Hedging of time discrete auto-regressive stochastic volatility options

Joan del Castillo[1] and **Juan-Pablo Ortega**[2]

[1]Universitat Autònoma de Barcelona, Spain.

[2]CNRS, Université de Franche-Comté, France

University of Calgary, 09/2011

Plan of the talk

1. Modeling of the underlying: the ARSV approach.
2. Volatility estimation methods.
3. Martingale measures: MMM and the Extended Girsanov Principle.
4. Pricing and hedging via local risk minimization.
5. Empirical results.

Some history

- Louis Bachelier and Brownian motion: "Théorie de la Spéculation" (1900)
- Black, Merton, Scholes formula (1973) and geometric Brownian motion. Merton, Scholes Nobel prize 1997. Prices always positive.

$$dS_t = S_t(\mu dt + \sigma dB_t) \Longleftrightarrow S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}$$

  - Empirical inaccuracies: leptokurtic distributions of returns and volatility clustering. The volatility smile.
- Proposals in continuous time:
  - More general stochastic processes: Levy finance, jumps...
  - Stochastic volatility: Hull and White (1987), Heston (1993).

Time series

- More general than strong Euler discretizations of SDEs.
- Possibility of incorporating in the pricing pure time series theoretical concepts: cointegration (Duan, Pliska (2001))
- Need to go beyond linear models. They are all homoscedastic.
- No stochastic volatility. The number of innovations remains constant.
- ARCH models: Engle (1982). Originally introduced to model the variance in the UK inflation.
- GARCH: Bollerslev (1986); more parsimonious.
- Asymmetric GARCH: Ding, Granger, and Engle (1993). It captures the different impact that positive and negative shocks have on volatility.

Generalized Autorregressive Conditional Heteroscedasticity

- The original Ding, Granger, and Engle (1993)-He and Terasvirta (1999) model:

$$\begin{aligned}\log\left(S_n/S_{n-1}\right) &=: r_n = \mu + \sigma_n \epsilon_n, \quad \mu \in \mathbb{R},\\ \sigma_n^\delta &= \omega + \sum_{i=1}^{p} \beta_i \sigma_{n-i}^\delta + \sum_{i=1}^{q} \alpha_i (|\overline{r}_{n-i}| - \gamma_i \overline{r}_{n-i})^\delta.\end{aligned}$$

- The Heston-Nandi model (2000). Picked because it yields a closed-form option pricing formula:

$$\begin{aligned}\log\left(S_n/S_{n-1}\right) &=: r_n = r + \lambda \sigma_n^2 + \sigma_n \epsilon_n, \quad \mu \in \mathbb{R},\\ \sigma_n^2 &= \omega + \sum_{i=1}^{p} \beta_i \sigma_{n-i}^2 + \sum_{i=1}^{q} \alpha_i (\epsilon_{n-i} - \gamma_i \sigma_{n-i})^2.\end{aligned}$$

GARCH advantages and rigidities

- Successful in capturing both leptokurticity and volatility clustering. For all these models leptokurticity and heteroscedasticity are linked. For example, for Gaussian innovations:

$$\kappa = \frac{E[\sigma_n^4 \epsilon_n^4]}{(E[\sigma_n^2 \epsilon_n^2])^2} = 3 + 3\frac{\mathrm{var}\left(\sigma_n^2\right)}{(E[\sigma_n^2])^2}.$$

- Kurtosis: Ling and McAleer (2002) contains a characterization for the existence of the fourth moment for asymmetric GARCH. This characterization is much needed in the optimization problem in the next point and for pricing via local risk minimization.

Necessary and sufficient condition for the existence of the moment of order $2m$ is that

$$\rho\left[E\left[A^{\otimes m}\right]\right] < 1, \tag{1}$$

where $\rho(B) = \max\{|\text{eigenvalues of the matrix } B|\}$, A is given by

$$A = \left(\begin{array}{ccc|ccc} \alpha_1 Z_t & \cdots & \alpha_p Z_t & \beta_1 Z_t & \cdots & \beta_q Z_t \\ & I_{(p-1)\times(p-1)} & 0_{(p-1)\times 1} & & 0_{(p-1)\times q} & \\ \hline \alpha_1 & \cdots & \alpha_p & \beta_1 & \cdots & \beta_q \\ & 0_{(q-1)\times p} & & & I_{(q-1)\times(q-1)} & 0_{(q-1)\times 1} \end{array}\right),$$

and $Z_t := (|\epsilon_t| - \gamma\epsilon_t)^2$. For $m = 1$, the condition (1) is the same as before. The kurtosis is finite whenever (1) holds with $m = 2$.

	Condition for finite kurtosis
GARCH(1,0)	$(1+\gamma^2)\alpha < 1.$
GARCH(2,0)	$1/2\,\left(1+\gamma^2\right)\left(\gamma^2{\alpha_1}^2+{\alpha_1}^2+2\,\alpha_2+\sqrt{\left(1+\gamma^2\right){\alpha_1}^2\left(\gamma^2{\alpha_1}^2+4\,\alpha_2+{\alpha_1}^2\right)}\right) < 1.$
GARCH(2,1)	$1/2\,\left(\gamma^4+2\,\gamma^2+1\right){\alpha_1}^2+1/2\,\left(2\,\beta+2\,\beta\,\gamma^2\right)\alpha_1+\gamma^2\alpha_2+1/2\,\beta^2+\alpha_2$ $+1/2\sqrt{\left(\left(1+\gamma^2\right)^2{\alpha_1}^2+\left(2\,\beta+2\,\beta\,\gamma^2\right)\alpha_1+4\,\alpha_2+4\,\gamma^2\alpha_2+\beta^2\right)\left(\left(1+\gamma^2\right)\alpha_1+\beta\right)^2}$
GARCH(2,2)	$1/2\,\left(\gamma^4+2\,\gamma^2+1\right){\alpha_1}^2+1/2\,\left(2\,\beta_1+2\,\beta_1\gamma^2\right)\alpha_1+1/2\,{\beta_1}^2+\alpha_2+\beta_2+\gamma^2\alpha_2$ $+1/2\sqrt{\left(\left(1+\gamma^2\right)\alpha_1+\beta_1\right)^2\left(\left(1+\gamma^2\right)^2{\alpha_1}^2+\left(2\,\beta_1+2\,\beta_1\gamma^2\right)\alpha_1+4\,\alpha_2+4\,\beta_2+4\,\gamma^2\alpha_2+\right.}$

The GARCH(1,1) case

- $p := \alpha + \beta$ persistence parameter.
- Finite kurtosis: $p^2 + \alpha^2(k_\epsilon - 1) < 1$.
- Relation between kurtosis and persistence parameter:

$$k_y = k_\epsilon \left(1 - \frac{\alpha^2(k_\epsilon - 1)}{1 - p^2}\right)^{-1}.$$

- Relation between autocorrelation of squares and persistence parameter:

$$\rho_2(1) = \frac{\alpha(1 - p^2 + p\alpha)}{1 - p^2 + \alpha^2}$$

- Poor cohabitation between finite kurtosis, volatility clustering (persistence), and autocorrelation of squares: elevated persistence implies small α which in turns makes $\rho_2(1)$ very small.

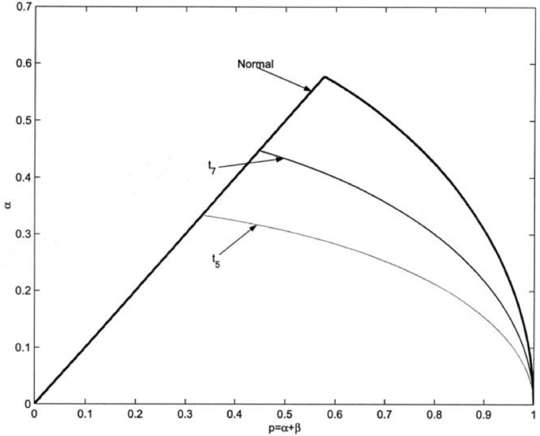

Figure: Regions with finite kurtosis (from Carnero et al (2004))

ARSV family. Taylor (1986)

- The model:

$$\begin{cases} y_t &= r + \sigma_t \epsilon_t, \qquad \{\epsilon_t\} \sim \mathrm{IIDN}(0,1) \\ b_t &= \gamma + \phi b_{t-1} + w_t, \qquad \{w_t\} \sim \mathrm{IIDN}(0,\sigma_w^2) \end{cases}$$

  where $b_t := \log(\sigma_t^2)$, γ is a real parameter, and $\phi \in (-1,1)$.
- The volatility $\{\sigma_t\}$ is a non-traded stochastic latent variable that is not a predictable process that can be written as a function of previous returns and volatilities.
- $\phi \in (-1,1)$ guarantees existence of stationary solution with finite moments of arbitrary order.

$$\mathrm{var}(y_t) = \mathrm{E}[\sigma_t^2] = \exp\left[\frac{\gamma}{1-\phi} + \frac{1}{2}\sigma_b^2\right], \quad \mathrm{kurtosis}\,(y_t) = 3\exp\left(\sigma_b^2\right).$$

Volatility estimation

State space representation and Kalman filtering. Write down the model using a state space representation by setting

$$l_t = \log\left(|y_t - r|\right), \qquad L_t = \log\left(\sigma_t\right).$$

This yields

$$\left\{ \begin{array}{rcl} l_t & = & L_t + \xi_t, \qquad \xi_t = \log\left(|\epsilon_t|\right), \\ (L_t - \alpha_K) & = & \phi(L_{t-1} - \alpha_K) + \eta_t, \qquad \eta_t = \frac{1}{2} w_t, \end{array} \right. \tag{2}$$

where $\alpha_K = E[L_t] = \frac{1}{2}E[b_t] = \frac{\gamma}{2(1-\phi)}$. Even though $\xi_t = \log\left(|\epsilon_t|\right)$ is not Gaussian, we treat it as such with mean $\mu_\xi := E[\xi_t] = -0.63518$ and variance $\sigma_\xi^2 = \mathrm{Var}[\xi_t] = \pi^2/8$.

Kalman provides one-step ahead linear forecasts $l_{t,1} := P[l_t|\mathcal{F}_{t-1}]$ that in turn produce predictable estimations σ_t^k of the volatilities σ_t by setting $\log\left(\sigma_t^k\right) := l_{t-1,1} - \mu_\xi$ and hence

$$\sigma_t^k = \exp(l_{t-1,1} - \mu_\xi). \tag{3}$$

The *h*-likelihood approach. Consists of carrying out a likelihood estimation while considering the volatilities as unobserved parameters that are part of the optimization problem. Let $\alpha := (\gamma, \phi, \sigma_w)$ be the parameters vector, $b = (b_1, \ldots, b_T)$ the vector that contains the unobserved $b_t := \log(\sigma_t^2)$, $z_t := y_t - r$:

$$h(z; b, \alpha) = -\frac{1}{2}\sum_{t=1}^{T}\left(z_t^2\exp(-b_t) + b_t + \frac{1}{\sigma_w^2}(b_t - \gamma - \phi b_{t-1})^2 + \log\sigma_w^2\right)$$

Availability of numerically efficient procedures to maximize the likelihood with respect to the variables b and α once the sample z has been specified.

This procedure has a natural one-step ahead forecast of the volatility associated using the variables $b_t := \log(\sigma_t^2)$ and alternating prediction and updating:

$$
\begin{aligned}
b_0 &:= \gamma/(1-\phi)\\
b_{1p} &= \gamma + \phi b_0\\
b_{1u} &= \arg\min_{b_1 \in \mathbb{R}} \left(z_1^2 \exp(-b_1) + b_1 + \frac{1}{\sigma_w^2}(b_1 - \gamma - \phi b_0)^2 \right)\\
b_{1p} &= \gamma + \phi b_0\\
\cdots &\quad \cdots \quad \cdots\\
b_{tp} &= \gamma + \phi b_{(t-1)u},\\
b_{tu} &= \arg\min_{b_t \in \mathbb{R}} \left(z_t^2 \exp(-b_t) + b_t + \frac{1}{\sigma_w^2}(b_t - \gamma - \phi b_{(t-1)u})^2 \right).
\end{aligned}
$$

These forecasts can be used to produce predictable estimations σ_t^h of the volatilities σ_t by setting

$$\sigma_t^h := \exp\left(\frac{1}{2} b_{tp}\right).$$

h-likelihood has a much wider range of applicability for it is not subjected to the rigidity of the state space representation and hence can be generalized to stochastic volatility models with complex link functions, to situations where the innovations are non-Gaussian, or there is a dependence between $\{\epsilon_t\}$ and $\{w_t\}$.

Incompleteness: the problem

A market model with a single risky asset modeled using a ARSV process driven by non-binomial innovations is incomplete, that is, not every payoff can be replicated via a self-financing portfolio.

The incompleteness comes from:

- Poor cohabitation between discreet time modeling and the infinite number of states of the innovations (already the case with GARCH; can be fixed by going to the continuous time limit)
- Use of two driving noises.

Generalized trading strategies

We abandon the use of self-financing portfolios we introduce the notion of **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.

Definition

A **generalized trading strategy** is a pair of stochastic processes (ξ^0, ξ) such that $\{\xi_n^0\}_{n\in\{0,\ldots,T\}}$ is adapted and $\{\xi_n\}_{n\in\{1,\ldots,T\}}$ is predictible. The **value process** V of (ξ^0, ξ) is defined as

$$V_0 := \xi_0^0, \qquad \text{and} \qquad V_n := \xi_n^0 + \xi_n \cdot S_n, \quad n \geq 1.$$

Definition

The **gains process** G:

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

The **cost process** C: $C_n := V_n - G_n$, $n = 0, \ldots, T$.

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

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

or, equivalently, if $V_0 = C_0 = C_1 = \ldots = C_T$.

Let r_t be the continuously composed risk-free interest rate paid on the currency of the underlying in the period that goes from time $t-1$ to t; we assume that $\{r_t\}$ is a predictable process. Denote by

$$R_t := \sum_{j=0}^{t} r_j.$$

The price at time t of the riskless asset S^0 such that $S^0_0 = 1$, is given by $S^0_t = \mathrm{e}^{R_t}$. The previous processes have discounted versions $\widetilde{V}_t$, $\widetilde{G}_t$, and $\widetilde{C}_t$ defined as:

$$\widetilde{V}_t := V_t \mathrm{e}^{-R_t}, \qquad \widetilde{G}_t := \sum_{k=1}^{t} \xi_k \cdot (\widetilde{S}_k - \widetilde{S}_{k-1}), \quad \text{and} \quad \widetilde{C}_t := \widetilde{V}_t - \widetilde{G}_t.$$

Definition

Assume that both H and the $\{S_n\}_{n\in\{0,\ldots,T\}}$ are $L^2(\Omega, P)$. A generalized trading strategy is called **admissible** for H whenever it is in $L^2(\Omega, P)$ and its associated value process is such that

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

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

Remark: since they are not self-financing these strategies may be available even for non-attainable payoffs!

Local risk minimizing strategies

Definition

The **local risk process** of an admissible strategy (ξ^0, ξ) is the process

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

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

$$R_t(\widehat{\xi}^0, \widehat{\xi}) \leq R_t(\xi^0, \xi), \quad \mathrm{P} \ a.s.$$

for all t and each admissible strategy (ξ^0, ξ).

Theorem (Föllmer, Schweizer, Sondermann)

An admissible strategy is local risk-minimizing if and only if the cost process is a P-martingale and it is strongly orthogonal to S, in the sense that $\mathrm{cov}_n(\widetilde{S}_{n+1} - \widetilde{S}_n, \widetilde{C}_{n+1} - \widetilde{C}_n) = 0$, P-a.s., for any $t = 0, \ldots, T-1$.

- An admissible strategy is local risk-minimizing for a fixed probability measure P. Usually the physical measure (not necessarily risk-neutral).
- Does not make the difference between shortfall and windfall.
- The local risk-minimizing strategy, if it exists, is unique and the payoff H can be decomposed as (Kunita-Watanabe)

$$\widetilde{H} = V_0 + \widetilde{G}_T + \widetilde{L}_T, \tag{4}$$

G_n gains process and $\widetilde{L}_n := \widetilde{C}_n - C_0$ the **discounted global risk process**.

General solution

Unlike its global counterpart, the solution of the local risk minimization problem, when it exists, can be explicitly written down:

$$V_T = H, \tag{5}$$

$$\xi_{t+1} = \frac{\mathrm{cov}^P(\widetilde{V}_{t+1}, \widetilde{S}_{t+1} - \widetilde{S}_t \mid \mathcal{F}_t)}{\mathrm{var}^P(\widetilde{S}_{t+1} - \widetilde{S}_t \mid \mathcal{F}_t)}, \tag{6}$$

$$\widetilde{V}_t = E^P\left[\widetilde{V}_{t+1} \mid \mathcal{F}_t\right] - \xi_{t+1} E^P\left[\left(\widetilde{S}_{t+1} - \widetilde{S}_t\right) \mid \mathcal{F}_t\right]. \tag{7}$$

Local risk minimization with respect to a martingale measure

The natural measure to be considered for local risk minimization is the physical measure: from a risk management perspective this is the natural measure that should be used in order to construct the local risk process. Two major difficulties:

- Resulting expressions numerically difficult to estimate due to the high variance of the associated Monte Carlo estimators.
- Unless there is a minimal martingale measure available, the option prices that result from this technique cannot be interpreted as arbitrage free prices.

This leads us to use equivalent martingale measures Q that additionally minimize the so called **remaining conditional risk**: $R_t^R(\xi^0, \xi) := E_t[(C_T - C_t)^2]$, $t = 0, \ldots, T$; this is in general not true outside the martingale setup.

$$
\begin{aligned}
V_T &= H, \\
\xi_{t+1} &= \frac{1}{\Sigma_{t+1}^2} E_t^Q \left[e^{-(R_T+R_t)} H(S_T) \left(S_{t+1} e^{-r_{t+1}} - S_t \right) \right], \\
V_t &= E_t^Q \left[e^{-(R_T-R_t)} H(S_T) \right],
\end{aligned}
$$

where

$$
\Sigma_{t+1}^2 := \mathrm{var}^Q(\widetilde{S}_{t+1} - \widetilde{S}_t \mid \mathcal{F}_t) = e^{-2R_t} \mathrm{var}^Q(S_{t+1} \mathrm{e}^{-r_{t+1}} \mid \mathcal{F}_t).
$$

V_0 has a clear interpretation

Changes in the hedging frequency

A major advantage of the local risk minimization hedging scheme is its adaptability to prescribed changes in the hedging frequency. Suppose that the life of the option H with maturity in T time steps is partitioned into identical time intervals of duration j; let k such that $kj = T$.

$$R_t^j(\xi^0, \xi) := E^P\left[(\widetilde{C}_{t+j}^j - \widetilde{C}_t^j)^2 \mid \mathcal{F}_t\right], \quad t = 0, j, 2j, \ldots, (k-1)j = T-j,$$

where $\{\widetilde{C}_t^j\}$ is a cost process constructed out of value and gains processes, $\{V_t^j\}$ and $\left\{G_t^j\right\}$ that only take into account the prices of the underlying assets at time steps $t = 0, j, 2j, \ldots, kj = T$, in particular, given an integer l such that $t = lj$

$$\widetilde{G}_t^j := \sum_{r=1}^{l} \xi_{rj} \cdot (\widetilde{S}_{rj} - \widetilde{S}_{(r-1)j}).$$

The solution of this local risk minimization problem with modified hedging frequency with respect to a martingale measure is given by the expressions:

$$V_T^j = H, \tag{8}$$

$$\xi_{t+j} = \mathrm{e}^{-(R_T - R_t)} \frac{E_t^Q\left[H(S_T)\left(S_{t+j}e^{-(R_{t+j}-R_t)} - S_t\right)\right]}{E_t^Q\left[S_{t+j}^2 \mathrm{e}^{-2(R_{t+j}-R_t)} - S_t^2\right]}, \tag{9}$$

$$V_t^j = E_t^Q\left[e^{-(R_T - R_t)} H(S_T)\right], \tag{10}$$

for any $t = 0, j, 2j, \ldots, (k-1)j = T - j$.

The minimal martingale measure

It is 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.

Minimal martingale measure: martingale measure $Q_{\min}$ equivalent to the physical probability P that satisfies the following two conditions: $E\left[(dQ_{\min}/dP)^2\right] < \infty$ and every P-martingale $M \in L^2(\Omega, P)$ that is strongly orthogonal to the price process s, is also a $Q_{\min}$-martingale. It satisfies an entropy minimizing property (Schweizer 2001) and the value process V_k can be expressed as

$$V_k = E^{Q_{\min}}[h],$$

which obviously yields the interpretation that we are looking for.

Theorem

In the ARSV setup, the minimal martingale measure is determined by the Radon-Nikodym derivative $dQ_{\min}/dP$ that is obtained by evaluating at time T the P-martingale $\{Z_t\}_{t\in\{1,\ldots,T\}}$ defined by

$$Z_t := \prod_{k=1}^{t} 1 + \frac{\left(e^{K^P_{\epsilon_t}(\sigma_k)} - 1\right)\left(e^{\sigma_k \epsilon_k} - e^{K^P_{\epsilon_t}(\sigma_k)}\right)}{e^{2K^P_{\epsilon_t}(\sigma_k)} - e^{K^P_{\epsilon_t}(2\sigma_k)}},$$

where $K^P_{\epsilon_t}$ is the **conditional cumulant functions** *of ϵ_t with respect to the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\{0,\ldots,T\}}$:*

$$K^P_{\epsilon_t}(u) = \log E^P\left[e^{u\epsilon_t} | \mathcal{F}_{t-1}\right], \qquad \textit{with } u \textit{ a random variable.}$$

- The measure $Q_{\min}$ *is in general signed*: as the random variable $\sigma_k \epsilon_k$ is $\mathcal{F}_k$-adapted and $K^P_{\epsilon_k}(\sigma_k)$ is $\mathcal{F}_k$-predictable, the term $\left(e^{\sigma_k \epsilon_k} - e^{K^P_{\epsilon_k}(\sigma_k)}\right)$ can take arbitrarily negative values that can force Z_t to become negative. Negative occurrences are extremely unlikely. The bias introduced by censoring paths that yield negative Radon-Nikodym derivatives and using $Q_{\min}$ as a well-defined positive measure is not noticeable.
- The quantity $K^P_{\epsilon_t}(\sigma_t) = \log E^P \left[e^{\sigma_t \epsilon_t} | \mathcal{F}_{t-1}\right]$ is difficult to compute. As

$$K^P_{\epsilon_t}(u) = \log E^P \left[e^{L^P_{\epsilon_t}(u)} \mid \mathcal{F}_{t-1}\right].$$

with $L^P_{\epsilon_t}(z) = \log E^P \left[e^{z \epsilon_t}\right]$, we use the estimations σ^k_t and σ^h_t to approximate $K^P_{\epsilon_t}(\sigma_t)$.

For example, if the innovations $\{\epsilon_t\}$ are Gaussian and we hence have $L^P_{\epsilon_t}(z) = z^2/2$, we approximate

$$K^P_{\epsilon_t}(\sigma_t) = \log E^P_{t-1}\left[e^{\frac{\sigma_t^2}{2}}\right]$$

by $(\sigma_t^k)^2/2$ or $(\sigma_t^h)^2/2$ depending on the technique (Kalman or h-likelihood, respectively) used to estimate the conditional volatility.

The value processes obtained when carrying out local risk minimization with respect to the physical and the minimal martingale measures are identical, however the hedges are in general **NOT** the same and consequently so are the hedging errors.

Proposition: let $\{\xi_t^{Q_{\min}}\}_{t=1}^T$ and $\{\xi_t^P\}_{t=1}^T$ be the local risk minimizing hedges associated to the minimal martingale measure $Q_{\min}$ and the physical measure P, respectively. Let $\left\{\widetilde{L}_t^P\right\}_{t=0}^T$ be the associated P-global risk process. Then, for any $t \in \{1, \ldots, T\}$:

$$\xi_t^{Q_{\min}} = \xi_t^P + \frac{E_{t-1}^{Q_{\min}}\left[\widetilde{L}_t^P(\widetilde{S}_t - \widetilde{S}_{t-1})\right]}{\operatorname{var}_{t-1}^{Q_{\min}}\left[\widetilde{S}_t - \widetilde{S}_{t-1}\right]}. \tag{11}$$

If the processes $\left\{\widetilde{L}_t^P\right\}_{t=0}^T$ and $\left\{\widetilde{S}_t\right\}_{t=0}^T$ are either $Q_{\min}$–strongly orthogonal or $Q_{\min}$–independent, then $\{\xi_t^{Q_{\min}}\}_{t=1}^T = \{\xi_t^P\}_{t=1}^T$.

The Extended Girsanov Principle

- Elliott, Madan (1998)
- The behavior of the process under the martingale measure coincides with that of its martingale component under the physical probability: mean corrected martingale measure (MMM).
- Widely used in the GARCH context (Badescu et al (2011))
- Local risk minimization with respect to this measure yields trading strategies that minimize the risk adjusted local cost with risk adjustment discount factor e^{μ_t} defined by

$$e^{\mu_t} := E_{t-1}^{P}\left[\frac{\widetilde{S}_t}{\widetilde{S}_{t-1}}\right].$$

In the ARSV setup the MCMM measure is determined by the stochastic discount factors of the form:

$$N_t = \frac{f_{\sigma_t \epsilon_t}(\sigma_t \epsilon_t + m_t - r + \log E^P_{t-1}\left[e^{\sigma_t \epsilon_t}\right])}{f_{\sigma_t \epsilon_t}(\sigma_t \epsilon_t)},$$

with $f_{\sigma_t \epsilon_t}$ the conditional density of $\sigma_t \epsilon_t$. As σ_t and ϵ_t are independent, Rohatgi's formula yields:

$$f_{\sigma_t \epsilon_t}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{\sigma_t}(x) f_{\epsilon_t}\left(\frac{z}{x}\right) dx.$$

Computationally heavy despite ad hoc methods (Glen et al (2004)).

MCMM GARCH-like proxy ARSV

Define the **market price of risk** process:

$$\rho_t = \frac{m_t + K^P_{\epsilon_t}(\sigma_t) - r}{\sigma_t}, \tag{12}$$

to which we associate the **stochastic discount factor**:

$$N_t(\epsilon_t, \rho_t) = \frac{f^P(\epsilon_t + \rho_t)}{f^P(\epsilon_t)},$$

f^P is the pdf of the innovations $\{\epsilon_t\}$ under P.

Theorem

(i) *The process* $Z_t := \prod_{k=1}^{t} N_k$*, is a* $(\mathcal{F}_t, P)$*-martingale such that* $E^P[Z_t] = 1$.
(ii) Z_T *defines an equivalent martingale measure* Q_{mc}.

Empirical study

The price generating ARSV process. We take the following ARSV parameters in:

$$r = 0.1/252, \qquad \gamma = -0.821, \qquad \phi = 0.9, \qquad \sigma_w = 0.675.$$

The resulting log-returns are very leptokurtic $(\text{kurtosis} = 33.00)$ and volatile (the daily volatility attains 63%!). We try to amplify the defects of the different hedging methods in order to better compare them.

The hedging methods.

- **Black-Scholes**: we forget that the price process is generated via an ARSV model and we handle the hedging using the standard Black-Scholes delta as if the underlying was a realization of a log-normal process..

- **Duan's static delta hedge**: this is a generalization of the Black-Scholes delta hedge that has been introduced in Duan (1995) in the GARCH context. In that result the author computes the derivative of a European call price with respect to the price of the underlying in an arbitrary incomplete situation in which that option price is simply obtained as the expectation of the discounted payoff with respect to a given pricing kernel Q, that is:

$$V_t = E^Q\left[e^{-(R_T - R_t)} H(S_T) \mid \mathcal{F}_t\right].$$

In those circumstances:

$$\xi_{t+1}^{\mathrm{Duan}} := \frac{\partial V_t}{\partial S_t} = e^{-(R_T - R_t)} E^Q\left[\frac{S_T}{S_t} \mathbf{1}_{S_T \geq K} \mid \mathcal{F}_t\right].$$

In our empirical study we will use this hedging technique using the various pricing kernels previously introduced.

- Local risk minimization using the minimal martingale measure and the Extended Girsanov Principle.
- In all these methods we implement volatility estimations both using Kalman and h-likelihood.

Exercise 1

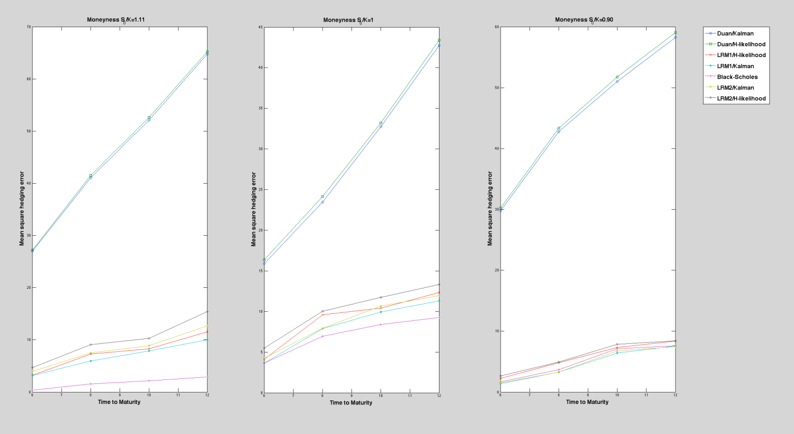

Figure: Experiment 1. All the hedging techniques explained used except for those based on the minimal martingale measure. The four maturities considered are 6, 8, 10, and 12 time steps. Hedging is carried out daily.

Exercise 2

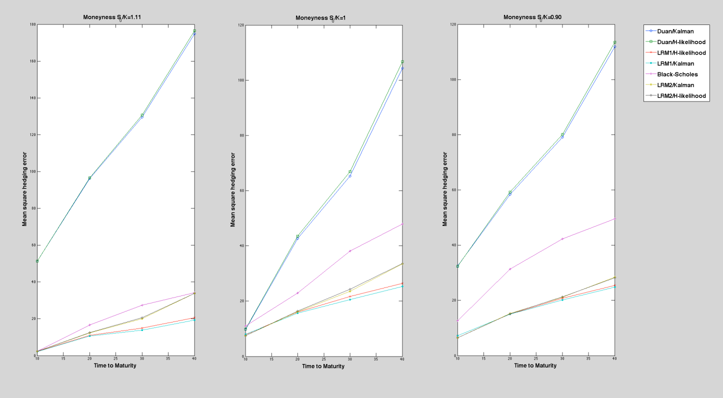

Figure: Experiment 2. All the hedging techniques explained in the paper are used except for those based on the minimal martingale measure. The four maturities considered are 10, 20, 30, and 40 time steps. Hedging is carried out every 10 days.

Exercises 3 and 4:

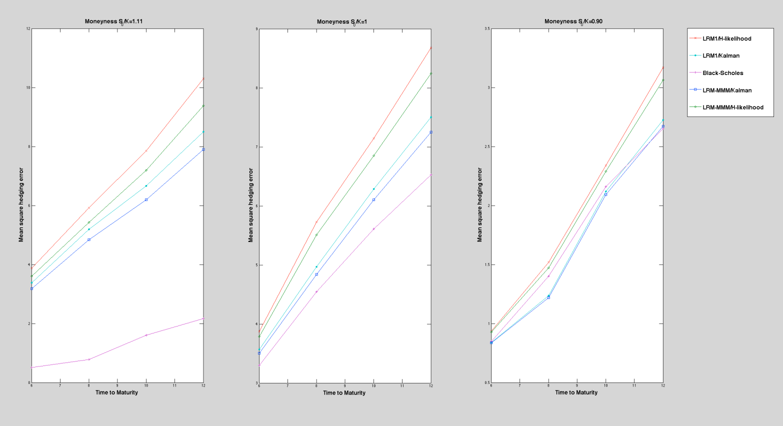

Figure: Experiment 3. The hedging techniques used are spelled out in the legend. The four maturities considered are 6, 8, 10, and 12 time steps. Hedging is carried out daily.

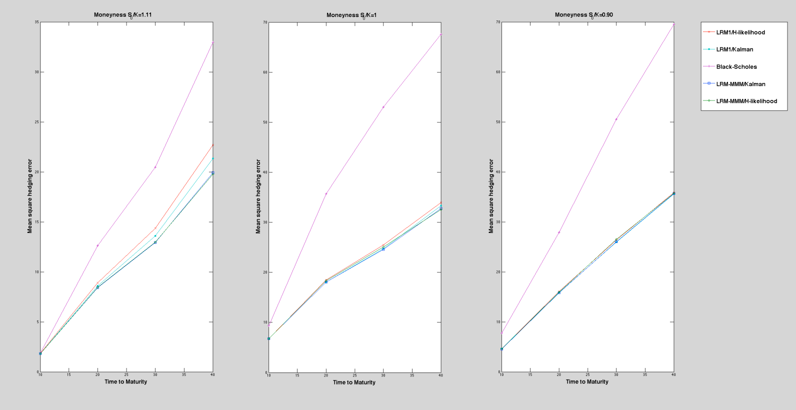

Figure: Experiment 4. The four maturities considered are 10, 20, 30, and 40 time steps. Hedging is carried out every 10 days.

Exercises 5 and 6:

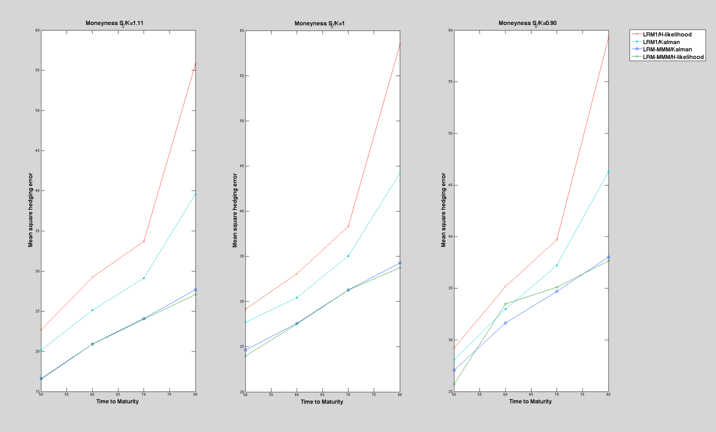

Figure: Experiment 5. The hedging techniques used are spelled out in the legend. The four maturities considered are 50, 60, 70, and 80 time steps. Hedging is carried out every 10 days.

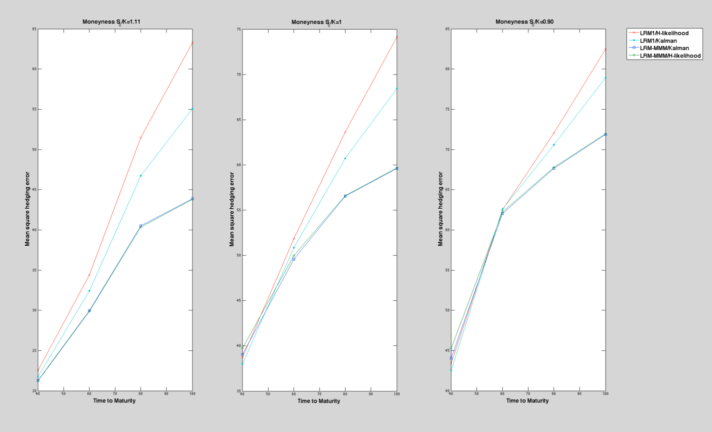

Figure: Experiment 6. The hedging techniques used are spelled out in the legend. The four maturities considered are 40, 60, 80, and 100 time steps. Hedging is carried out every 20 days.