Higher-order dependence in the general Power ARCH process and a special case

by

Changli He and Timo Teräsvirta

Department of Economic Statistics

Stockholm School of Economics

Box 6501, SE-113 83 Stockholm, Sweden

First draft: April 1999

Abstract. In this paper we consider a general rst-order power ARCH process and, in particular, a special case in which the power parameter approaches zero. These considerations give us the autocorrelation function of the logarithms of the squared observations for rst-order exponential and logarithmic GARCH processes. These autocorrelations decay exponentially with the lag and may be used for checking how well an estimated exponential or logarithmic GARCH model characterizes the corresponding autocorrelation structure of the observations. The results of the paper are also useful in illustrating differences in the autocorrelation structures of the classical rst-order GARCH and the exponential or logarithmic GARCH models.

Key Words. Box-Cox transformation, conditional heteroskedasticity, exponential GARCH, logarithmic GARCH, higher-order dependence

JEL Classi cation Code: C22

Acknowledgements. The rst author acknowledges nancial support from the Tore Browaldh s Foundation. The research of the second author has been supported by the Bank of Sweden Tercenterary Foundation.

1 Introduction

In a recent paper, Ding, Granger and Engle (1993) introduced a class of autoregressive conditional heteroskedastic models called Asymmetric Power Autoregressive Conditional Heteroskedastic (A-PARCH) models. The authors showed that this class contains as special cases a large number of well-known ARCH and GARCH models. The A-PARCH model contains a particular power parameter that makes the conditional variance equation nonlinear in parameters. Among other things, Ding, Granger and Engle showed that by letting the power parameter approach zero, the A-PARCH family of models also includes the logarithmic GARCH model as a special case. Hentschel (1995) defined a slightly extended A-PARCH model and showed that after this extension, the A-PARCH model also contains the exponential GARCH (EGARCH) model of Nelson (1991) as a special case as the power parameter approaches zero. Allowing this to happen in a general A-PARCH model forms a starting-point for our investigation.

A notable feature of the A-PARCH model is that, due to its parameterization, it is possible to find analytically only certain fractional moments of the absolute values of the original process related to the power parameter. Expressions for such moments were derived in He and Teräsvirta (1999a). In this paper we first define a slight generalization of the class of EGARCH models. Then we derive the autocorrelation function of squared and logarithmed observations for this class of models. For Nelson's EGARCH model it is possible to reconcile our results with those in Breidt, Crato and de Lima (1998). Furthermore, we show that this autocorrelation function follows as a limiting case from the autocorrelation function of some fractional powers of the absolute values of the original observations. On the other hand, if we want to derive the autocorrelation function of squares of the original observations and not their logarithms for the EGARCH model then the techniques applied in this paper do not apply. The solution to that problem will be discussed elsewhere. The autocorrelation functions considered here can be used for evaluating an estimated model by checking how well the model is able to reproduce stylized facts. In the present situation this means estimating the autocorrelation function of the squared and logarithmed observations from the data and comparing that with the corresponding values obtained by plugging the parameter estimates into the theoretical expressions of the autocorrelations derived here. An example can be found in He and Teräsvirta (1999a).

The plan of the paper is as follows. Section 2 de nes the class of models of interest and introduces notation. The main results appear in Section 3. Section 4 contains nal remarks. All proofs can be found in the appendix.

2 The model

Let $\{\varepsilon_t\}$ be a real-valued discrete time stochastic process generated by

$$\varepsilon_t = z_t h_t \tag{1}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with mean zero and unit variance, and h_t is a $\mathcal{F}_{t-1}$-measurable function, where $\mathcal{F}_{t-1}$ is the sigma-algebra generated by $\{z_{t-1}, z_{t-2}, z_{t-3}, ...\}$, and positive for all t with probability one. Let

$$h_t^{2\delta} = \alpha_0 + c_\delta(z_{t-1})h_{t-1}^{2\delta}, \;\; \delta > 0 \tag{2}$$

where α_0 is a positive scalar and $c_{\delta t} = c_\delta(z_t)$ is a well-de ned function of z_t. The sequence $\{c_{\delta t}\}$ is a sequence of independent identically distributed random variables such that each $c_{\delta t}$ is stochastically independent of $h_t^{2\delta}$. Function $c_{\delta t}$ contains parameters that determine the moment structure of $\{\varepsilon_t\}$. Constrains on these parameters are necessary to guarantee that $h_t^{2\delta}$ remains positive with probability 1. We call (1) and (2) a general power ARCH (GPARCH$(\delta, 1, 1)$) model. This model appeared in He and Teräsvirta (1999a) in a slightly more general form with $\alpha_0 = g(z_t)$ being a stochastic variable.

Setting $c_\delta(z_{t-1}) = \alpha(|z_t| - \phi z_t)^{2\delta} + \beta$ in equation (2), defines, together with equation (1), the Asymmetric Power ARCH (A-PARCH) (1,1) model of Ding, Granger and Engle (1993). Note that these authors use δ in place of 2δ in equation (2) but that does not affect the results. Hentschel (1995) also defined a parametric family of GARCH models similar to (1) and (2) for highlighting relations between different GARCH models and their treatment of asymmetry.

In this paper we are interested in the limiting case $\delta \to 0$. Taking logarithms of (1) yields

$$\ln \varepsilon_t^2 = \ln z_t^2 + \ln h_t^2. \tag{3}$$

On the other hand, equation (1) can be modified so that it relates the Box-Cox transformed ε_t^2, that is, $\varphi_\delta(\varepsilon_t^2) = (\varepsilon_t^{2\delta} - 1)/\delta$, to $(h_t^{2\delta} - 1)/\delta$. Then by applying l'Hôpital's rule it can be shown that letting $\delta \to 0$ in the modified equation also leads to (3). This entitles us to consider certain exponential GARCH models as limiting cases of the power ARCH model (1) and (2). In order to see that, rewrite (2) in terms of $(h_t^{2\delta} - 1)/\delta$ and define $c_\delta(z_t) = \delta g(z_t) + \beta$. It can be shown that under certain conditions, as $\delta \to 0$, equation (2) becomes

$$\ln h_t^2 = \alpha_0 + g(z_{t-1}) + \beta \ln h_{t-1}^2 \tag{4}$$

where $g(z_t)$ is a well-defined function of z_t. Equation (4) is thus nested in (2). Equations (1) and (4) or (3) and (4) define a class of GPARCH(0,1,1) models that contains certain well-known models as special cases. For example, setting $g(z_t) = \phi z_t + \psi(|z_t| - \mathsf{E}\,|z_t|)$ in

(4) yields

$$\ln h_t^2 = \alpha_0 + \phi z_{t-1} + \psi(|z_{t-1}| - \mathsf{E}\,|z_{t-1}|) + \beta \ln h_{t-1}^2 \tag{5}$$

which, jointly with (1), de nes the EGARCH(1,1) model of Nelson (1991). Similarly, we may set $c_\delta(z_t) = \alpha g_1^\delta(z_t) + \beta$ where $g_1(z_t) > 0$ for all t with probability one. Then, by l Hôpital s rule, (2) converges to

$$\ln h_t^2 = \alpha_0 + \alpha \ln g_1(z_{t-1}) + (\alpha + \beta) \ln h_{t-1}^2 \tag{6}$$

as $\delta \to 0$. Equations (1) and (6) de ne a class of logarithmic GARCH (LGARCH(1,1)) models. Setting $g_1(z_t) = z_t^2$ in (6) yields

$$\ln h_t^2 = \alpha_0 + \alpha \ln \varepsilon_{t-1}^2 + \beta \ln h_{t-1}^2 \tag{7}$$

which is the LGARCH(1,1)) model of Geweke (1986) and Pantula (1986). Since (4) and (6) have a similar structure we mainly consider results for the GPARCH(0,1,1) model (1) and (4). They can be easily modi ed to apply to the class of LGARCH(1,1) models.

3 The limiting results

In this section we derive the asymptotic moment structure of the GPARCH(1,1) model (1) and (2) as $\delta \to 0$ under the Box-Cox transformation. We rst give the moment structure of (1) and (2) for $\delta > 0$. Having done that we derive the moment structure of model (3) with (4). Finally, we show that this result may be also obtained as a limiting case of model (1) with (2) as $\delta \to 0$.

To formulate our rst result let $\gamma_\delta = \mathsf{E}c_{\delta t}$ and $\gamma_{2\delta} = \mathsf{E}c_{\delta t}^2$. We have

Lemma 1 *For the GPARCH(δ,1,1) model (1) with (2), a necessary and sufficient condition for the existence of the 4δ-th unconditional moment $\mu_{4\delta} = \mathsf{E}\,|\varepsilon_t|^{4\delta}$ of $\{\varepsilon_t\}$ is*

$$\gamma_{2\delta} < 1. \tag{8}$$

If (8) holds, then

$$\mu_{4\delta} = \alpha_0^2 \nu_{4\delta}(1+\gamma_\delta)/\{(1-\gamma_\delta)(1-\gamma_{2\delta})\} \tag{9}$$

where $\nu_{2\psi} = \mathsf{E}\,|z_t|^{2\psi}$, $\psi > 0$. *The autocorrelation function* $\rho_n(\delta) = \rho(|\varepsilon_t|^{2\delta}, |\varepsilon_{t-n}|^{2\delta})$, $n \geq 1$, *of* $\{|\varepsilon_t|^{2\delta}\}$ *has the form*

$$\rho_1(\delta) = \frac{\nu_{2\delta}[\overline{\gamma}_\delta(1-\gamma_\delta^2) - \nu_{2\delta}\gamma_\delta(1-\gamma_{2\delta})]}{\nu_{4\delta}(1-\gamma_\delta^2) - \nu_{2\delta}^2(1-\gamma_{2\delta})} \tag{10}$$

where $\overline{\gamma}_\delta = \mathsf{E}(|z_t|^{2\delta} c_{\delta t})$, *and* $\rho_n(\delta) = \gamma_\delta \rho_{n-1}(\delta)$, $n \geq 2$.

Let $\mathcal{M}_\delta(\mu_{4\delta}, \rho_n(\delta))$ be the analytic moment structure de ned by Lemma 1 for the GPARCH(δ,1,1) model (1) and (2). It is seen that $\mathcal{M}_\delta(\cdot)$ is a function of power parameter δ. Note that the autocorrelation function of $\{|\varepsilon_t|^{2\delta}\}$ is decaying exponentially with the discount factor γ_δ. In particular, setting $\delta = 1$ in equations (8) and (10) yields the existence condition of the fourth moment and the autocorrelation function of the squared observations of the standard GARCH(1,1) model (Bollerslev, 1986) with non-normal errors.

It is customary to also consider the kurtosis of any given GARCH process, see, for example, Bollerslev (1986) or He and Teräsvirta (1999b). In this case, the kurtosis of $|\varepsilon_t|^\delta$ or $\varphi_\delta(|\varepsilon_t|) = (|\varepsilon_t|^\delta - 1)/\delta$ may be de ned as

$$\begin{aligned} \square_4(\delta) &= \frac{\mathsf{E}(|\varepsilon_t|^\delta - \mathsf{E}\,|\varepsilon_t|^\delta)^4}{\{\mathsf{E}(|\varepsilon_t|^\delta - \mathsf{E}\,|\varepsilon_t|^\delta)^2\}^2} \\ &= \frac{\mathsf{E}(\varphi_\delta(|\varepsilon_t|) - \mathsf{E}\varphi_\delta(|\varepsilon_t|))^4}{\{\mathsf{E}(\varphi_\delta(|\varepsilon_t|) - \mathsf{E}\varphi_\delta(|\varepsilon_t|))^2\}^2} \end{aligned}$$

so that the limiting case

$$\lim_{\delta \to 0} \square_4(\delta) = \frac{\mathsf{E}(\ln|\varepsilon_t| - \mathsf{E}\ln|\varepsilon_t|)^4}{\{\mathsf{E}(\ln|\varepsilon_t| - \mathsf{E}\ln|\varepsilon_t|)^2\}^2}. \tag{11}$$

The kurtosis (11) is thus the limiting case of the kurtosis of the absolute-valued process $\{|\varepsilon_t|^\delta\}$. Computing it would require the expectations $\mathsf{E}(\ln|\varepsilon_t|)^4$ and $\mathsf{E}(\ln|\varepsilon_t|)^3$ or, alternatively, $\mathsf{E}(\ln\varepsilon_t^2)^4$ and $\mathsf{E}(\ln\varepsilon_t^2)^3$ for which no analytical expressions have been derived above. The kurtosis of $\ln|\varepsilon_t|$ is a concept quite different from that of ε_t, which is why it is not considered any further here.

For the GPARCH(0,1,1) process we obtain the following result:

Lemma 2 *For the GPARCH(0,1,1) process (3) and (4) the second unconditional moment of* $\ln\varepsilon_t^2$ *exists if and only if*

$$|\beta| < 1. \tag{12}$$

When (12) holds, this second moment can be expressed as

$$\mu_0 = \mathsf{E}(\ln\varepsilon_t^2)^2 = \frac{\Delta}{(1-\beta)(1-\beta^2)} \tag{13}$$

where $\Delta = \gamma_{(\ln z^2)^2}(1-\beta)(1-\beta^2) + 2\gamma_{\ln z^2}(\alpha_0+\gamma_g)(1-\beta^2) + [\alpha_0^2(1+\beta)+2\alpha_0(1+\beta)\gamma_g + 2\beta\gamma_g^2 + (1-\beta)\gamma_{g^2}]$ *and* $\gamma_{(\ln z^2)^2} = \mathsf{E}(\ln z_t^2)^2$, $\gamma_{\ln z^2} = \mathsf{E}\ln z_t^2$, $\gamma_g = \mathsf{E}g(z_t)$ *and* $\gamma_{g^2} = \mathsf{E}(g(z_t))^2$. *Furthermore, the autocorrelation function* $\rho_n^0 = \rho(\ln\varepsilon_t^2, \ln\varepsilon_{t-n}^2)$, $n \geq 1$, *of* $\{\ln\varepsilon_t^2\}$ *has the form*

$$\begin{aligned} \rho_1^0 &= \frac{(1-\beta^2)(\gamma_{g\ln z^2} - \gamma_g\gamma_{\ln z^2}) + \beta(\gamma_{g^2} - \gamma_g^2)}{(1-\beta^2)(\gamma_{(\ln z^2)^2} - \gamma_{\ln z^2}^2) + (\gamma_{g^2} - \gamma_g^2)}, \\ \rho_n^0 &= \rho_1^0\beta^{n-1}, n \geq 2, \end{aligned} \tag{14}$$

where $\gamma_{g\ln z^2} = \mathsf{E}(g(z_t)\ln z_t^2)$.

Nelson (1991) derived the autocovariance function of the logarithm of the conditional variance of the EGARCH process. Breidt, Crato and Lima (1998) obtained the autocorrelation function of $\{\ln\varepsilon_t^2\}$ for the EGARCH model. In both articles the authors made use of the in nite moving average representation of the logarithm of the conditional variance. Lemma 2 gives the corresponding result for the rst-order process directly in terms of the parameters of the original model, which is practical for model evaluation purposes.

Let $\mathcal{M}_0(\mu_0, \rho_n^0)$ be the moment structure de ned by Lemma 2 for the GPARCH(0,1,1) process (3) and (4). Next, let $I_\delta = (0, l)$, $l > 0$, be an open interval such that $\delta_0 \in I_\delta$ if and only if $\gamma_{2\delta_0} < 1$. We have

Theorem. *Assume that $\mathcal{M}_\delta(\cdot)$ is de ned on I_δ and the corresponding functions in $\mathcal{M}_\delta(\cdot)$ are continuous on I_δ and are twice differentiable with respect to δ. Then under the transformation $\varphi_\delta(\varepsilon_t^2) = (\varepsilon_t^{2\delta} - 1)/\delta$,*

$$\mathcal{M}_\delta(\mu_{4\delta}, \rho_n(\delta)) \to \mathcal{M}_0(\mu_0, \rho_n^0) \tag{15}$$

as $\delta \to 0$.

Remark. It has been pointed out above that, under the Box-Cox transformation $\varphi_\delta(\varepsilon_t^2) = (\varepsilon_t^{2\delta} - 1)/\delta$, equation (1), when appropriately modi ed, converges to equation (3) as $\delta \to 0$. The theorem then says that under this transformation the moment structure of the GPARCH(δ,1,1) model (1) and (2) approaches the moment structure of the GPARCH(0,1,1) model as $\delta \to 0$: $\mu_{4\delta} \to \mu_0$ and $\rho_n(\delta) \to \rho_n^0$. This convergence shows that the moment structure $\mathcal{M}_0(\cdot)$ belongs to the class of structures $\mathcal{M}_\delta(\cdot)$ as a boundary case. These moment structures are thus isomorphic. Besides, the parameter δ in the GPARCH(δ,1,1) process de nes a value for which the autocorrelation function $\rho(|\varepsilon_t|^\delta, |\varepsilon_{t-k}|^\delta)$, $k \geq 1$, decays exponentially with k.

To consider the practical value of these results suppose, for example, that $\gamma_4 < 1$. Then we have a class of GPARCH($\delta, 1, 1$) models with the same parameter values such that the available $\mathcal{M}_\delta(\cdot)$ is de ned on $[0, 1]$, that is, $\gamma_{2\delta} < 1$, δ □ 1. Practitioners may want to use these results to see what kind of moment implications GPARCH models they estimate may have. Results in $\mathcal{M}_\delta(\cdot)$ de ned on $[0, 1]$ may also be useful in checking how well different GPARCH models represent the reality, which is done by comparing parametric moment estimates from a GPARCH($\delta, 1, 1$) model with corresponding nonparametric ones obtained directly from the data. First-order LGARCH and EGARCH models may thus be compared with, say, a standard GARCH(1,1) model in this respect if both are estimated using the same data.

4 Final remarks

We have derived the autocorrelation structure of the logarithms of squared observations of a class of power ARCH processes. This structure may be obtained as a limiting case of a general power ARCH model. An interesting thing to notice is that the autocorrelation structure of the δth power of absolute-valued observations of this rst-order GPARCH process is exponential for all GPARCH$(\delta, 1, 1)$ processes such that the 4δth fractional moment exists. This property is retained at the limit as the power parameter approaches zero, which means that the autocorrelation function of the process of logarithms of squared observations also decay exponentially. It may be guessed that while this is true for the logarithmed squared observations of an LGARCH(1,1) or EGARCH(1,1) process it cannot simultaneously be true for the untransformed observations de ned by these processes. This turns out to be the case, and the explicit form of the autocorrelation function of squared observations of an LGARCH or an EGARCH process will be considered in another paper.

Conversely, if we have the original GARCH(1,1) [GPARCH(1,1,1)] process of Bollerslev (1986) with the autocorrelations of $\{\varepsilon_t^2\}$ decaying exponentially, the autocorrelation function of $\{\ln \varepsilon_t^2\}$ does not have this property. The practical value of these facts when discriminating between GARCH(1,1) and EGARCH(1,1) models is not clear, but they illustrate the theoretical differences in the higher-order dynamics between these two classes of models. Note that possible asymmetry is not an issue here. Nelson s EGARCH(1,1) model is a member of the GPARCH$(0, 1, 1)$ family independent of the value of the asymmetry parameter. Likewise, if the standard GARCH(1,1) process is generalized to an asymmetric GJR-GARCH(1,1) (Glosten, Jagannathan and Runkle, 1993) process the argument remains the same. This is because the GJR-GARCH model is still a member of the GPARCH(1,1,1) class; see Ding, Engle and Granger (1993) and He and Teräsvirta (1999b) for more discussion.

References

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

Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis* **9**, 121-131.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994). ARCH models. In: R.F. Engle and D.L. McFadden, eds. *Handbook of econometrics*, Vol. **4**, 2959-3038. Amsterdam: Elsevier.

Breidt, F.J., N. Crato and P. de Lima (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* **83**, 325-348.

Ding, Z., C.W.J. Granger and R.F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**, 83-106.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom in ation. *Econometrica* **50**, 987-1006.

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

Geweke, J. (1986). Modelling persistence of conditional variances. Comment. *Econometric Reviews* **5**, 57-61.

He, C. and T. Teräsvirta (1999a). Statistical properties of the Asymmetric Power ARCH process. In: R.F. Engle and H. White, eds. *Cointegration, causality, and forecasting. Festschrift in honour of Clive W.J Granger* (forthcoming).

He, C. and T. Teräsvirta (1999b). Properties of moments of a family of GARCH processes. *Journal of Econometrics* (forthcoming).

Hentschel, L. (1995). All in the family. Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* **39**, 71-104.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**, 347-370.

Pantula, S.G. (1986). Modelling persistence of conditional variances. Comment. *Econometric Reviews* **5**, 71-74.

Appendix

Proof. [Lemma 1] The proof can be found in He and Teräsvirta (1999a). ■

Proof. [Lemma 2] First, consider the moment condition (12). Repeatedly applying (4) to $\ln h_t^2$ gives

$$\ln h_t^2 = \alpha_0 \sum_{i=1}^{k} \beta^{i-1} + \sum_{i=1}^{k+1} \beta^{i-1} g(z_{t-1}) + \beta^{k+1} \ln h_{t-(k+1)}^2. \tag{A.1}$$

Assume that the process started at some finite value infinitely many periods ago. Taking expectations of both sides of (A.1) and letting $k \to \infty$ yield

$$\mathsf{E}(\ln h_t^2) = (\alpha_0 + \gamma_g)/(1-\beta) \tag{A.2}$$

if and only if $|\beta| < 1$. Similarly,

$$(\ln h_t^2)^2 = \beta^{2k} (\ln h_{t-k}^2)^2 + \sum_{j=1}^{2} \binom{2}{j} \sum_{i=1}^{k} \beta^{2(i-1)} (\alpha_0 + g(z_{t-i}))^j (\beta \ln h_{t-i}^2)^{2-j}.$$

Thus, by letting $k \to \infty$ and taking expectations

$$\begin{aligned} \mathsf{E}(\ln h_t^2)^2 &= [\alpha_0^2(1+\beta) + 2\alpha_0(1+\beta)\gamma_g + 2\beta\gamma_g^2 \\ &\quad +(1-\beta)\gamma_{g^2}]/[(1-\beta)(1-\beta^2)] \end{aligned} \tag{A.3}$$

if and only if (12) holds. It follows from formulas (3), (A.2) and (A.3) that expression (13) is valid.

Next, consider the n-th order autocorrelation of $\{\ln \varepsilon_t^2\}$

$$\rho_n^0 = \frac{\mathsf{E}(\ln \varepsilon_t^2 \ln \varepsilon_{t-n}^2) - (\mathsf{E}(\ln \varepsilon_t^2))^2}{\mathsf{E}(\ln \varepsilon_t^2)^2 - (\mathsf{E}(\ln \varepsilon_t^2))^2}. \tag{A.4}$$

We have

$$\ln \varepsilon_t^2 \ln \varepsilon_{t-n}^2 = \ln z_t^2 \ln z_{t-n}^2 + \ln z_t^2 \ln h_{t-n}^2 + \ln h_t^2 \ln z_{t-n}^2 + \ln h_t^2 \ln h_{t-n}^2. \tag{A.5}$$

It follows from (A.1) that

$$\ln h_t^2 \ln h_{t-n}^2 = \alpha_0 \sum_{i=1}^{n} \beta^{i-1} \ln h_{t-n}^2 + (\sum_{i=1}^{n} \beta^{i-1} g(z_{t-1})) \ln h_{t-n}^2 + \beta^n (\ln h_{t-n}^2)^2 \tag{A.6}$$

and

$$\begin{aligned} \ln h_t^2 \ln z_{t-n}^2 &= \alpha_0 \sum_{i=1}^{n} \beta^{i-1} \ln z_{t-n}^2 + (\sum_{i=1}^{n-1} \beta^{i-1} g(z_{t-1})) \ln z_{t-n}^2 \\ &\quad + \beta^{n-1} g(z_{t-n}) \ln z_{t-n}^2 + \beta^n \ln z_{t-n}^2 \ln h_{t-n}^2. \end{aligned} \tag{A.7}$$

Taking expectations of both sides of (A.6) and (A.7) and inserting them with (A.2) and (A.3) to (A.4) observing (A.5) yields (14). ■

Proof. [Theorem] For the ease of exposition, write (2) as

$$h_t^{2\delta} = \alpha_0^* + c_\delta(z_{t-1}) h_{t-1}^{2\delta} \tag{A.8}$$

where $c_\delta(z_{t-1}) = \delta g(z_{t-1}) + \beta$. Following Ding, Granger and Engle (1993), decompose α_0^* as

$$\alpha_0^* = (1 - \gamma_\delta)\omega^\delta \tag{A.9}$$

where $\gamma_\delta = \delta\gamma_g + \beta$ and $\omega^\delta = \mathsf{E} h_t^{2\delta}$, $\omega > 0$. Rewrite (A.8) as

$$\begin{aligned} (h_t^{2\delta} - 1)/\delta &= (\alpha_0^* + \beta - 1)/\delta + g(z_{t-1}) h_{t-1}^{2\delta} \\ &\quad + \beta(h_{t-1}^{2\delta} - 1)/\delta. \end{aligned} \tag{A.10}$$

Insert (A.9) into (A.10) and let $\delta \to 0$ on both sides of (A.10). Then, by l Hôpital s rule (A.10) converges to (4). In particular,

$$(\alpha_0^* + \beta - 1)/\delta \to \alpha_0 \tag{A.11}$$

where $\alpha_0 = (1-\beta)(\mathsf{E} \ln h_t^2) - \gamma_g$ is the constant term in (4). Besides, from (A.9) we have, as $\delta \to 0$,

$$\alpha_0^* \to 1 - \beta. \tag{A.12}$$

The convergence results (A.11) and (A.12) are used to prove the following results.

(i) We shall show that $\mu_{4\delta} \to \mu_0$ as $\delta \to 0$ under the Box-Cox transformation. From Lemma 1

$$\mu_{2\delta} = \mathsf{E}\varepsilon_t^{2\delta} = \frac{\alpha_0^* \nu_{2\delta}}{1-\gamma_\delta}. \tag{A.13}$$

Rewrite (A.13) as

$$\mathsf{E}\varphi_\delta(\varepsilon_t^2) = \frac{[\alpha_0^*(\nu_{2\delta}-1) + (\alpha_0^* + \beta - 1)]/\delta + \gamma_g}{1-\gamma_\delta}. \tag{A.14}$$

Letting $\delta \to 0$ on both sides of (A.14) and applying (A.11) and (A.12) to the right-hand side of (A.14) gives

$$\mu_2^0 = \mathsf{E}\ln\varepsilon_t^2 = \frac{\gamma_{\ln z^2}(1-\beta) + (\alpha_0 + \gamma_g)}{1-\beta}. \tag{A.15}$$

From (9) if follows that

$$\mathsf{E}(\varphi_\delta(\varepsilon_t^2))^2 = [\frac{\alpha_0^{*2}\nu_{4\delta}(1+\gamma_\delta)}{(1-\gamma_\delta)(1-\gamma_{2\delta})} - 2\mathsf{E}\varepsilon_t^{2\delta} + 1]/\delta^2. \tag{A.16}$$

Applying (A.13) it is seen that expression (A.16) is equivalent to

$$\begin{aligned}\mathsf{E}(\varphi_\delta(\varepsilon_t^2))^2 \quad = \quad & \frac{1}{(1-\gamma_\delta)(1-\gamma_{2\delta})}\{\frac{1}{\delta^2}[\alpha_0^{*2}\nu_{4\delta}(1+\beta) - 2\alpha_0^*\nu_{4\delta}(1-\beta^2) \\ & +1-\beta-\beta^2+\beta^3] + \frac{1}{\delta}[\alpha_0^{*2}\nu_{4\delta}\gamma_g + 4\alpha_0^*\beta\nu_{2\delta}\gamma_g \\ & -\gamma_g(1+2\beta-3\beta^2)] + [2\alpha_0^*\beta\nu_{2\delta}\gamma_{g^2} - \gamma_{g^2}(1-\beta) \\ & +2\beta\gamma_g^2] + \delta\gamma_g\gamma_{g^2}\}. \end{aligned} \tag{A.17}$$

Note that, as $\delta \to 0$, $\mathsf{E}(\varphi_\delta(\varepsilon_t^2))^2 \to \mathsf{E}(\ln\varepsilon_t^2)^2$, $(\nu_{4\delta} - 2\nu_{2\delta} + 1)/\delta^2 \to \mathsf{E}(\ln z_t^2)^2$ and $(\nu_{2\delta} - 1)/\delta \to \mathsf{E}(\ln z_t^2)$. Apply those facts and (A.11) and (A.12) to the right-hand side of (A.17) while letting $\delta \to 0$ on both sides of (A.17). It follows from l Hôpital s rule that (A.17) converges to

$$\mu_0 = \mathsf{E}(\ln\varepsilon_t^2)^2 = \frac{\Delta}{(1-\beta)(1-\beta^2)}. \tag{A.18}$$

Then $\mu_{4\delta} \to \mu_0$ holds in (15).

(ii) Second, we prove that $\lim_{\delta\to 0} \rho_n(\delta) = \rho_n^0$. Since $\lim_{\delta\to 0} \rho_n(\delta) = \lim_{\delta\to 0} \rho_1(\delta)\gamma_\delta^{n-1} = \beta^{n-1} \lim_{\delta\to 0} \rho_1(\delta)$, we have to prove that $\lim_{\delta\to 0} \rho_1(\delta) \to \rho_1^0$.

Let $\rho_1(\delta) = u/v$ in (10) where $u = \nu_{2\delta}[\overline{\gamma}_\delta(1-\gamma_\delta^2) - \nu_{2\delta}\gamma_\delta(1-\gamma_{2\delta})]$ and $v = \nu_{4\delta}(1-\gamma_\delta^2) - \nu_{2\delta}^2(1-\gamma_{2\delta})$. Since $\lim_{\delta\to 0} u = 0$ and $\lim_{\delta\to 0} v = 0$ we need to apply l Hôpital s rule in order to obtain $\lim_{\delta\to 0} \rho_1(\delta)$. Note that

$$\begin{aligned}
\frac{\partial}{\partial\delta}u &= (\nu_{2\delta} - \beta^2\nu_{2\delta} - \nu_{2\delta}^2 + \beta^2\nu_{2\delta}^2)\overline{\gamma}_g \\
&+\delta\frac{\partial}{\partial\delta}(\nu_{2\delta} - \beta^2\nu_{2\delta} - \nu_{2\delta}^2 + \beta^2\nu_{2\delta}^2)\overline{\gamma}_g \\
&+\frac{\partial}{\partial\delta}(-\delta^3\nu_{2\delta}\overline{\gamma}_g\gamma_g^2 - \delta^2\beta\nu_{2\delta}^2\gamma_g^2 - 2\delta^2\beta\nu_{2\delta}\overline{\gamma}_g\gamma_g \\
&+\delta^3\nu_{2\delta}^2\gamma_g^2\gamma_{g^2} + \delta^2\beta\nu_{2\delta}^2\gamma_{g^2} + 2\delta^2\beta\gamma_g^2)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial\delta}v &= \frac{\partial}{\partial\delta}(\nu_{4\delta}(1 - \delta^2\gamma_g^2 - 2\delta\beta\gamma_g - \beta^2) \\
&-\nu_{2\delta}^2(1 - \delta^2\gamma_{g^2} - 2\delta\beta\gamma_g - \beta^2))
\end{aligned}$$

imply that $\lim_{\delta\to 0} \frac{\partial}{\partial\delta}u = 0$ and $\lim_{\delta\to 0} \frac{\partial}{\partial\delta}v = 0$. Thus we have to calculate $\frac{\partial^2}{\partial\delta^2}u$ and $\frac{\partial^2}{\partial\delta^2}v$. We obtain

$$\begin{aligned}
\frac{\partial^2}{\partial\delta^2}u &= \frac{\partial}{\partial\delta}[(\nu_{2\delta} - \beta^2\nu_{2\delta} - \nu_{2\delta}^2 + \beta^2\nu_{2\delta}^2)\overline{\gamma}_g] \\
&+\frac{\partial}{\partial\delta}(\nu_{2\delta}\overline{\gamma}_g - \beta^2\nu_{2\delta}\overline{\gamma}_g - \nu_{2\delta}^2\overline{\gamma}_g + \beta^2\nu_{2\delta}^2\overline{\gamma}_g) \\
&+\delta\frac{\partial^2}{\partial\delta^2}(\nu_{2\delta}\overline{\gamma}_g - \beta^2\nu_{2\delta}\overline{\gamma}_g - \nu_{2\delta}^2\overline{\gamma}_g + \beta^2\nu_{2\delta}^2\overline{\gamma}_g) \\
&+\frac{\partial^2}{\partial\delta^2}(-\delta^3\nu_{2\delta}\overline{\gamma}_g\gamma_g^2 - \delta^2\beta\nu_{2\delta}^2\gamma_g^2 - 2\delta^2\beta\nu_{2\delta}\overline{\gamma}_g\gamma_g \\
&+\delta^3\nu_{2\delta}^2\gamma_g^2\gamma_{g^2} + \delta^2\beta\nu_{2\delta}^2\gamma_{g^2} + 2\delta^2\beta\gamma_g^2)
\end{aligned}$$

and

$$\frac{\partial^2}{\partial\delta^2}v = (1 - \delta^2\gamma_g^2 - 2\delta\beta\gamma_g - \beta^2)(\frac{\partial^2}{\partial\delta^2}\nu_{4\delta})$$

$$
\begin{aligned}
&-4(\delta\gamma_g^2-\beta\gamma_g)(\frac{\partial}{\partial\delta}\nu_{4\delta})-2\gamma_g^2\nu_{4\delta}\\
&-2\nu_\delta(1-\delta^2\gamma_{g^2}-2\delta\beta\gamma_g-\beta^2)(\frac{\partial^2}{\partial\delta^2}\nu_{2\delta})\\
&-2(1-\delta^2\gamma_{g^2}-2\delta\beta\gamma_g-\beta^2)(\frac{\partial}{\partial\delta}\nu_{2\delta})^2\\
&+4\nu_\delta(\delta\gamma_{g^2}-\beta\gamma_g)(\frac{\partial}{\partial\delta}\nu_{2\delta})+2\nu_{2\delta}^2\gamma_{g^2}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\lim_{\delta\to 0}\frac{\partial}{\partial\delta}\nu_{2\delta} &= \lim_{\delta\to 0}(\frac{\partial}{\partial\delta}\mathsf{E}z_t^{2\delta})=\lim_{\delta\to 0}\int\frac{\partial}{\partial\delta}x^{2\delta}f(x)dx\\
&= \int\lim_{\delta\to 0}x^{2\delta}(\ln x^2)f(x)dx=\mathsf{E}(\ln z_t^2). \qquad (A.19)\\
\lim_{\delta\to 0}\frac{\partial^2}{\partial\delta^2}\nu_{2\delta} &= \lim_{\delta\to 0}(\frac{\partial^2}{\partial\delta^2}\mathsf{E}z_t^{2\delta})=\lim_{\delta\to 0}\int\frac{\partial^2}{\partial\delta^2}x^{2\delta}f(x)dx\\
&= \int\lim_{\delta\to 0}x^{2\delta}(\ln x^2)^2f(x)dx=\mathsf{E}(\ln z_t^2)^2. \qquad (A.20)\\
\lim_{\delta\to 0}\frac{\partial}{\partial\delta}\overline{\gamma}_g &= \lim_{\delta\to 0}\frac{\partial}{\partial\delta}\mathsf{E}(z_t^{2\delta}g(z_t))=\lim_{\delta\to 0}\int\frac{\partial}{\partial\delta}(x^{2\delta}g(x))f(x)dx\\
&= \int\lim_{\delta\to 0}(x^{2\delta}g(x)\ln x^2)f(x)dx=\mathsf{E}(g(z_t)\ln z_t^2). \qquad (A.21)
\end{aligned}
$$

Applying (A.19)- (A.21) gives $\lim_{\delta\to 0}\frac{\partial^2}{\partial\delta^2}u$ and $\lim_{\delta\to 0}\frac{\partial^2}{\partial\delta^2}v$, respectively. We see that $\rho_1(\delta)\to\rho_1^0$ as $\delta\to 0$. ■