

University of Heidelberg

Department of Economics



Discussion Paper Series | No. 613

On the statistical properties of
multiplicative GARCH models

Christian Conrad and Onno Kleen

March 2016

On the statistical properties of multiplicative GARCH models*

Christian Conrad[†] and Onno Kleen[‡]

Heidelberg University, Germany

March 18, 2016

Abstract

We examine the statistical properties of multiplicative GARCH models. First, we show that in multiplicative models, returns have higher kurtosis and squared returns have a more persistent autocorrelation function than in the nested GARCH model. Second, we extend the results of Andersen and Bollerslev (1998) on the upper bound of the R^2 in a Mincer-Zarnowitz regression to the case of a multiplicative GARCH model, using squared returns as a proxy for the true but unobservable conditional variance. Our theoretical results imply that multiplicative GARCH models provide an explanation for stylized facts that cannot be captured by classical GARCH modeling.

Keywords: Forecast evaluation, GARCH-MIDAS, Mincer-Zarnowitz regression, volatility persistence, volatility component model, long-term volatility.

JEL Classification: C53, C58, G12

*We would like to thank Richard Baillie, Karin Loch, and Timo Teräsvirta for helpful comments and suggestions.

[†]Christian Conrad, Department of Economics, Heidelberg University, Bergheimer Strasse 58, 69115 Heidelberg, Germany, E-mail: christian.conrad@awi.uni-heidelberg.de; Phone: +49 6221 54 3173.

[‡]Onno Kleen, Department of Economics, Heidelberg University, Bergheimer Strasse 58, 69115 Heidelberg, Germany, E-mail: onno.kleen@awi.uni-heidelberg.de; Phone: +49 6221 54 2930.

1 Introduction

We analyze the fourth moment structure, the autocorrelation function (ACF) and the forecast evaluation properties of multiplicative GARCH models. We consider a multiplicative GARCH specification that consists of a short-term component, which is modeled as a standard GARCH(1,1), and a long-term component that varies smoothly over time. Hence, our model is closely related to the GARCH-MIDAS class of models that recently attracted considerable attention (see, e.g., Engle et al., 2013, Conrad and Loch, 2015, and Wang and Ghysels, 2015).¹ However, in our specification the short- and the long-term component are assumed to vary at the same frequency.

Assuming the long-term component to be covariance stationary, we show that the kurtosis of the returns is higher and that the ACF of the squared returns is more persistent in the multiplicative model than in the nested GARCH specification. The first result only requires a non-constant long-term component; the second one depends on the persistence of the long-term component. Both findings suggest a multiplicative volatility component structure as an explanation for the common failure of GARCH(1,1) models in adequately capturing the leptokurticity and extreme volatility persistence in observed returns. Our results are remarkably similar to the recent findings of Han (2015) on GARCH-X models, even though Han (2015) considers models with an additive explanatory variable in the conditional variance and focuses on the asymptotic limit of the sample kurtosis and the sample ACF.

Last, we show that the result of Andersen and Bollerslev (1998) on the upper bound of the R^2 in a Mincer-Zarnowitz regression, which is obtained when using squared returns as a proxy for the true but unobservable conditional variance, can directly be extended to the case of a multiplicative GARCH model. Again, the upper bound is given by one divided by the fourth moment of the innovation and will be reached as the variance of the long-term component tends to infinity.

The paper is organized as follows. In Section 2, the model and the properties of the kurtosis and the ACF are presented. In Section 3, we discuss forecast evaluation by means of Mincer-Zarnowitz regressions. All proofs are deferred to the Appendix.

¹See also Han and Kristensen (2015) for a semiparametric multiplicative component model with an exogenous covariate.

2 The Multiplicative GARCH Model

Let the sequence of (demeaned) returns be denoted by $(r_t)_t$, with $t \in T$ being the time index and let \mathcal{F}_t denote the information set up to time t . We define

$$r_t = \sigma_t Z_t, \quad (1)$$

where $(Z_t)_t$ is an *i.i.d.* innovation process with mean zero and variance one. We assume that σ_t^2 is measurable with respect to \mathcal{F}_{t-1} and, hence, represents the conditional variance of returns, i.e. $\text{Var}[r_t | \mathcal{F}_{t-1}] = \sigma_t^2$. Further, we assume that the conditional variance can be multiplicatively decomposed into two components:

$$\sigma_t^2 = g_t \tau_t. \quad (2)$$

We will refer to g_t as the short-term component and assume it to follow a GARCH(1,1) process:

$$g_t = \omega + \alpha \frac{r_{t-1}^2}{\tau_{t-1}} + \beta g_{t-1} = \omega + (\alpha Z_{t-1}^2 + \beta) g_{t-1}. \quad (3)$$

Therefore, g_t describes the well known day-to-day clustering of volatility.²

We make two assumptions regarding the innovation Z_t and the GARCH component.

Assumption 1. *Let Z_t be i.i.d. with $\mathbb{E}[Z_t] = 0$ and $\mathbb{E}[Z_t^2] = 1$. Further, Z_t^2 has a nondegenerate distribution and $\kappa = \mathbb{E}[Z_t^4] < \infty$.*

Assumption 2. *We assume that $\omega > 0$, $\alpha > 0$, $\beta \geq 0$ and $\alpha + \beta < 1$. Moreover, we assume that $\alpha^2 \kappa + 2\alpha\beta + \beta^2 < 1$.*

Note that Assumption 1 implies that $\kappa > 1$. Assumptions 1 and 2 imply that $\sqrt{g_t} Z_t$ is a covariance stationary GARCH(1,1) process. Furthermore, we have that the first- and second-order moment of the GARCH component exist, which are given by $\mathbb{E}[g_t] = \omega / (1 - \alpha - \beta)$ and

$$\mathbb{E}[g_t^2] = \frac{\omega^2 (1 + \alpha + \beta)}{(1 - \kappa \alpha^2 - \beta^2 - 2\alpha\beta)(1 - \alpha - \beta)}, \quad (4)$$

as well as that the fourth moment of $\sqrt{g_t} Z_t$ is finite.

The second component, τ_t , should be thought of as describing smooth movements in the conditional variance. Instead of explicitly specifying τ_t , we will simply assume that it follows a covariance stationary process.

²For simplicity we focus on a symmetric GARCH(1,1). However, our results can be easily extended to an asymmetric GARCH model.

Assumption 3. We assume that the long-term component $(\tau_t)_t$ is covariance stationary with $\mathbb{E}[\tau_t] = 1$ and that τ_t is measurable with respect to \mathcal{F}_{t-1} . Moreover, τ_t and Z_s are independent for all t and s .

For example, τ_t could be driven by the lagged values of an exogenous explanatory variable as in the GARCH-MIDAS model of Engle et al. (2013). In this case, τ_t may be written as $\tau_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-K})$ with $f(\cdot) > 0$, where X_t is some exogenous driver of the long-term component (see Conrad and Schienle, 2015).³

Assumptions 1, 2 and 3 imply that the returns r_t have mean zero, are uncorrelated and that their unconditional variance is given by $\text{Var}(r_t) = \mathbb{E}[g_t]$, which follows from the assumption that $\mathbb{E}[\tau_t] = \mathbb{E}[Z_t^2] = 1$. Moreover, the unconditional variance of the squared returns is well-defined, $\text{Var}(r_t^2) = \kappa \mathbb{E}[\tau_t^2] \mathbb{E}[g_t^2] - \mathbb{E}[g_t]^2$.

While the long-term component does not affect the unconditional variance of r_t , it does have an influence on the kurtosis and the ACF of the multiplicative process. Clearly, the model reduces to a GARCH(1,1) when $\tau_t = 1$.

2.1 Kurtosis

Financial returns are often found to be leptokurtic. As usual, we measure leptokurticity by means of the kurtosis coefficient. Under Assumptions 1, 2 and 3, the kurtosis of the returns defined in equation (1) is given by

$$\mathcal{K}^{MG} = \frac{\mathbb{E}[r_t^4]}{(\mathbb{E}[r_t^2])^2} = \frac{\mathbb{E}[\sigma_t^4]}{(\mathbb{E}[\sigma_t^2])^2} \kappa > \kappa.$$

Thus, the kurtosis of the multiplicative GARCH process is larger than the kurtosis of the innovations Z_t . This is a well known feature of GARCH-type processes. The following proposition relates the kurtosis \mathcal{K}^{MG} of the multiplicative GARCH to the kurtosis \mathcal{K}^{GA} of the nested GARCH(1,1).

Proposition 1. *The kurtosis \mathcal{K}^{MG} of a multiplicative GARCH process is given by*

$$\mathcal{K}^{MG} = \mathbb{E}[\tau_t^2] \cdot \mathcal{K}^{GA} > \mathcal{K}^{GA},$$

where $\mathcal{K}^{GA} = \kappa \cdot \mathbb{E}[g_t^2]/\mathbb{E}[g_t]^2$ is the kurtosis of the nested GARCH process.

³Han and Park (2014) and Han (2015) analyze the properties of a GARCH-X specification with an explanatory variable entering additively into the conditional variance equation. Obviously, this is an alternative approach for achieving that the volatility dynamics depend on exogenous information.

Hence, the kurtosis \mathcal{K}^{MG} is the product of \mathcal{K}^{GA} and the long-term component's second moment. When τ_t is constant (and equal to one), Proposition 1 nests the kurtosis of the standard GARCH model. In sharp contrast, for nearly integrated long-term components the kurtosis of a multiplicative GARCH process is much larger than the kurtosis of the nested GARCH model.⁴ Moreover, when estimating a GARCH model, it is often assumed that Z_t is standard normal so that $\kappa = 3$. Our result may explain why in empirical applications the volatility-adjusted residuals $r_t/\sqrt{g_t}$ often still exhibit 'excess' kurtosis. In the multiplicative model the kurtosis of $r_t/\sqrt{g_t}$ is given by $3 \cdot \mathbb{E}[\tau_t^2] > 3$.

2.2 Autocorrelation Function

Empirically, the ACF of squared returns is often found to decay more slowly than the exponentially decaying ACF implied by the simple GARCH(1,1) model. In literature on GARCH models, this is often interpreted as evidence for long-memory (see, e.g., Baillie, 1996), structural breaks (see, e.g., Hillebrand, 2005) or an omitted persistent covariate (see Han and Park, 2014) in the conditional variance. The following proposition shows that the theoretical ACF of the multiplicative GARCH process has a much slower decay than the ACF of the nested GARCH component if the long-term component is sufficiently persistent. Hence, the multiplicative structure provides an alternative explanation for the empirical observation of highly persistent ACF's of squared returns.

Proposition 2. *The ACF ρ_k^{MG} of a multiplicative GARCH process is given by*

$$\rho_k^{MG} = \rho_k^\tau \cdot \frac{\text{Var}(\tau_t)\mathbb{E}[g_t]^2}{\text{Var}(r_t^2)} + \rho_k^{GA} \cdot \frac{(\rho_k^\tau \text{Var}(\tau_t) + 1) \text{Var}(g_t)}{\text{Var}(r_t^2)} \quad (5)$$

where $\rho_k^\tau = \text{Corr}(\tau_t, \tau_{t-k})$ and

$$\rho_k^{GA} = \text{Corr}(g_t Z_t^2, g_{t-k} Z_{t-k}^2) = (\alpha + \beta)^{k-1} \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2}$$

is the ACF of the nested GARCH component (see Karanasos, 1999).

The ACF of multiplicative GARCH models is given by the sum of two components: The first term is given by the ACF of the long-term component ρ_k^τ times a constant,

⁴Han (2015) obtains a similar result for the sample kurtosis of the returns from a GARCH-X model with a covariate that can be either stationary or non-stationary.

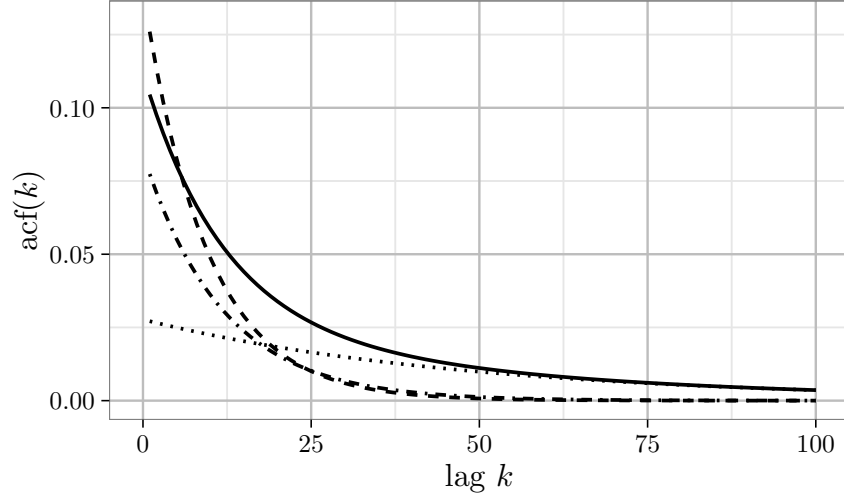


Figure 1: We depict the ACF of a multiplicative GARCH model (solid line), the first and second term in equation (5) (dotted and dot-dashed line), as well as the ACF of the nested GARCH(1,1) model (dashed line). The long-term component is defined as $\tau_t = \exp\left(X_{t-1} - \frac{\sigma_\varepsilon^2}{2(1-\phi^2)}\right)$ with exogenous covariate $X_t = \phi X_{t-1} + \varepsilon_t$, $\varepsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$, where $\phi = 0.98$ and $\sigma_\varepsilon^2 = 0.05^2$. The GARCH(1,1) parameters are $\omega = 0.02$, $\alpha = 0.06$ and $\beta = 0.86$. Moreover, we set $\kappa = 3$.

whereas the second term equals the exponentially decaying ACF of the nested GARCH model ρ_k^{GA} times a ratio that depends again on ρ_k^τ . Hence, if τ_t is sufficiently persistent, ρ_k^{MG} will essentially behave as ρ_k^τ for k large.⁵ For $\tau_t = 1$ constant, the first term in equation (5) is equal to zero and the second term reduces to the ACF of a GARCH(1,1).

The implications of Proposition 2 are depicted in Figure 1. Due to the additive structure of the correlation function, a highly persistent exogenous covariate causes the ACF of the multiplicative GARCH model to have a clearly distinct decay pattern in comparison to the ACF of its nested GARCH component. The ACF of the multiplicative GARCH behaves like the ACF typically observed for squared returns, i.e. it is below the ACF of the nested GARCH component for low lags but above for large lags. Figure 1 shows that the second term in equation (5) determines the decay behavior of ρ_k^{MG} when k is small, while the first term dominates when k is large. As for the kurtosis, our result may explain why in empirical applications the squared volatility-adjusted residuals r_t^2/g_t are still substantially autocorrelated. For the multiplicative GARCH model, the ACF of these residuals is given by $\rho_k^\tau \cdot \text{Var}(\tau_t)/(\kappa \mathbb{E}[\tau_t^2] - 1)$, which follows the rate

⁵Again, Han (2015) also obtains a two component structure for the sample ACF of the square returns from a GARCH-X model with a fractionally integrated covariate.

of decay of the long-term component. However, innovations with excess kurtosis reduce this effect. This may be a second explanation for why the volatility-adjusted residuals are often found to be leptokurtic when estimating a simple GARCH(1,1).

3 Forecast Evaluation and Mincer-Zarnowitz Regression

In this section, we assess the forecast evaluation of multiplicative GARCH models. We denote the one-step ahead volatility forecast coming from a certain volatility model by $h_{t|t-1}$ and compare it to a proxy $\hat{\sigma}_t^2$ for the true but unobservable volatility σ_t^2 . As a proxy, we use the squared returns, $\hat{\sigma}_t^2 = r_t^2$, which are conditionally unbiased, i.e. $\mathbb{E}[r_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2$. As shown in Patton (2011), the mean squared error (MSE) loss function which is given by the quadratic loss, $MSE(r_t^2, h_{t|t-1}) = (r_t^2 - h_{t|t-1})^2$, is *robust* to using a noisy but unbiased proxy for forecast evaluation. Therefore, the expected MSE leads to the same ranking of two competing forecasts whether the two are evaluated against the latent volatility or the unbiased proxy.

We will first show that – although the MSE is robust – using a noisy proxy for forecast evaluation can lead to a substantial overestimation of the true expected MSE (that would be based on the latent volatility). We then use this result to determine the R^2 in a Mincer-Zarnowitz regression.

Consider the model given by equation (1) and the corresponding expected MSE from evaluating a variance forecast $h_{t|t-1}$ against the noisy proxy r_t^2 . It is straightforward to show that

$$\mathbb{E}[MSE(r_t^2, h_{t|t-1})] = \mathbb{E}[MSE(\sigma_t^2, h_{t|t-1})] + (\kappa - 1)\mathbb{E}[\sigma_t^4].$$

That is, the expected MSE based on the noisy proxy equals the expected MSE based on the latent volatility plus a term that depends on the fourth moment of Z_t and the expected value of the squared conditional variance. Thus, even if the conditional variance forecast is *correctly specified*, i.e. $h_{t|t-1} = \sigma_t^2$ and so $\mathbb{E}[MSE(\sigma_t^2, h_{t|t-1})] = 0$, we have that

$$\mathbb{E}[MSE(r_t^2, h_{t|t-1})] = (\kappa - 1)\mathbb{E}[\sigma_t^4], \tag{6}$$

i.e. the expected MSE can be large when Z_t is leptokurtic or $\mathbb{E}[\sigma_t^4]$ is large.

Next, we illustrate the consequences of this result for evaluating the volatility forecast by means of the R^2 of a Mincer-Zarnowitz regression of the form

$$r_t^2 = \delta_0 + \delta_1 h_{t|t-1} + \eta_t. \quad (7)$$

Again, the dependent variable r_t^2 is used as a noisy proxy for the latent conditional variance σ_t^2 . When $h_{t|t-1} = \sigma_t^2$, the population parameters of the Mincer-Zarnowitz regression are given by $\delta_0 = 0$ and $\delta_1 = 1$ and, hence, the population R^2 can be written as:

$$R^2 = 1 - \frac{\mathbb{E}[MSE(r_t^2, h_{t|t-1})]}{\text{Var}[r_t^2]} = 1 - \frac{(\kappa - 1)\mathbb{E}[\sigma_t^4]}{\kappa\mathbb{E}[\sigma_t^4] - (\mathbb{E}[\sigma_t^2])^2} = \frac{1 - \frac{(\mathbb{E}[\sigma_t^2])^2}{\mathbb{E}[\sigma_t^4]}}{\kappa - \frac{(\mathbb{E}[\sigma_t^2])^2}{\mathbb{E}[\sigma_t^4]}} < \frac{1}{\kappa}.$$

The bound of $1/\kappa$ arises because the squared return is a noisy proxy for the latent volatility. This result nicely illustrates that a low R^2 is not necessarily evidence for model misspecification but can simply be due to using a noisy proxy. This point has been made before by Andersen and Bollerslev (1998), but for the special case that the true data generating process is a GARCH(1,1). Next, we derive an explicit expression for the Mincer-Zarnowitz R^2 when the data generating process is a multiplicative GARCH model.

Proposition 3. *If σ_t^2 follows a multiplicative GARCH, Assumptions 1, 2 and 3 are satisfied and $h_{t|t-1} = \sigma_t^2$, the population R^2 of the Mincer-Zarnowitz regression is given by*

$$R^2 = \frac{(1 - (\alpha + \beta)^2)\mathbb{E}[\tau_t^2] - (1 - \alpha^2\kappa - 2\alpha\beta - \beta^2)}{(1 - (\alpha + \beta)^2)\mathbb{E}[\tau_t^2]\kappa - (1 - \alpha^2\kappa - 2\alpha\beta - \beta^2)}. \quad (8)$$

Moreover, the R^2 increases monotonically in $\mathbb{E}[\tau_t^2]$ with

$$\lim_{\mathbb{E}[\tau_t^2] \rightarrow \infty} R^2 = 1/\kappa.$$

For $\tau_t = 1$ constant, equation (8) reduces to the expression in Andersen and Bollerslev (1998, p. 892), i.e. $R^2 = \alpha^2/(1 - 2\alpha\beta - \beta^2)$. Proposition 3 shows that the R^2 reaches the upper bound when τ_t is getting more volatile. At first, the result that the R^2 increases when τ_t gets more volatile may be puzzling because the expected MSE, $\mathbb{E}[MSE(r_t^2, g_t\tau_t)] = (\kappa - 1)\mathbb{E}[g_t^2]E[\tau_t^2]$, increases in such a situation. Although this observation is correct, the variance of the squared returns, $\text{Var}(r_t^2) = \kappa\mathbb{E}[\tau_t^2]\mathbb{E}[g_t^2] - \mathbb{E}[g_t]^2$, is increasing even faster, which leads to an overall increase in the R^2 .

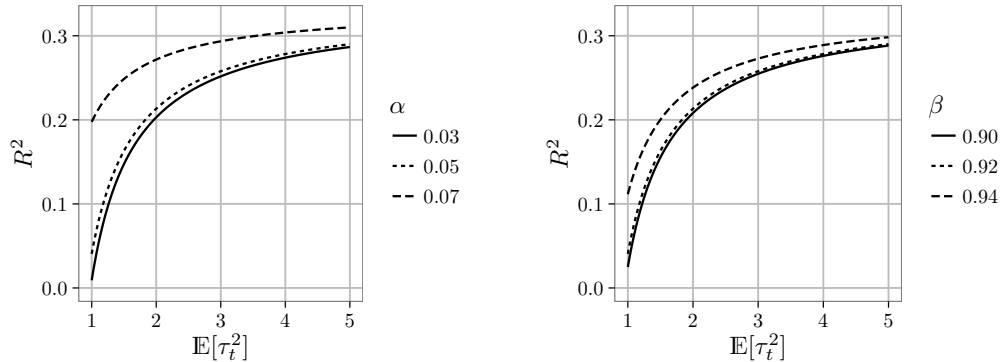


Figure 2: We depict the population R^2 of a Mincer-Zarnowitz regression as a function of $\mathbb{E}[\tau_t^2]$. In the left plot, β equals 0.92. In the right plot, we choose $\alpha = 0.05$. In all cases, we set $\kappa = 3$.

The effect of the long-term component on the Mincer-Zarnowitz R^2 is depicted in Figure 2 which shows the population R^2 as a function of $\mathbb{E}[\tau_t^2]$ for different values of α and β . For $\tau_t = 1$ being constant, the multiplicative GARCH model reduces to a GARCH(1,1). As can be seen, even a small increase in $\mathbb{E}[\tau_t^2]$ can cause a steep increase in the population R^2 . Also, for a given level of $\mathbb{E}[\tau_t^2]$, the population R^2 is the higher the more persistent (as measured by $\alpha + \beta$) the GARCH component is.

4 Discussion

We show that a multiplicative GARCH model leads to returns with higher kurtosis than in its nested GARCH specification. Similarly, the ACF of the squared returns can be much more persistent in the multiplicative model than in the nested GARCH model, provided that the long-term component is itself sufficiently persistent. Two empirically observed shortcomings of the GARCH model are that the squared standardized residuals still exhibit excess kurtosis and that the implied ACF decays too quickly in comparison with the empirical ACF of squared returns. Our results suggest that these shortcomings may be due to an omitted multiplicative long-term component. Also, note that both findings directly extend to a situation in which the GARCH(1,1) short-term component is replaced by a covariance stationary GARCH(p, q). Last, we show that the Mincer-Zarnowitz R^2 is bounded by one divided by the fourth moment of the innovation term. This finding generalizes the result presented in Andersen and Bollerslev (1998) for the GARCH(1,1) model to the setting of a multiplicative conditionally

heteroscedastic process.

References

- Andersen, T. G., Bollerslev, T., 1998. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, 39, 885-905.
- Baillie, R. T. (1996). Long Memory Processes and Fractional Integration in Econometrics. *Journal of Econometrics*, 73, 5-59.
- Conrad, C., Loch, K., 2015. Anticipating Long-Term Stock Market Volatility. *Journal of Applied Econometrics*, 30, 1090-1114.
- Conrad, C., Schienle, M., 2015. Misspecification Testing in GARCH-MIDAS Models. Department of Economics, Discussion Paper No. 597, University of Heidelberg.
- Engle, R. F., Ghysels, E., Sohn, B., 2013. Stock Market Volatility and Macroeconomic Fundamentals. *Review of Economics and Statistics*, 95, 776-797.
- Han, H., 2015. Asymptotic Properties of GARCH-X Processes. *Journal of Financial Econometrics*, 13, 188-221.
- Han, H., Kristensen, D., 2015. Semiparametric Multiplicative GARCH-X Model: Adopting Economic Variables to Explain Volatility. Working Paper.
- Han, H., Park, J.Y., 2014. GARCH with Omitted Persistent Covariate. *Economics Letters*, 124, 248-254.
- Hillebrand, E., 2005. Neglecting Parameter Changes in GARCH Models. *Journal of Econometrics*, 129, 121-138.
- Karanasos, M., 1999. The Second Moment and the Autocovariance Function of the Squared Errors of the GARCH Model. *Journal of Econometrics*, 90, 63-76.
- Patton, A., 2011. Volatility Forecast Comparison Using Imperfect Volatility Proxies. *Journal of Econometrics*, 160, 246-256.

Wang, F., Ghysels, E., 2015. Econometric Analysis of Volatility Component Models. *Econometric Theory* 31 (2), 362-393.

A Proofs

Proof of Proposition 1. The proof follows directly by applying the mutual independence of g_t , τ_t and Z_t and by noting that Assumption 3 implies that $\mathbb{E}[\tau_t^2] > 1$. \square

Proof of Proposition 2. First, note that under Assumptions 1, 2 and 3 the covariance $\text{Cov}(r_t^2, r_{t-k}^2)$ exists for every $k \in \mathbb{N}$ and is time-invariant. In the proof we use that τ_t and g_t are independent covariance stationary processes and that Z_t are *i.i.d.* innovations.

$$\begin{aligned} \rho_k^{MG} &= \frac{\text{Cov}(r_t^2, r_{t-k}^2)}{\sqrt{\text{Var}(r_t^2)} \sqrt{\text{Var}(r_{t-k}^2)}} \\ &= \frac{\mathbb{E}[\tau_t \tau_{t-k}] \mathbb{E}[g_t Z_t^2 g_{t-k} Z_{t-k}^2] - \mathbb{E}[g_t] \mathbb{E}[g_{t-k}]}{\text{Var}(r_t^2)} \\ &= \frac{(\mathbb{E}[\tau_t \tau_{t-k}] - 1) \mathbb{E}[g_t]^2}{\text{Var}(r_t^2)} + \frac{(\mathbb{E}[g_t Z_t^2 g_{t-k} Z_{t-k}^2] - \mathbb{E}[g_t] \mathbb{E}[g_{t-k}]) \mathbb{E}[\tau_t \tau_{t-k}]}{\text{Var}(r_t^2)} \\ &= \frac{\text{Cov}(\tau_t, \tau_{t-k}) \mathbb{E}[g_t]^2}{\text{Var}(r_t^2)} + \frac{\text{Cov}(g_t, g_{t-k}) (\text{Cov}(\tau_t, \tau_{t-k}) + 1)}{\text{Var}(r_t^2)} \end{aligned}$$

\square

Proof of Proposition 3. Using equation (4), we obtain

$$\begin{aligned} R^2 &= \frac{\text{Var}(g_t \tau_t)}{\text{Var}(r_t^2)} = \frac{\mathbb{E}[g_t^2] \mathbb{E}[\tau_t^2] - \mathbb{E}[g_t]^2}{\mathbb{E}[g_t^2] \mathbb{E}[\tau_t^2] \kappa - \mathbb{E}[g_t]^2} \\ &= \frac{(1 - (\alpha + \beta)^2) \mathbb{E}[\tau_t^2] - (1 - \alpha^2 \kappa - 2\alpha\beta - \beta^2)}{(1 - (\alpha + \beta)^2) \mathbb{E}[\tau_t^2] \kappa - (1 - \alpha^2 \kappa - 2\alpha\beta - \beta^2)}. \end{aligned}$$

When considering R^2 as a function of $\mathbb{E}[\tau_t^2]$, the monotonicity and the limit follow from rewriting R^2 as:

$$R^2 = 1 - \frac{(\kappa - 1) \mathbb{E}[g_t^2] \mathbb{E}[\tau_t^2]}{\kappa \mathbb{E}[\tau_t^2] \mathbb{E}[g_t^2] - \mathbb{E}[g_t]^2} = 1 - \frac{\kappa - 1}{\kappa - \frac{\mathbb{E}[g_t]^2}{\mathbb{E}[g_t^2] \mathbb{E}[\tau_t^2]}}$$

\square