GARCH options via local risk minimization

JUAN-PABLO ORTEGA*

Centre National de la Recherche Scientifique, Département de Mathématiques de Besançon,
Université de Franche-Comté, UFR des Sciences et Techniques,
16 route de Gray, F-25030 Besançon cedex, France

(Received 2 March 2009; in final form 13 May 2010)

We apply a quadratic hedging scheme developed by Föllmer, Schweizer, and Sondermann to European contingent products whose underlying asset is modeled using a GARCH process and show that local risk-minimizing strategies with respect to the physical measure do exist, even though an associated minimal martingale measure is only available in the presence of bounded innovations. More importantly, since those local risk-minimizing strategies are in general convoluted and difficult to evaluate, we introduce Girsanov-like risk-neutral measures for the log-prices that yield more tractable and useful results. Regarding this subject, we focus on GARCH time series models with Gaussian innovations and we provide specific sufficient conditions concerning the finiteness of the kurtosis, under which those martingale measures are appropriate in the context of quadratic hedging. When this equivalent martingale measure is adapted to the price representation we are able to recover the classical pricing formulas of Duan and Heston and Nandi, as well as hedging schemes that improve the performance of those proposed in the literature.

Keywords: GARCH models; Hedging techniques; Methodology of optimal hedging; Incomplete markets

1. Introduction

GARCH models (Engle 1982, Bollerslev 1986, Ding et al. 1993) have been introduced into the modeling of the time series obtained from financial stock prices with the objective of capturing, via a parametric and parsimonious family of processes, several features that have been empirically documented and that escape more elementary modeling tools. For example, the constant variance and drift time series model that one obtains from strong Euler discretization of the log-normal model that underlies the Black, Merton, Scholes (BMS) option valuation formula (Black and Scholes 1972, Merton 1976) is not able to account for the volatility clustering in the time series of the associated returns or for the leptokurtosis (fat tails) in their distribution. Moreover, the oversimplification in modeling the stock returns is a source of contradiction concerning the implications of the BMS pricing formula, such as the smile-shaped curve that one observes when the implied or implicit volatility is plotted as a function of either moneyness or maturity.

From the modeling point of view, the GARCH family is successful at reproducing the above mentioned empirically observed features. Moreover, these models are particularly attractive from a mathematical point of view since the conditions for the existence of stationary solutions can be simply formulated and, additionally, most of the standard techniques in the time series literature concerning model selection and calibration can be adapted (see, for instance, Hamilton 1994 and Gourieroux 1997 and other standard references therein).

The situation becomes more complicated when we try to price contingent products whose underlying asset is assumed to be a realization of a GARCH process. The discrete-time character of the model, together with the infinite state space usually assumed for the innovations, makes the associated market automatically incomplete, in the sense that there are payoffs that cannot be replicated using a self-financing portfolio consisting only of a bond and a risky asset. This difficulty has been treated extensively in the literature using various approaches.

The first way to address this problem (Duan 1995, Heston and Nandi 2000) consists of adding a term to the GARCH model in the spirit of the NGARCH and VGARCH models introduced by Engle (1993); when the
where we spell out the conditions under which there exists a hedging scheme in the GARCH context and the second result that shows the availability of the quadratic hedging approach developed by Øksendal and Taqqu (1998 and references therein). It is worth mentioning that tackling the problem in this way, Kallsen and Taqqu (1998) obtain results that are consistent with those of Duan (1995) as far as the pricing formula is concerned, but disagree on the associated hedging strategies (see Garcia and Renault 1999 for a discussion).

In this paper we focus on the hedging side of the problem and implement in the GARCH setting the quadratic hedging approach developed by Föllmer, Schweizer, and Sondermann (see Föllmer and Sondermann 1986, Föllmer and Schweizer 1990 and Schweizer 2001 and references therein). Given a probability measure, the theory developed in the quoted papers gives a prescription for the construction of a generalized trading strategy that minimizes the local quadratic hedging error. Quadratic hedging techniques are subject to improvement since they do not make a difference between hedging shortfalls and windfalls, which should obviously be treated differently as far as the associated risk is concerned. Even though this point has been addressed in a variety of studies (see Pham 2000 and references therein) the associated hedging and pricing problem is more convoluted; we will hence defer the use of these techniques in the GARCH context to future work.

The contents of the paper are organized as follows. Section 2 contains a quick review of the GARCH models as well as the notions on quadratic hedging that are used later in the paper. The last part of this section contains the first result that shows the availability of the quadratic hedging scheme in the GARCH context and the second where we spell out the conditions under which there exists a minimal martingale measure; whenever this measure exists, the value process of the local risk-minimizing strategy (with respect to the physical measure) admits an interpretation as an arbitrage-free price for the derivative product we are dealing with. Unfortunately, the range of situations in which the minimal martingale measure exists is rather limited and, as we will see, is constrained to GARCH models with bounded innovations; this limitation is, from the modeling point of view, not always appropriate. Moreover, the expressions that determine the optimal hedging strategy using this measure are convoluted and hence of limited practical applicability.

The situation we have just described motivated us to carry out in section 3 a local risk minimization program for a well chosen Girsanov-like equivalent martingale measure. We implement this program for GARCH models whose innovations are Gaussian. Quadratic hedging with respect to a martingale measure yields much simpler expressions, admits a clear arbitrage-free pricing interpretation and, additionally, the corresponding strategies minimize not only the local risk and the quadratic risk, but also the so-called remaining conditional risk (these concepts will be defined later). Moreover, we will prove that a linear Taylor expansion in the drift term of the local risk-minimizing value process with respect to this martingale measure coincides with the same expansion calculated with respect to the physical measure; consequently, since, in most cases, the drift term is very small, one can safely compute the risk-minimizing strategy with respect to the martingale measure, which is much more convenient, and one obtains practically the same value had one used the much more convoluted expressions in terms of the physical measure.

Even though the equivalent martingale measure always exists, the quadratic hedging scheme requires that the process modeling the underlying asset be square summable with respect to the pricing measure. In theorem 3.1 we show that a sufficient (but not necessary) condition for that to hold is the finiteness of the kurtosis with respect to the physical measure. It is worth mentioning that, with this change of measure, the independent Gaussian innovations of the original GARCH process automatically remain independent and Gaussian after risk neutralization and there is no need to impose this feature as an additional condition (compare with, for example, assumption 2 of Heston and Nandi 2000).

The developments explained above are formulated using log-prices as the risky asset. In section 3.1 we reformulate these results in the price representation using predictable drift terms. This degree of generality allows us to recover in section 3.2 the classical pricing formulas of Duan (1995) and Heston and Nandi (2000) in the context of the local risk-minimization scheme. It is worth mentioning that, apart from providing an alternative interpretation to existing pricing formulas and improving some theoretical issues (such as, for example, dropping the assumptions invoked by Heston and Nandi 2000 regarding the Gaussian nature of the innovations after risk neutralization), the scheme that we propose comes together with a hedging scheme that performs better than that proposed by Duan (1995) and is simply not available in the case of Heston and Nandi (2000). More explicitly, local risk-minimizing hedging strategies with respect to a martingale measure minimize by construction the mean square hedging error, when this error is measured with that martingale measure, and are hence from this point of
view preferable to the hedging ratios proposed by Duan (1995) and Kalens and Taqqu (1998). Even though there are no theoretical results that guarantee that this difference in hedging performance still exists when we move to the physical probability, we carry out numerical tests in section 3.3. that seem to indicate that this is indeed the case.

### 1.1. Conventions and notation

Throughout this paper we will use the riskless asset as numéraire in order not to use the riskless interest rate in our expressions. Given a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_n\}_{n \in \mathbb{N}})\) and \(X, Y\) two random variables, we will denote by \(E_n[X] := E[X | \mathcal{F}_n]\) the conditional expectation, \(\text{cov}_n(X, Y) := \text{cov}(X, Y | \mathcal{F}_n) := E_n[XY] - E_n[X]E_n[Y]\) the conditional covariance, and \(\text{var}_n(X) := E_n[X^2] - E_n[X]^2\) the conditional variance. A discrete-time stochastic process \(\{X_n\}_{n \in \mathbb{N}}\) is predictable when \(X_n\) is \(\mathcal{F}_{n-1}\)-measurable, for any \(n \in \mathbb{N}\).

#### 2. The GARCH family and pricing by local risk minimization

In this section we introduce the general family of times series that we will use for the modeling of the stock prices. We then briefly review the basics of quadratic hedging, and we finally prove the existence of this kind of strategy in the GARCH context.

##### 2.1. The GARCH models

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and \(\{\epsilon_n\}_{n \in \mathbb{N}} \sim \text{IIDN}(0,1)\) a sequence of zero-mean, square integrable, independent, and identically distributed random variables. We will denote by \(\{\mathcal{F}_n\}_{n \in \mathbb{N}}\) the filtration of \(\mathcal{F}\) generated by the elements of this family, that is \(\mathcal{F}_n := \sigma(\epsilon_1, \ldots, \epsilon_n), n \geq 1\), is the \(\sigma\)-algebra generated by \(\{\epsilon_1, \ldots, \epsilon_n\}\). We will assume that \(\mathcal{F}_0\) consists of \(\Omega\) and all the negligible events in \(\mathcal{F}\).

GARCH models were introduced by Bollerslev (1986) as a parsimonious generalization of the ARCH models used by Engle (1982) in the modeling of the dynamics of inflation in the UK. This parametric family has been modified in various forms to make it suitable for the modeling of stock prices and contains an important number of different models. Even though the treatment that we will carry out in the rest of the paper is valid for all the models in the literature, we will now pick one of them, namely the one introduced by Ding et al. (1993), to illustrate the features of these models that will be of relevance in the rest of the paper.

##### 2.1.1. An example: The asymmetric GARCH model

Let \(\{S_n\}_{n \in \mathbb{N}}\) be the sequence that describes the price of the risky asset that we are interested in. The asymmetric GARCH\((p, q)\) model (Ding et al. 1993; He and Terasvirta 1999) determines the dynamics of the prices \(\{S_n\}_{n \in \mathbb{N}}\) by prescribing the following dynamics for the log-returns \(r_n := \log(S_n/S_{n-1})\), which gives a recursive relation for the log-prices \(s_n := \log(S_n)\):

\[
r_n = s_n - s_{n-1} = \mu + \alpha_n \epsilon_n, \quad \mu \in \mathbb{R},
\]

where \(\alpha_n \geq 0 \) and \(|\gamma| < 1\) are constant real coefficients. Note that the fact of working with log-prices implies that the price process \(\{S_n\}_{n \in \mathbb{N}}\) is measured for all \(n \in \mathbb{N}\).

\[
\sigma_n^2 = \omega + \sum_{i=1}^{p} \alpha_i (\tau_{n-i} - \gamma \tau_{n-i})^2 + \sum_{i=1}^{q} \beta_i \sigma_{n-i}^2,
\]

where \(\tau_n = r_n - E[r_n] = r_n - \mu\), \(\{\epsilon_n\}_{n \in \mathbb{N}} \sim \text{IIDN}(0,1)\), \(\omega > 0\), and \(\alpha_i, \beta_i \geq 0 \) and \(|\gamma| < 1\) are constant real coefficients. Note that the fact of working with log-prices implies that the price process \(\{S_n\}_{n \in \mathbb{N}}\) determined by (2.1) and (2.2) is always positive. The parameter \(\gamma\) controls the asymmetric influence of shocks: If they are positive, negative past shocks increase the variance more than comparable positive shocks. This is an empirically observed feature of stock markets. The following proposition, the proof of which is sketched in the appendix, characterizes the constraints on the model parameters that ensure the existence and uniqueness of a weakly stationary solution for (2.1) and (2.2).

**Proposition 2.1:** Suppose that \(\omega > 0\), \(\alpha_i, \beta_i \geq 0 \) and \(|\gamma| < 1\). Then model (2.1) and (2.2) admits a unique weakly (second-order) stationary solution if and only if

\[
(1 + \gamma^2)(\alpha_1 + \cdots + \alpha_p) + \beta_1 + \cdots + \beta_q < 1,
\]

in which case

\[
\text{var}(r_n) = E[\sigma_n^2] = \frac{\omega}{1 - (1 + \gamma^2)(\alpha_1 + \cdots + \alpha_p) - (\beta_1 + \cdots + \beta_q)}.
\]
where $\otimes$ denotes the Kronecker product of matrices, $\rho(B) = \max\{|\text{eigenvalues of the matrix } B|\}$, $A$ is the matrix given by

$$A = \begin{pmatrix}
\alpha_1 Z_t & \cdots & \alpha_p Z_t & \beta_1 Z_t & \cdots & \beta_q Z_t

\vdots & \ddots & \vdots & \vdots & \ddots & \vdots

\alpha_1 & \cdots & \alpha_p & \beta_1 & \cdots & \beta_q

I_{p \times 1} & \cdots & I_{p \times 1} & 0_{(p-1) \times q} & \cdots & 0_{(p-1) \times q}

0_{(q-1) \times p} & \cdots & 0_{(q-1) \times p} & I_{(q-1) \times (q-1)} & \cdots & I_{(q-1) \times (q-1)}
\end{pmatrix},$$

$I_{p \times p}$ is the $p \times p$ identity matrix, and $Z_t := ((x_t - \gamma) - \gamma)^2$. For $m = 1$, condition (2.5) is the same as (2.3). The kurtosis is finite whenever (2.5) holds with $m = 2$. For example, in the case of a GARCH(1,1) model, (2.5) amounts to the following inequality relation among the model parameters:

$$\beta^2 + 2\beta(1 + \gamma^2) + 3\sigma^2[(1 + \gamma^2)^2 + 4\gamma^2] < 1.$$

Ling and McAleer (2002a) report the corresponding characterization for the finiteness of the kurtosis of other asymmetric GARCH(1,1) processes (such as GJR-GARCH) or those driven by non-normal innovations.

### 2.1.3. Volatility clustering and leptokurtosis

GARCH is successful in capturing these two features that one empirically observes in stock market log-returns. Actually, in the GARCH context, these two notions are intimately related in the sense that one can say that heteroscedasticity (volatility clustering) causes leptokurtosis (heavy tails) and vice versa. Indeed, since we are using Gaussian innovations, we have

$$E_{n-1}[\sigma_n^4 e_n^4] = 3\sigma_n^4 = 3(E_{n-1}[\sigma_n^2 e_n^2])^2.$$

This allows us to write the kurtosis (standardized fourth moment) as

$$\kappa = \frac{E[\sigma_n^4 e_n^4]}{(E[\sigma_n^2 e_n^2])^2} = \frac{3E[\sigma_n^2 e_n^2]^2 + 3E[(E_{n-1}[\sigma_n^2 e_n^2])^2] - 3E[\sigma_n^2 e_n^2]^4}{(E[\sigma_n^2 e_n^2])^2}$$

$$= 3 + \frac{E[(E_{n-1}[\sigma_n^2 e_n^2])^2] - E[E_{n-1}[\sigma_n^2 e_n^2])^2]}{(E[\sigma_n^2 e_n^2])^2}$$

$$= 3 + \frac{\text{var}(E_{n-1}[\sigma_n^2 e_n^2])}{(E[\sigma_n^2 e_n^2])^2} = 3 + \frac{\text{var}(\sigma_n^2)}{(E[\sigma_n^2])^2},$$

where $E[\sigma_n^2]$ is determined by (2.4). Note that this expression, due to Gourieroux, proves that the excess kurtosis is positive whenever the variance of the volatility is non-zero.

### 2.1.4. General GARCH models

The results that we will prove in this paper apply beyond time series models that follow exactly the functional prescription determined by expressions (2.1) and (2.2). In our discussion it will be enough to assume that the log-prices evolve according to

$$\log \left( \frac{S_n}{S_{n-1}} \right) = s_n = s_{n-1} = \mu_n + \sigma_n \varepsilon_n, \quad (2.6)$$

$$\sigma_n^2 = \sigma_n^2(\sigma_{n-1}, \ldots, \sigma_{n-\max(p,q)}, \varepsilon_{n-1}, \ldots, \varepsilon_{n-q}), \quad (2.7)$$

where $\{\varepsilon_n\}_{n \in \mathbb{N}} \sim \text{IID}(0, \sigma^2_n)$, $\{\mu_n\}_{n \in \mathbb{N}}$ is a predictable process (that is $\mu_n$ is measurable with respect to $\mathcal{F}_{n-1} := \sigma(\varepsilon_1, \ldots, \varepsilon_{n-1})$, for all $n \in \mathbb{N}$), and the function $\sigma_n^2(\sigma_{n-1}, \ldots, \sigma_{n-\max(p,q)}, \varepsilon_{n-1}, \ldots, \varepsilon_{n-q})$ is constructed so that the following two conditions hold:

- **GARCH1**: There exists a constant $\omega > 0$ such that $\sigma_n^2 \geq \omega$.
- **GARCH2**: The process $\{\sigma_n^2\}_{n \in \mathbb{N}}$ is weakly (autocovariance) stationary.

A process $\{\sigma_n^2\}_{n \in \mathbb{N}}$ determined by (2.6) and (2.7) will be generically called a GARCH($p,q$) process. Note that (2.7) implies that the time series $\{\sigma_n^2\}_{n \in \mathbb{N}}$ is predictable; this feature is the main difference between GARCH and the so-called stochastic volatility models.

### 2.2. Local risk-minimizing strategies

In the following paragraphs we briefly review the necessary concepts on pricing by local risk minimization that we will need in the sequel. The reader is encouraged to check with chapter 10 of the excellent monograph of Föllmer and Schied (2004) for a self-contained and comprehensive presentation of the subject.

Let $H(S_T)$ be a European contingent claim that depends on the terminal value of the risky asset $S_T$. In the context of an incomplete market, it will in general be impossible to replicate the payoff $H$ by using a self-financing portfolio. Therefore, we introduce the notion of a generalized trading strategy, in which the possibility of additional investment in the numéraire asset throughout the trading periods up to expiry time $T$ is allowed. All the following statements are made with respect to a fixed filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.

**Definition 2.2**: A generalized trading strategy is a pair of stochastic processes $(\xi_t^0, \xi_t^1)$ such that $\{\xi_t^0\}_{t \in [0,T]}$ is adapted and $\{\xi_t^1\}_{t \in [0,T]}$ is predictable. The value process $V$ of $(\xi_t^0, \xi_t^1)$ is defined as

$$V_0 := \xi_0, \quad V_n := \xi_n^0 + \xi_n^1 \cdot S_n, \quad n \geq 1.$$

The gains process $G$ of the generalized trading strategy $(\xi_t^0, \xi_t^1)$ is given by

$$G_0 := 0, \quad G_n := \sum_{k=1}^{n} \xi_k \cdot (S_k - S_{k-1}), \quad n = 0, \ldots, T,$$

and the cost process $C$ is defined by the difference

$$C_n := V_n - G_n, \quad n = 0, \ldots, T.$$

It is easy to check that the strategy $(\xi_t^0, \xi_t^1)$ is self-financing if and only if the value process takes the form

$$V_0 = \xi_0^0 + \xi_1^1 \cdot S_0, \quad V_n = V_0 + \sum_{k=1}^{n} \xi_k \cdot (S_k - S_{k-1})$$

$$= V_0 + G_n, \quad n = 1, \ldots, T, \quad (2.8)$$

or, equivalently, if

$$V_0 = C_0 = C_1 = \cdots = C_T. \quad (2.9)$$
Definition 2.3: Assume that both $H$ and $\{S_n\}_{n\in\{0,\ldots,T\}}$ are $L^2(\Omega,\mathbb{P})$. A generalized trading strategy is called admissible for $H$ whenever it is in $L^2(\Omega,\mathbb{P})$ and its associated value process is such that

$$V_T = H, \quad \mathbb{P}\text{-a.s.}, \quad V_t \in L^2(\Omega,\mathbb{P}), \text{ for each } t,$$

and its gain process $G_t \in L^2(\Omega,\mathbb{P})$, for each $t$.

The next definition introduces the strategies we are interested in.

Definition 2.4: The local risk process of an admissible strategy $(\xi^0, \xi)$ is the process

$$R_t(\xi^0, \xi) := E_t[(C_{t+1} - C_t)^2], \quad t = 0, \ldots, T - 1.$$

The admissible strategy $(\xi^0, \hat{\xi})$ is called local risk-minimizing if

$$R_t(\xi^0, \hat{\xi}) \leq R_t(\xi^0, \xi), \quad \mathbb{P}\text{-a.s.},$$

for all $t$ and each admissible strategy $(\xi^0, \xi)$.

It can be shown that (Föllmer and Schied 2004, theorem 10.9) an admissible strategy is local risk-minimizing if and only if the cost process is a $\mathbb{P}$-martingale and it is strongly orthogonal to $S$, in the sense that $\text{cov}_t(S_{n+1} - S_n, C_{n+1} - C_t) = 0$, $\mathbb{P}$-a.s., for any $t = 0, \ldots, T - 1$. An admissible strategy whose cost process is a $\mathbb{P}$-martingale is usually referred to as mean self-financing (recall (2.9) for the reason behind this terminology). Once a probability measure $\mathbb{P}$ has been fixed, if there exists a local risk-minimizing strategy $(\xi^0, \hat{\xi})$ with respect to it, then it is unique (Föllmer and Schied 2004, proposition 10.9) and the payoff $H$ can be decomposed as (Föllmer and Schied 2004, corollary 10.14)

$$H = V_0 + G_T + L_T, \quad (2.10)$$

with $G_n$ the gains process associated with $(\hat{\xi}, \hat{\xi})$ and $L_n := C_n - C_0, \quad n = 0, \ldots, T$. Since $(\xi^0, \hat{\xi})$ is local risk-minimizing, the sequence $(L_n)_{n\in\{0,\ldots,T\}}$ that we will call the global (hedging) risk process, is a square integrable martingale that is strongly orthogonal to $S$ and that satisfies $L_0 = 0$. The decomposition (2.10) and (2.8) shows that $L_T$ measures how far $H$ is from the terminal value of the self-financing portfolio uniquely determined by the initial investment $V_0$ and the trading strategy $\xi$ (Lamberton and Lapeyre 2008, proposition 1.1.3).

2.3. Local risk minimization in the GARCH context and minimal martingale measures

As pointed out in the previous section, the local risk-minimization approach to hedging demands selecting a particular probability measure in the problem. Given a contingent product on a GARCH-driven risky asset, the physical probability measure is the most conspicuous one since, from the econometrics point of view, it is the measure naturally used to calibrate the model.

Our next proposition shows that a local risk-minimizing strategy with respect to the physical measure does exist in the GARCH context. Given the specific form of (2.6) and (2.7), it is more convenient to reformulate the problem by finding a local risk-minimizing strategy in which we take the log-prices $s_n$ as the risky asset and $h(s_T) := H(\exp(s_T))$ as the payoff function.

Proposition 2.5: Consider a market with a single risky asset that evolves in time according to a GARCH process satisfying (2.6) and (2.7), driven by innovations $\{\epsilon_n\}_{n\in\mathbb{N}} \sim \text{IID}(0, \sigma_n^2)$. Suppose also that the GARCH process has a bounded drift $(\mu_n)_{n\in\mathbb{N}}$, that is $\mu_n < B$ for some $B \in \mathbb{R}$ and for all $n \in \mathbb{N}$. Let $h \in L^2(\Omega, \mathbb{P}, F_T)$ be a contingent product on $s = \log(S)$. Then there exists a unique local risk-minimizing strategy for $h$ with respect to the physical measure $\mathbb{P}$, uniquely determined by the following recursive relations:

$$\xi_k = \frac{1}{\sigma_k \sigma_{k+1}} E_{k-1} \left[ h \left( 1 - \frac{\mu_T}{\sigma_T} \epsilon_T \right) \left( 1 - \frac{\mu_{T-1}}{\sigma_{T-1}} \epsilon_{T-1} \right) \ldots \right. \left. \times \left( 1 - \frac{\mu_{k+1}}{\sigma_{k+1}} \epsilon_{k+1} \right) \right], \quad k = 1, \ldots, T - 1,$$

$$\hat{\xi}_T = \frac{1}{\sigma_T} E_{T-1}[h \epsilon_T], \quad (2.11)$$

$$V_k = E_k \left[ h \left( 1 - \frac{\mu_T}{\sigma_T} \epsilon_T \right) \left( 1 - \frac{\mu_{T-1}}{\sigma_{T-1}} \epsilon_{T-1} \right) \ldots \right. \left. \times \left( 1 - \frac{\mu_{k+1}}{\sigma_{k+1}} \epsilon_{k+1} \right) \right], \quad k = 0, \ldots, T - 1,$$

$$V_T = h. \quad (2.13)$$

The position on the riskless asset is given by $\xi^0_k := V_k - \hat{\xi}_k s_k$.

Remark 1: The condition on the boundedness of the drift is verified for most models considered in the literature in the context of derivatives pricing. In some instances the drift is just a constant (see, for example, Barone-Adesi et al. 2007). Another important situation where this hypothesis trivially holds is the model of Duan (1995), where the drift is given by $\mu_n := r + \lambda \sigma_n - \frac{1}{2} \sigma_n^2$, where $r$ and $\lambda$ are positive constants that account for the continuously compounded one time step risk-free interest rate and the unit risk premium, respectively. In this situation, it is clear that $\mu_n < r + \frac{1}{2} \lambda^2$ for any $n \in \mathbb{N}$.

Proof (proof of proposition 2.5): We start by noticing that since $\sigma_n^2$ is $F_{n-1}$-measurable, the relations (2.6) and (2.7) imply

$$E_{n-1}[\sigma_n s_n] = 0, \quad (2.15)$$

$$E_{n-1}[s_n - s_{n-1}] = \mu_n, \quad (2.16)$$

$$E_{n-1}[(s_n - s_{n-1})^2] = \mu_n^2 + \sigma_n^2 \sigma_n^2, \quad (2.17)$$

$$\text{var}_{n-1}[s_n - s_{n-1}] = \sigma_n^2 \sigma_n^2, \quad (2.18)$$

for any $n \in \{1, \ldots, T\}$. 


The first fact that we need to check is that the GARCH context fits the framework established by definition 2.3 to carry out hedging by local risk minimization. More explicitly, we have to verify that the log-prices $s$ are square integrable. This is a consequence of hypothesis GARCH2; indeed, for any $n \in \{1, \ldots, T\}$, $s_n = s_0 + \sum_{i=1}^{\sigma} \mu_i + \sigma_1 \epsilon_1 + \cdots + \sigma_n \epsilon_n$. Then

$$E[s_n^2] = E \left[ \left( s_0 + \sum_{i=1}^{n} \mu_i \right)^2 + \sum_{i=1}^{n} \sigma_i^2 \epsilon_i^2 + \sum_{i<j}^{n} \sigma_i \sigma_j \epsilon_i \epsilon_j \right].$$

Let $i < j$, then $E[\sigma_i \sigma_j \epsilon_i \epsilon_j] = E[\epsilon_i \epsilon_j] = E[\sigma_i \epsilon_i] E[\epsilon_j] = 0$. Since by hypothesis GARCH2, $E[\sigma_i^2 \epsilon_i^2] < \infty$ and the drift is bounded, we have that

$$E[s_n^2] = E \left[ \left( s_0 + \sum_{i=1}^{n} \mu_i \right)^2 + \sum_{i=1}^{n} E[\sigma_i^2 \epsilon_i^2] \leq (s_0 + nB)^2 + \sum_{i=1}^{n} E[\sigma_i^2 \epsilon_i^2] < \infty,$$

as required.

Now, according to Föllmer and Schied (2004, proposition 10.10) the existence and uniqueness of a local risk-minimizing strategy is guaranteed as long as we can find a constant $C$ such that $(E_{n-1}[s_n - s_{n-1}])^2 \leq C \cdot \text{var}_{n-1}[s_n - s_{n-1}]$, $\mathbb{P}$-a.s. for any $n$. In our case it suffices to take $C = B^2 / (\sigma^2 \omega)$. Indeed, with this choice and using (2.16) and (2.18),

$$\frac{(E_{n-1}[s_n - s_{n-1}])^2}{\text{var}_{n-1}[s_n - s_{n-1}]} = \frac{\mu^2}{\sigma^2 \sigma_n^2} \leq \frac{B^2}{\sigma^2 \omega} = C,$$

(2.19)

as required. The recursions (2.11)-(2.14) follow by rewriting expression (10.5) of Föllmer and Schied (2004) using the equalities (2.15)-(2.18).

Expressions (2.11)-(2.14) are convoluted and difficult to evaluate. Moreover, expression (2.13) does not allow us to interpret $V_k$ as an arbitrage free price for $h$ at time $k$. There are two possibilities to get around this problem: the first consists of dropping the physical probability and of choosing instead an equivalent martingale measure that has particularly good properties that make it a legitimate proxy for the original measure. This is the path that we will take in the next section.

As an alternative, one may want to look for an equivalent martingale measure for which the value process of the local risk-minimizing strategy with respect to the physical measure can be interpreted as an arbitrage free price for $h$. This is the motivation for introducing the so-called minimal martingale measure. This measure is defined as a martingale measure $\tilde{\mathbb{P}}$ that is equivalent to the physical probability $\mathbb{P}$ and satisfies the following two conditions: $E[\mathbb{P} \log \mathbb{P}] < \infty$ and every $\mathbb{P}$-martingale $M \in L^2(\mathbb{P})$ that is strongly orthogonal to the price process $s$ is also a $\tilde{\mathbb{P}}$-martingale. This measure satisfies an entropy minimizing property (Schweizer 2001, proposition 3.6) and if $\hat{E}$ denotes the expectation with respect to $\tilde{\mathbb{P}}$, then the value process $V_k$ in (2.13) can be expressed as (Föllmer and Schied 2004, theorem 10.22)

$$V_k = \hat{E}[h],$$

which obviously yields the interpretation that we are looking for.

As we see in the next proposition, minimal martingale measures exist in the GARCH context only when the innovations are bounded (for example, when the innovations are multinomial) and certain inequalities among the model parameters are respected.

**Proposition 2.6:** Using the same setup as in proposition 2.5, suppose that the innovations in the GARCH model are bounded, that is there exists $K > 0$ such that $\epsilon_k < K$, for all $k = 1, \ldots, T$, and that this bound is such that $K < \sigma^2 \sqrt{\omega} / B$, with $\omega > 0$ the constant such that $\sigma^2 \omega \geq 1$ (see condition GARCH1) and $B \in \mathbb{R}$ the upper bound for the drift. Then there exists a unique minimal martingale measure $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$. Conversely, if there exists a minimal martingale measure, then the innovations in the model are necessarily bounded. Whenever the minimal martingale measure exists, its Radon–Nikodym derivative is given by

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \prod_{k=1}^{T} \left( 1 - \frac{\mu_k \epsilon_k}{\sigma^2 \sigma_k} \right).$$

(2.20)

**Proof:** We start by recalling that, in the proof of proposition 2.5, we showed in (2.19) the existence of a constant $C$ such that

$$(E_{n-1}[s_n - s_{n-1}])^2 \leq C \cdot \text{var}_{n-1}[s_n - s_{n-1}],$$

for all $t = 1, \ldots, T$. In view of this and theorem 10.30 of Föllmer and Schied (2004), the existence and uniqueness of a minimal martingale measure $\tilde{\mathbb{P}}$ is guaranteed provided that the following inequality holds:

$$(s_n - s_{n-1})E_{n-1}[s_n - s_{n-1}] < E_{n-1}[(s_n - s_{n-1})^2].$$

(2.21)

By (2.16) and (2.17), this inequality is equivalent to $\epsilon_n < \sigma_n \mu_n$ and it obviously holds if the innovations are bounded and the bound satisfies $K < \sigma^2 \sqrt{\omega} / B$. Conversely, suppose that there exists a minimal martingale measure; corollary 10.29 of Föllmer and Schied (2004) implies that (2.21) holds and hence so does $\epsilon_n < \sigma_n \mu_n$. Given that $\sigma_n$ and $\mu_n$ are $\mathbb{F}_{n-1}$-measurable and $\epsilon_n$ is $\mathbb{F}_n$-measurable, this equality can only possibly hold whenever the innovations $\epsilon_n$ are bounded.

As to expression (2.20), it is a consequence of corollary 10.29 and theorem 10.30 of Föllmer and Schied (2004). According to those two results, the density $d\tilde{\mathbb{P}}/d\mathbb{P}$ is the evaluation at $T$ of the $\mathbb{P}$-martingale

$$Z_t := \prod_{k=1}^{T} (1 + \lambda_k \cdot (y_k - y_{k-1})),$$

where $\lambda_k := -E_{k-1}[s_k - s_{k-1}] \text{var}_{k-1}[s_k - s_{k-1}]$ and $y_k$ is the martingale part in the Doob decomposition of $s_k$ with respect to $\mathbb{P}$. Therefore, we have

$$y_k - y_{k-1} = s_k - s_{k-1} - E_{k-1}[s_k - s_{k-1}].$$
Using (2.16) and (2.18) in these expressions, the result follows.

3. GARCH with Gaussian innovations

The hedging strategies that come out of (2.11)–(2.14) are, in general, difficult to compute either explicitly or by Monte Carlo methods. Moreover, the interpretation of the values of the resulting local risk-minimizing portfolio as an arbitrage-free price for \( h \) needs a minimal martingale measure whose existence is not always available.

The approach that we take in this section consists of dropping the physical measure and of carrying out the local risk-minimizing program for a well chosen Girsanov-like equivalent martingale measure; we will justify below the use of this measure as a legitimate proxy for the physical probability. We will implement this program for GARCH models whose innovations are Gaussian, for which no minimal martingale measure exists according to proposition 2.6.

The use of a martingale measure for local risk minimization is particularly convenient since the formulas that determine the generalized trading strategy are particularly simple and admit a clear interpretation. Indeed, it is easy to show that, when written with respect to a martingale measure, the local risk-minimizing strategy is determined by

\[
\tilde{\xi}_k = \frac{1}{\sigma_k^2} E_{k-1}[h(\mu + \sigma_k \xi_k)], \quad k = 1, \ldots, T, \tag{3.22}
\]

\[
V_k = E_k[h], \quad k = 0, \ldots, T. \tag{3.23}
\]

The position on the riskless asset is given by \( \xi_0 \equiv V_0 - \tilde{\xi}_0 S_0 \). Moreover, local risk-minimizing trading strategies computed with respect to a martingale measure also minimize (Föllmer and Schied 2004, proposition 10.34) the so-called remaining conditional risk, defined as the process \( R^n(\theta, \xi) \equiv E_t[(C_T - C_t)^2] \), \( t = 0, \ldots, T \); this is generally not true outside the martingale setup (see Schweizer 2001, proposition 3.1 for a counterexample).

As we will see in proposition 3.2, apart from the computational convenience and the other arguments provided above, the chosen equivalent martingale measure has a particular legitimacy since a linear Taylor expansion in the drift term of the local risk-minimizing value process with respect to this measure coincides with the same expansion calculated with respect to the physical measure; consequently, since in most cases the drift term is very small, carrying out the risk-minimizing program with respect to the physical measure or the equivalent martingale measure that we introduce below yields virtually the same results.

Consider a GARCH process driven by Gaussian innovations, that is \( \{\epsilon_t\}_{t=1}^{\infty} \sim \text{IIDN}(0, 1) \). Since our intention is to carry out the quadratic hedging program, a challenge at the time of finding an equivalent martingale measure consists of making sure that, after the change of measure, we do not leave the square-summable category; as we will see in our next theorem, this will be ensured by working with processes with finite kurtosis.

Moreover, it is desirable that the innovations do not lose the Gaussian character in the new picture; this condition is sometimes imposed as a hypothesis (see, for example, assumption 2 of Heston and Nandi 2000). In the next theorem, this is naturally obtained as a consequence of the construction. The proof of the following result can be found in the appendix.

**Theorem 3.1:** Let \( (\Omega, \mathbb{P}, \mathcal{F}) \) be a probability space. Let \( \{\xi_0, s_1, \ldots, s_T\} \) be a GARCH process determined by a recursive relation of the type (2.6) and (2.7) and where the innovations \( \{\epsilon_t\}_{t=1}^{T} \sim \text{IIDN}(0, 1) \); let \( \mathcal{F}_t := \sigma(\epsilon_1, \ldots, \epsilon_t) \) be the associated filtration of \( \mathcal{F} \). Then:

(i) the process

\[
Z_n := \prod_{k=1}^{n} \exp \left( -\frac{\mu_k}{\sigma_k^2} \epsilon_k \right) \prod_{k=1}^{n} \exp \left( -\frac{1}{2} \frac{\mu_k^2}{\sigma_k^2} \right), \quad n = 1, \ldots, T,
\]

is a \( \mathbb{P} \)-martingale. If the drift process \( \{\mu_n\}_{n=0}^\infty \) is bounded, then \( \{Z_n\}_{n=0}^\infty \) is square summable;

(ii) \( Z_T \) defines an equivalent measure \( \mathbb{Q} \) such that \( Z_T = d\mathbb{Q}/d\mathbb{P} \);

(iii) the process

\[
\tilde{\epsilon}_n := \epsilon_n + \frac{\mu_n}{\sigma_n}, \quad n = 1, \ldots, T, \tag{3.24}
\]

forms an IIDN(0, 1) noise with respect to the new probability \( \mathbb{Q} \);

(iv) the log-prices \( \{s_0, s_1, \ldots, s_T\} \) form a martingale with respect to \( \mathbb{Q} \) and are fully determined by the relations

\[
s_n = s_0 + \sigma_1 \tilde{\epsilon}_1 + \cdots + \sigma_n \tilde{\epsilon}_n, \tag{3.25}
\]

\[
\sigma_n^2 = \tilde{\sigma}^2_n(\sigma_n, \ldots, \sigma_{n-\max(p,q)}, \tilde{\epsilon}_{n-1}, \ldots, \tilde{\epsilon}_{n-q}). \tag{3.26}
\]

The functions \( \tilde{\sigma}^2_n \) are the same as \( \sigma^2_n \) in (2.7) with \( \epsilon_{n-1}, \ldots, \epsilon_{n-q} \) written as a function of \( \tilde{\epsilon}_{n-1}, \ldots, \tilde{\epsilon}_{n-q} \) using (3.24). If the process \( \{\mu_n\}_{n=0}^\infty \) is bounded, then the martingale \( \{s_0, s_1, \ldots, s_T\} \) is square integrable with respect to \( \mathbb{Q} \); and

(v) the random variables in the process \( \{\sigma_n \tilde{\epsilon}_t\}_{t=1}^{T} \) are zero mean and uncorrelated with respect to \( \mathbb{Q} \).

**Remark 2:** The conclusion in part (iv) concerning the \( Q \)-square integrability of the martingale \( \{s_0, s_1, \ldots, s_T\} \) is very important, since it spells out clearly sufficient (but not necessary!) conditions, namely finite kurtosis and bounded drift under which we are entitled to use this new measure as a theoretical representation of the problem to compute hedging strategies via local risk minimization. In some market conditions, the finiteness of the kurtosis may be a rather restrictive condition, but it is well characterized in the GARCH context (see Li et al. 2002 and Ling and McAleer 2002a, b and references therein). The condition on the finiteness of the kurtosis can be weakened to require the process \( \{\sigma_n \tilde{\epsilon}_t\}_{t=1}^{T} \) to belong to \( L^{2+}\) for every possible \( (\Omega, \mathbb{P}, \mathcal{F}) \) with \( \epsilon > 0 \) arbitrarily small.
This result follows from using in the proof (available in the appendix) the fact that the elements of the process \( \{ Z_n \}_{n \in \mathbb{N}, 1 \leq T} \) do actually belong to \( L^q(\Omega, \mathbb{P}, F) \), for any \( q < \infty \), and by replacing the Cauchy–Schwarz inequality in (A16) by Hölder’s inequality.

The local risk-minimizing strategy associated with the martingale measure. Given a European claim \( H(S_T) \) on the risky asset \( S \), the martingale measure described in the previous theorem can be used to come up with a local risk-minimizing strategy by recasting the problem as a hedging problem where we consider the log-prices \( s_n \) as the risky asset and \( h(s_T) = H(\exp(s_T)) \) as the payoff function.

Suppose that the process \( \{ \sigma_n \varepsilon_n \}_{n \in \mathbb{N}, 1 \leq T} \) has finite kurtosis with respect to the physical measure \( \mathbb{P} \) and that the drift is bounded. Part (iv) of theorem 3.1 guarantees in that situation that the log-prices \( \{ s_n, s_T \} \) are square integrable martingales with respect to \( \mathbb{Q} \) and hence the local risk-minimization approach to hedging applies in this transformed setup. A straightforward computation using the elements in theorem 3.1 shows that, for any \( n \in \{ 1, \ldots, T \} \),

\[
\tilde{E}_{n-1}[s_n] = s_{n-1}, \quad \tilde{E}_{n-1}[\{ s_n - s_{n-1} \}^2] = \tilde{E}_{n-1}[\{ \sigma_n \varepsilon_n \}^2] = \sigma_n^2, \\
\hat{\varepsilon}_{n-1}[s_n - s_{n-1}] = \sigma_n^2.
\]

With these elements, the general local risk-minimizing strategy described in (2.11)–(2.14) becomes, with the use of this measure,

\[
\tilde{V}_k = \tilde{E}_k[h(s_T)], \quad k = 0, \ldots, T, \tag{3.27}
\]

\[
\hat{\varepsilon}_k = \frac{1}{\sigma_k} \tilde{E}_{k-1}[\varepsilon_k \tilde{V}_k] = \frac{1}{\sigma_k} \tilde{E}_{k-1}[\varepsilon_k \tilde{E}_k[h(s_T)]] = \frac{1}{\sigma_k} \tilde{E}_{k-1}[\varepsilon_k h(s_T)], \\
k = 1, \ldots, T, \tag{3.28}
\]

\[
L_T = C_T - C_0 = h(s_T) - \tilde{V}_0 - \sum_{k=1}^T \hat{\varepsilon}_k (s_k - s_{k-1}) = h(s_T) - \tilde{E}_1[h(s_T)] - \sum_{k=1}^T \hat{\varepsilon}_k \tilde{E}_{k-1}[\varepsilon_k h(s_T)]. \tag{3.29}
\]

The position on the riskless asset is given by \( \hat{\xi}_k := \tilde{V}_k - \hat{\varepsilon}_k \hat{\xi}_k \).

We conclude this section by showing that, since in practice the trend term \( \mu_n \) is usually very small when the time scale is days or weeks,\(^4\) the value process for the derivative product \( h \) obtained by risk minimization using the martingale measure just introduced and that computed using the physical measure are very similar. We make this explicit in the following statement, the proof of which is provided in the appendix.

**Proposition 3.2:** Let \( h \) be a derivative product whose underlying asset is modeled using a GARCH process with finite kurtosis and constant drift. Let \( V_k \) be the value process (2.13) of the local risk-minimizing strategy associated with \( h \) computed with the physical probability. Let \( \tilde{V}_k \) be the value process (3.27), this time computed with respect to the martingale measure introduced in theorem 3.1. The linear Taylor expansions of \( V_k \) and \( \tilde{V}_k \) in the drift term \( \mu \) coincide.

### 3.1. The martingale measure in the price representation and local risk minimization

The martingale measure presented in theorem 3.1 is constructed so that, after risk neutralization, log-prices become martingales driven by IIDN innovations. Obviously, this feature does not guarantee that the prices themselves share this attribute. However, theorem 3.1 can easily be modified so that an equivalent result is available for prices. Indeed, consider a European derivative product with payoff function \( H \), whose underlying asset is modeled using a general GARCH process as in (2.6) and (2.7), driven by IIDN(0, 1) innovations, that is

\[
\log \left( \frac{S_n}{S_{n-1}} \right) = s_n - s_{n-1} = \mu_n + \sigma_n \varepsilon_n, \tag{3.30}
\]

\[
\sigma_n^2 = \sigma_n^2, \tag{3.31}
\]

with \( \{ \varepsilon_n \}_{n \in \mathbb{N}} \sim \text{IIDN}(0, 1) \) and \( \{ \mu_n \}_{n \in \mathbb{N}} \) a predictable process. Rewrite (3.30) as

\[
s_n = s_{n-1} + \mu_n + \frac{1}{2} \sigma_n^2 = s_{n-1} + \mu_n + \frac{1}{2} \sigma_n^2 + \sigma_n \varepsilon_n = s_{n-1} + \frac{1}{2} \sigma_n^2 + \sigma_n \varepsilon_n, \tag{3.32}
\]

where

\[
\tilde{\mu}_n := \mu_n + \frac{1}{2} \sigma_n^2, \tag{3.33}
\]

and consider

\[
\tilde{\varepsilon}_n := \varepsilon_n - \frac{\tilde{\mu}_n}{\sigma_n}, \quad n = 1, \ldots, T. \tag{3.34}
\]

The proof of theorem 3.1 can be mimicked to show that the measure \( Q \) determined by the Radon–Nikodym derivative \( Z_T := dQ/d\mathbb{P} \) given by

\[
Z_T := \prod_{k=1}^T \exp \left( -\frac{\tilde{\mu}_k}{\sigma_k} \varepsilon_k \right) \exp \left( -\frac{1}{2} \frac{\tilde{\mu}_k^2}{\sigma_k^2} \right) \tag{3.34}
\]

is such that the process \( \{ \tilde{\varepsilon}_n \}_{n \in \{ 1, \ldots, T \} } \) is an IIDN(0, 1) noise with respect to it. Moreover, using \( \{ \tilde{\varepsilon}_n \}_{n \in \{ 1, \ldots, T \} } \) we have that (3.30) and (3.31) become

\[
s_n = s_{n-1} - \frac{1}{2} \sigma_n^2 + \sigma_n \tilde{\varepsilon}_n, \tag{3.35}
\]

\[
\sigma_n^2 = \sigma_n^2, \tag{3.36}
\]

As an example, consider the drift term \( \mu \) corresponding to the standard GARCH models with constant drift term historically calibrated to the daily log-returns of the following indices between the dates January 2, 2007 and December 31, 2008: Dow Jones Industrial Average, \(-6.99 \cdot 10^{-4}\); Nasdaq Composite, \(-8.53 \cdot 10^{-4}\); S&P 500, \(-8.94 \cdot 10^{-4}\); Euronext 100, \(-1.1 \cdot 10^{-3}\).
The importance of this modification is that the price process \( \{S_t\}_{t=0}^\infty \) forms a martingale with respect to \( Q \). Indeed, using the facts that \( S_0/S_{-1} = \exp(-\frac{1}{2}\sigma_n^2 + \sigma_n \tilde{e}_n) \), \( \{\tilde{e}_n\}_{n=1}^\infty \) is an IIDN(0, 1) noise with respect to \( Q \), and that \( \sigma_n \) is \( \mathcal{F}_{n-1} \)-measurable, we obtain

\[
\tilde{E}_{n-1} \left[ \frac{S_n}{S_{n-1}} \right] = \tilde{E}_{n-1} \left[ \exp \left( -\frac{1}{2} \sigma_n^2 + \sigma_n \tilde{e}_n \right) \right] = \exp \left( -\frac{1}{2} \sigma_n^2 \right) \tilde{E}_{n-1} \left[ \exp(\sigma_n \tilde{e}_n) \right] = 1,
\]

as required. The local risk-minimizing strategy associated with this martingale measure is easy to compute and is given by the expressions

\[
\tilde{V}_k = \tilde{E}_k[H], \quad k = 0, \ldots, T,
\]

and

\[
\tilde{\xi}_k = \frac{S_{k-1}}{\Sigma_k} \tilde{E}_{k-1} \left[ H \left( \exp \left( -\frac{1}{2} \sigma_k^2 + \sigma_k \tilde{e}_k \right) - 1 \right) \right], \quad k = 1, \ldots, T,
\]

where \( \Sigma_k^2 := \text{var}_{k-1}(S_k - S_{k-1}) = S_{k-1}^2 (\exp \tilde{e}_k - 1) \). The position on the riskless asset is given by \( \frac{\tilde{P}_k}{\tilde{E}_k} : = \tilde{V}_k - \tilde{\xi}_k S_k \).

### 3.2. Local risk minimization and the pricing formulas of Duan and Heston–Nandi

The martingale measure \( Q \) introduced in (3.34) can be used to provide an alternative risk-minimization interpretation to the pricing formula of Duan (1995), which was introduced using a utility maximization argument. The same holds for the formula introduced by Heston and Nandi (2000); in this case, our interpretation is even more valuable because, as far as I am aware, there is no other mathematical argument that supports the use of the formula suggested by the author, other than the fact that the result can be expressed using an extremely convenient closed-form expression. Moreover, after the change of measure, an additional hypothesis is needed by Heston and Nandi (2000) (see assumption 2, p. 590) in order to ensure that the innovations in the model remain IIDN after risk neutralization. We are able to drop that hypothesis since the IIDN character of those innovations is, in our case, part of the thesis of theorem 3.1.

As an additional bonus, in both cases the pricing via local risk-minimization comes together with an associated hedging strategy that does not exist for the model of Heston and Nandi (2000) and, as mentioned in the introduction, has been the subject of various conflicting proposals in the case of Duan’s model (see Garcia and Renault 1999 for a discussion).

By construction, the self-financing hedging strategy associated with the generalized local risk-minimizing strategy proposed here is also a mean–variance optimal strategy when the corresponding mean square hedging error is measured with the martingale measure. Nevertheless, the hedging risk is ‘left’ with the physical probability measure. Since there is no theoretical argument that proves that the optimality of the hedging ratios that we propose with respect to others in the literature survives when we change to that measure, we will carry out a numerical study in the next section that seems to indicate that this is indeed the case.

Formula (3.38), which provides the hedging ratios, generally needs to be evaluated via a Monte Carlo simulation. However, it is worth noting that it is readily available for any payoff \( H \), which yields a competitive advantage with respect to sensitivity methods that usually require the computation of derivatives of the option price and may prove to be an extremely convoluted task for exotic derivatives. The proof of the following proposition can be found in the appendix.

**Proposition 3.3:** The formulas proposed by Duan (1995) and Heston and Nandi (2000) for the pricing of European options that have underlying risky assets modeled via an N(NGARCH) and an asymmetric GARCH process, respectively, coincide with the prices of the local risk-minimizing strategies constructed using the martingale measures introduced in (3.34) associated with those two processes.

### 3.3. Numerical test of the hedging performance of the local risk-minimization (LRM) scheme

We see from proposition 3.3 that the local risk-minimizing approach recovers the pricing formulas of Duan (1995) and Heston and Nandi (2000) and we have mentioned that there is a general result (see, for example, Föllmer and Schied 2004, proposition 10.37) that ensures that the associated self-financing trading strategy (3.37)–(3.39) is variance optimal, that is

\[
\tilde{E} \left[ \left( H - \tilde{V}_0 - \sum_{k=1}^T \tilde{\xi}_k (S_k - S_{k-1}) \right)^2 \right] \leq \tilde{E} \left[ \left( H - V_0 - \sum_{k=1}^T \xi_k (S_k - S_{k-1}) \right)^2 \right],
\]

for any self-financing trading strategy \( \xi_1, \ldots, \xi_T \) with initial value \( V_0 \). Given that the hedging risk is measured with the physical probability, a question that needs to be addressed is the validity of the inequality (3.40) with respect to that measure, that is replacing the \( Q \)-expectation \( \tilde{E} \) by the \( P \)-expectation \( E \). An obvious continuity argument suggests that that statement is going to hold true provided that the modified drift term \( \{\tilde{m}_t\}_{t \in \mathbb{R}} \) is ‘sufficiently small’; however, there is no general result that we can invoke and we will hence carry out a numerical experiment to assess the performance of interest.

Since there is no proposal for hedging ratios in the work of Heston and Nandi (2000) we will limit our comparison to the hedging strategy proposal in Duan’s model for the European call option (Duan 1995, corollary 2.4) and the standard Black–Scholes model.

More specifically, we will consider a European call option whose underlying asset has as price \( \{S_t\}_{t \in \{0, \ldots, T\}} \) a realization of the model spelled out in (A21) and (A22)
Table 1. Mean square hedging errors (MSE) for European call options using the Black–Scholes scheme (BS), the Duan approach, and local risk-minimization (LRM). The MSEs are computed by simulating 1000 random price paths, using Duan’s model with respect to the physical probability; the Duan and LRM hedging ratios for each of the price paths are computed via Monte Carlo using each time 10,000 random paths simulated with the risk-neutralized version of Duan’s model, to which the EMS modification is subsequently applied. The standard deviations (Std) are obtained from 100 independent Monte Carlo hedging error estimates. The rows ‘Std A vs. B’ show the percentage increase in standard deviation experienced in the estimation when using method A instead of method B.

| $S_0/K$ | Maturity, six time steps | Maturity, 11 time steps | Maturity, 16 time steps | Maturity, 21 time steps |
|---------|--------------------------|-------------------------|------------------------|------------------------|
|         | 1.04 | 1 | 0.98 | 0.94 | 1.04 | 1 | 0.98 | 0.94 | 1.04 | 1 | 0.98 | 0.94 | 1.04 | 1 | 0.98 | 0.94 |
| **BE**  |      |   |     |     |      |   |     |     |      |   |     |     |      |   |     |     |
| MSE     | 0.0262 | 0.1681 | 0.1054 | 0.0111 | 0.0991 | 0.2402 | 0.1978 | 0.0457 | 0.1573 | 0.3250 | 0.2924 | 0.1124 | 0.2136 | 0.3803 | 0.3986 | 0.1545 |
| Std     | (0.0082) | (0.0118) | (0.0122) | (0.0071) | (0.0398) | (0.0237) | (0.0214) | (0.0188) | (0.0440) | (0.0271) | (0.0330) | (0.0327) | (0.0344) | (0.0460) | (0.1380) | (0.0285) |
| **Duan**|      |   |     |     |      |   |     |     |      |   |     |     |      |   |     |     |
| MSE     | 0.0253 | 0.1558 | 0.0998 | 0.0108 | 0.0894 | 0.2188 | 0.1844 | 0.0419 | 0.1457 | 0.2937 | 0.2675 | 0.1006 | 0.1999 | 0.3465 | 0.3411 | 0.1427 |
| Std     | (0.0063) | (0.0103) | (0.0106) | (0.0060) | (0.0233) | (0.0243) | (0.0173) | (0.0128) | (0.0290) | (0.0202) | (0.0239) | (0.0204) | (0.0301) | (0.0383) | (0.0428) | (0.0255) |
| **LRM** |      |   |     |     |      |   |     |     |      |   |     |     |      |   |     |     |
| MSE     | 0.0261 | 0.1544 | 0.0979 | 0.0101 | 0.0915 | 0.2201 | 0.1825 | 0.0409 | 0.1536 | 0.2965 | 0.2673 | 0.0991 | 0.2135 | 0.3545 | 0.3401 | 0.1404 |
| Std     | (0.0060) | (0.0091) | (0.0092) | (0.0052) | (0.0201) | (0.0221) | (0.0162) | (0.0112) | (0.0240) | (0.0184) | (0.0224) | (0.0183) | (0.0275) | (0.0356) | (0.0374) | (0.0240) |
| Std BS vs. Duan | 30.15% | 14.56% | 15.09% | 18.33% | 70.81% | -2.46% | 23.69% | 46.87% | 51.72% | 34.15% | 38.07% | 60.29% | 14.28% | 20.10% | 222.42% | 11.76% |
| Std BS vs. LRM | 36.66% | 29.67% | 32.60% | 36.53% | 98.00% | -7.23% | 32.09% | 67.87% | 83.33% | 47.28% | 47.32% | 78.68% | 25.09% | 29.21% | 268.98% | 18.75% |
| Std Duan vs. LRM | 5.00% | 13.18% | 15.21% | 15.38% | 15.92% | -9.95% | 6.79% | 14.28% | 20.83% | 9.78% | 6.69% | 11.47% | 9.45% | 7.58% | 14.43% | 6.25% |
with \( p = q = 1 \), \( a_0 = 0.00001 \), \( \alpha_1 = 0.2 \), \( \beta_1 = 0.7 \), and \( \lambda = 0.01 \). Additionally, the initial value equals \( S_0 = 100 \).

Our goal is to compare the three mean square hedging errors

\[
E\left[\frac{1}{C_0} (H - V_0 - \sum_{k=0}^T \hat{\xi}_k (S_k - S_{k-1}))^2\right],
\]

\[
E\left[\frac{1}{C_0} (H - V_0^B - \sum_{k=0}^T \hat{\xi}_k^B (S_k - S_{k-1}))^2\right],
\]

and

\[
E\left[\frac{1}{C_0} (H - V_0^{LRM} - \sum_{k=0}^T \hat{\xi}_k^{LRM} (S_k - S_{k-1}))^2\right],
\]

where:

- \((\hat{V}_0, \{\hat{\xi}_k\}_{k\in\{1, \ldots, T\}})\) is the self-financing trading strategy associated with the local risk-minimization scheme \((3.37)-(3.39)\) for Duan’s model with the parameter values just specified;
- \((V_0^B, \{\xi_k\}_{k\in\{1, \ldots, T\}})\) is the self-financing trading strategy associated with the Black–Scholes scheme for a log-normal model with constant volatility \(\sigma\) set equal to the stationary value of the volatility of Duan’s model under the physical probability, namely \(\sigma^2 = a_0/(1 - \alpha_1 - \beta_1)\); and
- \((\hat{V}_0^D, \{\hat{\xi}_k^D\}_{k\in\{1, \ldots, T\}})\) is the self-financing trading strategy given by Duan’s scheme and specified in corollary 2.4 of Duan (1995), namely

\[
\hat{\xi}_k^D := \frac{\hat{\xi}_k}{S_{k-1}} \left[ \frac{S_T}{S_{k-1}} 1_{X_k \in K} \right].
\]

where \(K\) is the exercise price of the European call option in question. By proposition 3.3, \(\hat{V}_0 = V^D\).

The mean square hedging errors are computed by simulating \(1000\) random price paths, using Duan’s model with respect to the physical probability \((A21)\) and \((A22)\), as the error is computed using the \(\mathbb{P}\)-expectation. Each of these price paths is hedged using the three different schemes listed above; the Duan and the local risk-minimization approaches require the estimation via Monte Carlo of the hedging ratios \((3.41)\) and \((3.38)\) at each time step. These \(Q\)-expectations are computed using each time \(10,000\) different random paths simulated with the risk-neutralized version \((A23)\) and \((A24)\) of Duan’s model. The martingale property is extremely important in this setup, consequently before the expectations are computed these paths are modified using the empirical martingale simulation (EMS) technique introduced by Duan and Simonato (1998) and Duan et al. (2001), which, additionally, reduces the variance of the estimation. The variance of the estimation of the hedging error is evaluated by repeating randomly \(100\) times the estimation of the mean square error.

The results of the numerical test are presented in table 1 and in figure 1, and show the following.

- Both Duan’s and the LRM hedging schemes yield a smaller mean square hedging error than the Black–Scholes scheme.

Figure 1. Mean square hedging errors for European call options using the Black–Scholes scheme (BS), the Duan approach, and local risk-minimization (LRM).
• The Duan and LRM hedging schemes produce mean square hedging errors that are not significantly very different, even though Duan performs slightly better than LRM for moneyness larger than one and LRM performs slightly better than Duan for moneyness smaller than one.

• The variance of the LRM estimation is always smaller than that obtained using the Duan and the Black–Scholes schemes.

4. Conclusions

In this paper, we have reported the applicability of the pricing/hedging scheme by local risk-minimization (LRM) to European options with one underlying asset that is modeled using a GARCH process. The main conclusions of the paper are the following:

• Local risk-minimizing strategies exist for options with square summable payoffs, even though it is basically only in the presence of models with bounded innovations that a minimal martingale measure is available in order to interpret the subsequent price as an arbitrage-free price.

• Since the conditions for the existence of a minimal martingale measure are too restrictive, we have carried out the LRM scheme with respect to a well-chosen equivalent martingale measure that produces the same results as the physical probability up to first order in the Taylor series expansion with respect to the drift.

• When this martingale measure is used in the context of the standard models of Duan (1995) and Heston and Nandi (2000) we recover their pricing formulas, which provides an alternative optimal hedging interpretation to the original utility maximization argument that motivated their introduction.

• The hedging strategy associated with the LRM scheme is additionally mean–variance optimal when the mean square hedging error is computed using the martingale measure.

• We have examined numerically if this hedging optimality survives when the error is measured using the physical probability by comparing the hedging performances of the standard Black–Scholes scheme, Duan’s scheme, and local risk minimization. Simulations show that the Duan and LRM hedging schemes produce mean square hedging errors that are not significantly very different, even though Duan performs slightly better than LRM for moneyness larger than one and LRM performs slightly better than Duan for moneyness smaller than one. The variance of the LRM estimation is always smaller than that obtained using both the Duan and the Black–Scholes schemes.

• The expression that provides the LRM hedging ratios generally needs to be evaluated via a Monte Carlo simulation, but it is readily available for any payoff $H$ and, unlike other hedging proposals based on sensitivity methods, it does not require the computation of derivatives of the option price, which may prove to be a very convoluted task for exotic derivatives.

Acknowledgements

I thank Josef Teichmann for generously spending his time in going carefully through the paper, giving it serious thought, and for providing me with relevant feedback that led to many improvements. I also profited from discussions with Stéphane Chrétien, who shared with me his vast knowledge of statistics and probability theory. Interesting discussions with Alberto Elices, Friedrich Hubalek, and Andreu Lázaro are also acknowledged, as well as the comments and criticisms of two anonymous referees that significantly improved the paper.

References

Barone-Adesi, G., Engle, R.F. and Mancini, L., A GARCH option pricing model in incomplete markets. Preprint, 2007.

Black, F. and Scholes, M., The valuation of option contracts and a test of market efficiency. J. Finance, 1972, 27, 399–417.

Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. J. Econometr., 1986, 31, 307–327.

Davis, M.H.A., Option pricing in incomplete markets. In Mathematics of Derivative Securities, edited by M.A.H. Dempster and S.R. Pliska, pp. 216–226, 1997 (Cambridge University Press: Cambridge).

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

Duan, J.-C., The GARCH option pricing model. Math. Finance, 1995, 5, 13–32.

Duan, J.-C., Augmented GARCH($p$, $q$) processes and its diffusion limit. J. Econometr., 1997, 79, 97–127.

Duan, J.-C., Gauthier, G. and Simonato, J.-G., Asymptotic distribution of the EMS option price estimator. Mgmt Sci., 2001, 47(8), 1122–1132.

Duan, J.-C. and Simonato, J.-G., Empirical martingale simulation of asset prices. Mgmt Sci., 1998, 44(9), 1218–1233.

Engle, R., Autoregressive conditional heteroskedasticity with estimates of the variance of the UK inflation. Econometrica, 1982, 50, 987–1008.

Engle, R. and Ng, V., Measuring and testing the impact of news on volatility. J. Finance, 1993, 43, 1749–1778.

Foata, D. and Fuchs, A., Calcul des Probabilités, 2nd ed., 2003 (Dunod).

Föllmer, H. and Schied, A., Stochastic Finance. An Introduction in Discrete Time, 2nd revised and extended ed., 2004 (De Gruyter Studies in Mathematics, vol. 27), 2004 (Walter de Gruyter: Berlin).

Föllmer, H. and Schweizer, M., Hedging of contingent claims under incomplete information. Appl. Stochast. Anal. London, Stochast. Monogr., 1990, 5, 389–414.

Föllmer, H. and Sondermann, D., Hedging of nonredundant contingent claims. In Contributions to Mathematical Economics, pp. 205–223, 1986 (North-Holland: Amsterdam).

García, R. and Renault, É., Note on hedging in ARCH-type option pricing models. Math. Finance, 1999, 8(2), 153–161.
Appendix A

Proof of proposition 2.1

Proof: The proof of the full statement in proposition 2.1 is lengthy and convoluted. The reader is encouraged to check with Ling and McAleer (2002b) and Li et al. (2002), and references therein. In the following we will content ourselves with checking that condition (2.3) implies the asymptotic weak stationarity of the solutions of the model and we will establish (2.4).

We start by noting that $E_n[r_n] = \mu = E[r_n]$, $\text{var}(r_n) = E_n[r_n] - E_n[r_n]^2 = \sigma_n^2$, and hence $\text{var}(r_n) = E[\text{var}(r_n)] + \text{var}(E_n[r_n]) = E[\sigma_n^2]$.

We now take expectations on both sides of (2.2), that is $\sigma_n^2 = \alpha + \sum_{i=1}^p \alpha_i (1 + \gamma^2) r_{n-i}^2 - 2 \gamma \alpha_i |r_{n-i}| r_{n-i} + \sum_{i=1}^q \beta_i \sigma_{n-i}^2$, taking into account that $E[r_n] = 0$, $E[r_n^2] = E[\sigma_n^2] = E[\sigma_n^2]$, and $E[|r_n|^2] = 0$. We obtain

$$E[\sigma_n^2] = \alpha + (1 + \gamma^2) A(L) E[\sigma_n^2] + B(L) E[\sigma_n^2],$$

where $A(L)$ and $B(L)$ are the polynomials $A(z) = \sum_{i=1}^p \alpha_i z^i$ and $B(z) = \sum_{i=1}^q \beta_i z^i$ on the one-step lag operator $L$. Equivalently,

$$E[\sigma_n^2] = \omega + [(1 + \gamma^2) A(L) + B(L)] E[\sigma_n^2].$$

This difference equation is stable (see, for instance, proposition 2.2 of Hamilton 1994, p. 34), that is it admits an asymptotic solution whenever the roots of the polynomial

$$1 - (1 + \gamma^2) A(z) - B(z) = 0 \quad (A1)$$

lie outside the unit circle, in which case expression (2.4) clearly holds. This condition on the roots of (A1) is equivalent to

$$(1 + \gamma^2) A(1) + B(1) < 1, \quad (A2)$$

which coincides with (2.3). Indeed, if $(1 + \gamma^2) A(1) + B(1) \geq 1$, we have that, since $(1 + \gamma^2) A(0) + B(0) = 0 < 1$, then (A1) necessarily has a real root between 0 and 1. Conversely, assume that (A2) holds and that $z_0$ is a root of (A1) such that $|z_0| < 1$. Then

$$1 = (1 + \gamma^2) A(z_0) + B(z_0) = \left(1 + \gamma^2\right) \sum_{i=1}^p \alpha_i z_0^i + \sum_{i=1}^q \beta_i z_0^i \leq (1 + \gamma^2) \sum_{i=1}^p \alpha_i |z_0|^i + \sum_{i=1}^q \beta_i |z_0|^i \leq (1 + \gamma^2) A(1) + B(1),$$

which contradicts our hypothesis.

Theorem 3.1

Proof: (i) We start by proving that

$$E[Z_n] = 1, \quad \text{for all } n = 1, \ldots, T. \quad (A3)$$

This equality will be needed later on and guarantees that $Z_n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Indeed, let $p(x) := (1/\sqrt{2\pi}) \exp(-x^2/2)$ be the standard normal distribution. Then

$$E[Z_n] = \int_{-\infty}^{\infty} dx_1 \cdots dx_n \exp \left(-\frac{\mu_1 x_1}{\sigma_1} - \frac{\mu_2}{2\sigma_2} \right) \cdots \times p(x_1) \cdots p(x_n)$$

$$= \int_{-\infty}^{\infty} dx_1 \cdots dx_{n-1} \exp \left(-\frac{\mu_1 x_1}{\sigma_1} - \frac{\mu_2}{2\sigma_2} \right) \cdots \times \frac{\mu_2}{\sigma_2(x_1, \ldots, x_{n-1})} \exp \left(-\frac{\mu_2}{2\sigma_2(x_1, \ldots, x_{n-1})} \right) \cdots \times p(x_{n-1}) \times p(x_n).$$

Given that

$$\int_{-\infty}^{\infty} dx_n \exp \left(-\frac{\mu_n x_n}{\sigma_n(x_{n-1}, \ldots, x_1)} - \frac{\mu^2_n}{2\sigma_n^2(x_{n-1}, \ldots, x_1)} \right) p(x_n) = 1,$$

and that we can repeat this integration procedure $n-1$ times more, we conclude that $E[Z_n] = 1$. We now recall...
that \(\sigma_1, \ldots, \sigma_n, \mu_1, \ldots, \mu_n\), as well as \(\epsilon_1, \ldots, \epsilon_{n-1}\) are \(\mathcal{F}_{n-1}\)-measurable and hence we can write

\[
E_{n-1}[Z_n] = E_{n-1}\left[\prod_{k=1}^{n} \exp\left(-\frac{\mu_k}{\sigma_k} e^k\right) \exp\left(-\frac{1}{2} \frac{\mu_k^2}{\sigma_k^2} e^k\right)\right]
\]

\[
= \prod_{k=1}^{n-1} \exp\left(-\frac{\mu_k}{\sigma_k} e_k\right) \exp\left(-\frac{1}{2} \frac{\mu_k^2}{\sigma_k^2} e_k\right)
\]

\[
\times \exp\left(-\frac{1}{2} \frac{\mu_n^2}{\sigma_n} e_n\right) E_{n-1}\left[\exp\left(-\frac{1}{2} \frac{\mu_n^2}{\sigma_n^2} e_n\right)\right]
\]

Since \(\sigma_n\) and \(\mu_n\) are \(\mathcal{F}_{n-1}\)-measurable and \(e_n\) is independent of \(\mathcal{F}_{n-1}\), this can be rewritten as (see, for example, proposition A.2.5 of Lamberton and Lapeyre 2008)

\[
E_{n-1}[Z_n] = Z_{n-1} \exp\left(-\frac{1}{2} \frac{\mu_n^2}{\sigma_n^2} Z_n\right) \exp\left(-\frac{\mu_n}{\sigma_n} x\right) dx = Z_{n-1},
\]

as required. We conclude by showing that \(Z_n\) is square integrable for all \(n = 1, \ldots, T\) whenever the drift term is bounded. Indeed, let \(B \geq 0\) be such that \(\mu_n \leq B\) for all \(n \in \mathbb{N}\), then

\[
E[Z_n^2] = \int_{-\infty}^{\infty} dx \cdots dx_n \exp\left(-\frac{2\mu_1 x_1}{\sigma_1} - \frac{\mu_1^2}{\sigma_1^2}\right) \times p(x_1) \cdots \exp\left(-\frac{2\mu_n x_n}{\sigma_n} - \frac{\mu_n^2}{\sigma_n^2}\right) p(x_n)
\]

\[
= \int_{-\infty}^{\infty} dx \cdots dx_{n-1} \exp\left(-\frac{2\mu_1 x_1}{\sigma_1} - \frac{\mu_1^2}{\sigma_1^2}\right) \times p(x_1) \cdots \exp\left(-\frac{2\mu_{n-1} x_{n-1}}{\sigma_{n-1}} - \frac{\mu_{n-1}^2}{\sigma_{n-1}^2}\right) p(x_{n-1})
\]

\[
\times \int_{-\infty}^{\infty} dx_n \exp\left(-\frac{2\mu_n x_n}{\sigma_n} - \frac{\mu_n^2}{\sigma_n^2}\right)
\]

\[
\times p(x_n).
\]

Given that

\[
\int_{-\infty}^{\infty} dx \exp\left(-\frac{2\mu_n x}{\sigma_n^2} - \frac{\mu_n^2}{\sigma_n^2}\right) p(x) = \exp(\mu_n^2/\sigma_n^2) \leq \exp(B^2/\sigma^2),
\]

where the inequality follows from hypothesis GARCH1 and the bounded character of the drift, we can conclude that

\[
E[Z_n^2] \leq \exp(B^2/\sigma^2) \int_{-\infty}^{\infty} dx \cdots dx_{n-1} \exp\left(-\frac{2\mu_1 x_1}{\sigma_1} - \frac{\mu_1^2}{\sigma_1^2}\right) \times p(x_1) \cdots \exp\left(-\frac{2\mu_{n-1} x_{n-1}}{\sigma_{n-1}} - \frac{\mu_{n-1}^2}{\sigma_{n-1}^2}\right) p(x_{n-1}).
\]

Using the inequality (A5) repeatedly in the previous formula we obtain

\[
E[Z_n^2] \leq \exp(nB^2/\sigma^2) < +\infty,
\]

as required.

(ii) \(Z_T\) is by construction non-negative and (A3) shows that \(E[Z_T] = \mathbb{P}(Z_T > 0) = 1\). This guarantees (see, for example, the remarks after theorem 4.2.1 of Lamberton and Lapeyre 2008) that \(Q\) is a probability measure equivalent to \(\mathbb{P}\).

(iii) Denote by \(\tilde{E}\) the expectations with respect to \(Q\). Then, for any \(u \in \mathbb{R}\) and \(n \in \{1, \ldots, T\}\), we will prove that

\[
\tilde{E}_n[\epsilon^{u\tilde{n}}] = \tilde{E}[\epsilon^{u\tilde{n}}] = e^{-u^2/2}.
\]

The first equality in (A6) together with proposition A.2.2 of Lamberton and Lapeyre (2008) show that the random variables \(\epsilon_1, \ldots, \epsilon_T\) are independent. The second equality, together with the uniqueness theorem for the characteristic function of a random variable (see, for instance, theorem 4.2. of Foata and Fuchs 2003) show that the random variables \(\epsilon_1, \ldots, \epsilon_T\) are normally distributed under \(Q\). Indeed, using the Bayes rule for conditional expectations and part (i), we have

\[
\tilde{E}_n[\epsilon^{u\tilde{n}}] = \frac{1}{E_{n-1}[Z_T]} E_{n-1}[Z_T e^{i\theta d\epsilon_n + (\mu_n/\sigma_n) x}]
\]

\[
= \frac{1}{Z_{n-1}} E_{n-1}[Z_T e^{i\theta d\epsilon_n + (\mu_n/\sigma_n) x}]
\]

\[
= \int_{-\infty}^{\infty} dx \cdots dx_T \exp\left(-\frac{\mu_n}{n\sigma_n} - \frac{\mu_n^2}{2\sigma_n^2}\right) p(x_T)
\]

\[
\cdots \exp\left(-\frac{\mu_T x_T}{\sigma_T} - \frac{\mu_T^2}{2\sigma_T^2}\right) e^{\theta d\epsilon_n + (\mu_n/\sigma_n) x}
\]

\[
= \int_{-\infty}^{\infty} dx \cdots dx_T \exp\left(-\frac{\mu_n}{n\sigma_n} - \frac{\mu_n^2}{2\sigma_n^2}\right) p(x_T)
\]

\[
\cdots \int_{-\infty}^{\infty} dx_T \exp\left(-\frac{\mu_T x_T}{\sigma_T} - \frac{\mu_T^2}{2\sigma_T^2}\right) p(x_T).
\]

Given that all the integrals

\[
\int_{-\infty}^{\infty} dx \exp\left(-\frac{\mu_n}{n\sigma_n} - \frac{\mu_n^2}{2\sigma_n^2}\right) p(x) = 1,
\]

the previous expression reduces to

\[
\tilde{E}_n[\epsilon^{u\tilde{n}}] = \int_{-\infty}^{\infty} dx \exp\left(-\frac{\mu_n}{n\sigma_n} - \frac{\mu_n^2}{2\sigma_n^2}\right) p(x_n)
\]

\[
\times e^{\theta d\epsilon_n + (\mu_n/\sigma_n) x} = e^{-u^2/2}.
\]

Regarding the second equality in (A6), we compute

\[
\tilde{E}[\epsilon^{u\tilde{n}}] = \int_{-\infty}^{\infty} dx \cdots dx_T \exp\left(-\frac{\mu_n}{n\sigma_n} - \frac{\mu_n^2}{2\sigma_n^2}\right) p(x_T)
\]

\[
\cdots \exp\left(-\frac{\mu_T x_T}{\sigma_T} - \frac{\mu_T^2}{2\sigma_T^2}\right) e^{\theta d\epsilon_n + (\mu_n/\sigma_n) x}
\]

\[
= \int_{-\infty}^{\infty} dx \exp\left(-\frac{\mu_n}{n\sigma_n} - \frac{\mu_n^2}{2\sigma_n^2}\right) p(x_n)
\]

\[
\cdots \int_{-\infty}^{\infty} dx_T \exp\left(-\frac{\mu_T x_T}{\sigma_T} - \frac{\mu_T^2}{2\sigma_T^2}\right) p(x_T).
\]
Again using (A7) and the second equality in (A8) we easily obtain that
\[ \tilde{E}[e^{	ilde{\eta}_i}] = e^{-\tilde{\eta}_i^2/2}, \]
as required.

(iv) Expressions (3.25) and (3.26) follow from substituting (3.24) in (2.6) and (2.7). Recall that
\[ E[\sigma_i^2 e_i^2] = E[E_i^2] = E[\sigma_i^2], \]
Hence, by hypothesis GARCH2 we have
\[ E[\sigma_i^2] = E[\sigma_i^2 e_i^2] < \infty. \quad (A9) \]
Now using (A9), part (i), and Bayes law of conditional probability, we have
\[ \tilde{E}[e_i|\tilde{\eta}_i] = \tilde{E}[\sigma_i|\tilde{e}_i = 1|\tilde{\eta}_i]] = \tilde{E}[\sigma_i|\tilde{e}_i = 1], \]
This inequality and (3.25) show that \( s_n \in L^1(\Omega, Q, \mathcal{F}) \). Indeed,
\[ E[|s_n|] = E[|s_0 + \sigma_1 \tilde{e}_1 + \cdots + \sigma_n \tilde{e}_n|] \leq E[|s_0|] + E[|\sigma_i|] \]
+ \cdots + E[|\sigma_n|] < \infty.
Finally,
\[ \tilde{E}_{n-1}[s_n] = \tilde{E}_{n-1}[s_n + \sigma \tilde{e}_n] = s_n + \sigma \tilde{E}_{n-1}[\tilde{e}_n] = s_n - 1, \]
which proves that \( \{s_0, s_1, \ldots, s_T\} \) forms a martingale with respect to \( Q \). Note that, in the last two equalities of the previous expression, we have used the conclusion of point (iii).

Suppose now that the variables \( \{\sigma_i\}_{i=1, \ldots, T} \) have finite kurtosis with respect to \( \mathbb{P} \) and that the drift process \( \{\mu_n\}_{n \in \mathbb{N}} \) is bounded. Then, for each \( i \in \{1, \ldots, T\} \),
\[ E[\sigma_i^4 e_i^4] < \infty. \quad (A11) \]
Then, since \( E[\sigma_i^4 e_i^4] = E[\sigma_i^4 E_i^2] = 3E[\sigma_i^4], \) we have
\[ E[\sigma_i^4] < \infty. \quad (A12) \]
We will proceed by showing first that (A11) and (A12) imply that
\[ E[s_n^4] < \infty, \quad (A13) \]
or, equivalently,
\[ E[s_n^4] = E\left[ \left( \sum_{i=1}^{T} \frac{1}{\sigma_i^4} \right)^{2} \right] < \infty. \quad (A14) \]
When the square inside the expectation is expanded, some algebra shows that \( E[s_n^4] \) is a finite sum of real numbers plus terms that, up to multiplication by finite constants, have the form:
\begin{itemize}
  \item \( E[\sigma_i^2 \sigma_j e_i e_j] = E[E_i e_j] = E[\sigma_i^2 e_j] = 0, \)
  where we assume, without loss of generality, that \( j < i; \)
  \item \( E[\sigma_i^2 e_i] < \infty, \) by hypothesis GARCH2;
  \item \( E[\sigma_i^2] < \infty, \) by (A11);
  \item also by (A11), the terms of the form
\end{itemize}
\[ E[\sigma_i^2 \sigma_j e_i e_j] \leq E[\sigma_i^2 e_i]^2/2 < \infty; \quad (A15) \]

This term is also finite because by (A15)
\[ E[\sigma_i^2 \sigma_j e_i e_j | e_i, e_j] < \infty; \]

This relation establishes (A14). We now use this relation to conclude that \( \{s_0, s_1, \ldots, s_T\} \) is square integrable with respect to \( Q \). Indeed, by part (i) of the theorem, \( \{Z_n\}_{n=1}^{T} \) is a square integrable martingale and hence
\[ E[s_n^2] = E[Z_n^2] < (E[s_n^4])^{1/2} < \infty, \quad (A16) \]
as required.

(v) Let \( n \in \{1, \ldots, T\} \). Then, by part (iii)
\[ \tilde{E}_{n-1}[s_n] = \tilde{E}_{n-1}[s_{n-1} + \sigma \tilde{e}_n] = s_n + \sigma \tilde{E}_{n-1}[\tilde{e}_n] = s_n - 1, \]
Now, as \( \tilde{E}[\sigma_n \tilde{e}_n] = \tilde{E}[\tilde{E}_{n-1}[\sigma_n \tilde{e}_n]] = \sigma_n \tilde{E}[\tilde{e}_n] = 0. \)
Consequently,
\[ \text{cov}(\sigma_n \tilde{e}_n, \sigma_j \tilde{e}_j) = \tilde{E}[\sigma_n \tilde{e}_n \sigma_j \tilde{e}_j] = \tilde{E}[\tilde{E}_{n-1}[\sigma_n \tilde{e}_n \sigma_j \tilde{e}_j]] = 0. \]

\[ \square \]

**Proof of proposition 3.2**

**Proof:** Let \( f_k(\mu) \) be the function defined by the value process (2.13) with respect to the physical measure, that is
\[ f_k(\mu) := E_k \left[ h \left( 1 - \frac{\mu}{\sigma_T} e_T \right) \left( 1 - \frac{\mu}{\sigma_{T-1}} e_{T-1} \right) \cdots \left( 1 - \frac{\mu}{\sigma_{k+1}} e_{k+1} \right) \right]. \]
A straightforward computation shows that
\[ f_k(0) = E_k[h], \quad f_k'(0) = - \sum_{j=k+1}^{T} E_k \left[ h \frac{e_j - E_k[e_j]}{\sigma_j} \right]. \quad (A17) \]
Consequently, the linear Taylor approximation $V^\text{lin}_k$ of $V_k$ is given by
\[
V^\text{lin}_k = E_k[h] - \mu \sum_{j=k+1}^T E_k \left[ \frac{h^j}{\sigma_j} \right]. \tag{A18}
\]

Now let $\tilde{f}_k(\mu)$ be the value process with respect to the martingale measure $Q$ in theorem 3.1. Using the martingale property of the process $Z_n$ that gives us the Radon–Nikodym derivative $dQ/dP$ we have
\[
\tilde{f}_k(\mu) := E_k[h] = \frac{1}{E_k[Z_T]} E_k[Z_T h] = \frac{1}{Z_k} E_k[Z_T h] = E_k \left[ \frac{Z_T}{Z_k} h \right]
\]
\[
= E_k \left[ \exp \left( -\frac{\mu}{\sigma} T - \frac{\mu^2}{2\sigma^2} \right) \cdots \exp \left( -\frac{\mu}{\sigma_{k+1}} \epsilon_{k+1} - \frac{\mu^2}{2\sigma_{k+1}^2} \right) \right]. \tag{A19}
\]

A straightforward computation shows that $\tilde{f}_k(0) = f_k(0)$ and $\tilde{f}'_k(0) = f'_k(0)$. Consequently, $V^\text{lin}_k = \tilde{V}^\text{lin}_k$, as required.

**Proof of proposition 3.3**

**Proof:** We start with Duan’s model, which is given by
\[
\log \left( \frac{S_n}{S_{n-1}} \right) = s_n - s_{n-1} = \lambda \sigma_n + \frac{1}{2} \sigma_n^2 + \sigma_n \epsilon_n, \tag{A21}
\]
\[
\sigma_n^2 = \alpha_0 + \sum_{j=1}^q \alpha_j \sigma_{n-j}^2 + \sum_{j=1}^p \beta_j \sigma_{n-j}^2, \quad \epsilon_{n,\text{Iid}} \sim \text{HIDN}(0,1), \tag{A22}
\]
where $\alpha_0 > 0$, $\alpha$, $\beta_j > 0$ and $\sum_{j=1}^q |\alpha_j| + \sum_{j=1}^p |\beta_j| < 1$ so that second-order stationarity is ensured and the coefficient $\lambda$ is interpreted as a unit risk premium. Using the notation introduced in (3.30) and (3.33), this is a general GARCH model with $\mu_n = \lambda \sigma_n - \frac{1}{2} \sigma_n^2$ and $\tilde{\mu}_n = \lambda \sigma_n$. Consequently, by (3.35) and (3.36), after risk neutralization this process is given by
\[
\log \left( \frac{S_n}{S_{n-1}} \right) = s_n - s_{n-1} = -\frac{1}{2} \sigma_n^2 + \sigma_n \tilde{\epsilon}_n, \quad \tilde{\epsilon}_{n,\text{Iid}} \sim \text{HIDN}(0,1), \tag{A23}
\]
\[
\sigma_n^2 = \alpha_0 + \sum_{i=1}^q \alpha_i (\tilde{\epsilon}_{n-i} - \lambda - \frac{1}{2} \sigma_n^2)^2 + \sum_{j=1}^p \beta_j \sigma_{n-j}^2, \tag{A24}
\]
and hence, according to (3.37), its local risk-minimizing price is given by $V_0 = \tilde{E}[H]$, which coincides with the formula proposed by Duan since the process (A23) and (A24) is identical to that obtained in theorem 2.2 of Duan (1995) from his *locally risk-neutral valuation relationship*.

As for Heston and Nandi, their process is given by
\[
\log \left( \frac{S_n}{S_{n-1}} \right) = s_n - s_{n-1} = \lambda \sigma_n^2 + \sigma_n \epsilon_n, \quad \epsilon_{n,\text{Iid}} \sim \text{HIDN}(0,1), \tag{A25}
\]
\[
\sigma_n^2 = \alpha_0 + \sum_{i=1}^q \alpha_i (\epsilon_{n-i} - \gamma \sigma_{n-i})^2 + \sum_{j=1}^p \beta_j \sigma_{n-j}^2, \tag{A26}
\]
where $\alpha_0 > 0$, $\alpha$, $\beta_j > 0$ and the roots of the polynomial $\lambda^2 - \sum_{j=1}^q (\beta_j + \alpha_j) \lambda + \alpha_j^2 \sigma_n^2$ lie inside the unit circle so that second-order stationarity is ensured. This time we have $\mu_n = \lambda \sigma_n^2$, $\tilde{\mu}_n = (\lambda + \frac{1}{2}) \sigma_n^2$, and
\[
\tilde{\epsilon}_n := \epsilon_n + \left( \lambda + \frac{1}{2} \right) \sigma_n. \tag{A27}
\]
Hence, the risk-neutralized version of (A25) and (A26) is
\[
\log \left( \frac{S_n}{S_{n-1}} \right) = s_n - s_{n-1} = -\frac{1}{2} \sigma_n^2 + \sigma_n \tilde{\epsilon}_n, \tag{A28}
\]
\[
\times \tilde{\epsilon}_{n,\text{Iid}} \sim \text{HIDN}(0,1), \tag{A29}
\]
which coincides with that proposed by proposition 1 of Heston and Nandi (2000). We emphasize that, by theorem 3.1, the risk-neutralized innovations $\tilde{\epsilon}_{n,\text{Iid}}$ are automatically IIDN(0,1) and, unlike in the treatment carried out by Heston and Nandi (2000), no additional assumption is needed.