Parametric and Nonparametric Volatility Measurement[*]

Torben G. Andersen[a], Tim Bollerslev[b], and Francis X. Diebold[c]

July 2002

[a] Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208, and NBER
        phone: 847-467-1285, e-mail: t-andersen@kellogg.northwestern.edu

[b] Departments of Economics and Finance, Duke University, Durham, NC 27708, and NBER
        phone: 919-660-1846, e-mail: boller@econ.duke.edu

[c] Departments of Economics, Finance and Statistics, University of Pennsylvania, Philadelphia, PA 19104, and NBER
        phone: 215-898-1507, e-mail: fdiebold@sas.upenn.edu

Table of Contents

Abstract

References

ABSTRACT

Volatility has been one of the most active areas of research in empirical finance and time series econometrics during the past decade. This chapter provides a unified continuous-time, frictionless, no-arbitrage framework for systematically categorizing the various volatility concepts, measurement procedures, and modeling procedures. We define three different volatility concepts: (i) the notional volatility corresponding to the ex-post sample-path return variability over a fixed time interval, (ii) the ex-ante expected volatility over a fixed time interval, and (iii) the instantaneous volatility corresponding to the strength of the volatility process at a point in time. The parametric procedures rely on explicit functional form assumptions regarding the expected and/or instantaneous volatility. In the discrete-time ARCH class of models, the expectations are formulated in terms of directly observable variables, while the discrete- and continuous-time stochastic volatility models involve latent state variable(s). The nonparametric procedures are generally free from such functional form assumptions and hence afford estimates of notional volatility that are flexible yet consistent (as the sampling frequency of the underlying returns increases). The nonparametric procedures include ARCH filters and smoothers designed to measure the volatility over infinitesimally short horizons, as well as the recently-popularized realized volatility measures for (non-trivial) fixed-length time intervals.

1. INTRODUCTION

Since Engle's (1982) seminal paper on ARCH models, the econometrics literature has focused considerable attention on time-varying volatility and the development of new tools for volatility measurement, modeling and forecasting. These advances have in large part been motivated by the empirical observation that *financial asset return volatility* is time-varying in a persistent fashion, across assets, asset classes, time periods, and countries.[1] Asset return volatility, moreover, is central to finance, whether in asset pricing, portfolio allocation, or risk management, and standard financial econometric methods and models take on a very different, *conditional*, flavor when volatility is properly recognized to be time-varying.

The combination of powerful methodological advances and tremendous relevance in empirical finance produced explosive growth in the financial econometrics of volatility dynamics, with the econometrics and finance literatures cross-fertilizing each other furiously. Initial developments were tightly parametric, but the recent literature has moved in less parametric, and even fully nonparametric, directions. Here we review and provide a unified framework for interpreting both the parametric and nonparametric approaches.

In section 2, we define three different volatility concepts: (i) the notional volatility corresponding to the ex-post sample-path return variability over a fixed time interval, (ii) the ex-ante expected volatility over a fixed time interval, and (iii) the instantaneous volatility corresponding to the strength of the volatility process at a point in time.

In section 3, we survey parametric approaches to volatility modeling, which are based on explicit functional form assumptions regarding the expected and/or instantaneous volatility. In the discrete-time ARCH class of models, the expectations are formulated in terms of directly observable variables, while the discrete- and continuous-time stochastic volatility models both involve latent state variable(s).

In section 4, we survey nonparametric approaches to volatility modeling, which are generally free from such functional form assumptions and hence afford estimates of notional volatility that are flexible yet consistent (as the sampling frequency of the underlying returns increases). The nonparametric approaches include ARCH filters and smoothers designed to measure the volatility over infinitesimally short horizons, as well as the recently-popularized realized volatility measures for (non-trivial) fixed-length time intervals.

We conclude in section 5 by highlighting promising directions for future research.

---

[1] See, for example, Bollerslev, Chou and Kroner (1992).

2. VOLATILITY DEFINITIONS

Here we introduce a unified framework for defining and classifying different notions of return volatility in a continuous-time no-arbitrage setting. We begin by outlining the minimal set of regularity conditions invoked on the price processes and establish the notation used for the decomposition of returns into an expected, or mean, return and an innovation component. The resulting characterization of the price process is central to the development of our different volatility measures, and we rely on the concepts and notation introduced in this section throughout the chapter.

2.1. Continuous-Time No-Arbitrage Price Processes

Measurement of return volatility requires determination of the component of a given price increment that represents a return innovation as opposed to an expected price movement. In a discrete-time setting this identification may only be achieved through a direct specification of the conditional mean return, for example through an asset pricing model, as economic principles impose few binding constraints on the price process. However, within a frictionless continuous-time framework the no-arbitrage requirement quite generally guarantees that the return innovation is an order of magnitude larger than the mean return. This result is not only critical to the characterization of arbitrage-free continuous-time price processes, but it also has important implications for the approach one may employ for measurement and modeling of volatility over short intradaily return horizons.

We take as given a univariate risky logarithmic price process $p(t)$ defined on a complete probability space, $(\Omega, \mathscr{F}, P)$. The price process evolves in continuous time over the interval $[0,T]$, where T is a (finite) integer. The associated natural filtration is denoted $(\mathscr{F}_t)_{t\in[0,T]} \subseteq \mathscr{F}$, where the information set, $\mathscr{F}_t$, contains the full history (up to time t) of the realized values of the asset price and other relevant (possibly latent) state variables, and is otherwise assumed to satisfy the usual conditions. It is sometimes useful to consider the information set generated by the asset price history alone. We refer to this coarser filtration, consisting of the initial conditions and the history of the asset prices only, by $(F_t)_{t\in[0,T]} \subseteq F \equiv F_T$, so that by definition, $F_t \subseteq \mathscr{F}_t$. Finally, we assume there is an asset guaranteeing an instantaneously risk-free rate of interest, although we shall not refer to this rate explicitly. Many more risky assets may, of course, be available but we explicitly retain a univariate focus for notational simplicity. The extension to the multivariate setting is conceptually straightforward as discussed in specific instances below.

The continuously compounded return over the time interval $[t-h,t]$ is then

$$r(t,h) = p(t) - p(t-h), \qquad 0 \leq h \leq t \leq T . \tag{2.1}$$

We also adopt the following short hand notation for the cumulative return up to time t, i.e., the return over the $[0,t]$ time interval:

$$r(t) \equiv r(t,t) = p(t) - p(0), \qquad 0 \leq t \leq T. \tag{2.2}$$

These definitions imply a simple relation between the period-by-period and the cumulative returns that we use repeatedly in the sequel:

$$r(t,h) = r(t) - r(t-h), \qquad 0 \leq h \leq t \leq T. \tag{2.3}$$

A maintained assumption throughout is that - almost surely *(P)* (henceforth denoted *(a.s.)*) - the asset price process remains strictly positive and finite, so that *p(t)* and *r(t)* are well defined over *[0,T] (a.s.)*. It follows that *r(t)* has only countably many jump points over *[0,T],* and we adopt the convention of equating functions that have identical left and right limits everywhere.

Defining $r(t-) \equiv lim_{\tau \to t, \ \tau < t} r(\tau)$ and $r(t+) \equiv lim_{\tau \to t, \ \tau > t} r(\tau),$ uniquely determines the right-continuous, left-limit (càdlàg) version of the process, for which *r(t)=r(t+) (a.s.)*, and the left-continuous, right-limit (càglàd) version, for which *r(t)=r(t-) (a.s.),* for all *t* in *[0,T]*. In the following, we assume without loss of generality that we are working with the càdlàg version of the various return processes and components.

The jumps in the cumulative price and return process are then

$$\Delta r(t) \equiv r(t) - r(t-), \qquad 0 \leq t \leq T. \tag{2.4}$$

Obviously, at continuity points for *r(t)*, we have *Δr(t) = 0*. Notice also that, because there are at most a countably infinite number of jumps, a jump occurrence is unusual in the sense that we generically have

$$P(\Delta r(t) \neq 0) = 0 , \tag{2.5}$$

for an arbitrarily chosen *t* in *[0,T]*. This does not imply that jumps necessarily are rare. In fact, equation (2.5) is consistent with there being a (countably) infinite number of jumps over any discrete interval - a phenomenon referred to as an explosion. Jump processes that do not explode are termed *regular*. For regular processes, the anticipated jump frequency is conveniently characterized by the instantaneous jump intensity, i.e., the probability of a jump over the next instant of time, and expressed in units that reflect the expected (and finite) number of jumps per unit time interval.

In the following, we invoke the standard assumptions of no arbitrage opportunities and a finite expected return. Within our frictionless setting, it is well known that these conditions imply that the log-price process must constitute a (special) semi-martingale (e.g., Back, 1991). This, in turn, affords the following unique canonical return decomposition (e.g., Protter, 1992).

PROPOSITION 1 - Return Decomposition

Any arbitrage-free logarithmic price process subject to the regularity conditions outlined above may be uniquely represented as

$$r(t) \ \equiv \ p(t) - p(0) = \ \mu(t) + M(t) = \ \mu(t) + M^c(t) + M^J(t), \qquad (2.6)$$

where $\mu(t)$ is a predictable and finite variation process, $M(t)$ is a local martingale which may be further decomposed into $M^c(t),$ a continuous sample path, infinite variation local martingale component, and $M^J(t),$ a compensated jump martingale. All components may be assumed to have initial conditions normalized such that $\mu(0) \equiv M(0) \equiv M^c(0) \equiv M^J(0) \equiv 0,$ which implies that $r(t) \equiv p(t).$

Proposition 1 provides a unique decomposition of the instantaneous return into an expected return component and a (martingale) innovation. Over discrete intervals, the relation becomes slightly more complex. Letting the expected returns over *[t-h,t]* be denoted by *m(t,h)*, equation (2.6) implies

$$m(t,h) \ \equiv \ E[r(t,h)|\mathscr{F}_{t-h}] \ = \ E[\mu(t,h)|\mathscr{F}_{t-h}] \,, \quad 0<h \leq t \leq T, \qquad (2.7)$$

where

$$\mu(t,h) \ \equiv \ \mu(t) - \mu(t-h) \,, \qquad 0<h \leq t \leq T, \qquad (2.8)$$

and the return innovation takes the form

$$r(t,h) - \ m(t,h) \ = \ (\ \mu(t,h) - m(t,h) \ ) \ + \ M(t,h) \,, \qquad 0<h \leq t \leq T. \qquad (2.9)$$

The first term on the right hand side of (2.9) signifies that the expected return process, even though it is (locally) predictable, may evolve stochastically over the *[t-h,t]* interval.[2] If $\mu(t,h)$ is predetermined (measurable with respect to $\mathscr{F}_{t-h}$), and thus known at time *t-h,* then the discrete-time return innovation reduces to $M(t,h) \equiv M(t) - M(t-h).$ However, any shift in the expected return process during the interval will generally render the initial term on the right hand side of (2.9) non-zero and thus contribute to the return innovation over *[t-h,t].*

Although the discrete-time return innovation incorporates two distinct terms, the martingale component, *M(t,h)*, is generally the dominant contributor to the return variation over short intervals, i.e., for *h* small. In order to discuss the intuition behind this result, which we formalize in the following section, it is convenient to decompose the expected return process into a purely continuous, predictable finite variation part, $\mu^c(t),$ and a purely predictable jump part, $\mu^J(t).$

---

[2] In other words, even though the conditional mean is *locally* predictable, all return components in the special semi-martingale decomposition are generally stochastic: not only volatility, but also the jump intensity, the jump size distribution and the conditional mean process may evolve randomly over a finite interval.

Because the continuous component, $\mu^c(t)$, is of finite variation it is locally an order of magnitude smaller than the corresponding contribution from the continuous component of the innovation term, $M^c(t)$. The reason is - loosely speaking - that an asset earning, say, a positive expected return over the risk-free rate must have innovations that are an order of magnitude larger than the expected return over infinitesimal intervals. Otherwise, a sustained long position (infinitely many periods over any interval) in the risky asset will tend to be perfectly diversified due to a Law of Large Numbers, as the martingale part is uncorrelated. Thus, the risk-return relation becomes unbalanced. Only if the innovations are large, preventing the Law of Large Numbers from becoming operative, will this not constitute a violation of the no-arbitrage condition (e.g., Maheswaran and Sims, 1993). The presence of a non-trivial $M^J(t)$ component may similarly serve to eliminate arbitrage and retain a balanced risk-return trade-off relationship.

Analogous considerations apply to the jump component for the expected return process, $\mu^J(t)$, if this factor is present. There cannot be a predictable jump in the mean - i.e., a perfectly anticipated jump in terms of both time and size - unless it is accompanied by large jump innovation risk as well, so that $Pr(\Delta M(t) \neq 0) > 0$. Again - intuitively - if there was a known, say, positive jump then this induces arbitrage (by going long the asset) unless there is offsetting (jump) innovation risk.[3] Most of the continuous-time asset pricing literature ignores predictable jumps, even if they are logically consistent with the framework. One reason may be that their existence is fragile in the following sense. A fully anticipated jump must be associated with release of new (price relevant) information at a given point in time. However, if there is any uncertainty about the timing of the announcement so that it is only known to occur within a given minute, or even a few seconds, then the timing of the jump is more aptly modelled by a continuous hazard function where the jump probability at each point in time is zero, and the predictable jump event is thus eliminated. In addition, even if there were predictable jumps associated with scheduled news releases, the size of the predictable component of the jump is likely much smaller than the size of the associated jump innovation, so that the descriptive power lost by ignoring the possibility of predictable jumps is minimal. Thus, rather than modify the standard setup to allow for the presence of predictable (but empirically negligible) jumps, we follow the tradition in the literature and assume away such jumps.

Although we will not discuss specific model classes at length until later sections, it may be useful to briefly consider two simple examples to illustrate the somewhat abstract definitions given in the current section.

EXAMPLE 1: Stochastic Volatility Jump Diffusion with Non-Zero Mean Jumps

---

[3] This point is perhaps most readily understood by analogy to a discrete-time setting. When there is a predictable jump at time t, the instant from $t-$ to t is effectively equivalent to a trading period, say from $t-1$ to t, within a discrete-time model. In that setting, no asset can earn a positive (or negative) excess return relative to the risk-free rate over $(t-1,t]$ without bearing genuine risk as this would otherwise imply a trivial arbitrage opportunity. A predictable price jump without an associated positive probability of a jump innovation works entirely analogous.

Consider the following continuous-time jump diffusion expressed in stochastic differential equation (sde) form,

$$dp(t) = (\mu + \beta \sigma^2(t)) \, dt + \sigma(t) \, dW(t) + \kappa(t) \, dq(t), \qquad 0 \le t \le T,$$

where $\sigma(t)$ is a strictly positive continuous sample path process (a.s.), $W(t)$ denotes a standard Brownian motion, $q(t)$ is a pure jump process with $q(t)=1$ corresponding to a jump at time t and $q(t)=0$ otherwise, while $\kappa(t)$ refers to the size of the corresponding jumps. We assume the jump size distribution has a constant mean of μ_κ and variance of σ_κ^2. Finally, the jump intensity is assumed constant (and finite) at a rate λ per unit time. In the notation of Proposition 1, we then have the return components,

$$\mu(t) = \mu^c(t) = \mu \cdot t + \beta \int_0^t \sigma^2(s) \, ds + \lambda \cdot \mu_\kappa \cdot t,$$

$$M^c(t) = \int_0^t \sigma(s) \, dW(s),$$

$$M^J(t) = \Sigma_{0 \le s \le t} \, \kappa(s) \, q(s) - \lambda \cdot \mu_\kappa \cdot t.$$

Notice that the last term of the mean representation captures the expected contribution coming from the jumps, while the corresponding term is subtracted from the jump innovation process to provide a unique (compensated) jump martingale representation for M^J.

EXAMPLE 2: Discrete-Time Stochastic Volatility (ARCH) Model

Consider the discrete-time (jump) process for $p(t)$ defined over the unit time interval,

$$p(t) = p(t-1) + \mu + \beta \sigma^2(t) + \sigma(t) z(t), \qquad t = 1, 2, ..., T,$$

where (implicitly),

$$p(t+\tau) \equiv p(t), \qquad\qquad t = 0, 1, ..., T-1, \qquad\qquad 0 \le \tau < 1,$$

and $z(t)$ denotes a martingale difference sequence with unit variance, while $\sigma(t)$ is a (possibly latent) positive (a.s.) stochastic process that is measurable with respect to $\mathscr{F}_{t-1}$. Of course, in this situation the continuous-time no-arbitrage arguments based on infinitely many long-short positions over fixed length time intervals are no longer operative. Nonetheless, we shall later argue that the orders of magnitude of the corresponding terms in the decomposition in Proposition 1 remain suggestive and useful from an empirical perspective. In fact, we may still think of the price process evolving in continuous time, but the market being open and prices observed only at discrete points in time. In this specific model we have, for $t = 1, 2, ..., T,$

$$\mu(t) = \mu \cdot t + \beta \, \Sigma_{s=1,...,t} \, \sigma^2(s),$$

$$M(t) \equiv M^J(t) = \Sigma_{s=1,\ldots,t} \, \sigma(s) \, z(s) \, .$$

A final comment is in order. We purposely express the price changes and associated returns in Proposition 1 over a discrete time interval. The concept of an instantaneous return employed in the formulation of continuous-time models, that are given in sde form (as in Example 1 above), is pure short-hand notation that is formally defined only through the corresponding integral representation, such as equation (2.6). Although this is a technical point, it has an important empirical analogy: real-time price data are not available at every instant and, due to pertinent microstructure features, prices are invariably constrained to lie on a discrete grid, both in the price and time dimension. Hence, there is no real-world counterpart to the notion of a continuous sample path martingale with infinite variation over arbitrarily small time intervals (say, less than a second). It is only feasible to measure return (and volatility) realizations over discrete time intervals. Moreover, sensible measures can typically only be constructed over much longer horizons than given by the minimal interval length for which consecutive trade prices or quotes are recorded. We return to this point later. For now, we simply note that our main conceptualization of volatility in the next section conforms directly with the focus on realizations measured over non-trivial discrete time intervals rather than vanishing, or instantaneous, interval lengths.

2.2. Notional, Expected, and Instantaneous Volatility

This section introduces three distinct volatility concepts that serve to formalize the process of measuring and modeling volatility within our frictionless, arbitrage-free setting.

By definition volatility seeks to capture the strength of the (unexpected) return variation over a given period of time. However, two distinct features importantly differentiate the construction of all (reasonable) volatility measures. First, given a set of actual return observations, how is the *realized* volatility computed? Here, the emphasis is explicitly on *ex-post* measurement of the volatility. Second, decision making often requires forecasts of future return volatility. The focus is then on *ex-ante* expected volatility. The latter concept naturally calls for a model that may be used to map the current information set into a volatility *forecast*. In contrast, the (ex-post) realized volatility may be computed (or approximated) without reference to any specific model, thus rendering the task of volatility measurement essentially a nonparametric procedure.

It is natural first to concentrate on the behavior of the martingale component in the return decomposition (2.6). However, a prerequisite for observing the $M(t)$ process is that we have access to a continuous record of price data. Such data are simply not available and, even for extremely liquid markets, microstructure effects (discrete price grids, bid-ask bounce effects, etc.) prevent us from ever getting really close to a true continuous sample path realization. Consequently, we focus on measures that represent the (average) volatility over a discrete time

interval, rather than the instantaneous (point-in-time) volatility.[4] This, in turn, suggests a natural and general notion of volatility based on the quadratic variation process for the local martingale component in the unique semi-martingale return decomposition.

Specifically, let *X(t)* denote any (special) semi-martingale. The unique *quadratic variation process, [X,X]*$_t$*, t∈[0,T],* associated with *X(t)* is then formally defined by

$$[X,X]_t \;\equiv\; X(t)^2 - 2 \int_0^t X(s_-)\, dX(s), \qquad 0 < t \le T, \qquad\qquad (2.10)$$

where the stochastic integral of the adapted càglàd process, *X(s₋)*, with respect to the càdlàg semi-martingale, *X(s)*, is well-defined (e.g., Protter, 1992). It follows directly that the quadratic variation, *[X,X]*, is an increasing stochastic process. Also, jumps in the sample path of the quadratic variation process necessarily occur concurrent with the jumps in the underlying semi-martingale process, *Δ[X,X] =(ΔX)²*.

Importantly, if *M* is a locally square integrable martingale, then the associated *(M² - [M,M])* process is a local martingale,

$$E[\, M(t,h)^2 - (\, [M,M]_t - [M,M]_{t-h}) \mid \mathscr{F}_{t-h}\,] = 0,\; 0 < h \le t \le T. \qquad\qquad (2.11)$$

This relation, along with the following well-known result, provide the key to the interpretation of the quadratic variation process as one of our volatility measures.

PROPOSITION 2 - Theory of Quadratic Variation

Let a sequence of possibly random partitions of *[0,T]*, (*τ*$_m$), be given s.t. (*τ*$_m$) ≡ {*τ*$_{m,j}$}$_{j≥0}$, *m = 1,2, ...* where *τ*$_{m,0}$ ≤ *τ*$_{m,1}$ ≤ *τ*$_{m,2}$ ≤ *...* satisfy, with probability one, for *m → ∞,*

$$\tau_{m,0} \to 0\,; \qquad sup_{j\ge1}\, \tau_{m,j} \to T; \qquad sup_{j\ge0}\,(\tau_{m,j+1} - \tau_{m,j}) \to 0\,.$$

Then, for *t ∈ [0,T]*,

$$lim_{m\to\infty}\{\, \Sigma_{j\ge1}\,(X(t\wedge\tau_{m,j}) - X(t\wedge\tau_{m,j-1}))^2\} \;\to\; [X,X]_t\,,$$

where *t∧τ ≡ min (t, τ)*, and the convergence is uniform in probability.

Intuitively, the proposition says that the quadratic variation process represents the (cumulative) realized sample path variability of *X(t)* over the *[0,t]* time interval. This observation, together

---

[4] Of course, by choosing the interval very small, one may in principle approximate the notion of point-in-time volatility, as discussed further below.

with the martingale property of the quadratic variation process in (2.11), immediately point to the following theoretical notion of *ex-post* return variability.

DEFINITION 1 - Notional Volatility

The *Notional Volatility* over *[t-h,t]*, *0<h ≤ t ≤ T*, is

$$\upsilon^2(t,h) \;\equiv\; [M,M]_t - [M,M]_{t-h} \;=\; [M^c,M^c]_t - [M^c,M^c]_{t-h} + \Sigma_{t-h<s\leq t}\, \Delta M^2(s)\;. \tag{2.12}$$

This same volatility concept has recently been highlighted in a series of papers by Andersen, Bollerslev, Diebold and Labys (2001a,b) and Barndorff-Nielsen and Shephard (2002a,b,c). The latter authors term the corresponding concept *Actual Volatility*.

Under the maintained assumption of no predictable jumps in the return process, and noting that the quadratic variation of any finite variation process, such as $\mu^c(t)$, is zero, we also have

$$\upsilon^2(t,h) \;\equiv\; [r,r]_t - [r,r]_{t-h} \;=\; [M^c,M^c]_t - [M^c,M^c]_{t-h} + \Sigma_{t-h<s\leq t}\, \Delta r^2(s)\;. \tag{2.13}$$

Consequently, the notional volatility equals (the increment to) the quadratic variation for the return series. Equation (2.13) and Proposition 2 also suggest that (ex-post) it is possible to approximate the notional volatility arbitrarily well through the accumulation of ever finely sampled high-frequency squared return, and that this approach remains consistent independent of the expected return process. We shall return to a much more detailed analysis of this idea in our discussion of nonparametric ex-post volatility measures in section 4 below.

Similarly, from (2.13) and Proposition 2 it is evident that the notional volatility, $\upsilon^2(t,h)$, directly captures the sample path variability of the log-price process over the *[t-h,t]* time interval. In particular, the notional volatility explicitly incorporates the effect of (realized) jumps in the price process: jumps contribute to the realized return variability and forecasts of volatility must account for the potential occurrence of such jumps. It also follows, from the properties of the quadratic variation process, that

$$E[\upsilon^2(t,h)|\mathscr{F}_{t-h}] \;=\; E[M(t,h)^2\,|\mathscr{F}_{t-h}] = E[M^2(t)|\mathscr{F}_{t-h}] - M^2(t-h), \qquad 0<h\leq t \leq T. \tag{2.14}$$

Hence, the *expected notional volatility* represents the expected future (cumulative) squared return innovation. As argued in section 2.1, this component is typically the dominant determinant of the expected return volatility.

For illustration, consider again the two examples introduced in section 2.1 above. Alternative more complicated specifications and issues related to longer horizon returns are considered in section 3.

EXAMPLE 1: Stochastic Volatility Jump Diffusion with Non-Zero Mean Jumps (Revisited)

The log-price process evolves according to

$$dp(t) = (\mu + \beta\sigma^2(t)) \, dt + \sigma(t) \, dW(t) + \kappa(t) \, dq(t), \qquad 0 \le t \le T.$$

The notional volatility is then

$$\upsilon^2(t,h) = \int_0^h \sigma^2(t-h+s) \, ds + \Sigma_{t-h<s\le t} \, \kappa^2(s).$$

The expected notional volatility involves taking the conditional expectation of this expression. Without an explicit model for the volatility process, this cannot be given in closed form. However, for small h the (expected) notional volatility is typically very close to the value attained if volatility is constant. In particular, to a first-order approximation,

$$E[\upsilon^2(t,h)|\mathscr{F}_{t-h}] \approx \sigma^2(t-h) \cdot h + \lambda \cdot h \cdot (\mu_\kappa^2 + \sigma_\kappa^2) \approx [\sigma^2(t-h) + \lambda(\mu_\kappa^2 + \sigma_\kappa^2)] \cdot h,$$

while

$$m(t,h) \approx [\mu + \beta \cdot \sigma^2(t-h) + \lambda \cdot \mu_\kappa] \cdot h.$$

Thus, the expected notional volatility is of order h, the expected return is of order h (and the variation of the mean return of order h^2), whereas the martingale (return innovation) is of the order $h^{1/2}$, and hence an order of magnitude larger for small h.

EXAMPLE 2: Discrete-Time Stochastic Volatility (ARCH) Model (Revisited)

The discrete-time (jump) process for $p(t)$ defined over the unit time interval is

$$p(t) = p(t-1) + \mu + \beta\sigma^2(t) + \sigma(t) \, z(t), \qquad t = 1, 2, ..., T.$$

The one-period notional volatility measure is then

$$\upsilon^2(t,1) = \sigma^2(t) \, z^2(t),$$

while, of course, the expected notional volatility is

$$E[\upsilon^2(t,1)|\mathscr{F}_{t-1}] = \sigma^2(t),$$

and the expected return is

$$m(t,1) = \mu + \beta\sigma^2(t).$$

Although direct comparison of the order of magnitude of the mean return relative to the return innovation is not feasible here, empirical studies based on daily and weekly returns invariably find the mean parameters, such as μ and β, to be very small relative to the expected notional volatility, $\sigma^2(t)$. In practice, the contribution of innovations in the mean process to the overall return variability is negligible over such horizons, as also suggested by Example 1.

It is obvious that - when volatility is stochastic - the ex-post (realized) notional volatility will not correspond to the ex-ante expected volatility. More importantly, equation (2.9) implies that even the ex-ante expected notional volatility generally is *not* identical to the usual notion of return volatility as an *ex-ante* characterization of future return variability over a discrete holding period. The fact that the latter quantity is highly relevant for financial decision making motivates the standard discrete-time expected volatility concept defined below.

DEFINITION 2 - Expected Volatility

The *Expected Volatility* over *[t-h,t]*, $0 < h \leq t \leq T$, is defined by

$$\vartheta^2(t,h) \equiv E[\,\{r(t,h) - E(\mu(t,h)|\mathscr{F}_{t-h})\}^2 \,|\mathscr{F}_{t-h}\,]$$

$$= E[\,\{r(t,h) - m(t,h)\}^2 \,|\mathscr{F}_{t-h}\,]. \tag{2.15}$$

If the $\mu(t,h)$ process is not measurable with respect to $\mathscr{F}_{t-h}$, the expected volatility will typically differ from the expected notional volatility in equation (2.14). Specifically, the future return variability in equation (2.15) reflects both genuine return innovations, as in equation (2.14), and intra-period innovations to the conditional mean process. Trivially, however, for models with an assumed constant mean return, or for one-period-ahead discrete-time volatility forecasts with given conditional mean representation (as in Example 2 above), the two concepts coincide.

Continuous-time models often portray the volatility process as perpetually evolving. From this perspective, the focus on volatility measurement over a fixed interval length, *h,* is ultimately arbitrary. A more natural theoretical concept is provided by the expected instantaneous volatility, measured as the current strength of the volatility process expressed per unit of time,

$$lim_{h\to 0}\, \vartheta^2(t,h)/h = lim_{h\to 0}\,[\,E\{\,([M,M]_t - [M,M]_{t-h})/h\}\,|\mathscr{F}_{t-h}\,]. \tag{2.16}$$

This is especially true when the underlying logarithmic price path is continuous, i.e., $M^J(t) \equiv 0$, in which case the (scaled) notional and expected instantaneous volatilities coincide,

$$lim_{h\to 0}\, \vartheta^2(t,h)/h = lim_{h\to 0}\, \upsilon^2(t,h)/h = lim_{h\to 0}\,[\,E\{\,([M^c,M^c]_t - [M^c,M^c]_{t-h})/h\}\,|\mathscr{F}_{t-h}\,].$$

Inspired by these results, we adopt the corresponding definition of instantaneous volatility.

DEFINITION 3 - Instantaneous Volatility

The *Instantaneous Volatility* at time *t*, $0 \leq t \leq T$, is

$$\sigma_t^2 \equiv lim_{h \to 0} \ [\ E\{ \ ([M^c,M^c]_t - [M^c,M^c]_{t-h})/h\} \ | \mathscr{F}_{t-h} \ ]. \tag{2.17}$$

This definition is consistent with the terminology commonly employed in the literature on continuous-time parametric stochastic volatility models.[5] Barndorff-Nielsen and Shephard (2002a,b,c), in a slightly different setting, refer to the corresponding concept as *Spot Volatility*.

Although the instantaneous volatility is a natural concept, and we refer to it frequently below, practical volatility measurement invariably takes place over discrete time intervals. In the sequel we stress the general relationship between the various volatility concepts for alternative parametric volatility models and nonparametric volatility measurements.

2.3. Volatility Modeling and Measurement

This section provides a qualitative overview of the different approaches to volatility modeling and measurement. Specific characterization of the volatility estimates and measurements are postponed to the following sections, where we present a more detailed study for each major class of models. The approaches for empirically quantifying volatility naturally falls into two separate categories, namely procedures based on estimation of parametric models and more direct nonparametric measurements. Within the parametric volatility classification, alternative models exploit different assumptions regarding the expected volatility, $\vartheta^2(t,h)$, through distinct functional forms and the nature of the variables in the information set, $\mathscr{F}_{t-h}$. In contrast, the data-driven or nonparametric volatility measurements typically quantify the notional volatility, $\upsilon^2(t,h)$, directly. Both set of procedures differ importantly in terms of the choice of time interval for which the volatility measure applies, e.g., a discrete interval, $h>0$, or a point-in-time (instantaneous) measure, obtained as the limiting case for $h \to 0$.

Within the discrete-time parametric models, the most significant distinction concerns the character of the variables in the information set, $\mathscr{F}_{t-h}$, which in turn governs the type of estimation and inference techniques that are required for their practical implementation. In the ARCH class of models, the expected volatility, $\vartheta^2(t,h)$, is parameterized as a function of past returns only, or F_{t-h}, although other observable variables could easily be included in $\mathscr{F}_{t-h}$. In contrast, the parameterized expectations in the Stochastic Volatility (SV) class of models explicitly rely on latent state variables. As we move to continuous-time parametric representations of either model, the assumption that all past returns are observable implies that the distinction between the two classes of models effectively vanishes, as the latent volatility

---

[5] A definition of instantaneous volatility similar to equation (2.17) is suggested by Comte and Renault (1998) in their discussion of alternative inference procedures for a continuous-time long-memory volatility model.

state variables may be extracted without error from the frictionless, continuous-time price record. The following definition formalizes these categorizations.

DEFINITION 4 - Parametric Volatility Models

Discrete-Time Parametric Volatility Models explicitly parameterize the expected volatility, $\vartheta^2(t,h)$, $h>0$, as a non-trivial function of the time *t-h* information set, $\mathcal{F}_{t-h}$. In the *ARCH* class of models, $\mathcal{F}_{t-h}$ depends on past returns and other directly observable variables only. In the *Stochastic Volatility (SV)* class of models $\mathcal{F}_{t-h}$ explicitly incorporates past returns as well as *latent* state variables. *Continuous-Time Volatility Models* provide an explicit parameterization of the instantaneous volatility, σ_t^2, as a (non-trivial) function of the $\mathcal{F}_t$ information set, with additional volatility dynamics possibly introduced through time variation in the process governing jumps in the price path.

In addition to these three separate model classes, so-called *Implied Volatility* approaches also figure prominently in the literature. The implied volatilities are based on a parametric volatility model for the returns, as defined above, along with an *asset pricing model* and an augmented information set consisting of options prices and/or term structure variables. Intuitively, if the number of derivatives prices at time *t-h* pertaining to the price of the asset at time *t* included in the augmented information set, $\mathcal{F}_{t-h}$, exceeds the number of latent state variables in the parametric model for the returns, it is possible to back out a value for $\vartheta^2(t,h)$ by inverting the theoretical asset pricing model.[6,7] Importantly, these procedures all depend (implicitly) on specific assumptions about the price of volatility risk. The statistical measurements surveyed in this chapter are void of any such assumptions; see e.g., Bates (1996b), Renault (1997) and the chapters by Garcia, Ghysels, and Renault (2002), and Piazzesi (2002) in this *Handbook* for a discussion of the extensive literature on options implied volatilities and related procedures.

In contrast to the parametric procedures categorized above, the nonparametric volatility measurements are generally void of any specific functional form assumptions about the stochastic process(es) governing the local martingale, *M(t)*, as well as the predictable and finite

---

[6] Most prominent among these procedures are, of course, the Black-Scholes option implied volatilities based on the assumption of an underlying continuous-time Random Walk model first analyzed empirically by Latané and Rendleman (1976). More detailed empirical analyses of Black-Scholes implied volatilities along with generalizations to allow for more realistic price dynamics have been the subject of an enormous literature, an incomplete list of which includes Bakshi, Cao and Chen (1997), Canina and Figlewski (1993), Chernov and Ghysels (2000), Christensen and Prabhala (1998), Day and Lewis (1992), Duan (1995), Dumas, Fleming and Whaley (1998), Fleming (1998), Heston (1993), Heston and Nandi (2000), Hull and White (1987), and Wiggins (1992).

[7] The 30-day VIX implied volatility index for the S&P100 and the more recent VXN index for the NASDAQ-100 traded on the Chicago Board Options Exchange (CBOE) are based on weighted averages of Black-Scholes put and call implied volatilities; see Fleming, Ostdiek and Whaley (1995) and Whaley (1993) for further discussion. A similar construct underlies the 45-day VDAX for the DAX index on the Deutsche Termin Börse (DTB).

variation process, $\mu(t)$, in the unique return decomposition. These procedures also differ importantly from the parametric models in their focus on providing measures of the notional volatility, $\upsilon^2(t,h)$, rather than the expected volatility, $\vartheta^2(t,h)$. In addition, the nonparametric procedures generally restrict the measurements to be functions of the coarser filtration, F_t, generated by the return on the asset only. In parallel to the parametric measures, the nonparametric procedures may be further differentiated depending upon whether they let $h\rightarrow 0$ and thus provide measures of instantaneous volatility, or whether they explicitly operate with a strictly positive $h>0$ resulting in realized volatility measures over a discrete time interval.

DEFINITION 5 - Nonparametric Volatility Measurement

Nonparametric Volatility Measurement utilizes the *ex-post* returns, or F_τ, in extracting measures of the notational volatility. *ARCH Filters and Smoothers* rely on continuous sample paths, or $M^J \equiv 0$, in measuring the instantaneous volatility $\sigma_t^2 = lim_{h\rightarrow 0} \vartheta^2(t+h,h)/h$. The filters only utilize information up to time $\tau=t$, while the smoothers are based on $\tau>t$. *Realized Volatility* measures directly quantify the notional volatility, $\upsilon^2(t,h)$, over (non-trivial) fixed-length time intervals, $h>0$.

Within the class of instantaneous volatility measures, the ARCH filters rely only on the past return record, typically through a weighted rolling regression, while the smoothers, or two-sided filters, also exploit (ex-post) the future price record. Realized volatility approaches may similarly be categorized according to whether the measurement of $\upsilon^2(t,h)$ exploits only price observations within the time interval *[t-h,t]* itself, or filtering/smoothing techniques are used to also incorporate return observations outside of *[t-h,t]*. An important advantage of realized volatility procedures, which exploit only interval specific information, is that they provide asymptotically unbiased measures, and therefore approximately serially uncorrelated measurement errors, under quite general conditions. A potential drawback is that useful information from adjacent intervals is ignored. Consistency of both ARCH filters and smoothers and realized volatility procedures generally require that the length of the underlying sampling interval for the returns within the *[t-h,t]* time interval approaches zero (even for the ARCH filters and smoothers where h itself is shrinking). We next turn to a more detailed discussion of these different procedures for modeling and measuring volatility within the context of the general setup in section 2.1.

3. PARAMETRIC METHODS

Parametric volatility models and their implementation constitute one of the cornerstones of modern empirical asset pricing, and a large econometrics and statistics literature has been devoted to the development and theoretical foundation of differently parameterized volatility models. A thorough review of this literature is beyond the scope of this chapter. Instead, we shall merely highlight the most important ideas as they relate specifically to the volatility

measurement for the different model classes.

3.1. Discrete-Time Models

Even if trading and pricing evolves in continuous time within the frictionless no-arbitrage setting outlined in section 2, it is sometimes more convenient to formulate a model for the associated discrete-time returns. Of course, such an approach is naturally motivated by situations in which the prices are only observed at regular fixed time intervals (e.g. daily closing prices, end-of-the-month prices). Alternatively, if trading is only feasible at the given discrete points in time, the relevant return distribution is captured fully by the conditional discrete-time dynamics. Either perspective allows us to embed the discrete-time ARCH and SV models in our basic continuous-time setting. Hence, for the remainder of this section, we assume that prices are only observed (and trades only possible) at discrete and equally spaced points in time, $t = 0, h, 2 \cdot h, \dots, T-h, T$.

The discrete-time models, at a minimum, assume that the correct specification of the h-step-ahead conditional mean and variances is known up to a low-dimensional parameter vector. That is, the models (parsimoniously) parameterize the first two conditional return moments,[8]

$$m(t,h) = E[\, r(t,h) \,|\mathscr{F}_{t-h}\,] = E[\mu(t,h)|\mathscr{F}_{t-h}\,] = \mu(t,h) \tag{3.1}$$

$$\vartheta^2(t,h) = E[\, (r(t,h) - m(t,h))^2 \,|\mathscr{F}_{t-h}\,] , \tag{3.2}$$

where $m(t,h)$ and $\mu(t,h)$ coincide because the h-step-ahead conditional mean is predictable. Of course, in general the first two conditional moments of the h-period returns do not fully characterize the dynamic return distribution, although this is true for continuous sample path diffusion models ($h \to 0$), which may be defined completely through the instantaneous drift and volatility coefficients alone.

The restriction of only observing prices at equidistant points in time is readily interpreted, within the continuous-time setting, as a pure jump process with known jump times but random jump sizes. In the notation of the previous section,

$$\Delta M(t) = r(t,h) - \mu(t,h) = r(t,h) - m(t,h), \qquad t = h, \dots, T . \tag{3.3}$$

As such, it follows directly from the definition of the notional volatility over $[t-h,t]$ that

$$v^2(t,h) \equiv [M,M]_t - [M,M]_{t-h} = \Delta M^2(t). \tag{3.4}$$

Moreover, the expected notional volatility over $[t-h,t]$ simply equals the conditional h-period-ahead variance as specified by the model,

---

[8] The information set, $\mathscr{F}_{t-h}$, is (implicitly) restricted to the corresponding discrete-time realizations of the process along with any other discrete-time (possibly latent) state variables.

$$\vartheta^2(t,h) \;=\; E[\upsilon^2(t,h)|\mathscr{F}_{t\text{-}h}] \;=\; E[\Delta M^2(t)|\mathscr{F}_{t\text{-}h}].\tag{3.5}$$

This same result is generally not true for multi-period forecasts, or volatilities over longer horizons, *[t-k· h, t]* where *k>1*. In this situation, any variation in the conditional mean process within the forecast horizon will contribute to the return variation, so the expected notional volatility typically is not equal to the expected volatility. However, as discussed further below, the contribution from the variation in the conditional mean will usually only be of second order importance unless the forecast horizon is very long. This same basic insights carry over to the continuous-time processes discussed in more detail in section 3.2 below.

In order to more explicitly clarify the relationship between the notional volatility and total return variability within the multi-period setting, recall the generic return decomposition for a discrete-time pure-jump process in Proposition 1,

$$r(t) \;=\; \mu(t) + M^J(t)\,, \qquad t = 0,\, h,\, 2\cdot h,\, \dots,\, T\text{-}h,\, T,\tag{3.6}$$

where

$$\mu(t) \;=\; \Sigma_{\tau=1,..,t/h}\; E(r(\tau\cdot h,h)|\mathscr{F}_{(\tau\text{-}1)\cdot h}) \;=\; \Sigma_{\tau=1,..,t/h}\; \mu(\tau\cdot h,h)\,,\tag{3.7}$$

$$M^J(t) \;=\; \Sigma_{\tau=1,..,t/h}\; \Delta M(\tau) \;=\; \Sigma_{\tau=1,..,t/h}\; (r(\tau\cdot h,h) - \mu(\tau\cdot h,h)).\tag{3.8}$$

Now, for any integer *k>0*,

$$\mu(t,k\cdot h) \;=\; \mu(t) - \mu(t\text{-}k\cdot h) \;=\; \Sigma_{\tau=1,...,k}\; \mu(t\text{-}(k\text{-}\tau)\cdot h,h).$$

This $\mu(t,k\cdot h)$ term represents the *cumulative conditional h-period-ahead expected return* and not the *conditional multi h-period* expected return. Specifically, for *k>1*, the term $\mu(t,k\cdot h)$ is generally not equal to $E(r(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}) = m(t,k\cdot h)$, even though

$$E[\mu(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}] \;=\; m(t,k\cdot h).\tag{3.9}$$

Hence, we obtain the decomposition of the *k*-period expected volatility over *[t-k·h,t]*,

$$
\begin{aligned}
\vartheta^2(t,k\cdot h) \;=\;& E[\,(r(t,k\cdot h)\text{-}m(t,k\cdot h))^2\,|\mathscr{F}_{t\text{-}k\cdot h}\,] \\[4pt]
=\;& E[M^2(t,k\cdot h) + (\mu(t,k\cdot h)\text{-} m(t,k\cdot h))^2 + 2\cdot M(t,k\cdot h)\cdot\mu(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}] \\[4pt]
=\;& Var[\,M(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}] \;+\; Var[\,\mu(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}] \\[4pt]
& +\; 2\cdot Cov[\,M(t,k\cdot h),\, \mu(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}] \\[4pt]
=\;& E(\,[M,M]_t - [M,M]_{t\text{-}k\cdot h}\,|\mathscr{F}_{t\text{-}k\cdot h}) \;+\; Var[\,\mu(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}]
\end{aligned}
\tag{3.10}
$$

$$+ \; 2 \cdot Cov[\; M(t,k\cdot h), \; \mu(t,k\cdot h)|\mathscr{F}_{t\text{-}k\cdot h}] \; .$$

Trivially, as noted above, for the one-period-ahead forecasts, or *k=1,* there cannot be any within-period variability in the conditional mean process, so that the last two terms in (3.10) vanishes, and the expected volatility equals the expected notional volatility. However, for multiple-period forecasts the stochastic evolution of the conditional mean within the interval contributes to the overall return variability, both through the variation in the conditional mean itself and through the covariance between the return innovations and future (within forecast horizon) changes in the conditional mean return. However, the period-by-period conditional mean is generally much smaller than the volatility, and the shifts in the conditional mean smaller yet. Hence, the expected notional volatility, or quadratic variation, remains the dominant component for the multi-period return variability in empirically realistic situations.[9]

Further, notice that the so-called *Leverage Effect* (e.g., Black, 1976) impacts only the expected notional volatility (quadratic variation), and none of the other terms. The hypothesis stipulates a (negative) correlation between the return innovations, *ΔM,* and the size of future return innovations, *(ΔM)²*, essentially predicting a left-skewed distribution for the return innovations. With no impact on the conditional mean, only the quadratic variation process is affected. The closely related *Volatility Feedback Effect* has an impact via the covariance term, but it remains limited by the size of the shifts in the conditional mean. Again, the hypothesis essentially implies a leftward skew in the return innovation distribution. Intuitively, given a positive volatility risk premium, large negative return innovations are magnified while large positive innovations are dampened due to the increase in the expected future return required to compensate for a positive and persistent shock to future volatility. Hence, technically, the volatility feedback tends to raise the expected volatility directly, but it also induces a negative correlation between the return innovations and the future expected mean returns.

Returning to the basic discrete-time setup, the conditional moments in equations (3.1) and (3.2) allow for relatively easy and consistent statistical inference concerning the unknown parameters by a standard Generalized Methods of Moments (GMM) estimator (Hansen, 1982), or for stochastic volatility and latent state variable(s), a Simulated Method of Moments (SMM) type estimator (Duffie and Singleton, 1993). Of course, simple method-of-moments estimators with ill chosen moment conditions may behave poorly, both asymptotically and in finite samples (Andersen, Chung and Sørensen, 1999), and much of the literature on discrete-time volatility models is concerned with the development of more efficient estimation procedures under auxiliary assumptions. In particular, assuming that the standardized innovations, *(r(t,h)- μ(t,h))/ϑ(t,h)*, belong to a specific parametric family of distributions, Maximum Likelihood Estimation (MLE) and corresponding Gaussian Quasi-MLE (QMLE) procedures (Bollerslev and Wooldridge, 1992) are both conceptually straightforward to implement for the ARCH

---

[9] Of course, the exact terms involved in the multiple-period volatility forecasts will depend upon the specific functional form and the underlying distributional assumptions. Their practical computation may not be trivial, or even feasible in closed form, necessitating the use of numerical simulation techniques (see, e.g., Geweke, 1989). We shall not be concerned with these more computationally oriented aspects of the problem in this chapter.

models, while more complicated procedures are required for stochastic volatility models.

Next, we briefly review some of the popular discrete-time parametric volatility models. The key distinguishing features for each class of models consist of the functional form for the conditional moments in equations (3.1) and (3.2), the variables in the information set $\mathcal{F}_{t-h}$, along with any additional distributional assumptions. The performance of the different models, such as the fit to the data and precision of forecasts, as well as the ease of computing parameter estimates and the various terms in the volatility forecast expressions, depend importantly on these features.

3.1.1. ARCH Models

The ARCH class of models was first introduced in the seminal paper by Engle (1982). It has since enjoyed unprecedented empirical success along with a myriad of extensions and further theoretical developments. Indeed, most of our empirical knowledge to date concerning the temporal dependencies in financial market volatility have arguably been gleaned from estimation and inference with ARCH type models. Several surveys of this burgeoning literature already exist (an incomplete list of which includes, Andersen and Bollerslev, 1998c; Bollerslev, Chou and Kroner, 1992; Bollerslev, Engle and Nelson, 1994; Diebold and Lopez, 1995; Engle, 1995; and Engle and Patton, 2001), and we will not attempt yet another comprehensive review. However, it is useful to briefly summarize the key developments and model formulations within the current framework.

The ARCH class of models differ from the discrete-time stochastic volatility models discussed below by the assumption that the filtration, $\mathcal{F}_{t-h}$, underlying the parameterized conditional expectations in equation (3.1) and (3.2) depends exclusively on observable variables. This assumption greatly facilitates statistical inference vis-a-vis stochastic volatility models, and the widespread empirical use of ARCH style models in part stems from the ease with which traditional (quasi-) maximum likelihood based procedures may be implemented.

Any time series model in which the conditional variance depends non-trivially on the time $t-h$ observable information set is now commonly referred to as an ARCH model. This terminology is explained by the particular parametric formulation first adapted by Engle (1982). Specifically, in the so-called ARCH(p) model, $\vartheta^2(t,h)$ is parameterized as an autoregressive distributed lag of p squared innovations,

$$\vartheta^2(t,h) = \omega + \sum_{j=1,\ldots,p} \alpha_j (r(t-j \cdot h,h) - \mu(t-j \cdot h,h))^2 \equiv \omega + \alpha(L,h)(r(t,h)-\mu(t,h))^2, \qquad (3.11)$$

where $\omega > 0$ and $\alpha_j \geq 0$ in order to ensure positivity of $\vartheta^2(t,h)$ (a.s.). A more parsimonious characterization of the intertemporal volatility dependencies is often obtained by the Generalized ARCH, or GARCH(p,q), model (Bollerslev, 1986),

$$\vartheta^2(t,h) = \omega + \sum_{j=1,\ldots,p} \alpha_j (r(t-j \cdot h,h) - \mu(t-j \cdot h,h))^2 + \sum_{i=1,\ldots,q} \beta_i \vartheta^2(t-j \cdot h,h)$$

$$\equiv \omega + \alpha(L,h) \ (r(t,h)-\mu(t,h))^2 + \beta(L,h) \ \vartheta^2(t,h).$$

For the popular GARCH(1,1) model, $\omega>0$, $\alpha_1 \geq 0$, and $\beta_1 \geq 0$ obviously guarantees positivity of $\vartheta^2(t,h)$. Corresponding conditions for the general case are presented in Nelson and Cao (1992). Rearranging the terms, the GARCH(p,q) model is readily interpreted as an ARMA model for $[r(t,h)-\mu(t,h)]^2$ in which the autoregressive and moving average polynomials are given by $[\alpha(L,h) + \beta(L,h)]$ and $[1 - \beta(L,h)]$, respectively. Hence, provided that all the roots of the characteristic equation, $\alpha(x,h)+\beta(x,h)=1$, have norm greater than one, the model is covariance stationary, and the unconditional h-period (one-period) variance equals $E[\vartheta^2(t,h)] = \omega(1-\alpha(1,h)+\beta(1,h))^{-1}$. Weaker conditions for strict stationarity have been derived by Nelson (1990b) and Bougerol and Picard (1992), while higher order moment conditions have recently been developed by Ling and McAleer (2002).

The leverage effect, briefly discussed earlier, stipulates a negative correlation between current return innovations and future expected conditional variances. The GJR-GARCH model (Glosten, Jagannathan and Runkle, 1993) in which the α_j coefficients in $\alpha(L,h)$ in equation (3.12) depend on the sign of the corresponding return innovations, $r(t-j \cdot h,h)-\mu(t-j \cdot h,h)$, was specifically designed to accommodate such asymmetries. A similar motivation underlies the EGARCH model in Nelson (1991). Defining the standardized innovations,

$$z(t,h) \equiv (r(t,h)-\mu(t,h))/\vartheta(t,h), \tag{3.13}$$

the EGARCH(p,q) model takes the form,

$$log[\vartheta^2(t,h)] = \omega + \alpha(L,h)\{\theta \cdot z(t,h) + \gamma \cdot [|z(t,h))| - E(|z(t,h))|)]\}$$
$$+ \beta(L,h) \ log[\vartheta^2(t,h)], \tag{3.14}$$

where as before $\alpha(L,h)$ and $\beta(L,h)$ denote pth and qth order lag polynomials respectively. The log-transform complicates the calculation of (unbiased) multi-step conditional variance forecasts, but conveniently avoids having to impose non-negativity constraints on the parameters. Obviously, for $\theta<0$ the model predicts a negative relation between current returns and future conditional variances.

Alternatively, as discussed above asymmetries in the return-volatility relationship may also be attributed to the so-called volatility feedback effect. This feature is captured by the ARCH-in-Mean type formulation (Engle, Lilien and Robins, 1987), in which the functional form for the conditional mean, $\mu(t,h)$, depends explicitly on the conditional variance of the process, $\vartheta^2(t,h)$. Which of these competing specifications is best able to capture the empirically observed asymmetry in equity return volatility has been the subject of several empirical studies (e.g., Bekaert and Wu, 2000, and Campbell and Hentschel, 1992).

Another important empirical finding concerns the strong degree of volatility persistence

estimated with most daily and weekly financial rates of return. This is manifest by the autoregressive polynomials in the GARCH(p,q) formulations, $1-\alpha(x,h)-\beta(x,h)$, and the EGARCH formulations, $1-\beta(x,h)$, having (their largest) roots very close to unity. The IGARCH model of Engle and Bollerslev (1986) directly imposes this condition; i.e., $\alpha(1,h)+\beta(1,h)=1$. However, the imposition of a unit-root in the conditional variance arguably exaggerates the true dynamic dependencies, and several alternative long-memory, or fractionally integrated, ARCH type formulations have recently been estimated and analyzed more formally in the literature (e.g., Baillie, Bollerslev and Mikkelsen, 1996; Bollerslev and Mikkelsen, 1996; Ding, Granger and Engle, 1993; Giraitis, Kokoszka and Leipus, 2000; Robinson, 1991, 2001; and Zumbach, 2001). Possible explanations for the apparent long-memory dependencies based on the aggregation of multiple volatility components and/or stochastic regime switching models have been explored by Andersen and Bollerslev (1997), Diebold and Inoue (2001), and Liu (2000) among others (see also the related component model in Engle and Lee, 1999). This remains a very active area of current research.

Our focus in this chapter has been almost exclusively univariate. Nonetheless, most interesting questions in asset pricing finance and risk management call for a multivariate framework involving not just conditional variances but also time-varying conditional covariances. From a conceptual view point, the extension of the univariate ARCH class of models to a multivariate setting presents few new issues. However, conditions to ensure that the parameterized conditional covariance matrices are positive definite (a.s.) and involve only a manageable (small) number of parameters are both important considerations from a practical perspective. In the Diagonal GARCH model of Bollerslev, Engle and Wooldridge (1988) the conditional variances and covariances are parameterized as univariate GARCH(p,q) processes; i.e., the ijth element in the conditional covariance matrix depends on a distributed lag of past values of the same element and the cross-products of the corresponding innovations. The related BEKK GARCH formulation (Engle and Kroner, 1995) guarantees that the covariance matrices are positive definite. The constant conditional correlation model in Bollerslev (1990) is empirically among the most frequently applied multivariate ARCH models. This model has recently been extended to incorporate parsimoneously parameterized time-varying conditional correlations by Engle (2001). Other multivariate formulations that allow for relatively easy implementation in large dimensions include the Flexible GARCH model of Ledoit, Santa-Clara and Wolf (2001) and the R-GARCH model in Gallant and Tauchen (2000).

Most industry applications entailing large scale covariance matrix measurements rely on J.P. Morgan's RiskMetrics (J.P. Morgan, 1997). The RiskMetrics procedure is based on exponential smoothing, and as such corresponds directly to a Diagonal IGARCH(1,1) model in which all of the intercepts in the conditional covariance matrix are fixed at zero and identical values of α and $\beta \equiv 1-\alpha$ are employed across all assets. The use of the same smoothing parameter ($\beta = 0.94$ with daily data) obviously facilitates the implementation and automatically guarantees that the covariance matrix measurements are positive definite. Nonetheless, when viewed as a data generating process as opposed to a filter, the RiskMetrics procedure is formally degenerate (Nelson, 1990b).

One major theoretical drawback to the GARCH class of models concerns their lack of closed-form aggregation. This is true both inter-temporally and cross-sectionally. For example, if daily asset returns follow a univariate GARCH(p,q) model, the corresponding weekly returns are not GARCH(p,q). Similarly, if a collection of asset returns follow a multivariate GARCH(p,q) model, (non-trivial) portfolio returns are not GARCH(p,q). The Weak GARCH class of models was explicitly introduced by Drost and Nijman (1993) and Nijman and Sentana (1996) to address this issue. In a Weak GARCH model, $\vartheta^2(t,h)$ has the interpretation of a parameterized linear projection for the squared innovation. In contrast to the conditional expectations underlying the standard ARCH formulations, the linear projections are closed under temporal (and in the multivariate case cross-sectional) aggregation. However, the linear projections do not easily translate into the volatility concepts in section 2 and, as emphasized by Meddahi and Renault (1996, 2002), asset pricing relationships are based on conditional expectations as opposed to linear projections. Thus, even though the difference between the linear projections and the true conditional expectations may be numerically small in empirical realistic situations, this limits the applicability and the formal interpretation of the Weak GARCH class of models. The discrete-time Square-Root Stochastic Autoregressive Volatility (SR-SARV) models provide an alternative formulation which circumvent these problems. We next turn to a discussion of this and other discrete-time stochastic volatility models.

3.1.2. Stochastic Volatility Models

The stochastic volatility models differ from the ARCH class of models in that the information set, $\mathscr{F}_{t-h}$, underlying the conditional expectations in equation (3.1) and (3.2) is not directly measurable with respect to the time $t-h$ observable filtration. This is typically the result of the inclusion of two separate stochastic innovations - one innovation term relating the conditional mean of the process to the actually observed return, a second innovation relating the *latent* volatility process to its conditional mean. This type of formulation is typically motivated by the Mixture-of-Distributions Hypothesis (MDH) and the idea of a latent information arrival process. The MDH was originally put forth by Clark (1973) as a way of conceptualizing the distributional characteristics of speculative returns, and the basic hypothesis has subsequently been extended and analyzed empirically by Epps and Epps (1976), Taylor (1982), Tauchen and Pitts (1983), Andersen (1996), Andersen and Bollerslev(1997), Ané and Geman (2000), among many others, to allow for more realistic temporal dependencies in the underlying latent information arrival process(es). We shall return to a discussion of these ideas in section 4.2 below. The actual parameterizations of the most popular discrete-time stochastic volatility models are often rationalized through the discretization of specific continuous-time stochastic volatility models. We do not provide an exhaustive review of the pertinent discrete-time stochastic volatility class of models here, but simply refer to the excellent surveys offered in Taylor (1994), Shephard (1996), and Ghysels, Harvey and Renault (1996).

In parallel to the GARCH and EGARCH class of models discussed above, most of the parametric stochastic volatility models employed in the literature are based on an autoregressive formulation for a continuous function of the (now) latent volatility process,

$$f[\vartheta^2(t,h)] \;=\; \omega + \beta(L,h)f[\vartheta^2(t,h)] + u(t,h), \tag{3.15}$$

where $\beta(L,h)$ denotes a *p*th order distributed lag polynomial, and $u(t,h)$ is a martingale difference sequence; i.e., $E[u(t,h)|\mathcal{F}_{t-h}]=0$. This class of models is commonly referred to as a Stochastic AutoRegresive Volatility, or SARV(p), model. Intuitively, it is the innovation term, $u(t,h)$, which distinguishes the stochastic volatility from the ARCH class of models.[10] Of course, analogous to the GARCH class of models discussed above, for the SARV(p) model in (3.15) to be well defined $\vartheta^2(t,h)$ must be positive (a.s.). Depending on the functional form for $f(\cdot)$, this restricts the admissible parameters in $\beta(L,h)$ and/or the support of $u(t,h)$. Most of the models estimated in the literature have included only a single lag in the $\beta(L,h)$ polynomial. Conditions to ensure ergodicity and stationarity for the general SARV(1) model are presented in Andersen (1994). The two leading cases are given by the lognormal stochastic autoregressive volatility model in which $f(x) \equiv log(x)$ and $u(t,h)$ is assumed to be Gaussian, and the square-root[11], or SR-SARV, model corresponding to $f(x) \equiv x$.

The lognormal stochastic volatility models was first analyzed by Taylor (1982) and subsequently popularized in influential papers by Harvey, Ruiz and Shephard (1994) and Jacquier, Polson and Rossi (1994). The logarithmic volatility model arises naturally from the standard return formulation, $r(t,h) = \mu(t,h) + \vartheta(t,h)\cdot z(t,h)$, in which $z(t,h)$ is an i.i.d. mean zero, unit variance, white noise process. Rearranging the terms, squaring both sides, and taking logarithms, it follows that

$$y(t,h) \;\equiv\; log[r(t,h) - \mu(t,h)]^2 \;=\; log[\vartheta^2(t,h)] \;+\; log[z(t,h)^2]. \tag{3.16}$$

Assuming the mean to be known, this may be interpreted as the measurement equation in a state space representation of the model, with corresponding transition equation defined by the parametric model for $log[\vartheta^2(t,h)]$. In particular, for the lognormal SARV(1) model,

$$log[\vartheta^2(t,h)] \;=\; \omega + \beta \cdot log[\vartheta^2(t-h,h)] + u(t,h). \tag{3.17}$$

In this situation, filtered and smoothed measurements of the latent $log[\vartheta^2(t,h)]$ volatility process are readily available by linear Kalman filtering which, as pointed out by Nelson (1988) and Harvey, Ruiz and Shephard (1994), in turn allows for relatively easy to compute Gaussian QMLE parameter estimates. Of course, the innovations in the measurement equation will generally not be Gaussian, so the Kalman may result in poor measurements of the latent

---

[10] Formal conditions under which the $u(t,h)$ term in the autoregressive formulation cannot be integrated out of the conditional expectations in (3.1) and (3.2), resulting in a genuine stochastic volatility model, are presented in Andersen (1992).

[11] This terminology derives from Andersen (1994) who parameterizes an AR(1) model for $f^{-1}[\vartheta(t,h)]$. Similarly, the lognormal stochastic volatility model is sometimes referred to as an Exponential SARV model.

volatility state variable and correspondingly highly inefficient parameter estimates.[12]

The lognormal SARV(1) formulation may also be justified as a discrete-time approximation to the popular Ornstein-Uhlenbeck diffusion for the logarithmic instantaneous volatility,

$$dlog(\sigma^2(t)) \ = - \beta \ (log(\sigma^2(t)) - \alpha) \ dt \ + \ \psi \ dV(t) \ , \tag{3.18}$$

where *V(t)* denotes a Standard Brownian Motion, or Wiener Process. We will return to a more detailed discussion of this and other continuous-time parametric stochastic volatility models in section 3.2.1 below. For now note that by a standard Euler scheme, the discrete-time version of the model in (3.18) takes the form

$$log[\vartheta^2(t,h)] \ = \ log[\vartheta^2(t-h,h)] - h \cdot \beta \cdot \{log[\vartheta^2(t-h,h) - \alpha] + h^{1/2} \cdot \psi \cdot [V(t,h) - V(t-h,h)].$$

The actual parameterization is, of course, different from the model in equation (3.17), but the structure corresponds exactly to that of the lognormal SARV(1) model. Interestingly, the continuous-time Ornstein-Uhlenbeck process in equation (3.18) also has the interpretation of being the diffusion limit of the discrete-time EGARCH model, in the sense that a sequence of appropriately parameterized EGARCH(1,1) models (as discussed in section 4.1 below) converges weakly to this model as the length of the sampling interval, *h*, approaches zero.

Although the latent logarithmic volatility in (3.17) takes the form of an AR(1) model, this translates into an ARMA(1,1) correlation structure for the demeaned logarithmic returns in (3.16), *y(t,h)*. Moreover, following Taylor (1986) and Harvey (1998) this same approximate correlation structure is present for any positive power transform of the squared returns; i.e., $exp[y(t,h)]^c$ for *c>0*. Hence, the shape of the autocorrelogram for the squared returns from the log-normal SARV(1) model mimics that of the empirically popular GARCH(1,1) model.

The second leading class of stochastic volatility models is given by the SR-SARV(p) model. Following Meddahi and Renault (2002), the SR-SARV(p) model for $\vartheta^2(t,h)$ is naturally defined by the marginalization of a *p*-dimensional latent VAR(1) process. As emphasized by Meddahi and Renault (1996, 2002), this class of models has the advantage of being closed under temporal (and in the multivariate setting cross sectional) aggregation. To appreciate this result, suppose that the true underlying continuous-time volatility is determined by the Constant Elasticity of Variance (CEV) diffusion (see section 3.2.1 below for further details),

$$d\sigma^2(t) \ = (\omega - \theta \cdot \sigma^2(t))dt \ + \sqrt{2} \cdot \alpha \cdot (\sigma^2(t))^\delta \ dV(t) \ , \tag{3.19}$$

where $\delta \geq \frac{1}{2}$ in order to ensure that the process for $\sigma^2(t)$ is stationary and non-negative. By standard arguments the *exact* discretization of the process then adheres to the basic SR-SARV(1) model structure,

---

$$\vartheta^2(t,h) \;=\; \omega + \beta \cdot \vartheta^2(t-h,h) + u(t,h) \,, \qquad\qquad\qquad (3.20)$$

where $E[u(t,h)|\mathscr{F}_{t-h}]=0$. Of course, the h-period time interval is arbitrary, so that the expected volatility for the temporally aggregated process, $\vartheta^2(t{\cdot}k,h{\cdot}k)$, where $k>1$ and $t = 0, 1, 2 \ldots$, must be governed the same AR(1) model structure. As discussed further in section 4.1 below, the CEV model in (3.19) with $\delta=1$ may also be interpreted as the diffusion limit of the GARCH(1,1) model.

The discrete-time AR(1) formulations in (3.17) and (3.20) are, of course, somewhat restrictive. In parallel to the developments within the parametric ARCH class of models discussed above, long-memory, or fractionally integrated stochastic volatility models, better suited at capturing the apparent long-run dependencies in the volatility, have recently been estimated by Breidt, Crato and de Lima (1998) and Harvey (1998). In a related context, Comte and Renault (1998) have proposed a continuous-time stochastic volatility model with discrete-time long-memory implications.

Direct extensions of the univariate discrete-time stochastic volatility models discussed above to a multivariate setting was first explored by King, Sentana and Wadhwani (1994), while earlier work by Diebold and Nerlove (1989) employed a related univariate latent ARCH factor structure in parameterizing time-varying conditional covariances. Recent proposals to allow for more flexible inference and the estimation of large-dimensional systems include Jacquier, Polson and Rossi (1999) and Chib, Nardari and Shephard (1999).

A more detailed discussion of statistical inference procedures for the stochastic volatility class of models (both univariate and multivariate) is provided by other chapters in the *Handbook* as well as the aforementioned surveys by Ghysels, Harvey and Renault (1996) and Shephard (1996). Intuitively, because of the latent information structure, any inference procedure must either rely on a (potentially noisy) proxy for the latent volatility or integrate out the latent stochastic variable(s) from the model. In addition to the method of moments, GMM, and SMM estimators implemented by Taylor (1986), Melino and Turnbull (1990), and Duffie and Singleton (1993), respectively, and the Gaussian QMLE procedures in Harvey, Ruiz and Shephard (1994) briefly discussed above, other noteworthy developments include the Bayesian Markov Chain Monte Carlo (MCMC) method in Jacquier, Polson and Rossi (1994) and Kim, Shephard and Chib (1998), the Efficient Method of Moments (EMM) procedure implemented by Gallant, Hsieh and Tauchen (1997), the simulated maximum likelihood technique of Danielsson (1994), the Monte Carlo maximum likelihood approach of Sandmann and Koopman (1998), the direct maximum likelihood estimation through recursive numerical integration in Fridman and Harris (1998), and the recent eigenfunction expansion of the latent volatility process in Meddahi (2001). Importantly, from the perspective of volatility measurements, many of these procedures in turn result in (approximately optimal) filtered and/or smoothed measurements of the functional latent volatility process, $f[\vartheta^2(t,h)]$, conditional on the underlying parametric model and the observable information.

Optimal measurements may in theory be obtained by non-linear filtering and smoothing

procedures (e.g., Kitagawa, 1987). However, the direct implementation of such procedures involves high-dimensional integration which is generally prohibitively expensive from a computational point of view (see e.g., the discussion in Shephard, 1996). Recent advances along these lines to allow for the practical numerical calculation and extraction of latent volatility measurements include the particle filters in Pitt and Shephard (1999) and the reprojection approach advocated by Gallant and Tauchen (1998).[13] We next turn to a discussion of analogous volatility modeling and measurement procedures based on continuous-time parametric formulations.

3.2. Continuous-Time Models

Much of the theoretical asset pricing literature is cast in continuous time. Within this tradition, the sample path of the price process is also commonly assumed to be continuous. This approach is convenient because the representation in Proposition 1 then ensures that - locally - the mean and variance are of the same order. Consequently, the framework effectively involves a dynamic mean-variance trade-off, which typically allows for a tractable analysis of asset pricing and portfolio choice problems. On the other hand, we usually do not observe a record of continuously evolving asset prices, and all but the very simplest specifications tend to imply intractable conditional return distributions for the corresponding discretely observed returns. This issue has historically inhibited empirical work on estimation and inference for realistic continuous-time asset price processes, although a burst of research activity in this area over the last few years has allowed important headway to be made. As a result, the parametric approach to continuous-time modeling is beginning to have a practical impact on return volatility modeling. We will not discuss estimation and inference techniques for this class of model in any detail, however, but rather outline the conceptual issues that distinguish this approach from the discrete-time modeling approached discussed above and the nonparametric volatility measurement discussed subsequently. Other chapters in this *Handbook* offer extensive coverage of parametric and semi-(non)parametric estimation techniques for diffusion processes (e.g., Aït-Sahalia, Hansen and Scheinkman, 2002; Bandi and Phillips, 2002; Bibby, Jacobsen and Sørensen 2002; Gallant and Tauchen, 2002; and Johannes and Polson, 2002)

The continuous-time parametric models are directly compatible with the no-arbitrage framework outlined in section 2, so the specific volatility concepts carry over without modification. However, the specifications of the models traditionally adapted in the literature differ from the general semi-martingale representation in section 2, as the models typically are expressed (in short-hand format) as stochastic differential equations driven by underlying Brownian motions and - in the case of discontinuities - Poisson jump processes.

Before exploring the manifestation of the different volatility concepts for some of the more

---

[13] In a related context, Alizadeh, Brandt and Diebold (2002) and Gallant, Hsu and Tauchen (1999) have recently proposed the use of intraperiod high-low, or range, data for more efficiently extracting discrete-time stochastic volatility measures.

popular model specifications, it is instructive to briefly reiterate the relationship between the concepts of expected notional volatility and expected volatility within a multi-period setting. Of course, for a continuous-time model, any volatility forecast over a discrete time interval invariably entails multi-period considerations (typically a continuum). Calculations similar to those presented in equation (3.10) (for details see Andersen, Bollerslev, Diebold and Labys, 2001b), show that the expected volatility may generally be expressed as

$$
\begin{aligned}
\vartheta^2(t,h) \;=\; & E[\,(r(t,h) - m(t,h))^2\,|\,\mathscr{F}_{t-h}\,] \\[6pt]
\;=\; & E[\upsilon^2(t,h)|\mathscr{F}_{t-h}] \;+\; Var[\,\mu(t,h)|\mathscr{F}_{t-h}] \;+\; 2\cdot Cov[M(t,h),\,\mu(t,h)|\mathscr{F}_{t-h}].
\end{aligned}
\tag{3.21}
$$

Again, the total expected return volatility involves the expected notional volatility (quadratic variation) as well as two terms induced by future within-forecast-period variation in the conditional mean. The intuition is identical to the one for the discrete-time setting discussed in section 3.1. The random variation in the mean component is a direct source of future return variation, and covariation between the return and conditional mean innovations will further impact the return variability. However, under standard conditions and moderate forecast horizons, the dominant factor is indisputably the expected notional volatility, as the innovations to the mean return process generally will be very small relative to the cumulative return innovations. Importantly, this does not rule out asymmetric effects from current return innovations to future return volatility, as the leverage and volatility feedback effects operate, respectively, exclusively or primarily through the impact on the notional volatility process.

3.2.1. Continuous Sample Path Diffusions

The continuous-time models in the theoretical asset and derivatives pricing literature frequently assume that the sample paths are continuous, with the corresponding diffusion processes given in the form of stochastic differential equations (as in Example 1 in section 2 above), rather than through (abstract) integral representations for continuous sample path semi-martingales along the lines of Proposition 1. This does not involve any loss of generality, as illustrated by the following well-known result (e.g., Protter, 1992; and Karatzas and Shreve, 1991).

PROPOSITION 3 - Martingale Representation Theorem

For any univariate, square-integrable, continuous sample path logarithmic price process, which is not locally riskless, there exists a representation such that for all $0 \le t \le T$, a.s.(P),

$$
r(t,h) \;=\; \mu(t,h) + M(t,h) \;=\; \int_0^h \mu(t-h+s)\,ds \;+\; \int_0^h \sigma(t-h+s)\,dW(s),
\tag{3.22}
$$

where $\mu(s)$ is an integrable, predictable and finite variation stochastic process, $\sigma(s)$ is a strictly positive càglàd stochastic process satisfying

$$P[\int_0^h \sigma^2(t\text{-}h\text{+}s)\,ds < \infty] \;=\; 1,$$

and *W(s)* is a standard Brownian motion.

The integral representation (3.22) is equivalent to the standard (short-hand) sde specification for the logarithmic price process,

$$dp(t) \;=\; \mu(t)\,dt \;+\; \sigma(t)\,dW(t), \qquad 0 \le t \le T. \tag{3.23}$$

Hence, within the class of continuous sample path semi-martingale (diffusion) models, there are no consequential restrictions involved in stating the model directly through a stochastic differential equation. In accordance with Definition 3 in section 2, the volatility coefficient process in this formulation, $\{\sigma_t^2\}_{t \in [0,T]}$, is usually termed the *Instantaneous Volatility Process*. We have the following direct link between these alternative volatility representations,

$$\sigma^2(t) \;=\; lim_{h \to 0}\,\sigma^2(t\text{-}h) \;=\; lim_{h \to 0}\,(\int_0^h \sigma^2(t\text{-}h\text{+}s)\,ds\,/\,h\,). \tag{3.24}$$

It is also immediately evident from Proposition 3 that in this situation the notional volatility (or the increment to the quadratic variation process) equals the so-called *Integrated Volatility*,

$$\upsilon^2(t,h) \;=\; [M,M]_t - [M,M]_{t\text{-}h} \;=\; \int_0^h \sigma^2(t\text{-}h\text{+}s)\,ds. \tag{3.25}$$

The (expected) integrated volatility plays a key role in the stochastic volatility option pricing literature. In particular, following Hull and White (1987) and ignoring the return variation associated with the conditional mean, options prices only depend on the expected notional volatility (see also Garcia, Ghysels and Renault, 2002; and Garcia, Lewis and Renault, 2001).

The number of alternative continuous-time specifications for asset returns employed in the literature is much too large for a comprehensive review at this point. For illustrative purposes, we simply consider the relevant volatility concepts implied by a few standard formulations.

The simplest possible case is provided by the time-invariant diffusion,

$$dp(t) \;=\; \mu\,dt \;+\; \sigma\,dW(t), \qquad 0 \le t \le T, \tag{3.26}$$

which underlies the Black-Scholes option pricing formula. Obviously, this process has a deterministic mean return so the expected return volatility trivially equals the expected notional volatility. Moreover, because the volatility is also constant, the expected notional volatility is identical to the notional volatility. Formally, we thus have for the Black-Scholes setting,

$$\vartheta^2(t,h) \;=\; E[\,(r(t,h) - m(t,h))^2 \,|\, \mathscr{F}_{t\text{-}h}\,] \;=\; \upsilon^2(t,h) \;=\; \int_0^h \sigma^2(t\text{-}h\text{+}s)\,ds \;=\; \sigma^2 \cdot h.$$

As discussed further in section 4.1 below, this model is also straightforward to estimate from discretely sampled data by, e.g., maximum likelihood, as the returns are *i.i.d.* and normally

distributed. Of course, the model is overwhelmingly rejected for moderately frequently sampled data (say, daily, weekly, or monthly), as it fails to accommodate the well-documented strong intertemporal volatility dependencies.

For some price series (notably real commodity prices and exchange rates) it is often sensible to postulate a stationary logarithmic price process. Popular models for such series - inspired by the interest rate literature - include the Ornstein-Uhlenbeck (OU) processes and the square-root, or Cox, Ingersoll and Ross (1985) (CIR), processes. These models take the general form

$$ dp(t) \;=\; \phi\,(\,\mu - p(t)\,)\,dt \;+\; \sigma(t)\,dW(t), \qquad 0 \le t \le T. \tag{3.27} $$

The drift specification ensures mean reversion in the process, given appropriate regularity conditions and a well-behaved diffusion (volatility) coefficient process. Letting $\sigma(s) \equiv \sigma$ results in the standard OU model, while having $\sigma(s) \equiv \sigma\,p^{\gamma}(s)$ produces a constant elasticity of variance (CEV) model, with the CIR model as a special case for $\gamma = \tfrac{1}{2}$. The CEV class of models was first proposed in the asset pricing literature by Cox and Ross (1976), and further popularized for interest rates by Chan, Karolyi, Longstaff and Sanders (1992). The attraction of the specific OU and CIR formulations stems primarily from the tractable distributions for discretely observed data, and the accompanying closed-form solutions for many related asset and derivatives pricing problems. Explicit solutions for the expected volatility and the expected notional volatility may be derived from existing results in the literature. One immediate observation is that these two volatility concepts now differ, as the return innovations will impact the mean process randomly over the forecast horizon. Nonetheless, the expected notional volatility will remain the dominant component in empirically realistic situations.

To exemplify the points above, consider the simple OU process which allows for explicit expressions for the different expected volatility components over the interval, *[t-h,t]*. For simplicity, and without loss of generality, we let *t-h=0*, corresponding to the interval *[0,h]* with return $r(h,h) \equiv r(h)$ and associated expected volatility $\vartheta^2(h,h) = E[\,(r(h) - m(h))^2\,|\,\mathscr{F}_0\,]$. The explicit solution to the OU stochastic differential equation may be written directly as

$$ r(h) \;=\; \mu(h) + M(h) $$

$$ =\; \{\,p(0)\,(exp(-\phi h)-1) + \sigma \int_0^h [\,exp(-\phi(h-s)) - 1\,]\,dW(s)\,\} + \sigma \int_0^h dW(s). \tag{3.28} $$

The martingale component is, of course, simply $M(h) = \sigma \int_0^h dW(s) = \sigma \cdot W(h)$, so the notional volatility equals $\upsilon^2(h,h) = \sigma^2 \cdot h$. The mean component accounts for the two first terms on the right hand side of (3.28). Notice that the drift coefficient, as required, only depends on the martingale innovation process through a weighted average of past realizations where, at time *s=h*, the current realization of *W(h)* receives a zero weight from $[\,exp(-\phi(h-s)) - 1\,]$. This representation reflects the negative correlation between the return innovations and the drift, as a (positive) shock to returns decreases the future drift coefficient via a mean-reverting term.

Given the (conditional) normality of the OU process, it is straightforward to compute the

various components of the expected volatility explicitly. We present the closed-form expressions along with simpler approximations that apply for a short horizon, h. First, for the overall expected volatility,

$$\vartheta^2(h,h) = [(1 - \exp(-2\phi h))/(2\phi)]\, \sigma^2 \approx \sigma^2 \cdot h - \phi h^2 \sigma^2 + (2/3) \cdot \sigma^2 \phi^2 h^3$$

$$\approx E[\upsilon^2(h,h)|\mathcal{F}_0] - \phi h^2 \sigma^2 + (2/3) \cdot \sigma^2 \phi^2 h^3.$$

(3.29)

The expected volatility is thus locally smaller than the notional volatility. This occurs because of the mean-reverting drift coefficient. Large return innovations will tend to be partially undone over the forecast horizon. However, to first order in h, expected volatility equals notional volatility, confirming the crucial role of expected notional volatility.[14]

Further, in reference to equation (3.21), straightforward calculations reveal that

$$Var[\mu(h)|\mathcal{F}_0] = [h + \{1 - \exp(-2\phi h)/(2\phi)\} - 2\{1 - \exp(-\phi h)/\phi\}] \cdot \sigma^2$$

$$\approx (\phi^2/3)\, h^3 \sigma^2,$$

and

$$Cov[M(h), \mu(h)|\mathcal{F}_0] = [\{1 - \exp(-\phi h)/\phi\} - h] \cdot \sigma^2 \approx -(\phi/2)\, h^2 \sigma^2 + (\phi^2/6)\, h^3 \sigma^2,$$

where the approximating expressions are computed to order h^3 to avoid the variation of the mean process from vanishing. Obviously, the contribution of the variation in the drift process is generally (locally) negligible. Hence, the main contribution to the expected volatility (beyond the notional volatility) stems from the covariance between return innovations and the future path of the mean process. The negative correlation between these components lowers the overall expected volatility (albeit the effect typically is small). Although the calculations are specific to the OU process, the orders of magnitude are indicative of the relative importance of the components governing the expected volatility. In fact, the OU process displays a very strong covariance between the return innovations and the expected returns process, suggesting that this example, if anything, overstates the typical contribution of the terms beyond the (expected) notional volatility in determining the expected volatility for many asset classes.

Unfortunately, the entire class of one-factor models covered by equation (3.27) falter dramatically when confronted with actual price or return data. In order to obtain more satisfactory empirical fits, the literature has moved towards multi-factor parametric formulations. A natural approach is to let the volatility process be governed by an independent

---

[14] A simple numerical example illustrates the orders of magnitude. The OU process is typically estimated, or calibrated, to capture slowly evolving long-run swings in the logarithmic price process (or interest rate) away from the unconditional mean. Such movements induce a relatively small degree of predictability in the short-term asset returns, but long-term mean reversion, as manifest by a small mean reversion parameter for data calibrated to an annual frequency, say $\phi=0.1$. At the daily frequency, or $h=1/250$, clearly $\phi \cdot h^2 \approx 0$, so that the difference between the expected volatility and the notional volatility is negligible. Even at the quarterly frequency, or $h=1/4$, the deviation is a modest *2.5%.* Of course, this number is somewhat sensitive to the assumed strength of the mean reversion.

source of random variation, leading to a (genuine) continuous-time stochastic volatility model. An influential specification is given by the square-root volatility model popularized by Heston (1993) corresponding to $\delta=\frac{1}{2}$ in the CEV diffusion,

$$d\sigma^2(t) \; = \; (\omega - \theta\,\sigma^2(t))\,dt \; + \; \eta\,(\sigma^2(t))^{\frac{1}{2}}\,dV(t), \qquad 0 \leq t \leq T, \tag{3.30}$$

where the standard Brownian motion process, $V(t)$, may be correlated with the $W(t)$ process driving the returns, thus introducing an asymmetric return-volatility relation into the asset price dynamics. This model is particularly attractive as it allows for closed-form solutions for option prices. An extensive analysis of multivariate square-root (or affine) processes in modeling term-structure dynamics is provided in Dai and Singleton (2000) (see also, Piazzesi, 2002 in this *Handbook*).

An alternative popular choice is to represent the logarithmic volatility process by the OU diffusion in equation (3.18). As discussed in section 3.1.2 above, this formulation corresponds to an (approximate) discrete-time SARV(1) model. In either case, the relation between the expected volatility and the expected notional volatility may be found from the general formula in equation (3.21). If the two Wiener innovation processes are correlated, all three terms become operative, although the expected notional volatility (expected quadratic variation) continues to dominate empirically.

The stochastic volatility diffusions above are considerably harder to estimate from discretely observed data than the classical one-factor models of the OU or CIR variety. However, recent progress has made relatively efficient inference possible through simulation-based procedures such as efficient method of moments (EMM) or Markov chain Monte Carlo (MCMC). The evidence indicates that these models provide major improvements over the traditional one-factor models although they continue to be decidedly rejected by both equity return and nominal interest rate data (e.g., Gallant and Tauchen, 1997; Andersen and Lund, 1997; Andersen, Benzoni and Lund, 2002; and Eraker, Johannes and Polson, 2002).

These failures have recently prompted a number of authors to add additional parametrically specified diffusion factors (e.g., Chernov, Gallant, Ghysels and Tauchen, 2001; and Andersen and Lund, 1998). In light of the general representation in equation (3.22), it is evident that such multi-factor models simply provide an alternative way of specifying the return dynamics that ultimately may be reduced to a single factor representation for the univariate process. The advantage is that the system may be defined through a sum of different factors, each following a simple dynamic process, rather than a single factor with a more complex specification. For example, one may approximate (apparent) long-range dependencies in the volatility process through a sum of multiple distinct AR(1) factors (e.g., Gallant, Hsu and Tauchen, 1999).[15]

---

[15] Alternatively, as shown by Chen, Hansen and Carrasco (2000) long-memory dependencies may also be represented through (appropriate) nonlinearities in the scalar diffusion process. This idea also underlies the eigenfunction stochastic volatility class of models recently developed by Meddahi (2001).

The ability to produce a simple parametric representation is extremely convenient, if not critical, for economic interpretation and implementation of tractable estimation strategies through standard (simulation based) likelihood and method of moments techniques. To illustrate this idea, consider the general *k*-factor model,

$$r(t) \equiv p(t) - p(0) = \int_0^h \mu(t-h+s)\,ds + \sum_{j=1,...,k} \int_0^h \sigma_j(t-h+s)\,dW_j(s),$$

where the $\sigma_j(t)$ refer to the jth volatility factor and $W(t) = (W_1(t), ..., W_k(t))$, denotes a *k*-dimensional vector process of independent standard Brownian motions. The notional volatility then follows straightforwardly as the sum of the integrated constituent components,

$$\upsilon^2(t,h) = \int_0^h \sigma^2(t-h+s)\,ds = \int_0^h \{\sum_{j=1,...,k} \sigma_j^2(t-h+s)\}\,ds = \sum_{j=1,...,k} \int_0^h \sigma_j^2(t-h+s)\,ds.$$

As such, none of the general principles change, but the requisite calculations for, say, the different terms in equation (3.21) may certainly become more involved.

It is arguably premature to judge the empirical performance of the parametric multi-factor continuous sample path (pure diffusion) volatility models for asset returns, as this work truly is in its infancy. It is clear, nonetheless, that such models serve as alternatives as well as complements for the parametric jump-diffusion models that we turn to next.

3.2.2. Jump Diffusions and Lévy Driven Processes

At the highest sampling frequencies, there is compelling evidence of the existence of jumps in asset price processes. Specifically, the arrival of important news such as macroeconomic announcements (at the aggregate level) or earnings reports (at the firm level) typically induce a discrete jump associated with an immediate revaluation of the asset; see, e.g., Andersen and Bollerslev (1998b) and Andersen, Bollerslev, Diebold and Vega (2002) for direct parametric modeling of such jumps or Johannes (2000) for nonparametric specification tests for the existence of jumps. Likewise, much evidence from the implied volatility literature - which extracts information about market expectations concerning the future return distribution directly from option prices - points towards the importance of incorporating discrete jump probabilities into the analysis of the return dynamics; see, e.g., Bates (1996a) and Bakshi, Cao and Chen (1997).

In the same way that the Brownian motion constitutes the basic building block of continuous-time martingales, the standard Poisson jump process serves as the basic building block for pure (compensated) jump martingales (e.g., Merton, 1982). Thus, one may accommodate the relevant jump features in an arbitrage-free continuous-time logarithmic price process by adding a Poisson jump component with appropriate time variation in the jump intensity and/or the jump distribution (as in Example 1 in section 2 above). In line with this reasoning, let *q(t)* denote a Poisson process, with *q(t)=1* indicating a jump at time *t*, and *q(t)=0* otherwise, and

(possibly time-varying) jump intensity denoted $\lambda(t)$.[16] With the random jump size denoted $\kappa(t)$, we then have the general representation

$$
\begin{aligned}
r(t,h) &= \mu(t,h) + M(t,h) \\
&= \int_0^h \mu(t-h+s)\,ds + \int_0^h \sigma(t-h+s)\,dW(s) + \Sigma_{t-h \le s \le t}\,\kappa(s)\cdot q(s).
\end{aligned}
\tag{3.31}
$$

The associated notional volatility process explicitly incorporates the jumps,

$$
\begin{aligned}
\upsilon^2(t,h) &\equiv [M,M]_t - [M,M]_{t-h} = [M^c,M^c]_t - [M^c,M^c]_{t-h} + \Sigma_{t-h \le s \le t}\,\Delta M^2(s) \\
&= \int_0^h \sigma^2(t-h+s)\,ds + \Sigma_{t-h \le s \le t}\,\kappa^2(s)\cdot q(s).
\end{aligned}
\tag{3.32}
$$

The computation of the corresponding expressions for the expected notional volatility and the expected volatility will depend on the specific parametric formulation.

To illustrate, consider the simple jump diffusion in Merton (1976) with constant mean and diffusion volatility coefficients as well as *i.i.d.* jumps; i.e., $\sigma(t) \equiv \sigma$, $\lambda(t) \equiv \lambda$. Also, denote the mean and the variance of the jump distribution by μ_κ and σ_κ^2, respectively. The notional volatility is now a stochastic variable, reflecting the random occurrence of jumps. Moreover, only if the mean jump size is zero, $\mu_\kappa = 0$, will the expected notional volatility coincide with the expected volatility (expected volatility does not contain the squared mean term below),

$$
E[\upsilon^2(t,h)|\mathscr{F}_{t-h}] = \sigma^2 \cdot h + E[\Sigma_{t-h \le s \le t}\,\Delta M^2(s)|\mathscr{F}_{t-h}] = \sigma^2 \cdot h + \lambda \cdot h \cdot (\mu_\kappa^2 + \sigma_\kappa^2).
$$

The corresponding expected instantaneous volatility, defined in equation (2.16), is

$$
lim_{h \to 0}\,\upsilon^2(t,h)/h = \sigma^2 + \lambda \cdot (\mu_\kappa^2 + \sigma_\kappa^2),
$$

which differs from the instantaneous volatility associated with the continuous martingale component, as defined in equation (2.17), and here restricted to be constant, $\sigma^2(t) \equiv \sigma^2$.

In this context, it is also important to recognize that the (implicit) càdlàg assumption on the $\sigma(t)$ process in equation (3.31), and the corresponding continuous sample path martingale representation in Proposition 3, does not rule out jumps in the $\sigma(t)$ process. Notice, however, that the presence of jumps in either $\sigma(t)$ and/or $r(t,h)$ invalidates the consistency arguments underlying the nonparametric ARCH filters and smoothers, but not for the so-called realized volatility measures, as discussed further in the next section.

As for the multi-factor parametric diffusion representations, the empirical evidence on jump

---

[16] Formally, $P[q(s)=0,\ t-h<s \le t] = 1 - \int_0^h \lambda(t-h+s)ds + o(h)$, $P[q(s)=1,\ t-h<s \le t] = \int_0^h \lambda(t-h+s)ds + o(h)$, and $P[q(s) \ge 2,\ t-h<s \le t] = o(h)$.

diffusion models for asset returns is still inconclusive. Early work estimated (overly) simple representations in line with the time-invariant jump diffusion discussed above (e.g., Press, 1967; Jarrow and Rosenfeld, 1984; Ball and Torous, 1985; and Akgiray and Booth, 1986), but these models are clearly at odds with the data. More realistic models have recently been explored by, e.g., Andersen, Benzoni and Lund (2002), Duffie, Pan and Singleton (2000), Eraker, Johannes Polson (2002), Pan (2001), and Eraker (2001). While these formulations improve dramatically on the fit of traditional univariate diffusion and standard stochastic volatility representations, a general consensus about the relative performance of the various alternative specifications remains elusive at this (early) point.

Another recent proposal is to retain the continuous sample path strategy for the asset returns, but model the volatility process as a non-Gaussian OU process driven by pure upward Lévy jumps (e.g., Barndorff-Nielsen and Shephard, 2001a). A primary motivation for this approach is to retain analytic tractability of the temporal aggregation process involved in the construction of volatility forecasts within a reasonably descriptive continuous-time setting.[17] Technically, the jumps in the volatility process introduces no new conceptual theoretical issues, as the Lévy processes are semi-martingales, and as such the general apparatus for diffusion processes discussed above applies directly. The empirical implementation of this approach is just beginning, but the results reported so far are intriguing (see also the *Handbook* chapter by Nguyen-Ngoc and Yor, 2002, for related theoretical developments).

4. NONPARAMETRIC METHODS

The data-driven, or nonparametric volatility measurements afford direct ex-post empirical appraisals of the notional volatility, $\upsilon^2(t,h)$, without any specific functional form assumptions. The most obvious such measure is, of course, given by the ex-post squared return spanning the *[t-h,t]* time interval. However, even though the (demeaned) squared return generally provides an unbiased estimator for $\upsilon^2(t,h)$, it is also a very noisy estimate. The nonparametric measurements more generally achieve consistency by measuring the volatility as (weighted) sample averages of increasingly finer sampled squared (or absolute) returns over (and possible outside) the *[t-h,t]* interval. This immediately raises important issues of efficiency, rates of convergence, and the (asymptotic) distributions for the measurement errors associated with different weighting schemes. At a more fundamental level, however, the nonparametric procedures differ importantly in their assumptions about the length of the time interval, *h*. The instantaneous volatility filters, or ARCH filters and smoothers, discussed next, are based on the assumption of ever more observations over ever finer time intervals (a double limit theory), while the realized volatility measures build on the idea of an increasing number of observations

---

[17] As noted by Barndorff-Nielsen and Shephard (2001a) this type of formulation also allows for the derivation of analytical option pricing formula under quite general conditions, as studied more thoroughly by Nicolato and Venardos (2001).

over fixed length time intervals (a single limit theory).[18]

4.1. ARCH Filters and Smoothers

Parametric ARCH models were designed to parsimoneously model the expected volatility as an explicit function of discretely observed returns; i.e., a parameterized conditional expectation, $\vartheta^2(t,h)=E[(r(t,h)-m(t,h))^2|F_{t-h}]$, where $h>0$ and F_{t-h} denotes the information set generated by the past returns $r(t-h,h), r(t-2h,h), ...$ However, as observed by Nelson (1992), these same discrete-time parametric models may alternatively be given a nonparametric interpretation as filters designed to extract information about the (latent) instantaneous volatility. In particular, assuming that the sample path of the price and the corresponding instantaneous volatility processes are both continuous, then - although formally misspecified at all discrete sampling frequencies, $h>0$ - an appropriately parameterized sequence of ARCH models, or expected (scaled) volatilities $\vartheta^2(t,h)/h$, will consistently (as $h\rightarrow\infty$) estimate the instantaneous volatility, $\sigma^2(t)$, at each point in time.

To grasp the intuition behind this powerful result, consider the simple continuous-time random walk model in equation (3.26), previously studied by Merton (1980) in this context,

$$dp(t) = \mu\,dt + \sigma\,dW(t), \qquad 0\le t\le T.$$

Suppose that observations are only available at $n+1$ equally spaced points over the $[t-h,t]$ time interval, where $0\le h<t\le T$; i.e., $t-h, t-h+(h/n), ... , t-h+(n-1)\cdot(h/n), t$. By the definition of the process, the corresponding sequence of $i = 1, 2, ..., n$, discrete (h/n)-period returns,

$$r(t-h+i\cdot(h/n), h/n) \equiv p(t-h+i\cdot(h/n)) - p(t-h+(i-1)\cdot(h/n)),$$

is then *i.i.d.* normally distributed with mean $\mu\cdot(h/n)$ and variance $\sigma^2\cdot(h/n)$. Hence, the MLE of the drift is simply given by the sample mean of the (scaled) returns,

$$\hat{\mu}_n \equiv n^{-1}\cdot\Sigma_{i=1,..,n}(h/n)^{-1}\cdot r(t-h+i\cdot(h/n), h/n) \equiv r(t,h)/h.$$

It follows immediately that

$$E(\hat{\mu}_n) = \mu.$$

This fixed-interval, or fill-in asymptotic, estimator for the drift only depends on h and not n. The sampling frequency is irrelevant, only the span of the data matters. Thus, although $\hat{\mu}_n$ is an unbiased estimator for μ, it is not consistent as $n\rightarrow\infty$.

---

[18] Of course, the discrete-time parametric ARCH and stochastic volatility models discussed in section 3.1 above may also be given the interpretation of fixed length interval filters for extracting $\vartheta^2(t,h)$, $h>0$. However, these type of filters are difficult to characterize and formally justify outside the realm of a specific parametric framework.

Consider now the (unadjusted) estimator for σ^2 defined by the sum of the (scaled) squared returns,

$$\hat{\sigma}_n^2 \equiv n^{-1} \cdot \Sigma_{i=1,..,n} (h/n)^{-1} \cdot r(t-h+i \cdot (h/n), h/n)^2 = h^{-1} \cdot \Sigma_{i=1,..,n} r(t-h+i \cdot (h/n), h/n)^2 .$$

Because

$$E[r(t-h+i \cdot (h/n), h/n)^2] = \sigma^2 \cdot (h/n) + \mu^2 \cdot (h/n)^2 ,$$

it follows readily that

$$E(\hat{\sigma}_n^2) = \sigma^2 + \mu^2 \cdot (h/n) .$$

Hence, the drift induces only a second order bias, or $O(n^{-1})$ term, in the estimation of σ^2 for $n \to \infty$. Moreover, this estimator for the diffusion coefficient is consistent as $n \to \infty$. To see this, note that

$$E[r(t-h+i \cdot (h/n), h/n)^3] = 3 \cdot \mu \cdot \sigma^2 \cdot (h/n)^2 + \mu^3 \cdot (h/n)^3 ,$$

$$E[r(t-h+i \cdot (h/n), h/n)^4] = 3 \cdot \sigma^4 \cdot (h/n)^2 + 6 \cdot \mu^2 \cdot \sigma^2 \cdot (h/n)^3 + \mu^4 \cdot (h/n)^4 ,$$

which along with the second moment given above, and the fact that the returns are *i.i.d.*, implies that

$$Var(\hat{\sigma}_n^2) = 2 \cdot \sigma^4 \cdot n^{-1} + 4 \cdot \mu^2 \cdot \sigma^2 \cdot n^{-2}.$$

Hence by a standard Law of Large Numbers,

$$plim_{n \to \infty} \hat{\sigma}_n^2 = \sigma^2.$$

The consistency result for the sample variance estimator for the time-invariant diffusion hinges on the true volatility being constant over *[t-h,t]*. Increasing the number of (scaled) squared return observations over the interval then produces an increasing number of unbiased and uncorrelated measures of σ^2, and simply averaging these yields a consistent estimator.

This basic idea may, given appropriate regularity conditions, be extended to the general class of continuous sample path diffusions considered in equation (3.23) above,

$$dp(t) = \mu(t) \, dt + \sigma(t) \, dW(t), \qquad 0 \leq t \leq T.$$

under the additional assumption that the sample path for the $\sigma(t)$ process also is continuous. The main difference between this general model and the time-invariant diffusion $\sigma(t) \equiv \sigma$ analyzed in detail above, is that the length of the sampling interval, *h*, now also must shrink to zero as the sampling intensity within the interval, *n*, increases.

At an intuitive level, by the assumed sample path continuity, the temporal variation in $\sigma^2(t)$ is readily bounded by restricting the length of the time interval, h, over which the variation is measured,

$$\forall \xi > 0,\ \exists h > 0 :\ \ sup_{t-h \leq \tau \leq t}\, |\sigma^2(\tau) - \sigma^2(t)| < \xi, \quad (a.s).$$

Using this result and refining the arguments above, it is possible to show that the analogous time t (unadjusted) sample variance estimator,

$$\hat{\sigma}^2_{n,h}(t) \ \equiv\ n^{-1} \cdot \Sigma_{i=1,..,n}\, (h/n)^{-1} \cdot r(t-h+i \cdot (h/n),\, h/n)^2, \tag{4.1}$$

consistently estimates the instantaneous volatility provided $h \to 0$ and $n \to \infty$ at the proper rates,

$$plim_{\,n \to \infty, h \to 0}\ \hat{\sigma}^2_{n,h}(t)\ =\ \sigma^2(t),$$

where the convergence is pointwise in probability.

The tradeoff between the length of the sampling interval, $h \to 0$, and the number of observations, $n \to \infty$, is analogous to the usual bias-variance tradeoff encountered in nonparametric kernel estimation. Similarly, the sample variance estimator in equation (4.1) corresponds to a flat kernel scheme and the efficiency of this estimator may generally be improved by using a weighted one- or two-sided average of squared returns. That is the motivation behind the ARCH filters and smoothers developed in a series of papers by Nelson (1992, 1996a,b), Nelson and Foster (1994, 1995) and Nelson and Schwartz (1992) (see also the discussion in Drost and Werker, 1996; Duan, 1997; Fornari and Mele, 2001; and Mele and Fornari, 2000).[19]

To illustrate, consider the GARCH(1,1) filter for the *(1/n)*-period returns defined in Nelson (1992),

$$\hat{\sigma}^2_n(t) \ =\ \omega_n\ +\ \alpha_n\, r(t,1/n)^2\ +\ \beta_n \cdot \hat{\sigma}^2_n(t\text{-}1/n)$$
$$=\ \omega_n \cdot (1-\beta_n)^{-1}\ +\ \Sigma_{i=0,..,\infty}\, \alpha_n \cdot \beta_n^i\, r(t-i/n, 1/n)^2, \tag{4.2}$$

where

$$\omega_n = \omega/n, \qquad \alpha_n = \alpha \cdot (1/n)^{\frac{1}{2}}, \qquad \beta_n = 1 - \alpha \cdot (1/n)^{\frac{1}{2}} - \theta/n,$$

and where $\omega>0$, $\alpha>0$, $\theta>0$, corresponding to $p=q=1$, $\vartheta^2(t,1/n) \equiv \hat{\sigma}^2_n(t)$, and $\mu(t,1/n) \equiv 0$ in equation (3.12) above. This filter again achieves consistency as $n \to \infty$ for $\sigma^2(t)$,

---

[19] For earlier work on nonparametric diffusion estimation based on stronger assumptions and different asymptotic arguments; see, e.g., Banon (1978), Dohnal (1987), Genon-Catalot, Laredo and Picard (1992), and Florens-Zmirou (1993).

$$plim_{n\to\infty}\ \hat{\sigma}_n^2(t)\ =\ \sigma^2(t),$$

and, as before, the convergence is pointwise in probability. Importantly, this result explicitly excludes any jumps, or discontinuities, in either the price or diffusion processes,[20] so that the sample path for the instantaneous volatility process, $\sigma^2(t)$, is continuous and coincides with that of the expected (scaled) instantaneous volatility, $lim_{h\to0}\upsilon^2(t,h)/h$.

Heuristically, the GARCH(1,1) filter works analogously to the sample variance estimator in equation (4.1) by utilizing an (infinite weighted) average of increasingly finer sampled squared returns ever closer to time t. However, the continuous record, or fill-in, asymptotics of ever more observations per interval, $n\to\infty$, over ever smaller time intervals, $h\to\infty$, is here achieved by a single asymptotic device dictating both the return sampling frequency and the simultaneous down-weighting of the more distant squared return observations, $r(t-i/n,1/n)^2$, for large values of i. Note that the parameter configuration in (4.2) underlying this result implies that

$$lim_{n\to\infty}(\ \alpha_n + \beta_n\ )\ =\ lim_{n\to\infty}(\ 1 - \theta/n\ )\ =\ 1,$$

so that the sequence of GARCH(1,1) filters approaches an IGARCH model, as discussed in section 3.1.1. above, in the limit.

Besides providing a consistent volatility filter, such sequences of GARCH models have other interesting and useful properties. For example, if the standardized returns,

$$z(t,1/n)\ \equiv\ n^{1/2}\cdot r(t,1/n)\ /\hat{\sigma}_n(t-1/n)\ ,\tag{4.3}$$

are *i.i.d.* normally distributed then, as shown by Nelson (1990a), the sequence of GARCH(1,1) models defined implicitly by (4.2) converges weakly to the following CEV diffusion process for the instantaneous volatility (Wong, 1964),

$$dp(t)\ =\ \sigma^2(t)\ dW(t),$$

$$d\sigma^2(t)\ =\ (\omega - \theta\cdot\sigma^2(t)\ )dt\ +\ \sqrt{2}\cdot\alpha\cdot\sigma^2(t)\ dV(t),\tag{4.4}$$

where the two Wiener processes are uncorrelated, $Corr(\ dW(t)\ ,\ dV(t)\ ) = 0$. Of course, it remains true, that when interpreted as a filter, the sequence of GARCH(1,1) models in (4.2) underlying this diffusion limit consistently extracts the instantaneous volatility, $\sigma^2(t)$, for *any* continuous sample path diffusion.

Many other appropriately parameterized ARCH models share this important property. Specifically, consider the sequence of EGARCH(0,1) models defined by

---

[20] Note, this rules out the Lévy driven Ornstein-Uhlenbeck volatility processes favored by Barndorff-Nielsen and Shephard (2001a).

$$log(\hat{\sigma}_n^2(t)) \;=\; \omega_n \;+\; \beta_n \cdot log(\hat{\sigma}_n^2(t\text{-}1/n)\,) \;+\; \theta_n \cdot z(t,1/n) \;+\; \gamma_n \cdot [\,|z(t,1/n)| - (2/\pi)^{1/2}\,]$$

$$\omega_n = \alpha \cdot \beta/n, \quad \beta_n = 1 - \beta/n, \quad \theta_n = \rho \cdot \psi \cdot (1/n)^{1/2}, \quad \gamma_n = \psi \cdot (1\text{-}\rho^2) \cdot (1\text{-}(2/\pi))^{-1/2} \cdot (1/n)^{1/2},$$

(4.5)

where $\beta > 0$, $\psi > 0$, and the standardized innovations are defined as in equation (4.3). Interpreted as a sequence of filters, this similarly provides consistent estimates (as $n \to \infty$) of the instantaneous volatility at each point in time for any continuous sample path diffusion of the general form in equation (3.23). In parallel to the consistent GARCH(1,1) filter,

$$lim_{\,n \to \infty} \; \beta_n \;=\; 1,$$

so that the root in the autoregressive polynomial dictating the exponential decay in the weights associated with the past absolute standardized returns approaches unity. Under the additional assumption of i.i.d. normally distributed standardized returns, the sequence of EGARCH(0,1) models defined by equation (4.5) converges weakly to the OU diffusion for $log(\sigma^2(t))$,

$$dp(t) \;=\; \sigma(t)\,dW(t)$$

$$dlog(\sigma^2(t)) \;=\; -\beta\,[log(\sigma^2(t)) - \alpha]\,dt \;+\; \psi\,dV(t)\,,$$

(4.6)

where the instantaneous correlation between the two Wiener processes is determined by the leverage parameter ρ; i.e., $Corr(\,dW(t)\,,\,dV(t)\,) = \rho\,dt$.

Because many candidate ARCH models may serve as consistent filters for the instantaneous volatility, this naturally raises the question of efficiency. The asymptotic distribution theory for the filter errors developed by Nelson and Foster (1994) and Nelson (1996a) allows for a formal analysis of this issue. Intuitively, in the diffusion limit (with continuous sample paths) the process is completely characterized by the first two conditional moments, and the optimal ARCH filter matches both of these. These results for continuous-time stochastic differential equations carry over to the design of optimal ARCH filters for the type of stochastic difference equations employed in the formulation of the discrete-time stochastic volatility models discussed in section 3.1.2. In this situation, if the conditional distribution of the innovations are sufficiently fat tailed, estimating $\sigma^2(t)$ by squaring a distributed lag of past absolute returns, as originally proposed by Taylor (1986) and Schwert (1989), may be more efficient than using a distributed lag of past squared returns. A detailed discussion of these results is beyond the scope of this chapter. However, it is worth noting that the comparisons in Nelson and Foster (1994) related to the diffusion in equation (4.6) show that asymptotically (for $n \to \infty$) the efficiency loss in extracting $log(\sigma^2(t))$ based on the lognormal SARV(1) model in equation (3.18) coupled with the (sub-optimal) linear Kalman filter can be substantial relative to the (asymptotically) optimal ARCH filter (which essentially looks like the EGARCH filter defined in equation 4.5). Of course, this still entails an efficiency loss relative to the optimal non-linear extraction filter (e.g., Kitagawa, 1987), but as discussed above, the numerical integration involved in the implementation of such filters is computationally much more demanding than the simple recursions underlying the filtered volatility estimates from ARCH models.

The ARCH filters explicitly restrict the information set used in the extraction of $\sigma^2(t)$ to past and current returns only; i.e., F_t. Asymptotic (for $n\to\infty$) optimal ARCH smoothers involving both lagged and future returns have been developed by Nelson (1996b). The basic idea behind the construction of optimal ARCH smoothers exploit principles similar to those involved in the extension of the Kalman filter to a Kalman smoother (e.g., Anderson and Moore, 1979). It is noteworthy that in contrast to the optimal ARCH filters, the resulting optimal ARCH smoothers do not necessarily match the first two conditional moments of the true distribution. An alternative asymptotic distribution theory for analyzing smoothed volatility measurements is provided by the rolling regression approach in Foster and Nelson (1996). We return to a discussion of some of these results in the following section.

4.2. Realized Volatility

The use of historical, ex-post sample variances computed from higher frequency return data as lower frequency volatility measures has many precedents within the empirical finance literature. For example, Poterba and Summers (1986), French, Schwert and Stambaugh (1987), and Schwert (1989) rely on monthly sample variances computed from daily returns, while Schwert (1990), Hsieh (1991) and Taylor and Xu (1997) exploit intraday data to produce daily sample return variance measures. In spite of the intuitive appeal of using sample variance estimators over fixed horizons as simple nonparametric volatility measures, they appear hard to justify theoretically if volatility truly is time-varying. However, by connecting the sample variances, termed *Realized Volatility* in financial economics, to the theory of quadratic variation it is possible to more formally justify and assess the properties of such measures. Moreover, this approach to volatility measurement has inspired promising and ongoing new research into volatility modeling based on general distributional assumptions. The formal definition is straightforward.

DEFINITION 6 - Realized Volatility

The *Realized Volatility* over *[t-h,t]*, for *$0<h \leq t \leq T$,* is defined by

$$v^2(t,h;n) \equiv \Sigma_{i=1,..,n} \, r(t-h+(i/n)\cdot h, h/n)^2. \tag{4.7}$$

The realized volatility is simply the second (uncentered) sample moment of the return process over a fixed interval of length h, scaled by the number of observations n (corresponding to the sampling frequency $1/n$), so that it provides a volatility measure calibrated to the h-period measurement interval. Although the definition is stated in terms of equally spaced observations, most results discussed below carry over to situations in which the realized volatility is based on the sum of unevenly, but increasingly finely sampled squared returns.

The realized volatility measure is closely related to, but different from, the theoretical volatility

concepts introduced in section 2. For example, if the mean return is zero, $\mu(t) \equiv 0$, the realized volatility represents the ex-post sample variance computed from n discretely sampled (h/n)-period returns over $[t-h,t]$. In this case the realized volatility is (ex-ante) unbiased for the expected volatility, $\vartheta^2(t,h)$. Formally, we have the following slight extension of equation (2.14) (see, e.g., Protter, 1992, Corollary 3 of Theorem 27, Chapter 2).

PROPOSITION 4 - Realized Volatility as an Unbiased Volatility Estimator

If the return process is square-integrable and $\mu(t) \equiv 0$, then for any value of $n \geq 1$ and $h > 0$,

$$\vartheta^2(t,h) = E[\upsilon^2(t,h)|\mathscr{F}_{t-h}] = E[M^2(t-h)|\mathscr{F}_{t-h}] = E[\upsilon^2(t,h;n)|\mathscr{F}_{t-h}]. \tag{4.8}$$

As such, the *ex-post* realized volatility is an unbiased estimator of *ex-ante* expected volatility. Of course, the zero mean assumption is highly restrictive but, as we discuss later, the result remains *approximately* true for a stochastically evolving mean return process over relevant horizons under weak auxiliary conditions, as long as the underlying returns are sampled at sufficiently high frequencies.

Another link to our previous discussion is provided by the theory of rolling sample variance estimators within the continuous sample path (diffusion) setting (Foster and Nelson, 1996). This theory implies that the realized volatility based on increasingly many return observations over finer and finer time intervals is consistent for the corresponding instantaneous volatility. That is, for $h \to 0$ and $n \to \infty$ (at proper rates),

$$plim_{n\to\infty,h\to0}\, \upsilon^2(t,h;n)/h \;=\; plim_{h\to0}\, \upsilon^2(t,h)/h \;=\; \sigma^2(t).$$

Although this result is of theoretical interest, it is less robust and less useful in practice.[21] One constraint is that the theory excludes jumps in both the return and volatility processes. More importantly, from a practical perspective, the result hinges on the length of the time interval going to zero and the number of observations going to infinity (over the vanishing interval) simultaneously. This construction is hard to mimic in any relevant sense. Market microstructure features invariably limit the number of (effectively) uncorrelated return observations, so even for highly liquid markets, it is not possible to measure returns (or volatilities) instantaneously. We discuss these practical issues in more detail below.

The rolling regression procedures and associated ARCH filters and smoothers for the instantaneous volatilities are also usually based on long (weighted) averages of the returns. Adjacent instantaneous volatility measures will therefore involve overlapping return

---

[21] Andreou and Ghysels (2002) provide some intriguing simulation-based efficiency comparisons of the realized volatility measure in Definition 6 with alternative ex-post volatility filters motivated by the Nelson and Foster (1996) rolling regression theory.

observations. This renders formal statistical analysis of the time-series properties of any such derived volatility series complex. The realized volatility approach explicitly seeks to avoid such difficulties by fixing $h>0$ and interpreting $\upsilon^2(t,h;n)$ as a measure of the overall volatility for the $[t-h,t]$ time interval. We turn now towards a general discussion of this approach.

The theoretical properties of realized volatility have been discussed from different perspectives in a number of recent studies including Andersen, Bollerslev, Diebold and Labys (2000b, 2001a,b), Barndorff-Nielsen and Shephard (2001a,2002a,b,c), and Comte and Renault (1998). A simple yet fundamental result follows directly by combining the theory of quadratic variation in Proposition 2 with the Definitions 1 and 6.

PROPOSITION 5 - Consistency of Realized Volatility

The *Realized Volatility* provides a consistent nonparametric measure of the *Notional Volatility*,

$$plim_{n \to \infty} \upsilon^2(t,h;n) \; = \; \upsilon^2(t,h), \qquad\qquad 0<h \leq t \leq T, \qquad\qquad (4.9)$$

where the convergence is uniform in probability.

The notional volatility plays a crucial role in the return dynamics. From the relation between expected notional volatility and expected volatility in equation (3.21), the ex-ante expected notional volatility is also the critical determinant of expected volatility. Any empirical measures of (ex-ante expected) notional volatility based on (3.21) will necessarily depend on the assumed parametric model structure. Proposition 5 implies that, in the limit for increasingly finely sampled returns, or $n \to \infty$, realized volatility is a consistent (nonparametric) estimator of the (realized) notional volatility over any fixed length time interval, $h > 0$.

It is natural to combine the unbiasedness property of realized volatility in Proposition 4 and the consistency result in Proposition 5 to think of ex-post realized volatility measures, in general, as approximately unbiased estimators, and the ex-ante expected values of the realized volatility measures as consistent estimators for the ex-ante expected notional volatility. These conjectures warrant some formal discussion. The former result is, indeed, only approximately correct, while the latter requires a minor auxiliary assumption to be generally valid.

PROPOSITION 6 - Consistency of Ex-Ante Expected Realized Volatility

The *Expected Realized Volatility* provides a consistent estimator of the *Expected Notional Volatility*,

$$plim_{n \to \infty} E[\upsilon^2(t,h;n)|\mathscr{F}_{t-h}] \; = \; E[\upsilon^2(t,h)|\mathscr{F}_{t-h}], 0<h \leq t \leq T, \qquad (4.10)$$

provided the return process is bounded and the expected return, $\mu(s)$, is of integrable variation for s in $[t-h,t]$.

The conclusion embodied in equation (4.10) follows from (4.9) if convergence in probability implies convergence in mean. This is not generally true, however, even if the result is valid for martingale processes, as stated in Proposition 4. The assumption of a bounded return process provides a simple sufficient condition for convergence in mean, although a weaker uniform integrability condition would suffice; e.g., Hoffmann-Jørgensen (1994), sections 3.22-3.25.

Proposition 6 establishes that the ex-post realized notional volatility provides an asymptotically (for $n \to \infty$) unbiased estimator of the ex-ante expected notional volatility, as long as the return is bounded over the relevant horizon. This condition is obviously not satisfied by all admittable price processes in Proposition 1, and will be violated for many typical return specifications, but still does not appear particularly restrictive. For example, one may imagine a bound on the return that prevents a small investment in the asset from ever producing a return that exceeds a (large) multiple of the expected value of all resources available in the world-wide economy. Moreover, it is only a sufficient condition, and the proposition will apply in most cases, even when the boundedness condition is violated. Nonetheless, some specifications are excluded (see, e.g., Barndorff-Nielsen and Shephard, 2002b, for an example).

Although the above consistency result for the realized volatility estimators leaves important considerations regarding the size of potential error terms and any finite-sample biases unanswered, the reasoning suggests that the simple realized volatility measure could serve as a constructive basis for volatility measurement (and modeling) under general conditions. We discuss the issue of the measurement errors involved in such an approach more formally below. It is instructive, however, first to decompose the realized volatility measure into the separate terms associated with the potential sources of error and bias. For that purpose, we apply the canonical decomposition to each return component of the realized volatility definition in equation (4.7) and simplify notation, so for $i = 1, \dots , n,$

$$r(t-h+(i/n)\cdot h, h/n) \quad = \quad \mu(t-h+(i/n)\cdot h, h/n) + M(t-h+(i/n)\cdot h, h/n)$$

$$\equiv \quad \mu_i + M_i .$$
(4.11)

In the frictionless arbitrage-free setting, the return on a risky asset over time intervals of length (h/n) has a martingale innovation of order $(h/n)^{1/2}$, while the corresponding mean component is at most of order (h/n). In particular, exploiting the notation introduced in equation (4.11), we have

$$\upsilon^2(t,h;n) \quad = \quad \Sigma_{i=1,\dots,n} \: [\mu_i^2 + 2 \cdot \mu_i M_i + M_i^2]$$

$$= \quad O_p(h^2/n) + O_p(h^{3/2}/ n^{1/2}) + \Sigma_{i=1,\dots,n} M_i^2$$
(4.12)

$$= \upsilon^2(t,h) + O_p(n^{-1}) + O_p(n^{-\frac{1}{2}}) + [\Sigma_{i=1,..,n} M_i^2 - \upsilon^2(t,h)].$$

It is apparent that the realized volatility may differ from the notional volatility for two distinct reasons. First, the second and third terms on the right-hand-side of the last equation in (4.12) reflect the mean returns which only truly vanishes in the limit for $n \to \infty$. However, the expected return over short intervals (large n) are necessarily small, so the contribution from these terms will be empirically negligible. This conclusion is only reinforced by noting that the component of the largest order, $O_p(n^{-\frac{1}{2}})$, represents a covariance term that is limited by the size of the innovations to the expected return over the *(h/n)* time interval, which typically will be very small. Second, the last term on the right-hand-side of (4.12) has the interpretation of a measurement error term, as Proposition 4 shows that the cumulative squared martingale innovations provide unbiased estimators for the corresponding notional volatility (quadratic variation). Hence, this term has a zero expected value. Nonetheless, for any given value of n it induces a measurement error that is unrelated to the mean return. This component is the source of empirically relevant deviations between realized volatility and (realized) notional volatility.

The actual size and exact distribution of the errors obviously depend on the particular return process and must be analyzed on a case-by-case basis. Barndorff-Nielsen and Shephard (2002a) provide specific evidence for the Ornstein-Uhlenbeck specification with a background-driving Lévy process. Similarly, Meddahi (2002) presents explicit expressions for the different terms in equation (4.12) for the class of eigenfunction stochastic volatility models, and goes on to numerically compare the size of the unconditional variance of the measurement error to the unconditional variance of the notional (integrated) volatility for some of the continuous-time diffusions of the general form in equation (3.23) that have been estimated in the existing literature.

The preceding discussion implies that realized volatility is approximately (apart from minor biases induced by the mean component) unbiased for the corresponding notional volatility. Importantly, it also follows from the local martingale property in (2.11) and the decomposition in (4.12), that the associated measurement errors are approximately uncorrelated, i.e.,

$$E[(\upsilon^2(t+j,h;n)-\upsilon^2(t+j,h)) \cdot (\upsilon^2(t,h;n)-\upsilon^2(t,h))] = O_p(n^{-1}), \qquad (4.13)$$

where $j \neq 0$. Again, the $O(n^{-1})$ term is identically equal to zero in the case of constant mean returns, and otherwise likely to be small in empirically realistic situations. This confirms that realized volatilities provide meaningful and theoretically well founded volatility measurements. Moreover, they constitute natural and convenient inputs into modeling and inference procedures concerning the expected notional volatility and, by extension, the expected return volatility.[22]

---

[22] This result also underlies the simple GMM estimation procedure for parametric continuous-time stochastic volatility models in Barndorff-Nielsen and Shephard (2002a) and Bollerslev and Zhou (2002) based on matching sample moments of realized volatility with corresponding model implied moments for notional (integrated) volatility.

It is worth reiterating that the fixed *h*, large *n* asymptotics, or realized volatility asymptotics, underlying the results discussed above, is pivotal in practice. In particular, in spite of the theoretical desirability of letting interval size, *h,* shrink indefinitely as an increasing number of high-frequency return observations is used within each (vanishing) interval (as in the ARCH filters and smoothers discussed in section 4.1), this idea is difficult (impossible) to mimic in practice. The number of data points, *n,* that adhere (approximately) to the underlying no-arbitrage semi-martingale property over short time intervals is severely limited by various market microstructure frictions. This invariably puts an effective (asset and/or market specific) lower bound on the highest sampling frequency that is applicable in empirical work, say *1/n >1/N*. We return to this important practical consideration below.

In summary, the realized volatility approach exploiting intraday return observations allow for directly observable return volatility measures that are consistent, approximately unbiased and have uncorrelated measurement errors. It is natural to exploit these properties by building a time series model directly for the observed realized volatility measures through standard ARMA style modeling. Importantly, such procedures sidestep the complex task of providing an appropriate model for the intraday volatility patterns while still exploiting the inherent information in the high-frequency data for lower frequency volatility movements. Of course, the use of nonparametric volatility measurements invariably entails a loss in statistical efficiency relative to the use of a fully (and by assumption correctly) specified parametric volatility model. We comment further on this issue, and the practical merits of reduced form realized volatility modeling below.

The imposition of additional restrictions on the return process allows for important additional insight into the size and asymptotic distribution of the realized volatility errors. In particular, consider the class of continuous sample path diffusions, characterized by the sde,

$$dp(t) = \mu(p(t), \sigma(t\text{-})) \, dt + \sigma(t\text{-}) \, dW(t) \, , \qquad 0 \le t \le T, \tag{4.14}$$

where $\mu(p(t), \sigma(t))$ is predictable and of finite variation and the $\sigma(t)$ process is independent of the Brownian motion, $W(t)$ (c.f., Proposition 3 and the corresponding sde in equation 3.23). Recent results by Barndorff-Nielsen and Shephard (2002a) then provide the asymptotic (for $n \to \infty$) distribution of the measurement errors in this setting.

PROPOSITION 7 - Asymptotic Mixed Normality of Realized Volatility

The *Realized Volatility Errors* for the continuous sample path diffusion in equation (4.14) is distributed as

$$[\upsilon^2(t,h;n) - \upsilon^2(t,h)] \cdot [\, \tfrac{2}{3} \cdot \upsilon^{[4]}(t,h;n) \,]^{-1/2} \to N(0,1), \tag{4.15}$$

for $n \to \infty$, where

$$\upsilon^{[4]}(t,h;n) \ \equiv \ \Sigma_{i=1,..,n} \, r(t-h+i\cdot(h/n)\,,\,h/n)^4. \tag{4.16}$$

This proposition strengthens the convergence of realized volatility to notional volatility (in probability) result considerably by providing the asymptotic distribution of the corresponding errors. Formally, the variance of the realized volatility errors is given by $\frac{2}{3}\cdot\int_{t-h}^{t}\sigma^4(\tau)d\tau$, which is consistently estimated by $\frac{2}{3}\cdot\upsilon^{[4]}(t,h;n)$ as defined in equation (4.16). Hence, the magnitude of the errors depends upon the level of the (latent) volatility. This result represents a fundamental extension of the corresponding expression for the variance of the (uncentered) sample variance for the continuous-time random walk model discussed in section 4.1 above, $2\cdot\sigma^4\cdot h^2\cdot n^{-1} + O(n^{-1})$.[23] Of course, the general expression for the variance of the realized volatility errors, or $\upsilon^{[4]}(t,h;n)$, is straightforward to calculate in practice.

The more powerful distributional result in Proposition 7, compared to the weak convergence in Proposition 4, comes at the cost of the stronger assumption on the underlying log-price process. Although the additional conditions, most notably the absence of jumps in the price path, likely are violated empirically at the highest sampling frequencies, the asymptotic distribution should nonetheless serve as a very useful theoretical benchmark for assessing the properties of the realized volatility measures and further assist in guiding empirical procedures.

In this regard, Barndorff-Nielsen and Shephard (2002c) find that an improved finite-sample (finite n) approximation may be obtained by the log-linearization,

$$[log(\upsilon^2(t,h;n)) - log(\upsilon^2(t,h)) +\tfrac{1}{2}\cdot s(t,h;n)^2 \,]\cdot s(t,h;n)^{-1} \ \sim \ N(0,1), \tag{4.17}$$
where
$$s(t,h;n)^2 \ \equiv \ max\{ \ \tfrac{2}{3}\,\upsilon^{[4]}(t,h;n)\cdot\upsilon^2(t,h;n)^{-2} \, , \, 2/n \ \}. \tag{4.18}$$

The upper bound of $2/n$ in equation (4.18) arises from imposing the theoretical lower bound for $n \to \infty$ on the first ratio. This approximation seems to work well, even for moderately sized n (say $n \geq 10$), in a (stylized) simulation setting. The improvement is related to the logarithm delivering a variance-stabilizing transformation. In that sense, the improved finite-sample distribution obtained by equations (4.17) and (4.18) is directly in line with the evidence for the parametric discrete-time stochastic volatility models discussed in section 3.1.2, for which the innovation process for the formulations involving the logarithmic volatility typically exhibits much reduced (conditional) heteroskedasticity.[24]

The above results speak to the precision in extracting information about the (realized) notional

---

[23] In an interesting paper Bai, Russell and Tiao (2000) extend this expression for the unconditional variance to allow for serially correlated, leptokurtic high-frequency returns subject to GARCH effects and intraday deterministic volatility patterns.

[24] This is also consistent with the recent empirical evidence in Andersen, Bollerslev, Diebold and Labys (2001a) and Andersen, Bollerslev, Diebold and Ebens (2001) suggesting that the unconditional distribution of realized volatility is approximately log-normal.

volatility from the realized volatility measures. Realizations of the notional volatility are, of course, of direct interest as indicators of return variability. However, they also provide an indication of the character of the underlying return distribution itself. In particular, it follows - under the identical conditions underlying Proposition 7 - that the returns, $r(t,h)$, conditional on the notional volatility (and the mean return) over the *[t-h,t]* return interval will be Gaussian.

PROPOSITION 8 - Normal Mixture Distribution

The discrete-time returns $r(t,h)$ over *[t-h,t], 0<h ≤ t ≤ T*, for the continuous sample path diffusion in equation (4.14) is distributed as a *normal mixture*,

$$r(t,h) \mid \sigma\{\mu(t,h), \upsilon^2(t,h)\} \quad \sim \quad N(\mu(t,h), \upsilon^2(t,h)). \tag{4.19}$$

Of course, the (ex-ante) mean return and the notional volatility is not directly observable. However, integrating out $\sigma\{\mu(t,h), \upsilon^2(t,h)\}$, the proposition implies that the return distribution conditional on time *t-h* information should be governed by a normal mixture distribution.[25] This is directly in line with the implications of the Mixture-of-Distributions Hypothesis (MDH) pioneered by Clark (1973) which, as discussed in section 3.1.2, has motivated the formulation of some of the most widely used empirical discrete-time stochastic volatility models.

The consistency of the realized volatility for the notional volatility in Proposition 4 along with the approximate log-normality of the realized volatility distribution and the normal mixture distribution in Proposition 8 suggest a simple alternative empirical return-volatility modeling strategy. Assume that the demeaned returns standardized by the realized volatilities, *[r(t,h) - μ(t,h)] ·υ²(t,h;n)⁻¹*, are (approximately) Gaussian, coupled with a simple reduced form (approximately) Gaussian time series model for the logarithmic realized volatilities, *log[υ²(t,h;n)]*. Effectively, this modeling strategy relies exclusively on forecasts for the distribution of the future notional volatilities through the observed realized volatilities, and as such is in principle straightforward to implement in practice.[26] Moreover, leverage effects, or asymmetries, in the notional volatility are easily incorporated by allowing the time series model for *log[υ²(t,h;n)]* to depend (non-trivially) on the level of the (past) returns. This empirical modeling framework has recently been pursued successfully by Andersen, Bollerslev, Diebold, and Labys (2001b). Related empirical work by Fleming, Kirby and Ostdiek (2002) and Martens and Zein (2002) suggests that important improvement can be obtained by using this realized volatility modeling approach in lieu of standard volatility

---

[25] The presence of jumps in the price process will generally render the corresponding distribution of the (standardized) returns non-normal. This could be exploited in the formulation of tests for (the importance of) jumps (e.g., Drost, Nijman and Werker, 1998).

[26] Of course, as discussed above the realized volatility invariably differs from the true notional volatility for finite *n*. However, the measurement errors are (approximately) serially uncorrelated, and therefore effectively averaged out in any reduced form time series model for $\upsilon^2(t,h;n)$.

modeling procedures for practical financial decision making as exemplified by better portfolio allocations and more accurate option pricing calculations.[27]

As mentioned repeatedly, the realized volatility approach, of holding $h > 0$ fixed, is motivated by the fact that it is undesirable, and due to the presence of market microstructure frictions indeed practically infeasible, to sample returns infinitely often ($n \to \infty$) over infinitesimally short time intervals ($h \to 0$). Model specific calculations and simulations by Andersen and Bollerslev (1998a), Andersen, Bollerslev and Lange (1999), Andreou and Ghysels (2002), Bai, Russell and Tiao (2000), Barndorff-Nielsen and Shephard (2002c), Barucci and Renò (2002), Bollerslev and Zhou (2002), and Zumbach, Corsi and Trapletti (2002), among others, illustrate the effects of finite n (and h) for a variety of settings. The discrepancies in the underlying model formulation and character of the assumed frictions render a general assessment of the results difficult. Moreover, the size of the measurement errors are often computed unconditionally rather than conditional on the realization of the realized volatility statistic. Nonetheless, it is evident that the measurement errors typically are non-trivial. In turn, this has motivated the development of alternative realized volatility measures designed to circumvent the microstructure biases. One common adjustment involves a correction for serial correlation in the high-frequency returns, supposedly induced by asynchronous trading and quoting, as well as bid-ask bounce.[28] This typically takes the form of a simple MA(1) type filter, examples of which include Zhou (1996), Andersen, Bollerslev, Diebold, and Ebens (2001), and Corsi, Zumbach, Müller and Dacorogna (2001).[29]

Alternatively, it may be desirable to focus on volatility estimators that are robust to microstructure noise. One example is the range. Alizadeh, Brandt and Diebold (2002) and Brandt and Diebold (2002) show that range-based volatility is robust both to bid-ask bounce and to asynchronous trading. This feature must be weighed against the fact that formal statistical analysis of range based estimators require specific distributional assumptions avoided with the realized volatility and power based variation measures discussed at length in the current section. A second example involves using power transformations of the absolute returns rather than the squared returns (Davidian and Carroll, 1987). For instance, as noted in section 4.1 above, the optimal ARCH filters for discrete-time stochastic volatility models may

---

[27] Other interesting recent empirical studies highlighting the potential benefits of realized volatility procedures, in addition to the ones already cites, include Areal and Taylor (2002), Galbraith and Zinde-Walsh (2001), Maheu and McCurdy (2002), and Thomakos and Wang (2002).

[28] This is manifest in the volatility signature plots in Andersen, Bollerslev, Diebold and Labys (2000a). In a multivariate setting, the so-called Epps effect and the breakdown of the sample autocorrelations at the highest intraday frequencies is also directly related to this phenomenon (Epps, 1979; and Lundin, Dacorogna and Müller, 1999).

[29] This mirrors (in many ways) earlier developments related to the estimation of CAPM beta's in the presence of asynchronous trading effects by Scholes and Williams (1977), and the adjustment to the sample variance in French, Schwert and Stambaugh (1987) obtained by including the cross-product between successive returns. Another recent approach developed in Malliavin and Mancino (2000) exploit frequency-domain methods, thereby avoiding any (arbitrary) assumptions about the length of the high-frequency return interval. However, this necessarily involves another (arbitrary) assumption about the (latent) price path between two successive observations (see also Barucci, Mancino and Renò, 2000; and Barucci and Renò, 2002).

entail a distributed lag of past absolute returns as opposed to the squared returns (Nelson and Foster, 1994). These results have a direct analogue for the continuous sample path diffusion in equation (4.14). The following two formal definitions of *Notional* and *Realized Power Variation* measures, adapted from Barndorff-Nielsen and Shephard (2001b), directly parallel the notional and realized volatility concepts in the general frictionless arbitrage-free setting.

DEFINITION 7 - Power Variation Measures

The *Notional s^{th} Order Power Variation* and the *Realized s^{th} Order Power Variation, s>0,* for the diffusion in equation (4.14) over *[t-h,t], 0<h ≤ t ≤ T,* are defined, respectively, by

$$\upsilon^{[s]}(t,h) \;\equiv\; \int^t_{t-h} \sigma^s(\tau)\,d\tau \tag{4.20}$$

and

$$\upsilon^{[s]}(t,h;n) \;\equiv\; \Sigma_{i=1,...,n}\,|r(t-h+i\cdot(h/n)\,,\,h/n)|^s\,. \tag{4.21}$$

It is apparent that, for *s=2,* the definitions correspond directly to the notional volatility and the realized volatility concepts discussed above; i.e., $\upsilon^{[2]}(t,h) \equiv \upsilon^2(t,h)$ and $\upsilon^{[2]}(t,h;n) \equiv \upsilon^2(t,h;n)$, respectively. However, other values of *s* may allow for more robust measurements. In particular, recent developments by Barndorff-Nielsen and Shephard (2001b) extend the distributional results for the (standard) realized volatility in Proposition 5 to the generalized power variation measures.

PROPOSITION 9 - Asymptotic Mixed Normality of Realized Power Variation

The *Realized s^{th} Order Power Variation Errors, s>0,* for the continuous sample path diffusion in equation (4.11) is distributed as

$$\kappa_s(h/n)^{s/2-1}\cdot[\kappa_s^{-1}(h/n)^{1-s/2}\cdot\upsilon^{[s]}(t,h;n) - \upsilon^{[s]}(t,h)]\cdot[\kappa_{2s}^{-1}\cdot\omega_s\cdot\upsilon^{[2s]}(t,h;n)\,]^{-1/2} \;\rightarrow\; N(0,1), \tag{4.22}$$

for *n→∞*, where $\kappa_s=E(|Z|^s)$, $\omega_s=Var(|Z|^s)$, and *Z* denotes a standard normal distribution.

The special case corresponding to *s=1* is naturally termed *Absolute Variation.* The realized absolute variation is, of course, simply constructed by the summation of the *n* absolute returns, $|\,r(t-h+i\cdot(h/n)\,,\,h/n)|$, *i = 1, 2, ..., n,* within the *[t-h,t]* time interval. From Proposition 9, the asymptotic (for *n → ∞)* distribution of the corresponding measurement error for the notional absolute variation satisfies

$$[(\pi/2)^{1/2}(h/n)^{1/2}\upsilon^{[1]}(t,h;n) - \upsilon^{[1]}(t,h)]\cdot[(h/n)(\pi/2 - 1)\upsilon^2(t,h;n)\,]^{-1/2} \;\rightarrow\; N(0,1).$$

This result holds the promise of improved volatility measurements and more accurate inference than the (standard) realized volatility (quadratic variation) measure in Proposition 7. The absolute variation measures also provide a theoretical basis for gauging the empirical results in Andersen and Bollerslev (1998b) among others based on $\upsilon^{[1]}(t,h;n)$. Similarly, the so-called Taylor Effect (e.g., Granger and Ding, 1995) according to which the autocorrelations of power transforms of the absolute returns are maximized (empirically) for values of s close to unity may be directly related to these new results.

Research in the realized volatility area has evolved rapidly over the past few years, and it is still too early to draw firm conclusions or consensus opinion about the preferred procedures. However, the theoretical and empirical results reported to date have been very promising. Current research into multivariate extensions of Propositions 7 and 9 may help form a firm theoretical foundation for new realized CAPM betas and factor loadings that are being explored empirically in ongoing work. Extensions of the continuous sample path diffusion in equation (4.14) to allow for more general processes (including asymmetries or leverage effects) within the frictionless, arbitrage-free setup is another theoretically challenging problem. As discussed repeatedly, the development of reliable empirical procedures for dealing with the inherent market microstructure frictions at the highest sampling frequencies across different assets and markets is also of utmost importance from a practical perspective (see also the surveys by Hasbrouck, 1996; Dacorogna et al., 2001; and Engle and Russell, 2002).

5. DIRECTIONS FOR FUTURE RESEARCH

In the last ten years, there has been a movement toward the use of newly-available high-frequency asset return data, and away from restrictive and hard-to-estimate parametric models toward flexible and computationally simple nonparametric approaches. Those trends will continue. Two related, directions for future research are apparent: (1) continued development of methods for exploiting the volatility information in high-frequency data, and (2) volatility modeling and forecasting in the high-dimensional multivariate environments of practical financial economic relevance. The realized volatility concept readily tackles both: it incorporates the powerful information in high-frequency data while dispensing with the need to actually model the high-frequency data, and it requires only the most trivial of computations, thereby bringing within reach the elusive goal of accurate and high-dimensional volatility measurement, modeling and forecasting. We look forward to realization of that goal in the not-too-distant future.

REFERENCES

Aït-Sahalia, Y., L.P. Hansen, J. Scheinkman (2002), "Discretely Sampled Diffusions," in Y.
       Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam:
       North Holland, forthcoming.

Akgiray, V. and G. Booth (1986), "Stock Price Processes with Discontinuous Time Paths: An
       Empirical Investigation," *The Financial Review*, 21, 163-184.

Alizadeh, S., M.W. Brandt, and F.X. Diebold (2002), "Range-Based Estimation of Stochastic
       Volatility Models," *Journal of Finance*, 57, 1047-1092.

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

Andersen, T.G. (1996), "Return Volatility and Trading Volume: An Information Flow
       Interpretation of Stochastic Volatility," *Journal of Finance*, 51, 169-204.

Andersen, T.G. (1992), "Volatility," Manuscript, Kellogg Graduate School of Management,
       Northwestern University.

Andersen, T.G., L. Benzoni and J. Lund (2002), "An Empirical Investigation of Continuous
       Time Equity Return Models," *Journal of Finance*, 57, 1239-1284.

Andersen, T.G. and T. Bollerslev (1997), "Heterogeneous Information Arrivals and Return
       Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of
       Finance*, 52, 975-1005.

Andersen, T.G. and T. Bollerslev (1998a), "Answering the Skeptics: Yes, Standard Volatility
       Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885-905.

Andersen, T.G. and T. Bollerslev (1998b), "Deutschemark-Dollar Volatility: Intraday Activity
       Patterns, Macroeconomic Announcements, and Longer Run Dependencies," *Journal of
       Finance*, 53, 219-265.

Andersen, T.G. and T. Bollerslev (1998c), "ARCH and GARCH Models," in S. Kotz, C.B.
       Read and D.L. Banks (eds), *Encyclopedia of Statistical Sciences, Vol.II* . New York:
       John Wiley and Sons.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001), "The Distribution of
       Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2000a), "Great Realizations," *Risk*,
       13, 105-108.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2000b), "Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian," *Multinational Finance Journal*, 4, 159-179.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001a), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42-55.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001b), "Modeling and Forecasting Realized Volatility," NBER Working Paper No. 8160.

Andersen, T.G., T. Bollerslev, F.X. Diebold and C. Vega (2002), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, forthcoming.

Andersen, T.G., T. Bollerslev and S. Lange (1999), "Forecasting Financial Market Volatility: Sample Frequency vis-a-vis Forecast Horizon," *Journal of Empirical Finance*, 6, 457-477.

Andersen, T.G., H.J. Chung and B.E. Sørensen (1999), "Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study," *Journal of Econometrics*, 91, 61-87.

Andersen, T.G., and J. Lund (1997), "Estimating Continuous-Time Stochastic Volatility Models of the Short Term Interest Rate," *Journal of Econometrics*, 77, 343-377.

Andersen, T.G., and J. Lund (1998), "Stochastic Volatility and Mean Drift in the Short Term Interest Rate Diffusion: Sources of Steepness, Level, and Curvature in the Yield Curve," Manuscript, Department of Finance, Kellogg School, Northwestern University.

Anderson, B.D.O. and J.B. Moore (1979), *Optimal Filtering*. Englewood Cliffs, NJ: Princeton Hall.

Andreou, E. and E. Ghysels (2002), "Rolling Sample Volatility Estimators: Some New Theoretical, Simulation and Empirical Results," *Journal of Business and Economic Statistics*, 20, 363-376.

Ané, T. and H. Geman (2000), "Order Flow, Transaction Clock, and Normality of Asset Returns," *Journal of Finance*, 55, 2259-2284.

Areal, N.M.P.C. and S.J. Taylor (2002), "The Realized Volatility of FTSE-100 Futures Prices," *Journal of Futures Markets*, forthcoming.

Back, K. (1991), "Asset Prices for General Processes," *Journal of Mathematical Economics*, 20, 371-395.

Bai, X., J.R. Russell and G.C. Tiao (2000), "Beyond Merton's Utopia: Effects of Non-Normality and Dependence on the Precision of Variance Estimates Using High-Frequency Financial Data," Manuscript, Graduate School of Business, University of Chicago.

Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 3-30.

Bakshi, G, C. Cao and Z. Chen (1997), "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003-2049.

Ball, C.A. and W.N. Torous (1985), "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing," *Journal of Finance*, 40, 155-173.

Bandi, F. and P.C.B. Phillips (2002), "Nonstationary Continuous-Time Models," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Banon, G. (1978), "Nonparametric Identification for Diffusion Processes," *SIAM Journal of Optimization and Control*, 16, 380-395.

Barndorff-Nielsen, O.E. and N. Shephard (2001a), "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society, Series B*, 63, 167-241.

Barndorff-Nielsen, O.E. and N. Shephard (2001b), "Realized Power Variation and Stochastic Volatility Models," Manuscript, Nuffield College, Oxford, United Kingdom.

Barndorff-Nielsen, O.E. and N. Shephard (2002a), "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, 253-280.

Barndorff-Nielsen, O.E. and N. Shephard (2002b), "Estimating Quadratic Variation Using Realised Variance," *Journal of Applied Econometrics*, forthcoming.

Barndorff-Nielsen, O.E. and N. Shephard (2002c), "How Accurate is the Asymptotic Approximation to the Distribution of Realized Volatility?" in D.W.F. Andrews, J.L. Powell, P.A. Ruud, and J.H. Stock (eds.), *Identification and Inference for Econometric Models*, forthcoming.

Barucci, E., M.E. Mancino, and R. Renò (2000), "Volatility Estimation via Fourier Analysis," Manuscript, University of Pisa.

Barucci, E., and R. Renò (2002), "On Measuring Volatility and the GARCH Forecasting Performance," *Journal of International Financial Markets, Institutions & Money*, 12, 183-200.

Bates, D. (1996a), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutschemark Options," *Review of Financial Studies*, 9, 69-107.

Bates (1996b), "Testing Option Pricing Models," in G.S. Maddala (ed.), *Handbook of Statistics Vol. 14, Statistical Methods in Finance*, 567-611. Amsterdam: North-Holland.

Bekaert, G. and G. Wu (2000), "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies*, 13, 1-42.

Bibby, B.M., M. Jacobsen and M. Sørensen (2002), "Estimating Functions for Diffusions," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

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

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

Bollerslev, T. (1990), "Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach," *Review of Economic and Statistics*, 72, 498-505.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992), "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5-59.

Bollerslev, T., R.F. Engle and J.M. Wooldridge (1988), "A Capital Asset Pricing Model with Time Varying Covariances," *Journal of Political Economy*, 96, 116-131.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994), "ARCH Models," in R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics, Volume IV*, 2959-3038. Amsterdam: North-Holland.

Bollerslev, T. and H.O. Mikkelsen (1996), "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometrics*, 73, 151-184.

Bollerslev, T. and J.M. Wooldridge (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Reviews*, 11, 143-172.

Bollerslev, T. and H. Zhou (2002), "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility," *Journal of Econometrics*, 109, 33-65.

Bougerol, P. and N. Picard (1992), "Stationarity of GARCH Processes and Some Non-Negative Time Series," *Journal of Econometrics*, 52, 115-128.

Brandt, M.W. and F.X. Diebold (2002), "A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations," Manuscript, University of Pennsylvania.

Breidt, F.J., N. Crato and P. de Lima (1998), "On the Detection and Estimation of Long Memory in Stochastic Volatility," *Journal of Econometrics*, 73, 325-334.

Campbell, J.Y. and L. Hentschel (1992), "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics*, 31, 281-318.

Canina, L. and S. Figlewski (1993), "The Information Content of Implied Volatility," *Review of Financial Studies*, 6, 659-681.

Chan, K.C., G.A. Karolyi, F.A. Longstaff and A. Sanders (1992), "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47, 1209-1227.

Chen, X., L.P. Hansen and M. Carrasco (2000), "Nonlinearity and Temporal Dependence," Manuscript, Department of Economics, University of Chicago.

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*, 57, 407-458.

Chernov, M., A.R. Gallant, E. Ghysels and G.E. Tauchen (2001), "Alternative Models for Stock Price Dynamics," Manuscript, Columbia University, Duke University and University of North Carolina.

Chib, S., F. Nardari and N. Shephard (1999), "Analysis of High Dimensional Multivariate Stochastic Volatility Models," Manuscript, Olin School of Business, Washington University.

Christensen, B.J. and N.R. Prabhala (1998), "The Relation Between Implied and Realized Volatility," *Journal of Financial Economics*, 37, 125-150.

Clark, P.K. (1973), "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135-155.

Comte, F. and E. Renault (1998), "Long Memory in Continuous Time Stochastic Volatility Models," *Mathematical Finance*, 8, 291-323.

Corsi, F., G. Zumbach, U. Müller, and M.M. Dacorogna (2001), "Consistent High-Precision Volatility from High-Frequency Data," *Economic Notes*, 30, 183-204.

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

Cox, J.C. and S.A. Ross (1976), "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 7, 229-263.

Dacorogna, M.M., R. Gencay, U. Müller, R.B. Olsen and O.V. Pictet (2001). *An Introduction to High-Frequency Finance*. San Diego, CA: Academic Press.

Dai, Q. and K.J. Singleton (2000), "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943-1978.

Danielsson, J. (1994), "Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood," *Journal of Econometrics*, 61, 375-400.

Davidian, M. and R.J. Carroll (1987), "Variance Function Estimation," *Journal of the American Statistical Association*, 82, 1079-1091.

Day, T.E. and C.M. Lewis (1992), "Stock Market Volatility and the Information Content of Stock Index Options," *Journal of Econometrics*, 52, 267-287.

Diebold, F.X. and A. Inoue (2001), "Long Memory and Structural Change," *Journal of Econometrics*, 105, 131-159.

Diebold, F.X. and J.A. Lopez (1995), "Modeling Volatility Dynamics," in K.V. Hoover (ed.), *Macroeconometrics: Developments, Tensions and Prospects*, 427-472. Boston: Kluwer Academic Press.

Diebold, F.X. and M. Nerlove (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics*, 4, 1-22.

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

Dohnal, G. (1987), "On Estimating the Diffusion Coefficient," *Journal of Applied Probability*, 24, 105-114.

Drost, F.C. and T.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.

Drost, F.C., T.E. Nijman and B.J.M Werker (1998), "Estimation and Testing in Models Containing Both Jumps and Conditional Heteroskedasticity," *Journal of Business and Economic Statistics*, 16, 237-243.

Duan, J.C. (1995), "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13-32.

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

Duffie, D. and K.J. Singleton (1993), "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 929-952.

Duffie, D., J. Pan and K. Singleton (2000), "Transform Analysis and Asset Pricing for Affine Jump Diffusions," *Econometrica*, 68, 1343-1376.

Dumas, B., J. Fleming and R. E. Whaley (1998), "Implied Volatility Functions: Empirical Tests," *Journal of Finance*, 53, 2059-2106.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987-1008.

Engle, R.F. (1995). *ARCH: Selected Readings*. Oxford, UK: Oxford University Press.

Engle, R.F. (2001), "Dynamic Conditional Correlations," Manuscript, Department of Economics, University of California, San Diego.

Engle, R.F. and T. Bollerslev (1986), "Modeling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1-50.

Engle, R.F. and Kroner, K. (1995), "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11, 122-150.

Engle, R.F. and G.G.J. Lee (1999), "A Permanent and Transitory Component Model of Stock Return Volatility," in R.F. Engle and H. White (eds.), *Cointegration, Causality, and*

Forecasting: A Festschrift in Honor of Clive W.J. Granger, 475-497. Oxford, UK: Oxford University Press.

Engle, R.F., D.M. Lilien and R.P. Robins (1987), "Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica*, 58, 525-542.

Engle, R.F. and A.J. Patton (2001), "What Good is a Volatility Model?" *Quantitative Finance*, 1, 237-245.

Engle, R.F. and J. Russell (2002), "High-Frequency and Transaction Data," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Epps, T.W. (1979), "Comovements in Stock Prices in the Very Short Run," *Journal of the American Statistical Association*, 74, 291-298.

Epps, T.W. and M.L. Epps (1976), "The Stochastic Dependence of Security Prices Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis," *Econometrica*, 44, 305-321.

Eraker, B. (2001), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," Manuscript, Department of Economics, Duke University.

Eraker, B., M. Johannes and N. Polson (2002), "The Impact of Jumps in Returns and Volatility," *Journal of Finance*, forthcoming.

Fleming, J. (1998), "The Quality of Market Volatility Forecasts Implied by S&P100 Index Options Prices," *Journal of Empirical Finance*, 5, 317-345.

Fleming, J., C. Kirby, and B. Ostdiek (2002), "The Economic Value of Volatility Timing Using Realized Volatility," *Journal of Financial Economics*, forthcoming.

Fleming, J., B. Ostdiek and R.E. Whaley (1995), "Predicting Stock Market Volatility: A New Measure," *Journal of Futures Market*, 15, 265-302.

Florens-Zmirou, D. (1993), "On Estimating the Diffusion Coefficient from Discrete Observations," *Journal of Applied Probability*, 30, 790-804.

Fornari, F. and A. Mele (2001), "Recovering the Probability Density Function of Asset Prices Using GARCH as Diffusion Approximations," *Journal of Empirical Finance*, 8, 83-109.

Foster, D. and D.B. Nelson (1996), "Continuous Record Asymptotics for Rolling Sample Estimators," *Econometrica*, 64, 139-174.

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

Fridman, M. and L. Harris (1998), "A Maximum Likelihood Approach for Non-Gaussian Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 16, 284-291.

Galbraith, J.W. and V. Zinde-Walsh (2001), "Properties of Estimates of Daily GARCH Parameters Based on Intraday Observations," Manuscript, Department of Economics, McGill University.

Gallant, A.R., D. Hsieh and G.E. Tauchen (1997), "Estimation of Stochastic Volatility Models with Diagnostics," *Journal of Econometrics*, 81, 159-192.

Gallant, A.R., C.T. Hsu and G.E. Tauchen (1999), "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance," *Review of Economics and Statistics*, 81, 617-631.

Gallant, A.R., and G.E. Tauchen (1997), "Estimation of Continuous-Time Models for Stock Returns and Interest Rates" *Macroeconomic Dynamics*, 1, 135-168.

Gallant, A.R., and G.E. Tauchen (1998), "Reprojecting Partially Observed Systems with Applications to Interest Rate Diffusions," *Journal of the American Statistical Association*, 93, 10-24.

Gallant, A.R., and G.E. Tauchen (2000), "SNP, A Program for Nonparametric Time Series Analysis: A Users Guide," available at ftp.econ.duke.edu, folder pub/arg/snp.

Gallant, A.R. and G.E. Tauchen (2002), "Simulated Score Methods and Indirect Inference for Continuous-Time Models," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Garcia, R., E. Ghysels and E. Renault (2002), "Option Pricing Models," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Garcia, R., M.A. Lewis and E. Renault (2001), "Estimation of Objective and Risk-Neutral Distributions Based on Moments of Integrated Volatility," Manuscript, Department of Economics, University of Montreal and CIRANO.

Genon-Catalot, V., C. Laredo and D. Picard (1992), "Non-parametric Estimation of the Diffusion Coefficient by Wavelets Methods," *Scandinavian Journal of Statistics*, 19, 317-335.

Geweke, J., (1989), "Exact Predictive Densities for Linear Models with ARCH Disturbances," *Journal of Econometrics*, 40, 63-86.

Ghysels, E., A. Harvey and E. Renault (1996), "Stochastic Volatility," in G.S. Maddala (ed.), *Handbook of Statistics Vol. 14, Statistical Methods in Finance*, 119-191. Amsterdam: North-Holland.

Giraitis, L., P. Kokoszka and R. Leipus (2000), "Stationary ARCH Models: Dependence Structure and Central Limit Theorem," *Econometric Theory*, 16, 3-22.

Glosten, L.R., R. Jagannathan and D. Runkle (1993), "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779-1801.

Granger, C.W.J. and Z. Ding (1995), "Some Properties of Absolute Returns, an Alternative Measure of Risk," *Annals d'Economie et de Statistique*, 40, 67-91.

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

Harvey, A.C. (1998), "Long Memory in Stochastic Volatility," in J. Knight and J. Satchell (eds.), *Forecasting Volatility in Financial Markets*. London, UK: Butterworth-Heinemann.

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

Hasbrouck, J. (1996), "Modeling Market Microstructure Time Series," in G.S. Maddala and C.R. Rao (eds.) *Handbook of Statistics, Vol.14*, 647-692. Amsterdam: North-Holland.

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, 327-343.

Heston, S.L., and S. Nandi (2000), "A Closed-Form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585-625.

Hoffmann-Jørgensen, J. (1994), *Probability with a View Toward Statistics,* Volume 1. Chapman and Hall Probability Series, New York: Chapman & Hall.

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

Hsieh, D.A. (1991), "Chaos and Nonlinear Dynamics: Application to Financial Markets," *Journal of Finance*, 46, 1839-1877.

J.P. Morgan (1997), *RiskMetrics, Technical Documents*, 4th Edition. New York.

Jacquier, E., N.G. Polson and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 12, 371-389.

Jacquier, E., N.G. Polson and P.E. Rossi (1999), "Stochastic Volatility: Univariate and Multivariate Extensions," Manuscript, Department of Finance, Boston College.

Jarrow, R. and E. Rosenfeld (1984), "Jump Risk and the Intertemporal Asset Pricing Model," *Journal of Business*, 57, 337-351.

Johannes, M. and N. Polson (2002), "Numerical Bayesian Methods for Estimating Continuous-Time Models," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Johannes, M.S. (2000), "A Nonparametric View of the Role of Jumps to Interest Rates, Manuscript, Columbia Business School.

Karatzas, I. and S.E. Shreve (1991), *Brownian Motion and Stochastic Calculus*, (Second Edition). New York: Springer-Verlag.

Kim, S., N. Shephard and S. Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparisons with ARCH Models," *Review of Economic Studies*, 65, 361-393.

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

Kitagawa, G. (1987), "Non-Gaussian State Space Modeling of Nonstationary Time Series," *Journal of the American Statistical Association*, 82, 1032-1063.

Latané, H. and R.J. Rendleman (1976), "Standard Deviations of Stock Price Ratios Implied in Option Prices," *Journal of Finance*, 31, 369-381.

Ledoit, O., P. Santa-Clara and M. Wolf (2001), "Flexible Multivariate GARCH Modeling with an Application to International Stock Markets," Manuscript, The Anderson School, UCLA.

Ling, S. and M. McAleer (2002), "Necessary and Sufficient Moment Conditions for the GARCH(r,s) and Asymmetric Power GARCH(r,s) Models," *Econometric Theory*, 18, 722-729.

Liu, M. (2000), "Modeling Long Memory in Stock Market Volatility," *Journal of Econometrics*, 99, 139-171.

Lundin, M., M.M. Dacorogna and U.A. Müller (1999), "Correlation of High-Frequency Financial Time Series" in P. Lequeux (ed.), *Financial Markets Tick by Tick*, 91-126. New York: John Wiley and Sons.

Maheu, J.M. and T.H. McCurdy (2002), "Nonlinear Features of Realized FX Volatility," *Review of Economics and Statistics*, forthcoming.

Maheswaran, S. and C.A. Sims (1993), "Empirical Implications of Arbitrage-Free Asset Markets," in P.C.B. Phillips (ed), *Models, Methods and Applications of Econometrics*, 301-316.

Malliavin, P. and M. Mancino (2000), "Fourier Series Method for Measurement of Multivariate Volatilities," *Finance and Stochastics*, 6, 49-61.

Martens, M. and J. Zein (2002), "Predicting Financial Volatility: High-Frequency Time-Series Forecasts vis-à-vis Implied Volatility," Manuscript, Econometric Institute and Department of Finance, Erasmus University, Rotterdam.

Meddahi, N. (2002), "A Theoretical Comparison Between Integrated and Realized Volatilities," *Journal of Applied Econometrics*, forthcoming.

Meddahi, N. (2001), "An Eigenfunction Approach for Volatility Modeling," Manuscript, Department of Economics, University of Montreal and CIRANO.

Meddahi, N. and E. Renault (2002), "Temporal Aggregation of Volatility Models," *Journal of Econometrics*, forthcoming.

Meddahi, N. and E. Renault (1996), "Aggregation and Marginalization of GARCH and Stochastic Volatility Models," Manuscript, Department of Economics, University of Montreal and CIRANO.

Mele, A. and F. Fornari (2000), *Stochastic Volatility in Financial Markets: Crossing the Bridge to Continuous Time*. Boston: Kluwer Academic Publishers.

Melino, A. and S.M. Turnbull (1990), "Pricing Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics*, 45, 239-265.

Merton, R.C. (1976), "Option Pricing when Underlying Stock Returns Are Discontinuous," *Journal of Financial Economics*, 3, 125-144.

Merton, R.C. (1980), "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323-361.

Merton, R.C. (1982), "On the Mathematics and Economics Assumptions of Continuous-Time Models," in W.F. Sharpe and C.M. Cootner (eds.), *Financial Economics: Essays in Honor of Paul Cootner*. Englewood Cliffs, New Jersey.

Nelson, D.B. (1988), "The Time Series Behavior of Stock Market Volatility and Returns," Unpublished Ph.D. Dissertation, Department of Economics, MIT.

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

Nelson, D.B. (1990b), "Stationarity and Persistence in the GARCH(1,1) Model," *Econometric Theory*, 6, 318-334.

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

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

Nelson, D.B. (1996a), "Asymptotic Filtering Theory for Multivariate ARCH Models," *Journal of Econometrics*, 71, 1-47.

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

Nelson, D.B. and C.Q. Cao (1992), "Inequality Constraints in the Univariate GARCH Model," *Journal of Business and Economic Statistics*, 10, 229-235.

Nelson, D.B. and D. Foster (1994), "Asymptotic Filtering Theory for Univariate ARCH Models," *Econometrica*, 62, 561-573.

Nelson, D.B. and D. Foster (1995), "Filtering and Forecasting with Misspecified ARCH Models II: Making the Right Forecast with the Wrong Model," *Journal of Econometrics*, 67, 303-335.

Nelson, D.B. and B. Schwartz (1992), "Filtering with ARCH Models: A Monte Carlo Investigation," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 1-6.

Nguyen-Ngoc, L. and M. Yor (2002), "Lévy Processes and Their Application to Option Pricing," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Nicolato, E. and E. Venardos (2001), "Option Pricing in Stochastic Volatility Models of the Ornstein-Uhlenbeck type with a Leverage Effect," Manuscript, Department of Mathematical Sciences, University of Aarhus.

Nijman, T.H., and E. Sentana (1996), "Marginalization and Contemporaneous Aggregation of Multivariate GARCH Processes," *Journal of Econometrics*, 71, 71-86.

Pan, J. (2002), "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3-50.

Piazzesi, M. (2002), "Affine Term Structure Models," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland, forthcoming.

Pitt, M.K. and N. Shephard (1999), "Filtering via Simulation: Auxiliary Particle Filter," *Journal of the American Statistical Association*, 94, 590-599.

Poterba, J. and L. Summers (1986), "The Persistence of Volatility and Stock Market Fluctuations," *American Economic Review*, 76, 1124-1141.

Press, J.S. (1967), "A Compound Events Model for Security Prices," *Journal of Business*, 40, 317-335.

Protter, P. (1992), *Stochastic Integration and Differential Equations: A New Approach* (Second Edition). New York: Springer Verlag.

Renault, E. (1997), "Econometric Models of Option Pricing Errors," in D.M. Kreps and K.F. Wallis (eds.) *Advances in Economics and Econometrics: Theory and Applications*, 223-278. Cambridge, UK: Cambridge University Press.

Robinson, P.M. (1991), "Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression," *Journal of Econometrics*, 47, 67-84.

Robinson, P.M. (2001), "The Memory of Stochastic Volatility Models," *Journal of Econometrics*, 101, 195-218.

Sandmann, G. and S.J. Koopman (1998), "Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood," *Journal of Econometrics*, 87, 271-301.

Scholes, M. and J. Williams (1977), "Estimating Beta from Non-Synchronous Data," *Journal of Financial Economics*, 5, 309-327.

Schwert, G.W. (1989), "Why Does Stock Market Volatility Change Over Time?," *Journal of Finance*, 44, 1115-1153.

Schwert, G.W. (1990), "Stock Volatility and the Crash of '87," *Review of Financial Studies*, 3, 77-102.

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

Tauchen, G.E., and M. Pitts (1983), "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51, 485-505.

Taylor, S.J. (1982), "Financial Returns Modeled by the Product of two Stochastic Processes - a Study of the Daily Sugar Prices 1961-75," in O.D. Anderson (ed.), *Time Series Analysis: Theory and Practice*, 203-226. Amsterdam: North-Holland.

Taylor, S.J. (1986), *Modeling Financial Time Series*. Chichester, UK: John Wiley and Sons.

Taylor, S.J. (1994), "Modeling Stochastic Volatility: A Review and Comparative Study," *Mathematical Finance*, 4, 183-204.

Taylor, S.J. and X. Xu (1997), "The Incremental Volatility Information in One Million Foreign Exchange Quotations," *Journal of Empirical Finance*, 4, 317-340.

Thomakos, D.D. and T. Wang (2002), "Realized Volatility in the Futures Markets," Manuscript, Department of Economics, Florida International University.

Whaley, R.E. (1993), "Derivatives on Market Volatility: Hedging Tools Long Overdue," *Journal of Derivatives*, 1, 71-84.

Wiggins, J.B. (1992), "Estimating the Volatility of S&P500 Futures Prices Using the Extreme Value Method," *Journal of Futures Market*, 12, 265-273.

Wong, E. (1964), "The Construction of a Class of Stationary Markoff Processes," in R. Bellman (ed.), *Sixteenth Symposium in Applied Mathematics - Stochastic Processes in Mathematical Physics and Engineering*, 264-276. Providence, RI: American Mathematical Society.

Zhou, B. (1996), "High-Frequency Data and Volatility in Foreign Exchange Rates," *Journal of Business and Economic Statistics*, 14, 45-52.

Zumbach, G. (2001), "Volatility Processes and Volatility Forecasts with Long Memory," Manuscript, Olsen & Associates, Zürich, Switzerland.

Zumbach, G., F. Corsi and A. Trapletti (2002), "Efficient Estimation of Volatility using High Frequency Data," Manuscript, Olsen & Associates, Zürich, Switzerland.