STOCHASTIC VOLATILITY IN FINANCIAL MARKETS

CROSSING THE BRIDGE TO CONTINUOUS TIME

Fabio Fornari
Bank of Italy

and

Antonio Mele
Université du Littoral
and THEMA, Université de Paris X

KLUWER ACADEMIC PUBLISHERS
Boston / Dordrecht / London

1

INTRODUCTION

1.1 Background and aims of the monograph

Financial returns, albeit unpredictable according to the definition of Sims (1984), display both temporal dependency in their second order moments and heavy-peaked and tailed distributions. While such a phenomenon was known at least since the pioneer work of Mandelbrot (1963) and Fama (1965), it was only with the introduction of the autoregressive conditionally heteroscedastic (ARCH) model of Engle (1982) and Bollerslev (1986) that econometric models of changing volatility have been intensively fitted to data. ARCH models have had a prominent role in the analysis of many aspects of financial econometrics, such as the term structure of interest rates, the pricing of options, the presence of time varying risk premia in the foreign exchange market: see Bollerslev et al. (1992), Bera and Higgins (1993), Bollerslev et al. (1994) or Palm (1996) for surveys. The quintessence of the ARCH model is to make volatility dependent on the variability of past observations. An alternative formulation initiated by Taylor (1986) makes volatility be driven by unobserved components, and has come to be known as the stochastic volatility (SV) model. As for the ARCH models, SV models have also been intensively used in the last decade, especially after the progress accomplished in the corresponding estimation techniques, as illustrated in the excellent surveys of Ghysels et al. (1996) and Shephard (1996). Early contributions that aimed at relating changes in volatility of asset returns to economic intuition include Clark (1973) and Tauchen and Pitts (1983), who assumed that a stochastic process of information arrival generates a random number of intraday changes of the asset price.

Parallel to this strand of empirical research, option pricing theory has expanded into generalizations of the celebrated Black and Scholes (1973) and Merton (1973) evaluation formulae of European options. The Black-Scholes model, for instance, assumes that the price of the asset underlying the option contract follows a geometric Brownian motion, and one of the most successful extensions has been the *continuous time* SV model originally introduced by Hull and White (1987), Johnson and Shanno (1987), Scott (1987) and Wiggins (1987) (more recent related work includes Amin and Ng (1993), Duan (1995), Kallsen and Taqqu (1998) and Hobson and Rogers (1998)). In these models, volatility is not a constant, as in the original Black-Scholes model; rather,

it is another random process typically driven by a Brownian motion that is imperfectly correlated with the Brownian motion driving the primitive asset price dynamics. Similar extensions have been introduced in the term structure literature.

In this monograph, we emphasize the use of ARCH models in formulating, estimating and testing the continuous time stochastic volatility models favored in the theoretical literature. The primary source of our research agenda came from work that Daniel B. Nelson published during the first half of nineties, and that is now collected in the second part of the book edited by Rossi (1996). In the first of his celebrated papers, Nelson (1990) was able to show that although ARCH processes are casted in terms of stochastic *difference* equations, they can be thought as reasonable approximations to the solutions of stochastic *differential* equations as the sampling frequency gets higher and higher. In technical terms, the volatility process generated within ARCH-type models converges in distribution towards a well defined solution of a stochastic differential equation as the sampling frequency increases. Since SV models are typically formulated in continuous time in the theoretical literature, Nelson's contribution appeared to many to be an important step towards 'bridging the gap' between the discrete time perspective followed by the applied econometrician and the continuous time perspective idealized by the theorist.

Yet, perhaps due to the great progress accomplished in the domain of the estimation of the parameters of stochastic differential equations through simulation-based methods expanding the early work of McFadden (1989), Pakes and Pollard (1989), Ingram and Lee (1991), Duffie and Singleton (1993), Smith (1990, 1993), Gouriéroux et al. (1993), Bansal et al. (1995) and Gallant and Tauchen (1996), Nelson's ideas were not pushed far enough in the subsequent empirical and statistical literatures; see Gouriéroux and Monfort (1996) for a systematic account of simulation-based econometric methods. One concomitant reason is that the continuous record asymptotics developed for ARCH models do *not* deliver a theory for the estimation of the relevant parameters; rather, such methods typically take the parameters as given, and study the limiting behavior of stochastic difference equations in correspondence of fixed, well-chosen (perhaps too well-chosen) sequences of parameters.

The methodology introduced by Nelson, however, revealed useful to show that appropriate sequences of ARCH models are able to estimate consistently the volatility of a given continuous time stochastic process as the sample frequency gets larger and larger, even in the presence of serious misspecifications: see Nelson (1992) and Nelson and Foster (1994) for the univariate cases, and Bollerslev and Rossi (1996) (p. xiii-xvii) for a very succinct primer on the filtering performances of ARCH models as applied to continuous time stochastic volatility models. As put by Bollerslev and Rossi (1996) (p. xiv),

> "one could regard the ARCH model as merely a device which can be used to perform filtering or smoothing estimation of unobserved volatilities".

We believe that this represents one of the most important aspects of Nelson's work. In addition to the point estimates of the parameters of stochastic differential equations systems, indeed, an essential ingredient for the practical implementation of any continuous time stochastic volatility model is obviously the knowledge of the volatility at some dates of interest. If one wishes to make use of an option pricing formula that takes stochastic volatility into account, for instance, one has to know not only the price of the asset underlying the contract, but also the instantaneous volatility of that price. However, volatility is obviously not observable—as it can instead be the case of a share price—, and obtaining estimates of it in continuous time is not an easy task; see, however, the recent work of Gallant and Tauchen (1998) that is based on reprojection techniques (previous work on the filtering techniques of stochastic volatility is succinctly reminded in section 1.2.1).[1]

Figure 1.1, taken from Fornari and Mele (1999*a*), visualizes one simulated path from which one can appreciate the 'typical' filtering of an ARCH model as applied to a restricted version of a theoretical short-term interest rate model presented in chapter 5:

$$\begin{aligned} \mathrm{d}r(t) &= (\iota - \theta r(t))\mathrm{d}t + r(t)^{1/2}\sigma(t)\mathrm{d}W^{(1)}(t) \\ \mathrm{d}\sigma(t) &= (\overline{\omega} - \varphi\sigma(t))\mathrm{d}t + \psi\sigma(t)\mathrm{d}W^{(2)}(t) \end{aligned} \tag{1.1}$$

where $W^{(i)}, i = 1, 2$, are standard Brownian motions, and $\iota, \theta, \overline{\omega}, \varphi$ and ψ are real parameters whose values have been fixed at the corresponding estimates obtained with US data (see chapter 5 for some additional details, and Fornari and Mele (1999*a*) for a more technical presentation). The single trajectory reported in the figure represents a weekly sampled trajectory of $\sigma(t)$ obtained by simulating (1.1); the dotted line represents instead the trajectory of the (rescaled) volatility obtained when an ARCH model is fitted to the simulated weekly sampled trajectory of $r(t)$. The strength of such a visual, informal evidence has been formally tested in the Monte Carlo experiment conducted in

[1] We are not considering here the possibility of using a theoretical model to extract volatility and/or estimate parameters by means of cross sectional information (e.g., option and/or bond prices). If such a theoretical model had a closed-form solution, this could be an interesting device. Since continuous time stochastic volatility models typically have *not* closed form solutions, using cross sectional information for filtering volatility and estimating all the model's parameters is for the moment nearly unfeaseable, requiring an extremely fine numerical integration of partial differential equations in correspondence of each candidate of the parameter values. See, however, Fornari and Mele (1999*a* and 1999*b*) for related work on both term-structure and European option pricing issues. For cross sectional methods applied to option pricing objectives similar to our 1999*b* paper, see Chernov and Ghysels (1999).

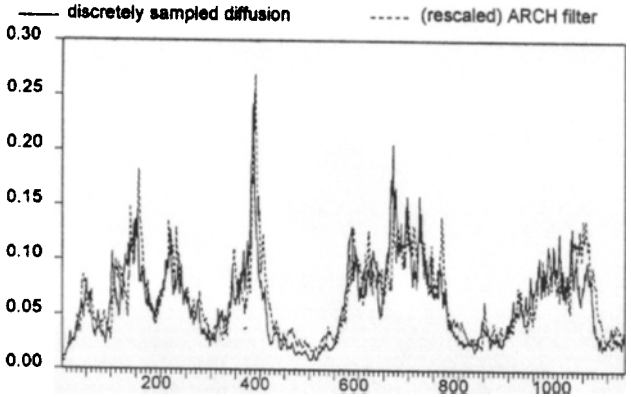

FIGURE 1.1. Typical filtering of the weekly sampled volatility diffusion $\sigma(t)$ in eqs. (1.1) by means of an ARCH model

Fornari and Mele (1999*a*), where a very low RMSE is shown to divide the two trajectories in thousands of simulations; notice that previous related work on Monte Carlo evidence concerning ARCH models as consistent volatility filters was already conducted as early as Schwartz et al. (1993).

Such results should reinforce the researcher's motivation on the use of ARCH models as approximations of diffusion processes. Yet, relatively little empirical work has been done in that direction. This point is also evidenced by Campbell et al. (1997) (p. 381), who notice that:

> "The empirical properties of [ARCH as approximators of continuous time stochastic volatility processes] have yet to be explored but will no doubt be the subject of future research."

The research presented in this monograph tries indeed to accomplish this task. We wish to outline two major steps in our research strategy. The first one consists in constructing continuous time economies displaying equilibrium dynamics to which ARCH models converge in distribution as the sample frequency gets infinite. This allows us to obtain a microeconomic foundation of the continuous time models that we take as the data generating mechanism. The utility of such an approach lies in the possibility of determining explicitly, within standard preference restrictions (e.g., CRRA utility functions), the risk-premia demanded by agents to be compensated for the fluctuations of the stochastic factors. Our primary field of application will be the theory of the term structure of interest rates with stochastic volatility, a field that is relatively less developed than the corresponding European option pricing domain

where, instead, we take a more data-oriented approach (see section 1.3 for further introductory details).

Naturally, there are other continuous time candidates than the economies that we consider here. As shown in chapters 4 and 5, however, only a minor change in notation (and in the corresponding computer codes!) is required to bring our models in touch with these alternatives. In our theoretical model of the term structure of interest rates, indeed, stochastic volatility is generated by the variability of the economic fundamentals and, as is well-known since the seminal contribution of Harrison and Kreps (1979), any arbitrage-free specification of asset prices can be sustained by a competitive equilibrium. In chapter 10 of the Duffie's (1996) textbook, for instance, the reader can see such a phenomenon at work within the standard univariate representation of the Cox et al. (1985*a*) model; as concerns stochastic volatility, we show in chapter 4 that the choice of the primitive measure space and subsequent factor restrictions crucially determine the final predictions of our equilibrium model of the term structure. To resolve for such an indeterminacy, two natural alternatives are possible. The first one consists in specifying the primitives so as to obtain a computationally (or even analytically) tractable model, as in the case of the original single factor Cox et al. (1985*a*) model. Such an idea is fully exploited by Duffie and Kan (1996), who construct a class of 'exponential-affine' models "specifying simple relationships among yields" (p. 380); see also Brown and Schaefer (1995) for an earlier treatment of related issues. Another possibility fully exploits the idea that, given the imperfections of any model, it is an empirical issue as to which primitives will serve best at generating a model of the term structure of interest rates. In this case, one can look for economies that support dynamics that are more or less consistent with past data analysis or even with informal observation. Our reverse strategy is thus justified by this second possibility.

The second step of our research program is devoted to a more concrete econometric analysis of continuous time stochastic volatility models. Precisely, the concern lies in the estimation of the parameters of the stochastic differential equations characterizing our equilibrium economies. In the estimation strategy that we suggest, one first uses the moment conditions under which ARCH models converge in distribution towards the theoretical models, obtaining a *direct*, preliminary estimate of the model's parameters. Since such estimates are obtained by means of discrete time models that are typically not closed under temporal aggregation (Drost and Werker (1996)), one then tests and corrects possible *disaggregation* biases by using, this time, ARCH models as *auxiliary* devices in simulation-based schemes. Fornari and Mele (1999*a*, 1999*b*) have already applied such a strategy to option prices and US interest rate data, finding strong evidence that the correction made by indirect inference is not significant. In addition to being an appropriate tool to filter out stochastic volatility, such results also make a strong case for the use of ARCH models

as approximating devices of the parameters of certain stochastic differential equations.

A related topic lies between the two above mentioned steps of our work: it consists of a better understanding of the functioning of ARCH models by resorting to the more easily tractable continuous time approach. This objective arises quite naturally, since ARCH models are intensively used (not only in mathematical finance as in this monograph), which would require a clear understanding of their theoretical characteristics. In this respect, too, the continuous time approach of Nelson (1990) offers an appropriate tool of analysis. Indeed, ARCH models are non linear stochastic difference equations, and some of their properties are quite cumbersome to establish. The task is easier when one examines their behavior as the sampling frequency tends to infinity. In this case, ARCH models are 'approximated' by stochastic differential equations, a relatively easier to study object: in other terms, one would use *diffusions as ARCH approximations* to get insights into the functioning of ARCH models. As an example, in his first contribution, Nelson (1990) showed that the GARCH(1,1) model of Bollerslev (1986) (see eq. (1.4) below) converges in distribution towards the following stochastic differential equation:

$$\mathrm{d}\sigma(t)^2 = (\overline{\omega} - \varphi\sigma(t)^2)\mathrm{d}t + \psi\sigma(t)^2\mathrm{d}W^\sigma(t), \tag{1.2}$$

where W^σ is a standard scalar Brownian motion, and $\overline{\omega}$, φ, and ψ are real valued, non stochastic parameters. Now such a result implies that in continuous time, (1) σ^2 follows a stationary distribution that is an inverted Gamma; and (2) the error process from a given observation model is (approximately) unconditionally Student's t distributed, even if it is conditionally normal. Such results would not be obtained in a discrete time setting.

Our own contribution in this field consisted in deepening some of the previous findings. Precisely, we showed (Fornari and Mele (1997*a*)) that the limiting results obtained for the GARCH(1,1) hold as well for a fairly general model previously introduced by Ding et al. (1993) (see eq. (1.6) below), and for the model of Fornari and Mele (1997*b*). The first scheme admits a 'diffusion limit' that has the following form:

$$\mathrm{d}\sigma(t)^\delta = (\overline{\omega} - \varphi\sigma(t)^\delta)\mathrm{d}t + \psi\sigma(t)^\delta\mathrm{d}W^\sigma(t),\ \delta \in \mathbb{R}_{++} \tag{1.3}$$

and generalizes both the volatility equation in (1.1) and eq. (1.2). In addition to provide a flexible specification that can be useful in applied work related to mathematical finance—notably via the introduction of a sort of 'volatility concept' σ^δ—, eq. (1.3) enabled us to carry out a detailed analysis concerning the distribution of errors terms from the corresponding (discrete time) observation model: see section 1.2.4 for more details. On the other hand, the derivation of a diffusion limit in correspondence of our 1997*b* model was instructive to us, since despite the nonlinearities that it was designed to capture (see section 1.2.3 below for details), the kind of conditions required to obtain the convergence result were of the same essence as those of Nelson.

The remainder of this chapter has been designed to be a more systematic introduction to all these research themes to which we gave a contribution.

Next section provides a very selective overview of ARCH models. Given the many existing surveys on ARCH models, we only constrain ourselves to present those aspects of ARCH models that are connected with our own contribution: in section 1.2.1, we present the very basic ARCH models as well as their (discrete time) SV competitors; in section 1.2.2, we present the extensions of these models that are designed to capture some observed non linear dynamics of volatility, notably the fact that volatility tends to react asymmetrically to past shocks of different sign; in section 1.2.3 we deepen the previous issue, by arguing that volatility asymmetries are subject to reversals, i.e. positive shocks may sometimes induce more volatility than negative shocks; section 1.2.4 provides some details and motivation concerning our contribution to the approximation results for ARCH models. Section 1.3 covers some of the most important economic implications of continuous time stochastic volatility: section 1.3.1 presents the central issue of the problem in the option pricing domain, and insists on market incompleteness, the related impossibility to implement pure hedging strategies, and the existence of partial hedging strategies that could be optimal according to certain criteria; section 1.3.2 describes some of the existing models of the term structure of interest rates with stochastic volatility. Section 1.4 is an introduction to the statistical methodology that we use to conduct inference in our continuous time models. Finally, section 1.5 provides information about the plan of the remainder of the monograph.

1.2 Empirical models in discrete time

1.2.1 The basic models

The original ARCH model posits the existence of a relation between past squared innovations of an observation assets returns changes model and their current conditional variances. Let $\{\varepsilon_t\}_{t=1}^\infty$ be the error process of some observation model; then, the GARCH(1,1) model assumes that $\{\varepsilon_t\}_{t=1}^\infty$ is conditionally normal with variance changing through time in a fashion which resembles a restricted ARMA process, i.e.:

$$\begin{aligned} \varepsilon_t \ / \ I_{t-1} &\sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \end{aligned} \tag{1.4}$$

where $\omega > 0$, $\alpha, \beta \geq 0$ are real, non-stochastic parameters and I_t is the information set dated t. In the original ARCH(p) model of Engle (1982), β was zero, and the volatility σ_t^2 was driven by a linear combination of p lagged squared

error terms: $\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$, $\alpha_i \geq 0$. The model can be estimated by standard maximum likelihood (ML) techniques. Very succinct presentations of the properties of such estimators are in the surveys cited at the beginning of the introductory section of this chapter.

A natural alternative to model (1.4) is given by the SV model, in which log-volatility typically follows an AR process. In the very first contributions, the corresponding estimation techniques relied typically on a simple method of moments (e.g., Scott (1987) and Wiggins (1987)) and the generalized method of moments of Hansen (1982) (e.g., Chesney and Scott (1989) and Melino and Turnbull (1990)). An alternative estimation technique is based on the Kalman filter (see Scott (1987) and Nelson (1988) for an early treatment of such issues): let r_t denote an asset return as of time t, and rewrite the observation equation

$$r_t = \sigma_t \cdot u_t, \; u_t \approx N(0,1)$$

as

$$\log r_t^2 = \log \sigma_t^2 + \xi_t, \; \xi_t \equiv \log u_t^2.$$

By postulating that the volatility propagation mechanism is

$$\log \sigma_t^2 = a + b \cdot \log \sigma_{t-1}^2 + \zeta_t, \; \zeta_t \approx N(0, \sigma_\zeta^2),$$

one sees that $\log r_t^2$ is written in a state space form; hence, the model can be estimated with the usual Kalman filtering techniques (see Harvey (1989) for a textbook on such techniques applied to economics). Due to non-normality of ξ, however, the likelihood function that results by the prediction error decomposition of the Kalman filter is not the exact one, but one can invoke the usual quasi-likelihood methods, as in Harvey et al. (1994). Jacquier et al. (1994) derive the exact filter, and extensions can be found in Jacquier et al. (1999). A comprehensive survey on related techniques can be found in Shephard (1996); notice that an important aspect of these methods consists in extracting volatility estimates, which is not the case for the method of moments. Furthermore, many of the preceding techniques were designed to estimate models that usually are naive discretization of the corresponding continuous time models favored in the theoretical literature; hence, they are likely to induce a discretization bias; see, for instance, Melino (1994) for one of the earliest discussion of this problem.

An alternative method to estimate a discrete time SV model can be based on the indirect inference methods that are succinctly described in section 1.4.

1.2.2 Extensions

A shortcoming of the GARCH model is that the sign of the forecast errors does not influence the conditional variance, which may contradict the observed dynamics of assets returns. Black (1976), for example, noted that volatility

tends to grow in reaction to bad news (excess returns lower than expected), and to fall in response to good news (excess returns higher than expected). The economic explanation given by Black is that negative (positive) excess returns make the equity value decrease and the leverage ratio (defined as debt/equity) of a given firm increase (fall), thus raising (lowering) its riskiness and the future volatility of its assets. This phenomenon has consequently come to be referred to as *leverage effect* (Pagan and Schwert (1990), Nelson (1991), Campbell and Hentschel (1992)). It has to be said, however, that the negative correlation between asset returns and volatility seems to be too strong to be explained on the basis of the leverage effect only; see Christie (1982) and Schwert (1989*a*).

The basic attempts to include such features into a convenient econometric framework are the Exponential GARCH (EGARCH) model of Nelson (1991), the Glosten et al. (1993) (GJR) model, the asymmetric power ARCH model of Ding et al. (1993), the threshold ARCH model of Rabemananjara and Zakoïan (1993) and Zakoïan (1994), the Quadratic ARCH of Sentana (1991), or the SV model of Harvey and Shephard (1993*a*, 1993*b*) and Harvey et al. (1994). All such models include the sign of past forecast errors as conditioning information for the current values of the conditional variance.

In the EGARCH(1,1), for instance, the following equation generates σ_t:

$$\log \sigma_t^2 = \omega + \gamma_0 u_{t-1} + \gamma_1 (|u_{t-1}| - E\,|u_{t-1}|) + \beta \log \sigma_{t-1}^2,$$

where $u_t = (\frac{\varepsilon}{\sigma})_t$ and the asymmetric behavior of the log-variance with respect to changes in past errors is captured by the terms multiplying γ_0, γ_1. Notice further that such a formulation also allows one to relax positivity constraints on the other parameters ω, β.

In the GJR model, to cite another example, the following is the volatility generating process:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma s_{t-1} \varepsilon_{t-1}^2, \tag{1.5}$$

with $\omega > 0$, $\alpha, \beta \geq 0$, and s_t is a dummy variable which equals one when ε_t is negative and is nil elsewhere. Here the asymmetry is captured by the term multiplying γ. When γ is negative, negative shocks ($\varepsilon < 0$) introduce more volatility than positive shocks of the same size in the subsequent period.

A fairly general model has been proposed by Ding et al. (1993). In this paper (see, also, Granger and Ding (1993, 1994)), the authors study US stock daily returns and show that the autocorrelation function (a.c.f.) of absolute returns raised to a positive power, say θ, is significantly different from zero and, further, is not the strongest one for $\theta = 2$ over a considerably wide range of lags.[2] Specifically, there are values of θ close to one which make such an a.c.f. the

[2] Granger and Ding (1993) attempted to link these findings to previous theoretical results of Luce (1980). Let x_t denote the series of returns on some speculative assets; let λ be a positive number. Then, Luce's results imply that the measure of risk per-

largest one for all the considered lags. Based on Monte Carlo evidence, Ding et al. (1993) also report that both the GARCH(1,1) of Bollerslev (1986) and the Taylor (1986) and Schwert's (1989*b*) model (TS henceforth) are able to reproduce the property that the a.c.f. of $|x_t|^\theta$ reaches its maximum in correspondence of $\theta \simeq 1$ for quite a high number of lags, despite the fact that the GARCH(1,1) makes the conditional variance a linear function of past squared innovations, and the TS model sets the conditional variance equal to the square of a linear function of past absolute innovations (see the third line in table 1). This motivated Ding et al. (1993) to propose the so-called Asymmetric Power ARCH (A-PARCH) model:

$$\sigma_t^\delta = \omega + \alpha(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta\sigma_{t-1}^\delta, \ \gamma \in (-1,1), \ (\omega,\alpha,\beta,\delta) \in \mathbf{R}_+^4. \quad (1.6)$$

The main difference between (1.6) and standard ARCH equations is the power δ to which σ and $(|\varepsilon| - \gamma\varepsilon)$ are raised. In standard applications, δ is 2 or 1, while the A-PARCH imposes δ as a Box-Cox power transform on σ, which has to be estimated. Ding et al. (1993) show that model (1.6) encompasses at least seven classes of ARCH models already proposed; see table 1.1 for some of these models. Naturally, the leverage effect is also captured within this model (hence its definition as 'asymmetric'), notably throughout the term multiplying γ.

TABLE 1.1

GARCH(1,1)	eq. (1.4)
GJR	eq. (1.5)
Taylor-Schwert	$\sigma_t = \omega + \alpha\,|\varepsilon_{t-1}| + \beta\sigma_{t-1}$
Nonlinear ARCH	$\sigma_t^\delta = \omega + \alpha\,|\varepsilon_{t-1}|^\delta + \beta\sigma_{t-1}^\delta$

One disadvantage of all the previous formulations is that volatility asymmetries are nested into the ARCH scheme in an abrupt manner. In the GJR model (1.5), for instance, asymmetries in volatility are introduced by adding a term to the standard GARCH(1,1) that is the product of the squared residual times a step function of the residual: denote such a step function as $\underline{\alpha}(\varepsilon_t) \equiv \gamma \cdot s_t$. The dotted line in figure 1.2 represents $\underline{\alpha}$ when $\gamma < 0$. In a private conversation held with Timo Teräsvirta in January 1995, we suggested a natural generalization that replaces $\gamma \cdot s_t$ with a function that changes smoothly in the domain of definition of ε. The straight line in figure 1.2 depicts one possible $\underline{\alpha}$; generally, a creative use of standard techniques from the smooth transition

ceived by agents can be proportional to $E_t\,|x_{t+1}|^\lambda$ where E_t denotes the expectation operator conditional on the information set as of time t. Under market consensus on λ, and under the assumptions that (1) x_t distributes normally with zero mean and variance σ_t^2 and (2) $E_t\sigma_{t+1}^\lambda$ is the appropriate measure of volatility, then, standard volatility measures may fail to capture an adequate dynamic characterization of risk, since λ is not necessarily 2; rather, it can take any positive value.

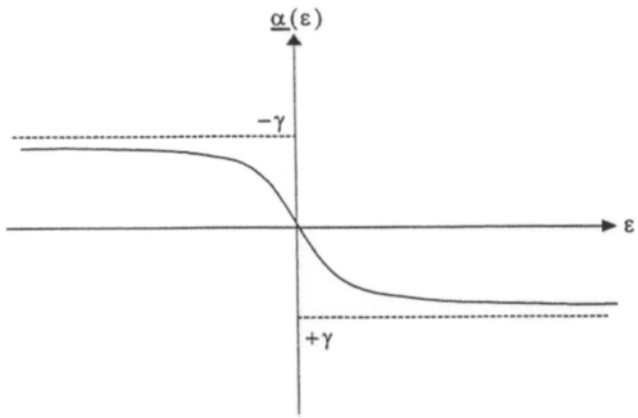

FIGURE 1.2. Smoothing volatility asymmetries ($\gamma < 0$)

modeling literature (see, e.g., Granger and Teräsvirta (1993)) can be employed to select an appropriate function $\underline{\alpha}$. To the best of our knowledge, Hagerud (1997) and Gonzalez-Rivera (1998) were the first papers that exploited such ideas by opening the route to the so-called 'smooth-transition' ARCH models.

1.2.3 VOLATILITY ASYMMETRIES: RAMIFICATIONS

According to empirical evidence originally reported in Rabemananjara and Zakoïan (1993), even the preceding asymmetric-type models might be unsuccessful in taking into account some nonlinearities of the volatility dynamics. In a study concerning disaggregated French returns, Rabemananjara and Zakoïan find that 'high' negative shocks increase future volatility more than high positive ones while at the same time 'small' positive shocks too often produce a stronger impact on future volatility than negative shocks of the same size. Thus, following the occurrence of a shock of a certain size, the asymmetric behavior of the volatility might become reversed; the modeling of such feature has also been the focus of two papers of ours (Fornari and Mele (1996, 1997*b*)).

Our first concern in Fornari and Mele (1997*b*) was to define the 'size' of the shock at which a volatility reversal may occur. The definition we adopted was based on the level of unexpected volatility generated by a shock at time $t-1$ (ε_{t-1}). Consider the information set dated $t-2$. The expected value of ε_{t-1}^2 is obviously σ_{t-1}^2; if, however, $\varepsilon_{t-1}^2 > \sigma_{t-1}^2$ ($\leq \sigma_{t-1}^2$), we said that ε_{t-1} generates (at time $t-1$) a level of volatility higher (lower) than expected (at time $t-2$). Consider now a very small negative shock at time $t-1$; if it produces a level of volatility at time $t-1$ lower than expected at time $t-2$, there should be no reason to believe that volatility at time t will increase as a consequence of

the leverage effect. Roughly speaking, a small negative shock which generates lower volatility than expected may be regarded as good news; at the same time, positive shocks which generate lower volatility than expected may be regarded as relatively bad news. This might be a possible explanation of the mechanism according to which a reversal originates. Furthermore, it gives a natural explanation for what has to be regarded as a 'high' or a 'small' shock: a high shock ε_t is the one for which $\varepsilon_t^2 > \sigma_t^2$ and viceversa.

In Fornari and Mele (1997*b*), we also presented an informal economic argument justifying the occurrence of volatility reversals. In particular, we noticed that the Black's (1976) arguments based on changes in the leverage ratio could be deepened. The observation was that a change in the leverage ratio is likely to be followed by a change in the expected performance of the firm, the latter being a function of the differential between the expected average performance of the sector in which the firm operates and the overall cost of debt. Suppose that the economy has K productive sectors; one has that in the kth sector, $k = 1, ..., K$,

$$i_j = \rho_k + (\rho_k - r)\frac{D}{S_j}, \tag{1.7}$$

where i_j and S_j are the expected profitability and the price of the stock, r the interest paid on debt, ρ_k the average performance of the kth sector, D the amount of debt, and $\frac{D}{S_j}$ represents the leverage ratio of the jth firm in the kth sector. Such a relation can be found in Modigliani and Miller (1958) (proposition II, p. 271). Consider now the case that $(\rho_k - r)$ is positive in eq. (1.7). Then, a negative shock on S_j may be regarded as more favorable than a positive shock; in fact it increases $\frac{D}{S_j}$ and the expected profitability of the firm: contrary to the celebrated Black's explanation, a positive shock may have a stronger impact on future volatility than a negative one of the same size. The Black's explanation, however, should be expected to be at work when the negative shock is very large. In this case, indeed, two things may be hypothesized to happen: first, economic agents may discount a recession of the kth sector, i.e. a fall of ρ_k; second, the cost of the debt might be thought to start rising sharply for the jth firm, which happens when r is positively related to $\frac{D}{S_j}$ (hence inversely related to S_j). Both events are likely to change the sign of $(\rho_k - r)$, hence causing volatility reversals.

Past empirical research had generally overlooked the impact of previous (unexpected or expected) volatility on its current expected level. Engle and Ng (1993), for example, propose to analyze the impact of news on the current conditional variance (i.e. on σ_t^2), keeping constant the information dated $t-2$ and earlier, with all the lagged conditional variances evaluated at their unconditional value. An important point in our paper was to define a sort of *response function of the future expected volatility to past unexpected volatility*. Let, for example, $v \equiv \delta_0(\varepsilon^2 - \sigma^2)$, where δ_0 is a real constant. A simple model that

takes into account the preceding remarks is:

$$\sigma_t^2 = w + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + s_{t-1}v_{t-1}. \tag{1.8}$$

Let us suppose indeed that $\delta_0 < 0$. If $v_{t-1} < 0$ (i.e., 'high' shocks have occurred) then, *cæteris paribus*, negative shocks generate more volatility than positive ones. However, if $v_{t-1} > 0$ (i.e., 'small' shocks have occurred) positive shocks increase volatility more than negative ones. Thus, model (1.8) can take into account situations where the asymmetric behavior of the volatility is reversed.

Unfortunately, model (1.8) is too simple. Furthermore, it imposes a fixed probability of occurrence of reversals that is approximately 0.68. To clarify this point, let χ_1^2 denote a chi-squared random variable, and notice that when $\delta_0 < 0$,

$$\Pr(\text{reversal at } t) = \Pr(\varepsilon_{t-1}^2 < \sigma_{t-1}^2) = \Pr(\chi_1^2 < 1) \simeq 0.68,$$

whereas we would like to be endowed with a more flexible specification. Fornari and Mele (1997*b*) took

$$\sigma_t^2 = w + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + s_{t-1}v_{t-1}, \tag{1.9}$$

where now

$$v_t \equiv \delta_0\varepsilon_t^2 - \delta_1\sigma_t^2 - \delta_2,$$

so that v_t is a constant plus a linear combination of the 'observed volatility' (ε_t^2) and its lagged expectation (σ_t^2), thus playing a role similar to that of an error correcting variable. Fornari and Mele (1997*b*) refer to model (1.9) as the Volatility-Switching (VS-) ARCH model.

Model (1.9) does not impose any a priori probability to the occurrence of reversals. Proceeding indeed with the same algebra as before, and assuming that $\delta_0 < 0$ (which was the case of our estimates),

$$\Pr(\text{reversal at } t) = \Pr(\varepsilon_{t-1}^2 < k_1\sigma_{t-1}^2 + k_2) = \Pr(\chi_1^2 < k_1 + k_2\sigma_t^{-2})$$

where $k_1 \equiv \frac{\delta_1}{\delta_0}$ and $k_2 \equiv \frac{\delta_2}{\delta_0}$. When $\delta_2 < 0$, the probability of reversals is an increasing function of the conditional precision process. This is in accordance with the explanation given above on the basis of relation (1.7): volatility reversals can occur when volatility is low. If, on the contrary, the estimates of (δ_1, δ_2) are such that $(-k_1, -k_2) \in \mathbb{R}_+^2$, the probability of reversal is nil; such a situation arises, for instance, once that $\delta_1 = \delta_2 = 0$, which is the restriction under which the VS-ARCH reduces to the GJR scheme. Thus, the VS-ARCH model allows a more detailed analysis of the asymmetric behavior of the volatility than the GJR model, enabling shifts in its direction, according to the size of past shocks. Furthermore, the existence of reversals can be

partly tested by ascertaining whether the VS-ARCH improves upon GJR.[3] Fornari and Mele (1997*b*) were looking for international evidence by fitting the VS-ARCH model to returns from seven stock exchanges. They found that on average the VS-ARCH model outperformed the GJR, and provided evidence of volatility reversals in almost all of the analyzed countries.[4]

Finally, it is worth noticing that a variant of VS-ARCH model has also been estimated in Fornari and Mele (1996); it takes the following form:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma s_{t-1}\varepsilon_{t-1}^2 + \delta(\frac{\varepsilon_{t-1}^2}{\sigma_{t-1}^2} - k)s_{t-1}. \tag{1.10}$$

The first three terms on the RHS are a standard GARCH(1,1) model. These, together with the fourth term, define a GJR(1,1) model. The last term is designed to capture the reversal of asymmetry that can be observed when $(\frac{\varepsilon}{\sigma})^2$ reaches k, the threshold value. Estimation results from this model were perfectly in line with those subsequently published in Fornari and Mele (1997*b*).

1.2.4 Diffusions as ARCH approximations

In addition to be useful devices in uncovering the filtering performances of ARCH and, as it will be described in section 1.4, in enriching the estimation techniques of continuous time stochastic volatility models, continuous time methods can also be understood as a guidance to get interesting insights into some properties of ARCH models in discrete time. In the introductory section of this chapter, we mentioned that in his seminal paper Nelson (1990) showed that the GARCH(1,1) error process follows (approximately) a Student's t, unconditionally. One of the tasks that we gave ourselves in our 1997*a* paper was to obtain results in correspondence of the GARCH(1,1) error process when the error process is not conditionally normal; also, we investigated what is the form of such a distribution when innovations are conditionally normal and the volatility propagation mechanism is eq. (1.6). More generally, we were looking for the stationary distribution of the error process when the error process is not conditionally normal and the volatility propagation process is eq. (1.6). Similar issues were dealt with in our 1997*b* paper.

In chapter 2 we attempt to present a scheme that explores these issues; we will draw heavily on the theoretical sections of our corresponding 1997 papers. In our scheme, the error process is taken to be conditionally *general error distributed* (g.e.d.). Such a choice is useful on an analytical standpoint,

[3]It is worth signalling that Engle and Ng (1993) fitted different asymmetric models (not the VS-ARCH) to stock returns from Japan and found the GJR to be the best parametric scheme to model asymmetries.

[4]In our original paper, the models were estimated via a normal likelihood profile. It is possible to show that our results are robust to a generalization of such a distributional assumption.

while allowing at the same time for a fairly instructive first approximation treatment of non-normality issues. One reason to deal with non-normal errors came from well-known empirical findings of the mid eighties: the standardized residuals obtained by using the conditional standard deviation of an ARCH-type model is typically not normal, despite the use of conditionally normal likelihood profiles.[5]

Nelson (1991) was the first author to make use of the g.e.d. in the attempt to fit stock returns; as is well-known, the g.e.d. is characterized by a 'tail-thickness' parameter, denoted hereafter as υ, that tunes the height of the tails; throughout this monograph, we will use the notation g.e.d.$_{(\upsilon)}$ to highlight the presence of such a parameter: see eqs. (2.1) for the corresponding analytical details. The g.e.d. encompasses a number of distributions such as the normal distribution or the Laplace distribution of the first kind; the Laplace distribution of the first kind was used by Granger and Ding (1993) to fit different ARCH models to stock returns data. It should be acknowledged, however, that Nelson (1991) reported that the g.e.d. did not fully account for all of the outliers using an EGARCH scheme for US stock market data. Alternatives to the g.e.d. are the Student's t used, for instance, in Engle and Bollerslev (1986), Bollerslev (1987), or Hsieh (1988) to model foreign exchange rates, or the generalized Student's t, used in Bollerslev et al. (1994) to model stock returns. The generalized Student's t distribution nests both the Student's t and the g.e.d., and is attractive since it has two shape parameters that take account of both the tails and the central part of the conditional distribution of the error process. Yet, Bollerslev et al. (1994) showed that even the likelihood function based on the generalized Student's t does not allow for a fully satisfactory treatment of tail events in US stock returns.

Our theoretical results can be summarized as follows. Under an appropriate discretization scheme, we first show that the solution of model (1.6) converges in distribution towards the solution of eq. (1.3). In a second step, we show that when the sampling frequency gets infinite, a conditionally g.e.d.$_{(\upsilon)}$ A-PARCH error process follows a stationary distribution that is a generalized Student's t distribution when $\delta = \upsilon$; the restriction $\delta = \upsilon$ was made for analytical convenience only, but the empirical section of Fornari and Mele (1997*a*) and Mele (1998) reported that it can hardly be rejected empirically. Obviously, the most interesting theoretical case is the general one in which $\delta \neq \upsilon$: while we do not have closed-form solutions in this case, numerical results suggest that *cæteris paribus*, low values of δ tend to: (1) raise the central part, and (2) increase the tails of the stationary distribution of the error process. Insofar as υ is considered, we find that it shapes such a stationary distribution in the same

[5]The failure of conditional normality was documented by Bollerslev (1987) or Hsieh (1988) in exchange rate studies, and by Bollerslev et al. (1994) in stock returns enquiries.

manner as it does with the conditional one. Furthermore, chapter 2 also derives a closed-form solution for the instantaneous correlation between a continuous time asset price process and its instantaneous volatility, as approximated by the A-PARCH.[6] We find that the correlation is constant, and that its modulus never reaches unity for a reasonably wide set of parameters' values.[7] Finally, chapter 2 present analogous results for the VS-ARCH eq. (1.9), although we do not deal there with correlation issues, and we confine ourselves to the simplest situation where the error process is conditionally normally distributed.

1.3 Theoretical issues

1.3.1 European option pricing

Among the standard absence-of-frictions hypotheses underlying the celebrated Black and Scholes (1973) formula[8] for the price of European-type option contracts, an important assumption was the complete market structure that can be generated by assuming that the price of the asset underlying the option contract is a geometric Brownian motion:

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W(t), \tag{1.11}$$

where μ and σ are two constants. Due to what Heston (1993*a*) (p. 933) figuratively terms "a surprising cancellation", the constant μ vanishes out from the final formula,[9] which is obviously not the case for the constant σ.

In addition to be inconsistent with time-series evidence, the constancy of σ also gives rises to empirical cross-sectional inconsistencies: when one compares the Black-Scholes formula with observed option prices and then inverts the formula to recover the σ in (1.11), one obtains that the resulting σ is in fact a U-shaped function of the strike of the option. Such a phenomenon has come to be known as the 'smile effect'.

To the best of our knowledge, Ball and Roma (1994) (p. 602) (see, also, Renault and Touzi (1996) for related work) were the first to point out that the smile effect emerges in a rather natural fashion when the data generating

[6]Such a result did not appear in the original 1997*a* paper, but in the 1999*a* one.

[7]In an independent work, Duan (1997) elegantly accomplishes similar tasks for models encompassing subclasses of the A-PARCH, but here we account for all the encompassed models, as well as more general distributional assumptions of the discrete model.

[8]See formula (3.16) in chapter 3.

[9]Heston (1993*a*) shows that the independence of the option pricing formula on the average appreciation rate of the underlying asset price is not a general property in correspondence of alternative asset price processes.

mechanism is a stochastic volatility one,

$$\begin{aligned} \mathrm{d}S(t) &= \mu\mathrm{d}t + \sigma(t)S(t)\mathrm{d}W^{(1)}(t) \\ \mathrm{d}\nu(t) &= \overline{\varphi}(t)\mathrm{d}t + \overline{\psi}(t)\mathrm{d}W^{\sigma}(t) \end{aligned} \tag{1.12}$$

and yet one insists on inverting the Black-Scholes formula to recover a constant volatility. In system (1.12), ν is a monotone function of σ; the functions $\overline{\varphi}, \overline{\psi}$ are taken to be measurable with respect to the information set generated by the Brownian motions $W^{(1)}, W^{\sigma}$, and guarantee the existence of a solution to (1.12) (see chapter 3 for a more technical presentation on existence issues). The assumptions of the authors were essentially the ones underlying the Hull and White (1987) model; Renault (1997) presents a survey on the state-of-the-art on this issue, and Das and Sundaran (1999) develop further analytical results concerning the case in which W^{σ} can be written as $W^{\sigma} = \rho W^{(1)} + (1-\rho^2)^{1/2}W^{(2)}$, where ρ is a constant in $(-1,1)$, and $W^{(2)}$ is another Brownian motion, finding however that standard SV models are not entirely satisfactory.

The Black-Scholes setting was of course modified to take account of stochastic volatility, as in (1.12); Hull and White (1987), Scott (1987), or Wiggins (1987) were early attempts in that direction. Explicit solutions have proved hard to derive: if one excludes the approximate solution provided by Hull and White (1987) or the analytical solution provided by Heston (1993*b*),[10] one typically needs to derive the price of the call by implementing Monte Carlo methods (e.g., Johnson and Shanno (1987), Engle and Mustafa (1992) or Lamoureux and Lastrapes (1993)), or by numerically solving a certain partial differential equation, as in the seminal papers of Wiggins (1987):[11] see the following chapters.

It became rapidly clear, however, that the very problem associated with stochastic volatility was not the mere computational effort needed to generalize the Black-Scholes option pricing formula. Rather, a conceptual problem

[10]Such a solution was provided on the assumption that stochastic volatility was the solution of a linear mean-reverting "square-root" process; in a square root process, the instantaneous variance of the process is proportional to the level reached by that process: in model (1.12), for instance, $\overline{\psi}(t) = \psi\nu(t)^{1/2}$, where ψ is a constant. In fact, by delving into the computation details of Heston (1993*b*), one realizes that the role played by the "square-root hypothesis" in obtaining a closed-form solution resembles very much the role played by "variance-affinity" in the affine term-structure literature (see chapter 4 for details concerning the term structure of interest rates with stochastic volatility).

[11]Given the increase in the computational speed of modern computers, the numerical integration of a partial differential equation is not a real limitation; rather, it would suffice to slightly change a computer code to compare rather different stochastic volatility models: see chapter 5 for an illustration of this fact to our class of term-structure model with stochastic volatility that is developed in chapter 4. By contrast, the "human" computational time needed to find closed-form solutions in correspondence of realistic models can sometimes increase in an unreasonable way!

emerged, which is the fact that the presence of stochastic volatility generates market incompleteness. Heuristically, market incompleteness means that agents cannot hedge against future contingencies; one of the formal reasons here is that the number of assets is too low to span the entire space of contingencies. In an economy with diffusion state variables, for instance, incompleteness arises when the number of assets is less than the dimension of the Brownian motion driving the state variables: this is exactly the case of model (1.12), where trading strategies involving only one asset are not sufficient to duplicate the value of the option. In other terms, when the option price H is *rationally* formed at time t, it will be of the form $H(t) = H(t, S(t), \nu(t))$, where $\nu(t)$ is not adapted to the filtration generated by $W^{(1)}(t)$: therefore, trading with only the primitive asset does not allow for a perfect replication of H, which is the argument required to obtain one, and only one, arbitrage-free option price: see chapter 3 for a technical presentation.

To summarize, the presence of stochastic volatility introduces two inextricable consequences: (1) perfect hedging strategies are impossible; (2) there is an infinity of option prices that are admissible with the absence of arbitrage opportunities.

One way to treat the second issue would consist in making a representative agent 'select' the appropriate pricing function. This was the case of the Wiggins (1987) and Heston (1993*b*) models, for instance, who use the representative agent framework of Cox et al. (1985*b*). Furthermore, the use of a representative agent is justified on a solid microeconomic basis, since in these models the representative agent typically can trade with two assets (the primitive asset and the option contract), thus having access to a complete market structure.[12] The price to be paid to select a pricing function via preference restrictions of a representative agent, however, is that the resulting price is obviously not preference-free. In a sense, Hull and White (1987) also followed such a selection-by-preferences-restrictions strategy: in their framework, indeed, the authors supposed that agents are not compensated for the fluctuations of stochastic volatility i.e, the volatility risk premium is nil:[13] such an assumption is not confirmed empirically (see, e.g., Lamoureux and Lastrapes (1993)), and has been formally shown to be equivalent to an economy with a representative agent with logarithmic-type utility function (Pham and Touzi (1996)), a fact that was roughly known since Wiggins (1987). In Fornari and Mele (1999*b*), we take a data-oriented approach, by estimating what we termed a *volatility risk-premium surface*; see the end of section 1.3.2 for further introductory

[12]Bajeux and Rochet (1996) derived the conditions under which market completeness is ensured in a stochastic volatility framework. Romano and Touzi (1998) extended that framework by allowing, inter alia, for the existence of a correlation process between the primitive asset price and its instantaneous volatility.

[13]As a Hull and White pointed out (p. 283, footnote 1), a change in notation is sufficient to switch their model to a model with a constant volatility risk premium.

discussion, and chapter 3 for technical details.

A related utility-based approach has been developed by Davis (1997);[14] Davis' results were not designed directly for the stochastic volatility framework. Davis proposed to select a pricing function by a marginal rate of substitution argument: an agent has initial wealth equal to x, from which she generates final wealth equal to $V^{x,\pi}(T)$ by means of some portfolio process π; her problem is $\max_\pi E(u(V^{x,\pi}(T)))$. Now suppose that our agent diverts a small amount of her initial wealth x to buy an unhedgeable claim $\widetilde{X}$ that costs p; define $Q(d,p,x) = \sup_\pi E(u(V^{x-d,\pi}(T) + \frac{d}{p}\widetilde{X}))$. Davis defines a 'fair' price as the solution $\widehat{p}(x)$ of the following first order condition: $\frac{\partial}{\partial d}Q(d,\widehat{p},x)\big|_{d=0} = 0$.

An useful complement to utility-based approaches can be made by concentrating on hedging strategies: indeed, despite the fact that there are no *perfect* hedging strategies in incomplete markets, one can always look for some sort of 'imperfect', but *optimal* strategies. The notion of optimality is then generated by some loss function. As an example, if one considers strategies that simply make the volatility of the resulting strategy value the best approximation (in projection terms) of the volatility of an unhedgeable claim value, one gets that the following imperfect, pure hedging strategy:

$$\widetilde{\pi}(t) = (\frac{\partial H}{\partial S}S + \frac{\partial H}{\partial \nu}\cdot\overline{\psi}\cdot\rho\cdot\sigma^{-1})(t), \tag{1.13}$$

has the property in question within model (1.12); such a result generalizes a previous one obtained by Hofmann et al. (1992) to the case of a non-zero correlation. If one takes model (1.3) as the volatility generating mechanism, with $W^\sigma = \rho W^{(1)} + (1-\rho^2)^{1/2}W^{(2)}$, for instance, relation (1.13) then becomes:

$$\widetilde{\pi}(t) = (\frac{\partial H}{\partial S}S)(t) + \psi\rho\cdot(\frac{\partial H}{\partial \sigma^\delta})(t)\cdot\sigma(t)^{\delta-1},$$

and can be numerically evaluated by usual methods, such as those exploited in the last section of Hofmann et al. (1992). In an empirical study, Chernov and Ghysels (1999) have already made use of formula (1.13) within the Heston's framework. In addition to derive such a formula in a general diffusion context, chapter 3 also adds a few results connecting strategies like (1.13) to the Hull and White hypothesis that the volatility risk premium is nil.

Naturally, one may wish to consider criteria that are more general than the preceding one, but the price to be paid is a more complex analysis that has recently received a somewhat detailed treatment in mathematical finance: portfolio selection strategies belong to a very old research activity, but techniques in continuous time economies with incomplete markets have been introduced relatively recently; see the introduction to chapter 3 for a list of

[14]Such an approach is very close to that developed by Mankiw (1986) in a different context.

some of the most important initial contributions and surveys papers in this domain. Fundamentally, two approaches have been formulated. In the first one, one searches over strategies that minimize a loss function, and a more general formulation than the one considered above is, for instance, the continuous time incomplete market problem considered by Duffie and Richardson (1991): $\widehat{\pi}(p) = \arg\min_\pi E(V^{p,\pi}(T) - \widetilde{X})^2$. In a series of papers that are cited in the introduction of chapter 3, Schweizer subsequently defines an *approximation price*, equal to $\widehat{p} = \arg\min_p E(V^{p,\widehat{\pi}(p)}(T) - \widetilde{X})^2$. The second approach follows the revolutionary perspective introduced by Bensaid et al. (1992) in the transaction costs literature, and identifies the bounds of a continuum of arbitrage-free prices of the claim; such bounds correspond to the so-called 'dominating' strategies rather than the standard duplicating strategies of the complete markets case.[15] Such an approach can be extended to any well behaved situation in which markets are incomplete; see, for instance, Cvitanic et al. (1997) for an application to models with stochastic volatility. All such more general issues are not treated here.

1.3.2 The term structure of interest rates

Despite the increased importance played by stochastic volatility in financial economics, only a few theoretical term structure models take into account such a phenomenon in the same fashion as one has observed for the European option pricing theory in the last decade.

A notable exception is the early equilibrium model of Longstaff and Schwartz (1992), in which the instantaneous interest rate is a linear combination of two factors, thus generating a two-factor model *à la* Cox et al. (1985*a*). Since volatility was driven there by the same Brownian motions driving the instantaneous interest rate, however, volatility acted in a way that is rather different from the one that is usually thought of in the traditional stochastic volatility literature, where volatility is typically *not* adapted to the filtration generated by the Brownian motion driving the observables. In fact, in one of the first empirical studies devoted to these issues, Andersen and Lund (1997*a*) convincingly propose to extend a model studied in Chan et al. (1992):

$$\mathrm{d}r(t) = (\iota - \theta r(t))\mathrm{d}t + \overline{\sigma} r(t)^\eta \mathrm{d}W(t), \tag{1.14}$$

where $\iota, \theta, \overline{\sigma}, \eta$ are real parameters, so as to incorporate a stochastic volatility factor in the following manner:

$$\begin{aligned} \mathrm{d}r(t) &= (\iota - \theta r(t))\mathrm{d}t + \sigma(t) r(t)^\eta \mathrm{d}W^{(1)}(t) \\ \mathrm{d}\log\sigma(t)^2 &= \kappa(\alpha - \log\sigma(t)^2)\mathrm{d}t + \psi \mathrm{d}W^{(2)}(t) \end{aligned} \tag{1.15}$$

[15] One bound is the minimum cost that is needed to obtain at least the same payoffs as those promised by the claim. The other bound is obtained by using a symmetric argument.

where κ, α, ψ are real parameters that guarantee the existence of a solution to (1.15).

Model (1.14) is a univariate generalization of the square-root model of Cox et al. (1985*a*) model, which sets $\eta \equiv \frac{1}{2}$. In general, $\eta > 0$ makes interest rate volatility increase with the interest rate level. This is the so-called 'level-effect': as is clear, the parameter η measures the sensitivity of the instantaneous volatility to the level of the interest rate; in fact, Chan et al. (1992) suggest that $\eta > 1$ is empirically more plausible than $\eta < 1$. The drawback of model (1.14), however, is that it can not accommodate for the autocorrelation of volatilities in the fashion described in the preceding section. In contrast, this is not the case of a model like (1.1) or (1.15): as Andersen and Lund pointed out, a representation such as (1.15) is ad hoc, but it is also in accordance with "similar formulations for general financial time series", such as those presented in the preceding subsection.

On a theoretical standpoint, Fong and Vasicek (1991), Fornari and Mele (1994, 1995) or Chen (1996) have primitives with a more traditional stochastic volatility flavor than the model of Longstaff and Schwartz (1992), for they propose that the short term interest rate is the solution of:[16]

$$\begin{aligned} \mathrm{d}r(t) &= (\iota - \theta r(t))\mathrm{d}t + \sigma(t) r(t)^\eta \mathrm{d}W^{(1)}(t) \\ \mathrm{d}\sigma(t)^2 &= \kappa(\alpha - \sigma(t)^2)\mathrm{d}t + \psi\sigma(t)\mathrm{d}W^{(2)}(t) \end{aligned} \tag{1.16}$$

or:

$$\begin{aligned} \mathrm{d}r(t) &= (\iota - \theta r(t))\mathrm{d}t + \sigma(t)\mathrm{d}W^{(1)}(t) \\ \mathrm{d}\sigma(t)^2 &= \kappa(\alpha - \sigma(t)^2)\mathrm{d}t + \psi\sigma(t)^2\mathrm{d}W^{(2)}(t) \end{aligned}$$

The latter formulation appeared in our 1994 and 1995 papers, and was motivated by the fact that it represents the diffusion limit of an AR(1)-GARCH(1,1) process of the short term interest rate. Its main inextricable disavantages are that: (1) the short term interest rate can attain negative values; and (2) the model does not take account of the level effect.[17] Furthermore, that model was estimated via identification techniques *à la* Nelson, without controlling the adequacy of the approximating model via the consistency tests that will be succinctly presented in chapter 5. Our papers were written much before the date they were published and we did not have the same powerful techniques of today for estimating the parameters of a system of stochastic differential equations: such techniques are going to be discussed succinctly in the following section (see chapter 5 for a technical presentation).

The common theoretical drawback of the models of Fong-Vasicek, Fornari-Mele and Chen is that they do not arise from an equilibrium theory determining

[16] Parametric restrictions in these models were as follows: in Fong-Vasicek, $\gamma \equiv 0$, and in Chen, $\gamma \equiv \frac{1}{2}$. Actually the Chen's model is more general than (1.16) for it allows $\frac{1}{\theta}$ to follow a third diffusion model.

[17] Brenner et al. (1996) presented a discrete-time ARCH model that took care of the level effect.

the risk-premia demanded by agents to be compensated for the stochastic fluctuations of (r, σ^2).[18]

In chapter 4, we present an equilibrium model that attempts to overcome the preceding critiques. Its concern is to find economies in which generalized versions of systems like (1.1), (1.15) or (1.16) can be viewed as the *equilibrium* data generating process, and as a by-product of this we provide the partial differential equation that has to be followed by the equilibrium price of a default free bond. The approach that we follow is close to the framework of Cox et al. (1985*b*), with the exception that we replace linear activities with stocks.[19] As in Longstaff and Schwartz (1992), the primitives of the economy include two factors, but instead of generating equilibria in which the instantaneous interest rate is a linear combination of the factors, we specify the factor dynamics in a way that allows for the equilibrium short term rate to be linear in the first factor only. Stochastic volatility of the kind considered in this section then simply emerges because the first factor exhibits stochastic volatility—interpreted as the second factor—that is not adapted to the filtration generated by the Brownian motion driving the first factor. The model is derived under standard preference restrictions of a representative agent. Specifically, we assume logarithmic utility and find an equilibrium by using standard dynamic programming techniques. Such assumption and techniques are made for pedagogical purposes only. In Fornari and Mele (1999*a*), the model is solved by assuming a CRRA utility function and utilizing martingale techniques within a more pronounced general equilibrium framework.

To conclude, we would notice that the approach that is favored here might appear to be in contrast with the approach that we are following in our European option pricing empirical studies, where we do not assume any prior pertaining to the preferences of agents; see chapter 3. It has to be reminded, however, that apart from the work of Longstaff and Schwartz (1992), we do not know of any paper dealing with continuous time stochastic volatility models of the term structure in a fully articulated equilibrium framework: as explained above, it is instead a common practice to impose specific, *non-flexible* functional forms to otherwise unidentified risk-premia. However, imposing ad hoc risk-premia that rule out arbitrage opportunities would be an interesting

[18]In a subsequent paper, Andersen and Lund (1997*b*) try to examinate the consequences of their model (1.15) on the term-structure of interest rates. As in the early contributions of Fong-Vasicek, Fornari-Mele and Chen, the authors's specification of the risk-premia is ad hoc. In addition to this, however, Andersen and Lund propose a three-factor model where the "long term" interest rate is another stochastic process—as in previous work of Chen—, and solve their model via Monte Carlo techniques.

[19]Such a modification essentially serves to jointly determine the dynamics of the stock price, which will be compared to the ones that are typically posited in the stochastic volatility pricing literature on European options.

agenda if the corresponding functional forms were general enough, especially when the final purpose of the researcher is to fit cross-sectional information. As noted, this is *not* the case of the above cited papers, but rather than undertaking such an objective, we preferred to pursue the more modest, yet logically preliminary task of developing more *theory* on the determinants of the risk premia within simple preference restrictions, which is for the moment more than an under-developed area. Naturally, one extension of our work would consist in applying a flexible, data-oriented approach to the determinants of the risk-premia, in the vein of our empirical work in the stochastic volatility option pricing area where, instead, it is now well-understood how risk-premia are affected by standard preference restrictions.

1.4 Statistical inference[20]

A recurrent difficulty arising in econometrics concerns the estimation of models giving rise to criterion functions that have no manageable analytical expressions. In modern finance theory, for instance, models are typically set in continuous time. Such a choice is justified by the fact that in continuous time, powerful mathematical techniques[21] exist, which have no counterparts in discrete time.[22] The resulting models are often diffusion processes, but jump diffusion processes are also part of a traditional research program. Furthermore, non-Markovian models have also been proposed, both in the term structure of interest rates (e.g., Heath et al. (1992), Comte and Renault (1996) (section 4.1)) and in the option pricing literatures (e.g., Comte and Renault (1998)). In this monograph, we constrain ourselves to Markovian models.

Apart from special cases such as the celebrated Black-Scholes or the Cox et al. (1985*a*) models, theoretical financial models typically give rise to transition and/or ergodic distributions for the observables that are not known explicitly, since these are solutions to parabolic partial differential equations that can only be solved numerically. Hence, the likelihood function implied by the measure induced by a discretely sampled diffusion can not be calculated explicitly.[23]

[20]This section is motivated by our applied research work on the estimation of the parameters of stochastic differential equations for the short term interest rates. Applications of the material discussed herein to stock returns, for instance, can be conceived in a straight forward manner. Furthermore, surveys that are more focussed on option pricing issues include Taylor (1994), Ghysels et al. (1996) and Shephard (1996).

[21]Essentially: stochastic calculus and elegant dynamic programming techniques and measure theory.

[22]See, however, Campbell et al. (1997) (chapter 11) for a survey of some interesting discrete time interpretations of continuous time models of the term-structure of interest rates.

[23]Following Lo (1988), ML estimation might also be feasible if the transition density

Accordingly, the proposed methods rely on nonparametric density estimation (e.g., Aït-Sahalia (1996*a*,*b*)) and/or closed-form approximations of the true (unknown) likelihood function of the discretely sampled diffusion (Aït-Sahalia (1998)), on generalized method of moments (e.g., Hansen and Scheinkman (1995), Conley et al. (1997)), or on the indirect inference principle.

Methods based on the indirect inference principle are particularly well suited to problems where the state is not fully observed, as it happens in the case of models with stochastic volatility. The quintessence of the methods relies on the simulation of the theoretical model. Its philosophy consists in comparing data simulated from the theoretical model with real data. If the model is a good description of reality (insofar as we are willing to accept the imperfections of *any* model!), then there should exist values of its parameters that make simulated data from that model 'resembling' to real data. On the statistical point of view, it is precisely the way how we think about comparing the two data sets that generates the so-called *auxiliary* criteria. In a classic contribution devoted to applied macroeconomics, for instance, Kydland and Prescott (1982) generated simulated moments from artificial economies corresponding to their models, from which they constructed reasonable 'confidence bands' that contained the sample-based moments corresponding to the US economy. The authors concluded that their model was a successful description of reality.

In a sense, the Kydland-Prescott procedure can be thought to be one of the latest antecedents of the modern simulation-based econometric techniques; Marcet (1994) has an excellent discussion concerning this point. The Kydland-Prescott techniques, however, did not insist on the formal statistical testing aspect of the story, which is of course a central issue of the modern methodology (see the debate of Kydland and Prescott (1996), Hansen and Heckman (1996) and Sims (1996)).

Back to finance, one of the first empirical study in which simulation-based methods were applied to estimating continuous time models of the short term interest rate was conducted by Broze et al. (1995*a*), who consider estimating a slightly more general version of the Chan et al. (1992) model (1.14): apart from methodology, one of the objectives of this study was to find empirically flexible functional forms of the diffusion of the short term interest rate, as opposed to the simple square root process of Cox et al. (1985*a*). It is instructive to remind that the issue of functional flexibility of the diffusion function of the short term interest rate has been pushed to the extreme by Aït-Sahalia (1996*a*), who estimated such a function nonparametrically. The basic idea can be explained

of the observables could be computed easily. When this is not the case, ML becomes computationally demanding. In the continuous time stochastic volatility case, for instance, ML would require to implement a numerical solution to a multi-dimensional partial differential equation at each iteration of the optimization algorithm. The likelihood would then be obtained by integrating out with respect to volatility.

as follows. Consider the following data generating process,

$$\mathrm{d}r(t) = \mu(r(t);\theta)\mathrm{d}t + \sigma(r(t))\mathrm{d}W(t)$$

where, in the preceding notation, only the drift $\mu(r;\theta)$ has been parametrized with θ: the function σ has to be estimated non-parametrically. Now it is well known that if r has a stationary distribution, denoted as π, that has support in the extended positive line and boundary condition $\pi(0) = 0$, π is then the solution of the following ordinary differential equation:

$$\mu(r;\theta)\pi(r) = \frac{1}{2}\frac{\partial}{\partial r}(\sigma(r)^2\pi(r)),\ \pi(0) = 0.$$

The key observation now is that in lieu of solving for π for a given σ^2, one may also choice to integrate the preceding equation for a given π and obtain:

$$\sigma(r)^2 = \frac{2}{\pi(r)}\int_0^r \mu(x;\theta)\pi(x)\mathrm{d}x.$$

After plugging in the preceding relation a non-parametric estimation of π obtained by standard methods (see, for instance, Härdle and Linton (1994)), one can obtain an estimate of σ^2 in a non-parametric way in correspondence of a given choice for μ. As concerns the drift function, Aït-Sahalia estimated a linear function in his original paper, but one can also add nonlinearities such as those considered in Aït-Sahalia (1996*b*) in a different context; see, also, Conley et al. (1997) or Stanton (1997) for related work. Issues pertaining to the nonlinearity of the drift function of the short term interest rate will be shortly presented in chapter 4.

As is clear, the preceding ideas are particularly interesting to apply to systems in which the state is completely observable: one of the main advantages of such an approach, indeed, is that it does not require any simulation of the system.[24] In contrast, this monograph deals with systems in which the state is partially observed due to the presence of stochastic volatility. This is one explanation for our choice of estimating continuous time stochastic volatility models via indirect inference.

A second explanation has been put forward in the introductory section of this chapter. Despite the great progress that has been made in the last decade in the estimation of the parameters of stochastic differential equations systems, one important aspect of our empirical research agenda is to understand to which extent ARCH models can be used as reliable approximators of continuous time stochastic volatility systems. Specifically, the quality of the filtering properties of ARCH models is now well-understood and, as stated in the introductory

[24] Hansen and Scheinkman (1995) and Aït-Sahalia (1998) also provide methods that are simulation-free. Again, such approaches are particularly well-suited to problems in which the state of the model is completely observable.

section, it has been further confirmed in a simulation study in Fornari and Mele (1999*a*) conducted in correspondence of a diffusion designed for the short-term interest rate dynamics (cf. eq. (1.1)). In contrast, there is no empirical work dealing with the quality of the approximation to the parameters of stochastic differential equations systems that is delivered by the moment conditions under which ARCH models converge to their continuous time counterparts. Now it turns out that one interesting way to address such an issue on a solid statistical-sounded basis just requires the simulation-based techniques that are associated with the indirect inference principle. Let us explain why.

As is clear, the preliminary step of any simulation-based method consists in an appropriate choice of the auxiliary criterion with which comparing real data with simulated data. In the context that is studied here, a natural auxiliary criterion can be based on the parameters' estimates of an ARCH model fitted to the available data.[25] In some cases, the estimation strategy would consist in finding parameters values of the continuous time model generating simulated data that, once sampled at the same frequency of the available data, can be fitted *with exactly the same ARCH model that fitted the real data*: this would be a just-identified problem. When, instead, the discrete time ARCH model has more parameters than the continuous time model, one obtains a classical over-identified problem, and the indirect inference estimator would now minimize an appropriate distance between the two sets of discrete time parameters (i.e., the parameters of the model applied to the observed data, and the parameters of the model applied on the simulated data).

The estimation strategy that we follow in our applied work focusses on the methodologically simple but empirically difficult just-identified case, in which the number of parameters of the discrete time model is equal to the number of parameters of the continuous time model. Naturally, our strategy does not spring out of nowhere and uses statistical techniques that were originally suggested in the seminal paper of Gouriéroux et al. (1993) (p. S108):

> "[Indirect inference] methods seem particularly promising when the criterion is based on approximations of the likelihood function, time discretization, range discretizations, linearizations, etc. In this case the method is simpler [...] and appears as an automatic correction for the asymptotic bias implied by the approximation."

It is clear how to identify the source of "the asymptotic bias implied by the

[25] Engle and Lee (1996) is the first paper we know in which ARCH models were exploited as direct and/or indirect devices for approximating stochastic differential equations. While the authors were concerned with stock returns data only, they also suggested (p. 352) to extend the Cox et al. (1985*a*) model to a setting with stochastic volatility of the kind described in the preceding section. As pointed out before, Andersen and Lund (1997*a*) is one of the first empirical studies that accomplished such a task.

approximation" in our context: as we reminded in the introductory section, most ARCH models are not closed under temporal aggregation, which suggests that using moment conditions ensuring the convergence towards a continuous time model should introduce an "asymptotic bias implied by the approximation" or, more correctly said, a *disaggregation* bias. Yet, ARCH models still have a natural interpretation in terms of the continuous time models that are supposed to approximate, since they are very close (in terms of the probability distributions generating them) to the continuous time models when the sampling frequency is high. Furthermore, it turns out that we are also endowed with a natural one-to-one interpretation of the sequence of the discrete time parameters of the auxiliary models in terms of the parameters of the continuous time model (see chapter 5 for technical details): as is clear, we exactly are in the position precognized by Gouriéroux et al. (1993), and we are only left with testing and correcting potential disaggregation biases.

The appropriate testing procedure has been designed within the logic of the indirect inference principle. It is based on testing procedures originally suggested by Gouriéroux et al. (1993) (section 4.2) that can be viewed as the natural substitutes of global specification tests in just-identified problems. Chapter 5 presents a technical description of the test, as well as the technical justification of it within our framework; it also succinctly describes the empirical results of Fornari and Mele (1999*a*), where it is shown that the disaggregation bias of fitting an ARCH model to weekly US interest rate data is not significant on the basis of that test. This is a particularly interesting empirical result. The simple reason is that Drost and Nijman (1993) constructively showed that ARCH models aggregate only when one weakens the concept of an ARCH model, which led the authors to introduce the so-called *weak*-ARCH process; more importantly, Drost and Werker (1996) generalized the Drost-Nijman setting and introduce the so-called GARCH diffusion which is, heuristically, the continuous time stochastic volatility process whose implied discrete differences form a weak-ARCH process. More precisely, a continuous time process $\{y(t)\}_{t\geq 0}$ is a GARCH diffusion if its implied differences process $\{({}_hy_{h(k+1)} - {}_hy_{hk})\}_{k=1}^{\infty}$, $kk \leq t < h(k+1)$, is weak-GARCH for any $h > 0$, i.e., if there exist a sequence of parameters (w_h, α_h, β_h) and a covariance-stationary process,

$$ {}_h\sigma^2_{hk} = w_h + \alpha_h \cdot {}_hy^2_{h(k-1)} + \beta_h \cdot {}_h\sigma^2_{h(k-1)}, \tag{1.17} $$

that is the best linear predictor of $({}_hy_{hk} - {}_hy_{h(k-1)})^2$ in terms of 1, ${}_h\sigma^2_0$ and lagged values of $({}_hy_{hk} - {}_hy_{h(k-1)})$ and $({}_hy_{hk} - {}_hy_{h(k-1)})^2$. Take, for instance, the following diffusion,

$$ \begin{aligned} \mathrm{d}y(t) &= \sigma(t)\mathrm{d}W^{(1)}(t) \\ \mathrm{d}\sigma(t)^2 &= \theta(\omega - \sigma(t)^2)\mathrm{d}t + \sqrt{2\lambda\theta}\sigma(t)^2\mathrm{d}W^{(2)}(t) \end{aligned} \tag{1.18} $$

where θ, ω, λ are parameters that satisfy $\theta > 0, \omega > 0$ and $\lambda \in (0,1)$. Drost and Werker (1996) (prop. 3.1 p. 37) then show that there is a continuous mapping

with an inverse from the parameters of the continuous time model (1.18) on to the parameters of the discrete time model (1.17).

We believe that the most natural interpretation of our empirical findings is that even though the ARCH models we use do not aggregate, they still remain, for a given frequency, an excellent approximation to the continuous time models towards which they converge in distribution, at least insofar as they are a natural proxy to the weak-ARCH models. Naturally, these are issues that deserve a deep theoretical investigation that we leave for further research.

A second way to implement the indirect inference principle has a rationale that is different from the one outlined above. Its main feature is to select, as an auxiliary device, an highly parametrized discrete time model that is used with the main purpose of calibration. Such an auxiliary model then generates a score (hence referred to as 'score generator'), and the objective becomes to search for the values of the parameters of interest that make such a score as close as possible to zero by using a long simulation of the theoretical model. Such a method has been introduced by Gallant and Tauchen (1996), and is referred to as efficient method of moments (EMM): heuristically said, the source of asymptotic efficiency comes here from the fact that if the true likelihood function is embedded in the density associated with the auxiliary model, then the EMM estimator achieves the same efficiency of the true ML estimator.[26] In practice, an embedding density can be built-up by providing additional parameters to the discrete time model with a semi-nonparametric (SNP) expansion of the distribution of the residuals by means of Hermite polynomials. In fact, as subsequently shown by Gallant and Long (1997), if the score generator is such an SNP, the efficiency of the EMM estimator can be made as close to the ML one as desired by taking the number of the auxiliary parameters large enough. One of the earliest applications of the EMM techniques to models of the stock prices with continuous time SV is in Gallant and Tauchen (1997), and the first application of EMM theory to continuous time SV models of the short term interest rate is in Andersen and Lund (1997*a*, *b*). Gallant and Tauchen (1997) also consider the application of EMM to interest rates models that have not stochastic volatility, while Gallant et al. (1997) apply the EMM technique to the discrete time SV models that have been succinctly presented in section 1.2.1. In all these applications, the score generator had a nonparametric density which also accommodated for an ARCH-type scale function.

[26] In this book, we are adopting the convention to include the EMM theory of Gallant and Tauchen (1996) as a part of the indirect inference principle.

1.5 Plan

Before giving the plan, it is useful to clarify what the following four chapters are and what are not: the rest of the monograph is intended as a succinct account of our past as well as ongoing research program in which we try to isolate our own contribution. Hence, the following chapters do *not* include extensive surveys on the state-of-the-art of the topics we treat. We only constrain ourselves to refer the reader to already published surveys or, when these are not available, provide a list of the papers that are related to our work, without however delving into the details.

Chapter 2 is devoted to a systematic presentation of our approximation results obtained in correspondence of some of the ARCH models presented in section 1.2. In addition to provide results that are useful when formulating and empirically implementing continuous time models with stochastic volatility, our objective also lies in finding results that can be useful up to a first order approximation treatment of the steady-state probabilistic properties of such models in discrete time.

Chapter 3 analyzes a few problems arising from the incomplete markets structure that is generated by the presence of continuous time stochastic volatility; our primary focus is on European-type options; we make use of a model with diffusion state variables. Although markets are incomplete insofar as one restricts attention to the primitive assets of the economy, the option itself can be taken to complete the markets; as a consequence of this, the risk premia demanded by agents to be compensated for the stochastic fluctuations of the state variables of the economy can be found via the preferences of a representative agent. As we mentioned in section 1.3.1, we then illustrate how we are currently attacking the problem: instead of imposing a functional form generated by a specific preference structure of a representative agent, as we do for the term structure model in chapter 4, we take the volatility risk-premium as a nonlinear function of the state variables of the model (i.e., a 'volatility risk-premium surface'), that can subsequently be estimated using cross-sectional information derived from option prices. One of the final objectives of the chapter is to provide a short description of hedging strategies that can be implemented within an economy with continuous time stochastic volatility. By delving into the simplest versions of the literature on risk-minimizing strategies—as opposed to the standard risk-neutralizing strategies *à la* Black-Scholes—we provide details concerning the construction of strategies for partial hedging in incomplete markets in the general version of the model, by focussing then on its stochastic volatility restrictions.

Chapter 4 presents a succinct overview of the theory of the term structure of interest rates within Markovian economies, and focusses essentially on the ramifications generated by 'injecting' stochastic volatility features into them. It then imposes restrictions to the model with diffusion state variables of chap-

ter 3, and develops a class of equilibrium models in which the instantaneous interest rate exhibits stochastic volatility that is imperfectly correlated with the instantaneous interest rate level itself. As concerns the statistical inference, this chapter also provides a very first illustration of simulation-based econometric techniques that can be applied to estimate continuous time models of the short term interest rates. Furthermore, it explains the role played by the linearity of the diffusion functions of the state variables of the economy to assist in getting tractable models: our discussion will thus concern a very special case of the well-known literature on affine models of the term structure.

Chapter 5 presents in detail the econometric techniques that are required to make estimation and testing procedures applied to the parameters of our theoretical model of the term structure of interest rates. These techniques are based on a combined use of the approximation results of chapter 2 and the indirect inference principle. This chapter also presents methodology to obtain the solution of our theoretical model of the term structure of interest rates with stochastic volatility. We follow two approaches. In the first one, we use the Crank-Nicholson scheme to numerically integrate two-dimensional partial differential equations that typically accommodate for stochastic volatility; a Matlab code to implement the solution of our model is available upon request (our code takes approximately 1 minute to obtain the solution with Matlab 4.2 on a Pentium II 366 MHz with 64 Mbytes of memory). While the code has been specifically designed for solving our term structure model, only minor changes are required for that code to be used to solve related problems (e.g., models with different drift or diffusion functions and/or computation of transition measures in continuous time). Finally, we show how to implement a second, less traditional approach that is based on a method of iterated approximations.

2

CONTINUOUS TIME BEHAVIOR OF NON LINEAR ARCH MODELS

2.1 Introduction

This chapter presents convergence results for the A-PARCH model (1.6) that was originally proposed by Ding et al. (1993). We remind that in addition to be a particular convenient tool to model volatility asymmetries, such a model imposes a sort of Box-Cox power transformation to the conditional standard deviation. According to this model, the 'volatility concept' is thus not imposed a priori by the modeler, but it has to be estimated from data. By assuming that such a transformation is the same at every sampling frequency, we derive continuous time results for model (1.6). Such results are useful for three main reasons: (1) they help formulating continuous time models that are flexible with respect to the choice of the volatility concept (see chapter 5); (2) they provide a simple identification device through which estimating the correlation process between a continuous time asset price process and its instantaneous volatility; (3) they help understanding the role played by the volatility concept in determining the long run behavior of the error process of the model.

The chapter is organized in the following manner. The approximation results for model (1.6) are in the following section; section 2.3 contains comments concerning the moment conditions that are needed to guarantee the convergence of the discrete time model; section 2.4 provides a primer on the connection between the approximation results and option pricing; section 2.5 is devoted to the study of the stationary distribution of the A-PARCH models innovations; section 2.6 provides continuous time results for the VS-ARCH model (1.9), but the analysis there is not as deep as the analysis conducted for model (1.6). The appendices contain technical material.

2.2 Approximation results for a general class of non linear ARCH models

If h denotes the sampling interval, we partition time in (1.6) in a way that allows for the corresponding solution $\{{}_hY_{hk}\}_{k=1}^{\infty} \equiv \{{}_hS_{h(k-1)}, {}_h\sigma_{hk}^{\delta}\}_{k=1}^{\infty}$ to