Journal of Empirical Finance 8 (2001) 83–110

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

Recovering the probability density function of asset prices using garch as diffusion approximations

Fabio Fornari [a,*], Antonio Mele [b]

[a] *Research Department, Bank of Italy, Via Nazionale 91, 00184 Rome, Italy*
[b] *Université de Paris X, Thema, Paris, France*

Accepted 2 February 2001

Abstract

This paper uses Garch models to estimate the objective and risk-neutral density functions of financial asset prices and by comparing their shapes, recover detailed information on economic agents' attitudes toward risk. It differs from recent papers investigating analogous issues because it uses Nelson's result that Garch schemes are approximations of the kind of differential equations typically employed in finance to describe the evolution of asset prices. This feature of Garch schemes usually has been overshadowed by their well-known role as simple econometric tools providing reliable estimates of unobserved conditional variances. We show instead that the diffusion approximation property of Garch gives good results and can be extended to situations with (i) non-standard distributions for the innovations of a conditional mean equation of asset price changes and (ii) volatility concepts different from the variance. The objective PDF of the asset price is recovered from the estimation of a nonlinear Garch fitted to the historical path of the asset price. The risk-neutral PDF is extracted from cross-sections of bond option prices, after introducing a volatility risk premium function. The direct comparison of the shapes of the two PDF$_s$ reveals the price attached by economic agents to the different states of nature. Applications

* Corresponding author. Tel.: +39-6-4792-3189; fax: +39-6-4792-4118.
E-mail address: f_fornari@hotmail.com (F. Fornari).

PII: S0927-5398(01)00021-4

are carried out with regard to the futures written on the Italian 10-year bond.

JEL classification: C22; G12; G13
Keywords: Option pricing; Stochastic volatility; Arch; Volatility risk premium

1. Introduction

It is now a common practice among researchers and practitioners to extract detailed forward-looking information from financial asset prices. Recently, much interest has been devoted to recovering the risk-neutral probability density function (henceforth PDF) of asset prices embodied in the prices of options. The characteristics of the risk-neutral density can be recovered from a panel of option prices by estimating the parameters of a system of stochastic differential equations, a field in which non-parametric or semi-non-parametric (SNP) schemes have found an increasing role. Examples are the non-parametric approach to the term structure modeling and option pricing of Aït-Sahalia (1996a,b), Aït-Sahalia and Lo (1998, 2000) and Aït-Sahalia et al. (1998) as well as the SNP model coupled with efficient method of moments (EMM) employed to estimate the parameters of a short term interest rate diffusion by Gallant and Tauchen (1998) and applied to an option pricing scheme by Chernov and Ghysels (2000).

Despite the flexibility of such estimators, in this paper we adopt a more parametric approach to extract the information contained in the objective and the risk neutral distribution functions of the futures contract written on the Italian 10-year benchmark government bond (henceforth BTP). We do this by adopting a Garch scheme as a tool to recover the coefficients of the underlying continuous time generating process, both under the objective and the risk-neutral measures.

Though the motivation and the contents of this paper are intrinsically linked to the alternative analyses mentioned above, there are many original aspects in our work: (i) compared to the non-parametric approach, we explicitly consider a true stochastic volatility case, as revealed by the use of a discrete time Garch model as a continuous time stochastic volatility filter, along the lines put forward by Nelson (1992, 1996) and Nelson and Foster (1994); (ii) compared to EMM estimation, our auxiliary model is not as highly parameterized as required by that technique: we use an auxiliary (discrete time) model that is close to the continuous time scheme and our need for indirect inference is only required as a correction for the disaggregation bias (see Section 2.1); (iii) from a more technical standpoint, unlike Chernov and Ghysels (2000) we do not impose linearity for the drift function of the volatility process under the risk neutral probability measure. Instead, we propose a risk premium volatility surface that is a polynomial function of the primitives of the model and we test and compare its ability to price options against the competing linear specification; (iv) finally, we investigate a central issue of financial analysis, i.e. whether volatility risk is priced and what is its shape. The paper is organized as follows: Section 2 shows how to use Garch models to estimate the density functions under the objective and risk neutral measures. Section 3 applies the techniques developed in Section 2. Section 4 compares the shapes of the two densities and recovers parameters of interest for risk management. Section 5 concludes.

2. Recovering the PDF

2.1. Objective PDF: DGP-based analysis

Derivatives are typically priced upon the assumption that the value of the underlying asset is a solution of a stochastic differential equation (SDE) with fixed variance such as

$$\mathrm{d}F_t = \mu F_t \mathrm{d}t + \sigma F_t \mathrm{d}W_t^{(1)}, \tag{1}$$

where F is the underlying price process (e.g. the price of a share or the price of futures written on bonds), $W^{(1)}$ is a standard Brownian motion, μ is a real parameter and σ is the volatility parameter.

The empirical evidence has strongly rejected the assumption of stability of σ in financial markets, showing instead that volatility evolves according to an unpredictable sequence of calm and turbulence, i.e. to heteroskedasticity. With heteroskedastic returns, the evaluation of derivative assets becomes more difficult than in the standard Black and Scholes (1973) world; also, the Black and Scholes model suffers severe distorsions (Hull and White, 1987). Allowing variance to be generated according to an autonomous SDE generally translates in closed-form solutions requiring highly restrictive assumptions; Hull and White, the typical reference model, assume that the price of the underlying asset follows Eq. (1) with σ^2 being the solution of another SDE

$$\mathrm{d}\sigma_t^2 = \phi\sigma_t^2 \mathrm{d}t + \zeta\sigma_t^2 \mathrm{d}W_t^{(\sigma)},$$

where $W^{(\sigma)}$ is a Brownian motion, and ϕ,ζ are real parameters. To derive a closed-form solution for the price of a call option based on this model,[1] they assume among other things that: (i) the instantaneous correlation between the two Brownian motions $W^{(1)}$ and $W^{(\sigma)}$ is nil; (ii) ϕ is nil and (iii) in pricing assets, agents do not need compensation for the fluctuations of volatility.[2] We will relax the above three assumptions throughout the paper and the basic pricing methodol-

---

[1] Other examples of closed-form solutions for the price of derivatives on assets with stochastic volatility have been provided by Heston (1993) or Leblanc (1995).

[2] It is well known that (iii) is equivalent to the property that traded risks are discounted martingales under the so-called minimal martingale measure introduced by Föllmer and Schweizer (1991).

ogy that we will follow is based on the simulation of the relevant pair of stochastic differential equations. This latter tool is necessary whenever one attempts to generalise the Hull and White model by resorting to more complex dynamics and analytic solutions are no longer available. Under such conditions, approximating schemes based on conditionally heteroskedastic autoregressive models represent, in our view and with the support of the data, a simple tool to recover the parameters of the continuous time model. As we will show, they can be used either as a direct device or as an indirect device within an indirect inference approach. Nelson (1990), for instance, showed that there exist versions of Ar(1)–Garch(1,1)[3] processes that converge in distribution to the solution of the Hull and White model.[4] Other schemes, better responding to real life, have also been suggested; in particular, if one supposes that the conditional volatility of an asset price follows the Power Arch scheme of the first order introduced by Ding et al. (1993), i.e.

$$\begin{aligned} &\sigma_n^{\ \delta} = \omega_1 + \alpha_1\left(|\varepsilon_{n-1}| - \gamma\varepsilon_{n-1}\right)^{\delta} + \beta_1\sigma_{n-1}^{\delta}, \\ &\gamma \in (-1,1), (\omega_1, \alpha_1, \beta_1, \delta) \in R_+^4, \end{aligned} \tag{2}$$

where the indexing $n = 0,1,\dots$ refers to consecutive observations sampled at the same frequency, and ε_n is a sequence of zero mean uncorrelated error terms coming from Eq. (3) below, then Fornari and Mele (1997, 2000a) have shown that Eq. (2), coupled with equations of the form

$$\begin{aligned} &F_n = F_{n-1} + (\iota_1 - \theta_1 \cdot F_{n-1}) + F_{n-1}^{\chi}\varepsilon_n, \\ &\varepsilon_n = (\sigma \cdot u)_n, u_n \sim i.i.d., \iota_1, \theta_1 \in R_+^2, \chi \in (0,1), \end{aligned} \tag{3}$$

are approximating processes converging in distribution toward the solution of the following stochastic differential system

$$\begin{aligned} &\mathrm{d}F_t = (\iota - \theta F_t)\mathrm{d}t + \sigma_t F_t^{\chi}\mathrm{d}W_t^{(1)}, \\ &\mathrm{d}\sigma_t^{\ \delta} = \left(\omega - \varphi\sigma_t^{\ \delta}\right)\mathrm{d}t + \zeta\sigma_t^{\ \delta}\mathrm{d}W_t^{(\sigma)}, \end{aligned}$$

where ι, θ, ω, φ, ζ, δ are non-negative parameters, $W^{(1)}$ is a standard Brownian motion, $W^{(\sigma)} = \rho W^{(1)} + \sqrt{1-\rho^2}\, W^{(2)}$, $W^{(2)}$ is another standard Brownian motion ($W^{(2)}$ independent of $W^{(1)}$) and $\rho \in (-1,1)$. In models such as Eqs. (2) and (3), the concept of volatility is not restricted a priori, as in the traditional Garch(1,1) scheme, but can be estimated from the data; for example, when $\delta = 1$ the relevant volatility concept is the standard deviation, while when $\delta = 2$, it is variance that matters. Note further that leaving δ free to be estimated will generally translate in

[3] By Ar(1)–Garch(1,1) we mean that, given a stationary series ε_t, its conditional mean evolves according to an autoregressive model of the first order, while its conditional variance is a Garch(1,1).

[4] See, however, Corradi (2000) for conditions under which the diffusion approximation property fails.

a greater flexibility for the estimation of both the simulated objective PDF of the asset price and the risk-neutral option pricing function

$$C_t = e^{-r(T-t)} \cdot E_t^Q (F_T - K)^+, \tag{4}$$

where K is the strike of the option, T is the expiration date, Q is the risk neutral PDF, and r is the (constant) risk-free interest rate. However, although the methods that we employ in this paper could allow in principle to obtain consistent estimates of χ, ι, θ, the computational effort would be considerably increased, so that we only restrict attention to the special case $\chi = 1$ and $\iota = \theta = 0$. The model that we take as the DGP in this paper is thus

$$\mathrm{d}F_t = \sigma_t F_t \mathrm{d}W_t^{(1)} \tag{5}$$

$$d\sigma_t^{\delta} = \left(\omega - \varphi \cdot \sigma_t^{\delta}\right)\mathrm{d}t + \zeta\sigma_t^{\delta}\mathrm{d}\left(\rho W_t^{(1)} + \sqrt{1-\rho^2} \cdot W_t^{(2)}\right),$$

whereas the discrete-time model that we use as a direct and/or indirect device to estimate the parameter vector of interest $\boldsymbol{a}_{\_} = (\omega, \varphi, \zeta, \rho, \delta)$ is

$$F_n = F_{n-1} + F_{n-1} \cdot \varepsilon_n, \quad \varepsilon_n = (\sigma u)_n, \tag{6}$$

with Eq. (2) as the volatility propagation equation.

In the discrete-time model one can easily allow for the case of a non-normal distribution, normality being unanimously rejected as a conditional or unconditional representation of the innovations. The specification that we adopt here is that u is generated by the *ged*[5]

$$u \sim ged_{\upsilon} = \frac{\upsilon \exp\left(-\frac{1}{2} \Xi_{\upsilon}^{-\upsilon} \cdot |u|^{\upsilon}\right)}{2^{1+\frac{1}{\upsilon}} \cdot \Xi_{\upsilon} \Gamma\left(\frac{1}{\upsilon}\right)}, \quad \Xi_{\upsilon}^{2} = \frac{\Gamma\left(\frac{1}{\upsilon}\right)}{2^{\frac{2}{\upsilon}} \Gamma\left(\frac{3}{\upsilon}\right)},$$

where $\Gamma(\cdot)$ is the Gamma function, and υ ($\upsilon > 0$) is the tail thickness parameter of the *ged*. Fornari and Mele (1997) showed that in some interesting cases the joint use of (δ, υ) can provide enough flexibility to fit complicated asset price dynamics.

To obtain an estimate of $\boldsymbol{a}_{\_}$ within this framework, one can start by estimating the parameters $b = (\omega_1, \alpha_1, \beta_1, \gamma, \delta)$ of the discrete time system (Eqs. (6) and (2)) by maximum likelihood (ML),

$$\hat{b} = \arg\max_{b} L_N(F; b),$$

where $L_N(\cdot)$ is the likelihood implied by the chosen model and N is the sample size.

[5] GED stands for the general error distribution.

In a preliminary step, one then recovers the parameter vector of interest $\boldsymbol{a}\_$ by using the moment conditions that guarantee the convergence in distribution of the discrete time system toward Eq. (5), i.e.

$$\begin{aligned}
&\omega = \Delta^{-3/2}\omega_1\\
&\varphi = \Delta^{-1}\Big(1 - n_{\delta,v}\cdot\Big[(1-\gamma)^{\delta} + (1+\gamma)^{\delta}\Big]\cdot\alpha_1 - \beta_1\Big)\\
&\zeta = \Delta^{-1/2}\sqrt{\Big(\big(m_{\delta,v} - n_{\delta,v}^2\big)\cdot\big(\gamma_*^{2\delta} + \gamma_{**}^{2\delta}\big) - 2\cdot n_{\delta,v}^2\cdot\gamma_*^{\delta}\cdot\gamma_{**}^{\delta}\Big)}\cdot\alpha_1\\
&\rho = \frac{2^{\frac{\delta-v+1}{v}}\,\Xi_v^{\delta+1}\Gamma\left(\frac{\delta+2}{v}\right)\left(\gamma_*^{\delta} - \gamma_{**}^{\delta}\right)}{\Gamma(v^{-1})\sqrt{\big(m_{\delta,v} - n_{\delta,v}^2\big)\cdot\big(\gamma_*^{2\delta} + \gamma_{**}^{2\delta}\big) - 2n_{\delta,v}^2\cdot\gamma_*^{\delta}\cdot\gamma_{**}^{\delta}}}\\
&\gamma_* = 1-\gamma\\
&\gamma_{**} = 1+\gamma\\
&n_{\delta,v} = \frac{2^{\frac{\delta}{v}-1}\,\Xi_v^{\delta}\Gamma\left(\frac{\delta+1}{v}\right)}{\Gamma\left(\frac{1}{v}\right)}\\
&m_{\delta,v} = \frac{2^{\frac{2\delta}{v}-1}\,\Xi_v^{2\delta}\Gamma\left(\frac{2\delta+1}{v}\right)}{\Gamma\left(\frac{1}{v}\right)},
\end{aligned} \tag{7}$$

where Δ is the fraction of the sample frequency to the "numéraire" period (e.g. $\Delta = 1/24$ if the sample frequency is hourly and the unit of time is expressed in days). The derivation of these formulae is detailed in Fornari and Mele (1997, 2000b). To get an intuition of them, let ω_h be a sequence of the form $(\omega_h)_{h\searrow 0}$,

$$\varphi_h = 1 - n_{\delta,v}\cdot[(1-\gamma)^{\delta} + (1+\gamma)^{\delta}]\cdot\alpha_h - \beta_h,$$

$$\zeta_h = \sqrt{\big(m_{\delta,v} - n_{\delta,v}^2\big)\cdot[(1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}] - 2n_{\delta,v}^2\cdot(1-\gamma)^{\delta}\cdot(1+\gamma)^{\delta}}\\ \cdot\alpha_h,$$

and

$${}_h\xi_{hk} = \frac{\left|\frac{{}_hu_{hk}}{\sqrt{h}}\right|^{\delta}(1-\gamma\cdot s_k)^{\delta} - E\left(\left|\frac{{}_hu_{hk}}{\sqrt{h}}\right|^{\delta}\cdot(1-\gamma\cdot s_k)^{\delta}\right)}{\sqrt{\big(m_{\delta,v} - n_{\delta,v}^2\big)[(1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}] - 2n_{\delta,v}^2\cdot(1-\gamma)^{\delta}(1+\gamma)^{\delta}}},$$

where $({}_h u_{hk})/(\sqrt{h})$ is *ged*. Then formulae (7) are identifying conditions: Eqs. (2) and (6) can be embedded in a scheme of the form

$$\begin{aligned} &{}_hF_{h(k+1)} - {}_hF_{hk} = {}_h\sigma_{h(k+1)} \cdot {}_hF_{hk} \cdot {}_hu_{h(k+1)} \\ &{}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} = \left(\omega_h - \varphi_h \cdot {}_h\sigma^{\delta}_{hk}\right) + \zeta_h \cdot {}_h\sigma^{\delta}_{hk} \cdot {}_h\xi_{hk}, \end{aligned} \tag{8}$$

when $h=1$; if there exist sequences of the form $(\omega_h)_{h\searrow 0},(\varphi_h)_{h\searrow 0},(\zeta_h)_{h\searrow 0}$ such that $\lim_{h\downarrow 0} h^{-1}\omega_h = \omega$, $\lim_{h\downarrow 0} h^{-1}\varphi_h = \varphi$,$\lim_{h\downarrow 0} h^{-1/2}\zeta_h = \zeta$, then Eq. (8) converges in distribution to the solution of Eq. (5) as $h\downarrow 0$. As regards the intuition behind ρ in Eq. (7), this can be based on the appealing result that

$$\zeta^{-1}\frac{1}{h}E\left(\frac{{}_hF_{hk} - {}_hF_{h(k-1)}}{{}_hF_{h(k-1)} \cdot {}_h\sigma_{hk}} \cdot \frac{{}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk}}{{}_h\sigma^{\delta}_{hk}} \Big/ \mathscr{F}_{hk}\right)_{h\downarrow 0} \to \rho,$$

where $\mathscr{F}_{hk}$ denotes the sigma-algebra generated by ${}_hF_0, {}_hF_h, {}_hF_{2h}, \ldots, {}_hF_{h(k-1)}$ and ${}_h\sigma_0^{\delta}, {}_h\sigma_h^{\delta}, {}_h\sigma_{2h}^{\delta}, \ldots, {}_h\sigma_{hk}^{\delta}$. The same approximating scheme holds trivially for a conditionally normal Garch, after setting $\delta = \upsilon = 2$, $\gamma = 0$.[6]

One general difficulty with the moment conditions reported in Eq. (7), however, is that the class of strong Garch models (those for which the rescaled innovations, ε_n/σ_n, are i.i.d.) is not closed under temporal aggregation (Drost and Nijman, 1993; Drost and Werker, 1996) so that when the continuous time parameters obtained via Eq. (7) are plugged into a Euler's discretization of Eq. (5), a disaggregation bias arises. Hence they must be regarded only as a starting point for an indirect inference approach which delivers their unbiased estimates for the chosen discrete frequency. Thus models (2) and (6), i.e.,

$$\begin{aligned} &F_{n+1} - F_n = \sigma_{n+1}F_nu_{n+1}, u_n \sim ged_{\upsilon}, \\ &\sigma_{n+1}^{\delta} - \sigma_n^{\delta} = \omega_1 - \left[\left(1 - n_{\delta,\upsilon}\cdot\left[(1-\gamma)^{\delta} + (1+\gamma)^{\delta}\right]\cdot\alpha_1 - \beta_1\right)\right]\cdot\sigma_n^{\delta} \\ &\qquad\qquad + \alpha_1\cdot\sigma_n^{\delta}\cdot\left[|u_n|^{\delta}\cdot(1-\gamma\cdot s_n)^{\delta} - E\left(|u_n|^{\delta}\cdot(1-\gamma\cdot s_n)^{\delta}\right)\right], \\ &s_n \equiv \operatorname{sign}(u_n), \end{aligned}$$

can be used as one of the possible discrete time counterparts of the DGP in an indirect inference scheme with simulations drawn from

$$\begin{aligned} &{}_hF_{h(k+1)} - {}_hF_{hk} = {}_h\sigma_{h(k+1)} \cdot {}_hF_{hk} \cdot {}_hu_{h(k+1)} \\ &{}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} = \left(\omega - \varphi \cdot {}_h\sigma^{\delta}_{hk}\right)\cdot h + \zeta \cdot {}_h\sigma^{\delta}_{hk} \cdot {}_h\xi_{hk}\cdot\sqrt{h}, \end{aligned} \tag{9}$$

for $h = N^{-d}$ with $d > 1/2$.[7] We also double the dimension of the high frequency generator by replacing ${}_h\xi_{hk}\sqrt{h}$ with ${}_h\tilde{\xi}h_{(k+1)}$, where $E({}_h\tilde{\xi}_{hk}) = 0$, $\operatorname{var}({}_h\tilde{\xi}_{hk}) = h$ and $\operatorname{corr}({}_hu_{hk}\cdot{}_h\tilde{\xi}_{hk}) = \rho h$, for all h.

[6] In a companion paper dealing with term-structure issues (Fornari and Mele, 2000a), we show that Eq. (7) indeed provide a reasonable approximation, in that the correction brought about by indirect inference is not important in terms of a specification test based on the ex-post adequacy of an approximating model having the form of Eqs. (2) and (6).

[7] Such a value for h avoids asymptotic bias due to simulations; see Broze et al. (1998).

We leave δ to be estimated by maximum likelihood, so that the parameter vector of interest to be recovered restricts to $\boldsymbol{a} = (\omega, \varphi, \zeta, \rho)$. The indirect inference estimation then runs as follows. After simulating Eq. (9) in correspondence of values of $\boldsymbol{a}$, we obtain ${}_{h,h}\tilde{F}^{(s)}(\boldsymbol{a}) = \{{}_h\tilde{F}^{(s)}_{hk}(a)\}_{k=0}$, $s = 1, \dots, S$, where S is the number of simulations. For each simulation, we retain the values of $\overset{(s)}{\tilde{F}}$ which correspond to integer indexes of time and estimate the auxiliary model on each simulated series to get

$$\hat{b}^{(h)}_{N,s}(\boldsymbol{a}) = \arg\max_b L_N\left({}_{1,h}\tilde{F}^{(s)}(\boldsymbol{a}); b\right), \qquad s = 1, \dots, S,$$

where ${}_{1,h}\tilde{F}^{(s)}(\cdot)$ denotes the set of the bond futures prices with integer indices of time under simulation s and time interval h. An indirect estimator of $\boldsymbol{a}$ is the solution ${}_h a_N$, say, of the system

$${}_h\hat{a}_N: 0 = \hat{b}_N - \frac{1}{S}\sum_{\tilde{s}=1}^{S} \hat{b}^{(h)}_{N,s}\left({}_h\hat{a}_N\right).$$

The asymptotics for ${}_h\hat{a}_N(\boldsymbol{a}_0)$ can be found in a straight forward manner by combining Gouriéroux et al. (1993) and Broze et al. (1998), specifically,

$$\sqrt{N}\left({}_h\hat{a}_N(\boldsymbol{a}_0) - \boldsymbol{a}_0\right)_{N\uparrow\infty, h\downarrow 0} \overset{d}{\rightarrow} N\left(0, \frac{S+1}{S}\left(\mathbf{V}_0^{-1}\mathbf{J}_0^{-1}\mathbf{I}_0\mathbf{J}_0^{-1}\mathbf{V}_0'^{-1}\right)\right), \qquad (10)$$

where $\boldsymbol{a}_0$ is the true parameter vector in Eq. (5), $\mathbf{V}_0$ is the 4×4 Jacobian matrix of the binding function (here, the derivatives of the discrete time parameters with respect to the continuous time parameters) evaluated at $\boldsymbol{a}_0$, $\mathbf{J}_0$ the pseudo-true Hessian and $\mathbf{I}_0$ the pseudo-true covariance matrix of the scores of the auxiliary parameters.

2.2. Risk-neutral PDF: option based analysis

In this subsection we present the scheme designed to recover the risk-neutral PDF. Given the model presented in the preceding subsection under Eq. (5), the methodology here consists in calibrating the parameters of the analogous continuous time bivariate diffusion which generates, under the risk-neutral measure, the changes in the underlying asset F_t and its volatility ${\sigma_t}^\delta$ in such a way as to match, as closely as possible, observed cross-sections of option prices with different strike prices and times to expiration. The assumption underlying our method is that the market is correctly pricing the options by formula (4). Under the risk-neutral probability measure, by Girsanov theorem system (5) can be written as

$$\begin{aligned} \mathrm{d}F_t &= -\lambda_t^{(1)}\sigma_t F_t \mathrm{d}t + \sigma_t F_t \mathrm{d}\hat{W}_t^{(1)} \\ \mathrm{d}{\sigma_t}^\delta &= \left(\omega - \varphi \cdot {\sigma_t}^\delta - \zeta \cdot \rho \cdot {\sigma_t}^\delta \cdot \lambda_t^{(1)} - \zeta \cdot \sqrt{1-\rho^2} \cdot {\sigma_t}^\delta \cdot \lambda_t^{(2)}\right)\mathrm{d}t \\ &\quad + \zeta {\sigma_t}^\delta\left(\rho \mathrm{d}\hat{W}_t^{(1)} + \sqrt{1-\rho^2}\,\mathrm{d}\hat{W}_t^{(2)}\right), \end{aligned} \qquad (11)$$

where $\hat{W}^{(1)}$ and $\hat{W}^{(2)}$ are standard Brownian motion under a probability measure Q that is absolutely continuous with respect to P, the primitive, objective measure; $\lambda^{(1)}$ and $\lambda^{(2)}$ are to be interpreted as risk-premia associated with the fluctuations of the (Eq. (5)) Brownian motions, measurable with respect to the filtration of $(W^{(1)},W^{(2)})$ under Q. To ensure existence for the measure Q, it is sufficient to impose a Novikov condition on $\lambda^{(1)}$, $\lambda^{(2)}$. Furthermore, the market incompleteness argument which affects schemes as Eq. (5) where the asset F has to cope with two sources of risk, loses strength if the option is taken to complete the markets in the Bajeux and Rochet (1996) sense and according to the further developments in Romano and Touzi (1997), as concerns an extension of the Bajeux and Rochet results to the case of a correlation process and general risk-premia of the kind that is considered here. This means essentially that $\lambda^{(1)}$ and $\lambda^{(2)}$ could be uniquely determined via preference restrictions of a representative agent. As regards $\lambda^{(1)}$, we have immediately by the martingale property of $e^{-rt}F_t$ under Q that it must satisfy: $\lambda^{(1)} = -r\sigma^{-1}$.[8] Substituting back into Eq. (11),

$$\begin{aligned}
\mathrm{d}F_t &= rF_t\mathrm{d}t + \sigma_t F_t \mathrm{d}\hat{W}_t^{(1)} \\
\mathrm{d}\sigma_t{}^{\delta} &= \left(\omega - \varphi\cdot\sigma_t{}^{\delta} + r\cdot\zeta\cdot\rho\cdot\sigma_t{}^{\delta-1} - \lambda_t^{(2)}\cdot\zeta\cdot\sigma_t{}^{\delta}\sqrt{1-\rho^2}\right)\mathrm{d}t \\
&\quad + \zeta\sigma_t{}^{\delta}\left(\rho\mathrm{d}\hat{W}_t^{(1)} + \sqrt{1-\rho^2}\,\mathrm{d}\hat{W}_t^{(2)}\right).
\end{aligned} \tag{12}$$

It remains to identify the risk-premium $\lambda^{(2)}$. Although this can be done through a representative agent argument, as put forward before, we assume here that it can be reasonably represented by means of a polynomial structure. This idea follows from $\lambda^{(2)}$ being measurable with respect to the filtration of $(W^{(1)},W^{(2)})$ and further, by combining a result of Harrison and Kreps (1979) with Romano and Touzi (1997), from the circumstance that the filtration of $(W^{(1)},W^{(2)})$ and the filtration of (F,σ) must coincide: then $\lambda^{(2)}$ is a functional of past and current values of (F,σ). We restrict attention to a Markovian structure and take the following form[9]

$$\lambda_t^{(2)} = \Lambda(F_t,\sigma_t) = p_1 + p_2\cdot\sigma_t{}^{\delta} + p_3\cdot\sigma_t{}^{2\delta} + \frac{p_4}{\sigma_t{}^{\delta}}. \tag{13}$$

We call the preceding function volatility risk premium surface. Our objective now is that of estimating the parameter vector $\boldsymbol{\vartheta} = (p_r)_{r=1}^4$ in Eq. (13) and recover the shape of such unknown function. To do this, we simulate system (12) from the

[8] The fact that $\lambda^{(1)} < 0$ does not necessarily imply that agents are risk-lovers in a model with (1) stochastic volatility; (2) negative correlation between volatility and the asset price; (3) positive risk-premia for volatility. A proof of this is available on request. All these conditions are met in our empirical implementation of the model; in Section 4, we show indeed that risk aversion is positive for a wide range of values of F.

[9] We started with a much more complicated functional form, but Eq. (13) gave the best results with the data employed in Section 3.

current period t up to maturities $T=[T_1,\ldots,T_{N1}]$, where T matches the maturities of observed cross-sections of N_1+N_2 traded options, where the additional N_2 options come from the observed structure of the strike prices. In practice, the risk-neutral PDF is easily obtained by matching the observed to the theoretical option prices, the latter being simulated from the bivariate data generating process of F and σ^{δ} under Q, where the vector of parameters $\boldsymbol{a}\equiv(\omega,\varphi,\zeta,\rho)$ is fixed at the indirect inference estimates ${}_h\hat{a}_N(\boldsymbol{a}_0)$ obtained before and the only unknown left to be determined are the coefficients in $\boldsymbol{\vartheta}$. For each simulation i, the call prices are evaluated for the observed strike prices as $C_{j,T_\ell}(\boldsymbol{\vartheta})=(1/\text{sim})\mathrm{e}^{-r(T_\ell-t)}\Sigma_{i=1}^{\text{sim}}(F_{T_\ell i}(\boldsymbol{\vartheta})-K_j)^+$, where $F_{T_\ell i}(\boldsymbol{\vartheta})$ denotes the price at time T_ℓ at the ith simulation obtained with the parameter vector $\boldsymbol{\vartheta}$, sim is the number of simulations and K_j is the jth strike price $(j=1,\ldots,N_2)$. The measure of distance between observed and fitted prices can be constructed as $D(\boldsymbol{\vartheta})=1/(N_1+N_2)\Sigma_{\ell=1}^{N1}\Sigma_{i=1}^{N2}[C_{i,T_\ell}-C_{i,T_\ell}(\boldsymbol{\vartheta})]^2$, where C_{i,T_ℓ} is the observed option price; an estimator of $\boldsymbol{\vartheta}$ is arg $\min_{\vartheta}[D(\boldsymbol{\vartheta})]$. More details of this are reported in Section 3.

3. Empirical analysis

So far, we showed the usefulness of Garch schemes in approximating models developed in continuous time as systems of stochastic differential equations. To recall, the discrete time coefficients of a Garch(1,1) or a Power Arch(1,1) are employed to recover the parameters of the continuous time model to which they converge via the moment conditions (Eq. (7)). Then such parameters are corrected for the disaggregation bias by indirect inference and used to simulate paths of the asset price $F_T(T>t)$. Under the risk-neutral measure, the corrected coefficients obtained for system (5) under the objective measure are still inserted in the simulation scheme (Eq. (12)) of the asset price and volatility under the risk-neutral measure; in addition explicit account is here made for the functional form of the volatility risk premium, whose parameters are identified by fitting option prices obtained averaging the discounted values of $(F_T-K)^+$ over the simulations. Though we are mainly interested in the density $f(F)$, it is worth pointing out that, under the objective measure, other distributions can be extracted from the simulation scheme; also, response analyses to modifications of the parameters of the SDE can be envisaged, which may turn out useful for applications dealing with the measurement of the so-called Value-at-Risk (VaR). For example, the densities $f(F,t,T)$ and $g(\sigma^2,t,T)$ or the joint density $q(F,\sigma^2)$ can be constructed through a Monte Carlo simulation of Eq. (5) performed by the Euler–Maruyama discretization scheme of the type reported in Eq. (9).

To start with the objective density function, Eqs. (2) and (6) are estimated with a *ged*-based likelihood function as illustrated in Section 2.1. These estimates are performed for the full sample 1 January 1991 to 20 January 1997 as well as for two subsamples having the same origin but ending on 22 July 1996 and 22 October 1996, respectively. Results for the three samples are reported in Table 1

Table 1
Parch models

Estimation from January 1, 1991 to July 22, 1996
$F_t = 117.59; \sigma_t^2 = 0.0000236608$
$u_t = \dfrac{F_t - F_{t-1}}{F_{t-1}} = \varepsilon_t; \varepsilon_t \| I_{t-1} \sim ged_{1.3188}$
$\sigma_t^{0.8607} = 1.7879 \cdot 10^{-4} + 0.08971(\|\varepsilon_t\| - 0.5354\varepsilon_{t-1})^{0.8607} + 0.9212\,\sigma_{t-1}^{0.8607}$
Estimation from January 1, 1991 to October 22, 1996
$F_t = 122.60; \sigma_t^2 = 0.0000202765$
$u_t = \dfrac{F_t - F_{t-1}}{F_{t-1}} = \varepsilon_t; \varepsilon_t \| I_{t-1} \sim ged_{1.3038}$
$\sigma_t^{0.9864} = 8.58153 \cdot 10^{-5} + 0.08876(\|\varepsilon_t\| - 0.44396\,\varepsilon_{t-1})^{0.9864} + 0.92189_{t-1}^{0.9864}$
Estimation from January 1, 1991 to January 20, 1997
$F_t = 131.47; \sigma_t^2 = 0.0000173162$
$r_t = \dfrac{F_t - F_{t-1}}{F_{t-1}} = \varepsilon_t; \varepsilon_t \| I_{t-1} \sim ged_{1.2685}$
$\sigma_t^{1.1870} = 2.53724 \cdot 10^{-5} + 0.08324(\|\varepsilon_t\| - 0.3932\,\varepsilon_{t-1})^{1.1870} + 0.92662\,\sigma_{t-1}^{1.1870}$

Note: Models are all Power Arch(1,1). F_t is the value of the Italian 10-year bond (BTP) futures; σ_t^2 is its conditional variance; I_t is the information set dated t; ged_v indicates the density defined in Section 2.1. All the parameters are significant according to the Bollerslev and Wooldridge's (1992) t-statistics. Sample size is 1208, 1274, and 1334 for the three subsamples, respectively.

for the Power Arch(1,1) case.[10] The conditional distribution of the errors departs from normality, as highlighted by υ, the tail-thickness parameter, which is

[10] The choice of these three samples follows from previous work investigating market's ability at forecasting official interest rate moves. For this reason, all of the three samples end the day before a change in the official interest rates. This circumstance has obviously no direct implication in the present work.

significantly different from two and more supportive of a Laplace assumption, even though this hypothesis has to be statistically rejected.[11] The estimate of δ is very close to unity, though increasing across the three dates (0.86,0.99,1.19); a standard volatility concept based on the variance (i.e. $\delta = 2$), for instance, is rejected. However, we repeat the estimates also for a Garch(1,1) with conditionally normal errors which, providing a discrete time approximation for the well-known Hull and White (1987) scheme, may provide a benchmark in the option pricing application reported below. The values of the parameters that maximise the likelihood function in this case are reported in Table 2.

Table 3 reports the four continuous time parameters computed using the moment conditions (Eq. (7)). These figures are corrected via the indirect inference principle as shown in Table 4. In the three Garch cases (Table 4 last three lines), the effect of the correction was sizeable, with seven out of the nine changes exceeding 20%. The drift parameter φ moves from 0.198, 0.232 and 0.27 to 0.15, 0.18 and 0.15 in the three samples, respectively, while the diffusion terms are strongly pushed downwards, from 0.94, 0.90 and 0.81 to 0.75, 0.60 and 0.60, respectively. The corrections are much lower in the case of the Power Arch scheme, with just one coefficient out of 12 changing by more than 20%. The volatility drift coefficients (φ) were 0.82, 0.74, 0.56 at the three dates and the corrected figures 0.90, 0.60 and 0.65; the analogous figures for the diffusion terms move from 0.51, 0.55 and 0.64 to 0.63, 0.55 and 0.57, respectively. The coefficient of conditional correlation remains stable after the correction.

We evaluated the t-ratios for the corrected coefficients with concern only for the case in which the Power Arch plays the role of auxiliary model (it will also be the only scheme upon which the subsequent analysis rests) and for the overall sample. For the vector of parameters (0.015, 0.65, 0.57, -0.56) (Table 4, last column, first four rows), the corresponding t-ratios evaluated according to the variance in Eq. (10) were found to be (23.41, 7.11, 3.34, -3.70), highly supportive of the significance of the estimates. It is also interesting to note that three out of the four coefficients obtained as concerns this sample through the moment conditions (Eq. (7)) (see Table 3, last column, first four rows) are within the 95% confidence bands defined by the corresponding indirect inference estimates and by the t-ratios reported above. The continuous time coefficient ω fails, albeit marginally (its value based on formulae (7) is 0.01269 against a lower value of the confidence band as of 0.0137) to reconcile with the indirect inference correction.[12]

[11] In the three cases, the closest-to-unity value that cannot be statistically rejected is approximately 1.1.

[12] Analogous results supporting the reliability of the continuous time coefficients implied by the Power Arch via closed formulae (7) were found in a closely related paper (Fornari and Mele, 2000a). See also Fornari and Mele (2000b) for further discussion.

Table 2
Garch models

Estimation from January 1, 1991 to July 22, 1996

$F_t = 117.59; \sigma_t^2 = 0.0000253876$

$$u_t = \frac{F_t - F_{t-1}}{F_{t-1}} = \varepsilon_t; \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = 3.00851 \cdot 10^{-7} + 0.08393 \varepsilon_{t-1}^2 + 0.9134 \sigma_{t-1}^2$$

Estimation from January 1, 1991 to October 22, 1996

$F_t = 122.60; \sigma_t^2 = 0.000023163$

$$u_t = \frac{F_t - F_{t-1}}{F_{t-1}} = \varepsilon_t; \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = 2.9634 \cdot 10^{-7} + 0.0804 \varepsilon_{t-1}^2 + 0.91592 \sigma_{t-1}^2$$

Estimation from January 1, 1991 to January 20, 1997

$F_t = 131.47; \sigma_t^2 = 0.000025465$

$$u_t = \frac{F_t - F_{t-1}}{F_{t-1}} = \varepsilon_t; \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = 2.9634 \cdot 10^{-7} + 0.07211 \varepsilon_{t-1}^2 + 0.92356 \sigma_{t-1}^2$$

Note: Models are all Garch(1,1). F_t is the value of the Italian 10-year bond (BTP) futures; σ_t^2 is its conditional variance; I_t is the information set dated t. All the parameters are significant according to the Bollerslev and Wooldridge's (1992) t-statistics. Sample size is 1208, 1274, and 1334 for the three subsamples, respectively.

The objective PDF at two dates, based on these coefficients, will be shown, along with the risk neutral PDF, in Section 4. As recalled before, apart from the density $f(F)$, one can obtain other information items from the simulation scheme based on system (5). To briefly illustrate this opportunity, some additional information about the objective data generating process has been evaluated for the first of the three dates analyzed, 22 July 1996, for the Power Arch case only. The

Table 3
Continuous time parameters implied by the discrete time parameters

	22 July 96	22 October 96	20 January 97
PARCH			
ω	8.94×10^{-2}	4.291×10^{-2}	1.269×10^{-2}
φ	0.81659	0.74186	0.55914
ζ	0.50796	0.55273	0.63806
ρ	-0.61187	-0.55486	-0.52994
GARCH			
ω	1.5044×10^{-4}	1.48184×10^{-4}	1.48184×10^{-4}
φ	0.19821	0.23184	0.27279
ζ	0.94211	0.90249	0.80943

Note: The parameters are obtained by replacing the discrete time parameters of Tables 1 and 2 into the moment conditions (7). For the Garch case, the instantaneous correlation coefficient has been fixed at zero.

closing price of the futures on that day was 117.59, the (not rescaled) conditional variance 2.36×10^{-5}. In Figs. 1–3 we show the distribution of the BTP futures price, $f(F)$, and the densities of the continuously compounded returns (u) evaluated at three horizons, 60, 20 and 5 days, i.e. $g(u,t,60)$, $g(u,t,20)$, $g(u,t,5)$, where with $g(u,t,z)$, we denote the probability of u at time $t+z$ $(z>0)$ given time t. The density of the returns is obtained as in Aït-Sahalia and Lo (2000), i.e.

Table 4
Continuous time parameters estimated via indirect inference

	22 July 96	22 October 96	20 January 97
PARCH			
ω	8.4×10^{-2} (-6.0)	3.5×10^{-2} (-18.4)	1.50×10^{-2} (18.2)
φ	0.90 (10.2)	0.60 (-19.1)	0.65 (16.2)
ζ	0.63 (24.0)	0.55 (0.00)	0.57 (-11.0)
ρ	-0.65 (-6.2)	-0.58 (-4.5)	-0.56 (5.7)
GARCH			
ω	4.5×10^{-5} (-70.0)	1.4×10^{-4} (-5.5)	1.1×10^{-4} (-25.7)
φ	0.15 (-11.0)	0.18 (-21.7)	0.15 (-44.4)
ζ	0.75 (-20.3)	0.60 (-33.5)	0.60 (-25.9)

Note: These estimates correct those derived through Eq. (7) for the (dis)aggregation bias due to the strong Garch class being not closed under temporal (dis)aggregation; see Section 2.1. In parentheses, the percentage change with respect to the values reported in Table 3. The correction is based on the indirect inference procedure described in Section 2.1. Data are simulated out of the auxiliary model (9) with a frequency of 20 per day; the overall length of the three simulations is thus equal to sample size times 20. The sample size for the three subsamples is 1208, 1276, and 1334, respectively.

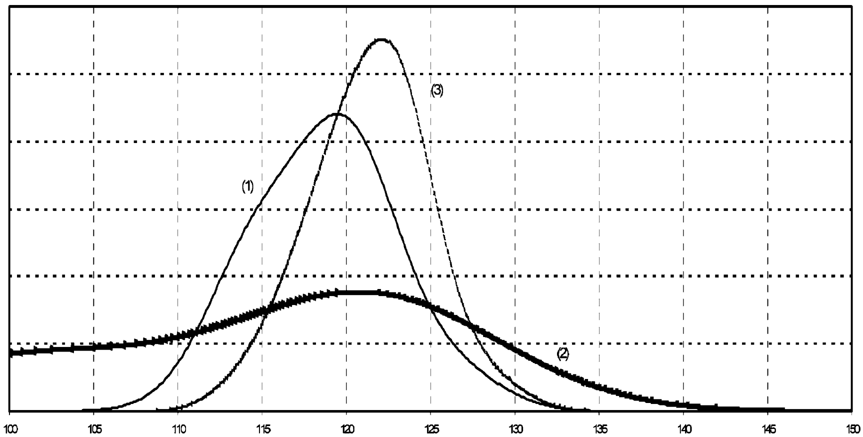

Fig. 1. Density functions of the BTP futures price at July 22, 1996. Note: (1)Density function of the BTP futures prices simulated according to the set of parameters reported in Table 4. (2)Same density but with the diffusion parameter of the volatility equation equal to the value in Table 4 plus one. (3)Drawn by setting the diffusion parameter of the conditional variance equation equal to the value in Table 4 minus 0.5, while the two coefficients in its drift term are scaled so to bring the autocorrelation of the volatility from 0.9857 to 0.9642.

$g(u,t,z)$ is recovered by first evaluating the $(\tau = T - t)$-period continuously compounded returns, $u_\tau = \log(F_T/F_t)$ and then constructing a kernel estimator of

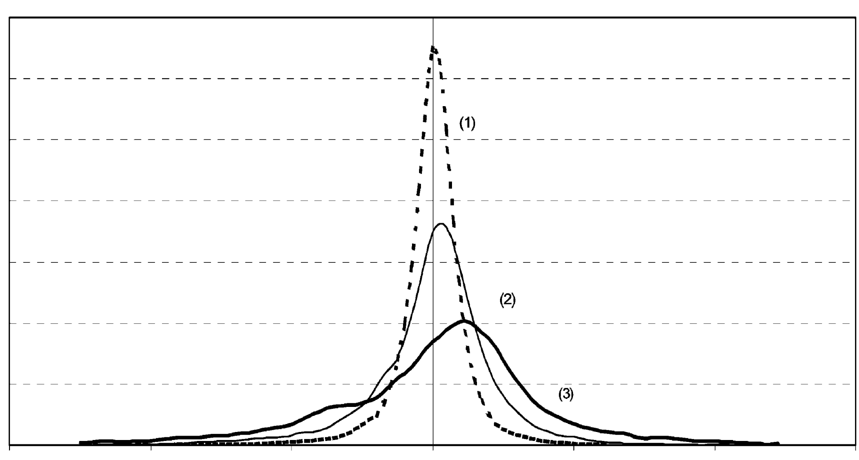

Fig. 2. Density functions of BTP futures returns as of July 22, 1996. Note: (1)$g(u,t,5)$; (2)$g(u,t,20)$; (3)$g(u,t,60)$. The prices of the BTP futures are simulated by plugging the parameters reported in Table 4 into system (5).

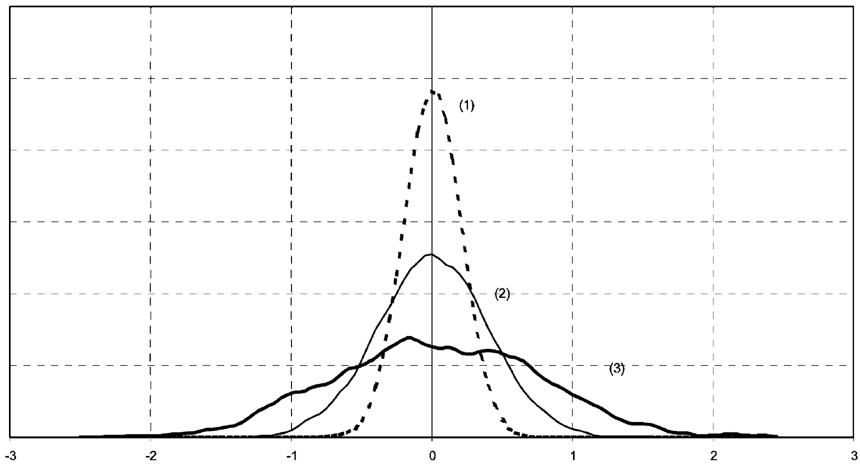

Fig. 3. Density functions of BTP futures returns as of July 22, 1996. Note: $^{(1)}g(u,t,5)$; $^{(2)}g(u,t,20)$; $^{(3)}g(u,t,60)$. The prices of the BTP futures are simulated by plugging the parameters reported in Table 4 into system (5) and lowering the diffusion coefficient of the volatility equation by 0.5 and reducing the autocorrelation of the volatility.

the density function $g(\cdot)$ of these returns.[13] The relation between $g(u_\tau)$ and $f(F,t,T)$ can be established by noting that

$$\Pr(F_T \le F) = \Pr(F_t e^{u_\tau} \le F) = \Pr(u_\tau \le \log(F/F_t)) = \int_{-\infty}^{\log(F/F_t)} g(u_\tau)\, du_\tau .$$

In a second step the price density $f(F,t,T)$ corresponding to the return density estimated in the first step is recoverd as

$$f(F,t,T) = \frac{\partial}{\partial F}\Pr(F_T \le F) = \frac{g(\log(F/F_t))}{F}. \tag{14}$$

This is the density that we use in Section 4; an analogous device will be used to get the risk-neutral density that is also used in Section 4.

The two types of densities (of prices and returns) are based on a single long simulation of 200,000 points drawn at the frequency of 20 per day, which have been successively re-sampled at the desired daily frequency, giving a total of 10,000 points. We also performed a sensitivity analysis exercise, in which the diffusion coefficient of the conditional volatility equation and the speed of reversion of the conditional volatility to its long run mean (which equals ω/φ as

[13] Details on kernel estimation can be found in Härdle and Linton (1994). The reason for adopting kernel estimation here is purely related to graphical purposes, i.e. to nicely represent the simulated PDF.

can be easily seen from Eq. (5)) are moved away from their original values (see Table 4, first column, Power Arch case); this last exercise would be helpful, within a VaR strategy, to anticipate how worse a portfolio's risk might become in a particularly turbulent environment. Fig. 1 shows three price densities: the distribution labelled with (1) corresponds to the simulation performed with the estimated parameters (reported in Table 4); label (2) indicates that the diffusion parameter in the volatility equation has been set to its original value of Table 4 plus one; label (3) indicates instead that the coefficient of autocorrelation of the volatility has been lowered from the estimated figure (0.9857) to 0.9642, while the diffusion parameter has been lowered by 0.5. The most evident thing to see is the larger area put especially in the left tail when the diffusion parameter of the volatility equation is increased, which would attach positive probability to outcomes earlier considered as unrealistic. The densities reported in Fig. 2 are drawn conditional on the estimated parameters; they highlight the range of possible outcomes, expressed in terms of returns, which one may expect 5, 20 and 60 days ahead. It is evident that the typical gaussian quantiles employed in most VaR applications would fail substantially in the definition of extreme event at the 60-day horizon, quite a typical one in most practical cases. The lower value of the diffusion parameter in the volatility equation, under which the densities of Fig. 3 are drawn, evidences the greater weight of central events rather than tail events, compared with Fig. 2. It is interesting to note, looking at Fig. 2, the negative skewness of the distribution which, in principle, one would not expect when F_t evolves like a martingale with fixed volatility or with stochastic volatility not correlated with F_t itself. In the present situation, i.e. with negative correlation between the two Brownian motions, one would instead expect positive skewness to appear, given that returns fall when the shock to the volatility (hence the volatility) increases. This kind of results, i.e. the emergence of a negative skewness, is however in accordance, for instance, with the findings of Heston (1993). What happens is that the negative correlation between the sources of risk increases the area in the left tail of the returns distribution, an occurrence which finds compensation in an additional probability mass in the positive range of returns, aimed at forcing the density to still integrate to one.

Coming to the risk neutral PDF, it is worth considering that if risk-premia were negligible, the values of the four relevant parameters obtained under the objective measure would deliver the optimal pricing scheme also for the call options under the equivalent probability. To test this, and indirectly to evidence the presence of a volatility risk premium, we evaluated the option prices by re-interpreting system (5) under the risk-neutral measure; in this scheme the price of each option is obtained as the average of 1000 simulated prices. The parameters which best priced the three samples of options are reported in Table 5 for both the Garch and the Power Arch case; options are evaluated according to Eq. (4). The fit of the two models at the three dates, 22 July 1996, 22 October 1996 and 20 January 1997, is based on cross-sections of 107, 140 and 135 options, respectively, with maturity

Table 5
Parameters of system (5) which minimize the option pricing error

	ρ	ω	φ	ζ	N	$\Sigma_{I=1}^{N}\varepsilon_i^2/N$
PARCH						
July 22, 1996	-0.069	8.6×10^{-4}	0.125	0.315	107	0.048
October 22, 1996	-0.162	4.75×10^{-4}	0.130	0.250	140	0.036
January 20, 1997	-0.170	1.0×10^{-3}	0.740	0.140	135	0.042
GARCH						
July 22, 1996	0.0	2.6×10^{-4}	0.52	0.135	107	0.112
October 22, 1996	0.0	2.8×10^{-4}	0.62	0.040	140	0.115
January 20, 1997	0.0	2.8×10^{-4}	0.64	0.035	135	0.090

Note: These estimates refer to system (5) interpreted under the risk-neutral measure. They differ from those obtained under the objective since they embody information concerning the risk premia related to the fluctuation of the two state variables (F,σ^2). N is the number of options employed in each of the three samples, ε_i is the difference between the observed and the predicted price of the ith option. The maturities of the options at the three dates ranged from 51 to 149 days.

ranging from 51 to 149 days; the constant short term interest rate has been fixed at 5.0% (per year).[14] There are sizeable differences between the objective and the hypothetical risk-neutral parameters, suggesting that volatility risk is priced by economic agents since, if this were not the case, the parameters should indeed stay unchanged under the two measures.[15] Under the Power Arch scheme the three sums of squared pricing errors were 4.79, 5.05 and 5.69 for the three dates, respectively, which amounts to a mean squared error slightly above 4% on average. The same measure nearly doubles with the Garch assumption and highlights the importance of adopting a proper parameterization of the volatility dynamics.

At this point we use Eq. (12) and fix the relevant vector of parameters at the values appearing under the objective measure; in addition, we let r and $\lambda_t^{(2)}$, as generated by Eq. (13), reveal the risk premia required for the fluctuations of the two state variables; as before, each option price is obtained as simple average of 1000 simulated prices.[16] From now on, we will work only with the Power Arch pricing scheme, based on its superior pricing performance, as evidenced by the preliminary test of Table 5. We use the functional form of Eq. (13) for $\lambda_t^{(2)}$ and we

[14] Option prices are closing prices. In all cases the regression of the 1000 simulated prices on a constant revealed that the latter was statistically significant, which supports the validity of the pricing procedure.

[15] Recent independent work by Chernov and Ghysels (2000) addresses the problem of employing past returns and option prices jointly in the estimation of the relevant parameters of the asset price diffusion.

[16] Since antithetic variates are employed, the simulated prices for each option are 4000, due to 1000 times the (four) combinations of the sign of the two sources of noise. In any case, we report the prices to be 1000 since 1000 is the true number of errors drawn at the beginning of the simulation.

Table 6
Coefficients of the volatility risk premium function

	p_1	p_2	p_3	p_4	$\Sigma_{I=1}^{N}\varepsilon_i^2/N$
PARCH					
July 22, 1996[1]	10.0	−9.6	−40.0	−0.045	0.044
October 22, 1996[1]	7.5	−9.6	−33.0	−0.017	0.006
January 20, 1997[1]	10.0	−9.6	−40.0	−0.035	0.025
Full sample[2]	7.0	−15.0	−40.0	−0.075	0.045
Full sample[3]	7.0	−15.0	−40.0	−0.075	0.085

Note: The term ε_i is the difference between observed and predicted option prices. N is the number of options employed in the evaluation of the pricing error (see Table 5 for the first three samples and notes (2), (3) to this Table for the other samples). The functional form of the volatility risk premium is: $\lambda_t^{(2)} = p_1 + p_2 \cdot \sigma_t^{\delta} + p_3 \cdot \sigma_t^{2\delta} + p_4 / \sigma_t^{\delta}$. [1]The maturity of the options at this date ranged between 51 and 149 days. [2]The full sample comprises 284 trading days between December 18, 1995 and January 31, 1997. The call options examined in this case were those with moneyness ranging from 0.97 to 1.03. There were 7621 of such options; their maturity ranged from 21 to 147 days. [3]The full sample comprises 284 trading days between December 18, 1995 and January 31, 1997. The options examined in this case were those with moneyness ranging from 0.93 to 1.07. There were 15,045 such options; their maturity went from 21 to 147 days.

manage to select values for p_1–p_4 which produce pricing errors analogous to those previously obtained in Table 5, when the objective parameters were free to vary and accommodated the presence of volatility risk premia. The coefficients of this functional form, which does not include any power of the first state variable, i.e. the futures price, while it includes three of the volatility, specifically $(\sigma^{\delta}, \sigma^{2\delta}, 1/\sigma^{\delta})$, in addition to a constant term, for the three dates employed throughout the paper, are shown in Table 6 along with the pricing errors; the shape of the three risk premium functions is reported in Fig. 4. A major concern we had

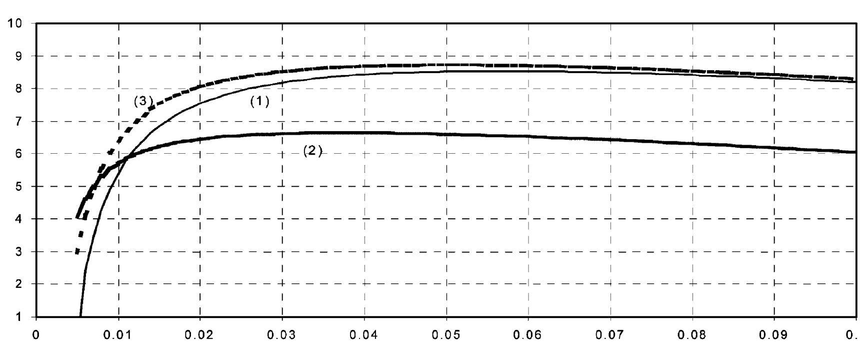

Fig. 4. Volatility risk premium. Note: The volatility risk premium (vrp) comes from the following specification: $vrp = p_1 + p_2 \sigma_t^{\delta} + p_3 \sigma^{2\delta} + p_4 / \sigma_t^{\delta}$. The parameters are reported in Table 6. [1]22 July 1996; [2]22 October 1996; [3]20 January 1997. The volatility reported on the x-axis is σ.

about the stability of this functional form did not turn out to be motivated: the curve appears noticeably stable across the three dates. This is quite an important and reassuring evidence, since the results of the following section, which focuses on the state price densities, rest on the stability of the prices of risk.

To have additional evidence on the stability of such estimated risk premium functions, we enlarged the analysis to a much wider sample. This is made up of 29,173 call options quoted at Liffe and collected between December 18, 1995 and January 31, 1997, embracing the three single dates analyzed before; the maturity of the options ranges from 21 to 147 days and their prices are daily closing prices. The identification of the volatility risk premium in this larger sample is carried out on two subsamples of options, both spreading the whole time sample but differing as concerns moneyness. The first subsample is made up of 7621 options, which are selected according to the criterion employed by Chernov and Ghysels (2000) based on discarding all the options whose moneyness falls below 0.97 and beyond 1.03. The second subsample considers instead the 15,045 options whose moneyness ranges between 0.93 and 1.07; it is used to test the adequacy of the nonlinear volatility risk premium function at pricing in- and out-of-the-money call options. For the first subsample we report, in the fourth line of Table 6, the volatility risk premium specification which delivers the best fit; surprisingly, it does not change significantly from the specification adopted in the first three single trading days.

The value of the pricing error over the 284 trading days and for all the options observed in a given day is reported in Fig. 5. It is analogous in size to the value recorded on July 22, 1996 and is worth noting that 0.045 (i.e. 4.5%) is the average

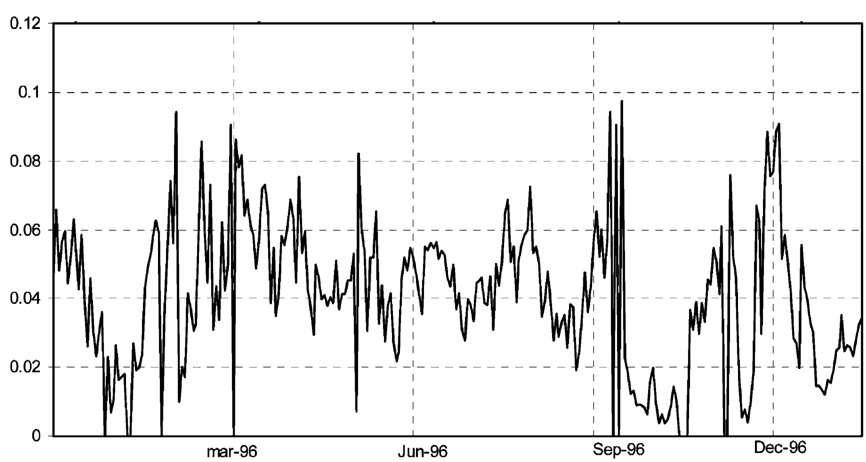

Fig. 5. Option pricing error. Note: The option pricing error reported in this graph refers to 284 trading days between December 18, 1995 and January 31, 1997, for a total of nearly 18,000 options. Each point is the average of the squared deviation between observed and theoretical prices of all options within a trading day.

value for all of the 284 trading days analyzed, including 11 days in which the error was slightly above 10%, which have not been reported in Fig. 5 for graphical reasons. This result is not worse than the figures reported in Chernov and Ghysels (2000), who found the equivalent of our pricing error to be, for options with moneyness ranging from 0.97 to 1.03 and for all the maturities observed in the market, 8.1% under the Black and Scholes pricing scheme and 6.3 under Heston's (1993) stochastic variance scheme. When the sample is further enlarged to consider out- and in-the-money options, yet the best functional form of the volatility risk premium does not change; in this case, however, the mean squared percentage error nearly doubles, to reach 8.5 percent (Table 6, last line). Analogous figures from Chernov and Ghysels are not available in this case, since the authors choose not to price options with moneyness beyond 1.03, given the reported lack of liquidity for such instruments; in any case, the errors they make in pricing options with all maturities and with moneyness ranging from 0.94 to 1.03 are 6.1% under Black and Scholes and 4.0 employing Heston's model. It is obvious to conclude that the greatest part of the additional errors made by our pricing scheme in the second larger sample (Table 6, last column, last line, against last column, sixth line) must be attributed to thinly traded in-the-money options.

4. Implications for risk management

Before examining the indications about risk which one can draw from the estimated density functions, we briefly illustrate why comparing the two types of PDF assumes particular relevance in finance. First of all, as outlined in Aït-Sahalia and Lo (2000), the ever expanding liquidity of financial markets and the increasing globalization calls for appropriate insurance of economic agents' portfolios, a necessity which has recently been popularized by the Value at Risk (VaR) strategy. If an investor did not consider the information contained in the risk-neutral measure but were to follow the indications deriving only from the objective, it is immediate to see, from what were explained in Sections 2.1 and 2.2, that his portfolio would be substantially unbalanced, since he could not include information about state prices. Second, the measure of risk aversion that can be extracted from estimated PDF$_s$ is important insofar as it sheds light on the preferences of investors themselves. By using a simple general equilibrium model à la He and Leland (1993), one has the standard result that[17]

$$\kappa_t(E_T) \equiv \frac{P^{\mathrm{RN}}(F_T)}{P^{\mathrm{OBJ}}(F_T)} = \kappa \frac{U_T'(C_T)}{U_t'(C_t)}, \tag{15}$$

with κ constant, where P^{RN} and P^{OBJ} are the risk-neutral and the objective density, respectively, $U_s(\cdot)$ is the instantaneous utility function as of time s of a

[17] See Rosenberg and Engle (1998) and Jackwerth (2000) for related work.

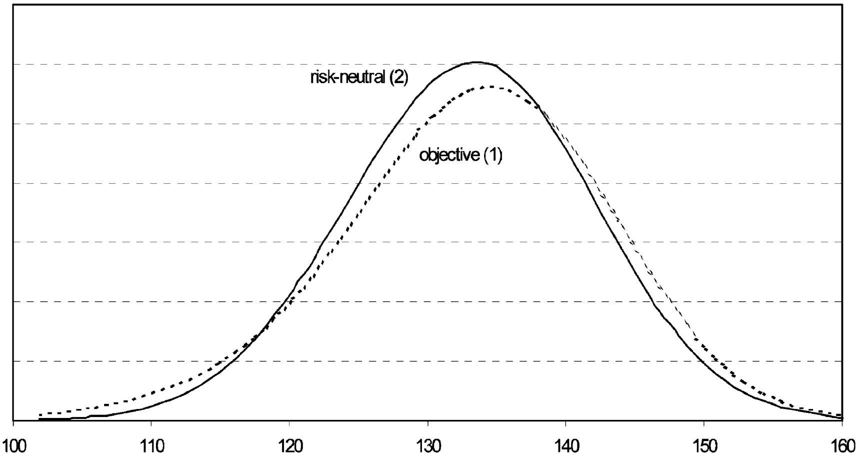

Fig. 6. January 20, 1997: objective and risk-neutral PDF. Note: (1) This is the objective density function of the BTP futures price at January 20, 1997. It is obtained by simulating Eq. (5) with the continuous time parameters reported in Table 4. (2) This is the risk neutral density function. It is obtained by simulating system (12) with the continuous time parameters reported in Table 4 and the volatility risk premium (13) with those reported in Table 6.

representative agent, and C_s is consumption which equals F_s at $s=T$. If $\kappa(\cdot)$ varies across prices one is led to conclude that significant information contained in the risk neutral density is not included in the objective. Hence, a direct measure represented by the ratio $P^{\mathrm{RN}}/P^{\mathrm{OBJ}}$ is a simple test for risk-neutrality. Also, Aït-Sahalia and Lo (2000) make the key insight that the coefficient of the Arrow–Pratt measure of absolute risk aversion—$\rho(F_T)$, say—can be expressed as $-\kappa_t'(F_T)/\kappa_t(F_T) = -U_T''(F_T)/U_T'(F_T) \equiv \rho(F_T)$, which finds empirical counterpart in the expression

$$\hat{\rho}(F_T) = \frac{\left(P_t^{\mathrm{OBJ}}\right)'(F_T)}{P_t^{\mathrm{OBJ}}(F_T)} - \frac{\left(P_t^{\mathrm{RN}}\right)'(F_T)}{P^{\mathrm{RN}}(F_T)}, \tag{16}$$

where primes indicate derivatives. In the following applications, first derivatives for both densities have been evaluated numerically.[18]

Fig. 6 shows the objective and the risk neutral densities estimated on January 20, 1997, while Fig. 7 reports the coefficient of absolute risk aversion (*ara*) obtained with formula (16) and the marginal rate of substitution (*mrs*)[19] obtained

[18] We can interpret the exercise of the present section as an experiment carried out in a partial equilibrium world, in which one investor holds bonds with maturity $T^{\wedge}$ until T ($T < T^{\wedge}$), sells them to another investor at T, and then consumes the receipts at T. While this interpretation is restrictive, it is also very helpful in assessing a fist order approximation of attitudes towards risk of market participants.

[19] Actually, *mrs* is a proxy for the marginal rate of substitution, since it does not disentangle the constant κ and the level of marginal utility. See formula (15).

as ratio of the risk-neutral to the objective PDF, as in Eq. (15); Figs. 8 and 9 show the analogous objects as of October 22, 1996. The different shapes and information contents of the two densities naturally reflect in the deviations of *ara* from zero and in the fact that *mrs* is not constant, which one would instead find under risk neutrality, i.e. under absence of risk premia; this is especially true for deep-in and deep-out of the money states, with discrepancies vanishing as long as the moneyness approaches unity. The estimated coefficients of absolute risk-aversion[20] at the two dates are slightly decreasing from out-of the-money to at-the-money states, which can in principle support the idea that investors are willing to pay less for insurance the higher their wealth. As of January 1997, the *ara* equals values as high as 0.3 for out-of-the-money states, then dropping toward zero as long as moneyness moves to unity (i.e. when F_t approximately equals 133); as of October 1996, this behavior is much more pronounced, with *ara* decaying very slowly towards nil to subsequently become largely negative for deep in-the-money states. In any case, the tail behavior of the absolute risk aversion is very different from the behavior in the remaining domain of the expected prices at expiration. Analogous indications come from the *mrs* being very variable across states (i.e. prices at expiration). These series of proxies for the *mrs* naturally embody indications on the expected returns from buying at time t an

[20] We remind some well-known basic facts, which may help the reader through the results presented below. Pratt (1964) shows that, given a utility function for money $u(x)$, the function $\rho(x) = -u''(x)/u'(x)$ is interpreted as a measure of local risk aversion or local propensity to insure (hence $-\rho(x)$ would measure propensity to gamble); it is a decreasing function of x if and only if the cash equivalent is a direct function of the economic agent's assets, and a negative function of the risk premium and the amount he would be willing to pay for insurance. Utility functions for which $\rho(x)$ is decreasing are candidates to use when describing the behavior of people who would pay less for insurance as far as their assets increase; however, there are no strong a-priori for not having utility functions for which $\rho(x)$ is first decreasing and then increasing. From an analytical standpoint, $\rho(x)$ is a also a measure of the concavity of $u(x)$ at x.
Starting from the definition of absolute risk aversion, that is $\rho(x) = \kappa = -u''(x)/u'(x) = -\mathrm{d}/\mathrm{d}x \log[u'(x)]$, it is easy to derive the utility associated to a given value of $r(x)$, by integrating, exponentiating and integrating again the above expression, i.e. with imprecise but usual notation, $-\int\rho = \log[u'] + c \Rightarrow u \sim \int e^{-\int\rho}$. If the local risk aversion is constant c, viz., $\rho(x) = c$ all x, then
$u(x) \sim x \text{ if } c = 0$
$u(x) \sim -\mathrm{e}^{-cx} \text{ if } c > 0$
$u(x) \sim \mathrm{e}^{-cx} \text{ if } c < 0.$
The three utilities are linear, strictly concave and strictly convex, respectively. Coming back to the economic meaningfulness of having decreasing absolute risk aversion (investors wish to gamble when their wealth increases), the utility $u(x) = -(b-x)^c, c > 1, x \leq b$, cannot have decreasing risk aversion. As Pratt (1964) shows, one can obtain decreasing and increasing absolute risk aversion for different ranges of x by considering the case where $u'(x) = (x^a + b)^{-c}(a > 0, c > 0)$, for which $r(x)$ is strictly decreasing when $x + bx^{1-a} \geq 0$. Pratt (1964) indicates a set of utility functions which have strictly decreasing risk aversion. Among these, $u(x) \sim \log(x+d), d \geq 0, u(x) = (x+d)^q, d \geq 0, 0 < q < 1$.

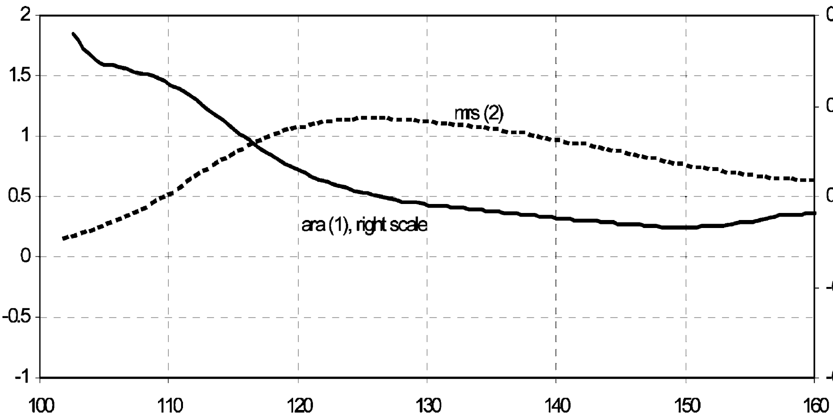

Fig. 7. January 20, 1997: marginal rate of substitution and absolute risk aversion. Note: $^{(1)}$*ara* is the coefficient of absolute risk aversion. It is obtained as $P'^{\mathrm{OBJ}}/P^{\mathrm{OBJ}} - P'^{\mathrm{RN}}/P^{\mathrm{RN}}$ where the symbol $'$ denotes derivatives and P^{RN} and P^{OBJ} are, respectively, the risk-neutral and the objective densities of Fig. 6. $^{(2)}$*mrs* proxies the marginal rate of substitution; it is obtained as $P^{\mathrm{RN}}/P^{\mathrm{OBJ}}$, where P^{RN} and P^{OBJ} are, respectively, the risk-neutral and the objective densities of Fig. 6.

Arrow–Debreu security paying one unit of money in a given state and selling it at time T (see, again, Aït-Sahalia and Lo, 2000); this type of information is fundamental in revealing the probability which the market attaches to the occur-

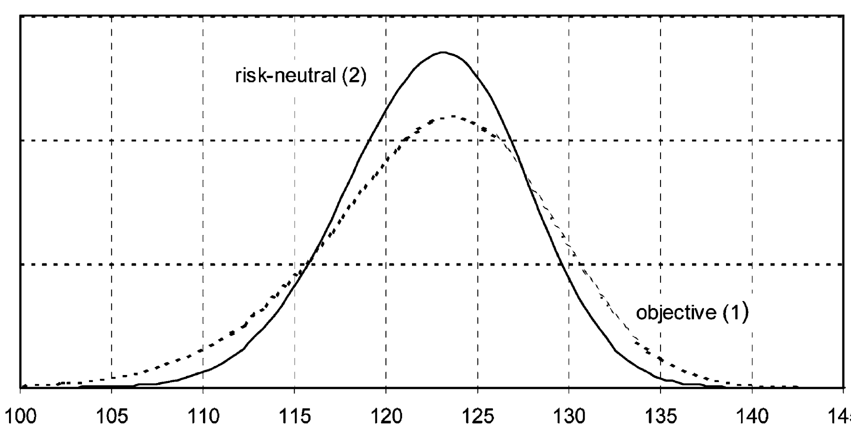

Fig. 8. October 22, 1996: objective and risk-neutral PDF. Note: $^{(1)}$This is the objective density function of the BTP futures price at October 22, 1996. It is obtained by simulating (5) with the continuous time parameters reported in Table 4. $^{(2)}$This is the risk neutral density function. It is obtained by simulating system (12) with the continuous time parameters reported in Table 4 and the volatility risk premium (13) with those reported in Table 6.

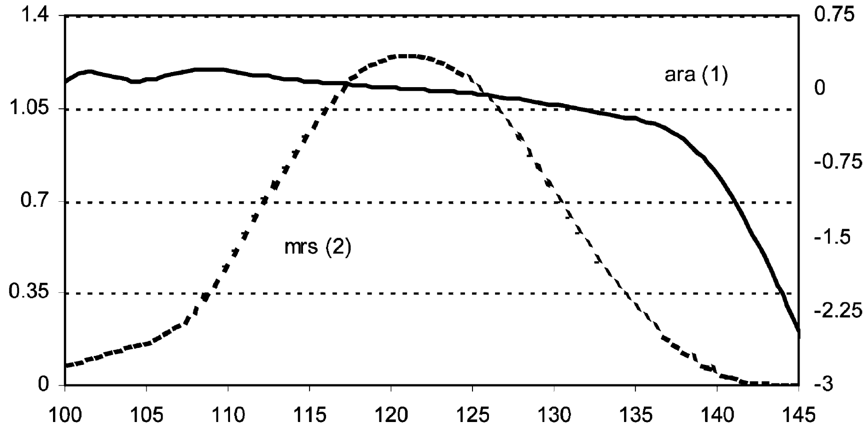

Fig. 9. October 22, 1996: marginal rate of substitution and absolute risk aversion. Note: [(1)]*ara* is the coefficient of absolute risk aversion. It is obtained as $P'^{\mathrm{OBJ}}/P^{\mathrm{OBJ}} - P'^{\mathrm{RN}}/P^{\mathrm{RN}}$ where the symbol $'$ denotes derivatives and P^{RN} and P^{OBJ} are, respectively, the risk-neutral and the objective densities of Fig. 8. [(2)]*mrs* proxies the marginal rate of substitution; it is obtained as $P^{\mathrm{RN}}/P^{\mathrm{OBJ}}$, where P^{RN} and P^{OBJ} are, respectively, the risk-neutral and the objective densities of Fig. 8.

rence of a given state which in turn helps to quantify the proper Value-at-Risk of a portfolio. In fact

$$r_\tau = \frac{P^{\mathrm{OBJ}}(F_T)}{\mathrm{e}^{-r\tau}P^{\mathrm{RN}}(F_T)} - 1, \tag{17}$$

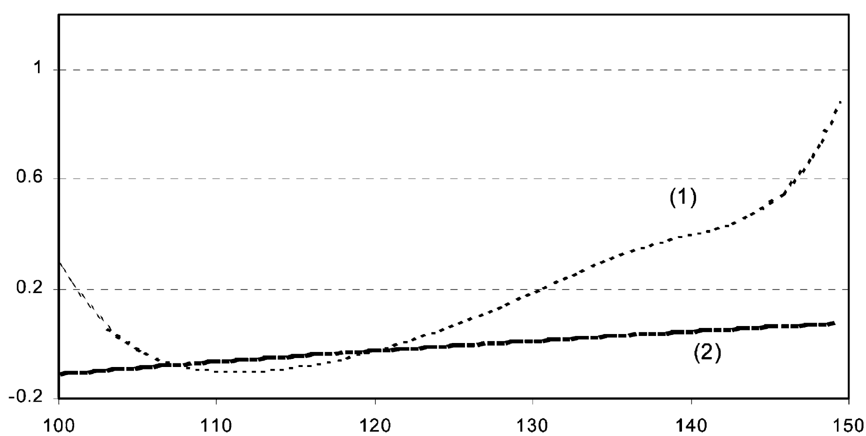

Fig. 10. January 20, 1997: expected returns of Arrow–Debreu securities. Note: [(1)]Expected returns of Arrow–Debreu securities evaluated according to the risk-neutral and the objective densities based on the simulation of Eqs. (5), (12) and (13), respectively, with the parameters reported in Tables 4 and 6. [(2)]Expected returns of Arrow–Debreu securities evaluated according to the objective and risk-neutral densities derived from the scheme of Black and Scholes.

is approximately the expected return of an Arrow–Debreu security written at t for the state F_T. It is then interesting to compare the indications coming out of the densities estimated in this paper to the analogous indications arising from the standard scheme of Black and Scholes. By fixing the standard deviation of the futures price changes to 9.1% per year (the historical value), we evaluated the objective and the risk-neutral PDF under the Black and Scholes assumption. Expected rates of return for various states of the futures price for both the Black and Scholes scheme and for model composed by Eqs. (5), (12) and (13) are reported in Fig. 10. They evidence, for the sample ending on January 20, 1997, remarkable differences between the prices to be paid for insurance against the occurrence of most of the price states reported there. Once again, the importance of an adequate characterization of the price and volatility dynamics emerges clearly.

5. Conclusions

This paper has shown how nonlinear Garch schemes can be effectively employed to approximate stochastic differential equations (SDE) typically adopted in finance as law of motion for the state variables. A practical application involved the estimation of the objective and risk-neutral densities of the BTP futures price. The objective PDF is obtained by first estimating a Power Arch(1,1) for the log-changes of the BTP price and then evaluating the continuous time equivalent of the estimated discrete time coefficients. Such parameters are then corrected via the indirect inference principle and subsequently employed to simulate price paths from the relevant SDE. The risk neutral density is simulated instead from the relevant SDE, which has the same parameters as the objective density, but includes a volatility risk premium which is a nonlinear function of the states (futures price and variability), whose parameters are chosen to fit observed cross sections of option prices as closely as possible. We have shown that the closed form relations of Nelson (1990) that provide the expression for the continuous time parameters are a reasonable approximation of the limiting process. This conclusion is more robust in the Power Arch(1,1) case, confirming the results obtained in a related work (Fornari and Mele, 2000a). This implies that in principle, one can employ daily data to recover the parameters according to which prices evolve in continuous time. We have determined the shape of the absolute risk aversion of market participants to BTP futures trading and have shown that risk averse behavior may emerge for out and in-the-money strike prices. This also provides information about economic agents' willingness to risk. With regard to the measurement of Value-at-Risk, we have shown that if economic agents rely on the Black and Scholes assumptions, insuring against the occurrence of a given state may be difficult.

Acknowledgements

We wish to thank the participants of the CEF99 Conference, Boston College, June 1999. Our gratitude goes to Andrea Beltratti, Jens Jackwerth, Loriana Pelizzon and a referee who raised helpful points. Responsibility for any views or errors contained in the paper rests only with the authors.

References

Aït-Sahalia, Y., 1996a. Nonparametric pricing of interest rate derivative securities. Econometrica 64, 527–560.

Aït-Sahalia, Y., 1996b. Testing continuous-time models of the spot interest rate. Review of Financial Studies 9, 385–426.

Aït-Sahalia, Y., Lo, A., 1998. Nonparametric estimation of state price densities implicit in financial asset prices. Journal of Finance 53, 499–547.

Aït-Sahalia, Y., Lo, A., 2000. Nonparametric risk management and implied risk aversion. Journal of Econometrics 94, 9–51.

Aït-Sahalia, Y., Wang, Y., Yared, F., 1998. Do option markets correctly price the probabilities of movements of the underlying asset? Journal of Econometrics, in preparation.

Bajeux, I., Rochet, J.C., 1996. Dynamic spanning: are options an appropriate instrument? Mathematical Finance 6, 1–16.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–659.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation of dynamic models with time varying covariances. Econometric Reviews 11, 143–172.

Broze, L., Scaillet, O., Zakoian, J.M., 1998. Quasi-indirect inference for diffusion processes. Econometric Theory 14, 161–186.

Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation of objective and risk-neutral measures for the purpose of options valuation. Journal of Financial Economics 56, 407–458.

Corradi, V., 2000. Reconsidering the diffusion limit of the GARCH(1,1) process. Journal of Econometrics 96, 145–153.

Ding, Z., Granger, C., Engle, R., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83–106.

Drost, F.C., Nijman, T.E., 1993. Temporal aggregation of GARCH processes. Econometrica 53, 385–407.

Drost, F.C., Werker, B.J.M., 1996. Closing the GARCH gap: continuous time GARCH modeling. Journal of Econometrics 74, 31–57.

Föllmer, H., Schweizer, M., 1991. Hedging of contingent claims under incomplete information. In: Davis, M., Elliott, R. (Eds.), Applied Stochastic Analysis. Gordon & Breach, New York.

Fornari, F., Mele, A., 1997. Weak convergence and distributional assumptions for a general class of nonlinear arch models. Econometric Reviews 16, 205–229.

Fornari, F., Mele, A., 2000a. An Equilibrium Model of the Term Structure with Stochastic Volatility, Working Paper, Université de Paris X-THEMA, 13.

Fornari, F., Mele, A., 2000b. Stochastic Volatility in Financial Markets—Crossing the Bridge to Continuous Time. Kluwer Academic Publishing, Boston.

Gallant, A.R., Tauchen, G., 1998. Reprojecting partially observed systems with applications to interest rate diffusions. Journal of the American Statistical Association 97, 10–24.

Gouriéroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, S85–S118.
Härdle, W., Linton, O., 1994. Applied nonparametric methods. In: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, vol. 4, Elsevier, Amsterdam.
Harrison, J.M., Kreps, D.M., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.
He, H., Leland, H., 1993. On equilibrium asset price processes. Review of Financial Studies 6, 593–617.
Heston, S., 1993. A closed form solution for options with stochastic volatility with application to bond and currency options. Review of Financial Studies 6, 327–344.
Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatility. Journal of Finance 42, 281–300.
Jackwerth, J.C., 2000. Recovering risk aversion from option prices and realized returns. Review of Financial Studies 13, 433–451.
Leblanc, B., 1995. Une Approche Unifiée pour une Forme Exacte du Prix d'une Option dans les Différents Modèles avec Volatilité Stochastique, INSEE Working Paper, No. 95–117.
Nelson, D., 1990. ARCH models as diffusion approximation. Journal of Econometrics 45, 7–38.
Nelson, D., 1992. Filtering and forecasting with misspecified ARCH models I: getting the right variance with the wrong model. Journal of Econometrics 52, 61–90.
Nelson, D., 1996. Asymptotic filtering theory for multivariate ARCH models. Journal of Econometrics 71, 1–47.
Nelson, D., Foster, D.P., 1994. Asymptotic filtering theory for univariate ARCH models. Econometrica 62, 1–41.
Pratt, J.W., 1964. Risk aversion in the small and in the large. Econometrica 32, 122–136.
Romano, M., Touzi, N., 1997. Contingent claims and market completeness in a stochastic volatility model. Mathematical Finance 7, 399–412.
Rosenberg, J., Engle, R.F., 1998. Empirical Pricing Kernels, New York University, Salomon Center, Working Paper, S-98-15.