BANCA D'ITALIA

Temi di discussione

del Servizio Studi

A Simple Approach to the Estimation of Continuous Time CEV Stochastic Volatility Models of the Short-Term Rate

by Fabio Fornari and Antonio Mele

Number 397 - February 2001

The purpose of the "Temi di discussione" series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

A SIMPLE APPROACH TO THE ESTIMATION OF CONTINUOUS TIME CEV STOCHASTIC VOLATILITY MODELS OF THE SHORT-TERM RATE

by Fabio Fornari * and Antonio Mele **

Abstract

Aim of this article is to judge the empirical performance of 'ARCH models as diffusion approximations' of models of the short-term rate with stochastic volatility. Our estimation strategy is based both on moment conditions needed to guarantee the convergence of the discrete time models and on the quasi indirect inference principle. Unlike previous literature in which standard ARCH models approximate only specific diffusion models (those in which the variance of volatility is proportional to the square of volatility), our estimation strategy relies on ARCH models that approximate any CEV-diffusion model for volatility. A Monte-Carlo study reveals that the filtering performances of these models are remarkably good, even in the presence of important misspecification. Finally, based on a natural substitute of a global specification test for just-identified problems designed within indirect inference methods, we provide strong empirical evidence that approximating diffusions with our models gives rise to a disaggregation bias that is not significant.

JEL classification: C15, E43, G12.

Keywords: stochastic volatility, CEV-ARCH, indirect inference, yield curve.

Contents

---

* Banca d'Italia, Research Department.

** Université de Paris X - Thema.

1. Introduction[1]

The increased importance played by conditional volatility in financial economics has led researchers (e.g., Hull and White, 1987; Wiggins, 1987; Longstaff and Schwartz, 1992; Heston, 1993) to extend early asset pricing theories (e.g., Black and Scholes, 1973; Merton, 1973; Vasicek, 1977) to the case in which volatility evolves in a stochastic manner. Empirically, time-varying volatility is well captured by the ARCH-type models introduced by Engle (1982) and Bollerslev (1986) (see, e.g., Bollerslev et al., 1994, for a survey). From a continuous time perspective, the initial contribution of Nelson (1990) established that some basic ARCH models can be reasonably considered as approximations of diffusion processes, which in turn are so frequently used to set up theoretical models; the major contribution of Nelson to this strand of research can be found in part II of the book edited by Rossi (1996).

The central objective of this article consists in extending the research agenda initiated by Nelson. As widely recognized, ARCH models are very appealing for statistical reasons, even though there exist alternative econometric formulations that are surveyed, for instance, in Ghysels et al. (1996) or in Shephard (1996). Despite the popularity of ARCH models and the celebrated work of Nelson, it is surprising that there has not yet been any empirical work assessing how well these models perform when they are taken as approximation to diffusion processes. As emphasized by Campbell et al. (1997, p. 381), the empirical properties of ARCH as approximations of continuous time stochastic volatility processes "have yet to be explored but will no doubt be the subject of future research". This is precisely what is attempted here.

Our focus is on formulating and estimating new stochastic volatility models of the short-term rate. Motivated by the above mentioned literature on the potential connection between ARCH models and the continuous time models that are typically used in finance, our primary concern lies in investigating whether ARCH-type models are useful devices to approximate

[1] This paper is a revised and extended version of Fornari and Mele (2000*b*). It was written while the first author was at the University of Cambridge and the second at Princeton University. We thank Gilles Dufrénot and Manfred Gilli for advice on numerical issues; Yacine Aït-Sahalia, Pippo Altissimo, Stephen Brown, Carl Chiarella, Ron Gallant, Michael Rockinger, José Scheinkman and seminar participants at Princeton University and Cambridge University, the 1998 Econometric Society European Meeting at Berlin and the 1999 Society for Computational Economics Conference at Boston College for helpful comments. We also thank three anonymous referees for valuable suggestions. The usual disclaimer applies: responsibility for any views or errors in the paper rests with the authors, who can be reached at the following e-mail addresses: *fornari.fabio@insedia.interbusiness.it* and *antonio.mele@u-paris10.fr*

and/or support the estimation of the parameters of stochastic differential equations. Unlike previous literature in which standard ARCH models approximated only specific diffusion models (namely, models in which the variance of volatility was proportional to the square of volatility), however, our estimation strategy relies on ARCH models that approximate any diffusion model for volatility with constant elasticity of variance (CEV henceforth), which we call CEV-ARCH models. To summarize, the class of models covered in this paper is a fairly general formulation that encompasses for example the continuous time version of the short-term rate model of Brenner et al. (1996).

2. Plan of the paper

Our first econometric objective is to make inference on the parameters of the stochastic differential equations which define our model. Of course, data are collected at discrete time points and it is well known that standard maximum likelihood (ML) techniques would not be suitable, since the likelihood function implied by the measure induced by our discretely sampled model is not known in closed-form. The econometric strategy that we implement is made up of two steps.

In the first one, we make use of the moment conditions that guarantee the weak convergence of ARCH models toward the theoretical model; in such a way we obtain 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 Nijman, 1993 and Drost and Werker, 1996), in a second step we test and correct potential 'disaggregation' biases using ARCH models viewed as *auxiliary* devices in simulation-based (indirect inference) schemes (see Gouriéroux and Monfort, 1996, for a full account of simulation-based inference methods). In applying such a research strategy to 3-month US Treasury Bill rates, we find that the correction made by indirect inference methods is not statistically significant. Such a result is obtained via a global specification test for just-identified models that was originally suggested by Gouriéroux et al. (1993).[2] Our empirical findings are obtained with the data set used in a frequently cited empirical work of Andersen and Lund (1997a). The authors make use of the efficient method of moments

[2] We obtain very similar findings in a companion paper (Fornari and Mele, 2001) in the stochastic volatility option pricing area. Naturally, the empirical success ARCH models have in approximating diffusion processes here does not invalidate simulation-based methods. On the contrary, exploring the validity of ARCH as approximators of diffusion processes has been possible due to the availability of simulation-based techniques.

(EMM) techniques developed by Gallant and Tauchen (1996), in which a highly parametrized discrete time model is used with the main purpose of calibration; precisely, the auxiliary model generates a score, and the EMM objective is then to minimize a chi-squared criterion that is a quadratic form in the expected score computed via a long simulation of the theoretical model. The advantage of the EMM estimator is that it achieves the same efficiency as the true (intractable) ML estimator when the auxiliary model generates a density that 'smoothly embeds' the true likelihood function of the discretely sampled diffusion. Following the results of Gallant and Long (1997), one can use a semi-nonparametric-based likelihood function to provide the additional parameters that increment the efficiency of the EMM estimator. One of the earliest applications of the EMM techniques to models of the stock prices with continuous time stochastic volatility is in Gallant and Tauchen (1997).[3]

It should be clear that the estimation strategy that we follow has a different rationale: instead of selecting a highly parametrized auxiliary model that has the scope of calibration, we just wish to ascertain whether our auxiliary model is a reasonable approximation of the continuous time model. In technical terms, we are going to focus on the empirically difficult just-identified case. Such a strategy was originally suggested in 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". In our context, indeed, "the asymptotic bias implied by the approximation" is given by a disaggregation bias. While not closed under temporal aggregation, ARCH models still have a natural interpretation in terms of the continuous time models that they approximate, since they are very close (in terms of the probability distributions generating them) to the approximated continuous time models when the sampling frequency is high. Furthermore, the auxiliary criteria that we construct are based on approximations that create a natural one-to-one interpretation of the sequence of the parameters of the auxiliary discrete time model in terms of the parameters of the continuous time model (see paragraph 4): as is clear, we are exactly in the situation originally put forward by Gouriéroux et al. (1993).

---

[3] Gallant and Tauchen (1997) also consider the application of EMM to interest rates models without stochastic volatility, while Gallant et al. (1997) apply the EMM technique to discrete time models with stochastic volatility.

In addition to the point estimates of the parameters of stochastic differential equation system, an essential ingredient of the practical implementation of any stochastic volatility model is obviously the knowledge of the volatility at the dates of interest. In pricing bonds in a stochastic volatility setting, for instance, one needs volatility estimates. Clearly, this is a challenging problem since the short-term rate volatility is not directly observable, and especially in continuous time, it is not trivial to obtain filtered estimates of the unobservable volatility; see however, the reprojection techniques implemented by Gallant and Tauchen (1998) in recent empirical work. In this respect, appropriate sequences of ARCH models are known to estimate consistently the volatility of a 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; Bollerslev and Rossi, 1996, (p. xiii-xvii) for a brief account on the filtering performances of ARCH models as applied to continuous time stochastic volatility models). In this case, as put by Bollerslev and Rossi (1996), “one could regard the ARCH model as merely a device which can be used to perform filtering or smoothing estimation of unobserved volatilities” (p. xiv). In addition, our motivation to use ARCH-type models to filter volatility is reinforced here, since we show that the desiderable filtering performances of *standard* ARCH models are also shared by the CEV-ARCH models, as it might be expected by a suitable interpretation of the theory (see Nelson and Foster, 1994, theorem 4.1). In a nutshell, the approach suggested in this paper allows one to filter volatility efficiently in any CEV-diffusion model for volatility.

The practical relevance of the filtering theory for ARCH models can be grasped very simply. Figure 1 depicts the typical filtering of an ARCH model as applied to a simplified version of our model. There, the straight line is one weekly sampled trajectory of the volatility $\sigma(\tau)$ simulated within the following model:

$$(1) \qquad \begin{cases} dr(\tau) &= (\iota - \theta \cdot r(\tau))d\tau + \sqrt{r(\tau)} \cdot \sigma(\tau) \cdot dW^{(1)}(\tau) \\ d\sigma(\tau) &= (\omega - \varphi \cdot \sigma(\tau))d\tau + \psi \cdot \sigma(\tau) \cdot dW^{(2)}(\tau) \end{cases}$$

where $W^{(i)}$, $i = 1, 2$, are standard Brownian motions, and $\iota, \theta, \omega, \varphi$ and ψ are real-valued parameters fixed at their estimates obtained with US data (see paragraph 5). The dotted line represents instead the (rescaled) volatility obtained via an ARCH model fitted to the weekly sampled trajectory of the short-term rate $r(\tau)$, as simulated by (1); of course, in estimating the ARCH model, we considered ourselves constrained to *only* knowing the realization of the simulated $r(\tau)$. In fact, figure 1 visualizes one of the simulations performed in the Monte Carlo experiment of paragraph 5, but such a performance is typical of the overall experiment; this can be gauged by the very tiny RMSE between the two trajectories computed over all the simulations (see section 5 and Schwartz et al., 1993, for previous related work on similar models).

The final contribution of the paper is to explore a term-structure extension of the model. The main objective is to understand the relationship between equilibrium bond prices and volatility within the framework of our estimated stochastic volatility model. While it is well known that a three-factor model is needed to explain level, slope and curvature of the term-structure (see, e.g., Andersen and Lund, 1997b, and Dai and Singleton, 2000), it will be argued that our two-factor model is already able to capture fundamental qualitative features of the relationship between bond prices and volatility, which is the only objective pursued here. Similar exercises were already performed by Chen (1996) and Andersen and Lund (1997b). These authors did not emphasize how to determine the risk-premia associated with the fluctuations of the uncertainty factors. To address this issue within a theoretically sound framework, we then study the compatibility of our data generating process with an equilibrium in which agents are endowed with a CRRA utility function. Our empirical results then imply that the term-structure of interest rates increases with volatility.

The paper is organized as follows. Next section presents the basic structure of our continuous time model; it also provides intuition and preliminary results on the estimation and filtering methods to be implemented with the help of ARCH models that do not constrain the elasticity of variance to one (the "CEV-ARCH models"). The econometric strategy is fully detailed in paragraph 4. Empirical results are in paragraph 5 and 6. Technical considerations and proofs are gathered in the appendices.

3. Modeling volatility as a CEV process

The basic structure of the continuous time models

The most salient feature of the model we consider in this paper is that the instantaneous volatility of the short-term rate is a process with constant elasticity of variance (CEV):

$$(2)\quad \begin{cases} dr(\tau) &= (\iota - \theta r(\tau))d\tau + \sigma(\tau)\sqrt{r(\tau)}dW^{(1)}(\tau) \\ d\sigma(\tau)^{\delta} &= (\omega - \varphi\sigma(\tau)^{\delta})d\tau + \psi\sigma(\tau)^{\delta\cdot\eta}d\left(\rho W^{(1)}(\tau) - \sqrt{1-\rho^2}W^{(2)}(\tau)\right), \end{cases}$$

where $a = (\iota, \theta, \delta, \omega, \varphi, \psi, \eta, \rho)$ is the parameter vector of interest, $W^{(i)}$, $i = 1, 2$, are standard Brownian motions, and $\delta \geq 1$.The $\sqrt{r(\cdot)}$-term included in the short-term rate diffusion equation constrains the short-term rate to take only positive values. With such a term, the model also captures an empirical regularity known as the 'level effect', i.e., *coeteris paribus*, the short-term rate volatility gets higher as the short-term rate level increases. Allowing for more general diffusion terms such as for instance $\sigma(\cdot)\,|r(\cdot)|^{d}$ $(d \geq 1/2)$ is possible, though it would not change dramatically our empirical results.

The central objective of the paper is to use ARCH-type models that allow for i) estimation of the continuous time parameters and ii) reconstruction of the unobserved short-term rate volatility process $\sigma(\cdot)$. A technical presentation of our methodology as opposed and/or related to other existing methodologies is deferred to the next paragraph. Here we give an heuristic motivation of the approach followed in this paper as well as preliminary evidence on its performances.

A class of ARCH models: the CEV-ARCH

Consider an Euler-Maruyama discrete time approximation of (2):

$$(3)\quad \begin{cases} {}_h r_{h(k+1)} - {}_h r_{hk} &= (\iota - \theta \cdot {}_h r_{hk})h + {}_h\sigma_{hk}\sqrt{{}_h r_{hk}} \cdot {}_h u_{h(k+1)} \\ {}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} &= (\omega - \varphi \cdot {}_h\sigma^{\delta}_{hk})h + \psi \cdot {}_h\sigma^{\delta\cdot\eta}_{hk}\sqrt{h}\,{}_h\widetilde{\xi}_{h(k+1)} \end{cases}$$

where h denotes the discretization step,

$$\begin{pmatrix} {}_h u_{hk} \\ {}_h\widetilde{\xi}_{hk} \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} h & \sqrt{h}\rho \\ \sqrt{h}\rho & 1 \end{pmatrix}\right),$$

and $({}_h r_{hk}, {}_h\,\sigma_{hk})_{k=1}^{\infty}$ are the discretized short-term rate and volatility processes.

It is well known that when $h \downarrow 0$ (3) converges weakly (or in distribution) to (2).[4] Hence, the higher the sampling frequency, the higher should be the accuracy of, say, ML estimates of a obtained with (3). Unfortunately (3) represents a discrete time stochastic variance model for which ML methods are quite cumbersome to implement. Used as an auxiliary model in a simulation-based framework, for instance, (3) would noticeably increase the computational burden. Considered as a potentially good approximation of (2), to mention a further example, (3) would lend itself to a computationally intensive testing strategy for the disaggregation bias. Third, there are no obvious techniques to filter out the actual volatility path with (3).

A natural alternative is provided by ARCH models. As noted in the introduction, ARCH models can be thought of as diffusion approximations. It is also well known, however, that not all diffusion models can be approximated by ARCH models. To get an intuition of this, consider the standard GARCH(1,1) model of Bollerslev (1986):

$$\sigma_{n+1}^2 = w + \beta\sigma_n^2 + \alpha\epsilon_n^2,\ \epsilon_n \equiv (u \cdot \sigma)_n,\ n = 0, 1, \cdots$$

where w, β and α are parameters, ϵ is the residual of an observation equation, and the index n is an abstract notation for sample points at discrete time intervals (a more precise notation will be introduced in the next section). Rewrite the preceding equation as:

$$\sigma_{n+1}^2 - \sigma_n^2 = w - \left(1 - \alpha E(u^2) - \beta\right)\sigma_n^2 + \alpha\sigma_n^2\left(u_n^2 - E(u^2)\right), \tag{4}$$

and suppose that $u \sim N(0,1)$. When we chop time so as to make $n : hk \leq n \leq h(k+1)$, $k = 1, 2, \cdots$, let the parameters w, β, α vary with h by introducing sequences w_h, β_h, α_h, and then let $h \downarrow 0$, the resulting volatility process converges in distribution to:

$$d\sigma(\tau)^2 = \left(\omega - \varphi\sigma(\tau)^2\right)d\tau + \psi\sigma(\tau)^2 dW^{(2)}(\tau), \tag{5}$$

[4] If (2) has a unique strong solution denoted as $\{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$, *weak convergence* of $\{{}_h r_{hk}, {}_h\,\sigma_{hk}^{\delta}\}_{k=1,2,\ldots}$ in (3) to $\{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$ means that the finite dimensional distributions of $\{{}_h r_{hk}, {}_h\,\sigma_{hk}^{\delta}\}_{k=1,2,\ldots}$ converge to those of $\{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$ as $h \downarrow 0$. See Stroock and Varadhan (1979). It turns out that the conditions demanded by Stroock and Varadhan (1979) are difficult to verify when studying the convergence of ARCH-type models. One then may wish to make reference to the conditions suggested by Nelson (1990).

where

$$\text{(6)} \qquad \begin{cases} \lim_{h\downarrow 0} h^{-1} w_h & = \omega \\ \lim_{h\downarrow 0} h^{-1}\left(1-\alpha_h-\beta_h\right) & = \varphi \\ \lim_{h\downarrow 0} h^{-1/2}\sqrt{2}\alpha_h & = \psi. \end{cases}$$

To obtain an intuition of this result, notice that the sequence $(\xi_n)_{n=1}^{\infty} \equiv (u_n^2 - E(u^2))_{n=1}^{\infty}$ is an i.i.d. sequence of centered chi-square variates with one degree of freedom and represents the discrete version of the Brownian motion increments $dW^{(2)}(\cdot)$. On the other side, the renormalizing $\sqrt{2}$-term in (6) is explained by the fact that $\xi = u^2 - E(u^2) = u^2 - 1$ is a chi-square variate with one degree of freedom and has a variance equal to two. Naturally, the normality assumption for u is not needed to obtain the convergence.

Equation (5) may correspond to the volatility dynamics in (2) when $\delta = 2$, $\eta = 1$ and $\rho = 0$. Similarly, it is possible to show that under conditions similar to (6), the so called Taylor-Schwert model:

$$\sigma_{n+1} - \sigma_n = w - \left(1 - \alpha E(|u|) - \beta\right)\sigma_n + \alpha\sigma_n\left(|u_n| - E(|u|)\right),$$

also converges in distribution to a diffusion limit with the following form:

$$\text{(7)} \qquad d\sigma(\tau) = (\omega - \varphi\sigma(\tau))d\tau + \psi\sigma(\tau)dW^{(2)}(\tau).$$

Equation (7) may now correspond to the volatility dynamics of (2) when $\delta = \eta = 1$ and $\rho = 0$.

As these two basic examples should make clear, standard ARCH models do not converge in distribution to any unrestricted CEV process. Rather, in their diffusion limit, ARCH models typically make the variance of volatility proportional to the square of volatility, thus restricting the elasticity of variance to unity. Motivated by this simple remark, we now describe a class of ARCH models that does not constrain the elasticity of variance to one.[5] Consider, for instance, the following model:

$$\text{(8)} \qquad \sigma_{n+1}^2 = w + \alpha\sigma_n^{2\eta}\left|u_n\right|^{2\eta} + \beta\sigma_n^2 + \alpha E(|u|^{2\eta})\left(\sigma_n^2 - \sigma_n^{2\eta}\right),$$

[5] This class of models can be shown to satisfy the most salient theoretical properties of an optimal volatility filter as developed earlier in the optimal filtering theory of Nelson and Foster (1994, theorems 4.1 and 5.2).

which can also be rewritten as:

$$\sigma_{n+1}^2 - \sigma_n^2 = w - \left(1 - \alpha E(|u|^{2\eta}) - \beta\right)\sigma_n^2 + \alpha\sigma_n^{2\eta}\left(|u_n|^{2\eta} - E(|u|^{2\eta})\right).$$

Clearly, this model collapses to the GARCH(1,1) model (4) when $\eta = 1$; yet it does not constrain η to that value: rather, η must be estimated from data. Furthermore, in the next section we show that under conditions similar to those of Nelson, this model converges in distribution to:

$$d\sigma(\tau)^2 = \left(\omega - \varphi\sigma(\tau)^2\right)d\tau + \psi\sigma(\tau)^{2\eta}dW^{(2)}(\tau).$$

Finally, to obtain convergence results closer to model (2), we shall be considering a generalization of (8) that sets the volatility propagation mechanism to:

$$\sigma_{n+1}^\delta = w + \alpha\sigma_n^{\delta\eta}\left|u_n\right|^{\delta\eta} + \beta\sigma_n^\delta + \alpha E(|u|^{\delta\eta})\left(\sigma_n^\delta - \sigma_n^{\delta\eta}\right). \tag{9}$$

As before, we will show that at a high sampling frequency, the volatility process in (9) converges in distribution to

$$d\sigma(\tau)^\delta = \left(\omega - \varphi\sigma(\tau)^\delta\right)d\tau + \psi\sigma(\tau)^{\delta\cdot\eta}dW^{(2)}(\tau),$$

which may correspond to the volatility dynamics in (2) when $\rho = 0$. Complications arising from the presence of correlation will be treated by introducing asymmetries in the volatility dynamics of (9). In the same way, our searching strategy can be used to introduce nonlinear volatility dynamics into discrete time models that match any desidered feature of the resulting diffusion limit. Consider, for instance, the following model:

$$\sigma_{n+1} = (1+w)\sigma_n - (1 - \alpha E(|u|) - \beta)\,\sigma_n^2 + \alpha\left(|u_n| - E(|u|)\right)\sigma_n^{3/2}.$$

Using the methods of paragraph 4, it can then be shown that this model converges in distribution towards:

$$d\sigma(\tau) = \{\sigma(\tau)\left(\omega - \varphi\sigma(\tau)\right)\}d\tau + \psi\sigma(\tau)^{3/2}dW^{(2)}(\tau),$$

as the sampling frequency gets higher and higher. Likewise, one can adjust both the short-term and the volatility equation to include both variables. In this paper, however, we will only test the adequacy of ARCH-type models in the estimation and filtering of system (2).

Filtering and invariance properties of the CEV-ARCH model: preliminary Monte Carlo evidence

Here we provide preliminary Monte Carlo evidence on the performance of the CEV-ARCH models previously introduced. Further Monte Carlo *and* empirical evidence will be provided in section 5. Our primary concern lies here in ascertaining whether the CEV-ARCH model (9) is able to deliver reliable parameter and filtering estimates. To this purpose, we consider model (2) and fix $\rho = 0$ (consistently with the empirical evidence provided in section 5), and the other coefficients at the values indicated in table 1 below. We then simulate (2) 1000 times with an Euler-Maruyama approximation device, and sample simulated data at a weekly frequency. All simulated samples have 1135 weekly points, which correspond to the actual sample size used in our empirical analysis (see section 5). Finally, all weekly simulated short-term rate data are fitted by a conditionally Gaussian AR(1) model of the form:

$$r_n = \phi_0 + \phi_1 r_{n-1} + \sqrt{r_{n-1}}\epsilon_n \quad (\phi_0,\, \phi_1 \text{ constants}),$$

with (9) as volatility propagation equation.

Table 1 reports the results of the experiment. We begin with the case related to the empirical evidence provided in section 5: there, we find that fitting (9) to actual US short-term rate data produces estimates of δ and η that are both statistically not distinguishable from 1. Now table 1 shows that when the data generating process in (2) has $\delta = \eta = 1$, then (9) also reproduces, on average, approximately the same ML estimates of δ and η (see section 5 for the implementation of experiments involving all parameters). Results not reported here reveal that the same phenomenon occurs with other possible combinations of δ and η. As an example, table 1 reports Monte Carlo results concerning the case in which $\delta = 2$ and $\eta = \frac{1}{2}$ in (9). This case emerges when the data generating process in (2) has a variance concept that follows a square root process. As is clear from table 1, model (9) plays in practice an excellent role in mimicking such characteristics of the data generating process. We call these preliminary properties as time-scale invariance properties of δ and η. Clearly, such properties should not be shared by the other parameters of the models: it is precisely the objective of section 4 to provide the necessary correction formulae introducing time-scale corrections.

The filtering performances of model (9) are also remarkably good. Volatility trajectories filtered with this equation are very close to volatility trajectories simulated from (2), and the resulting pattern of the two trajectories is very similar to the one shown in figure 1. Table 1 reports precise results assessing the volatility filtering performance of (9), by comparing simulated volatility paths with volatility paths filtered with (9). The common concept of volatility adopted to make comparisons is the standard deviation. The result is what we call the ‘volatility filtering error’, which is defined precisely in paragraph 5. The findings reported in table 1 are of the same order of magnitude as those of paragraph 5. Notice also that in order to be able to compare simulated with filtered volatility, the latter has to be rescaled for diffusions; techniques for treating this issue are introduced and explained in great detail in appendix C.

On probabilistic properties of the short-term rate volatility and comparisons with alternative formulations

Beyond providing a framework for CEV-type diffusion process for volatility, (2) differs significantly from previous stochastic volatility models, since it does not constrain the ‘volatility concept’ to be ‘variance’ or ‘standard deviation’; rather, in (2) δ is a new parameter that must be estimated from data. In the empirical section of the paper, for instance, we uncover evidence that $\delta \cong 1$. In that section, we also uncover evidence that $\eta \cong 1$.[6] With $\eta = 1$ and positive mean-reversion, the volatility process σ^{δ}, $\delta \geq 1$, has a steady state distribution that is an inverted Gamma with mean $\frac{\overline{\omega}}{\varphi}$ (e.g., lemma 3.1 p. 217 in Fornari and Mele, 1997a); the stationary distribution of σ is consequently given by

$$f_{\delta}(\sigma) \equiv \frac{\delta \cdot \left(\frac{2\omega}{\psi^2}\right)^{\frac{2\varphi+\psi^2}{\psi^2}}}{\Gamma\left(\frac{2\varphi+\psi^2}{\psi^2}\right)} \sigma^{-\frac{2\delta\varphi+(\delta+1)\psi^2}{\psi^2}} \exp\left(-\frac{2\omega}{\psi^2}\sigma^{-\delta}\right) \tag{10}$$

(see lemma A.2, p. 227, in Fornari and Mele, 1997a). As shown by Fornari and Mele (2000a) (chapter 5), the density $f_{\delta}(\cdot)$ tends to shrink to the left as δ decreases.

[6] Engle and Lee (1996) fitted a restricted version of the volatility equation of model (2) to stock returns, namely for $\delta = 2$, and supported a model in which the volatility of volatility raised linearly with the square of volatility, as our empirical findings do.

The volatility equation in (2) encompasses other formulations already encountered in the stochastic volatility literature (see, for instance, Ball and Roma, 1994, and Taylor, 1994, for a list of the typical models in the stochastic volatility option pricing area). This is the case, for instance, of the non-stationary models of Hull and White (1987) or Johnson and Shanno (1987), to which our volatility equation reduces when $\omega \equiv 0$: by Itô's lemma, indeed, $\mathcal{V} \equiv \log \sigma^2$ is solution of

(11)
$$d\mathcal{V}(\tau) = \left(-\frac{2\varphi + \psi^2}{\delta} + 2\frac{\omega}{\delta}\exp\left(-\frac{\delta}{2}\mathcal{V}(\tau)\right)\right)d\tau + \frac{2\psi}{\delta}d\left(\rho W^{(1)} + \sqrt{1-\rho^2}W^{(2)}\right).$$

In contrast, log-volatility mean-reverts in a *non-linear* manner when $\omega \neq 0$. Therefore, (11) is rather different from the *linear* mean-reverting process for the log-volatility adopted in the seminal paper of Wiggins (1987) in the stochastic volatility option pricing domain, and in the empirical work of Andersen and Lund (1997a) or Gallant and Tauchen (1998) in the interest rates field. To see this in more detail, consider the linear mean reverting model utilized in the study of Andersen and Lund,

$$d\mathcal{V}(\tau) = \left(\overline{\alpha} - \overline{\beta}\mathcal{V}(\tau)\right)d\tau + \overline{\xi}dW(\tau)$$

where W is a standard Brownian motion and $\overline{\alpha}, \overline{\beta}, \overline{\xi}$ are real constants. By Itô's lemma, in this model σ^δ is the solution of

$$(12) \qquad d\sigma(\tau)^\delta = \left(\frac{4\overline{\alpha}\delta + \overline{\xi}^2\delta^2}{8}\sigma(\tau)^\delta - \overline{\beta}\sigma(\tau)^\delta \cdot \log\sigma(\tau)^\delta\right)d\tau + \frac{\overline{\xi}\delta}{2}\sigma(\tau)^\delta dW(\tau),$$

which becomes of course also the starting point of Wiggins (1987 eq. (2) p. 353 and eq. (15) p. 361) when $\delta \equiv 1$. Although the volatility of volatility in (12) rises linearly with $\sigma^{2\delta}$, as in (2) when $\eta = 1$, the drift behaves rather differently in the two volatility equations.

Figure 2 depicts a comparison between the stationary densities that are generated by (11) and (12). The first is given by (10) and has been produced using the parameters estimates of section 5; the latter is just a log-normal density, and has been produced using the parameters estimates reported in Andersen and Lund (1997b). While the two models approximately put the same probability masses on low levels of volatility, our model puts relatively more masses on high values of volatility than the Andersen-Lund model. An explanation of such a phenomenon can be found by comparing the drift functions of the two models: as is clear from figure 3, the two drift functions are of the same order of magnitude when volatility is low; once volatility visits higher regions, however, the Andersen-Lund linear drift function pulls volatility towards its steady state expected value more rapidly than the drift function of our model. This implies that our model generates relatively more frequent episodes of high volatility than the Andersen-Lund model. Naturally, our model does not encompass the Andersen-Lund model, but it should be more flexible in practice due to the presence of the additional parameter δ in the volatility equation: should the volatility equation in (2) be misspecified, such an additional parameter might permit to better adjust the model to the statistical properties of the true volatility generating mechanism.

4. Statistical inference

Various methods have been recently proposed to estimate the parameters of a diffusion when sampling is not continuous. As is well known, the main difficulty of ML methods is that the likelihood function implied by the measure induced by a discretely sampled diffusion can not be calculated explicitly.[7] Alternative methods rely on nonparametric density estimation (Aït-Sahalia, 1996a and 1996b) and/or closed-form approximations of the true (unknown) likelihood function of the discretely sampled diffusion (Pedersen, 1995; Aït-Sahalia, 2000), on the generalized method of moments (Hansen and Scheinkman, 1995; Conley et al., 1997), or on the indirect inference principle,[8] whose main references have been cited in the introduction.

In this paper, we adopt the indirect inference principle, which is particularly well-suited to problems in which the state is partially observed. One important concern, however, is also to study whether a simple ARCH model fitted to high frequency data provides a reasonable approximation of (2). Accordingly, in section 4 we start with presenting a vary basic approach to obtain an initial estimate of the vector of parameter of interest, a. It consists in replacing the (intractable) likelihood function implied by the true measure induced by (2) with an approximation of it. Such an approximation may be based on a discretization of (2),

---

[7] Following Lo (1988), ML estimation might turn out to be feasible if the transition density of $\{r(\tau)\}_{\tau\geq 0}$ in (2) could be computed easily. Since this is not the case here — as in virtually all continuous time stochastic volatility models — ML is computationally demanding, since it 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 recovered by integrating out with respect to volatility.

[8] In this paper, we adopt the convention to include the EMM theory of Gallant and Tauchen (1996) as a part of the indirect inference principle.

but as the arguments of the previous paragraph should have clarified, even a standard Euler approximation of (2) yields a discrete time stochastic variance model, and eventually implies an approximated likelihood that does not simplify the problem in a noticeable way.[9]

One natural alternative is to make use of a (tractable) exact likelihood function of a class of approximated models. The main idea has been presented in the previous paragraph, and consists in resorting to a suitably chosen class of ARCH models converging in distribution to the solution of (2) as the sampling frequency gets infinite, following the strand of literature which shows the convergence in distribution of ARCH-type models to diffusion processes. Since the resulting likelihood function refers to a model converging in distribution to the solution of (2) that is *not* an Euler approximation of (2), however, we call the resulting criterion 'quasi'-approximated likelihood function.

The advantage of the quasi-approximated ML estimator is that it demands no computational efforts, and its main drawback is that it is not necessarily consistent. In fact, the ARCH models we use are typically not closed under temporal aggregation, which means that at least theoretically, there is not a one-to-one correspondence between convergence in distribution of the discrete time models and disaggregation from a diffusion. On a theoretical standpoint, such a correspondence exists only when the concept of an ARCH model is weakened (Drost and Nijman, 1993, and Drost and Werker, 1996).[10] Furthermore, Corradi (2000) recently criticized the conditions in Nelson (1990) , necesary to achieve the convergence of the basic GARCH(1,1) to a diffusion; in section 4, we show how to adapt Corradi's critique to our setup.

Recognising the presence of disaggregation bias and the possibility that the ARCH models we use may even fail to converge to any diffusion limit, in section 4 we show how to construct a very precise testing procedure of the validity of the moment conditions needed to guarantee the convergence to well-defined diffusion limits; as it turns out, such a testing procedure also gives information about the relevance of disaggregation biases. Our strategy

---

[9] See, for instance, Harvey et al. (1994) or Jacquier et al. (1994) for the estimation of discrete time stochastic variance models, Jacquier et al. (1999) for multivariate and distributional extensions, Gallant et al. (1997) for the EMM approach to discrete time stochastic variance models, and Shephard (1996) for a succinct survey of related methods.

[10] Drost and Werker (1996, p. 33) report that using ARCH as indirect approximators should be more efficient than using their identification procedures, since in this case the criterion function would be close to the true ML equations.

is based on the consistency test originally suggested by Gouriéroux et al. (1993, section 4.2), and it can be viewed as the natural substitute of a global specification test in just-identified problems.

Quasi-approximated likelihood functions

The rationale behind the quasi-approximated ML estimator that we propose lies in the weak convergence of a class of ARCH models towards the solution of (2). For ease of exposition, we start with considering the restricted version of (2) that sets $\eta \equiv 1$. Theorem 4.2 below treats the general case. With $\eta = 1$, a model approximating (2) can be a discrete time approximation of the short-term rate equation in (2) modified by introducing the so-called asymmetric-power ARCH model introduced by Ding et al. (1993):

$$
\text{(13)} \qquad \begin{cases} {}_\Delta r_{n+1} & = \; {}_\Delta r_n + \iota_\Delta - \theta_\Delta \cdot {}_\Delta r_n + {}_\Delta \sigma_{n+1} \sqrt{{}_\Delta r_n} \cdot {}_\Delta u_{n+1} \\ {}_\Delta \epsilon_n & = \; {}_\Delta u_n \cdot {}_\Delta \sigma_n, \quad \frac{{}_\Delta u_n}{\sqrt{\Delta}} \sim N(0,1) \\ {}_\Delta \sigma^\delta_{n+1} & = \; w_\Delta + \alpha_\Delta (|{}_\Delta \epsilon_n| - \gamma \cdot {}_\Delta \epsilon_n)^\delta + \beta_\Delta \cdot {}_\Delta \sigma^\delta_n \end{cases}
$$

where the indexing $n = 0, 1, \cdots$ refers to consecutive observations sampled at the same frequency Δ (weekly, say), $\iota_\Delta, \theta_\Delta, w_\Delta$ are of the form $x_\Delta = x^{(\Delta)} \cdot \Delta$, with $\iota^{(\Delta)}, \theta^{(\Delta)}$ real parameters and $w^{(\Delta)} > 0, \alpha_\Delta, \beta_\Delta \geq 0, \gamma \in (-1, 1), \delta > 0$. Finally, γ allows for the leverage effect originally observed by Black (1976), and incorporated by Nelson (1991) in ARCH-type models. To keep things relatively simple, we assumed a sort of time-scale invariance of (δ, γ) in the preceding approximation scheme. The invariance of δ is, however, strongly supported by the Monte Carlo experiments reported in paragraph 3.

To heuristically obtain the weak convergence towards the solution of (2), chop time as $hk \leq n \leq h(k+1)$:

$$
\text{(14)} \qquad \begin{cases} {}_h r_{h(k+1)} & = \; {}_h r_{hk} + \iota_h - \theta_h \cdot {}_h r_{hk} + {}_h \sigma_{h(k+1)} \sqrt{{}_h r_{hk}} \cdot {}_h u_{h(k+1)} \\ {}_h \epsilon_{hk} & = \; {}_h u_{hk} \cdot {}_h \sigma_{hk}, \quad \frac{{}_h u_{hk}}{\sqrt{h}} \sim N(0,1) \\ {}_h \sigma^\delta_{h(k+1)} - {}_h \sigma^\delta_{hk} & = \; w_h - (1 - \alpha_h \left| {}_h u_{hk} \right|^\delta (1 - \gamma s_k)^\delta h^{-\frac{\delta}{2}} - \beta_h) {}_h \sigma^\delta_{hk} \end{cases}
$$

(with $s_k = \text{sign}({}_h u_{hk})$ and, $\forall h > 0$, $(\{\iota_h\}, \{\theta_h\}\{w_h\}, \{\alpha_h\}, \{\beta_h\}) \in \mathbb{R}^5_+$ and $\gamma \in (-1, +1)$), and impose suitable Lipschitz conditions on the 'h-drift' as well as non-explosion conditions on the '$\sqrt{h}$-diffusion' terms of volatility.

Nelson (1996, p. 19) was one of the first to suggest a model of the kind of (14) as a discrete time approximation of a continuous time model for the short-term rate. More specifically, Nelson (1996) took $\delta \equiv 2$ and $\gamma \equiv 0$ in (14), and pointed out that the resulting scheme is the model of Brenner et al. (1996) altered slightly to admit a diffusion limit. While the empirical results of this paper suggest a simplification of (2) in which δ is one and ρ is nil, we provide here more general results that can be useful when applied to different data sets and/or related problems. As originally remarked by Nelson (1996), the kind of results that we are going to provide can be useful especially when a researcher is interested in the filtering performances of model (13) when ρ is not nil in (2). Also, we slightly complicate the analysis and allow standardized residuals to be general error distributed, but such a possibility is not considered in the empirical section of the paper:[11] in addition to the standard motivation that the normal distribution is not flexible enough empirically,[12] a second motivation for such a complication here came from some findings of Engle and Lee (1996) (see their tables 2 and 4), who obtained indirect estimates that seemed to be dependent on the distributional assumption made for the auxiliary model.

To save space, we shall be avoiding as much as possible unnecessary technical discussion referring to the construction of the measure space in (14): technical details can be found in Nelson (1990), and are those exploited in Fornari and Mele (1997a, b and 2000a) (see also Duan, 1997, for related work). We only introduce notation for the filtration generated by $\{{}_h r_{h(i-1)}, {}_h \sigma^\delta_{hi}\}^k_{i=1}$, which is $\mathcal{F}_{hk}$, and which will be used in appendix A. Let the symbol $\Rightarrow$ denote weak convergence. Recall that if a random variable x is general error distributed then its density is written as $\frac{\upsilon \exp(-\frac{1}{2}\nabla_\upsilon^{-\upsilon}|x|^\upsilon)}{2^{1+\upsilon^{-1}}\nabla_\upsilon \Gamma(\upsilon^{-1})}$, where $\nabla^2_\upsilon \equiv \frac{\Gamma(\upsilon^{-1})}{2^{2/\upsilon}\Gamma(3\upsilon^{-1})}, \upsilon > 0$ and $\Gamma(.)$ is the

---

[11] The reason why we did not implement the g.e.d.-based likelihood function in the empirical section is that doing so implies non-stationarity of the resulting model. However, by taking a normal-based likelihood function, we can always interpret the resulting estimates as qml estimates.

[12] As argued in Fornari and Mele (1997a), the *combination* of δ and υ should increase the flexibility of both the conditional and unconditional distributions of the error terms. In fact, while the conditionally normal GARCH gives rise to an invariant distribution of residuals that is a Student t, which is shaped by a single parameter, model (14) augmented with a conditional g.e.d. gives rise to an invariant distribution that is a *generalized* Student t when $\delta = \upsilon$, and a fairly general distribution when $\delta \neq \upsilon$, thus providing a potential better fit for the distribution of the residual.

Gamma function. The following convergence result is an extension of theorem 2.3 p. 211 in Fornari and Mele (1997a) that allows for the presence of the instantaneous correlation between $\{{}_h r_{hk}\}_{k=0,1,\cdots}$ and $\{{}_h\sigma^{\delta}_{hk}\}_{k=0,1,\cdots}$ as h shrinks to nil:

THEOREM 4.1: *Let* $m_{\delta,\upsilon} = \frac{2^{\frac{2\delta}{\upsilon}-1}\nabla_{\upsilon}^{2\delta}\Gamma(\frac{2\delta+1}{\upsilon})}{\Gamma(\upsilon^{-1})}$, $n_{\delta,\upsilon} = \frac{2^{\frac{\delta}{\upsilon}-1}\nabla_{\upsilon}^{\delta}\Gamma(\frac{\delta+1}{\upsilon})}{\Gamma(\upsilon^{-1})}$, *and let* $\frac{{}_h u_{hk}}{\sqrt{h}}$ *be general error distributed. Let*:

$$\varphi_h \equiv 1 - n_{\delta,\upsilon}((1-\gamma)^{\delta} + (1+\gamma)^{\delta})\alpha_h - \beta_h, \tag{15}$$

$$\psi_h \equiv \sqrt{(m_{\delta,\upsilon} - n^2_{\delta,\upsilon})((1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}) - 2n^2_{\delta,\upsilon}(1-\gamma)^{\delta}(1+\gamma)^{\delta}} \cdot \alpha_h, \tag{16}$$

$$\rho \equiv \frac{2^{\frac{\delta-\upsilon+1}{\upsilon}}\nabla_{\upsilon}^{\delta+1}\Gamma(\frac{\delta+2}{\upsilon}) \cdot ((1-\gamma)^{\delta} - (1+\gamma)^{\delta})}{\Gamma(\upsilon^{-1})\sqrt{(m_{\delta,\upsilon} - n^2_{\delta,\upsilon})((1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}) - 2n^2_{\delta,\upsilon}(1-\gamma)^{\delta}(1+\gamma)^{\delta}}} \tag{17}$$

and suppose that $\lim_{h\downarrow 0} h^{-1}\iota_h = \iota, \lim_{h\downarrow 0} h^{-1}\theta_h = \theta$ *and*:

$$\begin{array}{lcl} \lim_{h\downarrow 0} h^{-1} w_h & = & \omega \in (0,\infty), \\ \lim_{h\downarrow 0} h^{-1}\varphi_h & = & \varphi < \infty, \\ \lim_{h\downarrow 0} h^{-1/2}\psi_h & = & \psi < \infty. \end{array} \tag{18}$$

Then, $\{{}_h r_{h(k-1)}, {}_h\sigma^{\delta}_{hk}\}_{k=0,1,\cdots} \Rightarrow \{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$ *as* $h \downarrow 0$, *where* $\{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$ *are solutions of* (2) *when* $\eta \equiv 1$.

Let

$${}_h\xi_{hk} \equiv \frac{\left|\frac{{}_h u_{hk}}{\sqrt{h}}\right|^{\delta}(1-\gamma s_k)^{\delta} - E(\left|\frac{{}_h u_{hk}}{\sqrt{h}}\right|^{\delta}(1-\gamma s_k)^{\delta})}{\sqrt{(m_{\delta,\upsilon} - n^2_{\delta,\upsilon})((1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}) - 2n^2_{\delta,\upsilon}(1-\gamma)^{\delta}(1+\gamma)^{\delta}}}.$$

The preceding approximation result then says that when h shrinks to zero and the moment conditions in (18) are fulfilled, the distribution of the sample paths generated by the following model,

$$\left\{ \begin{array}{lcl} {}_h r_{h(k+1)} - {}_h r_{hk} & = & (\iota_h - \theta_h \cdot {}_h r_{hk}) + {}_h\sigma_{h(k+1)}\sqrt{{}_h r_{hk}} \cdot {}_h u_{h(k+1)} \\ {}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} & = & (w_h - \varphi_h \cdot {}_h\sigma^{\delta}_{hk}) + \psi_h \cdot {}_h\sigma^{\delta}_{hk} \cdot {}_h\xi_{hk} \end{array} \right. \tag{19}$$

gets 'closer and closer' to the distribution generated by the sample paths generated by (2), with ρ given by (17). Comparing (13) with (19) then suggests an estimator based on moment conditions; specifically, the *quasi-approximated* ML (q-aml) estimators of $\overline{\omega}, \varphi, \psi$ we propose are

$$\begin{array}{lcl} \omega_{\text{q-aml}} & \equiv & \Delta^{-3/2}\widehat{w}_{\Delta}, \\ \varphi_{\text{q-aml}} & \equiv & \Delta^{-1}\widehat{\varphi}_{\Delta}, \\ \psi_{\text{q-aml}} & \equiv & \Delta^{-1/2}\widehat{\psi}_{\Delta}, \end{array} \tag{20}$$

where $\widehat{\varphi}_{\Delta}, \widehat{\psi}_{\Delta}$ are obtained by means of (15)-(16) computed in correspondence of the qml estimator of model (13), $\widehat{w}_{\Delta}$ is the qml estimator of w_{Δ} of model (13). The q-aml estimator of δ is the qml estimator of δ in model (13), and the q-aml estimators of ι and θ are as those of ω and φ above. Finally, the q-aml estimator of ρ is obtained by plugging the qml estimators of $(\delta, \upsilon, \gamma)$ in formula (17).

The estimators in (20) are based on the moment conditions (18) and as we noted before, they may be affected by a disaggregation bias; furthermore, Corradi (2000) recently questioned the realism of the moment conditions that Nelson (1990) originally imposed to show the weak convergence of the GARCH(1,1) towards a continuous time stochastic volatility model. Her reasoning can be generalized here as follows. In the third equation of (14), the term generating the diffusion terms of volatility is proportional to $(h^{-\frac{\delta}{2}}\alpha_h) \cdot |_h u_{hk}|^{\delta}$, which is of course $O_p(\sqrt{h})$ under the third moment condition in (18). In other terms, a condition for a diffusion to be obtained is to scale the variance of $|_h u_{hk}|^{\delta}$ with a diverging sequence. In general, one would generate diffusion terms with $\alpha_h \cdot |_h u_{hk}|^{\delta}$, where $\alpha_h \approx O(h^q)$, $q \in \mathbb{R}$. This leaves three alternatives:

- a) $q = \frac{1-\delta}{2}$.
  b) $q < \frac{1-\delta}{2}$.
  c) $q > \frac{1-\delta}{2}$.

The first condition is another way to express the condition under which (14) has a well-defined diffusion limit; the second condition implies that (14) does not converge to any diffusion limit; and the third condition implies a 'degenerate' diffusion limit, i.e. with identically zero diffusion terms.

While recognizing that weak convergence results such as those contained in theorem 4.1 are obviously related to parametrization issues, in the empirical section we find that not only the parametrization in (14) provides a reasonably good picture of the volatility dynamics, consistently with the theoretical results of Nelson and Foster (1994), but it even passes the global consistency test that checks a posteriori the accuracy of the approximation in (15)-(16) and that we present below.

Finally, the previous conclusions remain perfectly the same when we generalize theorem 4.1 by freeing up η. As remarked in the previous section, there are no available ARCH models in the literature that can be used to obtain the convergence towards an unrestricted CEV volatility process. Consider however generalizing both (8) and (9) by means of the following model:

(21)
$$\begin{cases} {}_\Delta r_{n+1} & = & {}_\Delta r_n + \iota_\Delta - \theta_\Delta \cdot \ {}_\Delta r_n + \ {}_\Delta \sigma_{n+1} \sqrt{{}_\Delta r_n} \cdot \ {}_\Delta u_{n+1} \\ {}_\Delta \epsilon_n & = & {}_\Delta u_n \cdot {}_\Delta \sigma_n, \quad \frac{{}_\Delta u_n}{\sqrt{\Delta}} \sim ged_\upsilon \\ {}_\Delta \sigma^\delta_{n+1} & = & w_\Delta + \alpha_\Delta (|{}_\Delta \epsilon_n| - \gamma \cdot \ {}_\Delta \epsilon_n)^{\delta\eta} + \beta_\Delta \cdot \ {}_\Delta \sigma^\delta_n \\ & & + \alpha_\Delta \cdot E\left\{(|{}_\Delta u_n| - \gamma \cdot \ {}_\Delta u_n)^{\delta\eta}\right\} \cdot \left\{{}_\Delta \sigma^\delta_n - \ {}_\Delta \sigma^{\delta\eta}_n\right\}. \end{cases}$$

Chopping time in (21) as in (13), and rearranging, yields:

(22)
$$\begin{cases} {}_h r_{h(k+1)} & = & {}_h r_{hk} + \iota_h - \theta_h \cdot \ {}_h r_{hk} + \ {}_h \sigma_{h(k+1)} \sqrt{{}_h r_{hk}} \cdot \ {}_h u_{h(k+1)} \\ {}_h \epsilon_{hk} & = & {}_h u_{hk} \cdot {}_h \sigma_{hk}, \quad \frac{{}_h u_{hk}}{\sqrt{h}} \sim ged_\upsilon \\ {}_h \sigma^\delta_{h(k+1)} - {}_h \sigma^\delta_{hk} & = & w_h - \left(1 - h^{-\frac{\delta\eta}{2}} E\left\{\left|{}_h u_{hk}\right|^{\delta\eta} (1-\gamma s_k)^{\delta\eta}\right\} \alpha_h - \beta_h\right) \ {}_h \sigma^\delta_{hk} \\ & & + \alpha_h \cdot \left(\left|{}_h u_{hk}\right|^{\delta\eta} (1-\gamma s_k)^{\delta\eta} - E\left\{\left|{}_h u_{hk}\right|^{\delta\eta} (1-\gamma s_k)^{\delta\eta}\right\}\right) h^{-\frac{\delta\eta}{2}} \sigma^{\delta\eta}_{hk}. \end{cases}$$

We have:

THEOREM 4.2: *Let*

(23) $\varphi_h \quad \equiv \quad 1 - n_{\delta\eta,\upsilon}((1-\gamma)^{\delta} + (1+\gamma)^{\delta})\alpha_h - \beta_h,$

(24) $\psi_h \quad \equiv \quad \sqrt{(m_{\delta\eta,\upsilon} - n^2_{\delta\eta,\upsilon})((1-\gamma)^{2\delta\eta} + (1+\gamma)^{2\delta\eta}) - 2n^2_{\delta\eta,\upsilon}(1-\gamma)^{\delta\eta}(1+\gamma)^{\delta\eta}} \cdot \alpha_h,$

and

$$(25) \qquad \rho \equiv \frac{2^{\frac{\delta\eta-\upsilon+1}{\upsilon}} \nabla_{\upsilon}^{\delta\eta+1} \Gamma(\frac{\delta\eta+2}{\upsilon}) \cdot ((1-\gamma)^{\delta\eta} - (1+\gamma)^{\delta\eta})}{\Gamma(\upsilon^{-1})\sqrt{(m_{\delta\eta,\upsilon} - n^2_{\delta\eta,\upsilon})((1-\gamma)^{2\delta\eta} + (1+\gamma)^{2\delta\eta}) - 2n^2_{\delta\eta,\upsilon}(1-\gamma)^{\delta\eta}(1+\gamma)^{\delta\eta}}}.$$

Suppose that $\lim_{h\downarrow 0} h^{-1}\iota_h = \iota$, $\lim_{h\downarrow 0} h^{-1}\theta_h = \theta$ *and*:

$$(26) \qquad \begin{array}{lcl} \lim_{h\downarrow 0} h^{-1} w_h & = & \omega \in (0, \infty), \\ \lim_{h\downarrow 0} h^{-1}\varphi_h & = & \varphi < \infty, \\ \lim_{h\downarrow 0} h^{-1/2}\psi_h & = & \psi < \infty. \end{array}$$

Then, $\{_h r_{h(k-1),h}\, \sigma^{\delta}_{hk}\}_{k=0,1,\cdots} \Rightarrow \{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$ *as* $h \downarrow 0$, *where* $\{r(\tau), \sigma(\tau)^{\delta}\}_{\tau\geq 0}$ *are solutions of* (2) *and* $\{_h r_{h(k-1),h}\, \sigma^{\delta}_{hk}\}_{k=0,1,\cdots}$ *are solution of* (22).

Finally, notice that one can make a creative use of alternative asymmetric ARCH models to obtain convergence to models with correlated Brownian motions. An example of such a searching strategy is provided in appendix A.

Quasi indirect inference

We test and correct the potential disaggregation bias of the q-aml estimator with the indirect inference principle. The procedure that we follow is a natural generalization of Broze et al. (1995) and allows the volatility of the short-term rate to evolve in a stochastic and *autonomous* manner. Formally, if we replace the normality assumption with the g.e.d. assumption, the q-aml estimator of $b = (\Delta^{-1}\iota_{\Delta}$, $\Delta^{-1}\theta_{\Delta}$, $\Delta^{-3/2}w_{\Delta}$, $\Delta^{-1}\varphi_{\Delta}$, $\Delta^{-1/2}\psi_{\Delta}$, γ, δ, η, $\upsilon)'$ in (21) is:

$$a_{\text{q-aml}} \equiv \widehat{b}_N = \arg\max_b \mathfrak{L}_N(_{\Delta}r; b),$$

where $\mathfrak{L}_N(_{\Delta}r; b)$ is the likelihood function implied by (21), N is the sample size, and $_{\Delta}r$ is the observations set, which is supposed to be a discretely sampled diffusion from (2) when the true parameter vector is a_0. Note that dim$(b)>$ dim(a). In the empirical implementation below, however, we shall consider the Gaussian case in which $\upsilon\equiv 2$, and motivated by the Monte Carlo findings reported in section 3, we shall impose the time-scale invariance of δ and η. We also assume the same for γ; ascertaining whether such a time-scale invariance of γ is a reasonable assumption in practice is an open question that we leave for future research. Accordingly, now we re-interpret b as a vector in an open subset of $\mathbb{R}^5$ (with coordinates $\Delta^{-1}\iota_\Delta$, $\Delta^{-1}\theta_\Delta$, $\Delta^{-3/2}w_\Delta$, $\Delta^{-1}\varphi_\Delta$, $\Delta^{-1/2}\psi_\Delta$), $\mathfrak{L}_N(.)$ as a normal likelihood function with δ,η and γ fixed at prespecified values (e.g. at the preliminary qml estimates obtained by fitting model (21), as we actually do in the empirical part of the paper), and a as a vector in an open subset of $\mathbb{R}^5$, with coordinates $\iota,\theta,\overline{\omega},\varphi,\psi$.

It is well known that under standard regularity conditions (appendix B), one has asymptotic normality of the pseudo-ML estimator,

$$\sqrt{N}\left(\widehat{b}_N-b_0(a_0)\right)\xrightarrow{d}\mathrm{N}\left(0,\ddot{\mathfrak{L}}_\infty^{-1}(a_0;b_0(a_0))\cdot J(a_0)\cdot\ddot{\mathfrak{L}}_\infty^{-1}(a_0;b_0(a_0))\right),$$

where $\ddot{\mathfrak{L}}_\infty(.)$ and $J(.)$ are defined in appendix B, and $b_0(.)$ is the so-called *binding function*:

$$b_0(a_0)=\arg\max_b\mathfrak{L}_\infty(a_0;b),\text{ the limit problem.}$$

However, the true law of ${}_\Delta r$, as implied by the data generating mechanism, say $\ell_0({}_\Delta r)$, is such that

$$\ell_0({}_\Delta r)\notin\{\mathfrak{L}_N({}_\Delta r;b),\ b\text{ varying}\},$$

and the discrete time model is expected to behave in a way that allows for a discretization bias:

$$b(a_0)\neq a_0.$$

The reason why we may also refer to the preceding inequality as a 'discretization bias' is that when we chop time in (21) by creating sequences of the form $\{\iota_h,\theta_h,w_h,\alpha_h,\beta_h\}$, and substitute the moment conditions (23)-(26) of theorem 4.2 in (22), thereby creating a stochastic process $\{{}_hr_{hk},{}_h\,\sigma^{\delta}_{hk}\}_{k=0,1,\cdots}$ solution of:

$$
(27)\qquad \begin{cases} {}_hr_{h(k+1)} - {}_hr_{hk} &= (\iota - \theta\cdot {}_hr_{hk})h + {}_h\sigma_{h(k+1)}\sqrt{{}_hr_{hk}}\cdot {}_hu_{h(k+1)} \\ {}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} &= (\omega - \varphi\cdot {}_h\sigma^{\delta}_{hk})h + \psi\cdot {}_h\sigma^{\delta\eta}_{hk}\sqrt{h}\,{}_h\xi_{hk} \end{cases}
$$

then (21) is embedded in $\{{}_hr_{hk},{}_h\,\sigma^{\delta}_{hk}\}_{k=0,1,\cdots}$ (namely for $h \equiv \Delta$), and yet $\{{}_hr_{hk},{}_h\,\sigma^{\delta}_{hk}\}_{k=0,1,\cdots}$ converges weakly to the solution of (2) under the conditions given in theorem 4.2.

Indirect inference methods correct the preceding bias in the following manner. Consider simulating (27) for small h. This is accomplished by setting γ,δ to their ML estimates $\widehat{\gamma},\widehat{\delta}$, assigning values to $a = (\iota,\theta,\omega,\varphi,\psi)$, and drawing $\frac{{}_hu_{hk}}{\sqrt{h}}$ from the normal distribution; one obtains ${}_{h,h}\widetilde{r}^{(s)}(a) = \{{}_h\widetilde{r}^{(s)}_{hk}(a)\}_{k=0}^{N/h}$, $s = 1,\cdots,S$, where S is the number of simulations. For each simulation, just retain the (N) numbers ${}_h\widetilde{r}^{(s)}_{hk}(a)$ that correspond to integer indexes of time, and estimate the auxiliary model on each series of simulated data:

$$
\widehat{b}^{(h)}_{N,s}(a) = \arg\max_b \mathfrak{L}_N({}_{\Delta,h}\widetilde{r}^{(s)}(a);b), \qquad s = 1,\cdots,S,
$$

where ${}_{\Delta,h}\widetilde{r}^{(s)}(.)$ denotes the set of the simulated short-term rate with integer indexes of time at simulation s and interval h. In our specific just-identified problem ($\dim(a) = \dim(b)$), the indirect estimator of a is then the solution (provided it exists) of the following five-dimensional system:

$$
0 = \widehat{b}_N - \frac{1}{S}\sum_{s=1}^{S}\widehat{b}^{(h)}_{N,s}(a).
$$

Call ${}_h\widehat{a}_N(a_0)$ the solution of the preceding system. Heuristically, its asymptotic distribution can be obtained as follows. Expand the preceding system of equalities around a_0:

$$
\widehat{b}_N - \frac{1}{S}\sum_{s=1}^{S}\widehat{b}^{(h)}_{N,s}(a_0) = \left(\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widehat{b}^{(h)}_{N,s}}{\partial a}(a_0)\right)\left({}_h\widehat{a}_N(a_0) - a_0\right).
$$

For large N, the preceding is in fact an equality in distribution, and the covariance matrix of $(\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widehat{b}^{(h)}_{N,s}}{\partial a}(a_0))({}_h\widehat{a}_N(a_0) - a_0)$ is the covariance matrix of $\widehat{b}_N - \frac{1}{S}\sum_{s=1}^{S}\widehat{b}^{(h)}_{N,s}(a_0)$, i.e.

$(1+\frac{1}{S})\text{cov}(\widehat{b}_{N,s}^{(h)}(a_0))$, and one has:

(28) $$\sqrt{N}\left({}_h\widehat{a}_N(a_0)-a_0\right) \; N\uparrow\infty, \; h\downarrow 0 \stackrel{d}{\to} \mathrm{N}\left(0, \left(1+\frac{1}{S}\right) V_0^{-1}\Gamma_0 V_0^{\prime -1}\right),$$

where Γ_0 is the covariance matrix of the simulated estimator and $V_0 \equiv \frac{\partial b}{\partial a}(a_0)$, i.e. the Jacobian of the binding function evaluated at a_0. Broze et al. (1998) proved the preceding result in great generality — i.e. in the case of a general diffusion in $\mathbb{R}^l$ — and to avoid bias due to the discretization step used during the simulations (hence the label 'quasi'-indirect inference), the authors also suggested to take $h = N^{-d}$ with $d > \frac{1}{2}$. In appendix B, we check the conditions of Broze et al. (1998) that ensure that (28) holds for the scheme proposed here.

Notice also that (27) do *not* represent the Euler approximation of (2), but this is not a disturbing feature since it is known since Broze et al. (1998) that implementing the indirect inference estimator just requires the weak convergence of the high frequency simulator towards the solution of (2); see also appendix B. For reasons of comparisons, however, the empirical section also considers the case in which the high frequency simulator is the Euler-Maruyama approximation of (2) (i.e.,(3)).

Finally, it is easy now to implement a global specification testing procedure that controls the adequacy of the approximating model (21). It is sufficient to use the consistency test appearing in Gouriéroux et al. (1993, section 4.2 and appendix 3). Such a test is designed to verify the existence of a fixed point of the binding function:

$$\mathrm{H}_0 : a_0 = b(a_0).$$

Let I denote the identity matrix in $\mathbb{R}^{5\times 5}$. Under H_0, one has that:

$$\sqrt{N}\left(\widehat{b}_N - \frac{1}{S}\sum_{s=1}^{S}\widehat{b}_{N,s}^{(h)}(\widehat{b}_N)\right)$$

(29) $$\stackrel{d}{\to} \mathrm{N}\left(0, \left(I-\frac{\partial b}{\partial a}(a_0)\right) \ddot{\mathfrak{L}}_\infty^{-1} \; J \; \ddot{\mathfrak{L}}_\infty^{-1} \left(I-\frac{\partial b}{\partial a}'(a_0)\right) + \frac{1}{S}\,\ddot{\mathfrak{L}}_\infty^{-1} \; J \; \ddot{\mathfrak{L}}_\infty^{-1}\right).$$

5. Empirical analysis

The data

We use weekly data referring to 3-month Treasury Bill rates to approximate the short-term rate.[13] This is the same data set used by Andersen and Lund (1997a,b), but here we restrict attention to the sample spanning the period from May 30, 1973 to February 22, 1995, which has 1135 observations. The motivation for using weekly data lies in an attempt of avoiding problems raised by market microstructure effects. The motivation for restricting attention to such a particular sample lies in the possibility of estimating risk premia coefficients in a term-structure extension of the model (see paragraph 6), by fitting our resulting theoretical model to a target term structure that is closely related to the target term structure constructed by Aït-Sahalia (1996a) in correspondence of the same sample period. For reasons developed below, however, we did not use the short-term rate data set constructed by Aït-Sahalia (1996a, b).

Raw data are converted into instantaneous figures, hereafter referred to as r, and table 2 contains some preliminary statistics. Table 3 contains the estimated autocorrelation function, which shows a high amount of persistence in the data. Nonstationarity is formally tested by performing augmented Dickey-Fuller tests that indicate that data are borderline stationary. As an example, the statistic takes a value of -2.435 at lag 5, which is roughly the threshold value for rejecting nonstationarity with 90 percent probability; more generally, one rejects nonstationarity at the 85-90 percent to the extent of lag 15, but because the test has low power, even such a slight rejection can be symptomatic of stationarity in the data. It is worth noticing that the same kind of results holds for the full sample originally exploited by Andersen and Lund.

The auxiliary discrete time model

We start with estimating model (21). Consistently with previous results of Andersen and Lund (1997a), we do not find evidence that positive shocks introduce more volatility than negative shocks of the same size, i.e., the inverse of the leverage effect. At best, there is evidence that the opposite takes place, although the estimate of γ is not statistically significant. When we try to fit the same kind of models to weekly samples of the data used by Aït-Sahalia

---

[13] See Chapman et al. (1999) for an analysis concerning the validity of such an approximation.

(1996a, b), however, we find strong evidence of asymmetry having the 'right' direction, but also find that volatility dynamics is almost entirely driven by past errors, thus exhibiting a rather chaotic behavior. We thus pursue the analysis with the Andersen-Lund data, and estimate again model (21) dropping the asymmetry parameter. The parameters estimates are very close to the ones that we finally use as calibrating devices during the indirect inference procedure, and imply that the model gives rise to stable dynamics for the volatility process. As regards the estimates of δ and η, we find that they are 1.0326 and 1.0014, respectively, and that they are statistically not distinguishable from 1. This suggests the possibility of further simplifying the representation in (2), by fixing $\delta = \eta = 1$. Such a restriction, along with the restriction $\gamma = 0$, will propagate into an important simplification of the indirect inference phase. In the model that we select as an auxiliary device, we thus restrict $(\delta, \eta, \gamma) \equiv (1, 1, 0)$. Furthermore, please notice that due to numerical stability issues, model (21) was estimated without explicitly disentangling the sample frequency. When we estimated (21) with the restrictions $(\delta, \eta, \gamma) \equiv (1, 1, 0)$, for instance, we casted the model in the following format,

$$
(30) \qquad \left\{ \begin{array}{rcl} r_n & = & c_0 + c_1 r_{n-1} + r_{n-1}^{1/2} \cdot \epsilon_n, \quad \epsilon_n \equiv (u \cdot \sigma)_n, \quad u \sim NID(0,1) \\ \sigma_n & = & w + \alpha \left| \epsilon_{n-1} \right| + \beta \sigma_{n-1}, \quad n = 2, \cdots, N, \end{array} \right.
$$

where $\{r_n\}_{n=1}^N$ denotes the observed (weekly) series, and $(c_0, c_1, w, \alpha, \beta)$ are real parameters. The correspondence between the estimators of the parameters in (21) and (30) is not hard to write down:

$$
\widehat{b}_N \equiv a_{\text{q-aml}} = \Delta_0 + \Delta_1 \widehat{m}_N,
$$

where $\widehat{m}_N$ denotes the vector of the ML estimators of the parameters in (30), $\Delta_0 = \begin{pmatrix} 0 & \Delta^{-1} & 0 & \Delta^{-1} & 0 \end{pmatrix}'$, and

$$
\Delta_1 = \begin{pmatrix} \Delta^{-1} & 0 & 0 & 0 & 0 \\ 0 & -\Delta^{-1} & 0 & 0 & 0 \\ 0 & 0 & \Delta^{-3/2} & 0 & 0 \\ 0 & 0 & 0 & -0.798 \cdot \Delta^{-1} & -\Delta^{-1} \\ 0 & 0 & 0 & 0.603 \cdot \Delta^{-1/2} & 0 \end{pmatrix},
$$

with $\Delta = \frac{1}{52}$. Similarly, the Jacobian of the binding function that has been used to report the t-statistics and the consistency tests in table 6 is based on the set of parameters of the auxiliary model (30): to such a set of parameters is associated a binding function of the form $m = m(a)$, and the relationship between the Jacobians of b and m is

$$\frac{\partial b}{\partial a}(.) = \Delta_1 \frac{\partial m}{\partial a}(.). \tag{31}$$

Model (30) is the absolute-value model of Taylor (1986) and Schwert (1989) with normal errors, studied by Nelson and Foster (1994) and Fornari and Mele (1997a). Its main advantage over the more usual variance specifications is that it delivers estimates of volatility that are relatively more robust to the presence of possible outliers in the data. In this case, we also know that the invariant distribution of residuals is approximately a generalized Student's t when $\delta = \upsilon$ (theorem 3.3 p. 218 in Fornari and Mele (1997a)), which reduces to the celebrated Student's t result of Nelson (1990) when $\delta = \upsilon = 2$.

As mentioned in paragraph 4, we consider normally distributed errors only, since expanding into non-normality makes the resulting model non-stationary.[14] Hence, we are left with a specification in which $(\delta, \eta, \upsilon) = (1, 1, 2)$, and it is possible to show that in this case the invariant distribution of ϵ is more leptokurtic than the Student's t that obtains when $(\delta, \eta, \upsilon) = (2, 1, 2)$. Specifically, by applying theorem 3.5 p. 218 in Fornari and Mele (1997a), we obtain that the invariant distribution of residuals in model (30) is given by

$$P(\bar{\epsilon}) = \frac{\left(\frac{2\omega}{\psi^2}\right)^{\frac{2\varphi+\psi^2}{\psi^2}}}{\sqrt{2\pi}\cdot\Gamma\left(\frac{2\varphi+\psi^2}{\psi^2}\right)} \int_0^\infty x^{-\frac{2\varphi+3\psi^2}{\psi^2}} \exp\left(-\frac{1}{2}\bar{\epsilon}^2 x^{-2} - \frac{2\omega}{\psi^2}x^{-1}\right) dx,\ \bar{\epsilon} \equiv \frac{\epsilon}{\sqrt{h}}, \tag{32}$$

as $h \downarrow 0$. Figure 4 compares the density in (32) with a normal density with variance equal to $(\omega/\varphi)^2$ where ω, φ and ψ have been fixed at the values of the second column of table 6. The density in (32) should capture the usual stylized facts of the unpredictable parts of general financial time series and, following Gallant and Tauchen (1996), one might conjecture that the II estimator described in section 4 would be as efficient as the (intractable) ML estimator if the density in (32) were to form a smooth embedding of the invariant distribution associated with the discretely sampled diffusion (2).

---

[14] Such a phenomenon is also noted by Andersen and Lund (1997a), who show that a specification based on EGARCH-type models is more stable when the errors of the model are nonnormal. Motivated by further empirical findings of Andersen and Lund (1997a), we also tried to include further lags in the volatility equation, but we did not observe any significant improvements.

Unfortunately, such a conjecture seems to be hard to verify here, since the conditional distribution of the residuals that eventually generated (32) is just a normal distribution; nevertheless, the distribution (32) should play an excellent role in mimicking the long-run properties of the unpredictable part of the series that we study.

Table 4 reports the qml estimates of model (30). Notice that the condition for covariance-stationarity of this model is not violated. Theorem 4.1 suggests that the covariance-stationarity condition does not impose that $\alpha + \beta < 1$; rather, $2 \cdot n_{1,2} \alpha + \beta = 0.798 \cdot \alpha + \beta < 1$ has to hold here, which is effectively the case of the qml estimates reported in table 3. This implies a persistence of nearly 0.993 in the volatility propagating process.

Table 5 presents summary statistics of the volatility filtered by the model (not yet rescaled for diffusions), and figure 5 depicts its behavior in the sample. For reasons of comparisons, we also depict the first differences of r. The model appears to successfully capture some stylized features of the data, including the high volatility induced by the 'Monetary Experiment' of the early 80's. It is also worth noticing that perhaps due to such an isolated and yet relatively persistent episode, the long run volatility as implied by the parameter estimates attains the value of $\frac{w}{1-0.798 \cdot \alpha - \beta} = 15.458 \cdot 10^{-3}$, which is more than twice the average value of the filtered volatility for the whole sample. Because the estimated volatility wanders in a range of variation of about 0.026, however, such a difference is negligible: when we compute the ratio of the difference between the long run and average volatility to the range of variation, we find that it equals 0.321.

Correction of the discretization bias, consistency tests, and filtering

Following the program stated in paragraph 4, we begin with computing the q-aml estimates (see (18) and (20)). The second column in table 6 reports such figures. Then we proceed with correcting their potential disaggregation bias by means of indirect inference. To implement the indirect inference estimator, system (2) is simulated by means of the Euler-Maruyama approximation[15] (3) with $h^{-1} \equiv 1300$, which corresponds to generating 25 sub-intervals within a week. With an observations set of $N = 1135$, this implies that $h = N^{-l}$, with $l \simeq 1.0193 > \frac{1}{2}$: hence, we are fulfilling the conditions developed in Broze et al. (1998)

[15] Using (27) as simulation device does not alter our estimation results.

to avoid simulation biases. We use $S = 50$ simulations. The estimation results are in the third column of table 6. The correction made by indirect inference does not appear to be important. First, we find that none of the q-aml estimates are out of the usual 95 percent probability bands around the corresponding indirect inference estimates. Second, and more importantly, when we formally checked the adequacy of the auxiliary model through the consistency test described in paragraph 4, we found that the adjustment speed of the short-term rate is the only parameter that does not pass the test at the standard 95 percent level.

Such findings are of special interest here: as recalled in paragraph 4, 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 by introducing the so-called ARCH diffusion which is, heuristically, the continuous time stochastic volatility process whose implied discrete differences form a weak-ARCH process. A 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 corresponding (discrete time) weak-ARCH models. Naturally, these are issues that deserve a deep theoretical investigation that we leave for future research.

To check that the previous estimation results do not depend on the dimension of the simulation experiment ($S = 50$), we implement a sort of reverse exercise that consists at looking for the ARCH model that one can expect to estimate if the true data generating mechanism happens to be (2). Specifically, we simulate (2) with parameters fixed at the indirect inference estimates of table 6, sample the short-term rate at weekly frequency, and estimate model (30) with such sampled data. We repeat the experiment 5000 times, and remove the simulations for which there was not stationarity for the short-term rate and volatility (i.e., for persistence greater than one). Notice that as a by-product of such an experiment, we will also get an assessment of the filtering performance of model (30).

Table 7 provides some basic statistics of the estimates, and figure 6 displays their relative frequencies. The distributions of the estimates are concentrated around the values of the estimates reported in table 4: specifically, the standard 95 percent confidence bands of the

Monte Carlo estimates are sufficiently tight to ensure statistical significance; yet they contain the figures corresponding to the true estimates reported in table 4.

The filtering performance of the model is gauged in the following manner. Let $\sigma_{i,n}$ denote the volatility simulated at the ith simulation and sampled at n, and $\widehat{\sigma}_{i,n}$ is the corresponding (rescaled) ARCH estimate. We are interested in evaluating the average filtering error in all the simulations: $\{\mathcal{E}_i\}_{i=1}^{5000}$, where $\mathcal{E}_i \equiv \frac{1}{1135}\sum_{n=1}^{1135}(\sigma_{i,n} - \widehat{\sigma}_{i,n})$. Figure 7 displays the Monte Carlo distribution of the average filtering error. It has an average value of $9.610 \cdot 10^{-5}$ and a standard deviation of $3.275 \cdot 10^{-3}$ The RMSE, defined as $\sqrt{\frac{1}{5000}\sum_{i=1}^{5000}(\frac{1}{1135}\sum_{n=1}^{1135}(\sigma_{i,n} - \widehat{\sigma}_{i,n})^2)}$, is equal to 0.0209.

6. Volatility and the term-structure

What does theory say ?

An empirical issue that has received relatively little attention in the literature is the relationship between the short-term rate volatility and the whole term-structure of interest rates. In a recent paper, Mele (2000) provides a theoretical analysis of this problem. One of his main results is that when the risk-neutralized drift function of the short-term rate is increasing in volatility, the yield curve at *short* maturity dates increases with volatility. This phenomenon takes place irrespective of whether one considers two-factor models (such as the model this paper analyses) or models with, say, three factors that incorporate a stochastic central tendency factor.

To get an intuition of such a result, consider the following model:

$$(33) \qquad \begin{cases} dr(\tau) = b(r(\tau), z(\tau), l(\tau))d\tau + \sqrt{2\sigma(z(\tau)) \cdot a(r(\tau))} \cdot d\widetilde{W}(\tau), \ \tau \in (0, T] \\ dz(\tau) = \varphi(r(\tau), z(\tau))d\tau + \sqrt{2\psi(r(\tau), z(\tau))} \cdot d\widetilde{W}^z(\tau), \ \tau \in (0, T] \\ dl(\tau) = \varepsilon(l(\tau))d\tau + \sqrt{2\pi(l(\tau))} \cdot d\widetilde{W}^l(\tau), \ \tau \in (0, T] \\ r(0) = x, z(0) = s, l(0) = c \end{cases}$$

where $\widetilde{W}$, $\widetilde{W}^z$ and $\widetilde{W}^l$ are Brownian motion defined under the risk-neutral measure, and the various drift and diffusion functions above satisfy conditions guaranteeing that the previous system has a strong solution.

In this model, l can be interpreted as a stochastic central tendency factor: it enters the drift function of the short-term rate but not the diffusion coefficient $\sqrt{2\sigma a}$ in order to be distinguished from the stochastic volatility factor z. Now suppose for simplicity that the various Brownian motions are independent. Let the bond rational price function at $\tau = 0$ be $v(x, s, c, 0, T)$. As shown by Mele (2000, section 5), $\partial v(r, z, l, \tau, T)/\partial z$ is then the solution of the following partial differential equation:

(34)
$$\begin{cases} \left(\frac{\partial}{\partial \tau} + L - k\right) \frac{\partial v(r,z,l,\tau,T)}{\partial z} = -\left(\frac{\partial b(r,z)}{\partial z} \cdot \frac{\partial v(r,z,l,\tau,T)}{\partial r} + \sigma'(r) a(r) \frac{\partial^2 v(r,z,l,\tau,T)}{\partial r^2}\right), \\ \qquad \forall (r, z, l, \tau) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+ \times [0, T) \\ \frac{\partial v(r,z,l,T,T)}{\partial z} = 0, \ \forall (r, z, l) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+ \end{cases}$$

where L is a partial differential operator and k is a killing rate (see Mele, 2000, for further details). Mele shows that at short maturity dates, $\partial v(r, z, l, \tau, T)/\partial r$ is always negative and of a higher order than $\partial^2 v(r, z, l, \tau, T)/\partial r^2$. By the maximum principle, $\partial v(r, z, l, \tau, T)/\partial z$ is then always negative at short maturity dates whenever $\partial b(r, z)/\partial z > 0$.

Models with three factors of this kind have been introduced because the yield curve seems to be driven by three factors (see, e.g., Litterman and Scheinkman, 1991; Dai and Singleton, 2000; earlier proponents of models such as (33) were Balduzzi et al., 1996, Chen, 1996, Andersen and Lund, 1997b). Despite the rich dynamics that (33) can generate, however, (34) reveals that the qualitative behavior of $\partial v(r, z, l, \tau, T)/\partial z$ is the same as the qualitative behavior of bond prices in an economy *without* a stochastic central tendency factor. It is then appropriate to get a picture of the relationship between volatility and the term-structure within a two-factor model that does not display any central tendency of the kind of (33).

Motivated by these results, we now address issues concerning the relations between volatility and the term-structure by making reference to our model (2), with parameters set to the estimates obtained in the previous paragraph. Our primary interest lies in understanding whether the positive relationship between volatility and the term-structure that is predicted by the theory at *short* maturities also holds at higher ones. Naturally, we are not claiming to be correctly modeling the whole term-structure with the help of just two factors. The objective here is only a proper understanding of the relationships between volatility and term-structure: as explained before, appending a central tendency factor to model (2) would not significantly change any of the primary conclusions we shall obtain, as (34) reveals.

To correctly address the issues we are exploring in this paragraph, we need specifying the risk-premia demanded by agents as compensation for the fluctuation of the stochastic factors. In appendix D, we provide conditions under which a supporting equilibrium exists for a rational bond price function satisfying the following partial differential equation:

$$
\left\{
\begin{array}{lcl}
Lv(r,z,\tau,T) & = & \Lambda(r,z,\tau,T) + rv(r,z,\tau,T),\ \tau \in [0,T) \\
v(r,z,T,T) & = & 1,\ \forall (r,z) \in \mathbb{R}^2_{++}
\end{array}
\right. \tag{35}
$$

where

$$
\begin{array}{lcl}
Lv(r,z,\tau,T) & \equiv & v_\tau + v_r \cdot (\iota - \theta r) + v_z \cdot (\omega - \varphi z) \\
 & & +\frac{1}{2}\left(v_{rr}\cdot rz^2 + 2v_{rz}\cdot \psi\rho\sqrt{r}z^2 + v_{zz}\cdot \psi^2 z^2\right), \\
\Lambda(r,z,\tau,T) & \equiv & v_r \cdot \lambda_1 rz + v_z \cdot \left(\lambda_1\rho + \lambda_2\sqrt{1-\rho^2}\right)\cdot \psi z\sqrt{r},
\end{array}
$$

and λ_1, λ_2 are constants. As shown in appendix C, our supporting equilibrium extends Cox et al. (1985) and implies that the short-term rate is the solution of (2) when $\delta = \eta = 1$ and $\rho = 0$, which are the estimates found in the previous paragraph. Generalizing such a supporting equilibrium to cases in which $\rho \neq 0$, and δ and η take arbitrary (and admissible) values is straightforward. In any case, the two unit risk premia demanded by agents in this economy will always be given by:

$$
\overline{\lambda}_i(\tau) \equiv \lambda_i\sqrt{r(\tau)},\ i = 1,2. \tag{36}
$$

Calibration, comparative statics and misspecification issues

Previous work on the term-structure with stochastic volatility used to produce simulation exercises based on arbitrary functional forms of the risk premia with parameters fixed at similarly arbitrary numerical values.[16] In addition to provide an equilibrium justification of the functional forms of the risk-premia (see (36) and appendix D), we also calibrate model (35) to a target term structure. Our target term-structure is very close to the one used by Aït-Sahalia (1996a) (see Fornari and Mele, 2000b, section IV.E, for details). We fix the initial state at $(r(0),\sigma(0)) \cong (7.1\cdot 10^{-2}, 0.05)$, which corresponds to the sample average level of r and

[16] A lone exception is the work of Longstaff and Schwartz (1992).

the average (rescaled) σ as filtered by model (30). Then we calibrate model (35) by choosing the couple (λ_1, λ_2) as the one which minimizes the squared differences between the target term structure and the one predicted by our model. In such a search procedure, the remaining parameters of the model were fixed at the indirect inference estimates of table 6. The partial differential equation (35) was solved numerically with the Crank-Nicholson method.[17] We imposed the following transversality conditions: $\lim_{r\to\infty} B(r, \sigma, t) = 0 \; \forall(\sigma, t) \in \mathbb{R}_+ \times [0, T]$, and $\lim_{\sigma\to\infty} B(r, \sigma, t) = 0 \; \forall(r, t) \in \mathbb{R}_+ \times [0, T]$. We found that $(\lambda_1, \lambda_2) \cong (-0.581, 0.677)$, with a fairly good fit (see Fornari and Mele, 2000b, for further details).

Figure 8 depicts our fitted term-structure, which is the second curve starting from the top. It is consistent with well-known stylized facts of the US term structure in the analyzed sample: it is increasing, very steep until 5 years and relatively flat for higher maturities. Figure 8 also shows that the yield curve increases with volatility. This is in accordance with the theoretical predictions mentioned previously: due to the fact that λ_1 is negative, the risk-neutral drift of the short-term rate is increasing in volatility, according to the Girsanov's theorem, and (34) then says that bond prices are decreasing in volatility, at least at short-maturity dates. The new aspect that is important here is that the yield curve appears to be always increasing in volatility, even at medium-long maturity dates. In order to ascertain whether such a result is due to the transversality condition involving volatility, we then repeated the calibration procedure described before without imposing such a transversality condition. Of course we obtained different values for (λ_1, λ_2), but the qualitative features of figure 8 remained the same, although the fit deteriorated.

Naturally, the purpose of the previous exercises was not to test the restrictions imposed by the theoretical two-factor model (35). This is a kind a testing procedure that goes well beyond the central objectives of the paper. Indeed, it is well known at least since Litterman and Scheinkman (1991) that actual yield curve movements are driven by three factors corresponding to changes in level, steepness and curvature of the term structure. A three-factor model such as (33) seems then to be more appropriate for the purpose of bond pricing. Nevertheless, our intent here was to understand the relationship between volatility and the term-structure with the help of model (2). As already argued, theory suggests that our

[17] See Fornari and Mele (2000a, chapter 5) for technical details concerning the implementation of this method in models of the short-term rate with stochastic volatility.

findings can be utilized to obtain a reliable qualitative picture concerning such a relationship even when the data generating process comprises three factors, as in (33).

The last objective of this paragraph now consists in showing that even in the presence of misspecification, the kind of models considered in this paper still remain a valid reference, at least insofar as one considers volatility filtering issues. Suppose, in other terms, that the data generating process (under the objective measure) is a three-factor model including the short-term rate, stochastic volatility, and a stochastic central tendency factor. The question we want to answer to is: are the filtering results of this paper still valid when we attempt at extracting the (unobserved) stochastic volatility of such a data generating process? In addition to its obvious practical content, such a problem is directly related to previous theoretical work by Nelson (1992) and Nelson and Foster (1994). As mentioned in the Introduction, these authors produced many theoretical results based on more or less restrictive assumptions. The message of such results is that even in the presence of serious misspecification, ARCH models still remain robust volatility filters. Now we wish to ascertain whether such results hold in an experiment in which ARCH models are used to reconstruct the volatility dynamics of a three-factor data generating process.

To this end, we implement a Monte Carlo experiment in which we fit model (30) to 1,000 simulated trajectories of a three-factor model that extends in a natural way model (1) (see the equations in table 8). Table 8 provides the results. Even though model (30) is neglecting a factor (namely, the stochastic central tendency factor), it exhibits volatility filtering properties of exceptional interest. The Monte Carlo properties of the volatility filtering error display the same order of magnitude as those found in section 5 and, again, the resulting dynamics of simulated vis-a-vis filtered volatility trajectories display the same patterns as in figure 1. Considered as a (stochastic) volatility filter, model (30) would be hardly rejected as a remarkably useful tool of analysis, even in the presence of neglected factors.

7. Conclusion

The intent of this paper was to explore to which extent ARCH models can be used in practice for the purpose of providing parameter estimates and volatility filtering of diffusions processes. Since the *standard* ARCH models that have traditionally been used in the empirical literature do not approximate all diffusion models, we considered a reasonably wide class of

ARCH models, which we named CEV-ARCH, that converges toward any unrestricted CEV diffusion process as the sample frequency becomes larger and larger. While the searching strategy followed in this paper to the aim of approximating diffusions by means of ARCH can be used to construct ARCH sequences converging to yet more general diffusion processes, our central focus was the special case of volatility following a CEV-diffusion with linear drift.

Despite the fact that the CEV coefficient of volatility was unrestricted in this paper, we provided empirical evidence supporting a model in which the (stochastic) volatility process of the short-term rate follows a diffusion process with *unit* elasticity of variance. In addition, we made use of simulation-based techniques to implement a global specification test for just-identified problems and provided evidence that (suitably rescaled) ARCH estimates of relevant parameters are statistically not distinguishable from estimates that one obtains with, say, indirect inference methods. Finally, the volatility filtering performances of the models are excellent. Even in the presence of important misspecification, i.e. by extracting volatility from a three-factor model by means of a two-factor model only, the volatility filtering errors have the same order of magnitude as in absence of misspecification. This finding suggests very simple and yet efficient tools to extract (unobserved) volatility of a diffusion.

Tables and figures

Table 1

Monte Carlo study [a]

parameter	true	average	median	std.dev.
δ	1	1.0725	1.0206	0.1273
η	1	1.0849	1.0834	0.0961
volatility filtering error	NA	-1.1163$\cdot 10^{-4}$ [b]	-2.2082$\cdot 10^{-4}$	4.5025$\cdot 10^{-3}$
δ	2	2.0047	1.9737	0.2474
η	$\frac{1}{2}$	0.6178	0.6132	0.1320
volatility filtering error	NA	-1.4995$\cdot 10^{-3}$ [c]	-2.3333$\cdot 10^{-4}$	5.6091$\cdot 10^{-2}$

[a] The third column reports the average ML estimates of δ and η in (2) obtained by fitting an AR(1) model with volatility equation given by eq. (9) to 1000 simulated weekly sampled trajectories from the stochastic differential equation system (2). In these simulations, $\iota = 8 \cdot 10^{-3}$, $\theta = 0.11$, $\varphi = 0.38$ and δ and η are fixed at the values of the second column, with A) $\omega = 0.03$, $\psi = 0.8$ when $\delta = \eta = 1$, and B) $\omega = 2.36 \cdot 10^{-3}$, $\psi = 0.06$ when $\delta = 2$ and $\eta = \frac{1}{2}$. The fourth and fifth columns report the Monte Carlo median and standard deviation of such estimates. The case $\delta = \eta = 1$ corresponds to the actual estimates obtained in paragraph 5. The Table also reports the Monte Carlo average (with the RMSE and the steady state expectation of σ in parentheses), median and standard deviation of the volatility filtering error.

[b] (RMSE: $1.8609 \cdot 10^{-2}$) ; $(\omega / \varphi = 7.895 \cdot 10^{-2})$;

[c] (RMSE: $1.6692 \cdot 10^{-2}$) ; $\left(\sqrt{\left(\omega - \psi^2/4\right)/\varphi} = 6.1985 \cdot 10^{-2}\right)$.

Table 2

Summary statistics of r

mean	median	maximum	minimum	std. dev.	skewness	kurtosis
0.070	0.068	0.155	0.026	0.026	0.828	3.681

Table 3

Autocorrelation function of r

lag	1	2	3	4	5	10	30	50
autocorrelation	0.995	0.998	0.979	0.971	0.961	0.914	0.789	0.696

Table 4

QML estimates of (30)[a]

parameter	estimate	t-stat [b]
c_0	$1.555{\cdot}10^{-4}$	2.09 (2.58)
c_1	0.9979	$6.79{\cdot}10^2$ $(9.86{\cdot}10^2)$
w	$1.110{\cdot}10^{-4}$	4.36 (3.41)
α	0.1504	13.05 (13.76)
β	0.8728	93.97 (139.18)

[a] QML is the quasi-maximum likelihood estimation of the short rate dynamics. [b] Bollerslev-Wooldridge (1992) robust t-statistics in parentheses.

Table 5

Summary statistics of the conditional volatility σ as filtered by eq. (30) [a]

mean	median	maximum	minimum	std. dev.	skewness	kurtosis
$7.102{\cdot}10^{-3}$	$5.483{\cdot}10^{-3}$	$2.805{\cdot}10^{-2}$	$2.042{\cdot}10^{-3}$	$4.306{\cdot}10^{-3}$	1.761	6.048

[a] Not rescaled for diffusion.

Table 6

Parameter estimates [a]

parameter	q-aml	II	II t-stat	consistency tests
ι	0.0081	0.0082	3.04	-0.6727
θ	0.1067	0.1108	2.92	-2.0855
ω	0.0418	0.0301	2.98	-1.2177
φ	0.3736	0.3806	3.01	-0.1275
ψ	0.6540	0.8092	3.23	0.1390

[a] The second column reports the estimates of the parameters in (2) obtained with the moment conditions (18) and (20). The second column reports estimates obtained via the indirect inference (II) strategy explained in paragraph 5, and the third column gives the corresponding t-statistics computed using the variance in (28) and (31) as the Jacobian of the binding function. The last column reports the ratio of each element of $\widehat{b}_N - \frac{1}{S}\sum_{s=1}^{S}\widehat{b}_{N,s}^{(h)}(\widehat{b}_N)$ to the corresponding standard error computed from the variance in (29) and using (31) as the Jacobian of the binding function.

Table 7

Monte Carlo study [a]

parameter	average	median	std. dev.
c_0	$1.640{\cdot}10^{-4}$	$1.610{\cdot}10^{-4}$	$3.340{\cdot}10^{-5}$
c_1	0.9974	0.9976	$1.764{\cdot}10^{-3}$
w	$1.210{\cdot}10^{-4}$	$1.130{\cdot}10^{-4}$	$4.420{\cdot}10^{-5}$
α	0.1548	0.1544	$2.405{\cdot}10^{-2}$
β	0.8665	0.8669	$2.056{\cdot}10^{-2}$

[a] The second column reports the average qml estimates of the parameters in model (30) obtained by fitting model (30) to 5000 simulated weekly sampled trajectories from the stochastic differential equation system (2). In these simulations, parameters are set to their II estimates reported in the third column of table 6. The third and fourth columns report the Monte Carlo median and standard deviation of the simulated qml estimates.

Table 8

Monte Carlo study [a]

	average	median	std. dev.
volatility filtering error	$-3.6815 \cdot 10^{-5}$ [b]	$-6.0461{\cdot}10^{-5}$	$3.9666{\cdot}10^{-3}$

[a] The second column reports the average volatility filtering error defined in paragraph 5 (with the RMSE and the steady state expectation of σ in parentheses) obtained by fitting model (30) to 1000 simulated weekly sampled trajectories of the following three-factor model:

$$\begin{cases} dr(\tau) &= \theta\left(l(\tau) - r(\tau)\right)d\tau + \sqrt{r(\tau)}\sigma(\tau)dW^{(1)}(\tau) \\ d\sigma(\tau) &= \left(\omega - \varphi\sigma(\tau)\right)d\tau + \psi\sigma(\tau)dW^{(2)}(\tau) \\ dl(\tau) &= \left(b_1 - b_2 l(\tau)\right)d\tau + b_3\sqrt{l(\tau)}dW^{(3)} \end{cases}$$

where $W^{(i)}, i = 1, 2, 3$, are standard Brownian motions, θ, ω, φ and ψ are fixed at the indirect inference estimates of Table 6, and $b_i, i = 1, 2, 3$, are fixed at the values suggested by Andersen and Lund (1997b), i.e. $b_1 = 0.0078$, $b_2 = 0.1257$ and $b_3 = 0.0493$. The third and fourth columns report the Monte Carlo median and standard deviation of the volatility filtering error.

[b] (RMSE: 0.0285) ; $(\omega\,/\varphi \;=\; 7.895 \cdot 10^{-2})$.

Figure 1

Filtered weekly volatility diffusion $\sigma(t)$ in eqs.(1) by means of an ARCH model

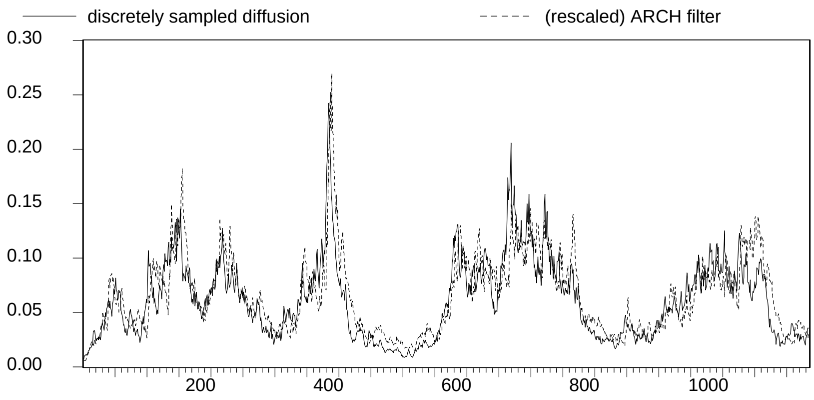
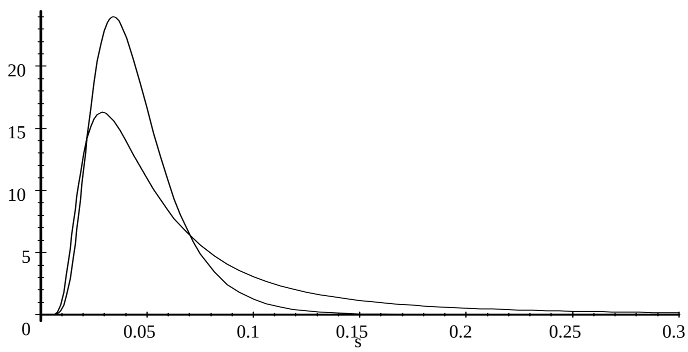


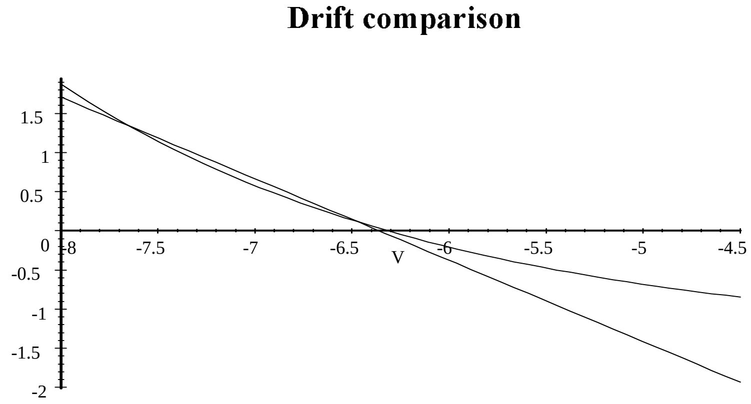
Figure 2

Density comparison



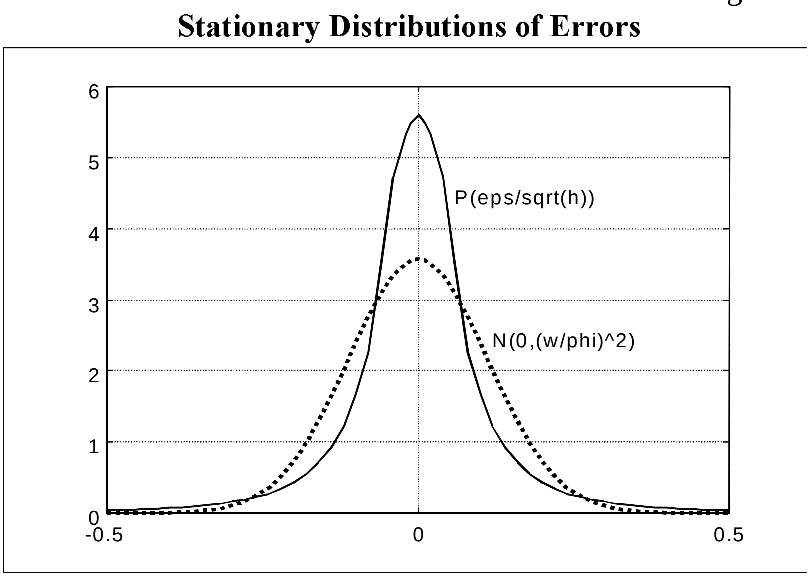
This Figure compares the log-normal density generated by the Andersen-Lund estimates (dotted line) and the density $f_1(\sigma)$ in (10) generated by the estimates of paragraph 5.

Figure 3



This Figure compares the linear drift function generated by the Andersen-Lund estimates (dotted line) and the nonlinear drift function in eq. (11) generated by the estimates of paragraph 5.

Figure 4



In this Figure, $P(\frac{\epsilon}{\sqrt{h}})$ is the approximate invariant distribution of the errors in model (30) rescaled by $\sqrt{h}$ (see eqs. (22)). $N(0, (\frac{\omega}{\varphi})^2)$ is instead a normal density with standard deviation fixed at the steady state expectation of the volatility process (see model (2)).

Figure 5

First difference of the interest rate and estimated volatility

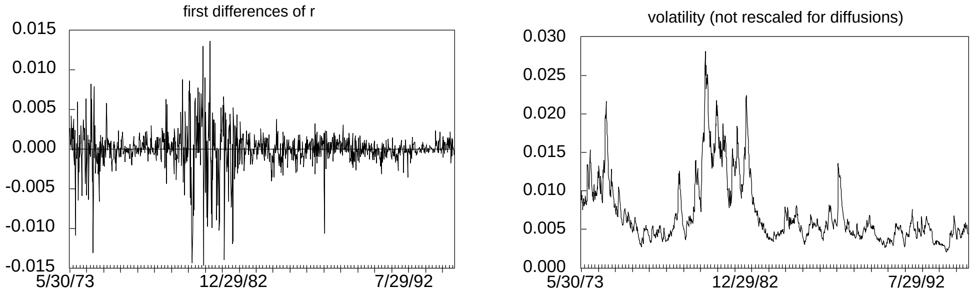


Figure 6

Monte Carlo densities of the ARCH parameters estimates

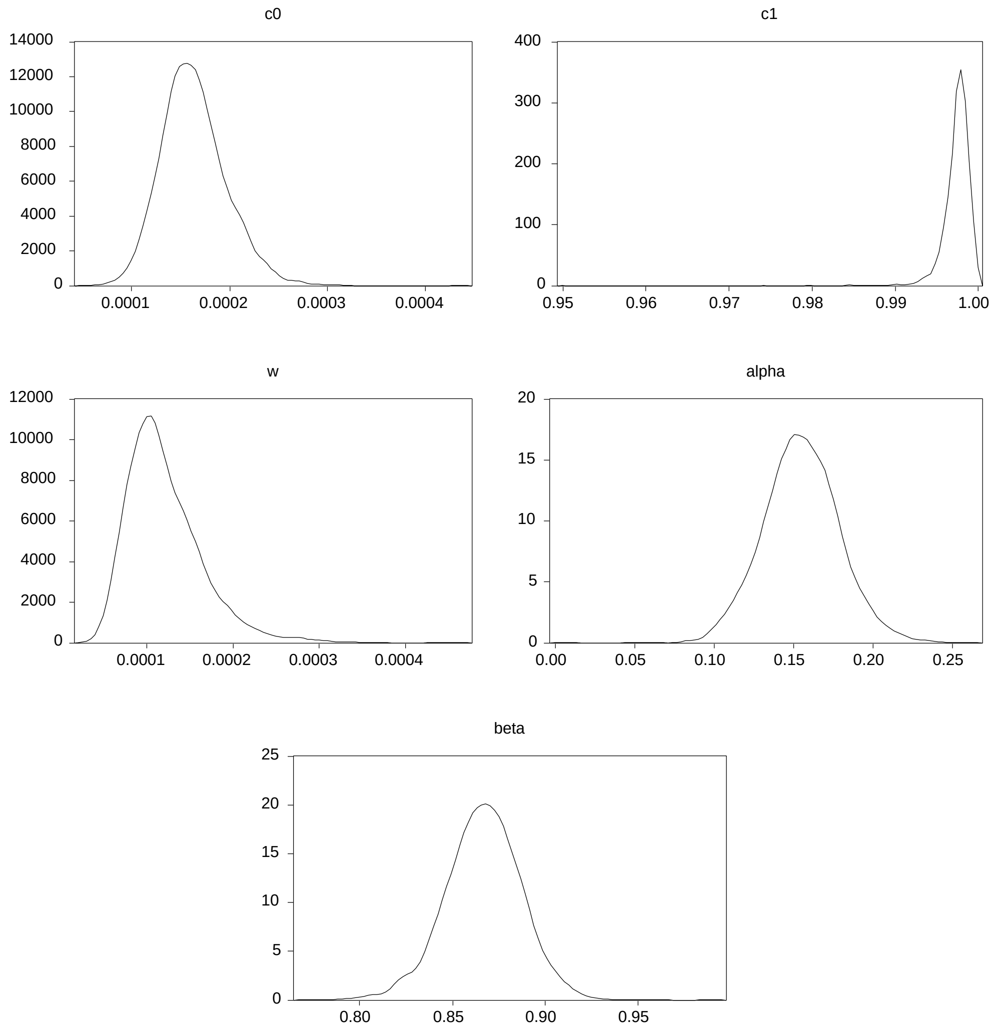
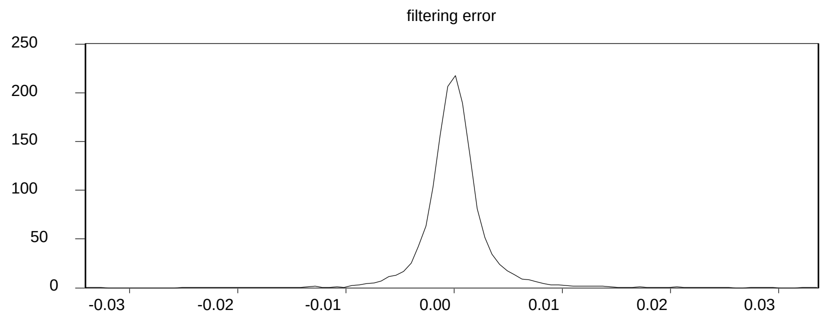


Figure 7

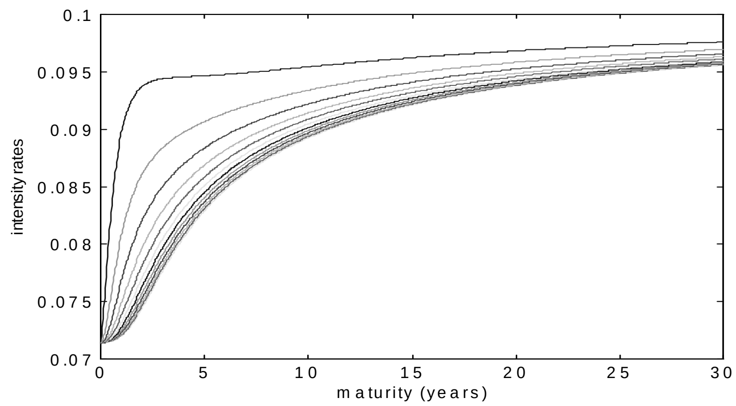
Monte Carlo fitering error



The filtering error of the conditional volatility is evaluated over 5,000 simulations as $\{E_i\}_{i=1}^{5000}$, where $E_i \equiv \frac{1}{1135}\sum_{n=1}^{1135}(\sigma_{i,n} - \widehat{\sigma}_{i,n})$ with 1,135 being sample size, $\sigma_{i,n}$ and $\widehat{\sigma}_{i,n}$ the true and the predicted volatility. The Monte Carlo distribution of the average filtering error has an average of $9.61 \cdot 10^{-5}$ and a standard deviation of $3.275{\cdot}10^{-3}$.

Figure 8

Term structure



From bottom to top, the curves correspond to values of $\sigma(0)$ equal to 0.0100, 0.0114, 0.0131, 0.0150, 0.0173, 0.0200, 0.0233, 0.0275, 0.0329, 0.0400, 0.0500, 0.0650. The average level of the (rescaled) volatility was 0.0490, and the curve corresponding to $\sigma(0) = 0.0500$ was fitted to the target term structure.

Appendix

Appendix A: Convergence results for section 4

Proof of theorem 4.1

Conditions (18) are sufficient to establish the weak convergence of the short-term rate and volatility processes toward the solutions of the following stochastic differential equations:

$$\begin{cases} dr(\tau) & = & (\iota - \theta r(\tau))d\tau + \sigma(\tau)\sqrt{r(\tau)}dW^{(1)}(\tau) \\ d\sigma(\tau)^{\delta} & = & (\omega - \varphi\sigma(\tau)^{\delta})d\tau + \psi\sigma(\tau)^{\delta}dW^{(3)}(\tau) \end{cases}$$

where $\{W^{(j)}(\tau)\}_{\tau\geq\mathbf{0}}$, $j = 1$ and 3, are $\mathcal{F}(\tau)$-Brownian motions. This has been shown in thm. 2.3 p. 209-211 of Fornari and Mele (1997a) in the case of a geometric Brownian motion, and the case of a square root process follows easily by an extension of another convergence result (see appendix B in Fornari and Mele (2000b) for further references).

It remains to show that $W^{(3)}(\tau)$ can be written as:

$$W^{(3)}(\tau) = \rho W^{(1)}(\tau) + \sqrt{1-\rho^2}W^{(2)}(\tau),\ \tau \geq 0$$

with $\{W^{(2)}(\tau)\}_{\tau\geq\mathbf{0}}$ another $\mathcal{F}(\tau)$-Brownian motion. It is sufficient to show that the limit:

$$\lim_{h\downarrow\mathbf{0}} h^{-1}E\left\{\left({}_h r_{hk} - {}_h r_{h(k-1)}\right)\left({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk}\right) \mid \mathcal{F}_{hk}\right\}$$

is not ill-behaved. After that, an identification argument will do the work.

By (18), and the fact that $\frac{{}_h u_{hk}}{\sqrt{h}}$ is g.e.d.$_{(\upsilon)}$ for each h,

$$\lim_{h\downarrow 0} h^{-1}E\left\{\left({}_h r_{hk} - {}_h r_{h(k-1)}\right)\left({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk}\right) \mid \mathcal{F}_{hk}\right\}$$

$$= \lim_{h\downarrow 0} h^{-1}E\left\{\left(\iota_h - \theta_h \cdot {}_h r_{h(k-1)} + {}_h\sigma_{hk}\sqrt{{}_h r_{h(k-1)}} \cdot {}_h u_{hk}\right)\right.$$
$$\left.\times\left(w_h + \left(\alpha_h \left|{}_h u_{hk}\right|^{\delta}(1-\gamma s_k)^{\delta} h^{-\frac{\delta}{2}} + \beta_h - 1\right) {}_h\sigma^{\delta}_{hk}\right) \mid \mathcal{F}_{hk}\right\}$$

$$= \lim_{h\downarrow 0} h^{-1}E\left\{{}_h u_{hk}\left(\alpha_h \left|{}_h u_{hk}\right|^{\delta}(1-\gamma s_k)^{\delta} h^{-\frac{\delta}{2}} + \beta_h - 1\right) \cdot {}_h\sigma^{\delta+1}_{hk} \mid \mathcal{F}_{hk}\right\} \cdot \sqrt{{}_h r_{h(k-1)}}$$

$$= \lim_{h\downarrow 0} h^{-1-\frac{\delta}{2}}\alpha_h \cdot E\left\{{}_h u_{hk}\left|{}_h u_{hk}\right|^{\delta}(1-\gamma s_k)^{\delta} \cdot {}_h\sigma^{\delta+1}_{hk} \mid \mathcal{F}_{hk}\right\} \cdot \sqrt{{}_h r_{h(k-1)}}$$

$$= \lim_{h\downarrow 0} \frac{\alpha_h}{\sqrt{h}}\left\{\left((1-\gamma)^{\delta} - (1+\gamma)^{\delta}\right)\int_{\mathbb{R}_+} x^{\delta+1}p(dx)\right\} \cdot {}_h\sigma^{\delta+1}_{hk} \cdot \sqrt{{}_h r_{h(k-1)}},$$

where $p(.)$ denotes the g.e.d.$_{(\upsilon)}$ density, or:

$$\lim_{h\downarrow 0} h^{-1}E\left\{({}_h r_{hk} - {}_h r_{h(k-1)})({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk}) \mid \mathcal{F}_{hk}\right\}$$

$$= \lim_{h\downarrow 0} \frac{\alpha_h}{\sqrt{h}}\left\{(1-\gamma)^{\delta} - (1+\gamma)^{\delta}\right\} K \cdot {}_h\sigma^{\delta+1}_{hk} \cdot \sqrt{{}_h r_{h(k-1)}};$$

here,

$$K = \frac{2^{\frac{\delta-\upsilon+1}{\upsilon}}\nabla_{\upsilon}^{\delta+1}\Gamma(\frac{\delta+2}{\upsilon})}{\Gamma(\upsilon^{-1})}.$$

By using (18),

$$\lim_{h\downarrow 0}\frac{\alpha_h}{\sqrt{h}} = \frac{\psi}{\sqrt{Z}},$$

where $Z \equiv (m_{\delta,\upsilon} - n^2_{\delta,\upsilon})\left((1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}\right) - 2n^2_{\delta,\upsilon}(1-\gamma)^{\delta}(1+\gamma)^{\delta}$.

Hence,

$$\lim_{h\downarrow 0} h^{-1}E\left\{\left({}_h r_{hk} - {}_h r_{h(k-1)}\right)\left({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk}\right) \mid \mathcal{F}_{hk}\right\} = \frac{\psi\overline{K}}{\sqrt{Z}}\,\sigma^{\delta+1} \cdot \sqrt{r},$$

where:

$$\overline{K} = \left((1-\gamma)^{\delta} - (1+\gamma)^{\delta}\right) K.$$

To identify ρ, we note that this has to solve the following equation: $\psi\rho = \frac{\psi}{\sqrt{Z}}\overline{K}$, which yields:

$$\rho = \frac{\overline{K}}{\sqrt{Z}}.$$

The proof is complete.

Proof of theorem 4.2

Nearly identical to the proof of theorem 4.1.

Construction of alternate converging asymmetric models

It is well known that in correspondence with a given diffusion model, there may exist many well-behaved discrete time models converging in distribution to the given continuous time model. Hence, we can find other examples of discrete time ARCH-type models converging to model (2). As an example, consider the following model:

$$\text{(A1)} \qquad \sigma_{n+1}^{\delta} = w + \beta\sigma_n^{\delta} + \alpha(1-\gamma s_n)^{\delta\eta}\left(|u_n|^{\delta\eta} - E\left(|u_n|^{\delta\eta}\right)\right)\sigma_n^{\delta\eta},\ \gamma \in (-1,1).$$

The main difference between model (21) and model (A1) is the way how asymmetries in volatility are modeled. Suppose for instance that $\gamma > 0$ in model (A1). In this case, 'large' negative shocks introduce more volatility than positive shocks of the same size, while 'small' negative shocks introduce less volatility than positive shocks of the same size. Such a phenomenon, referred to as 'volatility reversal' in Fornari and Mele (1997b), seems to be pervasive in many stock markets and in this respect, model (A1) represents another example of the volatility-switching ARCH models that were originally introduced by Fornari and Mele (1997b).

Our objective now is to give a sketch of the proof that (A1) converges in distribution to (2) as the sampling frequency gets higher and higher. Consider the following approximating scheme:

$$ {}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} = w_h - (1-\beta_h)\, {}_h\sigma^{\delta}_{hk} + \alpha_h (1-\gamma s_k)^{\delta\eta} \left\{ |{}_h u_{hk}|^{\delta\eta} - E\left(|{}_h u_{hk}|^{\delta\eta} \right) \right\} h^{-\frac{\delta\eta}{2}}\, {}_h\sigma^{\delta}_{hk}, $$

and introduce the following moment conditions:

$$ \text{(A2)} \qquad \begin{array}{lcl} \lim_{h\downarrow 0} h^{-1} w_h & = & \omega \in (0,\infty), \\ \lim_{h\downarrow 0} h^{-1}(1-\beta_h) & = & \varphi < \infty, \\ \lim_{h\downarrow 0} h^{-1/2} \left\{ (1+\gamma)^{2\delta\eta} + (1-\gamma)^{2\delta\eta} \right\} (m_{\delta\eta,\upsilon} - 2n^2_{\delta\eta,\upsilon})\alpha_h & = & \psi < \infty. \end{array} $$

For each h, we have that

$$ E\left\{ (1-\gamma s_k)^{\delta\eta} \left(|{}_h u_{hk}|^{\delta\eta} - E\left(|{}_h u_{hk}|^{\delta\eta} \right) \right) h^{-\frac{\delta\eta}{2}} \mid \mathcal{F}_{hk} \right\} = 0, $$

and so the drift per unit of time is:

$$ h^{-1} E\left({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} \mid \mathcal{F}_{hk} \right) = h^{-1} w_h - h^{-1}(1-\beta_h)\, {}_h\sigma^{\delta}_{hk}. $$

By taking limits for $h \downarrow 0$, and using the moment conditions (A2), we obtain the drift function of volatility in (2).

Now consider the second order moment per unit of time $h^{-1}E\{({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk})^2 \mid \mathcal{F}_{hk}\}$. By taking limits for $h \downarrow 0$, and using again the moment conditions in (A2), yields after tedious computations:

$$ \begin{aligned} & \lim_{h\downarrow 0} h^{-1} E\left\{ \left({}_h\sigma^{\delta}_{h(k+1)} - {}_h\sigma^{\delta}_{hk} \right)^2 \mid \mathcal{F}_{hk} \right\} \\ = \;& \lim_{h\downarrow 0} \left(\frac{\alpha_h}{\sqrt{h}} \right)^2 \left\{ E\left((1-\gamma s_k)^{2\delta\eta} \left| \frac{{}_h u_{hk}}{\sqrt{h}} \right|^{2\delta\eta} \right) + 4n^2_{\delta\eta,\upsilon} E\left((1-\gamma s_k)^{2\delta\eta} \right) \right. \\ & \qquad \left. - 4n_{\delta\eta,\upsilon} E\left((1-\gamma s_k)^{2\delta\eta} \left| \frac{{}_h u_{hk}}{\sqrt{h}} \right|^{\delta\eta} \right) \right\} {}_h\sigma^{2\delta\eta}_{hk} \\ = \;& \lim_{h\downarrow 0} \left(\frac{\alpha_h}{\sqrt{h}} \right)^2 \left\{ (1+\gamma)^{2\delta\eta} + (1-\gamma)^{2\delta\eta} \right\} (m_{\delta\eta,\upsilon} - 2n^2_{\delta\eta,\upsilon})\, {}_h\sigma^{2\delta\eta}_{hk}, \end{aligned} $$

which gives the diffusion function of volatility in (2).

As regards correlation issues, the proof is very similar to that of thm. 4.1:

$$\lim_{h\downarrow 0} h^{-1}E\left\{\left({}_hr_{hk}-\ {}_hr_{h(k-1)}\right)\left({}_h\sigma^{\delta}_{h(k+1)}-\ {}_h\sigma^{\delta}_{hk}\right)\mid \mathcal{F}_{hk}\right\}$$

$$= \lim_{h\downarrow 0}\frac{\alpha_h}{\sqrt{h}}\cdot E\left\{\frac{{}_hu_{hk}}{\sqrt{h}}(1-\gamma s_k)^{\delta\eta}\left(\left|\frac{{}_hu_{hk}}{\sqrt{h}}\right|^{\delta\eta}-2n_{\delta\eta,\upsilon}\right)\mid \mathcal{F}_{hk}\right\}\cdot\ {}_h\sigma^{\delta\eta+1}_{hk}\cdot\sqrt{{}_hr_{h(k-1)}},$$

where

$$\lim_{h\downarrow 0} E\left\{\frac{{}_hu_{hk}}{\sqrt{h}}(1-\gamma s_k)^{\delta\eta}\left(\left|\frac{{}_hu_{hk}}{\sqrt{h}}\right|^{\delta\eta}-2n_{\delta\eta,\upsilon}\right)\right\}$$

$$= E\left\{\widetilde{u}\,(1-\gamma\cdot \text{sign}(\widetilde{u}))^{\delta\eta}\left(|\widetilde{u}|^{\delta\eta}-2n_{\delta\eta,\upsilon}\right)\right\}$$

$$= \left\{(1-\gamma)^{\delta\eta}-(1+\gamma)^{\delta\eta}\right\}\frac{2^{\frac{\delta\eta-\upsilon+1}{2}}\nabla_{\upsilon}^{\delta\eta+1}}{\Gamma(\frac{1}{\upsilon})^2}\left\{\Gamma\left(\frac{\delta\eta+2}{\upsilon}\right)\Gamma\left(\frac{1}{\upsilon}\right)-\Gamma\left(\frac{\delta\eta+1}{\upsilon}\right)\Gamma\left(\frac{2}{\upsilon}\right)\right\},$$

and $\widetilde{u}$ is ged_{υ}.

Using an identification device as in the proof of theorem 4.1, we find that:

$$\rho=\frac{\left\{(1-\gamma)^{\delta\eta}-(1+\gamma)^{\delta\eta}\right\}\frac{2^{\frac{\delta\eta-\upsilon+1}{2}}\nabla_{\upsilon}^{\delta\eta+1}}{\Gamma(\frac{1}{\upsilon})^2}\left\{\Gamma\left(\frac{\delta\eta+2}{\upsilon}\right)\Gamma\left(\frac{1}{\upsilon}\right)-\Gamma\left(\frac{\delta\eta+1}{\upsilon}\right)\Gamma\left(\frac{2}{\upsilon}\right)\right\}}{\left\{(1+\gamma)^{2\delta\eta}+(1-\gamma)^{2\delta\eta}\right\}\left(m_{\delta\eta,\upsilon}-2n^2_{\delta\eta,\upsilon}\right)}.$$

In correspondence with reasonable values of δ,η and υ, the term $\{\Gamma\left(\frac{\delta\eta+2}{\upsilon}\right)\Gamma\left(\frac{1}{\upsilon}\right)-\Gamma\left(\frac{\delta\eta+1}{\upsilon}\right)\Gamma\left(\frac{2}{\upsilon}\right)\}$ is strictly positive, thus restricting sign(ρ) to be minus sign(γ), as in thms. 4.1 and 4.2.

Appendix B: Standard regularity conditions and the convergence of the criterion

ASSUMPTION B1.

-plim$_N$ $\mathfrak{L}_N({}_{\Delta}r;b)=\mathfrak{L}_{\infty}(a_0;b)$, say, uniformly in $b\in B\subset\mathbb{R}^5$.
-plim$_N$ $\frac{\partial^2\mathfrak{L}_N}{\partial b\partial b'}({}_{\Delta}r;b)=\ddot{\mathfrak{L}}_{\infty}(a_0;b)$, say, uniformly in $b\in B$. Further, $\ddot{\mathfrak{L}}_{\infty}(.)$ is invertible.
-$\left[\sqrt{N}\frac{\partial\mathfrak{L}_N}{\partial b}({}_{\Delta}r;b)\right]_{b=b_0(a_0)}\xrightarrow{d}N(0,J(a_0))$.

CONVERGENCE OF THE CRITERION (Sketch). We assume as in Broze et al. (1998) the continuity of the partial application $a \mapsto \widehat{b}_{N,s}^{(h)}(a)$, and for the case $S = 1$, we define $\widehat{b}_N^{(h)}(.) \equiv \widehat{b}_{N,1}^{(h)}(.)$ and ${}_\Delta\widetilde{r}(.) \equiv {}_{\Delta,h}\widetilde{r}^{(1)}(.)$. It is not hard to show that under conditions on $\mathfrak{L}_N({}_\Delta\widetilde{r}(a); b)$ that parallel those in assumption B1 stated above for the direct criterion $\mathfrak{L}_N({}_\Delta r; b)$, the simulated estimator is asymptotic normal:

$$\sqrt{N}\left(\widehat{b}_N^{(h)}(a) - b_0^{(h)}(a)\right) \xrightarrow{d} N\left(0, \ddot{\mathfrak{L}}_\infty^{(h)-1}(a; b_0^{(h)}(a)) \cdot J^{(h)}(a) \cdot \ddot{\mathfrak{L}}_\infty^{(h)-1}(a; b_0^{(h)}(a))\right),$$

where $b_0^{(h)}(a) = \arg\max_b \mathfrak{L}_\infty^{(h)}(a; b)$, the limit simulation problem, and $\ddot{\mathfrak{L}}_\infty^{(h)}(.)$ and $J^{(h)}(.)$ are defined similarly as $\ddot{\mathfrak{L}}_\infty(.)$ and $J(.)$. Now, it follows from thm. 4.2 that the solution of (22): $\{{}_h r_{hk}, {}_h\sigma_{hk}^\delta\}_{k=0,1,\cdots} \Rightarrow \{r(\tau), \sigma(\tau)^\delta\}_{\tau\geq 0}$ (the solution of (2)). By this, an extension of a result cited in Fornari and Mele (2000b) (appendix B) that shows that the solution of (22) is unique, stationary and ergodic (for fixed h), and assuming the uniform continuity of the criterion $\mathfrak{L}_N(.; b)$, it follows that $\mathfrak{L}_N({}_\Delta\widetilde{r}(a_0); b) \Rightarrow \mathfrak{L}_N({}_\Delta r(a_0); b)$ as $h \downarrow 0$, and we suppose, as in Broze et al. (1998), that the convergence is uniform in b. Finally, because $\text{plim}_N\, \mathfrak{L}_N({}_\Delta\widetilde{r}(a); b) = \mathfrak{L}_\infty^{(h)}(a; b)$ and $\text{plim}_N\, \mathfrak{L}_N({}_\Delta r; b) = \mathfrak{L}_\infty(a_0; b)$, uniformly in $b \in B$ (both by assumption), one can easily verify that for small h, this implies $b_0^{(h)}(a_0) = \arg\max_b \mathfrak{L}_\infty^{(h)}(a_0; b) = \arg\max_b \mathfrak{L}_\infty(a_0; b) = b_0(a_0)$. This is:

$$\lim_{h\downarrow 0} b_0^{(h)}(a_0) = b_0(a_0),$$

while for fixed h, it is assumed that there exists only one solution to the system $b_0^{(h)}(a) = b_0(a_0)$: this has the form $\mathcal{A}^{(h)}(a_0)$, with $\lim_{h\downarrow 0} \mathcal{A}^{(h)}(a_0) = a_0$. Now by proposition 6 in Broze et al. (1998), one has that $\sqrt{N}\left({}_h\widehat{a}_N(a_0) - \mathcal{A}^{(h)}(a_0)\right) \xrightarrow{d} \mathrm{N}\left(0, \Sigma^{(h)}\right)$ (for fixed h), where $\Sigma^{(h)}$ is such that $\lim_{h\downarrow 0} \Sigma^{(h)} = 2V_0^{-1}\Gamma_0 V_0'$, and (28) follows for $S = 1$. In the preceding expressions, Γ_0 is defined as the limit of $\Gamma_0^{(h)}$ as $h \downarrow 0$, V_0 is defined similarly, and $\Gamma_0^{(h)}$ is the limit of $\ddot{\mathfrak{L}}_\infty^{(h)-1}(a; b_0^{(h)}(a)) \cdot J^{(h)}(a) \cdot \ddot{\mathfrak{L}}_\infty^{(h)-1}(a; b_0^{(h)}(a))$ as $N \uparrow \infty$, whereas $V_0^{(h)}$ is the limit of $\left[\frac{\partial \widehat{b}_N^{(h)}}{\partial a}(a)\right]_{a=\mathcal{A}^{(h)}(a_0)}$ as $N \uparrow \infty$. The case $S > 1$ is similar.

Appendix C: How to rescale volatility for diffusions?

Here we provide details on how we rescaled ARCH-filtered volatility for diffusions. Let us rewrite the first equation of the Euler-Maruyama discrete approximation of (2) in (3) as:

(C1) $$r_n = \iota h + (1-\theta h) r_{n-1} + \sqrt{h}\sigma_{n-1}\sqrt{r_{n-1}}u_n,\ n = 1, \cdots, \widetilde{N},$$

where $\widetilde{N}$ denotes the total number of points generated by the simulations and u_n is $NID(0,1)$. Simulated data are sampled every ℓ points. Iterating (C1) leaves:

$$\begin{aligned} r_n \;=\; & \frac{\iota}{\theta}\left\{1-(1-\theta h)^{\ell}\right\} + (1-\theta h)^{\ell} r_{n-\ell} \\ & + \sqrt{h}\left\{\sigma_{n-1}\sqrt{r_{n-1}}u_n + (1-\theta h)\,\sigma_{n-2}\sqrt{r_{n-2}}u_{n-1} + (1-\theta h)^2\,\sigma_{n-3}\sqrt{r_{n-3}}u_{n-2} \right. \\ & \left. + \cdots + (1-\theta h)^{\ell-1}\,\sigma_{n-\ell}\sqrt{r_{n-\ell}}u_{n-(\ell-1)}\right\}. \end{aligned}$$

Because a diffusion is continuous with locally bounded paths, when h is low enough r and σ do not move too much within the unsampled ℓ subintervals. Let us denote with $\overline{r}^{\ell}_{n-1}$ and $\overline{\sigma}^{\ell}_{n-1}$ the (random) representative, fictious values of r and σ within the unsampled intervals that are such that the previous equation can be written approximately as:

$$\begin{aligned} r_n \;=\; & \frac{\iota}{\theta}\left\{1-(1-\theta h)^{\ell}\right\} + (1-\theta h)^{\ell} r_{n-\ell} \\ & + \overline{\sigma}^{\ell}_{n-1}\cdot\sqrt{h}\left\{u_n + (1-\theta h)\,u_{n-1} + \cdots + (1-\theta h)^{\ell-1}\,u_{n-(\ell-1)}\right\}\sqrt{\overline{r}^{\ell}_{n-1}}. \end{aligned}$$

Our objective is to estimate each point of the sequence $\left\{\overline{\sigma}^{\ell}_{j}\right\}_{j=\ell,2\ell,\cdots,\widetilde{N}/\ell}$ in order to use it to filter the actual (discretely sampled) volatility path generated by the second equation of the Euler-Maruyama discrete approximation of (2) in (3): $\{\sigma_j\}_{j=1}^{\widetilde{N}/\ell} = \{\sigma(\ell\cdot j)\}_{j=1}^{\widetilde{N}/\ell}$.

Rewrite the previous equation as:

$$\begin{aligned}
r_n &= \frac{\iota}{\theta}\left\{1-(1-\theta h)^{\ell}\right\}+(1-\theta h)^{\ell}\, r_{n-\ell} \\
&\quad +\overline{\sigma}_{n-1}^{\ell}\cdot\sqrt{h}\sqrt{1+(1-\theta h)^2+(1-\theta h)^4+\cdots+(1-\theta h)^{2(\ell-1)}}\cdot\sqrt{\overline{r}_{n-1}^{\ell}}\cdot\widetilde{u}_{n-1}^{\ell} \\
&= \frac{\iota}{\theta}\left\{1-(1-\theta h)^{\ell}\right\}+(1-\theta h)^{\ell}\, r_{n-\ell}+\overline{\sigma}_{n-1}^{\ell}\cdot\sqrt{\frac{h\left(1-(1-\theta h)^{2\ell}\right)}{1-(1-\theta h)^2}}\cdot\sqrt{\overline{r}_{n-1}^{\ell}}\cdot\widetilde{u}_{n}^{\ell},
\end{aligned}$$

where $\widetilde{u}^{\ell}$ is a standard Gaussian variate.

Now all the models we used in this paper actually deliver an estimate of

$$v_j \equiv \overline{\sigma}_j^{\ell}\cdot\sqrt{\frac{h\left(1-(1-\theta h)^{2\ell}\right)}{1-(1-\theta h)^2}},\ j=\ell,2\ell,\cdots,\widetilde{N}/\ell\,. \tag{C2}$$

Therefore, an estimate of each point of the sequence $\left\{\overline{\sigma}_j^{\ell}\right\}_{j=\ell,2\ell,\cdots,\widetilde{N}/\ell}$ is obtained by inverting formula (C2) to form the desired sequence:

$$\overline{\sigma}_j^{\ell} = \sqrt{\frac{1-(1-\theta h)^2}{h\left(1-(1-\theta h)^{2\ell}\right)}}\cdot v_j,\ j=\ell,2\ell,\cdots,\widetilde{N}/\ell\,. \tag{C3}$$

In this paper, we used:

$$h=\frac{1}{\Delta\cdot\ell},\ \Delta=52,\ \ell=25,$$

and the estimates of θ reported in table 6 of the main text are such that $\sqrt{\frac{1-(1-\theta h)^2}{h\left(1-(1-\theta h)^{2\ell}\right)}}$ is always close to 7.218.

The filtered series of volatility reported throughout the paper are based on formula (C3) (see, however, below for numerical improvements of this formula). To relate the number found before to the correction given in formula (20) for the intercept of the volatility equation, note that:

$$\omega_{\text{q-aml}} = \Delta^{-1}\cdot\Delta^{-1/2}\cdot\widehat{w}_{\Delta}.$$

Here the correcting term is $\Delta^{-1/2} = 7.211$, which in practice is very close to the conversion factor given above.

In addition to being based on the stability of volatility within unsampled periods, the conversion formula (C3) is based on the assumption that the (small) changes of $\sigma\sqrt{r}$ are not autocorrelated. Relaxing such an assumption requires a much more complicated approach with continuous updatings. A reliable alternative consists in finding numerically a conversion formula similar to (C3). In this paper, we proceeded in the following way. We simulated 5000 times the continuous time system (2) in correspondence of the parameter estimates found in paragraph 5. Then we defined:

$$\aleph \equiv \frac{1}{5000 \cdot \left(\widetilde{N}/\ell\right)} \sum_{i=1}^{5000} \sum_{j=1}^{\widetilde{N}/\ell} \frac{\sigma_i(\ell \cdot j)}{v_{ij}},$$

where $\sigma_i(\ell \cdot j)$ and v_{ij} are simulated volatility and filtered volatility as of time j obtained in the ith simulation. We found that $\aleph \approx 6.928$.

Appendix D: A supporting equilibrium for Paragraph 6

In this appendix we construct a supporting equilibrium for the model analysed in section V. To save space, only a sketch is provided of this construction; further details can be found in Fornari and Mele (2000b;appendix A).

Construction of an equilibrium state price density

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $T < \infty$, and $\mathbb{F} = \{\mathcal{F}(\tau)\}_{\tau\in[0,T]}$ the P-augmentation of the natural filtration $\mathcal{F}^W(\tau) = \sigma(W(s), s \le \tau)$ generated by a Brownian motion in $\mathbb{R}^2$: $W = \{W(\tau) = (W^{(1)}(\tau), W^{(2)}(\tau))'\}_{\tau\in[0,T]}$ (with $\mathcal{F} = \mathcal{F}(T)$). We consider a diffusion state process $y \equiv (y_1, y_2)$ solution of

$$\text{(D1)}\quad d\begin{pmatrix} y_1(\tau) \\ y_2(\tau) \end{pmatrix} = \begin{pmatrix} \kappa - \theta \cdot y_1(\tau) \\ \widetilde{\omega} - \varphi \cdot y_2(\tau) \end{pmatrix} d\tau + \begin{pmatrix} \sqrt{y_1(\tau)} \cdot y_2(\tau) & 0 \\ 0 & \psi \cdot y_2(\tau) \end{pmatrix} dW(\tau),$$

and assume that the various constants $\kappa, \theta, \widetilde{\omega}, \varphi$ and ψ are such that the preceding system admits a strong solution (Karatzas and Shreve, 1991, definition 2.1, p. 285). Let

$\{(M(\tau), S(\tau))'\}_{\tau \in [0,T]}$ be the $\mathcal{F}(\tau)$-adapted stochastic process representing the price of an accumulation factor (M) plus one primitive asset entitling to rights on the fruits, or dividends (the numéraire), of one tree—as in the discrete time model of Lucas (1978)—. We assume that the asset price is the solution of:

$$dS(\tau)/S(\tau) = a(\tau, y(\tau))d\tau + \sum_{j=1}^{2} \int_0^\tau \nu_j(s, y(s))dW^{(j)}(s),$$

where $a(\tau, y)$ is the total expected appreciation rate of the asset price, and is equal to a function $\widetilde{a}(\tau, y) + (\zeta/S)(\tau, y)$; ζ is the dividend rate; and $\widetilde{a}(\tau, y), \nu_j(\tau, y), j = 1, 2,$ are well specified functions. We assume that $\zeta(\tau) = (a - \widetilde{a})(\tau) \cdot S(\tau)$, where $(a - \widetilde{a})(\tau)$ is a deterministic function of time. This implies that

(D2) $$a_\zeta(\tau, y) = a_k(\tau) + a(\tau, y) \text{ and } \nu_\zeta(\tau, y) = \nu(\tau, y),$$

where a_ζ and ν_ζ denote drift and volatility-vector processes of $d\zeta/\zeta$, and a_k is the (deterministic) time-varying growth rate of $(a - \widetilde{a})$.

We now impose restrictions that represent a natural generalization of the Cox et al. (1985) single factor model:

(D3) $$\begin{cases} a(\tau) &= \widehat{a} \cdot y_1(\tau) \\ \nu(\tau) &= \widehat{\nu} \cdot \sqrt{y_1(\tau)} \end{cases}$$

where $\widehat{a}$ and $\widehat{\nu}$ are, respectively, a constant and a vector of constants in $\mathbb{R}^2$.

Let $\mu_{0,T}(\omega) \equiv M(\omega, T)^{-1}\xi(\omega, T)$ be the Arrow-Debreu pricing kernel of one unit of numéraire at T at the point $\omega \in \Omega$. In the notation of the preceding definition, $\xi(T)$ is the Radon-Nikodym derivative of any $Q \in \mathcal{Q}$ on $\mathcal{F}(T)$ with density process

(D4) $$\xi(\tau) \equiv \left.\left|\frac{dQ}{dP}\right|\right._{\mathcal{F}(\tau)} = \exp\left(-\int_0^\tau \overline{\lambda}'(s)dW(s) - \frac{1}{2}\int_0^\tau \left\|\overline{\lambda}(s)\right\|^2 du\right),$$

where $\{\overline{\lambda}(\tau)\}_{\tau \in [0,T]}$ is a $\mathcal{F}(\tau)$-adapted process that satisfies the Novikov's condition: $E\left(\exp\left(\frac{1}{2}\int_0^\tau \left\|\overline{\lambda}(s)\right\|^2 ds\right)\right) < \infty$, and has to be determined at the equilibrium.

Consider a representative agent maximizing $(1 - \ell)^{-1}E\left(\int_0^T e^{-g\cdot\tau}c(\tau)^{1-\ell}d\tau + x_+^{1-\ell}\right)$ under the constraint that $x = E\left(\mu_{0,T} \cdot x_+ + \int_0^T \mu_{0,\tau} \cdot c(\tau)d\tau\right)$, where x and x_+ are initital and final wealth, respectively, and g and ℓ are constants satisfying usual restrictions as well as viability restrictions to be determined below. Here markets are complete because our agent can trade with one stock and one bond the equilibrium price of which (v, say) will be determined below. After differentiating the first order conditions of this agent's problem we get:

$$d\log c(\tau) = \ell^{-1}\left(-g + r(\tau) + \frac{1}{2}\left\|\overline{\lambda}(\tau)\right\|^2\right)d\tau + \ell^{-1}\overline{\lambda}'(\tau)dW(\tau).$$

Using the equilibrium condition $c = \zeta$, and identifying both drift and diffusion terms of the preceding equation leaves:

$$\text{(D5)} \qquad \begin{cases} r(\tau, y(\tau)) &= g + \ell\left(a_k(\tau) + a(\tau, y(\tau))\right) - \frac{\ell(\ell+1)}{2}\left\|\nu(\tau, y(\tau))\right\|^2 \\ \overline{\lambda}'(\tau, y(\tau)) &= \ell \cdot \nu(\tau, y(\tau)). \end{cases}$$

Furthermore, by the martingale property of $M^{-1}S$ and $M^{-1}v$ under the measure Q that we are looking for, it must be the case that:

$$\text{(D6)} \qquad r(\tau, y(\tau)) = a(\tau, y(\tau)) - \nu(\tau, y(\tau)) \cdot \overline{\lambda}(\tau, y(\tau)),$$

as well as

$$\text{(D7)} \qquad r(\tau, y(\tau)) = a^b(\tau, y(\tau)) - \nu^b(\tau, y(\tau)) \cdot \overline{\lambda}(\tau, y(\tau)),$$

where a^b and ν^b are drift and volatility-vector processes of dv/v, and Q has density process as in (D4) with $\overline{\lambda}$ given by the second relation in (D5).

Substituting (D3) and the second relation of (D5) into (D6) leaves:

$$\text{(D8)} \qquad r(\tau) = \left(\widehat{a} - \ell \cdot \|\widehat{\nu}\|^2\right) \cdot y_1(\tau).$$

By differentiating r and using (D1), we get the first equation of system (1), and by differentiating the resulting volatility function we get the second equation in (1) whenever $\widehat{a} - \ell \cdot \|\widehat{\nu}\|^2 > 0$, in which case r is also positive. Finally, using again (D3) and the second equation in (D5),

$$\overline{\lambda}(\tau, y(\tau)) = \ell\widehat{\nu}'\sqrt{y_1(\tau)} = \ell\widehat{\nu}'\sqrt{1/\left(\widehat{a} - \ell\|\widehat{\nu}\|^2\right)}\sqrt{r(\tau)} \equiv \begin{pmatrix}\lambda_1 \\ \lambda_2\end{pmatrix} \cdot \sqrt{r(\tau)},$$

where λ_1 and λ_2 are two constants. Under all of our assumptions, $\{\overline{\lambda}(\tau)\}_{\tau\in[0,T]}$ satisfies the Novikov's condition. Now it follows by (D7), and Itô's lemma, that the bond price satisfies exactly the partial differential equation (35).

Viability restrictions

While deriving the previous results, we did not fully analyse what (D6) and system (D5) imply. Taking account of this imposes further restrictions that guarantee the internal consistency of the model, which we call "viability restrictions". The starting point is to note that comparing (D5) with (D6) implies that the following must hold:

$$g = (1-\ell)\left(a(\tau) - \frac{\ell}{2}\left\|\nu(\tau, y(\tau))\right\|^2\right) - \ell a_k(\tau).$$

By using (D3), the previous relation yields:

$$g = (1-\ell)\left(\widehat{a} - \frac{\ell}{2}\left\|\widehat{\nu}\right\|^2\right) y_1(\tau) - \ell a_k(\tau).$$

For each $\ell \neq 0$, the previous relationship can hold when $(1-\ell)\left(\widehat{a} - \frac{\ell}{2}\left\|\widehat{\nu}\right\|^2\right) = 0$, $a_k = 0$ and $g = 0$. In this case we are left with two possible choices:

– $\widehat{a} = \frac{\ell}{2}\left\|\widehat{\nu}\right\|^2$.
– $\ell = 1$ (logarithmic utility).

In the first case, relation (D8) implies that $r(\tau) = -\frac{\ell}{2}\cdot\left\|\widehat{\nu}\right\|^2\cdot y_1(\tau)$: except when $\ell < 0$, r is always negative in this case. In the second case, relation (D8) implies that:

$$r(\tau) = \left(\widehat{a} - \left\|\widehat{\nu}\right\|^2\right)\cdot y_1(\tau).$$

Hence, the results of the previous subsection can be fully supported by an economy with a representative agent with logarithmic utility and zero discount rate, and dividends on the stock price that are proportional to the share price, with a proportionality factor (see eq. (D2)) that is constant over time.

Consistency tests à la Walras

The supporting equilibrium for the model of paragraph 6 was found without explicitly dealing with the dynamic portfolio choices. In fact, it is possible to show that the equilibrium conditions of the previous subsections entail the equilibrium conditions in the securities markets. The argument is shown in Fornari and Mele (2000b; Appendix A).

References

Aït-Sahalia, Y. (1996a), "Nonparametric Pricing of Interest Rate Derivative Securities", *Econometrica,* Vol. 64, pp. 527-60.

__________ (1996b), "Testing Continuous-Time Models of the Spot Interest Rate", *Review of Financial Studies*, Vol. 9, pp. 385-426.

__________ (2000), "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach", *Econometrica*, forthcoming.

__________ and A. Lo (1998), "Nonparametric Estimation of State Price Densities Implicit in Financial Asset Prices", *Journal of Finance*, Vol. 53, pp. 499-547.

____________________ (2000), "Nonparametric Risk Management and Implied Risk Aversion", *Journal of Econometrics*, Vol. 94, pp. 9-51.

___________, Y. Wang and F. Yared (1998), "Do Option Markets Correctly Price the Probabilities of Movements of the Underlying Asset?", *Journal of Econometrics*, forthcoming.

Andersen, T.G. and J. Lund (1997*a*), "Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate", *Journal of Econometrics*, Vol. 77, pp. 343-77.

_______________________ (1997*b*), "Stochastic Volatility and Mean Drift in the Short Rate Diffusion: Sources of Steepness, Level, and Curvature in the Yield Curve", Northwestern University, mimeo.

Bajeux, I. and J. C. Rochet (1996), "Dynamic Spanning: Are Options an Appropriate Instrument?", *Mathematical Finance*, Vol. 6, pp. 1-16.

Balduzzi, P., S.R. Das, S. Foresi and R. K. Sundaram (1996), "A Simple Approach to Three Factor Affine Term Structure Models", *Journal of Fixed Income*, Vol. 6, pp. 43-53.

Ball, C.A. and A. Roma (1994), "Stochastic Volatility Option Pricing", *Journal of Financial and Quantitative Analysis*, Vol. 29, pp. 589-607.

Bates, D. S. (1991), "The Crash of '87: Was it Expected? The Evidence from Options Markets", *Journal of Finance*, Vol. 46, pp. 1009-44.

Black, F. (1976), "Studies of Stock Price Volatility Changes", *Proceedings of the 1976 Meeting of the American Statistical Association*, pp. 177-81.

_______ and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, Vol. 81, pp. 637-59.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 31, pp. 307-27.

_________ and P.E. Rossi (1996), "Introduction", in Rossi, P.E. (eds.), *Modeling Stock Market Volatility — Bridging the Gap to Continuous Time*, San Diego, Academic Press, pp. xi-xviii.

_________, R.F. Engle and D.B. Nelson (1994), "ARCH Models", in Engle, R.F. and D. McFadden (eds.), *Handbook of Econometrics*, Vol. 4, Amsterdam, North-Holland, pp. 2959-3038.

_________ and J. M. Wooldridge (1992), "Quasi-maximum Likelihood Estimation of Dynamic Models with Time Varying Covariances", *Econometric Reviews*, Vol. 11, pp. 143-72.

Breeden, D. and R. Litzenberger (1978), "Price of State Contingent Claims Implicit in Option Prices", *Journal of Business*, Vol. 51, pp. 621-51.

Brenner, R.J., R.H. Harjes and K.F. Kroner (1996), "Another Look at Alternative Models of the Short-Term Interest Rate", *Journal of Financial and Quantitative Analysis*, Vol. 31, pp. 85-107.

Broze, L., O. Scaillet and J. Zakoïan (1995), "Testing for Continuous Time Models of the Short Term Interest Rate", *Journal of Empirical Finance*, Vol. 2, pp. 199-223.

___________________________ (1998), "Quasi-Indirect Inference for Diffusion Processes", *Econometric Theory*, Vol. 14, pp. 161-86.

Campbell, J. Y., A.W. Lo and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton, Princeton University Press.

Chapman, D.A., J.B. Long and N.D. Pearson (1999), "Using Proxies for the Short Rate: When Are Three Months Like an Instant?", *Review of Financial Studies,* Vol. 12, pp. 763-806.

Chen, L. (1996), *Interest Rate Dynamics, Derivative Pricing, and Risk Management*, Lecture Notes in Economics and Mathematical Systems, No. 435, Springer, Berlin.

Chernov, M. and E. Ghysels (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*, Vol. 56, pp. 407-58.

Conley, T.G., L.P. Hansen, E.G.J. Luttmer and J.A. Scheinkman (1997), "Short-Term Interest Rates as Subordinated Diffusions", *Review of Financial Studies*, Vol. 10, pp. 525-77.

Corradi, V. (2000), "Reconsidering the Diffusion Limit of the GARCH(1,1) Process", *Journal of Econometrics*, Vol. 96, pp. 145-53.

Cox, J.C., J.E. Ingersoll and S.A. Ross (1985), "A Theory of the Term Structure of Interest Rates", *Econometrica*, Vol. 53, pp. 385-407.

Dai, Q. and K.J. Singleton (2000), "Specification Analysis of Affine Term Structure Models", *Journal of Finance*, Vol. 55, pp. 1943-78.

Ding, Z., C. Granger and R. Engle (1993), "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance*, Vol. 1, pp. 83-106.

Drost, F.C. and T.E. Nijman (1993), "Temporal Aggregation of GARCH Processes", *Econometrica*, Vol. 53, pp. 385-407.

________ and B. J. M. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling", *Journal of Econometrics*, Vol. 74, pp. 31-57.

Duan, J.C. (1997), "Augmented GARCH(p, q) Process and its Diffusion Limit", *Journal of Econometrics*, Vol. 79, pp. 97-127.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, Vol. 50, pp. 987-1008.

________ and G.G.J. Lee (1996), "Estimating Diffusion Models of Stochastic Volatility", in Rossi, P.E. (eds.), *Modeling Stock Market Volatility—Bridging the Gap to Continuous Time*, Academic Press, San Diego, pp. 333-55.

Föllmer, H. and M. Schweizer (1991), "Hedging of Contingent Claims under Incomplete Information", in Davis, M. and R. Elliott (eds.), *Applied Stochastic Analysis*, Gordon & Breach, New York.

Fornari, F. and A. Mele (1997a), "Weak Convergence and Distributional Assumptions for a General Class of Nonlinear Arch Models", *Econometric Reviews*, Vol. 16, pp. 205-29.

____________________ (1997b), "Sign- and Volatility- Switching ARCH Models: Theory and Applications to International Stock Markets", *Journal of Applied Econometrics*, Vol. 12, pp. 49-65.

____________________ (2000a), Stochastic Volatility in Financial Markets—Crossing the Bridge to Continuous Time, Kluwer Academic Publishers, Boston.

____________________ (2000b), "An Equilibrium Model of the Term Structure with Stochastic Volatility", THEMA, Université de. Cergy and Université de Paris X, Working Paper, No. 2000-13.

____________________ (2001), "Recovering the Probability Density Function of Asset Prices using GARCH Models as Diffusion Approximations", Banca d'Italia, Temi di Discussione, No. 396, forthcoming in Journal of Empirical Finance.

Gallant, A. R. and J.R. Long (1997), "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Squared", *Biometrika*, Vol. 84, pp. 125-41.

__________ and G. Tauchen (1996), "Which Moments to Match?", *Econometric Theory*, Vol. 12, pp. 657-81.

_______________________ (1997), "Estimation of Continuous Time Models for Stock Returns and Interest Rates", unpublished.

______________________(1998),"Reprojecting Partially Observed Systems with Applications to Interest Rate Diffusions", *Journal of American Statistical Association*, Vol. 93, pp.10-24.

Gallant, A.R., D. Hsieh and G. Tauchen (1997), "Estimation of Stochastic Volatility Models with Diagnostics", *Journal of Econometrics*, Vol. 81, pp. 159-92.

Ghysels, E., A. Harvey and E. Renault (1996), "Stochastic Volatility", in Maddala, G.S. and C.R. Rao (eds.), *Handbook of Statistics*, Vol.14, Statistical Methods in Finance, Amsterdam, North-Holland, pp. 119-91.

Gouriéroux, C. and A. Monfort (1996), *Simulation-Based Econometric Methods*, Oxford, Oxford University Press.

Gouriéroux, C., A. Monfort and E. Renault (1993), "Indirect Inference", *Journal of Applied Econometrics*, Vol. 8, pp. S85-S118.

Hansen, L.P. and J.A. Scheinkman (1995), "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes", *Econometrica*, Vol. 63, pp. 767-804.

Harvey, A., E. Ruiz and N. Shephard (1994), "Multivariate Stochastic Variance Models", *Review of Economic Studies*, Vol. 61, pp. 247-64.

Heston, S. (1993), "A Closed Form Solution for Options with Stochastic Volatility with Application to Bond and Currency Options", *Review of Financial Studies*, Vol. 6, pp. 327-44.

Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatility", *Journal of Finance*, Vol. 42, pp. 281-300.

Jacquier, E., N.G. Polson and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models (with discussion)", *Journal of Business and Economic Statistics*, Vol. 12, pp. 371-417.

______________________________ (1999), "Stochastic Volatility: Univariate and Multivariate Extensions", unpublished.

Johnson, H. and D. Shanno (1987), "Option Pricing when the Variance is Changing", *Journal of Financial and Quantitative Analysis*, Vol. 22, pp. 143-51.

Karatzas, I. and S. Shreve (1991), *Brownian Motion and Stochastic Calculus*, Berlin, Springer Verlag.

Litterman, R. and J. Scheinkman (1991), "Common Factors Affecting Bond Returns", *Journal of Fixed Income*, Vol. 1, pp. 54-61.

Lo, A.W. (1988), "Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data", *Econometric Theory*, Vol. 4, pp. 231-47.

Longstaff, F. and E. Schwartz (1992), "Interest-Rate Volatility and the Term Structure: A two Factor General Equilibrium Model", *Journal of Finance*, Vol. 47, pp.1259-82.

Lucas, R.E. (1978), "Asset Prices in an Exchange Economy", *Econometrica*,Vol. 46, pp. 1429-1445.

Mele, A. (2000), "Fundamental Properties of Bond Prices in Models of the Short-Term Rate", Working Paper, THEMA (Univ. Cergy and Univ. Paris X), No. 2000-39.

Merton, R. (1973), "Theory of Rational Option Pricing", *Bell Journal of Economics and Management Science*, Vol. 4, pp. 637-54.

Nelson, D.B. (1990), "ARCH Models as Diffusion Approximation", *Journal of Econometrics*, Vol. 45, pp. 7-38.

________ (1991), "Conditional Heteroskedasticity in Asset Returns: a New Approach", *Econometrica*, Vol. 59, pp. 347-70.

________ (1992), "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model", *Journal of Econometrics*, Vol. 52, pp. 61-90.

________ (1996), "Asymptotic Filtering Theory for Multivariate ARCH Models", *Journal of Econometrics*, Vol. 71, pp. 1-47.

________ and D.P. Foster (1994), "Asymptotic Filtering Theory for Univariate ARCH Models", *Econometrica*, Vol. 62, pp. 1-41.

Pedersen, A.R. (1995), "A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations", *Scandinavian Journal of Statistics*, Vol. 22, pp. 55-71.

Romano, M. and N. Touzi (1997), "Contingent Claims and Market Completeness in a Stochastic Volatility Model", *Mathematical Finance,* Vol. 7, pp. 399-412.

Rosenberg, J. and R.F. Engle (1998), "Empirical Pricing Kernels", New York University, Salomon Center, Working Paper, No. S-98-15.

Rossi, P.E. (1996), *Modeling Stock Market Volatility—Bridging the Gap to Continuous Time*, San Diego, Academic Press.

Schwartz, B.A., D.B. Nelson and D.P. Foster (1993), "Variance Filtering with ARCH Models: a Monte-Carlo Investigation", Tel Aviv University, Tel Aviv, mimeo.

Schwert, W. (1989), "Why Does Stock Market Volatility Change Over Time?", *Journal of Finance*, Vol. 44, pp. 1115-54.

Shephard, N. (1996), "Statistical Aspects of ARCH and Stochastic Volatility", in Cox, D.R., D.V. Hinkley and O.E. Barndorff-Nielsen (eds.), *Time Series Models in Econometrics, Finance and other Fields,* London, Chapman & Hall, pp. 1-67.

Stroock, D. and S. Varadhan (1979), *Multidimensional Diffusion Processes,* Berlin, Springer-Verlag.

Taylor, S. (1986), *Modeling Financial Time Series*, Chichester, UK, Wiley.

_______ (1994), "Modeling Stochastic Volatility: a Review and Comparative Study", *Mathematical Finance*, Vol. 4, pp. 183-204.

Vasicek, O. (1977), "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics,* Vol. 5, pp. 177-88.

Wiggins, J. (1987), "Option Values Under Stochastic Volatility. Theory and Empirical Estimates", *Journal of Financial Economics*, Vol. 19, pp. 351-72.

RECENTLY PUBLISHED "TEMI" (*)

No. 373 — *Tassazione e costo del lavoro nei paesi industriali, by* M. R. MARINO and R. RINALDI (June 2000).

No. 374 — *Strategic Monetary Policy with Non-Atomistic Wage-Setters, by* F. LIPPI (June 2000).

No. 375 — *Emu Fiscal Rules: is There a Gap?,* by F. BALASSONE and D. MONACELLI (June 2000).

No. 376 — *Do Better Institutions Mitigate Agency Problems? Evidence from Corporate Finance Choices,* by M. GIANNETTI (June 2000).

No. 377 — *The Italian Business Cycle: Coincident and Leading Indicators and Some Stylized Facts,* by F. ALTISSIMO, D. J. MARCHETTI and G. P. ONETO (October 2000).

No. 378 — *Stock Values and Fundamentals: Link or Irrationality?,* by F. FORNARI and M. PERICOLI (October 2000).

No. 379 — *Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries,* by A. B. ATKINSON and A. BRANDOLINI (October 2000).

No. 380 — *Bank Competition and Regulatory Reform: The Case of the Italian Banking Industry,* by P. ANGELINI and N. CETORELLI (October 2000).

No. 381 — *The Determinants of Cross-Border Bank Shareholdings: an Analysis with Bank-Level Data from OECD Countries,* by D. FOCARELLI and A. F. POZZOLO (October 2000).

No. 382 — *Endogenous Growth with Intertemporally Dependent Preferences,* by G. FERRAGUTO and P. PAGANO (October 2000).

No. 383 — *(Fractional) Beta Convergence,* by C. MICHELACCI and P. ZAFFARONI (October 2000).

No. 384 — *Will a Common European Monetary Policy Have Asymmetric Effects?,* by L. GUISO, A. K. KASHYAP, F. PANETTA and D. TERLIZZESE (October 2000).

No. 385 — *Testing for Stochastic Trends in Series with Structural Breaks,* by F. BUSETTI (October 2000).

No. 386 — *Revisiting the Case for a Populist Central Banke*r, by F. LIPPI (October 2000).

No. 387 — *The multimarket contacts theory: an application to italian banks,* by R. DE BONIS and A. FERRANDO (December 2000).

No. 388 — *La "credit view" in economia aperta: un'applicazione al caso italiano,* by P. CHIADES and L. GAMBACORTA (December 2000).

No. 389 — *The monetary trasmission mechanism: evidence from the industries of five OECD countries,* by L. DEDOLA and F. LIPPI (December 2000).

No. 390 — *Disuguaglianza dei redditi individuali e ruolo della famiglia in Italia,* by G. D'ALESSIO and L. F. SIGNORINI (December 2000).

No. 391 — *Expectations and information in second generation currency crises models,* by M. SBRACIA and A. ZAGHINI (December 2000).

No. 392 — *Unobserved Factor Utilization, Technology Shocks and Business Cycles,* by D. J. MARCHETTI and F. NUCCI (February 2001).

No. 393 — *The Stability of the Relation between the Stock Market and Macroeconomic Forces,* by F. PANETTA (February 2001).

No. 394 — *Firm Size Distribution and Growth,* by P. PAGANO and F. SCHIVARDI (February 2001).

No. 395 — *Macroeconomic Forecasting: Debunking a Few Old Wives' Tales,* by S. SIVIERO and D. TERLIZZESE (February 2001).

No. 396 — *Recovering the Probability Density Function of Asset Prices Using GARCH as Diffusion Approximations,* by F. FORNARI and A. MELE (February 2001).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi – Divisione Biblioteca e pubblicazioni – Via Nazionale, 91 – 00184 Rome (fax 0039 06 47922059). They are available on the Internet at www.bancaditalia.it