[1] Departamento de Producción Animal, Universidad de Buenos Aires, Avenida San Martín 4453, 1417 Buenos Aires, Argentina
[2] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

RODOLFO J. C. CANTET [1, 2,], MARÍA J. SUAREZ MUNILLA LEGUIZAMÓN [1], EDUARDO P. CAPPA [1], and ANA N. BIRCHMEIER [1]

Evidence for Autoregressive Conditional Heteroskedastic errors in growth traits of beef cattle[*]

Abstract

Evidence is presented for 'generalized autoregressive conditional heteroskedasticity' processes (GARCH(1,1)), in the residuals of beef cattle growth traits. This process can account for differences in variance at different time points, with the advantage of using a parsimonious parametrization. Data used were 10271 birth weights (BW), 19992 weaning weights (WW) and 9717 weight at 18 months (FW), from five herds registered in the national evaluation of the Brangus breed in Argentina. The residuals calculated from the 2005 genetic evaluation were regressed on Julian dates by least squares. From a second set of residuals out of the linear regression model, Maximum Likelihood estimation via the Fisher scoring algorithm was used to estimate the GARCH(1,1) parameters. Eight out of fifteen one-sided Lagrange multiplier statistics significantly ($P < 0.05$) rejected the hypothesis of null GARCH(1,1) parameters in the genetic evaluation residuals. Incorporating these effects in genetic evaluation is feasible due to the diagonal covariance matrix induced by the process on each trait, which simplifies building the mixed model equations.

Key Words: ARCH, GARCH, heteroskedasticity, growth, beef cattle

Zusammenfassung

Titel der Arbeit**: Beleg für autoregressiv bedingte Streuungsungleichheit bei Wachstumsmerkmalen von Fleischrindern**
Es werden Beweise für autoregressiv bedingte Streuungsungleichheit (GARCH (1,1)) in den Residuen bei Wachstumsmerkmalen von Fleischrindern vorgestellt. Dieser Prozess kann für eine Vereinfachung bei der Bestimmung einer zeitabhängigen Variation der Varianzen genutzt werden.
Der Datensatz enthält 10271 Geburtsgewichte, 19992 Absetzgewichte und 9717 Gewichte der Tiere am 18. Lebensmonat, welche in fünf Herden, die in der Nationalen Erfassungsstelle der Brangus Rasse Argentiniens erfasst sind, erhoben wurden. Die Residuals wurden 2005 nach genetischer Herkunft erfasst und nach ihren Kalenderdaten mit der Methode der kleinsten Quadrate berechnet. Aus einer zweiten Residualgesamtheit des linearen Regressionsmodells wurden mittels der Maximum-Likelihood-Schätzung über den Fisher Scoring Algorithmus, die Parameter des GARCH (1,1) Modells geschätzt. Acht der fünfzehn einseitigen Lagrangischen Multiplikationstests lehnten signifikant (P <0,05) die Hypothese von Null-GARCH (1,1) Parametern in den Residuen der genetischen Auswertung ab. Diese Effekte können durch die diagonale Kovarianz-Matrix, die durch diesen Prozess in jedem Wachstumsmerkmal induziert wird, in die genetische Auswertung einbezogen werden. Dadurch kann die Erstellung gemischter Modelle vereinfacht werden.

Schlüsselwörter: ARCH, GARCH, Varianzheterogenität, Wachstumsmerkmale, Fleischrind

Introduction

The presence of heterogeneous error variances in field data used for genetic evaluation has been extensively documented in different species and different traits. Fail to account for heterogeneous variance in models of genetic evaluation reduces intensity and accuracy of selection (HILL, 1984). The reason is that animals displaying higher

performances in highly variable herds tend to be selected over individuals that are high performers in herds with low variability (HILL, 1984; VAN VLECK, 1987). As a consequence, genetic progress will be reduced. On the other hand, if estimates of the heterogeneous variances are available, the mixed model equations for Best Linear Unbiased Prediction (HENDERSON, 1984) can be modified to account for heteroskedasticity into the model (GIANOLA, 1986). However, several covariance structures that account for heterogeneity of variances and covariances are either not amenable to computation (i.e. non-diagonal R covariance matrices of error effects), or are difficult to estimate due to the high number of parameters involved. J. L. Foulley and coworkers (FOULLEY et al, 1990, 1992; SAN CRISTOBAL et al., 1993) have developed models to detect causes of heteroskedasticity based on taking natural logs of residual variances and setting a linear model for the transformed dispersion parameters.

There may be instances in which the existing heterogeneity is time-dependent, and can not be ascribed to a single cause, such as contemporary group, age of dam, sex, etc. In these cases, a stochastic process may be envisaged at the herd level that makes the variance of observations from animals that are born and raised one or more days apart, as related to each other. The process is such that its unconditional variance does not display the heterogeneity, but the variance conditional on one or more previous errors does. Within the econometrics literature, ENGLE (1982) has described such a process, and called it 'autoregressive conditional heteroskedasticty'. The literature is abundant in related processes such as GARCH, IGARCH, ARCH-M, to quote some. ARCH-GARCH processes can take into account periods of low variance and periods where the variability observed is high. The objective of this research is to display evidence for the presence of GARCH processes in growth traits of beef cattle at the herd level.

Material and Methods

Data

Growth data used in the study consisted of 10271 birth weights (BW), 19992 weaning weights (WW) and 9717 weights at approximately 18 months of age (FW), of bulls and heifers registered in the animal genetic evaluation program (ERBra) of the Argentine Brangus Association. Out of the 56 herds that participated of ERBra 2005, five were chosen for the analysis. The criteria used in selecting the herds were 1) number of years of recording, 2) completeness of the data, and 3) regional representativity. The day in which the first recorded animal was born was arbitrarily set to day 1 for each herd, and the day of birth (t) of any animal born later was scored in Julian days.

Residuals for BW, WW and FW, were calculated from the multiple trait animal model evaluation used in the ERBra. Fixed effects included contemporary groups for all traits, sex of calf and age of dam for BW and WW, and age of weaning for WW and age at the day of measure for FW. In the ERBra, the parametrization of HILL (1982) is used to correct for heterosis in WW using mean dominance and mean additive maternal effects, whereas dominance is included in the model for FW. Random direct breeding values were included for the three traits, whereas WW also included maternal breeding values. Estimates of all fixed effects plus predicted breeding values from the 2005 evaluation were employed to obtain the estimated residuals.

Estimation and hypothesis testing
The characteristics and properties of ARCH and GARCH processes are described in Appendix A. The parameters of the GARCH(1,1) processes were estimated by maximum likelihood using the Fisher scoring algorithm. The estimating equations for the ARCH(1) process are described in Appendix B, whereas the derivatives needed for the GARCH(1,1) process are obtained in Appendix C.

In order to test for the presence of a GARCH (1,1) processes in the data, we used the Likelihood Multiplier (LM) statistics originally described by ENGLE (1982) to test for an ARCH(1) process. LEE (1991) showed that the LM test for an ARCH(1) is numerically identical to the LM test for the GARCH(1,1). Therefore, we used the LM statistics to test directly for the GARCH(1,1), where the null hypothesis is $\alpha_1 = 0$ and $\beta_1 = 0$. A further refinement consisted in using the more powerful one-sided version of the LM test discussed by DEMOS and SENTANA (1998), which will be explained below.

In the context of beef cattle growth traits, the LM procedure consists on fitting a linear regression model on the residuals from the genetic evaluation on julian dates, and then test for the presence of a GARCH process in the residuals from this regression. A problem with animal breeding data is that calvings occur irregularly through time, so that in any given date there may be several animals being born. As standard econometric techniques are based on fixed time interval analysis, ENGLE and RUSELL (1998) expanded ARCH-GARCH processes to account for variable time interval analysis into *autoregressive conditional duration* (ACD) models. However, ENGLE (2002) observed that standard ARCH-GARCH techniques can be used after transforming the ACD irregularly spaced random variable x_t (in our case Day of Birth, DOB, in Julian days) to $\sqrt{x_t}$, and then setting the mean to zero. Let $e_{(ij)t}$ be the residual of trait j from animal i with transformed DOB (denoted as t). The residual $e_{(ij)t}$ was regressed on the transformed DOB of individual i (t_i) such that:

$$e_{(ij)t} = \delta_0 + \delta_1 t_i + \varepsilon_{(ij)t} \tag{8}$$

where δ_0 is the intercept and δ_1 is the slope of the residuals through time, and $\varepsilon_{(ij)t}$ is the error term. It is expected that the 'regression residuals' from the genetic evaluation do not deviate from a zero expectation, i.e. $\delta_0 = 0$ and $\delta_1 = 0$. Moreover, if there is no evidence for an ARCH or GARCH process, then the term $\varepsilon_{(ij)t}$ is just 'white noise': normal independent random variables with zero mean and *constant* variance. Regression residuals were obtained with least squares estimates of the parameters in (8) as follows

$$\hat{\varepsilon}_{(ij)t} = e_{(ij)t} - \hat{\delta}_0 - \hat{\delta}_1 t_i \tag{9}$$

The LM test for a GARCH(1,1) process was implemented for each of the three traits and each of the five herds. The null hypothesis was H_0: $\alpha_1 = 0$ and $\beta_1 = 0$, and the test statistics is equal to

$$\frac{T f_0' F (F'F)^{-1} F' f_0}{f_0' f_0} = T R^2 \tag{10}$$

where $\boldsymbol{F} = [\boldsymbol{F}_1 : ... : \boldsymbol{F}_T], \boldsymbol{F}_i = \left[1, \hat{\varepsilon}^2_{(ij)t-1} \right]; \boldsymbol{f}_0$ is the column vector of $\left(\dfrac{\hat{\varepsilon}^2_{(ij)t}}{\sigma^2_{t0}} - 1 \right)$ under the

null hypothesis, and $\sigma^2_{t0} = \dfrac{\sum\limits_{i=1}^{T} \hat{\varepsilon}^2_{(ij)t}}{T}$. The R^2 is the squared multiple correlation between

$\boldsymbol{f}_0$ and $\boldsymbol{F}$. Under the null hypothesis of the one-sided test proposed by DEMOS and SENTANA (1998), the test statistics is asymptotically distributed as a 50:50 mixture of $\chi^2_{(0)}$ and $\chi^2_{(1)}$, which was the distribution used to calculate the p-values.

Results

Tables 1, 2, 3 display the LM tests for the hypothesis of the GARCH(1,1) process (H_0: $\alpha_1 = 0$ and $\beta_1 = 0$) on the errors of BW, WW, and FW, respectively.

Table 1
Results of the Lagrange multiplier tests for birth weight (Lagrangischer multipler Test für Geburtsgewichte)

Herd	Number of records	Number of Julian days	σ^2	LM statistics	p-value	Kurtosis
A	361	274	0.035	0.00	0.469	225.26
B	7,760	8,736	0.052	2.70	0.050	5.37
C	739	1,011	0.134	3.36	0.033	3.06
D	634	2,717	0.068	0.15	0.349	3.32
E	777	857	0.136	11.38	< 0.001	3.00

Table 2
Results of the Lagrange multiplier tests for weaning weight (Lagrangischer multipler Test für Absetzgewichte)

Herd	Number of records	Number of Julian days	σ^2	LM statistics	p-value	Kurtosis
A	5,703	11,061	1.189	187.64	< 0.001	4.05
B	11,350	9,036	1.018	3.27	0.035	2.49
C	758	1,011	1.977	0.01	0.454	2.22
D	1,275	2,793	1.735	15.31	< 0.001	2.01
E	836	857	1.968	0.11	0.371	1.92

Table 3
Results of the Lagrange multiplier tests for 18-months weight (Lagrangischer multipler Test für das 18 Monate-Gewicht)

Herd	Number of records	Number of Julian days	σ^2	LM statistics	p-value	Kurtosis
A	2,932	11,058	2.890	1.49	0.111	1.92
B	5,976	8,735	2.379	49.57	< 0.001	4.72
C	256	853	2.249	13.18	< 0.001	0.63
D	348	2,717	4.560	0.34	0.280	0.65
E	205	143	4.525	0.63	0.213	2.13

There were significant ($P < 0.05$) evidences that $\alpha_1 > 0$ and $\beta_1 > 0$ for all three traits in herd B, and at least one trait showed a value of α_1 and β_1 significantly ($P < 0.05$) greater than 0 in the other herds: WW for herd A, BW and FW for herd C, WW for herd D, and BW for herd E. Length of data collection (number of Julian dates) was not related to the result of the tests. Most values of the kurtosis for the unconditional distribution of error terms were greater than 3 whenever the LM test was significant, as expected by theory. The Figure depicts the distribution of errors for WW with time

in herd A. Observe that the period from 1 to 2000 days displays larger changes in variance than the interval between 2000 and 5000 days, although the mean does not deviate from 0 throughout the 30 years of data collection. Notice the typical GARCH effect by which large (small) absolute values are followed by large (small) values of unpredictable sign.
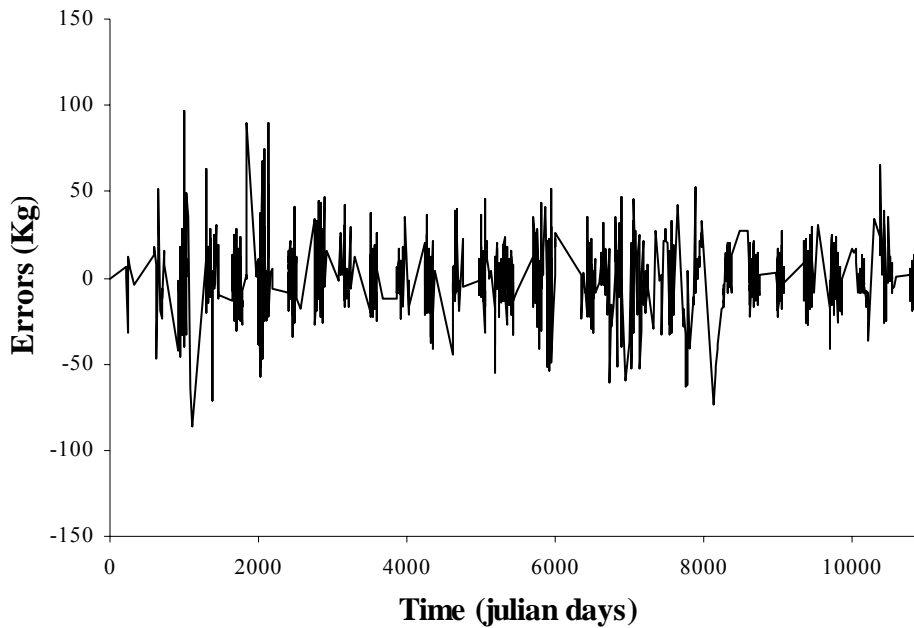

Figure: Distribution of errors for weaning weight with time for herd A. (Fehlerverteilung für die Absetzgewichte abhängig von den Tagen der Herde A)

Estimates of the parameters are showed in Tables 4, 5 and 6. Across all herds and traits that displayed significant LM statistics in favor of a GARCH(1,1) process, parameter estimates ranged from 0.009 to 0.182 for α_1, and 0.079 to 0.988 for β_1.

Table 4
Estimates of the GARCH(1,1) parameters for birth weight (Schätzung der GARCH (1,1) Parameter für Geburtsgewichte)

Herd	ω		α_1		β_1	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
A	0.0351	0.00051	0.0000	<0.00001	0.0028	0.00001
B	0.0001	0.00003	0.0087	0.00075	0.9882	0.00118
C	0.1002	0.01750	0.1719	0.04690	0.0788	0.13920
D	0.0501	0.03100	0.0875	0.03830	0.1814	0.44760
E	0.0016	0.00080	0.0618	0.01150	0.9273	0.01090

Table 5
Estimates of the GARCH(1,1) parameters for weaning weight (Schätzung der GARCH (1,1) Parameter für Absetzgewichte)

Herd	ω		α1		β1	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
A	0.0172	0.00211	0.0975	0.005694	0.8948	0.00523
B	0.0375	0.00740	0.0133	0.001824	0.9498	0.00840
C	1.9562	0.09810	0.0106	0.0318	0.0000	0.00000
D	0.2002	0.05440	0.1818	0.0265	0.7088	0.04550
E	1.9682	0.07180	0.0000	$< 10\text{-}10$	0.0000	0.00000

Table 6
Estimates of the GARCH(1,1) parameters for 18-month weight (Schätzung der GARCH (1,1) Parameter für 18 Monate-Gewichte)

Herd	ω		α1		β1	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
A	0.4795	0.21140	0.0399	0.0137	0.7945	0.0832
B	0.0245	0.00361	0.0187	0.0015	0.9709	0.0025
C	0.5855	0.49490	0.1004	0.0566	0.6383	0.2465
D	4.4625	0.32980	0.0213	0.0423	0.0000	$< 10\text{-}10$
E	4.5255	0.01600	0.0000	$< 10\text{-}10$	0.0000	0.0726

Discussion

In the current research, we found evidence that the environmental variances of growth traits in beef cattle can be modeled with GARCH(1,1) processes. Heteroskedasticity in growth traits have also been reported by ROBERT-GRANIE et al (2002); RODRIGUEZ-ALMEIDA et al (1995) and SAN CRISTOBAL et al (1993). HILL (1984) observed that animals from the more variable environments tend to be selected if heterogeneous environmental variances are not accounted for. He further indicated that heterogeneity of variances induces nonlinearity between true and predicted breeding value, a fact observed on simulated data by MEUWISSEN and VAN DER WERF (1993). Thus, accuracy of prediction of breeding value will decrease due to unaccounted variance heterogeneity.

There are two major advantages to the use of ARCH-GARCH processes for modeling heteroskedasticity. First, the parameters for the ARCH-GARCH process describe variance heterogeneity at the individual animal level, rather than to using groups of animals to model heteroskedasticity (say, by herd, by year, or by contemporary groups), being therefore more sensitive to differences in environmental variance. Second, the number of parameters introduced (2 per herd for ARCH(1) and 3 per herd for GARCH(1,1)) will generally be smaller than using one variance component per contemporary group. As an example, in the Argentinian Brangus evaluation the amount of information for a single parameter in the GARCH(1,1) process, is respectively, 3.3, 13.2 o 9.7 times greater for BW, WW and FW, as compared to using one variance parameter per contemporary group. We will deal with estimating the dispersion parameters in multiple-trait animal models elsewhere.

If estimates are available of the ARCH or GARCH parameters, the processes can be incorporated into an animal model of genetic evaluation. Let R be the $s \times s$ environmental covariance matrix among s traits. Then, R can be decomposed as $R = S C S$, where S is a diagonal matrix with non-zero diagonal entries equal to the environmental standard deviations of the s traits $\sigma_1, \ldots, \sigma_s$, whereas C is a a correlation matrix with diagonal elements equal to 1 and off-diagonals equal to r_{ij}, the

environmental correlation between traits i and j. To account for an ARCH-GARCH process, let $\boldsymbol{R}_{hit}$ be the covariance matrix of environmental effects for animal i within herd h, born at time t. Now, $\boldsymbol{R}_{iht} = \boldsymbol{S}_{iht}\,\boldsymbol{C}\,\boldsymbol{S}_{iht}$, with the diagonals of $\boldsymbol{S}_{iht}$ equal to σ_{ijht}: the square root of the variance in (3) for the ARCH or in (7) for the GARCH, scaled to the regular order of magnitud of $\boldsymbol{R}_{ii}$.

To build up the mixed model equations (MME) for a multi trait animal model, order the data by date of birth within the herd. Then, from previous solutions or a previous run, obtain the residuals for each animal within her and trait, and write them into the input data file. Next, calculate the variances σ^2_{ijht} while reading the data and the *ijht*-error term. Finally, read the data and the σ^2_{ijht}'s for each animal, create and invert (or obtain a g-inverse) of $\boldsymbol{R}_{ijht}$ and form the MME as usual.

In conclusion, heteroskedasticity of growth taits in some beef herds can be modeled by GARCH(1,1) processes. Incorporating these effects in genetic evaluation is feasible due to the diagonal covariance matrix induced by the process on each trait, which simplifies building the MME.

Appendix A. The ARCH and GARCH processes

The ARCH (q) process

Consider the set of random variables indexed by time $\mathfrak{I}_T = \{e_1,.., e_{t-1}, e_t,.., e_T\}$, and write

$$e_t = \epsilon_t \sqrt{\sigma^2_t} \tag{A1}$$

where ϵ_t are standard i.i.d. normal $(0,1)$ variables, and σ^2_t is the variance of the process at time t. The unconditional expectation of (1) is

$$\mathrm{E}\left(e_t\right) = \mathrm{E}\left(\epsilon_t\right)\sqrt{\sigma^2_t} = 0\sqrt{\sigma^2_t} = 0 \tag{A2}$$

If we condition on the set $\mathfrak{I}_{t-q}$, the conditional expected value is

$$\mathrm{E}\left(e_t \mid \mathfrak{I}_{t-q}\right) = \mathrm{E}\left(\epsilon_t \mid \mathfrak{I}_{t-q}\right)\sqrt{\sigma^2_t} = \mathrm{E}\left(\epsilon_t\right)\sqrt{\sigma^2_t} = 0$$

As put forward by ENGLE (1982), the variance in (1) is modeled conditional on $\mathfrak{I}_{t-q}$

$$\mathrm{Var}\left(e_t \mid \mathfrak{I}_{t-q}\right) = \mathrm{Var}\left(\epsilon_t\right)\sigma^2_t = \sigma^2_t = \omega + \alpha_1\,e^2_{t-1} + \alpha_2\,e^2_{t-2} + ... + \alpha_q\,e^2_{t-q} \tag{A3}$$

with the conditions $\omega > 0$ and $0 \le \alpha_i \le 1$, $i = 1, 2, ..., q$. This equation for the variance allows capturing the dynamics of the conditional heteroskedasticity, as the conditional variance is dependent upon the realized values of $e_{t-q},..., e_{t-1}$ squared. In turn, this conveys the idea of the process having 'memory' on the past $1, 2.., q$ realizations. A further requirement is needed for the variance to be stationary, so that it does not depend on the particular time t (BOLLERSLEV et al, 1993, page 2967):

$$\left(\alpha_1 + \alpha_2 + ... + \alpha_q\right) < 1$$

Notice that the unconditional variance of (1) is constant though, and can be obtained as follows

$$\sigma^2 = \mathrm{Var}\left(e_t\right) = \mathrm{E}\left(e^2_t\right) = \mathrm{E}\left(\mathrm{E}\left(e^2_t \mid \mathfrak{I}_{t-q}\right)\right) = \mathrm{E}\left(\sigma^2_t\right)$$

$$= \omega + \alpha_1\,\mathrm{E}\left(e^2_{t-1}\right) + ... + \alpha_q\,\mathrm{E}\left(e^2_{t-q}\right) = \omega + \alpha_1\,\sigma^2 + ... + \alpha_q\,\sigma^2$$

So that

$$\sigma^2 = \omega + \alpha_1 \, \sigma^2 + \ldots + \alpha_q \, \sigma^2 \Rightarrow \sigma^2 \left(1 - \alpha_1 - \ldots - \alpha_q \right) = \omega$$

$$\Rightarrow \quad \sigma^2 = \mathrm{Var}\left(e_t \right) = \frac{\omega}{1 - \alpha_1 - \ldots - \alpha_q}$$

Also, the ARCH(q) variables are uncorrelated at different times:

$$\mathrm{Cov}\left(e_t, e_{t-q} \right) = \mathrm{E}\left(e_t \, e_{t-q} \right) = \mathrm{E}\left(\mathrm{E}\left(e_t \, e_{t-q} \mid \mathfrak{I}_{t-q} \right) \right) = $$
$$\mathrm{E}\left(e_{t-q} \, \mathrm{E}\left(e_t \mid \mathfrak{I}_{t-q} \right) \right) = \mathrm{E}\left(e_{t-q} (0) \right) = 0 \tag{A4}$$

Although the e's are serially uncorrelated, they are not independent through time. In words of BOLLERSLEV et al (1993) "there is a tendency for large (small) absolute values of the process to be followed by other large (small) values of unpredictable sign". After (A2), (A3) and (A4), the conditional distribution of the ARCH(q) random variable in (1) is equal to

$$e_t \mid \mathfrak{I}_{t-q} \sim N\left(0, \sigma_t^2 \right) \qquad p\left(e_t \mid \mathfrak{I}_{t-q} \right) = \frac{1}{\sqrt{2 \, \pi \, \sigma_t^2}} \, \exp\left[-\frac{e_t^2}{2 \, \sigma_t^2} \right]. \tag{A5}$$

Note that in ARCH(q) processes the conditional probability of $e_1, .., e_{q-1}$ is unknown, as $\mathfrak{I}_{q-1}$ is undefined. Although the conditional distribution of e_t is time-invariant, the unconditional distribution for e_t will tend to have 'fatter tails' than the normal density of ϵ_t. For example, BOLLERSLEV et al (1993) observed that the kurtosis of an ARCH(1) process with conditionally normally distributed errors is equal to:

$$\mathrm{Kurtosis}\left(e_t \right) = \frac{\mathrm{E}\left(e_t^4 \right)}{\left[\mathrm{E}\left(e_t^2 \right) \right]^2} = 3 \, \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} \geq 3 \tag{A6}$$

Therefore, the ARCH(1) process is leptokurtic: the curvature is high in the center of the density and the tails are 'fatter' than those of the normal distribution.

The GARCH (p, q) process
BOLLERSLEV (1986) generalized ARCH(q) processes by allowing the conditional variance to be also a function on the p previous variances $\sigma_{t-p}^2, \sigma_{t-p+1}^2, \ldots, \sigma_{t-1}^2$, besides the q previous realizations of $e_{t-q}, \ldots, e_{t-1}$. Thus, the variance in the GARCH(p, q) process is

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \, e_{t-i}^2 + \sum_{i=1}^{p} \beta_i \, \sigma_{t-i}^2 \tag{A7}$$

The requirement for the variance to be stationary (HAMILTON, 1994, page 666) is as follows

$$\left(\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i \right) < 1$$

For the GARCH(1,1) process, BOLLERSLEV et al (1993) observed that the conditions for positivity of the conditional variance (i.e. $\sigma_t^2 > 0$) are $\omega \geq 0$, $\alpha_1 \geq 0$, and $\beta_1 \geq 0$.

Appendix B: Maximum likelihood estimation

By conditioning on ε_1 from the first observation, at the start of the recursive process the likelihood does not depend on the parameters (BOLLERSLEV et al, 1993). From (A4) and (A5), the log-likelihood function can be written as

$$\mathcal{L}\left(\varepsilon_t, \sigma_t^2\right) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{T}\left(\log\left(\sigma_t^2\right) + \frac{\varepsilon_t^2}{\sigma_t^2}\right) \tag{B1}$$

By taking first derivatives of (B1) the scores are:

$$\frac{\partial \mathcal{L}}{\partial \omega} = \left(\frac{\partial \mathcal{L}}{\partial \sigma_t^2}\right)\left(\frac{\partial \sigma_t^2}{\partial \omega}\right) = \sum_{t=2}^{T}\left(\frac{1}{2\,\sigma_t^2}\left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\right) \qquad \frac{\partial \mathcal{L}}{\partial \alpha_1} = \left(\frac{\partial \mathcal{L}}{\partial \sigma_t^2}\right)\left(\frac{\partial \sigma_t^2}{\partial \alpha_1}\right) = \sum_{t=2}^{T}\left(\frac{\varepsilon_{t-1}^2}{2\,\sigma_t^2}\left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\right) \tag{B2}$$

Whereas second derivatives are equal to:

$$\frac{\partial^2 \mathcal{L}}{\left(\partial\omega\right)^2} = \sum_{t=2}^{T}\left[-\frac{1}{2\,\sigma_t^4}\left(2\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\right] \qquad \frac{\partial^2 \mathcal{L}}{\left(\partial\alpha_1\right)^2} = \sum_{t=2}^{T}\left[-\frac{\varepsilon_{t-1}^4}{2\,\sigma_t^4}\left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\right]$$

$$\frac{\partial^2 \mathcal{L}}{\partial\omega\,\partial\alpha_1} = \sum_{t=2}^{T}\left[-\frac{\varepsilon_{t-1}^2}{2\,\sigma_t^4}\left(2\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\right] \tag{B3}$$

The expected values of the second derivatives are needed for the Fisher scoring algorithm. Now, observe that after (1) we can write $\epsilon_t^2 = \varepsilon_t^2\sigma_t^{-2}$. As the ϵ_t are $N(0,1)$, we have that $E\left(\epsilon_t^2\right) = 1$. Consequently, $E\left(\frac{\varepsilon_t^2}{\sigma_t^2}\right) = 1$ and $E\left(2\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right) = 1$, so that the expectations of second derivatives are

$$E\left[\frac{\partial^2 \mathcal{L}}{\left(\partial\omega\right)^2}\right] = -\frac{1}{2}\sum_{t=2}^{T}E\left[\frac{1}{\sigma_t^4}\right] \qquad E\left[\frac{\partial^2 \mathcal{L}}{\left(\partial\alpha_1\right)^2}\right] = -\frac{1}{2}E\left[\frac{\varepsilon_{t-1}^4}{\sigma_t^4}\right] \qquad E\left[\frac{\partial^2 \mathcal{L}}{\partial\omega\,\partial\alpha_1}\right] = -\frac{1}{2}E\left[\frac{\varepsilon_{t-1}^2}{\sigma_t^4}\right]$$

ENGLE (1982) observed that these expectations are consistently estimated by

$$\hat{E}\left[\frac{\partial^2 \mathcal{L}}{\left(\partial\omega\right)^2}\right] = -\frac{1}{2(T-1)}\sum_{t=2}^{T}\frac{1}{\hat\sigma_t^4} \qquad \hat{E}\left[\frac{\partial^2 \mathcal{L}}{\left(\partial\alpha_1\right)^2}\right] = -\frac{1}{2(T-1)}\sum_{t=2}^{T}\frac{\hat\varepsilon_{t-1}^4}{\hat\sigma_t^4}$$

$$\hat{E}\left[\frac{\partial^2 \mathcal{L}}{\partial\omega\,\partial\alpha_1}\right] = -\frac{1}{2(T-1)}\sum_{t=2}^{T}\frac{\hat\varepsilon_{t-1}^2}{\hat\sigma_t^4} \tag{B5}$$

with $\hat\sigma_t^2 = \dfrac{\sum_{t=2}^{T}\left(\hat\omega + \hat\alpha_1\,\hat\varepsilon_t\right)^2}{(T-1)}$. The scalars $\hat\omega$ and $\hat\alpha_1$ are the estimated dispersion parameters, and $\hat\varepsilon_t$ is the residual from the previous iteration of the Fisher scoring algorithm. The estimating equations are:

$$\begin{bmatrix} \omega^{[k]} \\ \alpha_1^{[k]} \end{bmatrix} = \begin{bmatrix} \omega^{[k-1]} \\ \alpha_1^{[k-1]} \end{bmatrix} + \begin{bmatrix} \hat{E}\left[\frac{\partial^2 \mathcal{L}}{\left(\partial\omega\right)^2}\right] & \hat{E}\left[\frac{\partial^2 \mathcal{L}}{\partial\omega\,\partial\alpha_1}\right] \\ \hat{E}\left[\frac{\partial^2 \mathcal{L}}{\partial\omega\,\partial\alpha_1}\right] & \hat{E}\left[\frac{\partial^2 \mathcal{L}}{\left(\partial\alpha_1\right)^2}\right] \end{bmatrix}^{-1} \frac{1}{T}\begin{bmatrix} \frac{\partial \mathcal{L}}{\partial\omega} \\ \frac{\partial \mathcal{L}}{\partial\alpha_1} \end{bmatrix} \tag{B6}$$

The superindex k indicates the iteration number. Convergence was assumed to be attained when the difference between the estimated parameter vector in (B6) from two

successive iterations was lower than a preset value (for example, 10^{-6}) for both ω and α_1. The equivalent system to (B6) for the GARCH(1,1) is 3×3 due to the added equation for β_1.

Appendix C: Analytic derivatives of the GARCH(1,1) log-likelihood

By combining expressions (3) and (4) in FIORENTINI et al (1996), the log-likelihood can be written as

$$\mathcal{L}\left(\varepsilon_t, \sigma_t^2\right) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log\left(\sigma_t^2\right) + \frac{\varepsilon_t^2}{\sigma_t^2}\right) \tag{C1}$$

To take care of the first observation, Fiorentini *et al* (1996) employed an estimate of the unconditional expectation, so when $t \le 0$

$$\sigma_t^2 = \varepsilon_t^2 = \frac{1}{T}\sum_{s=1}^{T}\varepsilon_s^2$$

Now, write $\sigma_t^2 = z_t'\theta$, with the parameter vector being equal to $\theta = \begin{bmatrix} \omega & \alpha_1 & \beta_1 \end{bmatrix}'$, and $z_t' = \begin{bmatrix} 1 & \varepsilon_{t-1}^2 & \sigma_{t-1}^2 \end{bmatrix}$.

On taking derivatives in (C1) produces

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = -\frac{1}{2\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \theta_i} + \frac{1}{2}\frac{\partial \sigma_t^2}{\partial \theta_i}\frac{\varepsilon_t^2}{\sigma_t^2} = \frac{1}{2\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \theta_i}\left[\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right] \tag{C2}$$

Now, we need an expression for the derivatives of σ_t^2 with respect to the elements of θ:

$$\frac{\partial \sigma_t^2}{\partial \theta_i} = \frac{\partial\left(z_t'\theta\right)}{\partial \theta_i} = z_{t(i)} + \frac{\partial \sigma_{t-1}^2}{\partial \theta_i}\beta_1 \tag{C3}$$

Expression (C3) is recursive as the derivative at time t depends on to the derivatives of the variance at previous times. Thus, for $t = 1, 2, 3$, we have that

$$\frac{\partial \sigma_1^2}{\partial \theta_i} = z_{t(i)} + \beta_1\frac{\sigma_0^2}{\partial \theta_i} = z_{t(i)}$$

$$\frac{\partial \sigma_2^2}{\partial \theta_i} = z_{t(i)} + \beta_1\frac{\sigma_1^2}{\partial \theta_i} = z_{t(i)} + \beta_1 z_{t(i)}$$

$$\frac{\partial \sigma_3^2}{\partial \theta_i} = z_{t(i)} + \beta_1\frac{\sigma_2^2}{\partial \theta_i} = z_{t(i)} + \beta_1\left(z_{t(i)} + \beta_1 z_{t(i)}\right) = z_{t(i)} + \beta_1 z_{t(i)} + \beta_1^2 z_{t(i)} = z_{t(i)}\left(1 + \beta_1 + \beta_1^2\right)$$

The emerging pattern is such that

$$\frac{\partial \sigma_t^2}{\partial \theta_i} = z_{t(i)}\left(1 + \beta_1 + \beta_1^2 + \ldots + \beta_1^{t-1}\right) \tag{C4}$$

Let $h_t\left(\beta_1\right) = 1 + \beta_1 + \beta_1^2 + \ldots + \beta_1^{t-1}$. Using (C4) we have

$$\frac{\partial \sigma_t^2}{\partial \omega} = h\left(\beta_1\right) \qquad\qquad \frac{\partial \sigma_t^2}{\partial \beta_1} = \varepsilon_{t-1}^2\, h\left(\beta_1\right)$$

$$\tag{C5}$$

$$\frac{\partial \sigma_t^2}{\partial \beta_1} = \sigma_{t-1}^2 + \beta_1 \sigma_{t-2}^2 + \beta_1^2 \sigma_{t-3}^2 + \ldots + \beta_1^{t-1}\sigma_0^2 = \sum_{i=0}^{t-1}\beta_1^i \sigma_{t-i-1}^2$$

On replacing with (C5) in (C2), first derivatives of σ_t^2 with respect to $\mathcal{L}$ are respectively equal to

$$\frac{\partial \mathcal{L}}{\partial \omega} = \frac{1}{2\sigma_t^2} h_t(\beta_1) \left[\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right] \tag{C6}$$

$$\frac{\partial \mathcal{L}}{\partial \alpha_1} = \frac{\varepsilon_{t-1}^2}{2\sigma_t^2} h_t(\beta_1) \left[\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right] \tag{C7}$$

$$\frac{\partial \mathcal{L}}{\partial \beta_1} = \frac{1}{2\sigma_t^2} \left[\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right] \left(\sum_{i=0}^{t-1} \beta_1^i \, \sigma_{t-i-1}^2 \right) \tag{C8}$$

In order to obtain the second derivatives, we take derivatives in (C1) to get

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_i \partial \theta_{i'}} = \frac{\partial}{\partial \theta_{i'}} \left\{ \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_i} \left[\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right] \right\} =$$

$$\left[\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right] \left(\frac{1}{2\sigma_t^2} \left(\frac{\partial^2 \sigma_t^2}{\partial \theta_i \partial \theta_{i'}} \right) - \frac{1}{2\sigma_t^4} \left(\frac{\partial \sigma_t^2}{\partial \theta_i} \right) \left(\frac{\partial \sigma_t^2}{\partial \theta_{i'}} \right) \right) - \frac{1}{2\sigma_t^4} \left(\frac{\partial \sigma_t^2}{\partial \theta_i} \right) \left(\frac{\partial \sigma_t^2}{\partial \theta_{i'}} \right) \left(\frac{\varepsilon_t^2}{\sigma_t^2} \right) \tag{C9}$$

And, after some algebra second derivatives are equal to:

$$\frac{\partial^2 \mathcal{L}}{(\partial \omega)^2} = -\frac{h_t^2(\beta_1)}{2\sigma_t^4} \left(\frac{2\varepsilon_t^2}{\sigma_t^2} - 1 \right) \tag{C10}$$

$$\frac{\partial^2 \mathcal{L}}{(\partial \alpha_1)^2} = -\frac{\varepsilon_{t-1}^4 \, h_t^2(\beta_1)}{2\sigma_t^4} \left(\frac{2\varepsilon_t^2}{\sigma_t^2} - 1 \right) \tag{C11}$$

$$\frac{\partial^2 \mathcal{L}}{(\partial \beta_1)^2} = \frac{\left(i\beta_1^{i-1} \sigma_{t-i-1}^2 + \beta_1^i \sigma_{t-i-2}^2 \, h_{t-i}(\beta_1) \right)}{2\sigma_t^2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) - \frac{\left(\sum_{i=0}^{t-1} \beta_1^i \sigma_{t-i-1}^2 \right)^2}{2\sigma_t^4} \left(\frac{2\varepsilon_t^2}{\sigma_t^2} - 1 \right) \tag{C12}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \omega \, \partial \alpha_1} = -\frac{\varepsilon_{t-1}^2 \, h_t(\beta_1)}{2\sigma_t^4} \left(\frac{2\varepsilon_t^2}{\sigma_t^2} - 1 \right) \tag{C13}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \omega \, \partial \beta_1} = \frac{h_t'(\beta_1)}{2\sigma_t^2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) - \frac{h_t(\beta_1) \left(\sum_{i=0}^{t-1} \beta_1^i \sigma_{t-i-1}^2 \right)}{2\sigma_t^4} \left(\frac{2\varepsilon_t^2}{\sigma_t^2} - 1 \right) \tag{C14}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha_1 \, \partial \beta_1} = \frac{\varepsilon_{t-1}^2 \, h_t'(\beta_1)}{2\sigma_t^2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) - \frac{\varepsilon_{t-1}^2 \, h_t(\beta_1) \left(\sum_{i=0}^{t-1} \beta_1^i \sigma_{t-i-1}^2 \right)}{2\sigma_t^4} \left(\frac{2\varepsilon_t^2}{\sigma_t^2} - 1 \right) \tag{C15}$$

References

BOLLERSLEV, T.:
Generalized autoregressive conditional hetoskedaticity. Journal of Econometrics. **31** (1986), 307-327

BOLLERSLEV, T.; ENGLE, R. F.; NELSON, D. B.:
ARCH Models. In: *Handbook of Exonometrics* vol. 4. ed. Engle, R.F., and McFadden, D.L., Elsevier, Amsterdam, (1993) 2959-3040.

DEMOS A.; SENTANA, E.:
Testing for GARCH effects: A one-sided approach. Journal of Econometrics. **86** (1993), 97-127

ENGLE, R.F.:
Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica. **50** (1982), 987-1007

ENGLE, R.F.:
>   New frontiers for ARCH models. Journal of Applied Econometrics. **17** (2002) 425-446
ENGLE, R. F.; RUSSELL, J. R.:
>   Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica. **66** (1998), 1127-1162
FIORENTINI, G; CALZOLARI, G.; PANATTONI, L.:
>   Analytic Derivatives and the computation of GARCH estimates. Journal of Applied Econometrics. **11** (1996), 399-417
FOULLEY, J. L.; GIANOLA, D.; SAN CRISTOBAL, M.; IM, S.:
>   A method for assessing extent and sources of heterogeneity of residual variances in mixed linear models. Journal of Dairy Science. **73** (1990), 1612-1624
FOULLEY, J. L.; SAN CRISTOBAL, M. ; GIANOLA, D.; IM, S.:
>   Marginal likelihood and Bayesian approaches to the analysis of heterogeneous residual variances in mixed linear Gaussian models. Computational Statistics and Data Analysis. **13** (1992), 291-305
GIANOLA, D.:
>   On selection criteria and estimation of parameters when the variance is heterogeneous. Theoretical and Applied Genetics. **5** (1986) 671-677
HAMILTON, J. D.:
>   Time series analysis. Princeton University Press. USA. (1994)
HENDERSON, C. R.:
>   Applications of linear models in animal breeding. University of Guelph. Canada. (1984)
HILL, W. G.:
>   Dominance and epistasis as components of heterosis. Zeitschrift für Tierzüchtung und Züchtungsbiologie **99** (1982), 161-167
HILL, W.G.:
>   On selection among groups with heterogeneous variance. Animal Production. **39** (1984), 473-477
LEE, J. H. H.:
>   A Lagrange multiplier test for GARCH models. Economics Letters. **37** (1991), 265-271.
MEUWISSEN, T. H. E.; VAN DER WERF, J. H. J.:
>   Impact of heterogeneous within herd variances on dairy cattle breeding schemes: a simulation study. Livestock Production Science. **33** (1993), 31-41
ROBERT-GRANIE, C.; HEUDE, B.; FOULLEY, J. L. :
>   Modeling the growth curve of Maine-Anjou beef cattle using heteroskedastic random coefficients models. Genetics, Selection, Evolution **43** (2002), 423-445
RODRIGUEZ-ALMEIDA, F. A.; VAN VLECK, L. D.; CUNDIFF, L. V.; KACHMAN., S. D.:
>   Heterogeneity of variance by sire breed, sex, and dam breed in 200-day and 365-day weights of beef-cattle from a top cross experiment. Journal of Animal Science **73** (1995), 2579-2588
SAN CRISTOBAL, M.; FOULLEY, J. L; MANFREDI, E.:
>   Inference about multiplicative heteroskedastic components of variance in a mixed linear gaussian model with an application to beef-cattle breeding. Genetics Selection Evolution **25** (1993), 3-30
VAN VLECK, L. D. 1987.:
>   Selection when traits have different genetic and phenotypic variances in different environments. Journal of Dairy Science **70** (1987), 337-344

Corresponding Author
RODOLFO J. C. CANTET, Departamento de Producción Animal, Universidad de Buenos Aires, Avenida San Martín 4453,
1417 BUENOS AIRES, ARGENTINA

E-Mail: rcantet@agro.uba.ar)