Supplement to "Conditionally Heteroskedastic Factor Models with Skewness and Leverage Effects"

Prosper DOVONON*
Concordia University, CIREQ and CIRANO

February 9, 2012

Abstract

This supplemental material contains further details on the listed examples of models with affine conditional leverage, the Monte Carlo experiments assessing the finite sample performance of the GMM estimation of the conditionally heteroskedastic factor model with asymmetries (ACHF) proposed in "Conditionally Heteroskedastic Factor Models with Skewness and Leverage Effects" along with the proofs of the propositions stated in the same paper.

1 Further details on the supporting examples for affine the conditional leverage

This section gives further details on the listed examples in Section 3.2 of the main paper. In the following examples, all of the random variables with index t are assumed to be J_t-adapted:

Example 1 *The $\mathbb{A}_1(3)$-affine family processes*[1] *(Dai and Singleton (2000), Singleton (2001)). Let f_{t+1} be defined by*

$$\begin{aligned} f_{t+1} &= \sqrt{\alpha+v_t}\varepsilon_{1,t+1}+\sigma_1\eta\sqrt{v_t}\varepsilon_{2,t+1}, \\ v_{t+1} &= \xi\bar{v}+(1-\xi)v_t+\eta\sqrt{v_t}\varepsilon_{2,t+1}, \\ E\left(\varepsilon_{j,t+1}|J_t\right) &= E\left(\varepsilon_{1,t+1}\varepsilon_{2,t+1}|J_t\right)=0,\ j=1,2, \\ E\left(\varepsilon^2_{j,t+1}|J_t\right) &= 1,\ j=1,2, \end{aligned}$$

where $(\alpha,\eta,\xi,\bar{v},\sigma_1)\in\mathbb{D}$, a conveniently restricted subset of $\mathbb{R}^5$. It follows that $\sigma_t^2\equiv Var\left(f_{t+1}|J_t\right)=\alpha+(1+\sigma_1^2\eta^2)v_t$ and the affine process complies with Equation (5) in the main paper since:

$$Cov(f_{t+1},\,\sigma^2_{t+1}|J_t)=-\sigma_1\eta^2\alpha+\sigma_1\eta^2\sigma_t^2=\pi_0+\pi_1\sigma_t^2.$$

*Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8 Canada; tel: (514) 848-2424 (ext. 3479); email: prosper.dovonon@concordia.ca.

[1] Backus, Foresi and Telmer(2001) use the Euler scheme discrete time version of the Cox, Ingersoll and Ross' (1985) diffusion process to propose an affine model of currency. The affine process nests the square-root process of Heston (1993) and Cox, Ingersoll and Ross (1985).

Example 2 *The Quadratic GARCH (QGARCH(1,1)) of Sentana (1995). Let* f_{t+1} *be given by*

$$f_{t+1} = \sigma_t \, \eta_{t+1}, \qquad \eta_{t+1}|J_t \sim \mathcal{D}\,(0,\,1)\,; \quad \sigma_t^2 = \theta + \beta\,\sigma_{t-1}^2 + \alpha\,(f_t - \xi)^2\,, \quad (\theta,\,\beta,\,\alpha,\,\xi) \in \mathbb{D},$$

where $\mathcal{D}$ *is any symmetric distribution. It follows that the QGARCH(1,1) dynamics satisfies the affine representation since:*

$$Cov(f_{t+1},\, \sigma_{t+1}^2|J_t) = -2\alpha\xi\sigma_t^2 = \pi_1\sigma_t^2.$$

Example 3 *Heston-Nandi's (2000) GARCH process. Let* f_{t+1} *be given by*

$$f_{t+1} = \sigma_t \, \eta_{t+1}, \qquad \eta_{t+1}|J_t \sim \mathcal{N}\,(0,\,1)\,; \quad \sigma_{t+1}^2 = \omega + \beta\,\sigma_t^2 + \alpha\,(\eta_{t+1} - \gamma\sigma_t)^2\,, \quad (\omega,\,\beta,\,\alpha,\,\gamma) \in \mathbb{D}.$$

We can show that this model also satisfies the affine dynamic since:

$$Cov(f_{t+1},\, \sigma_{t+1}^2|J_t) = -2\alpha\gamma\sigma_t^2 = \pi_1\sigma_t^2,$$

Example 4 *The Inverse Gaussian GARCH(1,1) of Christoffersen, Heston and Jacobs (2006). The Inverse-Gaussian-GARCH(1,1) model proposed by Christoffersen, Heston and Jacobs (2006) is shown by the authors to be embedded in the class of SR-SARV(1) models and allows for a leverage effect that is an affine function of past conditional variance. In particular, they show for such a process* f_{t+1}*, that:* $Cov(f_{t+1},\, \sigma_{t+1}^2|J_t) = \pi_1\sigma_t^2$ *for some* π_1 *where* σ_t^2 *is the conditional variance of* f_{t+1}*.*

2 Monte Carlo results

The main goal of this section is to assess the finite sample performance of our estimation procedure for different values of the factor volatility persistence. Because the GMM inference results are known to be sensitive to the set of valid instruments that are used (see e.g. Andersen and Sørensen (1996)), we first investigate the relative performance of three sets of valid instruments. We evaluate the performance of each instrument set by the simulated bias, the root mean square error (RMSE), the median and the least absolute deviation (LAD) of the estimates it provides for the conditionally heteroskedastic factor model with asymmetries. The best of these instrument sets is subsequently used in our experiments for assessing the sensitivity of our inference procedure to the factor volatility persistence.

We simulate samples of three asset excess returns with null risk premia ($\mu = 0$) from a single factor model. The model considered is the following:

$$Y_t = \lambda f_t + U_t,$$

with $\lambda = (1, 1, 1)'$ and $U_t \sim i.i.d.N(0, \omega Id_3)$, $\omega = 0.35$, Id_3 is the identity matrix of size 3. In this model, the signal to noise ratio λ_i/ω is 2.86 which roughly matches the average signal to noise ratio found in previous empirical researches (see e.g. Fiorentini, Sentana and Shephard (2004)). The factor process f_t in all our experiments has a GARCH$(1,1)$ dynamics i.e. $f_t = \sigma_{t-1}\eta_t$, $\eta_t|J_{t-1} \sim (0,1)$ and $\sigma_t^2 = 1-\alpha-\beta+\alpha f_t^2+\beta\sigma_{t-1}^2$, $0 < \alpha+\beta < 1$. This GARCH$(1,1)$ process corresponds to an SR-SARV(1) process with persistence parameter $\gamma = \alpha + \beta$.

We simulate η_t from two different distributions that distinguish two Monte Carlo designs. In *Design 1*, $\eta_t \sim NID(0,1)$ and hence the factor has a standard Gaussian GARCH$(1,1)$ dynamics. In *Design 2*, $\eta_t = (\sigma_{t-1}^2 - X_t)/\sigma_{t-1}$ where $X_t|J_{t-1} \sim$ Gamma$(\sigma_{t-1}^2, 1)$. In this case, $E(f_{t+1}^3|J_t) = -2\sigma_t^2$ and $Cov(f_{t+1}, \sigma_{t+1}^2|J_t) = -2\alpha\sigma_t^2$. Hence, *Design 2* fits with the occurrence of conditional skewness and leverage in $Y_{i,t}$, $i = 1,2,3$ while the conditional skewness and leverage are both 0 in *Design 1*.

For each design, we consider four volatility persistence configurations: $\gamma = 0.7$ ($\alpha = 0.2$, $\beta = 0.5$); $\gamma = 0.8$ ($\alpha = 0.2$, $\beta = 0.6$); $\gamma = 0.9$ ($\alpha = 0.2$, $\beta = 0.7$) and $\gamma = 0.95$ ($\alpha = 0.2$, $\beta = 0.75$). These values for the volatility persistence are chosen to reflect the range of estimates in empirical researches. In particular, the estimated volatility persistence in our empirical application in Section 5 range from 0.68 to 0.87 for daily data. Also, $\gamma = 0.80$ matches approximately the factor volatility persistence estimate by Fiorentini, Sentana and Shephard (2004) for monthly U.K. index excess returns while $\gamma = 0.90$ and $\gamma = 0.95$ are the usual range of the standard GARCH volatility persistence estimate in the empirical literature for daily returns (see e.g. Harvey and Siddique (1999)). We set the number of replications to 1,000, and the sample sizes that we consider are: $T = 500$, $T = 1,000$, and $T = 2,000$.

We perform the inference by the normalization approach described in Section 4 setting the first asset's factor loading to $\bar{\lambda} = 1$. The moment conditions (12)-(13)-(14)-(15)-(16)-(17)-(18) in the main text are used in the estimation process with specified instruments. In *Design 1*, $s_1 = s_2 = s_3 = s = 0$, $h_1 = 0$ and $\pi_0 = \pi_1 = 0$ and in *Design 2*, $s_1 = s_2 = s_3 = s = 0$, $h_1 = -2.0$, $\pi_0 = 0$ and $\pi_1 = -0.4$. Therefore, as far as the Monte Carlo designs are concerned, λ_2, λ_3, ω, γ, s, h_1, π_0 and π_1 are the only relevant parameters of our conditionally heteroskedastic factor model with asymmetries.

We first assess the relative performance of the estimation procedure for our model in terms of the instruments used. In particular, we consider three sets of instruments: $z_{1,t} = (1, Y_{1,t}^2, Y_{1,t-1}^2, Y_{1,t-2}^2)$, $z_{2,t} = (1, |Y_{i,t}|, |Y_{i,t-1}|, |Y_{i,t-2}|)$, and $z_{3,t} = (1, \sum_{i=1}^3 |Y_{i,t}|, \sum_{i=1}^3 |Y_{i,t-1}|, \sum_{i=1}^3 |Y_{i,t-2}|)$. For each set of instruments, we simulate data from *Design 1* with $\gamma = 0.9$ and evaluate the simulated bias, Root mean square error (RMSE), median and least absolute deviation (LAD) for the parameter estimates. The results contained in Table MC1 reveal that the three sets of instruments lead to estimates of smaller bias, RMSE and LAD as the sample size increases. The factor loadings are estimated with very small bias throughout, even in the smallest samples. The volatility persistence suffers more of bias in small samples while the conditional third moment parameters are slightly less precise as they display larger

RMSE and LAD. This table also allows some ranking of the instruments sets. It appears that z_{1t} leads to estimates less precise and with larger bias than z_{2t} which in turn is marginally better than z_{3t}. An intuition as to why z_{2t} and z_{3t} dominate z_{1t} may be related to the fact that we are in presence of conditionally heteroskedastic processes which, in general are not guaranteed to have finite higher moments. Both z_{2t} and z_{3t} are made of $|Y_{i,t}|$-like components while z_{1t} is made of $Y_{i,t}^2$-like components. This latter, even though very attractive for the ARMA structure in the square factor process, requires that higher moments of the processes being bounded to perform well in comparison with z_{2t} and z_{3t}.

Comparing z_{2t} with z_{3t}, we first observe that z_{3t}–by construction–encapsulates more extensively the heteroskedastic directions of each series and is qualitatively better than z_{2t}. The marginal out-performance of z_{2t} over z_{3t} does not offset in our opinion this interest. We keep z_{3t} as instrument for the next Monte Carlo experiment where we evaluate the effect of increasing persistence and for our empirical applications as well.

The results of the second Monte Carlo experiment are displayed by Table MC2. In this experiment, we increase the value of volatility persistence for both *Design 1* and *Design 2* and check the distribution of the estimates through the usual indicators. It appears that, for the same persistence, the estimates from *Design 1* have a smaller bias and RMSE as well as a smaller LAD than the estimates from the heavy tailed distribution which is *Design 2*. It is worth pointing out that, as the persistence increases, the precision of all the estimates sharply decreases.

This Monte Carlo experiment suggests that our inference procedure is reliable in finite samples particularly when the volatility persistence is not close to 1, whereas the inference could be inaccurate for persistence values larger than 0.95. This observation seems to confirm a well known drawback of the GMM's application in volatility literature which delivers bad results when the volatility persistence is close to 1 (see e.g. Broto and Ruiz (2004)).

Table MC1

Simulated bias, root mean square error (RMSE), median and least absolute deviation (LAD) of GMM parameter estimates of the conditionally heteroskedastic factor model with asymmetries. We report the results from GMM estimates using 3 different sets of valid instruments:
$z_{1,t} = (1, Y_{1,t}^2, Y_{1,t-1}^2, Y_{1,t-2}^2)$,
$z_{2,t} = (1, |Y_{1,t}|, |Y_{1,t-1}|, |Y_{1,t-2}|)$, and
$z_{3,t} = (1, \sum_{i=1}^{3} |Y_{i,t}|, \sum_{i=1}^{3} |Y_{i,t-1}|, \sum_{i=1}^{3} |Y_{i,t-2}|)$.
The simulated data are obtained from *Design 1* for $\gamma = 0.9$. The true parameter values are $\lambda_2 = \lambda_3 = 1, \omega = 0.35, s = h = \pi_0 = \pi_1 = 0$.

T	500				1000				2000			
	Bias	Rmse	Med.	LAD	Bias	Rmse	Med.	LAD	Bias	Rmse	Med.	LAD
					z_{1t}							
λ_2	-0.002	0.003	0.999	0.042	-0.002	0.001	0.998	0.028	-0.003	0.002	0.997	0.021
λ_3	-0.001	0.003	0.997	0.041	-0.001	0.001	0.998	0.028	-0.003	0.002	0.997	0.020
γ	-0.147	0.087	0.826	0.200	-0.077	0.036	0.867	0.135	-0.030	0.014	0.893	0.091
ω	-0.024	0.002	0.326	0.030	-0.014	0.000	0.336	0.017	-0.007	0.000	0.343	0.010
s	0.001	3.199	-0.041	1.201	-0.023	2.208	-0.064	0.974	-0.002	1.337	-0.018	0.776
h	0.015	3.592	0.056	1.515	0.026	2.466	0.096	1.167	0.003	1.573	0.026	0.909
π_0	0.027	1.207	-0.030	0.550	0.026	0.524	0.035	0.438	0.005	0.330	-0.004	0.352
π_1	0.002	0.729	0.053	0.641	-0.042	0.506	-0.031	0.507	-0.015	0.325	0.003	0.405
					z_{2t}							
λ_2	-0.003	0.002	0.996	0.039	-0.002	0.001	0.997	0.026	-0.003	0.001	0.997	0.018
λ_3	-0.002	0.002	0.996	0.039	-0.002	0.001	0.997	0.026	-0.002	0.000	0.998	0.017
γ	-0.082	0.054	0.911	0.156	-0.044	0.023	0.902	0.111	-0.014	0.011	0.912	0.082
ω	-0.023	0.001	0.327	0.025	-0.013	0.000	0.337	0.015	-0.007	0.000	0.343	0.009
s	-0.014	1.412	0.003	0.930	-0.012	0.970	-0.020	0.732	-0.010	0.651	-0.011	0.581
h	0.026	2.510	0.053	1.284	0.026	1.504	0.040	0.942	0.014	0.908	0.007	0.717
π_0	-0.007	0.330	-0.011	0.417	0.043	0.225	0.036	0.350	0.013	0.213	0.004	0.294
π_1	0.011	0.529	0.036	0.539	-0.047	0.343	-0.041	0.433	-0.017	0.265	-0.010	0.353
					z_{3t}							
λ_2	-0.002	0.002	0.997	0.038	-0.002	0.001	0.997	0.025	-0.002	0.000	0.997	0.017
λ_3	0.000	0.002	0.998	0.038	-0.001	0.001	0.999	0.025	-0.002	0.000	0.999	0.016
γ	-0.087	0.062	0.916	0.160	-0.054	0.033	0.901	0.121	-0.018	0.013	0.903	0.082
ω	-0.023	0.001	0.328	0.025	-0.013	0.000	0.337	0.015	-0.007	0.000	0.344	0.009
s	-0.041	1.862	-0.058	1.059	-0.011	1.331	0.002	0.827	0.033	1.325	0.011	0.680
h	0.066	3.275	0.117	1.474	0.020	2.073	0.020	1.089	-0.027	1.401	-0.033	0.837
π_0	0.005	0.494	-0.016	0.480	0.033	0.295	0.022	0.379	0.020	0.261	0.003	0.309
π_1	-0.002	0.711	0.027	0.620	-0.040	0.444	-0.038	0.479	-0.024	0.326	-0.019	0.377

Table MC2

Simulated Bias, root mean square error (RMSE), median and least absolute deviation (LAD) of GMM parameter estimates of the conditionally heteroskedastic factor model with asymmetries. We report the results from GMM estimates using $z_{3,t} = (1, \sum_{i=1}^{3} |Y_{i,t}|, \sum_{i=1}^{3} |Y_{i,t-1}|, \sum_{i=1}^{3} |Y_{i,t-2}|)$ as instrument. The data are generated according to *Design 1* and *Design 2*. In both designs, $\lambda_2 = \lambda_3 = 1, \omega = 0.35$, and $s = \pi_0 = 0$ while $h_1 = \pi_1 = 0$ for *Design 1* and $h_1 = -2.0$ and $\pi_1 = -0.4$ for *Design 2*. The values of γ are 0.7, 0.8, 0.9 and 0.95; $T = 2,000$.

	Design 1				*Design 2*			
	Bias	Rmse	Med.	LAD	Bias	Rmse	Med.	LAD
				$\gamma = 0.70$				
λ_2	-0.002	0.000	0.998	0.017	-0.004	0.006	0.999	0.021
λ_3	-0.002	0.000	0.998	0.017	-0.009	0.017	0.997	0.025
γ	-0.021	0.030	0.687	0.133	-0.033	0.033	0.675	0.137
ω	-0.007	0.000	0.343	0.009	-0.006	0.001	0.343	0.010
s	0.038	0.665	0.084	0.657	-0.134	1.993	-0.186	0.995
h	-0.048	0.903	-0.091	0.772	0.438	3.294	-1.483	1.318
π_0	0.025	0.168	0.015	0.300	-0.083	0.412	-0.127	0.470
π_1	-0.032	0.221	-0.032	0.345	0.211	0.693	-0.126	0.631
				$\gamma = 0.80$				
λ_2	-0.002	0.000	0.997	0.017	-0.011	0.024	0.999	0.029
λ_3	-0.002	0.000	0.999	0.017	-0.011	0.024	1.000	0.029
γ	-0.018	0.022	0.798	0.114	-0.031	0.027	0.786	0.123
ω	-0.007	0.000	0.344	0.009	-0.005	0.001	0.343	0.011
s	0.021	0.671	0.036	0.647	-0.110	1.763	-0.132	0.924
h	-0.027	0.938	-0.054	0.777	0.430	3.229	-1.520	1.294
π_0	0.022	0.179	0.007	0.296	-0.081	0.342	-0.119	0.414
π_1	-0.029	0.241	-0.025	0.350	0.218	0.604	-0.106	0.584
				$\gamma = 0.90$				
λ_2	-0.002	0.000	0.997	0.017	-0.046	0.112	0.997	0.064
λ_3	-0.002	0.000	0.999	0.016	-0.040	0.097	0.999	0.059
γ	-0.018	0.013	0.903	0.082	-0.026	0.011	0.889	0.082
ω	-0.007	0.000	0.344	0.009	0.004	0.006	0.344	0.019
s	0.033	1.325	0.011	0.680	0.003	1.956	-0.123	0.881
h	-0.027	1.401	-0.033	0.837	0.373	3.436	-1.458	1.314
π_0	0.020	0.261	0.003	0.309	-0.051	0.314	-0.093	0.387
π_1	-0.024	0.326	-0.019	0.377	0.198	0.621	-0.137	0.591
				$\gamma = 0.95$				
λ_2	-0.003	0.001	0.997	0.019	-0.084	0.193	0.996	0.104
λ_3	-0.002	0.000	0.998	0.017	-0.080	0.186	0.997	0.102
γ	-0.021	0.005	0.968	0.051	-0.021	0.007	0.967	0.052
ω	-0.007	0.000	0.344	0.009	0.025	0.034	0.344	0.041
s	-0.001	1.657	0.000	0.737	0.064	4.551	-0.115	0.905
h	-0.002	2.053	-0.038	0.963	0.475	4.723	-1.307	1.429
π_0	0.017	0.288	0.013	0.316	-0.015	0.461	-0.080	0.357
π_1	-0.026	0.368	-0.021	0.409	0.204	0.733	-0.087	0.614

3 Proofs of Propositions

Proof of Proposition 3.1: The expression given in (9) is obvious and arises from the sum of (1) over the time period: $\tau = (t-1)m+1$ through tm with the respective aggregation coefficients α_l and $\mu(J_t) = \mu$. Let $\left(F_{(t+1)m}^{(m)}\right)$ and $\left(U_{(t+1)m}^{(m)}\right)$ be the resulting factor and the idiosyncratic shocks and let $D_{tm}^{(m)}$ be the $J_{tm}^{(m)}$-conditional variance of this factor. We have

$$D^{(m)}_{tm} = E\left(F^{(m)}_{(t+1)m}F^{(m)'}_{(t+1)m}|J^{(m)}_{tm}\right) - E\left(F^{(m)}_{(t+1)m}|J^{(m)}_{tm}\right)E\left(F^{(m)'}_{(t+1)m}|J^{(m)}_{tm}\right) = E\left(F^{(m)}_{(t+1)m}F^{(m)'}_{(t+1)m}|J^{(m)}_{tm}\right).$$

We can actually show that $E\left(F^{(m)}_{(t+1)m}|J^{(m)}_{tm}\right) = 0$ as we do below.

$$\begin{aligned} E\left(F^{(m)}_{(t+1)m}|J_{tm}\right) &= E\left(\left(\sum_{l=1}^m \alpha_l F_{tm+l}\right)|J_{tm}\right) = \sum_{l=1}^m \alpha_l E\left(F_{tm+l}|J_{tm}\right) \\ &= \sum_{l=1}^m \alpha_l E\left(E\left(F_{tm+l}|J_{tm+l-1}\right)|J_{tm}\right) = 0. \end{aligned}$$

The third equality holds by the law of iterated expectations and the last one comes from (2). Since $J^{(m)}_{tm}$ is included in J_{tm} by definition, the law of iterated expectations also applies the following way: $E(X|J^{(m)}_{tm}) = E\left(E(X|J_{tm})|J^{(m)}_{tm}\right)$ for any measurable variable X. Therefore, $E\left(F^{(m)}_{(t+1)m}|J^{(m)}_{tm}\right) = 0$.

We now show that $D^{(m)}_{tm}$ is diagonal. Let us now consider k and k' such that $k \neq k'$.

$$\begin{aligned} E\left(F^{(m)}_{k,(t+1)m}F^{(m)}_{k',(t+1)m}|J^{(m)}_{tm}\right) &= E\left(\left(\sum_{l=1}^m \alpha_l F_{k,tm+l}\right)\left(\sum_{l=1}^m \alpha_l F_{k',tm+l}\right)|J^{(m)}_{tm}\right) \\ &= E\Big[\sum_{l<l';l,l'=1}^m \alpha_l\alpha_{l'}\left(F_{k,tm+l}F_{k',tm+l'} + F_{k',tm+l}F_{k,tm+l'}\right) \\ &\quad + \sum_{l=1}^m \alpha_l^2 F_{k,tm+l}F_{k',tm+l}|J^{(m)}_{tm}\Big]. \end{aligned}$$

But, from the law of iterated expectations and (2), for $l < l'$,
$E\left(F_{k,tm+l}F_{k',tm+l'}|J_{tm}\right) = E\left(F_{k,tm+l}E\left(F_{k',tm+l'}|J_{tm+l'-1}\right)|J_{tm}\right) = 0$ and in addition, as D_t is diagonal for all t from (2), $E\left(F_{k,tm+l}F_{k',tm+l}|J_{tm}\right) = E\left(E\left(F_{k,tm+l}F_{k',tm+l}|J_{tm+l-1}\right)|J_{tm}\right) = E\left(D_{k,k',tm+l-1}|J_{tm}\right) = 0$.
By the law of iterated expectations as above we can deduce that $E\left(F^{(m)}_{k,(t+1)m}F^{(m)}_{k',(t+1)m}|J^{(m)}_{tm}\right) = 0$ and therefore $D^{(m)}_{tm}$ is diagonal.

Regarding the aggregated idiosyncratic shocks, by the law of iterated expectations and simple product expansion, we can show that $E\left(U^{(m)}_{(t+1)m}|J^{(m)}_{tm}\right) = 0$ and $E\left(U^{(m)}_{(t+1)m}F^{(m)'}_{(t+1)m}|J^{(m)}_{tm}\right) = 0$. Also,

$$\begin{aligned} Var\left(U^{(m)}_{(t+1)m}|J^{(m)}_{tm}\right) &= E\left(\left(\sum_{l=1}^m \alpha_l U_{tm+l}\right)\left(\sum_{l=1}^m \alpha_l U_{tm+l}\right)'|J^{(m)}_{tm}\right) \\ &= E\Big(\sum_{l<l';l,l'=1}^m \alpha_l\alpha_{l'}\left(U_{tm+l}U'_{tm+l'} + U_{tm+l'}U'_{tm+l}\right) \\ &\quad + \sum_{l=1}^m \alpha_l^2 U_{tm+l}U'_{tm+l}|J^{(m)}_{tm}\Big) \end{aligned}$$

For $l < l'$, $E\left(U_{tm+l}U'_{tm+l'}|J_{tm}\right) = E\left(U_{tm+l}E\left(U'_{tm+l'}|J_{tm+l'-1}\right)|J_{tm}\right) = 0$
and $E\left(U_{tm+l}U'_{tm+l}|J_{tm}\right) = E\left(E\left(U_{tm+l}U'_{tm+l}|J_{tm+l-1}\right)|J_{tm}\right) = E\left(\Omega|J_{tm}\right) = \Omega$ thus, $Var\left(U^{(m)}_{(t+1)m}|J^{(m)}_{tm}\right) = \Omega\sum_{l=1}^m \alpha_l^2$ completing the proof of Proposition 3.1□

Proof of Proposition 3.2: Since (f_{t+1}) has a SR-SARV(1) dynamic, with $v_{t+1} \equiv \sigma^2_{t+1} - (1-\gamma) - \gamma\sigma^2_t$, we have: $E(v_{t+1}|J_t) = 0$. For $l = 1$, the first conclusion of the proposition is obvious since σ^2_{tm} is J_{tm}-measurable. For $l \geq 2$, by writing v_{t+1} for different time and making some simple substitutions, we can write:

$$\sigma^2_{tm+l-1} = (1-\gamma)(1+\gamma+\gamma^2+\cdots+\gamma^{l-2}) + \gamma^{l-1}\sigma^2_{tm} + \gamma^{l-2}v_{tm+1} + \gamma^{l-3}v_{tm+2} + \cdots + v_{tm+l-1}.$$

By taking the expectation conditionally on J_{tm} and by the law of iterated expectations, we have:

$$\begin{aligned} E(\sigma^2_{tm+l-1}|J_{tm}) &= (1-\gamma)(1+\gamma+\gamma^2+\cdots+\gamma^{l-2}) + \gamma^{l-1}\sigma^2_{tm} \\ &= (1-\gamma)\tfrac{1-\gamma^{l-1}}{1-\gamma} + \gamma^{l-1}\sigma^2_{tm} \\ &= 1-\gamma^{l-1} + \gamma^{l-1}\sigma^2_{tm}. \end{aligned}$$

The first conclusion is then established. Moreover,

$$
\begin{aligned}
E\left((f_{(t+1)m}^{(m)})^2|J_{tm}\right) &= E\left(\sum_{l=1}^m(\alpha_l f_{tm+l})^2|J_{tm}\right), \\
&\quad \text{since } f_t \text{ is conditionally non-autocorrelated} \\
&= \sum_{l=1}^m \alpha_l^2 E(f_{tm+l}^2|J_{tm}) \\
&= \sum_{l=1}^m \alpha_l^2 E(E(f_{tm+l}^2|J_{tm+l-1})|J_{tm}) \\
&= \sum_{l=1}^m \alpha_l^2 E(\sigma_{tm+l-1}^2|J_{tm}) \\
&= \sum_{l=1}^m \alpha_l^2[(1-\gamma^{l-1})+\gamma^{l-1}\sigma_{tm}^2]
\end{aligned}
$$

Since σ_{tm}^2 is $J_{tm}^{(m)}$-measurable, $E\left((f_{(t+1)m}^{(m)})^2|J_{tm}^{(m)}\right) = \sum_{l=1}^m \alpha_l^2[(1-\gamma^{l-1})+\gamma^{l-1}\sigma_{tm}^2]$. Hence,

$$
\sigma_{tm}^{(m)^2} \equiv Var\left(f_{(t+1)m}^{(m)}|J_{tm}^{(m)}\right) = S_1^{(m)} + S_2^{(m)}\sigma_{tm}^2
$$

with $S_1^{(m)} = \sum_{l=1}^m \alpha_l^2[(1-\gamma^{l-1})$ and $S_2^{(m)} = \sum_{l=1}^m \alpha_l^2\gamma^{l-1}$. This completes the proof of Proposition 3.2. □

Proof of Proposition 3.3:

$$
\begin{aligned}
Cov\left(f_{(t+1)m}^{(m)}, \sigma_{(t+1)m}^{(m)^2}|J_{tm}\right) &= \sum_{l=1}^m \alpha_l Cov\left(f_{tm+l}, \sigma_{(t+1)m}^{(m)^2}|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l Cov\left(f_{tm+l}, S_1^{(m)} + S_2^{(m)}\sigma_{(t+1)m}^2|J_{tm}\right) \\
&\quad \text{(from Proposition 3.2)} \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)} Cov\left(f_{tm+l}, \sigma_{(t+1)m}^2|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)} E\left(f_{tm+l}\sigma_{(t+1)m}^2|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)} E\left(f_{tm+l} E\left(\sigma_{(t+1)m}^2|J_{tm+l}\right)|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)} E\left(f_{tm+l}\left(1-\gamma^{m-l}+\gamma^{m-l}\sigma_{tm+l}^2\right)|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)}\gamma^{m-l} E\left(f_{tm+l}\sigma_{tm+l}^2|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)}\gamma^{m-l} E\left(E\left(f_{tm+l}\sigma_{tm+l}^2|J_{tm+l-1}\right)|J_{tm}\right) \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)}\gamma^{m-l} E\left(\pi_0 + \pi_1\sigma_{tm+l-1}^2|J_{tm}\right) \\
&\quad \text{(from the dynamics of the conditional leverage in Equation (5))} \\
&= \sum_{l=1}^m \alpha_l S_2^{(m)}\gamma^{m-l}\left(\pi_0 + \pi_1 E\left(\sigma_{tm+l-1}^2|J_{tm}\right)\right) \\
&\equiv \ell_1^{(m)} + \ell_2^{(m)}\sigma_{tm}^2, \text{ (Proposition 3.2); } \ell_1^{(m)} \text{ and } \ell_2^{(m)} \text{ are two scalars.}
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
Cov\left(f_{(t+1)m}^{(m)}, \sigma_{(t+1)m}^{(m)^2}|J_{tm}^{(m)}\right) &= E\left(f_{(t+1)m}^{(m)}\sigma_{(t+1)m}^{(m)^2}|J_{tm}^{(m)}\right) = E\left[E\left(f_{(t+1)m}^{(m)}\sigma_{(t+1)m}^{(m)^2}|J_{tm}\right)|J_{tm}^{(m)}\right] \\
&= E\left[\ell_1^{(m)} + \ell_2^{(m)}\sigma_{tm}^2|J_{tm}^{(m)}\right].
\end{aligned}
$$

Since σ_{tm}^2 is $J_{tm}^{(m)}$-measurable, $Cov\left(f_{(t+1)m}^{(m)}, \sigma_{(t+1)m}^{(m)^2}|J_{tm}^{(m)}\right) = \ell_1^{(m)} + \ell_2^{(m)}\sigma_{tm}^2$ and from the one to one mapping between σ_{tm}^2 and $\sigma_{tm}^{(m)^2}$ from Proposition 3.2 we can deduce that there exists two scalars $\pi_0^{(m)}$ and $\pi_1^{(m)}$ such that $Cov\left(f_{(t+1)m}^{(m)}, \sigma_{(t+1)m}^{(m)^2}|J_{tm}^{(m)}\right) \equiv \pi_0^{(m)} + \pi_1^{(m)}\sigma_{tm}^{(m)^2}$ □

Proof of Proposition 3.4: In the following, $E_t(\star)$ denotes $E(\star|J_t)$.

$$
\begin{aligned}
E_{tm}\left(\left(f_{(t+1)m}^{(m)}\right)^3\right) &= E_{tm}\left(\left(\sum_{l=1}^m \alpha_l f_{tm+l}\right)^3\right) \\
&= \sum_{l=1}^m \alpha_l^3 E_{tm}(f_{tm+l}^3) + 3\times\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2 E_{tm}(f_{tm+l}f_{tm+l'}^2), \\
&\quad \text{by the specifications in (2).} \\
&= \sum_{l=1}^m \alpha_l^3 E_{tm}(f_{tm+l}^3) + 3\times\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2 E_{tm}\left(f_{tm+l}E_{tm+l'-1}(f_{k,tm+l'}^2)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{l=1}^m \alpha_l^3 E_{tm}(f_{tm+l}^3) + 3\times\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2 E_{tm}\left(f_{tm+l}E_{tm+l}\sigma_{tm+l'-1}^2\right) \\
&= \sum_{l=1}^m \alpha_l^3 E_{tm}(f_{tm+l}^3) + 3\times\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2 E_{tm}\left(f_{tm+l}\gamma^{l'-l-1}\sigma_{tm+l}^2\right) \\
&= \sum_{l=1}^m \alpha_l^3 E_{tm}(E_{tm+l-1}f_{tm+l}^3) + 3\times\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2\gamma^{l'-l-1} E_{tm}(E_{tm+l-1}(f_{tm+l}\sigma_{tm+l}^2)) \\
&= \sum_{l=1}^m \alpha_l^3 E_{tm}(h_0 + h_1\sigma_{tm+l-1}^2) + 3\times\sum\alpha_l\alpha_{l'}^2\gamma^{l'-l-1}E_{tm}\left(\pi_0+\pi_1\sigma_{tm+l-1}^2\right) \\
&= \sum_{l=1}^m \alpha_l^3[h_0 + h_1(1-\gamma^{l-1}+\gamma^{l-1}\sigma_{tm}^2)] \\
&\quad +3\times\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2\gamma^{l'-l-1}\left[\pi_0+\pi_1\left(1-\gamma^{l-1}+\gamma^{l-1}\sigma_{tm}^2\right)\right], \\
&\quad \text{from Proposition 3.2.} \\
&= \sum_{l=1}^m \alpha_l^3\left[h_0+(1-\gamma^{l-1})h_1\right] + 3\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2\gamma^{l'-l-1}[\pi_0+\pi_1(1-\gamma^{l-1})] \\
&\quad + \left[h_1\sum_{l=1}^m\alpha_l^3\gamma^{l-1} + 3\pi_1\sum_{1\le l<l'\le m}\alpha_l\alpha_{l'}^2\gamma^{l'-2}\right]\sigma_{tm}^2 \\
&\equiv B_0^{(m)} + B_1^{(m)}\sigma_{tm}^2
\end{aligned}
$$

Since σ_{tm}^2 is $J_{tm}^{(m)}$-measurable, the law of iterated expectations implies that $E\left((f_{(t+1)m}^{(m)})^3|J_{tm}^{(m)}\right) = B_0^{(m)} + B_1^{(m)}\sigma_{tm}^2$. From Proposition 3.2, $E\left((f_{(t+1)m}^{(m)})^3|J_{tm}^{(m)}\right) = h_0^{(m)} + h_1^{(m)}\sigma_{tm}^{(m)^2}$ with $h_0^{(m)} = B_0^{(m)} - h_1^{(m)}S_1^{(m)}$ and $h_1^{(m)} = B_1^{(m)}/S_2^{(m)}$; $S_1^{(m)}$ and $S_2^{(m)}$ are defined as in Proposition 3.2.

Also,

$$
\begin{aligned}
E\left(\left(U_{i,(t+1)m}^{(m)}\right)^3|J_{tm}\right) &= E\left(\left(\sum_{l=1}^m\alpha_l^3U_{i,tm+l}^3\right)|J_{tm}\right) \text{ from the specifications in (2)} \\
&= E\left(\left(\sum_{l=1}^m\alpha_l^3E(U_{i,tm+l}^3|J_{tm+l-1})\right)|J_{tm}\right) \\
&= E\left(\left(\sum_{l=1}^m\alpha_l^3 s_i^0\right)|J_{tm}\right) = s_i^0\left(\sum_{l=1}^m\alpha_l^3\right), \text{ from Assumption 2-}\textit{(ii)}
\end{aligned}
$$

hence,

$$
E\left(\left(U_{i,(t+1)m}^{(m)}\right)^3|J_{tm}^{(m)}\right) = s_i^0\left(\sum_{l=1}^m\alpha_l^3\right).
$$

From Assumptions 1 and 2 and the specifications in (2),

$$
\begin{aligned}
E\left(\left(Y_{i,(t+1)m}^{(m)}\right)^3 |J_{tm}\right) &= E\left(\left(\lambda_i f_{(t+1)m}^{(m)}\right)^3 + \left(U_{i,(t+1)m}^{(m)}\right)^3 |J_{tm}\right) \\
&= \lambda_i^3 E\left(\left(f_{(t+1)m}^{(m)}\right)^3 |J_{tm}\right) + E\left(\left(U_{i,(t+1)m}^{(m)}\right)^3 |J_{tm}\right)
\end{aligned}
$$

Thus, $E\left(\left(Y_{i,(t+1)m}^{(m)}\right)^3 |J_{tm}^{(m)}\right) = \lambda_i^3 h_1^{(m)} \sigma_{tm}^2 + \lambda_i^3 h_0 + s_i^0 \left(\sum_{l=1}^m \alpha_l^3\right)$ for all i and t □

References

Andersen, T. G. and Sørensen, B. E. (1996). gmm estimation of a stochastic volatility model: A monte carlo study. *Journal of Business and Economic Statistics*, 14:328–352.

Backus, D., Foresi, S., and Telmer, C. (2001). Affine models of currency pricing: Accounting for the forward premium anomaly. *Journal of Finance*, 56:279–304.

Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys*, 18:613–649.

Christoffersen, P., Heston, S., and Jacobs, K. (2006). Option valuation with conditional skewness. *Journal of Econometrics*, 131:253–284.

Cox, J. C., J. E. Ingersoll, J., and Ross, S. A. (1985). A theory of term structure of interest rates. *Econometrica*, 53-2:385–408.

Dai, Q. and Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55:1943–1978.

Fiorentini, G., Sentana, E., and Shephard, N. (2004). Likelihood-based estimation of generalised arch structures. *Econometrica*, 72:1481–1517.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6-2:327–343.

Heston, S. L. and Nandi, S. (2000). A closed-form garch option pricing model. *Review of Financial Studies*, 13-3:585–625.

Sentana, E. (1995). Quadratic garch models. *Review of Economic Studies*, 62:639–661.

Singleton, K. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102:111–141.