Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Regular Variation of Popular GARCH Processes Allowing for Distributional Asymmetry

Todd Prono

2017-095

Please cite this paper as:
Prono, Todd (2017). "Regular Variation of Popular GARCH Processes Allowing for Distributional Asymmetry," Finance and Economics Discussion Series 2017-095. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.095.

Regular Variation of Popular GARCH Processes Allowing for Distributional Asymmetry[1]

Todd Prono[2]

First Version: August 2017

Abstract

Linear GARCH$(1,1)$ and threshold GARCH$(1,1)$ processes are established as regularly varying, meaning their heavy tails are Pareto like, under conditions that allow the innovations from the, respective, processes to be skewed. Skewness is considered a stylized fact for many financial returns assumed to follow GARCH-type processes. The result in this note aids in establishing the asymptotic properties of certain GARCH estimators proposed in the literature.

Keywords: GARCH, threshold GARCH, heavy tail, Pareto tail, regular variation. JEL codes:C20, C22, C53, C58.

[1] Gratitude is owed to Dong Hwan Oh for providing detailed comments. The views expressed in this paper are those of the author and do not necessarily reflect those of the Federal Reserve Board.

[2] Federal Reserve Board. (202) 973-6955, todd.a.prono@frb.gov.

1. Introduction

Generalized autoregressive conditional heteroskedastic (GARCH) models are a workhorse for conditional variance forecasting in financial economics. The linear GARCH(1,1) model of Bollerslev (1986) is a popular choice amongst practitioners, owing, in part, to its (relative) simplicity, but also to its strong forecasting performance, generally, and superior performance, specifically, on foreign exchange rate returns against more complicated alternatives (see; e.g., Hansen and Lunde, 2005). It is widely recognized that the conditional variance of equity returns tends to be asymmetric.[3] This feature, sometimes referred to as a "leverage effect," is captured by threshold GARCH models, Glosten, Jagannathan, and Runkle (1993), hereafter GJR GARCH, being one-such example. In out-of-sample forecast evaluations using equity returns, GJR GARCH$(1,1)$ is shown to improve upon the linear GARCH$(1,1)$ specification (Hansen and Lunde, 2005). As a consequence, the linear GARCH$(1,1)$ and GJR GARCH$(1,1)$ models represent (very) popular choices among academics and practitioners alike for characterizing the conditional variance of financial returns.

Linear GARCH processes are shown to be regularly varying (see Basrak, Davis, and Mikosch, 2002), meaning their tails are heavy and Pareto like. Mikosch and Stărică (2000) study the linear GARCH$(1,1)$ case in detail and demonstrate it to be regularly varying under the condition that the model's innovations follow a symmetric distribution.[4] They do not consider the GJR GARCH$(1,1)$ specification. Table 1 summarizes the skewness statistics on various (very) high frequency equity and foreign exchange rate returns. Evident from the table is that these statistics tend to be large in absolute terms and (highly) statistically significant, a tendency sufficiently prevalent to render skewness a stylized fact for many financial returns. Under either the linear GARCH$(1,1)$ or GJR GARCH$(1,1)$ model, skewness in returns necessarily sources to the given model's innovations. Skewness in these innovations conflicts with the aforementioned demonstration that a linear GARCH$(1,1)$ process is regularly varying. Moreover (to the best of my knowledge), such a demonstration (regardless of the treatment of the model's innovations) is not extended to the GJR GARCH$(1,1)$ case. As a consequence, this note establishes linear GARCH$(1,1)$ and GJR GARCH$(1,1)$ processes as regularly varying, where this result does not depend on the given model's innovations being symmetrically distributed. Besides being interesting in its own right, this result also aids in establishing the large-sample properties of the linear GARCH$(1,1)$ estimators discussed

[3] That is, tomorrow's variance tends to be higher (all else equal) if today's return is negative.

[4] Davis and Mikosch (1998) conduct an equally-detailed study of the linear ARCH(1) case, demonstrating it to be regularly varying under the same condition.

in Mikosch and Straumann (2002), Kristensen and Linton (2006), and Vaynmann and Beare (2014).

2. Regular Variation

Consider the GARCH model of

$$Y_t = \sigma_t \epsilon_t, \qquad \epsilon_t \sim i.i.d.\ D\left(0,\ 1\right), \tag{1}$$

where

$$\sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 \times I_{\left\{Y_{t-1} \geq 0\right\}} + \alpha_2 Y_{t-1}^2 \times I_{\left\{Y_{t-1} < 0\right\}} + \beta \sigma_{t-1}^2. \tag{2}$$

Given (1), the general model under consideration follows the Drost and Nijman (1993) definition of a strong GARCH process. Given (2), if $\alpha_1 \neq \alpha_2$, the specific model is GJR GARCH$(1,1)$. Under the special case where $\alpha_1 = \alpha_2$, the specific model is linear GARCH$(1,1)$. Recasting (2) as

$$\begin{aligned} \sigma_t^2 &= \omega + \sigma_{t-1}^2 \left(\gamma_{t-1} \epsilon_{t-1}^2 + \beta\right), \qquad \gamma_{t-1} = \alpha_1 \times I_{\left\{Y_{t-1} \geq 0\right\}} + \alpha_2 \times I_{\left\{Y_{t-1} < 0\right\}}, \\ &= \omega + \sigma_{t-1}^2 A_t \end{aligned}$$

represents the GARCH process as a stochastic recurrence equation (SRE), which is important for establishing $\left\{\left(Y_t,\ \sigma_t\right)\right\}_{t \in \mathbb{Z}}$ as regularly varying.

For a fixed and non-negative h, let

$$\mathbf{Y} = \mathbf{Y}_t = \Big(\ \left(Y_t, \sigma_t\right),\ \ \ldots,\ \ \left(Y_{t+h}, \sigma_{t+h}\right)\ \Big).$$

This section demonstrates that $\mathbf{Y}$ is regularly varying with (tail) index κ, or, using shorthand notation, $\mathbf{Y}$ is RV(κ). That is, there exists a sequence of constants $\{a_n\}$ such that

$$nP\left(\left|\mathbf{Y}\right| > a_n\right) \longrightarrow 1, \qquad n \to \infty,$$

where $|\cdot|$ denotes the max norm, $a_n = n^{1/\kappa} L\left(n\right)$, and $L\left(\cdot\right)$ is slowly-varying at ∞.

ASSUMPTION A1: *The distribution D has an unbounded support, and $E\left|\epsilon_t\right|^{i+\delta} < \infty$ for $i \geq 2$ and some $\delta > 0$.*

ASSUMPTION A2: $\omega \geq \underline{\omega} > 0$, $\alpha_j > 0$ for $j = 1, 2$, and $\beta > 0$.

ASSUMPTION A3: $E\left(A_t^l\right) < 1$ *for* $l \in \left[1,\ \frac{i}{2}\right]$, where i is defined in A1.

The moment existence condition in A1 is (fairly) standard (see; e.g., Lee and Hansen, 1994, Berkes, Horváth, and Kokoszka, 2003, and Berkes and Horváth, 2004). The novelty of A2 is the strictly positive, lower-bound for ω (see Kristensen and Rahbek, 2005). Establishing regular variation for $\{Y_t\}$ relies on taking a first-order Taylor Expansion around this lower bound; see (7). Notice, as well, the strict positivity of all model parameters, thus excluding the ARCH(1) case. In order to establish $\{Y_t\}$ as regularly varying in the special case where $\beta = 0$, see Prono (2016). Under A3, (at least) $E\left(Y_t^2\right) < \infty$ (see; e.g., Loretan and Phillips, 1994, for empirical evidence supporting this condition for various stock and foreign exchange rate returns). A3 is sufficient for $\{(Y_t,\ \sigma_t)\}$ to be strictly stationary (see; e.g., Mikosch, 1999, Corollary 1.4.38 and Remark 1.4.39). $\{(Y_t,\ \sigma_t)\}$ is also strong mixing by Carrasco and Chen (2002, Corollary 6) when $\alpha_1 = \alpha_2$, and Carrasco and Chen (2002, Corollary 10), otherwise.

A1 and A3 together distinguish $\{\epsilon_t\}$ as being thinner tailed than $\{\sigma_t\}$. As a consequence, regular variation of $\{Y_t\}$ stems directly from $\{\sigma_t\}$, as is also the case in Davis and Mikosch (1998), Mikosch and Stărică (2000), and Basrak et al. (2002). The generality of A1 and A3 includes the baseline case of a covariance-stationary GARCH$(1,1)$ process, but also covers the higher-moment existence conditions necessary in Mikosch and Straumann (2002), Kristensen and Linton (2006), and Vaynaman and Beare (2014).

PROPOSITION. *For the GARCH model of (1) and (2), let Assumptions A1–A3 hold, and consider*

$$\mathbf{Y} = \Big(\ (Y_0, \sigma_0)\,,\quad \ldots,\quad (Y_h, \sigma_h)\ \Big)$$

for a fixed $h \geq 0$. *Then* $\mathbf{Y}$ *is RV*(κ).

REMARK. *In the proof that follows, C denotes a generic constant that can assume different values in different places.*

Proof. Since A_t is (strictly) positive $\forall\ t$,

$$P\left(\sigma > x\right) \sim c x^{-\kappa}, \quad x \to \infty, \tag{3}$$

where $c = c\left(\omega, \alpha_1, \alpha_2, \beta\right)$, the precise value of which is given in Goldie (1991), and $\kappa \in (2,\ \overline{\kappa}]$, where

$\overline{\kappa}$ is an upper bound, is the unique solution to

$$E\left(A\right)^{\kappa/2}=1,$$

by Kesten (1973, Theorem 4). Next, for

$$\theta=\left(\omega,\alpha_1,\alpha_2,\beta\right),\qquad \underline{\theta}=\left(\underline{\omega},\alpha_1,\alpha_2,\beta\right),$$

define

$$\sigma_t^2\equiv\sigma_t^2\left(\theta\right)=\omega+\gamma_{t-1}Y_{t-1}^2+\beta\sigma_{t-1}^2,$$

and $\underline{\sigma}_t^2\equiv\sigma_t^2\left(\underline{\theta}\right)$ analogously. Also define

$$\overline{\alpha}\equiv\max\left(\alpha_1,\ \alpha_2\right)\geq\gamma_{t-1}\quad\forall\ t. \tag{4}$$

Then

$$\frac{\partial\sigma_t}{\partial\omega}=\frac{\partial\sqrt{\sigma_t^2}}{\partial\omega}=\frac{1}{2}\times\sigma_t^{-1}\times\frac{\partial\sigma_t^2}{\partial\omega}\leq\frac{1}{2}\times\sigma_t^{-1}\times\frac{1}{1-\beta}, \tag{5}$$

where the inequality follows from Lumsdaine (1996, Lemma 1, A1.2). Also, using recursive substitution,

$$\sigma_t^2-\underline{\sigma}_t^2=\left(\omega-\underline{\omega}\right)\sum_{i=0}^{t-1}\beta^i+\left(\sigma_0^2-\underline{\sigma}_0^2\right)\beta^t\leq\frac{\omega-\underline{\omega}}{1-\beta}. \tag{6}$$

Consider a first-order Taylor Expansion of σ_t around $\underline{\omega}$ such that

$$\begin{aligned}
\sigma_t &= \underline{\sigma}_t+\frac{\partial\underline{\sigma}_t}{\partial\omega}\left(\omega-\underline{\omega}\right) \\
&\leq \frac{\gamma_{t-1}Y_{t-1}^2+\beta\underline{\sigma}_{t-1}^2}{\underline{\sigma}_t}+C\underline{\sigma}_t^{-1} \\
&\leq C\times\left(\frac{\gamma_{t-1}Y_{t-1}^2+\beta\underline{\sigma}_{t-1}^2}{\underline{\sigma}_t}\right) \\
&\leq C\times\left(\frac{\gamma_{t-1}\left(\underline{\sigma}_{t-1}^2+C\right)\epsilon_{t-1}^2+\beta\underline{\sigma}_{t-1}^2}{\underline{\sigma}_t}\right) \\
&\leq C\times\left(\frac{\underline{\sigma}_{t-1}^2\left(\overline{\alpha}\epsilon_{t-1}^2+\beta\right)}{\underline{\sigma}_t}+C\times\frac{\epsilon_{t-1}^2}{\underline{\sigma}_t}\right) \\
&\leq C\times\left(\underline{\sigma}_{t-1}\overline{A}_t+C\times\frac{\epsilon_{t-1}^2}{\underline{\sigma}_t}\right),
\end{aligned} \tag{7}$$

where the first inequality relies on (5) , the second on $\underline{\sigma}_t^{-1}$ being bounded and $\beta > 0$, the third on (6), the fourth on (4), and the fifth on $\underline{\sigma}_t > \beta^{1/2}\underline{\sigma}_{t-1}$. Consider next

$$\frac{\underline{\sigma}_{t-1}}{\sigma_0} \leq \frac{\sigma_{t-1}}{\sigma_0}.$$

For

$$\begin{aligned}
\frac{\underline{\sigma}_1}{\sigma_0} &\leq \frac{\sigma_1}{\sigma_0} \\
&\leq \frac{\left(\omega + \sigma_0^2 \overline{A}_1\right)^{1/2}}{\sigma_0} \\
&\leq \frac{\omega + \sigma_0 \overline{A}_1^{1/2}}{\sigma_0} \\
&\leq C \times \overline{A}_1^{1/2},
\end{aligned}$$

where the third inequality follows from the Triangle Inequality, and the fourth from σ_0^{-1} being bounded and $\beta > 0$. Parallel reasoning produces

$$\begin{aligned}
\frac{\underline{\sigma}_2}{\sigma_0} &\leq \frac{\sigma_2}{\sigma_0} \\
&\leq \frac{\left(\omega + \sigma_1^2 \overline{A}_2\right)^{1/2}}{\sigma_0} \\
&\leq C \times \left(\frac{\sigma_1}{\sigma_0}\right) \times \overline{A}_2^{1/2} \\
&\leq C \times \overline{A}_1^{1/2} \times \overline{A}_2^{1/2}.
\end{aligned}$$

Suppose then that

$$\frac{\underline{\sigma}_{t-2}}{\sigma_0} \leq \frac{\sigma_{t-2}}{\sigma_0} \leq C \times \prod_{i=1}^{t-2} \overline{A}_i^{1/2}. \tag{8}$$

From (8) follows that

$$\begin{aligned}
\frac{\underline{\sigma}_{t-1}}{\sigma_0} &\leq \frac{\sigma_{t-1}}{\sigma_0} \\
&\leq \frac{\omega^{1/2}}{\sigma_0} + \left(\frac{\sigma_{t-2}}{\sigma_0}\right) \times \overline{A}_{t-1}^{1/2} \\
&\leq C \times \prod_{i=1}^{t-1} \overline{A}_i^{1/2}.
\end{aligned} \tag{9}$$

Then

$$
\begin{aligned}
\mathbf{Y} &= \Big(\ \sigma_0 (\epsilon_0, 1)\,, \quad \sigma_1 (\epsilon_1, 1)\,, \quad \sigma_2 (\epsilon_2, 1)\,, \quad \ldots, \quad \sigma_h (\epsilon_h, 1)\,, \ \Big) \\
&\leq \sigma_0 \times \Big(\ (\epsilon_0, 1)\,, \quad C \times \left(\tfrac{\sigma_0}{\sigma_0}\right) \times \overline{A}_1 (\epsilon_1, 1)\,, \quad C \times \left(\tfrac{\sigma_1}{\sigma_0}\right) \times \overline{A}_2 (\epsilon_2, 1)\,, \quad \ldots, \quad C \times \left(\tfrac{\sigma_{h-1}}{\sigma_0}\right) \times \overline{A}_h (\epsilon_h, 1)\,, \ \Big) \\
&\quad +\mathbf{R} \\
&\leq \sigma_0 \times \mathbf{E} + \mathbf{R},
\end{aligned}
$$

where

$$
\mathbf{E} = \left(\ (\epsilon_0, 1)\,, \quad C \times \overline{A}_1 (\epsilon_1, 1)\,, \quad C \times \overline{A}_1^{1/2} \overline{A}_2 (\epsilon_2, 1)\,, \quad \ldots, \quad C \times \left(\prod_{i=1}^{h-1} \overline{A}_i^{1/2} \right) \overline{A}_h (\epsilon_h, 1) \ \right),
$$

with

$$
\overline{A}_h = \overline{\alpha} \epsilon_{h-1}^2 + \beta \quad \forall\ h \geq 1,
$$

given (7) and (9), and

$$
\mathbf{R} = \Big(\ 0, \quad \tfrac{C}{\underline{\sigma}_1} \times \epsilon_0^2 (\epsilon_1, 1)\,, \quad \tfrac{C}{\underline{\sigma}_2} \times \epsilon_1^2 (\epsilon_2, 1)\,, \quad \ldots, \quad \tfrac{C}{\underline{\sigma}_h} \times \epsilon_{h-1}^2 (\epsilon_h, 1)\,, \ \Big)
$$

given (7). Let $\mathbf{Z} = \sigma_0 \times \mathbf{E} + \mathbf{R}$. Because $\underline{\sigma}_h^{-1}$ is bounded $\forall\ h$, the tail of $\mathbf{R}$ is 'light' relative to the tail of $\sigma_0 \times \mathbf{E}$. As a consequence, the tail of $\mathbf{Z}$ is determined by the tail of $\sigma_0 \times \mathbf{E}$. Then, since $E\left(|E_h|^{\kappa+\varepsilon}\right) < \infty\ \forall\ h$ and some $\varepsilon > 0$, $\sigma_0 \times \mathbf{E}$ is RV(κ) by (3) and Basrak et al. (2002, Corollary A.2) with $d = 1$, which means that the tail of $\mathbf{Z}$ is determined by the tail of σ_0. Since $\mathbf{Y} = \sigma_0 \times \mathbf{D} \leq \sigma_0 \times \mathbf{E} + \mathbf{R}$, the tail of $\mathbf{Y}$ is also determined by the tail of σ_0, which implies, then, that $\mathbf{Y}$ is RV(κ). ■

Let

$$
\mathbf{Y}^2 = \Big(\ (Y_0^2, \sigma_0^2)\,, \quad \ldots, \quad (Y_h^2, \sigma_h^2) \ \Big).
$$

The (general) method of proof behind the Proposition is comparable to those methods used to establish $\mathbf{Y}^2$ as RV($\kappa/2$) in Davis and Mikosch (1998, Lemma A.1) and Mikosch and Stărică (2000, Theorem 2.3). In contrast to these two aforementioned results, however, moving to establish $\mathbf{Y}$ as RV(κ) does not require a symmetric D. As a consequence, the Proposition is consistent with the empirical features (see Table 1) of many financial returns to which the model of (1) and (2) gets applied and is complementary to Basrak et al. (2002). Moreover, the Proposition explicitly covers a threshold GARCH$(1,1)$ model under empirically-relevant cases.

References

[1] Basrak, B., R.A Davis & T. Mikosch (2002) Regular variation of garch processes. *Stochastic Processes and Their Applications* 99, 95-115.

[2] Berkes, I., & L. Horváth (2004) The efficiency of the estimators of the parameters in garch processes. *Annals of Statistics* 32, 633-655.

[3] Berkes, I., L. Horváth & P. Kokoska (2003) GARCH processes: structure and estimation. *Bernoulli* 9, 201-227.

[4] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.

[5] Campbell, J.Y., A.W. Lo & A.C. MacKinlay (1997) *The Econometrics of Financial Markets.* Princeton University Press, Princeton.

[6] Carrasco, M., & X. Chen (2002) Mixing and moment properties of various garch and stochastic volatility models. *Econometric Theory* 18, 17-39.

[7] Davis, R.A., & T. Mikosch (1998) The sample autocorrelations of heavy-tailed processes with applications to arch. *The Annals of Statistics* 26, 2049-2080.

[8] Drost, F.C., & T.E. Nijman (1993) Temporal aggregation of garch processes. *Econometrica* 61, 909-927.

[9] Glosten, L.R., R. Jagannathan & D.E. Runkle (1993) On the relation between expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779-1801.

[10] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* 1, 126-166.

[11] Hansen, P.R., & A. Lunde (2005) A forecast comparison of volatility models: does anything beat a garch(1,1)? *Journal of Applied Econometrics* 20, 873-889.

[12] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta. Math.* 131, 207-248.

[13] Kristensen, D., & O. Linton (2006) A closed form estimator for the garch(1,1)-model. *Econometric Theory* 22, 323-327.

[14] Kristensen, D., & A. Rahbek (2005) Asymptotics of the qmle for a class of arch(q) models. *Econometric Theory* 21, 946-961.

[15] Lee, S.W., & B.E. Hansen (1994) Asymptotic theory for the garch(1,1) quasi-maximum likelihood estimator. *Econometric Theory* 10, 29-52.

[16] Loretan, M., & P.C.B Phillips (1994) Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance* 1, 211-248.

[17] Lumsdaine, R.L. (1996) Consistency and asymptotic normality of the quasi-maximum likelihood estimator in igarch(1,1) and covariance stationary garch(1,1) models. *Econometrica* 64, 575-596.

[18] Mikosch, T. (1999) Regular Variation, Subexponentiality and their applications in probability theory. Lecture notes for the workshop "Heavy Tails and Queques," EURANDOM, Eindhoven, Netherlands.

[19] Mikosch, T., & C. Stărică (2000) Limit theory for the sample autocorrelations and extremes of a garch(1,1) process. *The Annals of Statistics* 28, 1427-1451.

[20] Mikosch, T., & D. Straumann (2002) Whittle estimation in a heavy-tailed garch(1,1) model. Stochastic Processes and Their Applications 100, 187-222.

[21] Prono, T. (2016) Closed-form estimation of finite-order arch models: asymptotic theory and finite-sample performance. Finance and Economics Discussion Series 2016-083. Washington: Board of Governors of the Federal Reserve System.

[22] Vaynman, I. & B.K. Beare (2014) Stable limit theory for the variance targeting estimator, in Y. Chang, T.B. Fomby & J.Y. Park (eds), *Essays in Honor of Peter C.B. Phillips,* vol. 33 of *Advances in Econometrics:* Emerald Group Publishing Limited, chapter 24, 639-672.

TABLE 1

	CHF		EUR		JPY		DJIA		SPX	
freq.	obs.	skew.	obs.	skew.	obs.	skew.	obs.	skew.	obs.	skew.
1-min	174,741	0.41	190,338	-1.27	190,058	-1.59	46,557	-1.25	46,551	-1.75
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
5-min	34,973	0.35	38,081	-0.30	38,035	-1.20	9,315	-2.68	9,312	-3.17
		(0.01)		(0.01)		(0.01)		(0.03)		(0.03)
10-min	17,489	0.55	19,044	-1.26	19,021	-0.75				
		(0.02)		(0.02)		(0.02)				
15-min	11,660	0.14	12,699	-0.78	12,680	-0.73				
		(0.02)		(0.02)		(0.02)				
20-min	8,747	-0.05	9,525	-0.50	9,512	-0.49				
		(0.03)		(0.03)		(0.03)				

Notes to Tables 1. All data source to Bloomberg LP. The date range for the Swiss Franc (CHF) spot return series is 1/16/2015–7/1/2015. The date range for the Euro (EUR) and Japanese Yen (JPY) spot return series is 1/1/2015–7/1/2015. The date range for the Dow Jones Industrial Average (DJIA) and S&P 500 (SPX) spot return series is 7/19/2015–12/31/2015. Skew is an estimate of the (unconditionally) standardized third moment. The standard error for this estimate is in parentheses and is measured against a null of normality, as in Cambell, Lo, and MacKinlay (1997).