Back to the past: option pricing using realized volatility

F. Corsi[§] N. Fusari[✠] D. La Vecchia[*]

[§]Swiss Finance Institute and University of Siena
[✠]Swiss Finance Institute, University of Lugano
[*]University of Lugano

12th October, Kellogg School of Management

Outline

Outline

Motivation

- High frequency data are nowdays wide available and easily tractable;
- considerable interest has been recently devoted to the use of high frequency (HF) data for measuring and forcasting volatility;
- volatility is a key ingredient of every option pricing model;
- only few papers try to employ high frequency data to price options or other derivatives.

GOAL: definition of a discrete-time model for option pricing using Realized Volatility (RV) as a proxy for the unobservable underlying volatility.

Outline

Previous work

Realized volatility measure and forcast:

- Andersen et al. (2001) and (2003),
- Barndorff-Nielsen and Shephard (2001) (2002a) (2002b) and (2005),
- Comte and Renault (1998);

Option pricing using RV:

- Stentoft (2008).

Outline

General idea

A well established result in the financial econometrics literature is that, daily log–returns do not have a Gaussian distribution: leptokurtic $\rightarrow$ heavy-tailed distributions.

Clark (1973) and Ane and Geman (2000): for an underlying semi-martingale process, rescaling the log–returns by an appropriate measure of the market activity allows to recover the standard Gaussian distribution.

Here we focus on **Realized Volatility** as a measure of market activity (which is per se not observable).

Outline

Model set-up

Methodological approach:

We cast our model in the general framework proposed by Bertholon, Monfort and Pegoraro (2008).
Specifically, in pricing derivatives, one can follow three different approaches:

- direct modeling ($\mathbb{P}$ and SDF);
- Risk-neutral constrained modeling ($\mathbb{Q}$ and $\mathbb{P}$);
- back modeling ($\mathbb{Q}$ and SDF).

Here we follow the first alternative (**direct modeling**).

The model under the historical measure

Log-returns

We assume the following dynamics for the log–returns:

$$\ln\left(\frac{S_{t+1}}{S_t}\right) := y_{t+1} = \mu_{t+1} + \sqrt{RV_{t+1}}\epsilon_{t+1}, \qquad (1)$$

$$\epsilon_{t+1} | RV_{t+1} \sim N(0,1).$$

In our notation, S_{t+1}, y_{t+1}, and RV_{t+1} are respectively the price, the log–returns and the RV at time $t+1$. For the drift of the log–returns (i.e. μ_{t+1}) we propose the following specification:

$$\mu_{t+1} = r + (\tilde{\gamma} - 1/2)\, RV_{t+1}.$$

The model under the historical measure (cont'd)

Realized volatility dynamics

Which process to describe the RV dynamics?

- Empirically we observe that the main feature of the RV is the strong persistence.

$\Longrightarrow$Therefore we follow Corsi (2009) and we model the conditional mean of the RV (given its past values) using the conditional expected value of an HAR (Heterogeneous AutoRegressive) process.

The model under the historical measure (cont'd)

Realized volatility dynamics

We model the RV as an Auto Regressive Gamma process (ARG) (see Gourieroux and Jasiack 2006). This implies that:

$$RV_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\delta, \beta'(\mathbf{RV}_t, L_t), c)$$

δ and c describe, respectively, the shape and the scale of the distribution, whereas $\beta'(\mathbf{RV}_t, L_t)$ is the location parameter given by:

$$\beta'(\mathbf{RV}_t, L_t) = \beta_1 RV_t + \beta_2(\sum_{i=1}^{4} RV_{t-i}) + \beta_3(\sum_{i=5}^{21} RV_{t-i}) + \Lambda L_t. \quad (2)$$

We label this model **Heterogeneous AutoRegressive Gamma (HARG)** process.

Outline

Model simulation

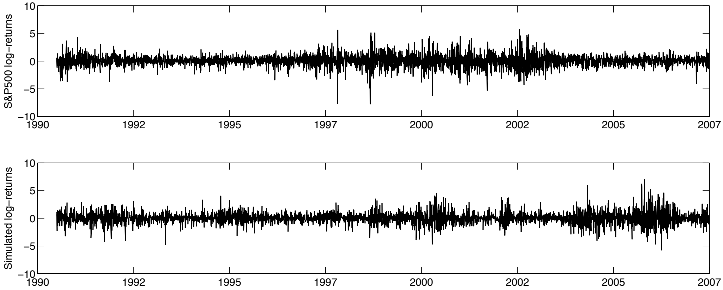
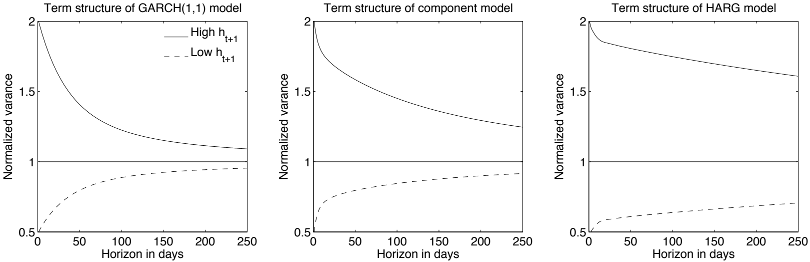
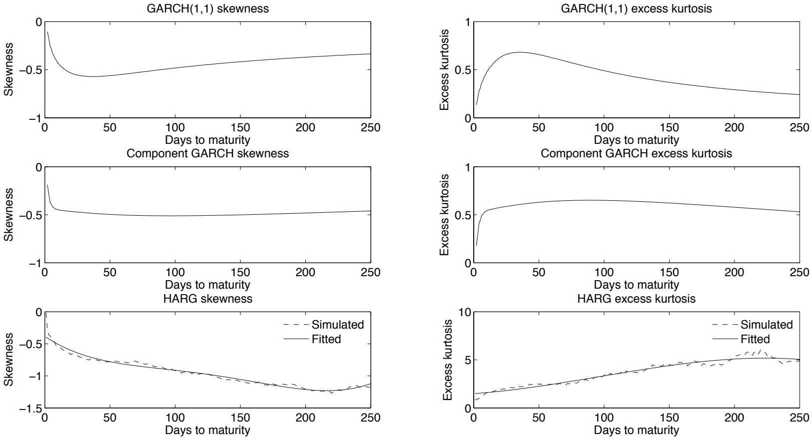


Figure: **Log-returns.** Top panel: S&P 500 log-returns from January 1st 1990 to December 31th 2007 (4218 observations). Bottom panel: log-returns simulated path from HARG(3).

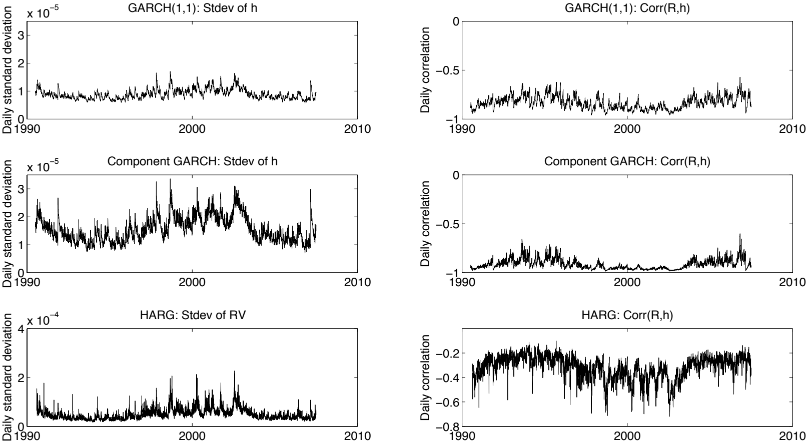
Variance term structure



Skewness-Kurtosis term structure



Volatility and Correlation



Outline

The stochastic discount factor

The HARG(3) is an "affine" process under $\mathbb{P}$.

In order to maintain the analytical tractability of the model under $\mathbb{Q}$ we set up an **exponential affine stochastic discount factor (SDF)**:

$$M_{t,t+1} = \exp(-\nu_0 - \nu_1 RV_{t+1} - \nu_2 RV_t^d - \nu_3 RV_{t-1}^w - \nu_4 RV_{t-5}^m - \nu_5 y_{t+1} - \nu_6 L_t).$$

$\Rightarrow$This SDF is compatible with the no-arbitrage conditions

(parameter restrictions - in the paper).

The (resulting) model under the risk neutral measure

Under $\mathbb{Q}$, the log–returns follow a discrete-time stochastic model, with risk premium $\gamma^* = -1/2$. The RV is an $HARG(3)$ process, featuring a transition density given by a non–central gamma, with parameters ρ^*, δ^*, μ^*.

$$\rho^* = \frac{c\beta}{(1 + c\lambda)^2}, \tag{3}$$

$$\delta^* = \delta, \tag{4}$$

$$c^* = \frac{c}{1 + c\lambda}, \tag{5}$$

with $\lambda = \nu_1 + \frac{\gamma^2}{2} - \frac{1}{8}$.

Dataset

- We employ a large dataset of options taken from OptionMetrics from 1st January 1996 to 27th December 2004.
- We consider only OTM Call-Put options with volatility less then 70%, maturity between 10 and 365 days and whose price is not less than 0.05 dollars.

Outline

Competing models

The direct modeling approach allows to estimate the HARG(3) model under the historical measure using the time series of log-returns and realized volatility.

Just one parameter (ν_1) has to be calibrated. From a practical point of view we use the most liquid ATM Call and Put (minimizing the $RMSE_{IV}$).

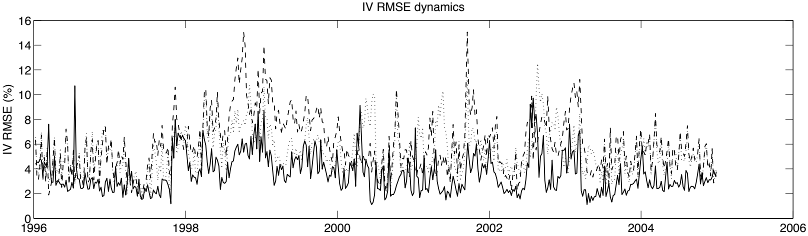
Competitor models: HN GARCH(1,1), Component GARCH.

Option pricing performace - Global

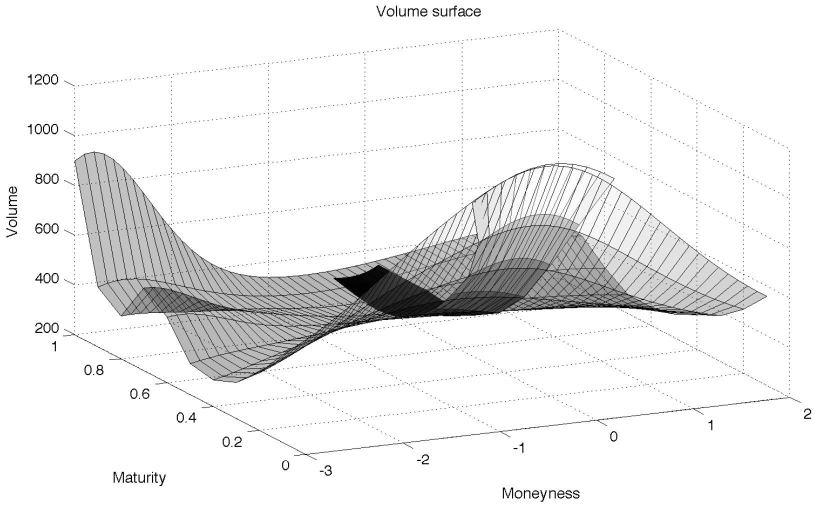
	HARG	*HN GARCH(1,1)*	*Component GARCH*
$RMSE_p$	4.2618	0.5708	0.6573
$RMSE_{IV}$	5.9838	0.6469	0.8126

Table: **Option pricing performance.**

RMSE dynamics



Nonparametric volume surface



HARG vs HN GARCH(1,1)

Moneyness	Maturity: Less then 20	20 to 60	60 to 160	More then 160
< -2	0.4591	0.5915	0.4318	0.3236
$[-2, -1)$	0.7771	0.7160	0.5286	0.4000
$[-1, -0.5)$	0.7542	0.8089	0.6328	0.4786
$[-0.5, 0.5)$	0.7367	0.8751	0.8087	0.6284
$[0.5, 1)$	0.9104	1.0530	1.1544	1.1529
$>= 1$	0.7406	0.7974	0.6444	0.9624

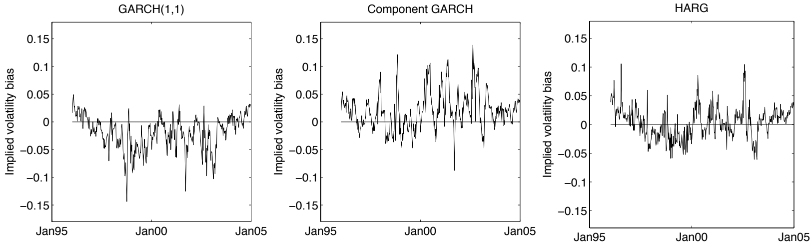
Table: **HARG(3) vs HN GARCH(1,1).**

HARG vs Component GARCH

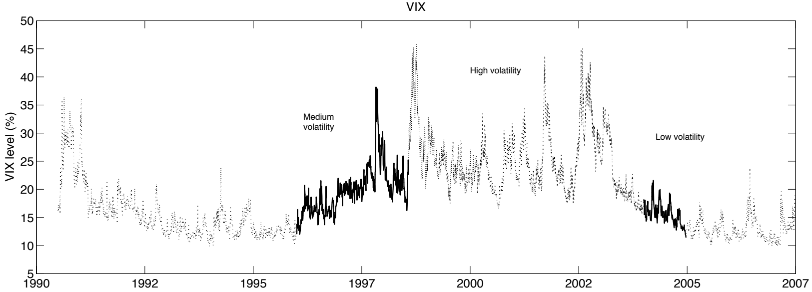
Moneyness	Maturity: Less then 20	20 to 60	60 to 160	More then 160
< -2	0.4535	0.6013	0.5022	0.4014
$[-2, -1)$	0.7786	0.7665	0.6507	0.5318
$[-1, -0.5)$	0.7702	0.8223	0.8044	0.6740
$[-0.5, 0.5)$	0.7864	0.7857	0.9132	0.9217
$[0.5, 1)$	1.0449	0.8144	0.8944	1.3573
$>= 1$	0.7141	0.9400	0.8419	1.2395

Table: **HARG(3) vs Component GARCH.**

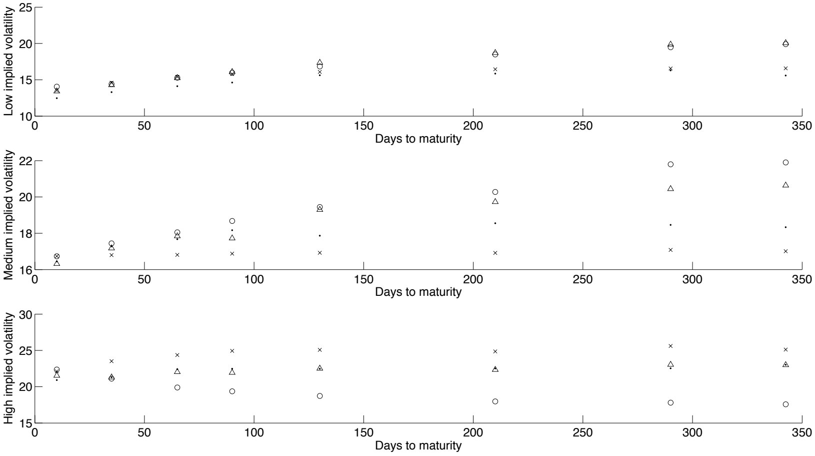
Option pricing performance - ATM options bias



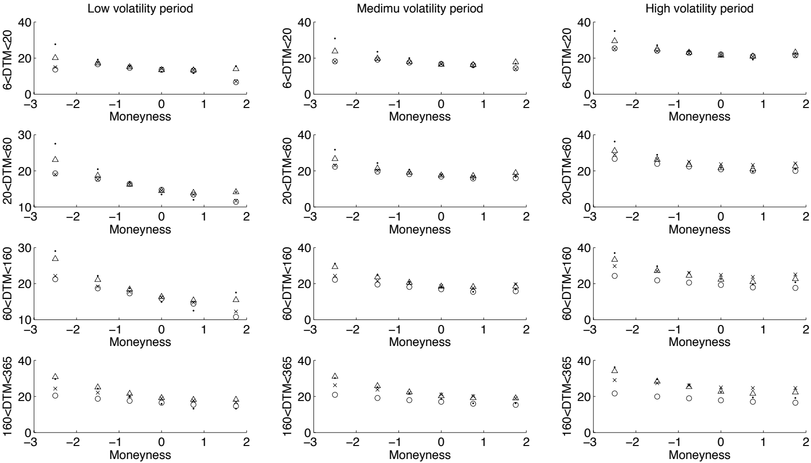
Volatility regimes



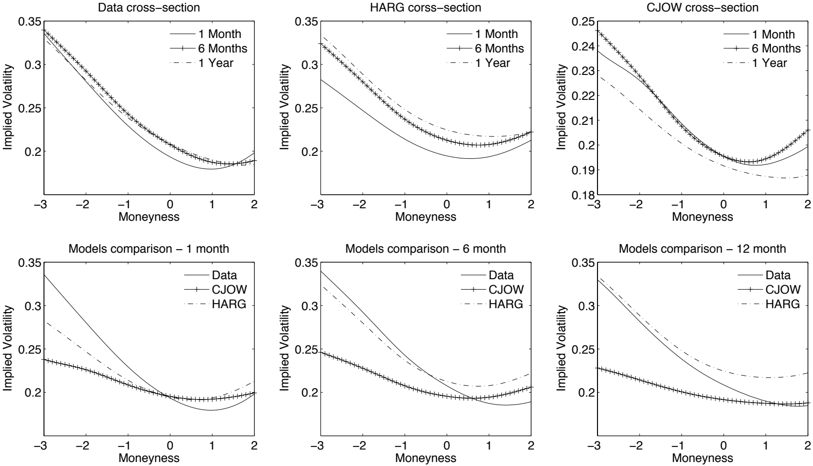
Option pricing performance - ATM options term structure



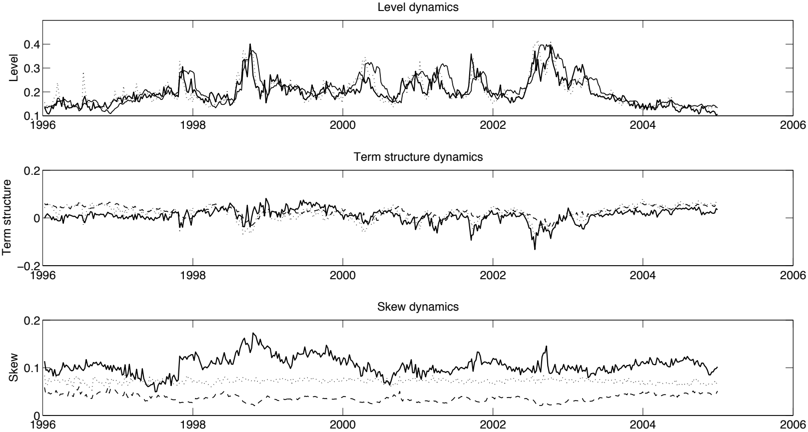
Option pricing performance - cross-section



Option pricing performance - cross-section



Option pricing performance - IV surface dynamics



Conclusions

- We develop a stochastic volatility option pricing model that exploits the historical information contained in the HF data.
- Straightforward estimation.
- No need to calibrate all the parameters.
- The HARG(3) outperform competing GARCH models.