AN OVERVIEW OF PROBABILISTIC AND TIME SERIES MODELS IN FINANCE

Alejandro Balbás
Dept. of Business Administration, Universidad Carlos III de Madrid

Rosario Romera
Dept. of Statistics and Econometrics, Universidad Carlos III de Madrid

Esther Ruiz
Dept. of Statistics and Econometrics, Universidad Carlos III de Madrid

Abstract
In this paper, we partially review probabilistic and time series models in finance. Both discrete and continuous-time models are described. The characterization of the No-Arbitrage paradigm is extensively studied in several financial market contexts. As the probabilistic models become more and more complex to be realistic, the Econometrics needed to estimate them are more difficult. Consequently, there is still much research to be done on the link between probabilistic and time series models.

Keywords:
Asset Pricing, CAPM, Choquet integral, Diffusion process, GARCH, Stochastic Volatility, Term Structure, Value at Risk.

2.1 Introduction
Uncertainty plays a central role in financial theory and its empirical implementation. The objective of this paper is to review the connection between the theory and the empirical analysis in the area of Finance. It is obvious that the scope of the subject is too wide and, consequently, we will not be able to cover all contributions in the area. Therefore, in the framework of probabilistic models, we focus on those pricing models reflecting the absence of arbitrage and free-lunch. The problem of valuation and hedging of contingent claims (risks) presents important difficulties when markets imperfections are met. The characterization of No-Arbitrage (NA) is extensively studied in section 2. Pricing of contingent claims when markets are subject to portfolio constraints, transactions costs and taxes as well as new results for nonlinear pricing along with
a universal framework for pricing financial and insurance risks are reviewed in this section.

Section 3 reviews the main time series models devoted to the analysis of financial returns. We start describing models for the conditional mean usually fitted to test whether financial prices are predictable. In this sense, it is generally accepted that asset returns are close to be martingale difference processes. However, they are not independent because of the often observed dependence of some transformations related with second moments. Consequently, we then describe models to represent the dynamic evolution of conditional variances and covariances of high frequency returns. Finally, section 3 reviews the models recently proposed to represent the main empirical properties of ultra high frequency (intra-daily) returns.

In section 4, we focus on the link between probabilistic models and Financial Econometrics. We show that the estimation of realistic financial models for asset prices are, in general, difficult and much research remains to be done in this area. In particular, in this section, we describe the empirical implementation of the CAPM as well as the estimation procedures of the term structure, the VaR and continuous time diffusions.

The paper finishes in section 5 with a summary of the main conclusions.

2.2 Probabilistic models for finance

A classical problem in mathematical finance is the pricing of financial assets. The usual solution of this problem involves the so-called Fundamental Theorem of Asset Pricing. This result ensures that the assumption of NA is essentially equivalent to the existence of an equivalent martingale measure, in a perfect financial market. The NA assumption amounts to saying that there is no plan yielding some profit without a countervailing threat of loss. It prevents the existence of zero cost portfolios with positive return. The problem of fair pricing of financial assets is then reduced to taking their expected values with respect to equivalent martingale measures. Initial results on the Fundamental Theorem of Asset Pricing hold in the case of finite number of assets and a finite discrete time models; see Harrison and Kreps (1979) and Harrison and Pliska (1981).

Various generalizations are now available in the literature. For discrete infinite or continuous time, the notion of “no free lunch” or “no free lunch with bounded (vanishing) risk” is needed, which is a slightly stronger version of the non-arbitrage condition; see, for example, Dalang et al. (1989), Back and Pliska (1991) and Schachermayer (1992). In these generalizations, securities markets are assumed to be frictionless, i.e. without considering transaction costs. For discrete infinite case see Schachermayer (1994). For continuous time models see Delbaen (1992) or Delbaen and Schachermayer (1994, 1998);
2.2.1 The Fundamental Theorem of Asset Pricing

The mathematical translation of this concept uses martingale theory and stochastic analysis. Under the assumption that the $\mathbb{R}^d$-valued price process $\{S_t\}_{t \in \mathbb{R}^+}$ reflect economically meaningful ideas and does not generate arbitrage profits, the Fundamental Theorem of Asset Pricing allows the probability measure $P$ on the underlying probability space $(\Omega, \mathcal{F}, P)$ to be replaced by an equivalent measure $Q$ such that $\{S_t\}_{t \in \mathbb{R}^+}$ becomes a (local) martingale under the new measure. The information structure is given by a filtration $(\mathcal{F}_t)_{t \in T}$. Following Delbaen and Schachermayer (1994, 1998), there should be no trading strategy $H$ for the process $S$, such that the final payoff described by the stochastic integral $(H.S)_\infty$, is a nonnegative function, strictly positive with positive probability.

A buy-and-hold strategy can be described, from the mathematical point of view, as an integrand of the form $H = f.1_{(T_1,T_2)}$, where $T_1 \leq T_2$ are stopping times and $f$ is $\mathcal{F}_T$-measurable. The interpretation of this integrands is clear: when time $T_1(w)$ comes up, buy $f(w)$ units of the financial asset, keep them until time $T_2(w)$ and sell. Stopping times are interpreted as signals coming from available information and this is one reason why, in mathematical finance, the filtration and further concepts such as predictable processes, are so relevant. Even if the process $S$ is not a semi-martingale, the stochastic integral $(H.S)$ for a buy-and-hold strategy $H$ can be defined as the process $(H.S)_t = (S_{\min(t,T_2)} - S_{\min(t,T_1)})$. A linear combination of buy-and-hold strategies is called a simple integrand. In the general case simple integrands are not sufficient to characterize these processes that admit an equivalent martingale measure. On the other hand the use of general integrands leads the problem of the existence of $(H.S)$. The so called admissible integrands avoid all of these pathologies.

Formally, if $S$ denotes an $\mathbb{R}^d$-valued semi-martingale, defined on the filtered probability space $(\Omega, \{\mathcal{F}_t\}_{t \in \mathbb{R}^+}, P)$, an $\mathbb{R}^d$-valued predictable process $H$ is called $\alpha$-admissible if it is $S$-integrable, if $H_0 = 0$, if the stochastic integral satisfies $H.S \geq -\alpha$ and if the $\lim_{t \to \infty}(H.S)_t$ exists a.s. If $H$ is admissible for some $\alpha$, then is simply call admissible.

In order to characterize mathematically the NA and the No Free Lunch (NFL) properties, we need to consider the following vector spaces. Let us denote by $L^0$ the vector space of all real-valued measurable functions defined on $\Omega$. Endowed with the topology of convergence in probability, this space becomes a Fréchet space (i.e. a complete and metrisable vector space). $L^\infty$ denotes the subspace of $L^0$ of all bounded functions. It is remarkable that the two
spaces $L^0$ and $L^\infty$ are, among the $L^p$ spaces, the only two spaces that remain the same when the original probability measure is replaced by an equivalent one. Let us introduce the following sets:

$$\Phi = \{(H.S)_\infty / H \text{ is admissible}\},$$

$$\Phi_\alpha = \{(H.S)_\alpha / H \text{ is } \alpha\text{-admissible}\},$$

$$\Upsilon_0 = \Phi - L^0_+, \quad \Upsilon = \Phi_0 \cap L^\infty.$$

In all papers dealing with the Fundamental Theorem of Asset Pricing (with simple integrands), the assumption of NA or NFL essentially amounts to saying that the set $\Phi$ does not contain any non-negative random variable except the null one.

Formally, we say that the process $S$ satisfies the NA property if:

$$(\Phi \cap L^0_+) = \{0\}$$

which is equivalent to the expression

$$(\Upsilon \cap L^\infty_+) = \{0\}.$$

The process $S$ satisfies the NFL property if

$$(\Upsilon \cap L^\infty) = \{0\},$$

where the bar denotes closure in the norm topology of $L^\infty$.

The NFL is an old expression used in the early days of the finance literature. The NA postulates that the set of random variables which can be achieved by a zero cost portfolio does not include any positive random variable. The NFL condition, postulates the same on the topological closure of the previous set. The following technical definition is due to Kreps (1981). Let $S$ be a bounded process and let us denote by $\Phi^*$ the set of all outcomes with respect to bounded simple integrands. $\Upsilon^*$ is defined in the same way $\Upsilon^* = (\Phi^* - L^0_+) \cap L^\infty$.

Then, an adapted process $S$ satisfies the NFL property, as above, if the corresponding set of outcomes does not contain any non-negative random variable except the null, $\Upsilon^* \cap L^\infty_+ = \{0\}$, where the tilde denotes weak closure. Dealing with the weak closure it may happen that an element of this set can only be obtained by an unbounded generalized sequence. Unfortunately the economic interpretation of this unbounded objects is unclear. However requirements of NA and NFL in expressions (1) and (2) are very strong. We assume that $S$ is a semi-martingale and there is an equivalent martingale measure for the process $S$. On the other hand we need a definition for the set of outcomes with respect to general admissible integrands. The following theorem from Delbaen and Schachermayer (1998), characterizes the NFL concept through a boundedness property in $L^0$. 


THEOREM 1 The process $S$ satisfies the property NFL (2) if and only if it satisfies

1. the NA property (1) and

2. $\Phi_1$ is bounded in the space $L^0$

They remark that the boundedness of the set $\Phi_1$ has the following economic interpretation: for outcomes that have a maximal loss bounded by 1, the profit is bounded in probability, this means that the probability of making a big profit can be estimated from above, uniformly over all such outcomes.

For further characterization of the NFL property and related results for locally bounded semi-martingales $S$, see Delbaen and Schachermayer (1994, 1998).

A recent projective system approach to the martingale characterization of the absence of arbitrage is provided by Balbás et al. (2002). The equivalence between the absence of arbitrage and the existence of an equivalent martingale measure fails when an infinite number of trading dates is considered. Thus, enlarging the set of states of nature and the probability measure through a projective system of perfect measure space, the authors characterize the absence of arbitrage when the time set is countable.

The martingale characterization can be extended in the context of imperfect financial models, mainly financial models with proportional transaction costs, short sale constraints, convex cone constraints, etc.

We can observe three main lines of research generalizing these initial results. The first one applies in the context of imperfect financial markets for a model with transaction costs. The second line of research expands the restricted feasible portfolio case, usually cone constraints. The third research direction and the most recent one is based on the assumption that the price is non-linear with respect to the portfolio. Then the subaditivity property is needed and the Choquet integral is a powerful tool to be used in this context. The asset pricing problem is then solved as a Choquet integral of the future returns with respect to a new capacity introduced by Chateauneuf et al. (1994, 1996).

Currently there is a pressing need for a universal framework for the determination of the fair value of financial and insurance risks. In the financial services industry, this pressing need is evidenced by the recent Basel Accords on regulatory risk management that require fair value, analogous to market prices, to be applied to all assets or losses, whether traded or not. More recently Wang (2000, 2001) presents a universal framework for pricing financial and insurance risks.
2.2.2 Asset Pricing in Imperfect Financial Markets

In the classical setting, the financial market is modeled in a “frictionless” way which is a clear idealization of the real world. Therefore models with transaction costs have been increasingly studied in the literature; see Davis and Norman (1990) or Striker (1990). Jouini and Kallal (1995a) characterize the assumption of NFL in a model with transaction costs and give fair pricing intervals for contingent claims in such a model. As for other imperfection, Jouini and Kallal (1995b, 1999) consider the case of short sale constraints or shortselling costs with possibly different rates for borrowing and lending rates. The problem of hedging contingent claims, in continuous time, is study by Cvitanic and Karatzas (1996). They propose a diffusion model (with one bond and one risky asset) with proportional transaction costs, and give a dual formulation for the so-called super-replication price of a contingent claim (i.e. the minimum initial wealth needed to hedge the contingent claim, or in other words, to obtain, through the investment opportunities available on the market, at least the contingent claim). Delbaen et al. (1998) generalize this result to the multivariate case, in discrete as well as in continuous time, and with a semi-martingale price process. In these models too, typically there is a “bond” which serves as numeraire asset. The usual assumption is that, at final date $T$, all the positions in the other traded assets are liquidated, i.e., converted into units of the bond.

More recently, Jouini and Napp (2002) generalize existing results in the following ways: first, they do not assume that there exists a numéraire available to investors and allowing them to transfer money from one date to another; this enables to consider any type of friction on the numéraire-like no borrowing, different borrowing and lending rates, bonds with default risk, etc. These setting also take into account the fact that all investors are not equal with regard to borrowing and lending, namely some investors may enjoy special borrowing facilities while others may not; second, they are led to introduce a new notion of NFL, which is the classical concept in finite time but does not exclude a free lunch at infinite and is therefore may be more economically meaningful; last, they characterize the NFL assumption for very general investments, which enables to consider investment opportunities that are not necessary related to a market model and, to generalize the results obtained for imperfect markets and to obtain them all in a unified way. Technically, all investment opportunities are described in terms of cash flow. Therefore, separation techniques in more complex spaces to obtain the Fundamental Theorem of Asset Pricing are needed. Let consider their main Assumption $A$.

**Definition 2** An investment is an $(F_t)_{t \in T}$-adapted process $H = (H_t)_{t \in T}$, null outside a finite number of dates, i.e. there exists $(t_1^H, \ldots, t_N^H)$ such that $H_t = 0$ for all $t \notin (t_i^H)_{i=1}^N$, and such that $H_t$ is in $L^1(\Omega, F_t, P)$ for all $t \in T$. 
DEFINITION 3 (ASSUMPTION A) There exists a sequence \( d = (d_n)_{n \geq 0} \) such that for all \( t^* \geq 0 \), for all \( B_{t^*} \in F_{t^*} \) of positive probability, there exists \( H \) in the convex cone of investment opportunities \( J \), of the form \( H_{t^*} = 0 \) outside \( B_{t^*} \), \( H_t = 0 \) for all \( t < t^* \), \( H_t \geq 0 \) for all \( t > t^* \), and there exists \( d_n \in d \), \( P[H_{d_n} > 0] > 0 \).

Roughly, Assumption A corresponds to the possibility of transferring “some money” from any date and event to some particular date. This assumption is not too restrictive: it is satisfied if we can buy at every date and event a bond with a given maturity even if this bond is defaultable and even if there is no secondary market for that bond (i.e., we have to wait until maturity in order to recover any money with a positive probability, which may be different from 1); this includes market models with frictions on the numeraire like no borrowing, different borrowing and lending rates, bonds with default risk, different borrowing facilities among the investors. More generally, it is satisfied if there is at least one asset whose price cannot be negative (which is usually the case for stocks or for options, defaultable bonds, etc.).

Then a characterization of the NA property in a model with flows is given by Jouini and Napp (2002) in the following theorem.

THEOREM 4 Let \( J \) denote a convex cone of investments satisfying Assumption A. There is NFL for \( J \) if and only if there exists a process \( g = (g_t)_{t \in T} \) satisfying for all \( t \in T \), \( P[0 < g_t < M] \) for some \( M \in R_+ \), and such that

\[
E[\sum_{t \in T} g_t H_t] \leq 0 \quad \text{for all} \quad H = (H_t)_{t \in T} \in J.
\]

Moreover, the process \( g \) can be taken \((F_t)_{t \in T}\)-adapted.

In other words, there is NFL for a convex cone of available investments satisfying Assumption A if and only if a given convex set of “admissible” discount processes is non-void. The theorem ensures the existence of a “discount process” such that, using this process as deflator, all available investments have non-positive present value; this means that there exists a term structure such that the market consisting of the primitive investment opportunities and of the additional borrowing and lending facilities is still “arbitrage-free”. Besides, the existence of such a discount process prevents from any arbitrage opportunity. Notice that Assumption A is not needed to obtain this result if the set of investment opportunities is related to a countable set of dates.

Since most market models with frictions can fit in the model with flows for a specific convex cone of available investments, the model in Jouini and Napp (2002) provides a unified framework for the study of the characterization of the absence of FL in such imperfect market models. However this model with flows does not stand for economies with fixed transaction costs, since the set of available investments is not a cone.

Kabanov (1999, 2001) develops a mathematical theory of currency markets with transaction costs based on ideas of convex geometry. He proposed an
appealing framework to model financial markets in a numéraire-free way for both frictionless markets and markets with transaction costs. This approach turns out to be conceptually interesting, even in the frictionless case, as it allows for a new look on the wealth processes, arising in financial modelling, without explicitly using stochastic integration: expressing portfolios in terms of the number of physical units of the assets, as opposed to the values of the assets in terms of some numéraire, opens new perspectives. Basically, the financial market is modelled by a $d \times d$ matrix-valued stochastic process specifying the mutual bid and ask prices between $d$-assets. The terms of trade at time $t$ are modeled via an $F_t$-measurable non-negative $d \times d$ matrix-valued map $\omega \rightarrow \Sigma_t(\omega)$ denoting the bid and ask prices for the exchange between the $d$ assets. The entry $\sigma_{ij}^t$ of $\Sigma_t$ denotes the number of units of asset $i$ from which an agent can trade in one unit of asset $j$ in terms of the asset $i$ bid-ask processes are defined as adapted processes taking values in the set of bid-ask-matrices. $\sigma_{ij}^t = 1$ a.s. for all $1 \leq i, j \leq d$ and $t = 0, ..., T$ in the frictionless case.

Kabanov et al. (2001) introduce the bid-ask process in a somewhat indirect way. They start with a $d$-dimensional price process which models the prices of the $d$ assets without transaction cost in terms of some numéraire (it may be a traded asset or not). One then defines a non-negative $d \times d$-matrix $\Lambda = (\lambda_{ij})_{1 \leq i, j \leq d}$ of transaction cost non-negative coefficients $\lambda_{ij}$, modelling the proportionally factor one has to pay in transaction costs, when exchanging the $i$'th asset into the $j$'th asset. Then the bid-ask process is obtained as

$$\sum_t(\omega) = \operatorname{Diag}(S_t(\omega))^{-1}(1 + \Lambda_t(\omega))\operatorname{Diag}(S_t(\omega)), \quad (2.3)$$

where $I$ denotes the unit matrix (not to be confused with the identity matrix).

Schachermayer (2002) presents a direct modelization of the bid-ask process $\Sigma = (\Sigma_t)_{t=0}^T$ without first defining $(S_t)_{t=0}^T$ and $\Lambda$. It seems more natural, from an economic point of view, as in a market with friction an agent is certainly faced with a bid-and an ask-price. But these prices are not necessarily decomposed into a “frictionless” price and additional transaction costs.

The notion of consistent price system (resp. strictly consistent) introduced by Kabanov and his co-authors extends the notion of equivalent martingale measures. Similar notions are in Schachermayer (2002).

**Definition 5** An adapted $\mathbb{R}_+^d$ valued-process $Z = (Z_t)_{t=0}^T$ is called a consistent (resp. strictly consistent) price process for the bid-ask process $\Sigma_t$ if $Z$ is a martingale under $P$, and $Z_t(\omega)$ lies in $K_t^*(\omega) \setminus \{0\}$ (resp. in the relative interior of $K_t^*(\omega)$) a.s. for each $t=0, ..., T$. 


An Overview of Probabilistic and Time Series Models in Finance

\[ K_t^*(\Sigma) = \{ \omega \in \mathbb{R}^d : <v, \omega > \geq 0, \text{ for } v \in K(\Sigma) \} \]
is the polar of \(-K(\Sigma)\), and \(K(\Sigma)\) is the solvency cone, i.e., the convex cone in \(\mathbb{R}^d\) spanned by the unit vectors \(e_i\), \(1 \leq i \leq n\), and the vectors \(\sigma e_i - e_j\), \(1 \leq i, j \leq d\).

The cone \(K^*(\Sigma)\) has a nice economic interpretation, eluded by the term “consistent price system”. A vector \(\omega \neq 0\) is in \(K^*(\Sigma)\) if it defines a friction-less pricing system for the assets \(1, \ldots, d\) which is consistent with the bid-ask-matrix \(\Sigma\) in the following sense: if the price of asset \(i\) (denoted in terms of some numéraire) equals \(\omega^i\), then the friction-less exchange rates, denoted by \(\tau^{ij}\), clearly equal

\[ \tau^{ij} = \frac{\omega^j}{\omega^i}, 1 \leq i, j \leq d. \]

>From the economical point of view, a consistent price system \(\omega = (\omega^i)_{i=1}^d\) is strictly consistent if, for all \(1 \leq i, j \leq d\), the exchange rate \(\tau^{ij} = \frac{\omega^j}{\omega^i}\) is in the relative interior of the bid-ask spread \([\frac{1}{\sigma^{ij}}, \sigma^{ij}]\).

The main theorem in Kabanov et al. (2001) is the following version of the Fundamental Theorem of Asset Pricing: under an additional assumption, a bid-ask process \(\Sigma\) satisfies the strict NA condition, if there is a strictly consistent price system \(Z\) for \(\Sigma\). The additional assumption is called “efficient friction” and requires that \(F_t(\omega) = \{0\}\), a.s., for all \(t = 0, \ldots, T\). It was asked by these authors whether this additional assumption can be dropped. Schachermayer (2002) gives an example of a bid-ask process \(\Sigma\), with \(d = 5\) and \(T = 2\), showing that, in general, the answer to this question is no. In the same paper a slight strengthening of the notion strict NA, called the robust no arbitrage \(\text{NA}^R\) is introduced. A subsequent Fundamental Theorem of Asset Pricing as a main result is then formulated.

2.2.3 Asset Pricing with Cone Constraints

Pham and Touzi (1999) addresses the problem of characterization of NA in the presence of frictions in a discrete-time financial market model. They extend the Fundamental Theorem of Asset Pricing with cone constraints on the trading strategies under a nondegeneracy assumption. In the presence of transaction costs and under a nondegeneracy condition on the risky assets price process, they also prove that the NFL and the NA conditions are locally equivalent i.e. when trading is restricted to some period \([t - 1, t]\). Their main result states the equivalence of the no local arbitrage condition and the existence of an equivalent probability measure satisfying a further generalization of the martingale property. They do not provide a multiperiod version of this result. For a more general setting of convex constraints see Brannath (1997).
2.2.4 Nonlinear Asset Pricing

On financial markets without frictions, no-arbitrage pricing allows to price non-marketed redundant assets using the equilibrium prices of the marketed assets. Assets are then valued by a linear function of their payoffs (mathematical expectation). The equilibrium prices of the marketed assets determine a set of risk neutral probability distributions such that the equilibrium price of a redundant asset equals the mathematical expectation of its discounted payoff with respect to these probability distributions. This pricing rule is consistent with equilibrium in the sense that, introducing a redundant asset at its no-arbitrage price does not affect the equilibrium allocation; see, for example, Harrison and Kreps (1979). In markets with frictions, pricing rules may be non-linear. Two portfolios yielding the same payoffs need not have the same formation cost (net of transaction cost), but the difference may not imply the existence of a free lunch because of frictions. Consider for example bid-ask spreads or transaction costs. Then clearly prices (as a function of asset payoffs) are non-linear, since the price an agent has to pay for buying an asset is strictly larger than the price an agent receives for selling it. Therefore equilibrium asset prices cannot be represented by the mathematical expectation of their discounted payoff with respect to a probability measure.

Asset valuation by a Choquet integral is introduced in Chateauneuf et al. (1996). They introduce a nonlinear valuation formula similar to the usual expectation with respect to the risk-adjusted probability measure. This formula expresses the asset’s selling and buying prices set by dealers as the Choquet integrals of their random payoffs. In this paper bid-ask spreads are considered. Bid-ask spreads is one of many types of friction prevailing in financial markets which differs from the traditional formalization of proportional transaction costs.

Let consider the following situation pointed out by Chateauneuf et al. (1996): assumed that a dealer sells an asset \( Y \) (defined by its flow of payoffs) at a price \( q(Y) \) and that she buys it a price \(-q(-Y)\) such that she makes the positive profit \( q(Y) + q(-Y) > 0 \). Then, because \( 0 = q(0) = q(Y - Y) \), \( q \) cannot be linear, hence it cannot be calculated as \( q(Y) = \int_S Y d\mu \), where \( S \) is the set of random states and \( \mu \) is some risk-adjusted probability over \( S \). In these settings, the paper imposes certain axioms on prices (generalizing the usual no-arbitrage conditions) and deduces from them a result on the structure of prices (representation as Choquet integral: an expectation with respect to a concave capacity). Capacities were introduced by Schmeidler (1989) in individual decision theory. Formally, a capacity on a measurable space \( (S, F_S) \) is a set of functions \( \nu : F_S \to [0, 1] \) satisfying \( \nu(F_S) = 1, \nu(\emptyset) = 0 \). Furthermore \( \nu \) is said to be convex (resp. concave or supermodular) if
An Overview of Probabilistic and Time Series Models in Finance

In this context, a convex capacity is interpreted as a representation of risk (uncertainty) aversion. This characterization of uncertainty aversion has been used in single-agents models for which convex capacities are representations of individual behaviors. In contrast, Chateauneuf et al. (1996) use a model for which agents are price takers and the concave capacity is derived from prices.

Formally, the model uncertainty they consider is described by the measurable state space \((S, F_S)\) where \(F_S\) is a given \(\sigma\)-algebra of events of \(S\). An asset is defined by the random variable \(X\) of its payoffs. Bounded assets are considered. These assets are sold and bought by a dealer to agents. Hence, all traded assets have a bid and an ask price fixed by the dealer. These prices are described by \(q(Y)\) and \(-q(-Y)\) respectively, i.e., the prices at which the dealer sells asset \(Y\) to agents and buy asset \(Y\) from agents. Three axioms on prices which generalize the usual NA conditions to market with a dealer are then imposed. The first is the usual NFL. The second one, as is usually done in pricing models, assumes no transaction costs on riskless assets. The third axiom replaces the (usually implicit) tight markets condition. Traditionally, two portfolios yielding the same payoffs must have the same price, implying that price functional is linear. Taking into account potential reduction of risks when portfolio \(X + Y\) is sold instead of \(X\) or \(Y\) alone induces the dealer to sell \(X + Y\) at a discount to \(X\) and \(Y\).

A typical example where hedging effects occur and \(X\) and \(Y\) are not comonotone (comonotonicity := for all \(s, s' \in S, [X(s) - X(s')] [Y(s) - Y(s')] \geq 0\)), is the following one from Chateneauf et al. (1996). Suppose that \(X\) offers 1000 if even \(B\) occurs, 5000 otherwise, \(Y\) offers 5000 if \(B\) occurs, 1000 otherwise. Clearly \(X\) and \(Y\) are not comonotone and \(X\) (resp \(Y\)) is a hedge against \(Y\) (resp \(X\)) since \(X + Y\) is riskless: it offers 6000 with certainty. So, subadditivity for \(q : q(X + Y) \leq q(X) + q(Y)\) is required. Notice that, consequently, no discount will be offered by the dealer when \(X\) and \(Y\) are comonotone; i.e., \(q(X + Y) = q(X) + q(Y)\) if \(X\) and \(Y\) are comonotone. Then the third axiom (Comonotonicity Premium) expresses for all \(X, Y \in A : q(X + Y) \leq q(X) + q(Y)\) equality holds if \(X\) and \(Y\) are comonotone. Their main result is the so-called Choquet Sublinear Pricing Theorem. Under the three axioms as above this theorem asserts that there exists a unique concave capacity \(\nu\) on the set of states \(S\) such that the value of an asset \(X\) is defined by \(q(X) = \max\{\int_S X d\mu; \mu\text{ is an additive probability s.t. } \mu \leq \nu\}\). The price of \(X\) is the Choquet integral of its payoffs: 

\[
\int_S X d\nu = \int_{\mathbb{R}_-} [\nu(X \geq t) - 1]dt + \int_{\mathbb{R}_+} [\nu(X \geq t)]dt,
\]

\(\nu(A \cup B) + \nu(A \cap B) \geq (\text{resp. } \leq) \nu(A) + \nu(B), \text{ for all } A, B \in F_S.\)
and $q$ is sublinear (i.e. subadditive and positively homogeneous, and indeed $q$

Application to pricing “primes” and “scores” are given in the paper of
Chateauneuf et al. (1996).

In these settings De Waegenaere et al. (1996) propose a pricing rule for the
valuation of assets on financial markets with intermediaries. They assume that
the non-linearity arises from the fact that dealers charge a price for their inter-
mediation between buyer and seller. The price of an asset equals the signed
Choquet integral of its discounted payoff with respect to a concave signed ca-
pacity. Furthermore, they show that this pricing rule is consistent with equilib-
rium and equilibria satisfy a notion of constrained Pareto optimality.

On the other hand, a universal framework for pricing financial and insurance
risks has been introduced recently by Wang (2000) who proposes a pricing
method based on the following transformation $F^*(x) = \phi[\phi^{-1}(F(x)) + \lambda]$, where $\phi$ is the standard normal cumulative distribution. The key parameter $\lambda$
is called the market price of risk, reflecting the level of systematic risk. For a
given asset $X$ with $F(x) = \Pr\{X \leq x\}$, the Wang transform will produce
a “risk-adjusted” cumulative probability distribution $F^*(x)$. The mean value
under $F^*(x)$ will define a risk-adjusted “fair value” of $X$ at time $T$, which
can be further discounted to time zero, using the risk-free interest rate. This
approach is partly inspired in the work of Venter (1991) and Butsic (1999).

2.3 Time series models

In this section, we revise the literature on the time series models usually
fitted to financial data. As this is a very broad area, the focus is only on the
main branches of the literature with special attention to the most recent devel-
opments. Campbell et al. (1997) and Tsay (2002) present excellent textbook
reviews of Financial Econometrics and Bollerslev (2001) and Engel (2001,
2002a) have very interesting discussions on past developments and future per-
spectives in this area.

Traditionally, the two main motivations to use time series models to ana-
lyze financial data are to represent the empirical properties often observed in
real prices and to estimate and test the financial models described in section 2.
In this section, we describe models proposed mainly to represent the empiri-
cal properties of financial prices while section 4 is devoted to the relationship
between time series models and Finance theory.

The empirical properties of financial prices depend crucially on the fre-
quency of observation. We consider three main classes of frequencies. First,
it is possible to observe prices at very high frequencies as, for example, tick
by tick or hourly prices. These observations are called Ultra-high-frequency
(UHF) data by Engle (2000) and they are usually characterized by unequally
An Overview of Probabilistic and Time Series Models in Finance

spaced and discrete-value observations. Another important property is the presence of strong daily patterns with highest volatility at the open and toward the close of the day. On top of this intraday volatility pattern, UHF returns are characterized by highly persistent conditionally heteroscedastic components along with discrete information arrival effects; see Andersen and Bollerslev (1997a, 1997b, 1998), Müller et al. (1997) and Andersen et al. (2001). Finally, it is possible to have multiple transactions within a single second.

Prices can also be observed at high frequencies, as for example, daily or weekly. This frequency is the most extensively analyzed in the empirical literature. There is a vast number of papers that show that high frequency returns are nearly non-correlated although they are not independent because there are non-linear transformations, as squares or absolute values, that have significant autocorrelations. Furthermore, these autocorrelations are usually small and decay very slowly towards zero. The significant autocorrelations of squared returns are often related with the presence of volatility clustering, i.e. periods of low volatility are usually followed by periods of low volatility and vice-versa. Furthermore, the slow decay is usually interpreted as the presence of long-memory in the volatility; see Lobato and Savin (1998) and Granger et al. (2000) and the references therein. On the other hand, high frequency returns are often leptokurtic and, consequently, non-Gaussian. The heavy tails property of returns can also be related with the dynamic evolution of volatility.

Finally, prices are sometimes observed at very low frequencies as, for example, monthly. Tsay (2002) shows that monthly returns still have excess kurtosis although smaller than in lower frequencies. On the other hand, monthly returns seem to have more serial correlations than daily returns. Given that low frequencies are not in general of interest for asset pricing models, the focus in this section is on UHF and high frequency observations.

The rest of the section is organized as follows. Subsections 3.1 to 3.3 deal with models for high frequency observations. In subsections 3.1 and 3.2, we describe the models usually fitted to represent expected returns and volatilities respectively. In subsection 3.3, we consider multivariate models for systems of returns. Finally, in subsection 3.4, we describe models for UHF data.

2.3.1 Models for the conditional mean

One of the central questions in the Financial Econometrics literature is whether financial prices are predictable and this is still a topic of controversy; see, for example, the special issue of the Journal of Empirical Finance, 8 (2001). In this section we describe univariate models and, consequently, the problem is whether future prices can be predicted with information contained in their own past. The main hypothesis that have often been tested are the martingale
and the random walk hypothesis. The martingale hypothesis can be expressed as follows:

\[ E[P_t \mid P_{t-1}, P_{t-2}, \ldots] = P_{t-1} \]  \hspace{1cm} (3.1)

Therefore, given the prices up to time \( t - 1 \), the price at time \( t \) is expected to be equal to the price at time \( t - 1 \). The martingale hypothesis places a restriction on expected returns but does not take into account the risk. However, as said in section 2, once asset returns are properly adjusted for risk, the martingale hypothesis holds for rationally determined asset prices; see Harrison and Kreps (1979). It is known that, the risk-adjusted martingale property is the basis of many financial derivatives as, for example, options and swaps; see, for example, Merton (1990) and Campbell et al. (1997).

The second hypothesis often tested in the financial literature is whether prices are generated by a random walk plus drift model given by:

\[ P_t = \mu + P_{t-1} + \epsilon_t \]  \hspace{1cm} (3.2)

where \( \epsilon_t \) is an independent process with zero mean and variance \( \sigma^2 \) and \( \mu \) is the expected price change. In model (5), if the distribution of the errors \( \epsilon_t \) is, for example, Gaussian, there is a positive probability that prices can be negative, violating limited liability. Therefore, it is usual to assume the random walk model not for prices but for logarithmic prices, i.e.

\[ \log(P_t) = \mu + \log(P_{t-1}) + \epsilon_t \]  \hspace{1cm} (3.3)

In model (6) any arbitrary transformation of prices is unforecastable using any arbitrary transformation of past prices. However, it is usual to assume that the errors \( \epsilon_t \) are merely uncorrelated instead of independent allowing, for example, for the presence of conditional Heteroscedasticity. As we have mentioned before, this is a property often observed in high frequency returns. Consequently, we will focus on tests of the random walk hypothesis where \( \epsilon_t \) is uncorrelated.

When testing the null hypothesis that the autocorrelation coefficients of returns, \( r_t = \Delta \log(P_t) \), are all zero, it is important to take into account that \( \epsilon_t \) is not independent because, usually, \( \epsilon_t^2 \) is correlated. Therefore, the traditional tests for uncorrelatedness should be adequately modified; see Romano and Thombs (1996) and Lobato et al. (2001) among others.

Alternatively, the random walk hypothesis can be tested using the Variance Ratio (VR) statistic. This test is based on the property that the variance of random walk increments is a linear function of time interval; see Campbell et al. (1997) for a detailed description of the VR test.

The implementation of the previous tests to financial prices, seems to suggest that financial asset returns are predictable; see the special issue of the
Journal of Empirical Finance, 8 (2001) and the references therein. There are several alternative explanations for this predictability. For example, Campbell et al. (1997) and Lo and MacKinlay (1990) show that nonsynchronous trading can introduce negative autocorrelations in returns. The bid-ask spread can also introduce negative autocorrelations in asset returns; see, among others, Campbell et al. (1997). Other possible explanations are time-varying risk premiums as in Harvey (2001) and Bekaert et al. (2001), irrational behavior of market participants in Hong and Stein (1999), Benartzi and Thaler (1995), Barberis et al. (2001) and Epstein and Zin (2001), market frictions as transaction costs or agency problems or fluke due to statistical inference.

2.3.2 Models for the conditional variance

Although, it is generally accepted that asset returns appear to be close to a martingale difference process, there is an overwhelming evidence that they are not independent due to autocorrelated squares. Assuming that returns have zero mean and are serially uncorrelated, they can be represented by the following model:

\[ r_t = \sigma_t \varepsilon_t \]  

(3.4)

where \( \varepsilon_t \) is an independent and identically distributed (i.i.d.) process with zero mean and unity variance independent of the volatility, \( \sigma_t \). There are two main proposals in the literature to represent the dynamic evolution of \( \sigma_t \): Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and Stochastic Volatility (SV) models.

GARCH models, originally proposed by Engle (1982) and Bollerslev (1986), are based on modelling the volatility as the variance of returns conditional on past observations. There is a pleyade of papers where GARCH models are investigated from a theoretical point of view or are applied to the empirical analysis of financial time series. The main properties of GARCH models have been reviewed, among others, by Bollerslev et al. (1995) and Carnero et al. (2001a). Although the original motivation of GARCH models was mainly empirical, Nelson (1992) shows that even when mispecified, ARCH models may serve as consistent filters for the continuous–time stochastic volatility diffusions often employed in the asset pricing literature. Furthermore, Nelson (1990, 1994) and Nelson and Foster (1994) provide some important links between GARCH and the corresponding continuous–time models.

The original GARCH model has been extended in a huge number of directions. Two of the main extensions from the empirical point of view, are models to represent the asymmetric response of volatility to positive and negative returns and to represent the effect of the volatility on the return of a stock. The first effect is known as leverage effect and was introduced by Black (1986).
The first model proposed to represent the leverage effect was the Exponential GARCH (EGARCH) model of Nelson (1991). Later, Hentschel (1995), Duan (1997) and He and Terasvirta (1999) have proposed models general enough to unify many of the main previous ARCH-type models. With respect to the effect of volatility on the expected return, Engle et al. (1987) introduced the GARCH in mean (GARCH-M) model given by

\[ r_t = \mu + c\sigma_t^2 + a_t \]
\[ a_t = \sigma_t \varepsilon_t \]
\[ \sigma_t^2 = \omega + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2 \]  

(3.5)

The parameter $c$ is known as the risk premium parameter. Returns generated by the GARCH-M model are autocorrelated because of the autocorrelations of the volatility, $\sigma_t^2$.

There are many other generalizations of the original GARCH model. For example, Zakodian (1994) allows for regime switching where volatility persistence can take different values depending on whether returns are in a high or a low volatility regime. To represent the long memory property of squared returns, Baillie et al. (1996) introduce the Fractionally Integrated GARCH (FIGARCH) model. Although the FIGARCH model has been fitted in several empirical applications, it is not stationary in covariance and, consequently, the properties of the corresponding estimators and tests are generally unknown. Finally, Engle and Lee (1999) have proposed a GARCH model with two components in volatility: one which is nearly nonstationary and another that is much less persistent.

All GARCH models have the attractive that can be easily estimated by Maximum Likelihood techniques. However, Terasvirta (1996) and Carnero et al. (2001b) show that the basic GARCH(1,1) model is not flexible enough to represent adequately the properties often observed in real time series of returns.

Alternatively, the volatility, $\sigma_t^2$, can be modelled using SV models that introduce an additional noise in its equation. Therefore, the volatility is a latent variable composed of a predictable component, that depends on past returns, plus an unexpected component. SV models were originally proposed by Taylor (1986) and their properties have been reviewed by Taylor (1994), Ghysels et al. (1996) and Shephard (1996). The introduction of the unobserved component in the representation of the volatility, gives more flexibility to SV models to represent the empirical properties often observed in real time series of returns; see Carnero et al. (2001b). However, the estimation of these models present some added difficulties over the estimation of GARCH models. The likelihood function has not a close form and, consequently, most estimation methods proposed in the literature are based on numerical approximations of the likelihood or on transformations of the observations. Although, there is
not still a consensus about which are the most adequate methods to estimate SV models, recently there has been important progress towards methods that are computationally feasible and, at the same time, have properties similar to the Maximum Likelihood estimators; see Broto and Ruiz (2002) for a detailed description of estimation methods for SV models.

Recently, Chib et al. (2002) have proposed the following SV model where returns can contain a jump component to allow for large, transient movements,

\[ r_t = x_t' \beta + k_t q_t + w_t' \sigma_t \varepsilon_t \]
\[ \log \sigma_t^2 = \mu + z_t' \alpha + \phi (\log \sigma_{t-1}^2 - \mu) + \eta_t \]

where \( x_t, w_t \) and \( z_t \) are covariates and \( \gamma \) denotes the level effect. The covariate \( w_t \) is a non-negative process as, for example, lagged interest rates; see Andersen and Lund (1997). The noises \( \varepsilon_t \) and \( \eta_t \) are mutually independent Student-t and Gaussian white noise processes respectively, both with zero mean and variances one and \( \sigma_{\eta}^2 \). Finally, with respect to the jump component, \( q_t \) is a Bernoulli random variable that takes value one with probability \( \kappa \) and \( k_t \) is the size of the jump distributed as \( \log(1 + k_t) \sim N(-0.5\delta^2, \delta^2) \). They argue that model (9) without the jump component can be thought of as an Euler discretization of a Student-t Lévy process with additional stochastic volatility effects. This process has been used in the continuous time options and risk assessment literature; see, for example, Barndorff-Nielsen and Shephard (2002b), Eberlein (2002) and Eberlein and Prause (2002). On the other hand, models with jumps have also been frequently applied in continuous time models of financial asset pricing; see, for example, Merton (1976), Ball and Torous (1985), Bates (1996), Duffie et al. (2000) and Barndorff-Nielsen and Shephard (2001). From the point of view of the Financial Econometrics literature, SV models with jumps have been previously considered by Chernov et al. (2000), Barndorff-Nielsen and Shephard (2002a) and Eraker et al. (2003).

As in the case of GARCH models, SV models have also been extended to represent the asymmetric response of volatility to negative and positive returns and the response of expected returns to volatility by Harvey and Shephard (1996) and Koopman and Uspensky (2002) respectively. Another extension of SV models considered in the literature is to allow for long memory in volatility; see Harvey (1998) and Breidt et al. (1998).

### Models for conditional covariances

Multivariate models have been often used to represent financial series of returns related, for example, with the Asset Pricing Theory (APT), asset allocation, estimation of time-varying betas or Value at Risk (VaR). However, although numerous multivariate models for returns have been proposed, there
is not yet a consensus about which models are better mainly due to a dimensionality problem. The literature on multivariate GARCH models is often related with the lack of parsimony of these models and the constraints needed to guarantee that the conditional covariance matrix, $\Sigma_t$, is positive definite; see Engle (2002a,b) who revises the most popular multivariate models proposed in the context of GARCH. The dimensionality becomes very quickly a problem because the conditional covariance matrix of a k-dimensional return series has $k(k+1)/2$ distinct quantities. To keep the number of parameters low, Bollerslev (1990) considers a multivariate GARCH model with constant correlations that always satisfies the positive-definite condition of $\Sigma_t$. The constant correlation hypothesis can be tested using the Lagrange multiplier test proposed by Tse (2000). Because of its computational simplicity, the constant correlation model of Bollerslev (1990) has been widely used in the empirical analysis of financial data. However, if the correlations evolve over time, this model is inadequate and can give incorrect inferences. Very recently, there have been different proposals of multivariate GARCH models with time-varying conditional correlations. For example, Tsay (2002) proposes two alternative ways of dealing with the conditional covariance matrix. The first one consists of modeling directly the evolution of the autocorrelation and the second is based on the Cholesky decomposition of $\Sigma_t$. The attractive of the second alternative is that it does not require any constraint to ensure the positive definiteness of $\Sigma_t$. Alternatively, Tse and Tsui (2002) propose a multivariate GARCH (MGARCH) model with time-varying correlations where the constraints required to ensure positive definite covariance matrix can be imposed during the optimization procedure. Finally, Engle (2002b) proposes a nonlinear Dynamic Conditional Correlation (DCC) model that can be estimated in two steps from univariate GARCH models. Alternatively, Ledoit et al. (2003a) also propose a two step estimation procedure of the original unrestricted diagonal-Vech multivariate GARCH(1,1) model of Bollerslev et al. (1988) given by

$$ Cov(r_{it}, r_{jt} \mid \Omega_{t-1}) = h_{ij,t} = c_{ij} + a_{ij} r_{it-1} r_{jt-1} + b_{ij} h_{ij,t-1} \quad (3.7) $$

In the first step, the parameters are estimated separately by estimating the two-dimensional or one-dimensional equations in (10). Then, the estimated matrices are transformed to guarantee positive semi-definiteness.

An extensive and detailed comparison between the alternative models to represent time-varying correlations is still to be done.

Another completely different approach to simplify the dynamic structure of a multivariate volatility process is to use factor models. Multivariate factor models provide a way of dealing with the APT; see, for example, Campbell et al. (1997) for a very simple exposition. Denoting by $y_t$ the $N \times 1$ vector of returns at time $t$, it is given by
where $D$ is a diagonal matrix, $B$ is the matrix of factor loadings and $f_t$ is a $K$ dimensional vector of factors. The APT says that, as the dimension of $r_t$ increases (approximating the market), then $\alpha \simeq \iota r + B\lambda$, where $\iota$ is the riskless interest rate, $r$ is a vector of ones and $\lambda$ is a vector representing the factor risk premium associated with the factors often identified as the variances of the factors. However, the normality assumption in (11) is usually inadequate for high frequency series of returns. Consequently, this assumption has been relaxed in the consequent literature. Diebold and Nerlove (1989) and King et al. (1994) analyze factor models where the factors and idiosyncratic errors follow their own ARCH process. Sentana and Fiorentini (2001) show that the identifiability restrictions for conditionally heteroscedastic factor models are less severe than in static factor models.

In the context of SV models, the first multivariate model was originally proposed by Harvey et al. (1994) who allow the variances and covariances to evolve through time with possibly common trends. Later, Ray and Tsay (2000) used the same model to study common long memory components in daily stock volatilities of groups of companies. However, the multivariate SV model of Harvey et al. (1994) restricts the correlations to be constant over time. Later, Jacquier et al. (1995) propose a factor SV model given by

$$r_t = \alpha + B f_t + \varepsilon_t$$

$$\varepsilon_t \sim NID(0, I)$$

$$f_t \sim SV(\phi; \sigma^f_t; 0), i = 1, ..., K$$


2.3.4 Models for intradaily data

The analysis of UHF data is closely related with what is known as Market Microstructure and is one of the most active research areas in Financial Econometrics. However, traditional econometric tools may not be appropriate as tick by tick observations are not equally spaced and discrete valued. In this case, it is possible to use market point processes or continuous time methods in which the sampling frequency is determined by some notion of time.
deformation; see, for example, Andersen (1996). With respect to using UHF data to estimate the volatility, Andersen and Bollerslev (1998) show that the precision of volatility forecast is improved if the data are sampled more frequently. However, UHF data are affected by problems as the bid-ask spread or non-synchronous trading that, as previously mentioned, can generate autocorrelations in returns. Andersen et al. (2001) develop new robust methods for inference in the UHF data setting. Their approach is based on an extension of the Fourier Flexible Form (FFF) regression framework.

Hausman et al. (1992) proposes an ordered probit model to study price movements in transactions data where the explanatory variables are the duration between trades, the bid-ask spread, the lagged values of price change and volume, the return of the S&P500 index and an indicator variable that depends on the bid and ask prices. Alternatively, Rydberg and Shephard (2003) propose to decompose the price change into three components: an indicator for the price change, the direction of the change and the size of the change.

Finally, when analyzing UHF data, it is important to model not only the trades but also the timing between trades. In this sense, Engle and Russell (1998) propose the Autoregressive Conditional Duration (ACD) model that estimates the distribution of the time between events conditional on past information. Later, Dufour and Engle (2000) show that the more frequent the transactions, the greater the volatility. Furthermore, they show that transaction arrivals are predictable based on economic variables as the bid-ask spread. Zhang et al. (2001) extend the ACD model to account for nonlinearity and structural breaks in the data. Finally, Tsay (2002) introduces the Price Change and Duration (PCD) model to describe the multivariate dynamics of prices changes and associate durations.

2.4 Applications of time series to financial models

Summarizing the literature described in sections 2 and 3, it seems rather clear that there is a gap between the theoretical asset pricing and the Financial Econometrics literature. First, although continuous time methods and no-arbitrage arguments are prominent in the asset pricing literature, most influential contributions have been derived under very restrictive assumptions about the underlying process. For example, the Black-Scholes option valuation formula assumes constant volatility when, it is generally accepted empirically, that volatility evolves over time. However, recently, some authors have proposed more realistic continuous time processes with time varying volatilities; see, for example, Hull and White (1987), Heston (1993), Duffie and Kan (1996) and Dai and Singleton (2000). Engle (2001) suggests that the use of UHF data potentially could provide information on the more appropriate class of diffusion models to use for pricing both underlying and derivative assets.
On the other hand, the Financial Econometrics literature has many challenges to provide instruments adequate to represent the behavior of asset prices. The econometrics of, for example, jump diffusion or affine models are difficult. Bollerslev (2001) points out that recent research on the link between the probability distributions of actual asset prices and the corresponding risk-neutral probability distributions implied by derivative prices has just started and that much research remains to be done. Some relevant references in this sense are Aït-Sahaila and Lo (2000), Andersen et al. (2002), Chernov and Ghysels (2000) and Duffie et al. (2000). Also, it is very useful the guest editorial by Ghysels and Tauchen (2003) and all the papers within the special issue of the *Journal of Econometrics* on the intersection between Financial Econometrics and Financial Engineering.

### 2.4.1 Estimation of the CAPM

Two classical pricing models arise in the financial literature. **Capital Asset Pricing Model (CAPM)** is a set of predictions concerning equilibrium expected return on assets; see, for example, Sharpe (1964) or Lintner (1965). Classic CAPM assumes that all investors have the same one-period horizon, and asset returns have multivariate normal distributions. For a fixed time horizon, let $R_i$ and $R_M$ be the returns of asset $i$ and of the market portfolio $M$, respectively. Classic CAPM, sometimes called Sharpe-Lintner CAPM, asserts that

$$E[R_i] = r + \beta_i \{E[R_M] - r\}$$

(4.1)

where $r$ is the risk-free return and $\beta_i = \frac{\text{cov}(R_i, R_M)}{\sigma^2_M}$ is the beta of asset $i$.

Assuming that asset returns are normally distributed and the time horizon is one period (e.g., one year), a key concept in financial economics is the **market price of risk**, given by $\lambda_i = \frac{E[R_i] - r}{\sigma_i}$. In asset portfolio management, this is also called the **Sharpe Ratio**, after William Sharpe.

In terms of market price of risk, CAPM can be restated as follows:

$$\lambda_i = \frac{E[R_i] - r}{\sigma_i} = \frac{\text{cov}(R_i, R_M)}{\sigma_i \sigma_M} \cdot \frac{E[R_M] - r}{\sigma_M} = \rho_{i,M} \lambda_M,$$

(4.2)

where $\rho_{i,M}$ is the linear correlation coefficient between $R_i$ and $R_M$. In other words, the market price of risk for asset $i$ is directly proportional to the correlation coefficient between asset $i$ and the market portfolio $M$.

CAPM automatically prices assets in the set of all linear combinations of basic assets according to this linearity rule, as long as the market portfolio used in the CAPM is the mean-variance efficient portfolio of risky assets (alternative termed the Markowitz portfolio). CAPM provides a powerful insight regarding
the risk-return relationship, where only systematic risk deserves an extra risk premium in an efficient market. However, CAPM and the concept of “market price of risk” were developed under the assumption of normal multivariate distributions for asset returns, and in practice the underwriting beta can be difficult to estimate.

On the other hand, a common practice pricing non-marketed assets is to infer the price applying the CAPM formula to this asset as well, by simply entering the random payoff $B$ corresponding to the non-marketed asset into the CAPM formula. Technically, the new price has a systematic relationship to the prices of the basic assets, more precisely, it is the price of the marketed asset that best approximates the random payoff $B$ in the sense of minimum expected squared error. Following geometric and statistical considerations, Luenberger (2002a) proposes a correlation pricing formula similar to the CAPM formula, which expresses the price of a non-marketed asset in terms of a priced asset that is the most correlated with the non-marketed asset, rather than in terms of the marked portfolio. The method has accuracy advantages when values in the formula must be estimated. Beyond the NA principle, Luenberger (2002b) derives a pricing method for non-marketed assets determining the price such that an investor with a specific utility function will elect to include the new asset in his/her portfolio at the zero level. The idea of zero-level pricing of a non-marketed payoff is to find the price such that a certain investor will elect to neither purchase nor short it. At this price the investor is indifferent to the inclusion of the considered payoff. Conditions ensuring for such a price to be unique are given in Luenberger (2002b).

Besides CAPM, another major financial pricing paradigm is modern option pricing theory, first developed by Black and Scholes (1973). Unfortunately, the Black-Scholes formula only applies to lognormal distributions of market returns. Options pricing is performed in a world of Q-measure, where the available data consists of observed market prices for related financial assets. On the other hand, actuarial pricing takes place in a world of P-measure, where the available data consists of projected losses, whose amounts and likelihood need to be converted to a “fair value” price; see Panjer (1998). Because of this difference in types of data available, modern option pricing is mostly concerned with the minimal cost of setting up a hedging portfolio, whereas actuarial pricing is based on actuarial present value of costs, with additional adjustments for correlation risk, parameter uncertainty and cost of capital. In these setting new research directions are proposed in the recent literature.

The statistical framework for estimation and testing for the classical CAPM is the Maximum Likelihood (ML) approach; see Campbell et al. (1997), Gibbons et al. (1989) and Bollerslev et al. (1988).

Inferences when there are deviations from the assumption that returns are jointly normal and iid through time have been developed. Tests which accom-
moderate non-normality, heteroscedasticity, and temporal dependence returns are of interest for two reasons. First, while the normality assumption is sufficient, it is not necessary to derive the CAPM as a theoretical model. Rather, the normality assumption is adopted for statistical purposes. Without this assumption, the finite sample properties of asset pricing model tests are difficult to derive. Second, departures of monthly security returns from normality have been documented. As we have pointed out in this review, there is also abundant evidence of heteroscedasticity and temporal dependence in stock returns. It is therefore of interest to consider the effects of relaxing these statistical hypothesis. Robust tests of the CAPM can be constructed using a Generalized Method of Moments (GMM). Within the GMM framework, the distribution of returns conditional on the market return can be both serially dependent and conditionally heteroscedastic. The only assumption is that excess asset returns are stationary and ergodic with finite fourth moments. GMM procedure to estimate time-varying term premia and a consumption based asset pricing model are used in Hansen and Singleton (1982) and Hansen and Scheikman (1995).

Other lines of research are also of interest. One important topic is the extension of the framework to test conditional versions of the CAPM, in which the model holds conditional on state variables that describe the state of the economy. Econometric methods from section 3 are suitable for testing the conditional CAPM.

Another important subject is Bayesian analysis of mean-variance efficiency and the CAPM. Bayesian analysis allows the introduction of prior information. Harvey and Zhou (1990) and Kandel et al. (1995) are examples of work with this perspective.

There is a controversy about the statistical evidence against the CAPM in the past 30 years. Some authors argue that the CAPM should be replaced by multifactor models with several sources of risk; others argue that the evidence against the CAPM is overstated because of mismeasurement of the market portfolio, improper neglect of conditional information, data snooping, or sample-selection bias; and yet others claim that no risk-based model can explain the anomalies of stock-market behavior. Campbell et al. (1997) explore multifactor asset pricing models.

2.4.2 Estimation of the term structure

There is a vast literature devoted to the estimation of dynamic models of the term structure that describe the evolution of yields at all maturities. One of the main problems in this area is that the theoretical models need to be complex enough as to represent adequately the empirical complexity often observed. However, as the complexity of the models increases, their estimation becomes more difficult.
Models of the term structure focus mainly on affine models, characterized originally by Duffie and Kan (1996), that assume that the market price of risk is a multiple of the interest rate volatility and that the state variables are independent. Under these assumptions, ML estimation of the parameters is feasible. However, many empirical studies have shown that this model has fundamental limitations; see, for example, Ghysels and Ng (1998) and Dai and Singleton (2000) between many others. To overcome these limitations, Dai and Singleton (2000) propose the multivariate affine term structure models while Ahn et al. (2002) propose the quadratic term structure models. However, neither of these models is able to track adequately the dynamic evolution of volatility. Recently, Ahn et al. (2003) investigates whether an hybrid model between affine, quadratic and nonlinear models is able to outperform each of the individual models. However, they conclude that, in general, this is not the case. Dai and Singleton (2003) is an excellent review on models of the term structure described from the point of view of their empirical implementation. They focus on the fit of the theoretical specifications of dynamic structure models to the historical shapes of the yield curves.

On the other hand, as we mentioned before, the estimation of these more complex models becomes difficult as the likelihood does not have, in general, a close form. One of the most popular methods in this context is the Efficient Method of Moments (EMM) of Gallant and Tauchen (1996). Duffee and Stanton (2003) estimate a multifactor term structure model with correlated factors, nonlinear dynamics and flexible price of interest rate risk, using both the EMM and an approximate Kalman filter. They conclude that the best results are obtained when the latter procedure is used to estimate the model although it is not asymptotically optimal. However, their results reveal severe biases in the parameter estimates regardless of the estimation method; see also Duan and Simonato (1999) and Chen and Scott (2002) for other authors that have also used the Kalman filter to estimate the term structure.

2.4.3 Estimation of the VaR

Regulators and risk managers are interested in obtaining measures of the Value at Risk (VaR), defined as the expected loss of a portfolio after a given period of time (usually 10 days) corresponding to the $\alpha$% quantile (usually 1%). This interest has motivate new methods designed to estimate the tails of the distribution of returns. There are several methods to estimate the VaR. The early VaR parametric models impose a known theoretical distribution to price changes. Usually it is assumed that the density function of risk factors influencing asset returns is a multivariate normal distribution as, for example, in J.P. Morgan (1996). The most popular parametric methods are variance-covariance models and Monte Carlo simulation. However, excess kurtosis of
these factors will cause losses greater than VaR to occur more frequently and be more extreme than those predicted by the Gaussian distribution. Consequently, several authors propose to use nonparametric (historical simulation) and semiparametric models that avoid to assume a particular distribution of price increments although they usually assume independent increments; see, for example, Danielsson and de Vries (1998). Finally, some authors propose to use extreme value theory estimation of tail shapes to estimate the VaR; see, for example, Embrechts et al. (1997) and McNeil and Frey (2000). In relation with these methods, Pearson and Smithson (2002) describe refinements which increase computational speed and improve accuracy.

However, as described in previous sections, financial returns are often characterized by volatility clustering and non-Gaussianity. Therefore, several authors have considered extensions of the previous approaches that allow for time-varying volatilities. The most popular approach is to estimate the VaR based on Conditional Gaussian GARCH models; see, for example, Christoffersen and Diebold (2000) and Christoffersen et al. (2001). Guermat and Harris (2002) even extend further the GARCH approach to allow for kurtosis clustering.

Recently, Engle and Manganelli (1999) have proposed a conditional quantile estimation based on the CaViar model given by

$$\text{VaR}_t = \beta_0 + \beta_1 \text{VaR}_{t-1} + \beta_2 |y_{t-1}| \quad (4.3)$$

Gourieroux and Jasiak (2001) describe several alternative methods to estimate the VaR, focusing on their main advantages and limitations. Tsay (2002) also describe several of these methods and compare their performance to estimate the VaR of daily returns of IBM stocks. In particular, he compares the RiskMetrics methodology developed by J.P. Morgan, GARCH models, non-parametric estimation, quantile regression and extreme value, finding substantial differences among the approaches.

Given that, as we have mentioned already, the distribution of high frequency price increments is non-Gaussian, and even in many cases the conditional distribution of GARCH models is not Gaussian, many authors suggest using bootstrap techniques to avoid particular assumptions on the distribution of factors beyond stationarity of the distribution of returns; see, for example, Barone-Adessi et al. (1999), Barone-Adessi and Giannopoulos (2001) and Vlaar (2000). Ruiz and Pascual (2002) review the use of bootstrap methods to estimate the VaR.

Although there is a huge number of papers devoted to analyze methods to estimate the VaR as a measure of financial risk, this measure is not without criticisms; see, for example, Szego (2002) and the papers contained in the especial number of the Journal of Banking and Finance, 26. There are several
new measures of risk proposed as remedy for the deficiencies of VaR as, for example, Conditional VaR (CVaR) and Expected Shortfall.

### 2.4.4 Estimation of diffusion processes

There are two relatively independent lines in financial modeling: continuous–time models typically used in theoretical finance and discrete-time models favored for empirical work. The continuous–time models are dominated by the diffusion approach. In contrast to stochastic differential equations used in discrete-time models, stochastic differential equations are widely used to describe continuous–time models in the theoretical finance literature. The stochastic processes characterized by the stochastic differential equations are Itô processes, and continuous–time model assumes that a security price \( S_t \) follows the stochastic differential equation:

\[
dS_t = \mu_t S_t dt + \sigma_t S_t dW_t \quad t \in [0, T]
\]

where \( W_t \) is a standard Wiener process, \( \mu_t \) is called diffusion drift in probability or instantaneous mean rate of return in finance and \( \sigma_t^2 \) is called diffusion variance in probability or instantaneous conditional variance (or volatility). The celebrated Black-Scholes model corresponds to (16) with constants \( \mu_t \) and \( \sigma_t \). Given that financial time series tend to be highly heteroscedastic, the general modelization assumes that \( \sigma_t^2 \) is random and itself is governed by another stochastic differential equation.

For continuous–time models, the “no arbitrage” condition, as we have extensively developed in section 2, can be characterized by a martingale measure, that is, a probability law under which \( S_t \) is a martingale. Prices of options and derivatives are then the conditional expectation of certain functionals of \( S \) under this measure. The calculations and derivations can be manipulated by tools as the Itô lemma and Girsanov theorem; see Karatzas and Shreve (1991) or the overviews in Dixit (1993) and Merton (1990).

The log price process \( X_t = \log(S_t) \) after the Itô lemma and from (16) follows the diffusion model

\[
dX_t = (\mu_t + \sigma_t^2/2)dt + \sigma_t dW_t, \tag{4.5}
\]

where the drift for \( X_t \) has a term \( \sigma_t^2 \). GARCH models are used to represent statistically the increments of the log price process, so from the diffusion point of view, (17) is also a natural parametrization of the GARCH drift \( \mu_k \).

While the models are written in continuous–time, the available data are mostly sampled discretely in time. Ignoring this difference can result in inconsistent estimators (see, e.g., Merton (1980)). A number of statistical/econometric methods have been recently developed to estimate the parameters of a
An Overview of Probabilistic and Time Series Models in Finance

The methods of moments together with simulation estimations have been used by Gouriéroux et al. (1993) and Gallant and Tauchen (1996). A forceful criticism of simulation-based method-of-moments estimation has been that this method does not provide a representation of the observables in terms of their own past as do maximum likelihood based on a conditional density and time series methods such as ARIMA, ARCH and GARCH modeling; see Jacquier et al. (1994). Gallant and Tauchen (1998) use the notion of reprojection to let a representation of the observed process in terms of observables that incorporates the dynamics implied by the possibly nonlinear system under consideration. They propose a methodology for estimation and diagnostic assessment of several diffusion models of the short rate expressed as a partially observed system of stochastic differential equations. The theoretical support of the projection method was provided by Gallant and Long (1997) who showed that it achieves the same efficiency as ML.

Nonparametric density-matching methods have been applied in Aït-Sahalia (1996a, 1996b). Discretely observed diffusions have also been fit by estimating functions; see Kessler and Sørensen (1999) and Kessler (2000). A Monte Carlo Markov Chain (MCMC) based method is proposed in Eraker (2001). The method is applied to the estimation of parameters in one-factor interest-rate models and a two-factor model with a latent stochastic volatility component.

Elerian et al. (2001) propose a new method for dealing with the estimation problem of stochastic differential equations that is likelihood based, can handle nonstationarity, and is not dependent on finding an appropriate auxiliary model. As they point out, their idea is simply to treat the values of the diffusion between any two discrete measurements as missing data and then to apply tuned MCMC methods based on the Metropolis-Hasting algorithm to learn about the missing data and the parameters.

As in most contexts, provided one trusts the parametric specification in the diffusion, ML is the method of choice. The major caveat in the present context is that the likelihood function for discrete observations generated by the parametric stochastic differential equation cannot be determined explicitly for most models. Since the transition density is generally unknown, one is forced to approximate it. The simulation-based approach suggested by Pedersen (1995), has great theoretical appeal but its implementation is computationally costly. Durham and Gallant (2002) examine a variety of numerical techniques designed to improve the performance of this approach.

If sampling of the process were continuous, the situation would be simpler. First, the likelihood function for a continuous record can be obtained by means of a classical absolutely continuous change of measure. Second, when the sam-
pling interval goes to zero, expansions of the transition function “in small time” are available in the statistical literature and some calculate expressions for the transition function in terms of functionals of a Brownian Bridge. Available methods to compute the likelihood function in the case of discrete-time sampling, involve either solving numerically the Fokker-Plank-Kolmogorov partial differential equation (see Lo (1988)) or simulating a large number of sample paths along with the process is sampled very finely (see Pedersen (1995)). Neither methods produces a closed-form expression to be maximized over the parameter: the criterion function takes either the form of an implicit solution to a partial differential equation, that could be approximated by a sum over the outcome of the simulations. Using Hermite polynomials, Aït-Sahalia (2002) provides an explicit sequence of closed-form functions. It is shown that it converges to the true (but unknown) likelihood function. It is also documented that maximizing the sequence results in an estimator that converges to the true ML estimator and shares its asymptotic properties.

As we have pointed out in section 3, high-frequency financial data are not only discretely sampled in time but the time separating successive observations is often random. Aït-Sahalia and Mykland (2003) analyzes the consequences of this dual feature of the data when estimating a continuous–time model. More precisely, they measure the additional effect of the randomness of the sampling intervals over and beyond those due to the discreteness of the data. They also examine the effect of simply ignoring the sampling randomness and find that in many situations the randomness of the sampling has larger impact than the discreteness of the data.

As we have described previously, continuous–time models, dominated by the diffusion approach, are typically favored in the theoretical finance while discrete-time models, mainly of the ARCH type, are the focus of empirical research. Nelson (1990) tried for the first time to reconcile both approaches, showing that GARCH processes weakly converge to some bivariate diffusions as the length of the discrete time interval goes to zero. Later, Duan (1997) proposed an augmented GARCH model and derived its diffusion limit. These authors link the two types of models by weak convergence. Consequently, it is rather common to apply the statistical inferences derived under the GARCH model to its diffusion limit. However, recently Wang (2002), using the Le Cam’s deficiency distance, shows that the GARCH model and its diffusion limit are asymptotically equivalent only under deterministic volatility. He concludes that, for modelling stochastic volatility, if a diffusion model is preferred, it is statistically more efficient to fit data directly to the diffusion model and carry out the inference.
2.5 Conclusions

Throughout the paper we have summarized several applications of probabilistic and time series models in finance. We have specially focused on those pricing models reflecting the absence of arbitrage and free-lunch. Almost all of them are characterized by the existence of equivalent martingale probability measures (or risk-neutral measures). Thus the martingale property permits to price, hedge, speculate or compose efficient portfolios since future prices must verify the random walk assumption.

However, there are still many open problems that will merit future research. So, the absence of arbitrage (free-lunch) does not always lead to martingales, even it one focuses on perfect markets. When dealing with incomplete markets there are infinitely many risk-neutral measures and it is necessary to establish coherent criteria in order to choose the adequate one. For imperfect markets we will never have a unique risk-neutral measure and it is also necessary to find appropriate instruments in order to relate risk-neutral measures and hedging or efficient strategies.

Most of the concrete pricing models applied in practice are characterized by stochastic differential equations reflecting the market dynamic behavior. By manipulating the stochastic equation it is possible to obtain the partial differential equation or the risk-neutral measure leading to pricing or hedging rules, as well as, to those usual topics of asset pricing theory. Time Series and Econometric Models are the key when designing these pricing models and calibrating or evaluating its empirical possibilities. Furthermore, the growing complexity of real markets, characterized by more and more connections amongst them all, higher and higher volatilities, more and more complex risks and securities, and a increasing number of investors, make it rather necessary to improve those models usually applied when dealing with pricing issues or interest-rate linked topics.

Summarizing, probabilistic and time series approaches play a crucial role in finance, and it is emphasized if one focuses on arbitrage pricing theory. Moreover, the level of development of current markets makes it essential to improve and enlarge our knowledge about all the involved fields, from theoretical foundations to empirical applications.

 References
An Overview of Probabilistic and Time Series Models in Finance


Engle, R. (2002a), New frontiers for ARCH models, forthcoming in *Journal of Applied Econometrics*


Recent Advances in Applied Probability
Baeza-Yates, R.; Glaz, J.; Gzyl, H.; Hüsler, J.; Palacios, J.L.
(Eds.)
2005, XIII, 497 p., Hardcover