{2}$ days respectively in physical time at the busiest trading periods. The 6 types of traders are in Müller, Dacorogna, Davé, Olsen, Pictet & Weizsäcker (1996) given the following tentative descriptions; 1. short-term, intra-day traders; 2. intra-day traders with only a few trades per day; 3. traders with over-night positions; 4. few traders (no explanation); 5. medium-term traders with no intra-day trading; and 6. long-term traders, investors and central banks.

4.1.2 The binomial model

Another popular type of discrete time model is the binomial model. A binomial model can simply be described by a tree where the vertices at deterministic time points split into two. Whereas all the ARCH type models are build on statistical knowledge of the financial market the binomial model is rather stylized and more of a toy model, since stock prices change randomly at random time points rather than move to one of two levels (or a countably finite number of levels, then we have a multinomial model) at deterministic time points.

Despite the simple form of the binomial model it enables one to understand fundamental features about how basic structures of markets work. See for instance Baxter & Rennie (1996) where the binomial model is presented as the first, very small, step towards a realistic model of the real financial market. An appealing feature of the binomial models is that they can be rolled up backwards for easy determination of pricing and hedging strategies. To hedge a position is to secure it against changes in the market. How this is done is not the topic of this paper but can be found in standard textbooks on mathematical finance, such as Dothan (1990) and Duffie (1992). In Figure 12 the structures of a binomial branch model, a two period binomial model and a recombinant two period binomial model are displayed.

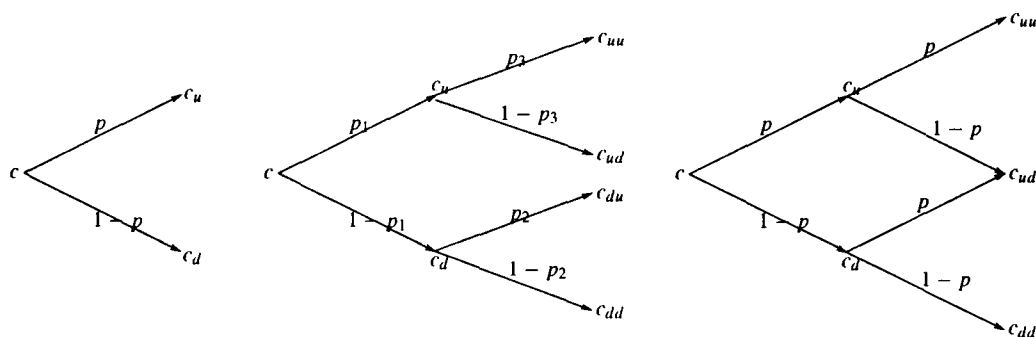


Figure 12. Examples of binomial models.

Here the p 's are the probabilities of the price going up by some prespecified percentage u and $1 - p$ is the probability of the price going down by some prespecified percentage d . So the probability of ending up with the price c_{uu} in the two period model is $p_1 p_2$, since there is one path and the

probabilities of the edges are p_1 and p_3 . Furthermore, the value of the stock after two ups is $u_1 u_3 c$. In the recombinant tree the probability of ending up with the price c_{ud} is $2p(1 - p)$ since there are two possible ways of getting c_{ud} . If the number of periods tends to infinity it can be shown that a binomial model converges to the Black–Scholes model.

4.1.3 Stochastic volatility models

Stochastic volatility models can be either discrete or continuous. But both the discrete and the continuous models share the common feature that the variance of the distribution of the log returns is unobserved and specified by some stochastic process. These types of models could also appropriately be referred to as “parameter driven” models in contrast to the ARCH models which are “observation driven” (See Cox (1981) for the definition of this terminology). What we consider to be a stochastic volatility model is a model of the type

$$y_t = \varepsilon_t f(\sigma_t).$$

Here the ε_t 's are independent identically distributed normal and independent of the positive process σ_t . Furthermore, σ_t is unobserved and f some function such that $f(\sigma_t) > 0$. The most well-known discrete time stochastic volatility model is the log-normal stochastic volatility model by Taylor (1986) which is given by

$$y_t = \varepsilon_t \exp(\sigma_t/2)$$

and

$$\sigma_t = \alpha_0 + \alpha_1 \sigma_{t-1} + \eta_t,$$

where the η_t 's are independent identically distributed normal random variables. Recently more elaborate stochastic volatility models have been proposed. Among these are the models by Barndorff-Nielsen (1997) and Barndorff-Nielsen (1998), who use normal inverse Gaussian distributions as building blocks. These types of distributions are discussed later in the sections about Lévy and diffusion models. Stochastic volatility models also exist in diffusion process formulations. Here the most famous is the one by Hull & White (1987) where

$$\begin{cases} dS_t = \mu S_t dt + \sigma_t S_t dW_t^{(1)}, \\ d\sigma_t = \beta \sigma_t (\theta - \sigma_t) dt + \gamma \sigma_t dW_t^{(2)}, \\ \text{Cov}(W_t^{(1)}, W_t^{(2)}) = d. \end{cases}$$

For examples of other specifications of stochastic volatility models of diffusion type and further discussion see; Hull & White (1988), Stein & Stein (1991), Ball (1991) and Ball & Roma (1994). For an exhaustive survey of stochastic volatility models see Ghysels, Harvey & Renault (1996). For problems related to the pricing in stochastic volatility models, see Section 5.

The problem of using the stochastic volatility models is that they are incomplete since it is not possible to observe the volatility process σ_t . Incompleteness (It can be shown that a stochastic volatility model can be completed by observing the prices of European options. See Romano & Touzi (1993).) will be discussed later in relation to Lévy models in Section 4.2.2. Lévy models can themselves be interpreted as stochastic volatility models.

4.2 Continuous Time Models

In continuous time models the models are described by stochastic differential equations in contrast to the stochastic difference equations employed in the discrete time models. As mentioned above you could roughly state that discrete time models belong to econometrics and that continuous time

models belong to mathematical finance. As correctly noted by Nelson (1990): "Relatively little work has been done on the relation between the continuous time nonlinear stochastic differential systems, used in so much of the theoretical literature, and the ARCH stochastic difference equation systems, favored by empiricists. Indeed, the two literatures have developed quite independently, with little attempt to reconcile the discrete and continuous models".

It is beyond the scope of this paper to attempt this reconciliation, but one should note that from a theoretical statistical point of view there are interesting things to be found in both worlds.

4.2.1 The Black and Scholes model and option pricing formula

The most celebrated of all models used in finance is probably the Black and Scholes model, suggested in Black & Scholes (1973). In the Black and Scholes model the stock simply follows a geometric Brownian motion, i.e.,

$$S_t = \mu S_t dt + \sigma S_t dW_t,$$

which implies that log returns $\log(S_t/S_{t-1})$ are normally distributed with mean $\mu - \frac{1}{2}\sigma^2$ and variance σ^2 . The reason for the success of the Black and Scholes is not to be found in its statistical properties, as is clear from the preceding text (e.g. no fat tails), but because in this model it is easy to give closed form expressions for the prices of derivative assets such as European options. A European call option is an asset which gives the buyer the right, but not the obligation, to purchase the stock S_t at a prenegotiated price K at time of maturity T . In mathematics the value of the European call option, at time of maturity, is given by

$$\max(0, S_T - K).$$

The price of this derivative is given by the Black and Scholes option pricing formula which is

$$C(T, K) = S_0 \Phi(d) - K e^{-rT} \Phi(d - \sigma \sqrt{T})$$

where

$$d = \frac{\log(S_0/K) + (r + \sigma^2/2)T}{\sigma \sqrt{T}}$$

and Φ denotes the standard normal cumulative distribution function.

4.2.2 Lévy type models

Before and after Black & Scholes (1973) several people have worked on models which focus on the distribution of log returns. Mandelbrot (1963) explained the non-normality of log returns by stable laws for which the variance does not exist. This was built on the idea that returns are independent and that the stable laws are the only possible limit laws for properly normed sums of independent and identically distributed random variables. As can be seen from the above discussion of stylized features such a statement is now strongly disputed, especially the independence claim.

However, several other models have been proposed, which also build on independent increments of the returns. Praetz (1972) presents a scaled t -distribution and this distribution is shown to give a good fit to weekly observations from the Sydney Stock Exchange for the period 1958–1966, a total of 462 observations. This type of model can also be viewed as a subordinate normal model, where the conditional distribution $y|\sigma^2$ is a normal distribution and σ^2 is an inverted gamma type of distribution.

In Praetz (1972) a possible explanation of the stochastic σ^2 is said to be the change in activity over the years, however, it is not justified by any statistical test. Clark (1973) suggests another

subordinated model for weekly returns. Here the process used for subordination was not only an extra stochastic element. The extra stochasticity was used as a model for trading volume. The model was chosen in this way based on examination of the relationship between the trading volume and the changes in variance. The suggested model implies normal-log-normally distributed returns, i.e. the variance is modelled as a stochastic variable which is log-normally distributed. Epps & Epps (1976) again discuss a mixture distribution for the log returns and support Clark (1973) in the finding that the log price depends on the trading volume. However, they suggest a different functional relation between log price and trading volume.

Lately, other subordinate models have been suggested. Many of the suggested models fall in the class of generalized hyperbolic Lévy models. These models are characterized by independent increments which belong to the class of generalized hyperbolic distributions. This class of distributions can be described as normal variance-mean mixtures where the mixing distribution is a generalized inverse Gaussian distribution, i.e.,

$$H(x; \lambda, \alpha, \beta, \mu, \delta) = N(\mu + \beta z, z) \int_z GIG(\lambda, \delta^2, \alpha^2 - \beta^2),$$

where H denotes the generalized hyperbolic distribution, N denotes the normal distribution and GIG is the generalized inverse Gaussian distribution. The density function of the GIG is given by,

$$gig(z; \lambda, \chi, \psi) = \frac{\left(\frac{\psi}{\chi}\right)^{\frac{1}{2}}}{2K_{\lambda}(\sqrt{\chi\psi})} z^{\lambda-1} \exp\left(\frac{1}{2}(\chi z^{-1} + \psi z)\right), \quad z > 0,$$

and the domain of variation of (λ, χ, ψ) is given by

$$\begin{array}{lll} \chi \geq 0 & \psi > 0 & \text{if } \lambda > 0, \\ \chi > 0 & \psi > 0 & \text{if } \lambda = 0, \\ \chi > 0 & \psi \geq 0 & \text{if } \lambda < 0. \end{array}$$

When $\chi = 0$ or $\psi = 0$ the density is found as the limit. For $\lambda > 0$ and $\chi = 0$ we get a gamma density. For $\lambda = -1/2$ the $GIG(\lambda, \chi, \psi)$ is the inverse Gaussian density $IG(\chi, \psi)$, which can be given a probabilistic interpretation as the first hitting time to the level $\sqrt{\chi}$ of a Brownian motion with drift $\sqrt{\psi}$ and diffusion coefficient 1.

The previously mentioned model by Praetz (1972) belongs to this class of models since the inverted gamma distribution is a GIG distribution with parameters $(\lambda, \chi, 0)$. In Madan & Seneta (1990), Madan & Milne (1991) and Madan & Chang (1996) the variance gamma (V.G.) model, which is a normal gamma Lévy process, is studied as a model for share market returns. Here the distribution of the log returns is a mixture of a normal and a gamma distribution.

In Eberlein & Keller (1995) the hyperbolic Lévy model is proposed as a model for German stock prices and it is shown to give an extremely good fit. The hyperbolic distribution is the special case of the generalized hyperbolic distributions where $\lambda = 1$. Furthermore, the model is used to price options. Option pricing in models based on Lévy models is discussed in Section 5.

In Barndorff-Nielsen (1995) and Barndorff-Nielsen (1998) the normal inverse Gaussian Lévy model is suggested, and in Rydberg (1997a) it is shown to perform well in modelling German stock prices.

4.2.3 Diffusion type models

Several different diffusion models have been formulated. One of the features often incorporated in these models is a mean-reverting effect. Support for a mean-reverting effect is found if the first lagged autocorrelation is positive and the second negative and of comparable size. A way of incorporating

this effect is by choosing the stochastic differential equation to be of the form

$$dS_t = \alpha(x - S_t)dt + \sigma(S_t)dW_t,$$

where W is a Wiener process and σ is some sufficiently well-behaved function of S_t .

The above form is used in the Cox–Ingersoll–Ross model (Cox, Ingersoll & Ross (1985)) for interest rates where $\sigma(x)$ is of the form $c\sqrt{x}$. However, the attempt to incorporate statistical features in the models are limited, one of the reasons is perhaps that the goal is to price derivative assets and this becomes rather complicated as soon as we leave the Black and Scholes framework (see Section 5).

One of the few early attempts to create a more elaborate model is by Cox (1975) and Cox & Ross (1976), who suggests the “constant elasticity of variance” diffusion model. This model is based on the empirical finding that the variance of stocks tends to decrease when the stock price increases (e.g. see Black (1976)). This effect is referred to as leverage (see the preceding section about ARCH models). The model incorporates this fact by letting the volatility be a function of the stock price in the following way

$$dS_t = \mu S_t dt + \sigma S_t^{\gamma} dW_t.$$

Again W is a Wiener process and $\gamma < 2$. For this particular diffusion Cox (1975) also developed an option pricing formula. This formula is, however, rather complicated compared to the Black and Scholes formula (see Section 4.2.1).

Some of the later diffusion type models have, in spite of the problems of pricing, tried to incorporate some of the statistical aspects of financial data. Among these are the hyperbolic diffusion model proposed for Danish stock prices in Bibby & Sørensen (1997) and the generalized hyperbolic diffusion models, proposed in Rydberg (1999). In Rydberg (1999) one of the models is the normal inverse Gaussian diffusion model which is used to model American stock prices. The generalized hyperbolic diffusion models again use the generalized hyperbolic distributions as building blocks like the generalized hyperbolic Lévy processes discussed in Section 4.2.2.

However, the diffusion models and the Lévy models are very different in nature since the diffusion processes are constructed in such a way that they have a stationary distribution that is a generalized hyperbolic distribution whilst the Lévy models are by nature highly non-stationary. Also, when it comes to pricing derivatives the models do not have much in common. Although pricing in the diffusion models is not straightforward, as can be seen from Rydberg (1997b), these models do not generate a whole continuum of equivalent martingale measures like the Lévy models do.

4.2.4 Models for high frequency financial data

In almost all financial markets prices are restricted to lie on a discrete grid. For example this is the case for stocks traded at the New York Stock Exchange which can be seen from Figure 2. In 1995 the tick size was 1/8th of a dollar, in 1997 it changed to 1/16th of a dollar. O’Hara (1995, Ch. 1) provides an introduction to markets and market making on the NYSE.

The discreteness of the prices, which is enhanced at the intra-daily level, has been the basis for a very recent line of research which is concerned with the timing of the actual trades and the movements of the price at trades. When studying the stock prices at this ultra high frequency it becomes apparent that it is important to model the duration between trades as well as price moves.

Engle & Russell (1998) developed the Autoregressive Conditional Duration (ACD) model, in which the arrivals of transactions are described by a counting process and the duration between events follows a process of the following type:

$$\psi_i = \omega + \alpha x_{i-1} + \beta \psi_{i-1} \quad \text{for } \alpha, \beta \geq 0, \omega > 0, \forall i, i = 1, \dots, N.$$

Here x_i denotes the duration between events at time t_{i-1} and t_i , i.e. $x_i = t_i - t_{i-1}$. So in some sense the process is discrete and the intensity with which the jumps occurs can only change at the discrete time points. This model was based on the assumption that the density of the standardized durations x_i/ψ_i are independent and identically distributed with density

$$g(x_i/\psi_i|x_{i-1}, \dots, x_1; \theta) = g(x_i/\psi_i; \theta)$$

and

$$E(x_i|x_{i-1}, \dots, x_1; \theta) = \psi_i = \psi(x_{i-1}, \dots, x_{i-p}, \psi_{i-1}, \dots, \psi_{i-p}; \theta).$$

This type of process has been used in the insurance literature and was previously discussed in Wold (1948) and Wold (1949).

This approach is quite different from the one found in a recent paper by Rydberg & Shephard (1998b) who proposed a continuous time framework for the evolution of the prices and times. In order to model the fact that the trades occur at random time points a compound process was suggested. Let $p(t)$ denote the price of the asset at time t , then allow

$$p(t) = p(0) + \sum_{k=1}^{N(t)} Z_k, \quad (2)$$

where $N(t)$ is the number of trades recorded up until time t and Z_k is the price movement or change associated with the k -th trade. In Rydberg & Shephard (1998b) and Rydberg & Shephard (1999) $N(t)$ is modeled as a counting process with new arrivals being generated by a Cox process. In contrast to Engle & Russell (1998) they allowed the intensity of the Cox process to be a continuous process.

The duration models constitute only half of the description of the price formation. To get a description for the price process as such the price changes at a trade should also be modelled. This is the topic of Russell & Engle (1998) and Rydberg & Shephard (1998a).

Russell & Engle (1998) suggest modelling the price movements using a conditional multinomial distribution. Their paper can be viewed as a time series extension of a multivariate probit analysis of transaction data proposed by Hausman, Lo & MacKinlay (1992). But where Hausman, Lo & MacKinlay (1992) use the probit function as their link, Russell & Engle (1998) prefer to work with logistic functions. In order to use the multinomial model it is necessary to assume that the price movements can only exist on a fixed number of states in this case it is assumed that they live on $-2, -1, 0, 1, 2$ which is the case for 98.6% of all the trades in the data set considered in Russell & Engle (1998). Their autoregressive conditional multinomial models are shown to give a very good description of the price changes and it is easy to extend to include many types of explanatory variables.

The approach in Rydberg & Shephard (1998a) is somewhat different in that they use a decomposition technique which enables them to specifically model the activity, the direction and the size of a given trade. If Z_k denotes the k -th trade then the decomposition can be written as

$$Z_k = A_k D_k S_k.$$

Where A_k takes on only two values: 0, 1. If $A_k = 0$ then $D_k = S_k = 0$. Otherwise, when $A_k = 1$, D_k and S_k live on the structure:

$$D_k = -1, 1 \quad \text{and} \quad S_k = 1, 2, \dots$$

Thus if A_k is zero then Z_k must be zero. This means the price does not move or, in other words, is *In-Active*. If $A_k = 1$ then there are *Active* price movements. The non-zero price movement must be $Z_k = D_k S_k$. Likewise, if we assume $A_k = 1$, then D_k controls the *Direction* of the price move. If $D_k = 1$ the price moves upwards, else it moves downwards. Finally, S_k controls the *Size* of price

movements.

Hasbrouck (1996) introduced a dynamic model for the evolution of quotes of bid and ask prices. Here continuous models are suggested as models for, respectively, the theoretical efficient price in the market, the ask and bid costs. Hasbrouck proposes a model where the bid price is the underlying theoretical efficient price minus the bid costs rounded down to the closest tick whilst the ask price is the underlying theoretical efficient price plus the ask costs rounded up to the closest tick. Related papers include Bollerslev & Melvin (1994) and Harris (1994). Manrique & Shephard (1997) have studied the implied econometrics of this type of model.

The Hasbrouck (1996) bid/ask model is not immediately applicable to transaction data, but the principle of using a continuous time model which is then truncated in some way is potentially useful if combined (perhaps) with the Hausman, Lo & MacKinlay (1992) static model of clustering.

There are several market microstructure questions that can be answered by the models described above. One of the goals in modelling the intra-daily data is to gain insight into how different types of trades affect the long term price levels (here *long term* should be understood in terms of number of trades). Forecasts can be calculated on the basis of both the models developed in Russell & Engle (1998) and Rydberg & Shephard (1998a). Furthermore, given the structures of the models it is easy to include explanatory variables which make it possible to test for the effects of, e.g. volume and seasonal effects. Questions about how activity in the market affects the price can also be answered.

5 Pricing and Hedging

The Black and Scholes formula for the pricing of options, where the underlying asset is modelled by the geometric Brownian motion, gives a closed form solution to the pricing problem. This is only possible because of the simple structure of the geometric Brownian motion. More generally fair prices of a derivative assets such as a European or American call option, may be calculated as the mean value of the discounted payoff function of the derivative, under the unique equivalent martingale measure provided that this measure does in fact exist. Existence and uniqueness hold in particular for the Black–Scholes model as was seen previously.

In other, perhaps more realistic, diffusion models such as the hyperbolic diffusion model proposed by Bibby & Sørensen (1997) or the normal inverse Gaussian diffusion model proposed by Rydberg (1999), it may be difficult or perhaps even impossible to show existence by standard conditions such as the sufficient, but not necessary, conditions proposed by Novikov and Kazamaki. The Novikov and Kazamaki conditions are discussed, for instance in Revuz & Yor (1994, Ch. VIII). However, even if this unique equivalent martingale measure does in fact exist, the price will in most instances have to be calculated based on a numerical procedure. We will briefly go through some of the mathematics, in the continuous SDE case, as this might be helpful for understanding the problem.

Let σ and b be Borel-measurable functions $[0, T] \times \mathbf{I} \mapsto \mathbf{I}$, $\mathbf{I} = (l, r) \subseteq \mathbf{R}$. Assume that X is a solution of the SDE given by

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x,$$

where W_t is a Wiener process defined on a probability space (Ω, \mathcal{F}, P) equipped with the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$. For simplicity let \mathcal{F}_t be the natural filtration generated by X_t .

Let (Y, W^*) (W^* is a P^* -Wiener process) be a solution to the SDE

$$dY_t = \sigma(t, Y_t)dW_t^*, \quad Y_0 = y.$$

Then the Girsanov theorem implies that the two laws P and P^* are equivalent on $[0, T]$ if the process given by

$$N_t = \exp \left(-\frac{1}{2} \int_0^t \frac{b^2(s, X_s)}{\sigma^2(s, X_s)} ds - \int_0^t \frac{b(s, X_s)}{\sigma(s, X_s)} dW_s \right),$$

is a martingale.

The financial market described by the discounted price process X which is a solution to a SDE of the above form, is complete (see Hansen (1995)) if: the self financing trading strategies (A trading strategy is self financing if no other capital is supplied than the initial.) are well defined and the contingent claim H , $H \in \mathcal{F}_T$, $H > 0$, such that $E^P [N_T H] < \infty$, is attainable at the fair price given by

$$\pi(H) = E^P [N_T H].$$

If the unique equivalent martingale measure, P^* , exists then

$$\pi(H) = E^P [N_T H] = E^{P^*} [H],$$

is the arbitrage free price.

The situation for diffusion models is simpler than for Lévy models since diffusion models only allow for one candidate for the riskless equivalent martingale measure, while some Lévy models might allow for a continuum of measures. Due to the discontinuities of the Lévy processes, excluding the Wiener process, a unique equivalent martingale measure does not exist, there may in fact, be a whole continuum of such measures if the Lévy measure is continuous.

This is the case for generalized hyperbolic Lévy processes in general and in particular the hyperbolic Lévy process. In Eberlein & Keller (1995) this problem is overcome by using exponential tilting, also called the Esscher transform, to distinguish a particular measure. However, it is not clear whether this approach is the right one, i.e. finds the correct measure, seen from an economists point of view. Here it should be noted that finding the right probability measure is important since it can be shown, see Eberlein & Jacod (1997), that the arbitrage free (A price is arbitrage free if it is based on an arbitrage free market. That a market is arbitrage free means that there is no way of making riskless profits.) option prices calculated by use of the riskless equivalent martingale measures (The existence of at least one riskless equivalent martingale measures implies no-arbitrage. The other way around is not as straight forward.) fill the whole interval $((S_0 - e^{-rT} K)^+, S_0)$.

In Rydberg (1997a) a way to overcome the problem of finding unique prices is also put forward. The method is based on a discretized version of the continuous Lévy measure, which delimits the number of possible probability measures from infinity to a finite number n , such that hedging and pricing is possible if $n + 1$ assets are traded.

Bladt & Rydberg (1998) take an alternative approach. Here the pricing is carried out under the physical measure, whereby one unique measure is chosen. Perhaps due to the incompleteness of the Lévy models they are not the most favored types of models when it comes to pricing options even though they might give a good statistical description of the price changes (quite often they would lack correlation structure).

Alternative approaches on how to price options, but based on stochastic volatility models, can be found in Hull & White (1987), Heston (1993) and Chernov & Ghysels (1998). From a statistician's point of view these works are much more interesting as they are based on models which are developed to describe the financial data and not only models where pricing is possible.

Option pricing based on ARCH-type models is complicated by their discrete nature, which causes them to be incomplete. However, based on, for instance, equilibrium arguments it is possible to overcome the incompleteness and deduce a pricing formula. This has been studied in e.g. Duan (1995) and Kallsen & Taqqu (1998).

The problem of pricing options is not the only issue. Maybe even more important is the hedging problem. To hedge a position means to secure it against the risk of market movements. In simple models such as the Black and Scholes model this can be solved exactly and the exact replicating (To replicate a payoff is to generate the same payoff by holding a portfolio of bonds and stocks. For the interested reader we refer to Baxter & Rennie (1996).) strategy can be calculated. Again, as soon as

we move away from simple models the problem becomes more difficult and is somewhat unsolved.

For both the pricing and the hedging problem it seems that new thinking is needed. Even though we have mentioned some attempts they are by no means exhaustive. For a very mathematical approach to the problem Delbaen & Schachermayer (1994) is a good source, while Chernov & Ghysels (1998) provide a statistically oriented approach and include many references to related work.

6 Conclusion

This paper has dealt with problems in the area of interplay between statistics and financial econometrics. It has described most of the features that characterize financial data. The aim has been to give an idea why building good models is a difficult project and takes great statistical skills. Hopefully it has also transpired why it is important to have good statistical models. Just to recapture the main reasons. Misspecified models may lead to mispricing of derivatives, hedge strategies will not be adequate in covering all the portfolio risk, and other economic questions given wrong answers. Furthermore, being able to hedge and assess risk is not only of interest to a small elite of financial speculators: it is of interest indirectly to almost everyone via pension schemes and mortgages. See Karatzas & Shreve (1998) for other examples.

Some of the complications in applying the more sophisticated models are outlined. They include problems related to incompleteness and arbitrage. In order to resolve such problems an interaction between the fields of econometrics and mathematical finance will be beneficial.

Acknowledgements

I am grateful to Carsten Tanggaard at the Aarhus School of Business for providing the NYSE data and to Marc Hoffmann for making the French translation of the summary. Thanks to my colleagues Susannah Morris and Neil Shephard for discussions and comments. Also thanks to the associate editor and anonymous referees for helpful corrections and suggestions. Funding via EU grant on "Econometric inference using simulation techniques" is gratefully acknowledged.

The paper was specially prepared for the 51st Session of the International Statistical Institute held in Istanbul August '97.

References

- Andersen, T.G. & Bollerslev, T. (1998). Towards a unified framework for high and low frequency return volatility modeling. *Statistica Neerlandica*, **52**, 273–302.
- Bachelier, L. (1900). 'Théorie de la Speculation'. *Annales scientifiques de l'école normale supérieure* 3d ser. 17:21–88, Translation in *The Random Character of Stock Market Prices*, Ed. P. Cootner, (1964) pp. 17–79. Cambridge, Mass.: MIT Press.
- Ball, C.A. (1991). A Review of Stochastic Volatility Models with Application to Option Pricing. *Financial Markets: Institutions & Instruments*, **2**, 55–71.
- Ball, C.A. & Roma, A. (1994). Stochastic Volatility Option Pricing. *Journal of Financial and Quantitative Analysis*, **29**, 589–607.
- Barndorff-Nielsen, O.E. (1995). *Normal\Inverse Gaussian Processes and the Modelling of Stock Returns*. Research Report 300, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.
- Barndorff-Nielsen, O.E. (1997). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scand. J. Statist.*, **24**, 1–13.
- Barndorff-Nielsen, O.E. (1998). Processes of normal inverse Gaussian type. *Finance Stochast.*, **2**, 41–68.
- Baxter, M.W. & Rennie, A.J.O. (1996). *Financial Calculus—An introduction to derivative pricing*. Cambridge: Cambridge University Press.
- Bibby, B.M. & Sørensen, M. (1995). Martingale Estimating Functions for Discretely Observed Diffusion Processes. *Bernoulli*, **1**, 17–39.
- Bibby, B.M. & Sørensen, M. (1997). A Hyperbolic Diffusion Model for Stock Prices. *Finance and Stochast.*, **1**, 25–41.
- Black, F. (1976). Studies of Stock Price Volatility Changes. In *Proceedings of the 1976 meetings of the American Statistical Association, Business and Economic Session*.
- Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**, 637–654.

- Bladt, M. & Rydberg, T.H. (1998). An actuarial approach to option pricing under the physical measure and without market assumptions. *Insurance: Mathematics and Economics*, **22**, 65–73.
- Bollerslev, T. (1986). Generalised Autoregressive Conditional Heteroskedasticity. *J. Econometrics*, **51**, 307–327.
- Bollerslev, T., Chou, R.Y. & Kroner, K.F. (1992). ARCH Modeling in Finance. *J. Econometrics*, **52**, 5–59.
- Bollerslev, T., Engle, R.F. & Nelson, D.B. (1994). ARCH Models. In *Handbook of Econometrics*, Eds. R.F. Engle and D.L. McFadden, Volume IV, Chapter 49. Amsterdam: Elsevier Science B.V.
- Bollerslev, T. & Melvin, M. (1994). Bid-ask spreads in the foreign exchange market: an empirical analysis. *Journal of International Economics*, **36**, 355–372.
- Campbell, J.Y., Lo, A.W. & MacKinlay, A.C. (1997). *The Econometrics of Financial Markets*. Princeton, New Jersey: Princeton University Press.
- Chernov, M. & Ghysels, E. (1998). What data should be used to price options. Unpublished working paper, downloadable from http://papers.ssrn.com/sol3/paper.taf?ABSTRACT_ID=120028.
- Clark, P.K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, **41**, 135–155.
- Cox, D.R. (1981). Statistical Analysis of Time Series: Some Recent Developments. *Scand. J. Statist.*, **8**, 93–115.
- Cox, J.C. (1975). *Notes on Option Pricing I: Constant Elasticity of Variance Diffusions*. Working paper, Graduate School of Business, Stanford University.
- Cox, J.C., Ingersoll, J.E. & Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, **53**, 385–407.
- Cox, J.C. & Ross, S.A. (1976). The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics*, **3**, 145–166.
- Dacorogna, M.M., Gauvreau, C.L., Müller, U.A., Olsen, R.B. & Pictet, O.V. (1994). Changing Time Scale for Short-term Forecasting in Financial Markets. Preprint MMD. 1993–08–09, Olsen & Associates Ltd., Research Institute for Applied Economics, <http://www.olsen.ch/>.
- Davis, R.A. & Mikosch, T. (2000). Limit theory for the sample acf of stationary processes with heavy tails with applications to ARCH. *Ann. Statist.* To appear.
- Delbaen, F. & Schachermayer, W. (1994). The fundamental asset pricing theorem. *Mathematische Annalen*, **300**, 463–520.
- Ding, Z., Granger, C.W.J. & Engle, R.F. (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, **1**, 83–106.
- Dothan, M.U. (1990). *Prices in Financial Markets*. New York; Oxford: Oxford University Press.
- Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance*, **5**, 13–32.
- Duffie, D. (1992). *Dynamic Asset Pricing Theory*. Princeton, New Jersey: Princeton University Press.
- Eberlein, E. & Jacod, J. (1997). On the Range of Option Prices. *Finance and Stochastics*, **1**, 131–140.
- Eberlein, E. & Keller, U. (1995). Hyperbolic Distributions in Finance. *Bernoulli*, **1**, 281–299.
- Elerian, O., Chib, S. & Shephard, N. (1998). Likelihood inference for discretely observed non-linear diffusions. Unpublished paper: Nuffield College, Oxford.
- Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation. *Econometrica*, **50**, 987–1007.
- Engle, R.F. & Russell, J.R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, **66**, 1127–1162.
- Epps, T.W. & Epps, M.L. (1976). The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypotheses. *Econometrica*, **44**, 305–321.
- Gallant, A.R. & Long, J.R. (1997). Estimating stochastic differential equations efficiently by minimum chi-square. *Biometrika*, **84**, 125–141.
- Gallant, A.R. & Tauchen, G. (1996). Which moments to match. *Econometric Theory*, **12**, 657–681.
- Ghysels, E., Gouriéroux, C. & Jasiak, J. (1998). High frequency financial time series data: Some stylized facts and models of stochastic volatility. In *Nonlinear Modelling of High Frequency Financial Time Series*, Financial Econometrics and Quantitative Analysis, Eds. C. Dunis and B. Zhou, pp. 127–159. Chichester: Wiley & Sons.
- Ghysels, E., Harvey, A.C. & Renault, E. (1996). Stochastic volatility. In *Statistical Methods in Finance*, Eds. C.R. Rao and G.S. Maddala, pp. 119–191. Amsterdam: North-Holland.
- Guillaume, D.M., Dacorogna, M.M., Davé, R.D., Müller, U.A., Olsen, R.B. & Pictet, O.V. (1997). From the Bird's Eye to the Microscope—a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics*, **1**, 95–129.
- Hansen, A.T. (1995). Complete Market Pricing in the Wiener-Filtration: A Classification of Regularities by Application of the local non-Arbitrage Condition. Publication no: 10/95, Department of Operations Research, Institute of Mathematics, University of Aarhus.
- Harris, L.E. (1994). Minimum price variation, discrete bid-ask spreads and quotation sizes. *Review of Financial Studies*, **7**, 149–178.
- Hasbrouck, J. (1996). The dynamics of discrete bid and ask quotes. Unpublished paper: Stern Business School, New York University.
- Hausman, J., Lo, A.W. & MacKinlay, A.C. (1992). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics*, **31**, 319–330.
- Heston, S.L. (1993). A closed-form solution for options with stochastic volatility, with applications to bond and currency options. *Rev. Financial Studies*, **6**, 327–343.
- Hull, J. & White, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finance*, **42**, 281–300.
- Hull, J. & White, A. (1988). An Analysis of the Bias in Option Prices Caused by Stochastic Volatility. *Advances in Futures and Options Pricing Research*, **3**, 29–61.

- Kallsen, J. & Taqqu, M.S. (1998). Option pricing in ARCH-type models. *Mathematical Finance*, **8**, 13–26.
- Karatzas, I. & Shreve, S.E. (1998). *Methods of Mathematical Finance*, Volume 39 of *Applications of Mathematics: Stochastic Modelling and Applied Probability*. Berlin; Heidelberg; New York: Springer-Verlag.
- Madan, D.B. & Chang, E.C. (1996). Volatility Smiles, Skewness Premia and Risk Metrics: Applications of a four Parameter Closed form Generalization of Geometric Brownian Motion to the Pricing of Options. Presented at a Conference on Mathematical Finance—I, University of Aarhus, Denmark.
- Madan, D.B. & Milne, F. (1991). Option Pricing with V.G. Martingale Components. *Mathematical Finance*, **1**(4), 39–55.
- Madan, D.B. & Seneta, E. (1990). The Variance Gamma (V.G.) Model for Share Market Returns. *The Journal of Business*, **63**, 511–524.
- Mandelbrot, B.B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, **36**, 394–419.
- Manrique, A. & Shephard, N. (1997). Likelihood analysis of a discrete bid/ask price model for a common stock quoted on the NYSE. Unpublished paper: Nuffield College, Oxford.
- Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V. & Weizsäcker, J.E.V. (1996). Volatilities of Different Time Resolutions—Analyzing the Dynamics of Market Components. Preprint UAM.1995–01–12, Olsen & Associates Ltd., Research Institute for Applied Economics, <http://www.olsen.ch/>.
- Müller, U.A., Dacorogna, M.M. & Pictet, O.V. (1996). Heavy tails in high-frequency financial data. Preprint August 9, 1996, Olsen & Associates Ltd., Research Institute for Applied Economics, <http://www.olsen.ch/>.
- Nelson, D.B. (1990). ARCH Models as Diffusion Approximations. *Journal of Econometrics*, **45**, 7–38.
- Nelson, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347–370.
- O'Hara, M. (1995). *Market Microstructure Theory*. Oxford: Blackwell Publishers.
- Olsen, R.B. & Associates (Eds.) (1995). *Proceedings of the conference on High Frequency Data in Finance, HFDF-I*. Zürich, Switzerland, March 29–31, 1995.
- Pedersen, A.R. (1995). A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations. *Scand. J. Statist.*, **22**, 55–71.
- Praetz, P.D. (1972). The Distribution of Share Prices. *The Journal of Business*, **45**, 49–55.
- Revuz, D. & Yor, M. (1994). *Continuous Martingales and Brownian Motion* (Second ed.), Volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Berlin: Springer-Verlag.
- Romano, M. & Touzi, N. (1993). Contingent claims and market completeness in a stochastic volatility model. Discussion paper, CREST, Paris.
- Russell, J.R. & Engle, R.F. (1998). Econometric analysis of discrete-valued, irregularly-spaced financial transactions data using a new autoregressive conditional multinomial model. Unpublished paper: Graduate School of Business, University of Chicago. Presented at the Second international conference on high frequency data in finance, Zürich, Switzerland, April.
- Rydgberg, T.H. (1997a). The normal inverse Gaussian Lévy process: simulation and approximation. *Communications in Statistics: Stochastic Models*, **13**, 887–910.
- Rydgberg, T.H. (1997b). A Note on the Existence of Unique Equivalent Martingale Measures in a Markovian Setting. *Finance and Stochastics*, **1**, 251–257.
- Rydgberg, T.H. (1999). Generalized hyperbolic diffusions processes with applications in finance. *Mathematical Finance*, **9**, 183–201.
- Rydgberg, T.H. & Shephard, N. (1998a). Dynamics of trade-by-trade price movements: decomposition and models. Working paper, Nuffield College, Oxford. Presented at Workshop on Econometrics and Finance, Isaac Newton Institute, Cambridge University, October 1998.
- Rydgberg, T.H. & Shephard, N. (1998b). A modelling framework for the prices and times of trades made on the NYSE. To appear in *Nonlinear and nonstationary signal processing*, Eds. W.J. Fitzgerald, R.L. Smith, A.T. Walden and P. C. Young. Cambridge University Press, 2000.
- Rydgberg, T.H. & Shephard, N. (1999). Bin models for trade-by-trade data: Modelling the number of trades in a fixed interval of time. Working paper, Nuffield College, Oxford.
- Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility. In *Time Series Models—in econometrics, finance and other fields*, Eds. D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen, pp. 1–67. London: Chapman and Hall.
- Stein, E.M. & Stein, J.C. (1991). Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *The Review of Financial Studies*, **4**, 727–752.
- Taylor, S.J. (1986). *Modelling Financial Time Series*. Chichester: John Wiley & Sons.
- Taylor, S.J. (1994). Modeling Stochastic Volatility: A Review and Comparative Study. *Mathematical Finance*, **4**, 183–204.
- Wold, H. (1948). On stationary point processes and Markov chains. *Skand. Aktuar*, **31**, 229–240.
- Wold, H. (1949). Sur les processus stationnaires ponctuels. *Colloques Internationaux du CNRS*, **13**, 75–86.

Résumé

Le but de cet exposé et d'essayer de comprendre pourquoi les données financières sont intéressantes du point de vue de la statistique. J'essaierai en particulier de décrire ce que l'on cherche à modéliser et je présenterai certains des modèles les plus populaires ainsi que des modèles nouveaux. Pour faciliter la discussion, on analysera plusieurs ensembles de données. C'est pourquoi le nombre de graphiques sera important et le nombre de formules modéré. En particulier, on discutera en détail les

données contenant tous les échanges américains sur les stocks IBM en 1995 au NYSE.

[Received September 1997, accepted August 1999]