Temporal Aggregation of Volatility Models*

Nour Meddahi † ‡ and Eric Renault §

PLEASE DO NOT CIRCULATE

Abstract

In this paper, we consider temporal aggregation of volatility models. We introduce a semi-parametric class of volatility models termed square-root stochastic autoregressive volatility (SR-SARV) and characterized by an autoregressive dynamic of the stochastic variance. Our class encompass the usual GARCH models of Bollerslev (1986), the asymmetric GARCH models of Glosten, Jagannathan and Runkle (1989) and Engle and Ng (1993). Moreover, when the volatility is stochastic, that is there is a second source of randomness, the considered models are characterized by observable multi-period conditional moment restrictions (Hansen, 1985). The SR-SARV class is a natural extension of the weak GARCH models of Drost and Nijman (1993). Our extension has two advantages: i) We allow for asymmetries (skewness, leverage effect) that are excluded by the weak GARCH models; ii) we derive observable conditional moment restrictions which are useful for (non linear) inference.

Keywords: GARCH, Stochastic Volatility, State-Space, SR-SARV, Aggregation, Asset Returns, Diffusion Processes.

---

*This is a revision of a part of Meddahi and Renault (1996), "Aggregations and Marginalization of GARCH and Stochastic Volatility Models." Some other results of this manuscript are now included in a companion paper Meddahi and Renault (1999) entitled "Conditioning Information in Volatility Models."

†Corresponding author.

‡Département de sciences économiques, Université de Montréal, CRDE, CIRANO and CEPR. Address: C.P. 6128, succursale Centre-ville, Montréal (Québec), H3C 3J7, Canada. E-mail: nour.meddahi@umontreal.ca

§CREST-ENSAI, Rennes, France. E-mail: erenault@ensai.fr

1 Introduction

In this paper, we propose a class of Conditional Heteroskedastic (CH) models which is robust to temporal aggregation. There are two main characteristics of the class of models we consider: the variance is **linear** and **stochastic**. The observable restrictions implied by our model are that the squared residual process is an ARMA process. This ARMA structure is the main characteristic of GARCH models (Engle 1982, Bollerslev 1986, Drost and Nijman 1993).[1] For instance, when a process ε is GARCH(1,1), ε^2 is ARMA(1,1). This is probably the main reason of the widespread use of such models in Financial Econometrics. In particular, the well known volatility clustering effect in financial data that GARCH models account for amounts to the ARMA structure of the squared residuals. Moreover, an ARMA representation for squared innovations is a general parameterization of their temporal autocorrelation, since by the Wold decomposition theorem, any regular second order stationary process admits a MA(∞) representation (see, e.g., Brockwell and Davis, 1990).

On the other hand, modern financial economics is defined in continuous time while financial data (interest rates, stock price returns, indexes) are typically observed at ad hoc frequencies (hourly, daily, weekly...). Thus, there are systematic sampling and temporal aggregation effects (see Weiss, 1984). Hence, when one specifies a model, on does it either for the observable frequency and has to be sure that it is correct for this given frequency (which is testable), or specifies a model for a high frequency, e.g. continuous time, then derives the observable restrictions for a given frequency. Typically, models as in Engle (1982), Bollerslev (1986), Nelson (1991), among others, are from the first class, while Drost and Nijman (1993), Drost and Werker (1996) and Hansen and Scheinkman (1995) stem from the second approach.[2]

Drost and Nijman (1993) consider the temporal aggregation of CH models. They introduce the weak GARCH models which shape is maintained whatever the sampling frequency.[3] Their model is characterized by the weak ARMA structure of the squared innovation process, i.e. ARMA whose innovation process is serially uncorrelated (weak white noise). They do this since temporal aggregation literature of ARIMA models (see Granger, 1990, for a review) learns us that after temporal aggregation, a weak ARIMA is still a weak ARIMA for both flow and stock variables. This result does not hold in general for semi-strong ARMA, i.e. ARMA whose innovation process is a martingale difference sequence (m.d.s.).[4] This is why Drost and Nijman (1993) show that the semi-strong GARCH class is not closed under temporal aggregation.[5]

Since weak GARCH models are characterized by the weak ARMA structure of the squared innovation process, only linear projections are considered, that is conditional expectations are

[1] For a review of GARCH models, see, Bollerslev, Engle and Nelson (1994), and Diebold and Lopez (1995).

[2] Nelson bridges the gap between discrete time ARCH models and continuous time models by taking an approximating, filtering or smoothing approach: Nelson (1990, 1992, 1996), Nelson and Foster (1994).

[3] Hansen and Scheinkman (1995) consider continuous time stochastic differential equations and derive moment restrictions for a given frequency of data.

[4] However, homoskedastic gaussian ARMA class of processes is closed under temporal aggregation.

[5] Two additional reasons are given later: perfect linear correlation between squared innovations and conditional variance, and affine relationship between conditional variance and the squared conditional variance. However, semi-strong ARCH(1) class is closed under temporal aggregation.

not characterized. It is an important drawback for both financial interpretations (e.g., the conditional variance which is the relevant measure of risk is not characterized[6]) and statistical ones (e.g., the QMLE setting is violated). Furthermore, for flow variables (returns and interest rates), Drost and Nijman (1993) need for temporal aggregation a symmetry assumption[7] which excludes skewed innovations and the so-called leverage effect (Black, 1976, Nelson, 1991).

In this paper, we follow the main idea of Drost and Nijman (1993) by considering ARMA structure of the squared innovations. However, our approach is based on linear state-space modeling, that is, according to financial econometrics terminology, stochastic volatility (SV) modeling.[8] We consider the Square-Root Stochastic AutoRegressive Volatility (SR-SARV) models characterized by AR dynamics of the conditional variance process. This allows us to relax the symmetry assumption and to keep conditional restrictions for financial and statistical interpretations. Besides, we prove that any symmetric SR-SARV model (excluding leverage effect) is in the weak GARCH class. Hence, weak GARCH are stochastic volatility processes rather than standard GARCH and our results generalize those of Drost and Nijman (1993).

Several models in the CH literature share the property of autoregression of the variance: GARCH models, structural GARCH models of Harvey, Ruiz and Sentana (1992), and the SR-SARV models of Andersen (1994).[9] Our class of models is closely related to the Andersen (1994) SR-SARV, and we adopt his terminology. However, while Andersen (1994) adopts a complete parametric setting, here we take a semiparametric one since we do not specify the probability distributions. We do this because distributional assumptions are not robust to temporal aggregation.[10] In the stochastic volatility and factor GARCH literature, it is usual and indeed necessary to specify the complete probability distribution which is required, e.g. for inference or forecasting, in the presence of nonlinear transformation of latent variables (see, e.g., Gourieroux and Jasiak, 1999). However, we consider here **linear models** and so we do not require distributional assumptions. In particular, we derive, for inference purposes, **observable** multi-period conditional moment restrictions (Hansen, 1985).

The time series literature, see e.g. Aoki (1990), shows that there is an equivalence between ARMA and state-space models. In particular, given an ARMA model, we can find a state-space model, generally not unique, such that the observable restrictions are the same for both models. Consider the case of an ARMA(1,1), say $z_t = \varepsilon_t^2$, which corresponds to GARCH(1,1) for ε_t. One way to represent z_t in state-space form is by assuming that it is the sum of a first

[6] In addition a linear projection of a positive random variable may be negative.

[7] The intuition is as follows. For simplicity, assume that we consider the process $\{\varepsilon_t, t \in \mathbf{Z}\}$ which is weak GARCH (ε_t^2 is a weak ARMA). We observe the process $\{\varepsilon_{2t}^{(2)} = \varepsilon_{2t} + \varepsilon_{2t-1}, t \in \mathbf{Z}\}$. To see if the process $(\varepsilon_{2t}^{(2)})^2 = \varepsilon_{2t}^2 + \varepsilon_{2t-1}^2 + 2\varepsilon_{2t}\varepsilon_{2t-1}$ is a weak ARMA ($\varepsilon_{2t}^{(2)}$ weak GARCH), Drost and Nijman (1993) need that the process $\{\varepsilon_{2t}\varepsilon_{2t-1}, t \in \mathbf{Z}\}$ is not serially correlated, that is $E[\varepsilon_t \varepsilon_{t-i} \varepsilon_{t-j} \varepsilon_{t-k}] = 0$ for $\forall\ 0 \leq i \leq j \leq k, i \neq 0$ or $j \neq k$, which is a symmetry assumption.

[8] See Ghysels, Harvey and Renault (1996) and Shephard (1996) for a review.

[9] Several multivariate models in factor GARCH literature also share this property: Diebold and Nerlove (1989), Engle, Ng and Rothschild (1990), King, Sentana and Wadhwani (1994).

[10] This is conformable to Drost and Nijman (1993) who show that the **homo**kurtosis assumption is not robust to temporal aggregation. This is why strong ARCH(1) are not closed under temporal aggregation.

order autoregressive process g_t and a white noise v_t, that is

$$z_t = g_{t-1} + v_t, \quad g_t = \omega + \gamma g_{t-1} + \eta_t, \quad \text{with } |\gamma| < 1.$$

While the class of semi-strong ARMA is not robust to temporal aggregation, the class of semi-strong AR(1) is. Hence, the previous **particular state-space representation** of the ARMA(1,1) is robust to temporal aggregation. This representation implies[11]

$$E[z_t - \omega - \gamma z_{t-1} \mid z_\tau, \tau \leq t-2] = 0,$$

which is a multi-period conditional moment restriction. Such conditional moment restriction is very useful for inference (see Hansen and Singleton, 1996). Note that it implies that z_t is a weak ARMA but not a semi-strong ARMA: it is in between. This is mainly the reason why semi-strong GARCH are not closed w.r.t. temporal aggregation, since we will see later that the process ε_t^2 is a semi-strong ARMA(1,1) if and only if ε_t is a semi-strong GARCH(1,1).

In our case the process z_t is the squared value of the residual ε_t, the former is positive. On the other hand, the process g_{t-1} is the conditional mean of ε_t^2 given the large past information $J_{t-1} = \sigma(\varepsilon_\tau, g_\tau, \tau < t)$.[12] Thus, the process g_t is also positive, while it is an autoregressive process. This is not very restrictive. For instance, in a GARCH(1,1), the conditional variance, which is positive by definition, is an AR(1) process. Further, consider any AR(1) process x_t (e.g., gaussian), then the nonnegative process $z_t = x_t^2$ is also AR(1). Finally, as we will show later, the exact discretization of the so-called Constant Elasticity of Variance (CEV) are semi-strong AR(1) processes. These CEV processes, introduced by Cox (1975) are now widespread in continuous time interest rates literature due to their nonnegativity (Chan et al., 1992). As a consequence, **exact discretization** of continuous time SV models where the variance follows a CEV process (e.g., Heston, 1993) is in our class of SV discrete time models.

Starting from the SR-SARV(1) class characterized by the AR(1) dynamics of the conditional variance process, we propose several extensions. In the spirit of GARCH (p,p) modeling, we introduce the SR-SARV(p) class: the variance process is the sum of the components (marginalization) of a positive multivariate VAR(1) of size p. We adopt this particular state-space representation since the class of semi-strong VAR(1) is closed under temporal aggregation.[13] This class contains the usual GARCH(p,p) model. Besides, the squared value of a SR-SARV(p) process is an ARMA(p,p) with multi-period conditional moment restrictions of p lags. In continuous time, this leads up to consider a SV model where the variance is a marginalization of a positive vector of size p, that is a multi-factor model for the variance.[14] Exact discretization of such models is SR-SARV(p) and, hence, the process of squared residuals is an ARMA(p,p).

In a companion paper, Meddahi and Renault (1999), we show that our class of SR-SARV is robust to information reduction (filtering, heterogenous information between agents) and,

[11] We have: $E[z_t \mid z_\tau, \tau \leq t-1] = E[g_{t-1} + v_t \mid z_\tau, \tau \leq t-1] = \omega + \gamma E[g_{t-2} \mid z_\tau, \tau \leq t-1] = \omega + \gamma E[z_{t-1} - v_{t-1} \mid z_\tau, \tau \leq t-1]$, which implies the result.

[12] More precisely, the processes g_t and v_t are defined by $g_{t-1} \equiv E[\varepsilon_t^2 \mid J_{t-1}]$ and $v_t \equiv \varepsilon_t^2 - g_{t-1}$.

[13] Moreover, this allows us to maintain some markovian properties which are useful in finance.

[14] More recently, Gallant, Hsu and Tauchen (1998) adopted this approach.

following Nijman and Sentana (1994), to marginalization and contemporaneous aggregation (of portfolio, e.g.). Furthermore, we establish the relations in terms of conditioning information between the various concepts of factor GARCH models (Diebold and Nerlove, 1989; Engle, Ng and Rothschild, 1990; King, Sentana and Wadhwani, 1994). Moreover, following Engle, Lilien and Robins (1987), we consider SR-SARV-M models where conditional mean and variance may share some common factors in both discrete and continuous time. We prove that this class is closed under temporal aggregation. Note that no such result is available for weak GARCH. We also show that temporal aggregation of these models creates **automatically** a correlation between the innovation and the variance processes.[15]

The rest of the paper is organized as follows. In section 2, we summarize some classical results of ARMA theory. In particular, we stress the equivalence between ARMA and state-space representation. Furthermore, the latter is used to derive multi-period conditional moment restrictions, exact discretization of continuous time models and temporal aggregation properties. This allows us to introduce in section 3 the SR-SARV(p) model in both discrete and continuous time settings. Temporal aggregation, exact discretization and multi-period conditional moment restrictions for volatility models are then deduced from the state-space representation. In particular, we characterize the relations between SR-SARV(p), semi-strong GARCH(p,p), weak GARCH(p,p) and ARMA(p,p) representations for squared innovations. Section 4 focuses more specifically on SR-SARV(1) processes to go further on the characterization of the subclass of GARCH(1,1) and to discuss asymmetry issues (leverage effect and skewness). Section 5 concludes the paper.

2 Lessons from ARMA theory

In this section, we revisit standard ARMA theory to enhance its main lessons for volatility modeling. We particularly focus on the state-space representation of an ARMA model, the implied conditional moments restrictions (for inference purpose) and the discretization of autoregressive continuous time models. Finally, we give several examples of autoregressive processes distributed on the positive real line and therefore meaningful for volatility processes.

2.1 State-space representation

In the time series literature (see, e.g., Brockwell and Davis, 1990), two types of ARMA processes are generally studied. The first one considers independent and identically distributed (i.i.d.) innovations with a finite variance $D(0, \sigma^2)$ (strong white noise) while in the second one the innovations are only assumed to be serially uncorrelated (weak white noise). On the other hand, several economic models are defined in terms of conditional moment restrictions (first order conditions, rational expectations...). Thus econometric literature often focuses on an intermediate type of ARMA models based on conditional expectations: the innovation process

[15] This is consistent with Black (1976) and French, Schwert and Stambaugh (1987) who conclude that leverage is probably not the sole explanation for the negative correlation between stock returns and volatility.

is a martingale difference sequence (semi-strong white noise).

Definition 2.1: Strong, semi-strong and weak ARMA:
Let $\{z_t, t \in \mathbf{Z}\}$ be a second order stationary process such that $P(L)z_t = \omega + Q(L)\eta_t$, with $P(L) = 1 - \sum_{i=1}^{p} a_i L^i$, $Q(l) = 1 - \sum_{j=1}^{q} b_i L^i$, where L is the lag operator. We assume that $a_p \neq 0,\ b_q \neq 0$ and the polynomials $P(L)$ and $Q(L)$ have different roots which are outside the unit circle. We say that:
i) z_t is a strong ARMA(p,q) if the process η_t is i.i.d. $D(0, \sigma^2)$;
ii) z_t is a semi-strong ARMA(p,q) if η_t is a m.d.s. ($E[\eta_t \mid \eta_\tau, \tau \leq t-1] = 0$);
iii) z_t is a weak ARMA(p,q) if $E[\eta_t] = 0$ and $Cov[\eta_t, \eta_{t-h}] = 0$ for $h \geq 1$.

Note that the three notions are nested (i)$\Rightarrow$ ii) $\Rightarrow$ iii)) and are , under normality, equivalent. Another approach to describe time series is based on state-space modeling:

Definition 2.2 State-space representation:
We say that a second-order stationary process $\{z_t, t \in \mathbf{Z}\}$ admits the state-space representation of order p $\{G_t, \eta_t, t \in \mathbf{Z}\}$ if G_t and η_t are second order stationary processes such as:

$$z_t = g_{t-1} + \eta_t, \text{ with} \tag{2.1}$$

$$g_t = e' G_t, \tag{2.2}$$

$$G_t = - \ + \Gamma G_{t-1} + V_t, \tag{2.3}$$

$$E[\eta_t \mid z_\tau, G_\tau, \tau \leq t-1] = 0, \tag{2.4}$$

$$E[V_t \mid z_\tau, G_\tau, \tau \leq t-1] = 0, \tag{2.5}$$

where $e \in I\!R^p$ and the eigenvalues of Γ are assumed to be smaller than one in modulus.

In such models, the dynamic of the process z_t is defined through the process g_t which is a marginalization of the VAR(1) process G_t of size p. Since g_t is a marginalization of a VAR(1), it is a weak ARMA(p,p-1) (see, e.g., Lutkepohl (1991)).[16] The process G_t is possibly unobservable by the economic agent or by the econometrician. For instance, g_t can represents the process anticipated by the economic agent of the variable z_t.[17] In the latter case, g is observable by the economic agent and not by the econometrician. As we already mentioned, ARMA models can be represented by state-space models. More precisely:

Proposition 2.1 State-space representation of a semi-strong ARMA

[16]Note however that the above definition in terms of a VAR(1) process G_t of state variables is not tantamount to a definition directly in terms of a state process g_t ARMA(p,p-1); the important difference relies on the conditioning information set.

[17]For instance, Campbell (1990) stressed that the dynamics of rational expectations can be characterized without presumption that the relevant information set contains only the history of past asset returns and a large number of variables may be useful in forecasting.

Let $\{z_t, t \in \mathbf{Z}\}$ be a semi-strong ARMA(p,p) with a corresponding representation $P(L)z_t = \omega + Q(L)\eta_t$ with $P(L) = 1 - \sum_{i=1}^{p} a_i L^i$ and $Q(L) = 1 - \sum_{i=1}^{p} b_i L^i$. Define the processes $\{G_t, g_t, v_t, t \in \mathbf{Z}\}$ by

$$G_{t-1} \equiv (E[z_{t+p-1} \mid I_{t-1}], E[z_{t+p-2} \mid I_{t-1}], .., E[z_t \mid I_{t-1}])', \tag{2.6}$$

$$g_{t-1} \equiv (0, 0.., 0, 1)G_{t-1}, \tag{2.7}$$

and $v_t \equiv z_t - g_{t-1}$ where $I_t = \sigma(z_\tau, \tau \leq t)$. Then z_t admits the state-space representation $\{G_t, v_t, t \in \mathbf{Z}\}$. Moreover, $G_t = - + \Gamma G_{t-1} + V_t$ where $- = (\omega, 0, 0 \cdots 0)'$ and

$$\Gamma = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ . & \cdots & \cdots & . \\ . & \cdots & \cdots & . \\ 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{2.8}$$

Note that this result concerns semi-strong ARMA and therefore strong ARMA. However, it is not true for weak ARMA because the weak noise property is too poor to provide conditional moment restrictions like (2.4) and (2.5). These conditional moment restrictions are the only binding restrictions w.r.t. the Wold representation setting. Of course, any ARMA(p,q) can be written as an ARMA(p',p') with p'=max(p,q). But the above property shows that the state-space representation requires something intermediate between weak and semi-strong ARMA(p,p) properties, characterized by the following multi-period conditional moment restrictions of order p on observable:

Proposition 2.2 State-space representation and multi-period conditional moments *A stationary process $\{z_t, t \in \mathbf{Z}\}$ admits a state-space representation of order p iff there exist $(p+1)$ reals ω, a_1,..,a_p, such as the roots of $1 - \sum_{i=1}^{p} a_i L^i$ are outside the unit circle and*

$$E[z_t - \omega - \sum_{i=1}^{p} a_i z_{t-i} \mid z_\tau, \tau \leq t - p - 1] = 0. \tag{2.9}$$

To summarize, the state-space representation of order p characterizes a class of processes which contains strictly the class of semi-strong and strong ARMA(p,p) and is strictly included in the weak ARMA(p,p) one. Multi-period conditional moment restrictions like (2.9) occur in several economic contexts: models of multi-period returns of securities (Hansen and Hodrick, 1980), in the presence of time-averaged data (Grossman, Melino and Shiller, 1987). Optimal instruments issue for (2.9) is addressed in Hayashi and Sims (1983) and in several papers by Hansen and co-authors (Hansen, 1985, Hansen, Heaton and Ogaki, 1988, Hansen and Singleton, 1996).

On the other hand, temporal aggregation properties of ARMA models are obtained **only** for weak ARMA,[18] see e.g., Palm and Nijman (1984) and see Granger (1990) for a survey. However, the class of semi-strong VAR(1) is closed under temporal aggregation. Thus, the state-space representation of order p, where G_t is an VAR(1), is closed under temporal aggregation:

[18] Of course, under normality, temporal aggregation holds for semi-strong and strong ARMA since they are equivalent to weak ARMA.

Proposition 2.3 Temporal aggregation of the state-space representation
Let $\{z_t, t \in \mathbf{Z}\}$ a stationary process which admits a state-space representation of order p $\{G_t, \eta_t, t \in \mathbf{Z}\}$. Define for a given integer m and real numbers $(a_0, a_1, .., a_{m-1})$ the process $\{z_{tm}^{(m)}, t \in \mathbf{Z}\}$ by $z_{tm}^{(m)} \equiv \sum_{i=0}^{m-1} a_i z_{tm-i}$. Then $\{z_{tm}^{(m)}, t \in \mathbf{Z}\}$ admits a state-space representation of order p. More precisely, we have $z_{tm}^{(m)} = g_{tm-m}^{(m)} + \eta_{tm}^{(m)}$ where

$$g_{tm-m}^{(m)} \equiv E[z_{tm}^{(m)} \mid z_{tm-m-\tau}, G_{tm-m-\tau}, \tau \geq 0] = e'(A^{(m)} G_{tm-m} + B^{(m)}), \tag{2.10}$$

$$\text{with } A^{(m)} = \sum_{i=0}^{m-1} a_i \Gamma^{m-i-1} \text{ and } B^{(m)} = (\sum_{i=0}^{m-1} a_i (\sum_{k=0}^{m-i-2} \Gamma^k))\text{-} \ . \tag{2.11}$$

Assume that $e' A^{(m)} \neq 0$, i.e. $z_{tm}^{(m)}$ is not a constant almost surely, then $g_{tm}^{(m)} = e^{(m)\prime} G_{tm}^{(m)}$ with

$$e^{(m)} = A^{(m)\prime} e, \quad G_{tm}^{(m)} = G_{tm} + e^{(m)} (e^{(m)\prime} e^{(m)})^{-1} e' B^{(m)}. \tag{2.12}$$

Besides, $G_{tm}^{(m)}$ is a VAR(1) process whose autoregressive matrix $\Gamma^{(m)}$ is given by

$$\Gamma^{(m)} = \Gamma^m. \tag{2.13}$$

This proposition means that while semi-strong ARMA class is not closed under temporal aggregation, the particular state-space representation that we consider is. Thus, the class of ARMA processes defined by the multi-period conditional moment restrictions (2.9) is closed under temporal aggregation. Note that this class is endowed with richer properties than weak ARMA which makes it more interesting for both financial and statistical purposes (see below). Let stress at this stage some interpretations of the above results which will make even more sense for temporal aggregation of volatility models hereinafter. The resulting variable $g_{tm}^{(m)}$ at the low frequency is the conditional mean of the aggregated process $\varepsilon_{(t+1)m}^{(m)}$ given the information at the **high frequency** $\sigma(z_{tm-\tau}, G_{tm-\tau}, \tau \geq 0)$ (see the first part of (2.10)). Thus, it is an affine function of the initial state variable G_{tm} (second part of (2.10)).[19] Then, assuming that $e' A^{(m)} \neq 0$,[20] we can rewrite $g_{tm}^{(m)}$ as a marginalization of a new state variable $G_{tm}^{(m)}$ which is indeed the original one plus a constant (see (2.12)).[21] Therefore, it is also a VAR(1) and the low frequency autoregressive coefficient is equal to the high frequency coefficient to the power m (see (2.13)). Thus persistence increases exponentially with the frequency.

2.2 From continuous time to discrete time

Several models in financial economics are defined in continuous time. However, the data are typically available in discrete time. Therefore, for inference purposes, it is necessarily to derive

[19] This is due to the markovianity in mean and to the autoregressive form of g_{tm}.

[20] This assumption is not restrictive. Indeed, $e' A^m = 0$ means that the process $z_{tm}^{(m)}$ is a constant ($e' B^{(m)}$). Therefore, it is a degenerate state-space model.

[21] As usual, the state space representation is not unique. For instance, we can consider $\tilde{G}_{tm}^{(m)} = A^{(m)} G_{tm} + B^{(m)}$ as a state variable. In this case, if one assumes that the matrix $A^{(m)}$ is non singular, then $\tilde{G}_{tm}^{(m)}$ is also a VAR(1) whose autoregressive matrix is also $\Gamma^{(m)}$. In other words, we prefer the state $G_{tm}^{(m)}$ rather than $\tilde{G}_{tm}^{(m)}$ because we have to assume only that $e' A^{(m)} \neq 0$ rather then $A^{(m)}$ non singular.

the implied restrictions fulfilled by the data. A natural approach is to derive the exact likelihood of the data from a fully parametric continuous time model. However this likelihood does not admit in general a closed form expression. Therefore alternative approaches are developed like non parametric methods (Ait-Sahalia, 1996), simulated methods (see Gourieroux and Monfort, 1996), bayesian methods (Elarian, Chib and Shephard, 1998) or GMM method (Hansen and Sheinkman, 1995).
However, the template of continuous time model which allows us to derive the exact likelihood for the discrete time data is the Ornstein-Uhlenbeck process:

$$dY_t = K(\Theta - Y_t)dt + \Sigma dW_t \tag{2.14}$$

where $Y_t \in I\!R^p$, $\Theta \in I\!R^p$, K is a matrix of size $(p \times p)$ and dW_t is p-variate standard Wiener process. In this case, for any $h > 0$, the process $\{Y_{\tau h}, \tau \in \mathbf{Z}\}$ is a conditionally gaussian VAR(1) process whose conditional mean is $(Id - e^{-Kh})\Theta + e^{-Kh}Y_{(\tau-1)h}$.
Note however that the VAR structure of the conditional mean is indeed only due to the linearity structure of the drift. Therefore, given a process $\{Y_t, t \in I\!R^p\}$ defined by

$$dY_t = K(\Theta - Y_t)dt + \Sigma_t dW_t \tag{2.15}$$

where the matrix Σ_t can depend on Y_t or on additional variables F_t,[22] the process $\{Y_{\tau h}, \tau \in \mathbf{Z}\}$ is a semi-strong VAR(1) process (see Appendix A), that is:

$$E[Y_{th} \mid Y_{\tau h}, \tau \leq t-1] = (Id - e^{-Kh})\Theta + e^{-Kh}Y_{(t-1)h}. \tag{2.16}$$

Such processes will be of interest in our paper. Furthermore, we also require their positivity.

2.3 Autoregression and positivity

In this section, we propose three examples of autoregressive **and** nonnegative processes.
• **Example 1:** Let us consider the process $\{z_t, t \in I\!R\}$ which is the stationary solution of:

$$dz_t = k(\theta - z_t)dt + \delta(z_t)^\lambda dW_t, \tag{2.17}$$

where $1/2 \leq \lambda \leq 1$ ensures that there exists a nonnegative stationary process solution of (2.17).[23] This is the class of CEV processes. When $\lambda = 1$, we say that it is the square-root process. Note that (2.17) is the univariate version of (2.15). Thus, from (2.16) we deduce that the **nonnegative** process $\{z_{\tau h}, \tau \in \mathbf{Z}\}$ is a semi-strong **AR**(1) process.
• **Example 2:** Let $\{x_t, t \in \mathbf{Z}\}$ the process defined by $x_t = ax_{t-1} + u_t$ where $\mid a \mid < 1$ and u_t i.i.d. $\mathcal{N}(0, \sigma^2)$. Define $\{z_t, t \in \mathbf{Z}\}$ by $z_t \equiv x_t^2$, then it is straightforward to prove that

$$E[z_t \mid z_\tau, \tau < t] = a^2 z_{t-1} + \sigma^2 \quad \text{and} \quad Var[z_t \mid z_\tau, \tau < t] = 2\sigma^2(\sigma^2 + 2a^2 z_{t-1}). \tag{2.18}$$

[22]Of course, the choice of Σ_t is such that there exists a unique stationary solution of (2.15).
[23]Note however that the existence of a stationary solution can be guaranteed without the restriction $\lambda \leq 1$; see Conley et al. (1995).

The first part of (2.18) means that the nonnegative process z_t is an AR(1) while the second one means that it is conditionally heteroskedastic. Note however that the conditional heteroskedasticity is not necessary to ensure the nonnegativity of an AR process (see the following example). We can also adapt this example in continuous time. More precisely, consider $\{x_t, t \in I\!R\}$ the stationary solution of the stochastic differential equation $dx_t = -kx_t dt + \sigma dW_t$, with $k > 0$, and define the process $\{z_t, t \in I\!R\}$ by $z_t \equiv x_t^2$. Then by the Ito Lemma we have $dz_t = (\sigma^2 - 2kz_t)dt + 2x_t\sigma dW_t$, which can be rewritten as $dz_t = (\sigma^2 - 2kz_t)dt + 2\sqrt{z_t}\sigma d\tilde{W}_t$. In other words, z_t is a constrained square-root process.

• **Example 3:** Let $\{z_t, t \in \mathbf{Z}\}$ the process defined by $z_t = \omega + \rho z_{t-1} + v_t$ where, $0 < \omega$, $0 \leq \rho < 1$ and v_t i.i.d. $\mathcal{D}(0, \sigma^2)$. The process z_t has the following MA(∞) representation: $z_t = \sum_{i=0}^{+\infty} \rho^i (v_t + \omega)$. Thus, nonnegativity of z_t is ensured when $v_t \geq -\omega$ almost surely.

3 SR-SARV(p) model

In this section we introduce the Square Root Stochastic AutoRegressive Volatility model of order p (SR-SARV(p)) in discrete and continuous times. The main idea of these models is a state-space representation of order p for the squared (innovation) process. We prove the consistency between these two models by showing that exact discretization of continuous time SR-SARV(p) model is a discrete time SR-SARV(p) model. This suggests that this class of models is closed under temporal aggregation and we therefore prove the aggregation result. Then we derive observable restrictions of our model. It provides multi-period conditional moment restrictions of p lags which hold for the squared process. This ensures an ARMA structure for the squared innovation process which is intermediate between weak and semi-strong. Finally we recall the definitions of semi-strong GARCH and weak GARCH and their links with the ARMA structure of the squared innovations.

3.1 The model

3.1.1 Discrete time SR-SARV(p) model

Definition 3.1 Discrete time SR-SARV(p) model:
A stationary squared integrable process $\{\varepsilon_t, t \in \mathbf{Z}\}$ is called a **SR-SARV(p) process** with respect to an increasing filtration $J_t, t \in \mathbf{Z}$, if:
i) the process ε_t is adapted w.r.t. J_t, that is $I_t \subset J_t$ where $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$;
ii) ε_t is a martingale difference sequence w.r.t. J_{t-1}, that is $E[\varepsilon_t \mid J_{t-1}] = 0$;
iii) the conditional variance process f_{t-1} of ε_t given J_{t-1} is a marginalization of a stationary VAR(1) of size p:

$$f_{t-1} \equiv Var[\varepsilon_t \mid J_{t-1}] = e' F_{t-1}, \tag{3.1}$$

$$F_t = - \; + \Gamma F_{t-1} + V_t, \quad \text{with } E[V_t \mid J_{t-1}] = 0, \tag{3.2}$$

where $e \in I\!R_+^p$, $- \; \in I\!R_+^p$ and the eigenvalues of Γ have modulus smaller than one.

Observe that our definition is defined for a given information set J_t. J_t can strictly contain the information set $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$, which is the minimal information. In particular, J_t may contain macroeconomic variables, informations about others assets and markets, the volume of transactions, the spread, the order book and so on.[24] Indeed, we never assume that the econometrician observes the full information set J_t even if the economic agent does. Thus, the model is a Stochastic Volatility (SV) model since the conditional variance process is a function of possibly latent variables.

The process of interest ε_t is assumed to be a martingale difference sequence w.r.t the large information J_t and therefore w.r.t. I_t. Typically, ε_t could be the log-return of a given asset whose price at time t is denoted by S_t: $\varepsilon_t = Log(S_t/S_{t-1})$. This assumption of m.d.s. is widespread in financial economics in relation with a property of informational efficiency of asset markets. However, we do not preclude predictable log-returns; in this case, our ε_t should be interpreted as the innovation process.[25] In addition, ε_t^2 admits the state-space representation $\varepsilon_t^2 = e' F_{t-1} + (\varepsilon_t^2 - E[\varepsilon_t^2 \mid J_{t-1}])$. Therefore, ε_t^2 is endowed with the state-space model properties like, e.g., multi-period conditional moment restrictions and ARMA structure (see below).

This model is related to the Andersen (1994) SR-SARV and indeed we adopt his terminology. However, there are several differences between our model and the Andersen (1994) one.[26] More precisely, Andersen (1994) considers a fully parametric model by specifying the complete distribution of the process $(\varepsilon_t, F_t')'$. Temporal aggregation requirement prevents us from completely specifying the probability distributions. Indeed, neither distributional assumptions nor homo-conditional moments restrictions (homoskewness, homokurtosis) are closed under temporal aggregation (see below). Furthermore, Andersen (1994) excludes leverage effect while we do not. Note that we do specify neither this leverage effect $Cov[\varepsilon_t, F_t]$ (which can be interpreted as a multivariate leverage effect) nor the high order moments of ε_t (third, fourth...) and V_t.[27] To summarize, we consider here a semiparametric SV model. Of course, an apparent weakness of our model is the autoregressive structure of the variance process instead of its logarithm as in the most standard SV models. However, we already proposed such processes in the previous section. In addition, we now consider continuous time stochastic volatility models which are popular in finance due to their positivity; then, we prove that exact discretization of these processes are discrete time SR-SARV(p) ones.

[24] Note also that $\sigma(\varepsilon_\tau, f_\tau, \tau \leq t) \subset J_t$ since the process f_t is adapted w.r.t. J_t.

[25] In section 5, we address more specifically the issue of SV processes whose conditional mean includes the constant and dynamic variables.

[26] Andersen (1994) considers the general class of SARV models where a function of the conditional variance process is a polynomial of an AR(1) markov process. When this function is the square-root, Andersen (1994) calls it Square-Root (SR) SARV while he terms Exponential SARV when this function is the exponential one, corresponding to the Taylor (1986) and Harvey, Ruiz and Shephard (1994) lognormal SV model.

[27] On the other hand, Andersen (1994) considers only one factor so his model is related to a SR-SARV(1). However, he defined the volatility process as a function of a polynomial, say of degree p, of an AR(1) state-variable K_t. Thus, it is a marginalization of the vector $(K_t, K_t^2, ..., K_t^p)'$ which is indeed a VAR(1) of size p, see section 2. In other words, Andersen (1994) considers also a particular case of SARV(p) model. But he does not stress this point and he only considers in his empirical study the case $p = 1$.

3.1.2 Continuous time SR-SARV(p) model

Definition 3.2 Continuous time SR-SARV(p) model
A continuous time stationary process $\{y_t, t \in I\!R\}$ is called a **SR-SARV(p)** process with respect to an increasing filtration $J_t, t \in I\!R$, if and only if there exists a p-variate process F_t^c such that y_t is the stationary solution of

$$d\begin{pmatrix} y_t \\ F_t^c \end{pmatrix} = \begin{pmatrix} 0 \\ K(\Theta - F_t^c) \end{pmatrix} dt + R_t\, dW_t, \tag{3.3}$$

where W_t is a $(p+1)$-variate standard Wiener process adapted w.r.t J_t, K is $p \times p$ matrix whose eigenvalues have positive real parts[28] and R_t is a $(p+1) \times (p+1)$ lower triangular matrix, such that the north-west coefficient $r_{11,t}$ is the square-root of

$$r_{11,t}^2 \equiv \sigma_t^2 = e' F_t^c, \quad \text{with } e \in I\!R_+^p. \tag{3.4}$$

The instantaneous conditional variance of (y_t, F_t^c) given J_t is $R_t R_t'$. The matrix R_t is lower triangular,[29] therefore the conditional variance of y_t given J_t is $r_{11,t}^2$. In other words, we follow the main idea of the discrete time SR-SARV(p) model, that is the conditional variance is a marginalization of a p dimensional VAR(1) positive process F_t^c. Note that as for the discrete time model, we have a semiparametric SV model since we do not define completely the matrix R_t. In particular, we allow for leverage effect. Of course, the matrix R_t has to fulfill conditions ensuring existence and unicity of a stationary solution of the stochastic differential equation (3.3). For instance, this is consistent with the Duffie and Kan (1996) setting of a multivariate square-root process such that each coefficient of $R_t R_t'$ is of the form $(1, F_t^{c'})\tilde{e}$ with $\tilde{e} \in I\!R^{p+1}$.[30] For $p = 1$, we can consider a CEV process[31]

$$d\sigma_t^2 = k(\theta - \sigma_t^2)dt + \delta(\sigma_t^2)^\lambda dW_{2,t}, \text{ with } 1/2 \le \lambda \le 1. \tag{3.5}$$

Finally, note that we can also consider a model where there are additional factors in R_t.
We will now prove that the two previous definitions are indeed consistent since exact discretization of continuous time SR-SARV(p) model is a discrete time SR-SARV(p) one:

Proposition 3.1 Exact discretization of continuous time SR-SARV(p)
Let $\{y_t, t \in \Re\}$ a continuous time SR-SARV(p) process with a corresponding factor process $\{F_t^c, t \in \Re\}$. Then, for any sampling interval h, the associated discrete time process $\varepsilon_{th}^{(h)} = y_{th} - y_{(t-1)h}$, $t \in \mathbf{Z}$, is a SR-SARV(p) process w.r.t. $J_{th}^{(h)}$, $J_{th}^{(h)} = \sigma(\varepsilon_{\tau h}^{(h)}, F_{\tau h}^c, \tau \le t, \tau \in \mathbf{Z})$. The corresponding conditional variance process $f_{(t-1)h}^{(h)} \equiv Var[\varepsilon_{th}^{(h)} \mid J_{(t-1)h}^{(h)}]$ is given by $f_{th}^{(h)} = e' F_{th}^{(h)}$ with $F_{th}^{(h)} = A^{(h)} F_{th}^c + B^{(h)}$, where $A^{(h)} = K^{-1}(Id - e^{-Kh})$ and $B^{(h)} = (hId - A^{(h)})\Theta$.

[28] Indeed, a usual assumption, see e.g. Bergstrom (1990), page 53, is that the eigenvalues of K are distinct. Therefore K is diagonalisable, i.e. there exist a matrix H such that $HKH^{-1} = Diag(\lambda_1, ..., \lambda_p) \equiv \Lambda$. As a consequence, for $u > 0$, $He^{-uK}H^{-1} = e^{-u\Lambda} = Diag(e^{-u\lambda_1}, ..., e^{-u\lambda_p})$ with $e^Z = \sum_{i=0}^{\infty} \frac{Z^i}{i!}$. The positivity of the real parts of the eigenvalues K ensures the existence of e^{-uK} $\forall u > 0$.

[29] This Gramm-Schmidt normalization rule is standard and can be maintained without loss of generality.

[30] See Duffie and Kan (1996) for sufficient conditions of existence of a stationary solution of (3.3) in this case.

[31] Since there is only one factor, we change the notations by taking $F_t^c \equiv \sigma_t^2$, $W_t = (W_{1,t}, W_{2,t})'$.

This proposition[32] means that exact discretization of the factor or Stochastic volatility models of Duffie and Kan (1996), Constantinides (1992), Heston (1993), are also factor or SV models. Moreover, as we will see later, such models imply conditional moment restrictions based only on the observable.
More recently, Barndorff-Nielsen and Shephard (1998) consider a new class of continuous time stochastic volatility models (termed Background Driving Lévy Processes (BDLP)) which can be simulated without any form of discretization error. They also characterize the moments, autocorrelation functions and spectrum of squared returns. Indeed we can prove that exact discretization of BDLP processes are also SR-SARV (see Appendix B).

The previous result suggests that the SR-SARV(p) class is closed under temporal aggregation. This is the main contribution of our paper and the purpose of the next subsection.

3.2 Temporal aggregation of SR-SARV(p) models

We consider here general temporal aggregation of a given process. More precisely, let consider a process $\{\varepsilon_t, t \in \mathbf{Z}\}$, we assume that we observe the process $\{\varepsilon_{tm}^{(m)}, t \in \mathbf{Z}\}$ defined by

$$\varepsilon_{tm}^{(m)} = \sum_{k=0}^{m-1} a_k \varepsilon_{tm-k}, \tag{3.6}$$

where $m \in \mathbf{N}^*$, $a = (a_0, a_1, .., a_{m-1})' \in I\!R^m$. Typically, temporal aggregation of stock variables observed at the dates m, 2m, 3m,.., Tm, corresponds to $a = (1, 0, 0..., 0)'$ while for flow variables $a = (1, .., 1)'$. This later case is particularly suitable for log-returns and continuously compounded interest rates.

Proposition 3.2 Temporal aggregation of SR-SARV(p) model
Let ε_t a SR-SARV(p) process w.r.t. an increasing filtration J_t and a conditional variance process $f_t = e'F_t$. For a given integer m, the process $\varepsilon_{tm}^{(m)}$ defined by (3.6) is a SR-SARV(p) w.r.t. $J_{tm}^{(m)} = \sigma(\varepsilon_{\tau m}^{(m)}, F_{\tau m}, \tau \leq t)$. More precisely, we have:

$$f_{tm-m}^{(m)} \equiv Var[\varepsilon_{tm}^{(m)} \mid J_{tm-m}] = e'(A^{(m)} F_{tm-m} + B^{(m)}), \tag{3.7}$$

$$\text{where } A^{(m)} = \sum_{k=0}^{m-1} a_k^2 \Gamma^{m-k-1} \text{ and } B^{(m)} = (\sum_{k=0}^{m-1} a_k^2 (\sum_{i=0}^{m-k-2} \Gamma^i))\text{-} \ . \tag{3.8}$$

Assume that $e'A^{(m)} \neq 0$, then $f_{tm}^{(m)} = e^{(m)\prime} F_{tm}^{(m)}$ with

$$e^{(m)} = A^{(m)\prime} e, \quad F_{tm}^{(m)} = F_{tm} + e^{(m)} (e^{(m)\prime} e^{(m)})^{-1} e' B^{(m)}. \tag{3.9}$$

Besides, $F_{tm}^{(m)}$ is a VAR(1) process whose autoregressive matrix $\Gamma^{(m)}$ is given by

$$\Gamma^{(m)} = \Gamma^m. \tag{3.10}$$

[32]Note that in this proposition, the discrete time state variable is $A^{(h)} F_{th}^c + B^{(h)}$ and not, as in Proposition 2.3, F_{th}^c plus a constant. The reason is that we are sure that the matrix $A^{(h)}$ is not singular (see footnote 21).

Therefore, models where the conditional variance is a marginalization of a VAR(1) process are robust to temporal aggregation. Note that this result is not a direct application of the temporal aggregation property of the state-space representation (Proposition 2.3): Here we consider the dynamics of $(\sum_{i=0}^{m} a_i \varepsilon_{tm-i})^2$ while in the previous case we had characterized the dynamics of $\sum_{i=0}^{m} a_i z_{tm-i}$. Actually, in section 2, the process z has a state-space representation of order p while here it is ε^2 and not ε. However, the intuition of the two results are the same. Consider the initial process ε_t with the information J_t at high frequency and define the process at low frequency $\varepsilon_{tm}^{(m)}$ by (3.6). Define $f_{tm}^{(m)}$ as the conditional variance of $\varepsilon_{(t+1)m}^{(m)}$ given the information at **high frequency** J_{tm} (first part of 3.7). This information is generally not observable either by the agent or by the econometrician and thus the variance is stochastic. But by something like a markovian property,[33] the conditional variance $f_{tm}^{(m)}$ is a function of F_{tm}. Thanks to the linearity of the model, this function is indeed affine (second part of (3.7)). Define the information at low frequency by $J_{tm}^{(m)} \equiv \sigma(\varepsilon_{\tau m}^{(m)}, F_{\tau m}^{(m)}, \tau \leq t)$. Then $\varepsilon_{tm}^{(m)}$ is still a m.d.s. with respect to $J_{tm}^{(m)}$ since $E[\varepsilon_{tm+m}^{(m)} \mid J_{tm}] = 0$ and $J_{tm}^{(m)} \subset J_{tm}$. Of course, by definition, the conditional variance $f_{tm}^{(m)}$ of $\varepsilon_{tm+m}^{(m)}$ given $J_{tm}^{(m)}$ is positive. Then assuming that $e'A^{(m)} \neq 0$,[34] we can rewrite this conditional variance as a marginalization of a new state variable $F_{tm}^{(m)}$. The latter is a VAR(1) since it is the sum of a VAR(1) and a constant. Thus, $\varepsilon_{tm}^{(m)}$ is a SR-SARV(1) w.r.t. $J_{tm}^{(m)}$. Finally the autoregressive parameter of the VAR(1) F_{tm}^{m} is equal to the autoregressive parameter of the high frequency vector F_t to power m (3.10). It means that the persistence increases exponentially with the frequency. Conversely, conditional heteroskedasticity vanishes when the frequency is low. This corresponds to a well-documented empirical evidence and was already pointed out by Diebold (1988) and Drost and Nijman (1993).

Temporal aggregation of conditional heteroskedastic models was already considered for GARCH models by Drost and Nijman (1993) as well as the links between continuous time SV model and GARCH by Drost and Werker (1996). In the next subsection, we will recall these results and show the links between their weak GARCH model and our SR-SARV model.

3.3 Observable restrictions

3.3.1 Multi-period conditional moment restrictions

The SR-SARV is defined w.r.t. an increasing filtration J_t, which may not be observable by the economic agent or the econometrician. However, as in the previous section, since a SR-SARV(p) implies that ε_t^2 has a state-space representation of order p, we can derive conditional moments fulfilled by the observable process ε_t given the minimal information $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$. This means that the information generated by the returns should belong, in any case, to the

[33]If one has in mind an underlying continuous time representation like (3.3), the low frequency process $(y_{tm}^{(m)}, F_{tm}^{(m)})$ is markovian. More generally, our setting ensures that the conditional variance $f_{tm}^{(m)}$ depends on past information only through F_{tm}.

[34]As in the previous section, this assumption is not restrictive. The equality $e'A^{(m)} = 0$ means that the process $\varepsilon_{tm}^{(m)}$ is homoskedastic which is a degenerate SR-SARV model. In other words, temporal aggregation cancels the heteroskedastic effect.

econometrician information set. These restrictions are of course multi-period ones of order p:

Proposition 3.3 SR-SARV and ARMA
Let $\{\varepsilon_t, t \in \mathbf{Z}\}$ a stationnary process. It admits a SR-SARV(p) representation w.r.t. an increasing filtration J_t if and only if there exist p+1 reals ω, γ_1,..,γ_p, such that the roots of $1-\sum_{i=1}^{p}\gamma_i L^i$ are outside the unit circle and

$$E[\varepsilon_t^2 - \omega - \sum_{i=1}^{p}\gamma_i \varepsilon_{t-i}^2 \mid \varepsilon_\tau, \tau \leq t-p-1] = 0. \tag{3.11}$$

This result is an application of Proposition 2.2. Therefore, ε_t^2 is an ARMA(p,p) defined by (3.11), that is an ARMA property stronger than the weak one but weaker than the semi-strong one. The (semi-strong) ARMA structure was the main idea of the ARCH models introduced by Engle (1982) and generalized (GARCH) by Bollerslev (1986). In particular, the well known clustering effect in financial data that these models account for is directly related to the ARMA structure of the squared residuals. On the other hand, for temporal aggregation purposes, Drost and Nijman (1993) introduce the weak GARCH models where the squared residuals process is weak ARMA. Following the Drost and Nijman (1993) terminology, we precisely define below the various concepts and show how they are nested.

3.3.2 GARCH(p,q)

Definition 3.3 GARCH(p,q):
Let a stationary process $\{\varepsilon_t, t \in \mathbf{Z}\}$ and define the processes $\{h_t, u_t, t \in \mathbf{Z}\}$ by the stationary solution of

$$B(L)h_t = \omega + A(L)\varepsilon_t^2 \tag{3.12}$$

and $u_t = \varepsilon_t/\sqrt{h_t}$, with $A(L) = \sum_{i=1}^{q}\alpha_i L^i$, $B(L) = 1-\sum_{i=1}^{p}\beta_i L^i$ where the roots of $B(L)-A(L)$ and $B(L)$ are assumed to be outside the unit circle. We say that:
i) ε_t is a strong GARCH(p,q) if the process u_t is i.i.d. $D(0,1)$;
ii) ε_t is a semi-strong GARCH(p,q) if the process u_t is such that

$$E[u_t \mid \varepsilon_\tau, \tau \leq t-1] = 0 \text{ and } Var[u_t \mid \varepsilon_\tau, \tau \leq t-1] = 1; \tag{3.13}$$

iii) ε_t is a weak GARCH(p,q) if

$$EL[\varepsilon_t \mid H_{t-1}] = 0 \text{ and } EL[\varepsilon_t^2 \mid H_{t-1}] = h_t, \tag{3.14}$$

where $EL[x_t \mid H_{t-1}]$ denotes the best linear predictor of x_t on the Hilbert space spanned by $\{1, \varepsilon_\tau, \varepsilon_\tau^2, \tau \leq t-1\}$, H_{t-1}, that is

$$E[(x_t - EL[x_t \mid H_{t-1}])\varepsilon_{t-i}^r] = 0 \text{ for } i \geq 1 \text{ and } r = 0,1,2. \tag{3.15}$$

Proposition 3.4 Semi-strong GARCH and ARMA
Let $\{\varepsilon_t, t \in \mathbf{Z}\}$ a m.d.s. ($E[\varepsilon_t \mid \varepsilon_\tau, \tau \leq t-1]$). It is a semi-strong GARCH(p,q) if and only if ε_t^2 is a semi-strong ARMA(max$\{p,q\}$,p).

This equivalence result was already stated by Bollerslev (1988). Note that strong GARCH implies semi-strong ARMA and not strong one: When ε_t^2/h_t is i.i.d., the ARMA process ε_t^2 should in general be conditionally heteroskedastic. On the other hand, since ε_t^2 is a semi-strong ARMA, it implies a multi-period conditional moment restriction of order max(p,q).[35] Therefore, Proposition 3.3 implies that ε_t admits a SR-SARV(max(p,q)) representation. Furthermore, the continuous time SR-SARV(p) is like GARCH(p,p) model. To the best of our knowledge, the relationship between GARCH(p,p) modeling of higher order ($p > 1$) and continuous time stochastic volatility models was not clearly stated before in the literature whatever the approach of diffusion approximating (Nelson, 1990), filtering (Nelson and Foster, 1994) or closing the GARCH Gap (Drost and Werker, 1993). Finally, temporal aggregation of GARCH model is a SR-SARV model. In other words, to close the class of GARCH processes, we have to go into the stochastic volatility models. This is not a surprising result since we know from ARMA theory that semi-strong ARMA are not closed under temporal aggregation. Further, it is an additional reason[36] for the parsimony of stochastic volatility models.

In the next section, we give additional insights about the reasons why GARCH models are not robust to temporal aggregation. Drost and Nijman (1993) already focused on this weakness of standard GARCH models. They give examples of strong and semi-strong GARCH which are not closed under temporal aggregation. Then, they introduce the weak GARCH model where the squared residuals are weak ARMA in order to benefit from the aforementioned temporal aggregation of the weak ARMA structure.

Proposition 3.5 Weak GARCH and ARMA
If ε_t is a weak GARCH(p,q) process, then ε_t^2 is a weak stationary ARMA(max$\{p,q\}$,p) process. Conversely, if ε_t^2 is a weak stationary ARMA(q,p) process: $A(L)\varepsilon_t^2 = \omega + B(L)\eta_t$, with η_t weak white noise , ε_t is a weak GARCH(p,q) if and only if:

$$Cov(\eta_t, \varepsilon_\tau) = 0, \ \forall \tau < t. \tag{3.16}$$

In this case, $\gamma = \alpha + \beta$ and $EL[\varepsilon_t^2 \mid H_{t-1}] = EL[\varepsilon_t^2 \mid H_{t-1}^s]$ where H_{t-1}^s is the Hilbert space spanned by $\{1, \varepsilon_\tau^2, \tau \leq t-1\}$.

Thus, the weak GARCH property is slightly more restrictive than the weak ARMA assumption for the squared residuals. In particular, (3.16) is like a symmetry assumption, which is implied by the maintained condition m.d.s. for ε_t when assuming semi-strong GARCH. In fact, Drost and Nijman (1993) take a "consistent" definition in the sense that they project both the residual and its square onto the same space H_{t-1}. However, ARMA structure of the squared residuals was the main idea of the weak GARCH.[37] As we can already see, the class of weak

[35] More precisely, a semi-strong ARMA($\tilde{q}, \tilde{p}$) implies a multi-period conditional moment restrictions of order equal to max$\{\tilde{q}, \tilde{p}\}$. Thus a semi-strong ARMA(max$\{p,q\}$,p) implies a multi-period restriction of order max$\{p,q\}$.

[36] See Ghysels, Harvey and Renault (1996) for the other reasons.

[37] We thank Feike Drost for confirming this point to us in a private communication. In addition, when Nijman and Sentana (1996) and Drost and Weker (1996) prove respectively that a marginalization of a multivariate GARCH and that the discretization of (3.3) for p=1 under (3.5) are weak GARCH, they only deal with the ARMA property of squared residuals.

ARMA strictly contains the multi-period type ARMA models. Therefore, it means that weak GARCH are **Stochastic Volatility** models. In other words, Drost and Nijman (1993) go also into the stochastic volatility models to close the GARCH models.[38]

However, to show that weak GARCH are closed under temporal aggregation, Drost and Nijman (1993) make a restrictive technical assumption by assuming one of the following symmetry conditions:

$$\forall h \in \mathbf{N}^*, \forall (a_k)_{1\leq k\leq h} \in \{-1,1\}^h, (\varepsilon_{t+k})_{1\leq k\leq h} \; = \; (a_k\varepsilon_{t+k})_{1\leq k\leq h} \text{ in distribution, or} \qquad (3.17)$$

$$\forall 0 \leq i \leq j \; E[\varepsilon_t\varepsilon_{t-i}\varepsilon_{t-j}] = 0 \text{ and } \forall \; 0 \leq i \leq j \leq k, i \neq 0 \text{ or } j \neq k \; E[\varepsilon_t\varepsilon_{t-i}\varepsilon_{t-j}\varepsilon_{t-k}] = 0. \quad (3.18)$$

Such symmetry restrictions are indeed very restrictive both in theoretical and empirical points of view. They preclude two types of asymmetry, both of which appear relevant for financial data. First, even in the strong GARCH setting, the probability distribution of the standardized innovations $\varepsilon_t/\sqrt{h_t}$ may be skewed. Second, since the weak GARCH models go into the SR-SARV setting (outside the standard GARCH class), another type of asymmetry (termed leverage effect by Black, 1976, and popularized by Nelson, 1991) may matter. A clear distinction between these two types of asymmetric behavior of a general SR-SARV(1) process will be made in section 4 below. Equivalently, leverage effect can be introduced in the continuous time setting by allowing the volatility matrix R_t to be non-diagonal, unlike the case considered by Drost and Werker (1996). Finally, note that our results of temporal aggregation and exact discretization are consistent with those of Drost and Nijman (1993) and Drost and Werker (1996).[39] In particular, the restrictions on the parameters are the same ($\Gamma^{(m)} = \Gamma^m$).[40]

4 SR-SARV(1)

4.1 SR-SARV and GARCH

The GARCH(1,1) model is nowadays dominant w.r.t. any other ARCH or GARCH type model in the empirical finance literature. In this section, we discuss in more details its relationships with general SR-SARV(1). In the previous section, we proved that a semi-strong GARCH(p,q) is also a SR-SARV($\max\{p,q\}$). In the following proposition we characterize the SR-SARV(1) processes which are also semi-strong GARCH(1,1):

Proposition 4.1 Semi-strong GARCH(1,1) and SR-SARV(1)
Let $\{\varepsilon_t, t \in \mathbf{Z}\}$ a SR-SARV(1) process with a conditional variance process f_t. ε_t is a semi-strong GARCH(1,1) if and only if:
i) ε_t^2 and f_t are conditionally perfectly positively correlated given J_{t-1} (conditional linear correlation equal to 1);
ii) the ratio $Var[f_t \mid J_{t-1}]/Var[\varepsilon_t^2 \mid J_{t-1}]$ is constant and smaller or equal to γ^2.
In this case: $h_{t+1} = f_t$, $J_t = I_t$ and $\beta = \gamma - \alpha$ with $\alpha = \sqrt{Var[f_t \mid J_{t-1}]/Var[\varepsilon_t^2 \mid J_{t-1}]}$.

[38]See the following section where we establish the exact links between SR-SARV and weak GARCH.
[39]Nevertheless, Drost and Werker (1993) consider only the one factor case.
[40]For more details, see Meddahi and Renault (1996).

The first restriction is related to the common idea that ARCH models correspond to the degenerate case where there are no exogenous sources of randomness in the conditional variance dynamics. This degeneracy corresponds to GARCH only if it is a perfect linear conditional correlation. The second restriction is less known even though it was already coined by Nelson and Foster (1994). They observed that the most commonly used ARCH models assume that the variance of the variance rises linearly with the square of the variance, which is the main drawback of GARCH models in approximating SV models in continuous time. Thus, semi-strong or strong GARCH setting implies nontrivial restrictions on conditional kurtosis dynamics.

In the literature, there are additional ARCH-type models characterized by an autoregressive dynamic of the volatility. For instance, Glosten, Jagannathan and Runkle (1989, GJR) introduce a model based on a GARCH formulation but accounting for the sign of the past residuals. This symmetry is a limitation of GARCH models with respect to EGARCH ones. Such asymmetric models where studied and compared theoretically and empirically by Engle and Ng (1993). More precisely, they consider the following models:[41]

$$\mathrm{GJR}: \mathrm{h_t} = \omega + \alpha\varepsilon_{\mathrm{t-1}}^2 + \beta \mathrm{h_{t-1}} + \gamma\ \mathrm{S_{t-1}}\varepsilon_{\mathrm{t-1}}^2,\ \text{where } \mathrm{S_t} = 1 \text{ if } \varepsilon_{\mathrm{t}} < 0,\ \mathrm{S_t} = 0 \text{ otherwise;} \tag{4.1}$$

$$\text{Asymmetric GARCH}: \ h_t = \omega + \alpha(\varepsilon_{t-1} + \gamma)^2 + \beta h_{t-1}; \tag{4.2}$$

$$\text{Nonlinear Asymmetric GARCH}: \ h_t = \omega + \alpha(\varepsilon_{t-1} + \gamma\sqrt{h_{t-1}})^2 + \beta h_{t-1}; \tag{4.3}$$

$$\text{VGARCH}: \ h_t = \omega + \alpha(\varepsilon_{t-1}/\sqrt{h_{t-1}} + \gamma)^2 + \beta h_{t-1}; \tag{4.4}$$

Let us also consider a related model considered by Heston and Nandi (1999):[42]

$$\text{Heston and Nandi } h_t = \omega + \alpha(\varepsilon_{t-1}/\sqrt{h_{t-1}} - \gamma\sqrt{h_{t-1}})^2 + \beta h_{t-1}. \tag{4.5}$$

In the following proposition, we show that all the previous models are SR-SARV(1) ones:

Proposition 4.2 Asymmetric GARCH and SR-SARV(1)
Let $\{\varepsilon_t, t \in \mathbf{Z}\}$ such that $E[\varepsilon_t \mid \varepsilon_\tau, \tau \leq t-1] = 0$ and define h_t the conditional variance of ε_t, i.e. $h_t \equiv Var[\varepsilon_t \mid \varepsilon_\tau, \tau \leq t-1]$. Assume that h_t is given by (4.1), (4.2), (4.3), (4.4), or by (4.5), then ε_t is a SR-SARV(1) model.

4.2 SR-SARV and weak GARCH

In the following, we will focus on the relationships between SR-SARV and weak GARCH. As we say above, Drost and Nijman (1993) prove the temporal aggregation property of symmetric weak GARCH (assuming (3.17) or (3.18)) which excludes leverage effect and all the previous

[41] The Asymmetric GARCH was introduced by Engle (1990) while both the Nonlinear Asymmetric GARCH and VGARCH models were intoduced by Engle and Ng (1993).

[42] Heston and Nandi (1999) show that the limit diffuion of (4.5) is the stationary solution of (3.5) with $\lambda = 1/2$, i.e. the model considered by Heston (1993). Both the discrete time and the continuous time models present closed-form option pricing formulas.

asymmetric models (e.g. GJR). Now we precise two kinds of asymmetries for the SR-SARV model:

Definition 4.1 Leverage effect and skewness:
Let $\{\varepsilon_t, t \in \mathbf{Z}\}$ a SR-SARV(1) process w.r.t. an increasing filtration J_t with corresponding processes $\{f_t, u_t, \nu_t\}$. We say that:
i) ε_t does not present leverage effect w.r.t. J_t if and only if

$$E[u_t \nu_t \mid J_{t-1}] = 0; \tag{4.6}$$

ii) ε_t does not present skewness w.r.t. J_t if and only if

$$E[u_t^3 \mid J_{t-1}] = 0. \tag{4.7}$$

With the above definitions, we can now show that a SR-SARV model without leverage effect and skewness is a weak GARCH:

Proposition 4.3 Weak GARCH(1,1) and SR-SARV(1)
If ε_t is a SR-SARV(1) process without leverage effect and skewness, that is if (4.6) and (4.7) hold, then ε_t is a weak GARCH(1,1) process.

Therefore, there is no major difference between symmetric weak GARCH and symmetric SR-SARV. However, we do not prove an equivalence result and it is clear that the class of symmetric weak GARCH is larger than the symmetric SR-SARV. Indeed, one can interpret the weak GARCH as a stochastic volatility model, but a model which is so weak, that is there is not enough restrictions which can be useful for interpretations and inference. In addition, we proved in section 3 that this weakness is not needed to close the GARCH gap with continuous time as in Drost and Werker (1996). In a sense, by introducing the SR-SARV, we reached the weak GARCH models by adding useful restrictions for financial and statistical interpretations. Another advantage of SR-SARV is that they allow for asymmetries like leverage effect and skewness. Furthermore, these symmetric assumptions are closed under temporal aggregation:

Proposition 4.4 Temporal aggregation, leverage effect and skewness
Let $\{\varepsilon_t, t \in \mathbf{Z}\}$ a SR-SARV process w.r.t. an increasing filtration J_t with corresponding processes $\{f_t, u_t, \nu_t, t \in \mathbf{Z}\}$. Define $\varepsilon_{tm}^{(m)}$ by (3.6) and the corresponding SR-SARV(1) representation of Proposition 3.2 $J_{tm}^{(m)}$, $\{f_{tm}^{(m)}, u_{tm}^{(m)}, \nu_{tm}^{(m)}, t \in \mathbf{Z}\}$. Then the absence of leverage effect and skewness are robust to temporal aggregation, that is symmetric SR-SARV class is also closed under temporal aggregation:

$$E[u_t \nu_t \mid J_{t-1}] = 0 \Longrightarrow E[u_{tm}^{(m)} \nu_{tm}^{(m)} \mid J_{tm-m}^{(m)}] = 0, \tag{4.8}$$

$$E[u_t \nu_t \mid J_{t-1}] = E[u_t^3 \mid J_{t-1}] = 0 \Longrightarrow E[u_{tm}^{(m)} \nu_{tm}^{(m)} \mid J_{tm-m}^{(m)}] = E[(u_{tm}^{(m)})^3 \mid J_{tm-m}^{(m)}] = 0. \tag{4.9}$$

This proposition means that our results generalize those of Drost and Nijman (1993) and Drost and Werker (1996), since symmetric SR-SARV are weak GARCH and are closed under temporal aggregation. Besides, the relationships between parameters at various frequencies, already stressed by these authors (particularly the persistence parameter) are maintained in our SR-SARV setting. On the other hand this proposition implies that a symmetry assumption about the standardized innovation (a no-skewness effect) cannot be alleged without precluding leverage effect as well (see 4.9). Therefore, when one observes significant skewness at a low frequency, it may be due either to genuine skewness or to leverage effect at the high frequency, while the presence of leverage effect at a low frequency implies the same feature at the high frequency.

Proposition 4.5 Observable restrictions of leverage effect and skewness
Let ε_t a SR-SARV(1) w.r.t. an increasing filtration J_t.
i) ε_t is without leverage effect ((4.6) holds) if and only if

$$E[\varepsilon_t^2 \varepsilon_{t-1} \mid I_{t-2}] = 0. \tag{4.10}$$

ii) ε_t is without skewness ((4.7) if and only if holds), then

$$E[\varepsilon_t^3 \mid I_{t-1}] = 0. \tag{4.11}$$

Therefore we can derive moments restrictions based on observable data which can be used to test the absence of leverage effect or skewness. Moreover, usual GARCH allows for leverage affect as soon as there is skewness since the conditions (4.6) and (4.7) are equivalent in this case.

5 Conclusion

We have proposed in this paper a new concept of semiparametric stochastic volatility model which appears to be the good framework for structural interpretations of times series models with conditional heteroskedasticity. Actually, if one wants to consider time series of conditionally heteroskedastic asset returns, there was no framework available until now to capture in the same setting temporal aggregation or portfolios of these returns. On the one hand, it is well known that the usual GARCH setting is not robust with respect to temporal aggregation. On the other hand, the only robust setting already suggested in the literature, that is the Drost and Nijman (1993) weak GARCH one, suffers from several drawbacks:

In this paper, we consider temporal aggregation of volatility models. We introduce a semiparametric class of volatility models termed square-root stochastic autoregressive volatility (SR-SARV) and characterized by an autoregressive dynamic of the stochastic variance. Our class encompass the usual GARCH models of Bollerslev (1986), the asymmetric GARCH models of Glosten, Jagannathan and Runkle (1989) and Engle and Ng (1993). Moreover, when the volatility is stochastic, that is there is a second source of randomness, the considered models are characterized by observable multi-period conditional moment restrictions (Hansen, 1985). The SR-SARV class is a natural extension of the weak GARCH models of Drost and Nijman (1993). Our extension has advantages since we allow for asymmetries (skewness, leverage effect) that are excluded by the weak GARCH models and we derive observable conditional moment restrictions which are useful for inference.

References

Ait-Sahalia, Y. (1996), "Nonparametric Pricing of Interest Rate derivative Securities," *Econometrica*, 64, 527-560.

Andersen, T.G. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling", *Mathematical Finance*, 4, 75-102.

Aoki, M. (1990), *State Space Modeling of Time Series,* Second Edition.

Barndorff-Nielsen, O.E. and N. Shephard (1998), "Aggregation and Model Construction for Volatility Models," unpublished paper, Nuffield College, Oxford, presentend at the Second HFDF Conference, Zurich, April 1998.

Black, F. (1976), "Studies in Stock Price Volatility Changes", *Proceedings of the 1976 Business Meeting of the Business and Statistics Section, American Statistical Association*, 177-181.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. (1988), "On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Processes", *Journal of Times Series Analysis*, 9, 121-131.

Bollerslev, T., R. Engle and D.B. Nelson (1994), "ARCH Models", *Handbook of Econometrics*, Vol IV.

Brockwell, P.J. and R.A. Davis (1990), "Time Series: Theory and Methods," Second Edition, Springer Verlag.

Campbell, J.Y. (1990), "Measuring the Persistence of expected Returns," *American Economic Review*, 80, 43-47.

Chan, K.C, G.A. Karolyi, F.A. Longstaff and A.B. Sanders (1992), "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, XLVII, 1209-1227.

Conley, T., L.P. Hansen, E. Luttmer and J. Scheinkman (1997), "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies,* 10, 525-577.

Constantinides, J. (1992), "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies,* 5, 531-552.

Cox, J.C. (1975), "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions", Discussion Paper, Stanford University.

Diebold, F.X. (1988), *Empirical Modeling of Exchange Rates*, Springer Verlag.

Diebold, F.X. and J.A. Lopez (1995), "Modeling Volatility Dynamics", in Macroeconometrics, 427-466, edited by K.V. Hoover, Kluwer Academic Publishers.

Diebold, F.X. and M. Nerlove (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent ARCH Model", *Journal of Applied Econometrics*, 4, 1-21.

Drost, F.C. and TH.E. Nijman (1993), "Temporal Aggregation of GARCH processes", *Econometrica*, 61, 909-927.

Drost, F.C. and B.J.M. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling", *Journal of Econometrics*, 74, 31-58.

Duffie, D. and R. Kan (1996), "A Yield-Factor Model of Interest Rates", *Mathematical Finance*, 6, 379-406.

Elarian, O., S. Chib and N. Shephard (1998), "Likelihood Inference for Discretely Observed Non-Linear Diffusions," unpublished manuscript, Nuffield College, Oxford.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 987-1007.

Engle, R.F., D.M. Lilien and R.P. Robins (1987), "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Models", *Econometrica*, 55, 391-407.

Engle, R.F., V.M. Ng and M. Rothschild (1990), "Asset Pricing with a Factor ARCH Structure: Empirical estimates for Treasury Bills," *Journal of Econometrics*, 45, 213-237.

French, K.R., G.W. Schwert and R.F. Stambaugh (1987), "Expected Stock Returns and Volatility", *Journal of Financial Economics*, 19, 3-29.

Gallant, A.R, C.T. Hsu and G. Tauchen (1998), "Calibrating Volatility Diffusions and Extracting Integrated Volatility," unpublished paper, Duke, presentend at the Second HFDF Conference, Zurich, April 1998.

Granger, C.W.J (1990), "Aggregation of Time Series Variables: A Survey," in T. Barker and H. Pesaran (Eds), *Disaggregation in Econometric Modelling,* pp. 17-34, Ratlege, New-York.

Ghysels, E., A.C. Harvey and E. Renault (1995), "Stochastic Volatility", *in Handbook of Statistics*, vol 14, forthcoming.

Glosten, L.R., R. Jagannathan and D. Runkle (1993), "On the Relation between the Expected Value of the Volatility of the Nominal Excess Return on Stocks", *The Journal of Finance*, Vol XLVIII, 1779-1801.

Gouriéroux, C. and J. Jasiak (1999), "Dynamic Factor Models", CREST DP.

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

Grossman, S., A. Melino and R.J. Shiller (1988), "Estimating the Continuous Time Consumption-Based Asset Pricing Model," *Journal of Business and Economic Statistics,* 5, 315-327.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029-1054.

Hansen, L.P. (1985), "A Method for Calculating Bounds in the Asymptotic Covariance Matrix of Generalized Method of Moments Estimators", *Journal of Econometrics*, 30, 203-238.

Hansen, L.P, J.C. Heaton and M. Ogaki (1988), "Efficiency Bounds Implied by Multiperiod Conditional Moment Restrictions", *Journal of the American Statistical Association*, 83, 863-871.

Hansen, L.P and R.J. Hodrick (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829-863.

Hansen, L.P. and J. Scheinkman (1985), "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes," *Econometrica*, 63, 767-804.

Hansen, L.P. and K.J. Singleton (1996), "Efficient Estimation of Linear Asset-Pricing Models With Moving Average Errors", *Journal of Business and Economic Statistics*, 14, 53-68.

Harvey, A.C., E. Ruiz and E. Sentana (1992), “Unobserved Component time series models with ARCH disturbances”, *Journal of Econometrics*, 52, 129-157.

Harvey, A.C., E. Ruiz and N. Shephard (1994), “Multivariate Stochastic Variance Models”, *Review of Economic Studies*, 61, 247-264.

Heston, S.L. (1993), “A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”, *Review of Financial Studies*, 6, 327-344.

Heston, S.L. and S. Nandi (1999), “A Closed-Form GARCH Option Pricing Model,” unpublished manuscript, Fereral Reserve Bank of Atlanta.

Hull, J. and A. White (1987), “The Pricing of Options on Assets with Stochastic Volatilities”, *The Journal of Finance*, Vol XLII, 281-300.

King, M., E. Sentana and S. Wadhwani (1994), “Volatility and Links between National Stock Markets”, *Econometrica*, 62, 901-933.

Lutkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer Verlag.

Meddahi, N. and E. Renault (1996), “Aggregation and Marginalization of GARCH and Stochastic Volatility Models”, Discussion paper, GREMAQ, Univesité de Toulouse.

Meddahi, N. and E. Renault (1999a), “Conditionning Information in Volatility Models,” mimeo, Université de Montréal.

Nelson, D.B. (1990), “ARCH Models as Diffusion Approximations,” *Journal of Econometrics*, 45, 7-39.

Nelson, D.B. (1991), “Conditional Heteroskedasticity in Asset Returns: A New Approach”, *Econometrica*, 59, 347-370.

Nelson, D.B. (1992), “Filtering and Forecasting aith Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model,” *Journal of Econometrics*, 52, 61-90.

Nelson, D.B. (1996), “Asymptotically Optimal Smoothing with ARCH Models”, *Econometrica*, 64, 561-573.

Nelson, D.B. and D.P. Foster (1994), “Asymptotic Filtering Theory for Univariate ARCH models”, *Econometrica*, 62, 1-41.

Nijman, TH. and E. Sentana (1996), “Marginalization and Contemporaneous Aggregation of Multivariate GARCH Processes”, *Journal of Econometrics*, 71, 71-87.

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

Taylor, S.J. (1986), *Modeling Financial Time Series*, John Wiley.

Weiss, A. (1984), “Systematic Sampling and Temporal Aggregation in Time Series Models,” *Journal of Econometrics*, 26, 271-281.