財金論文叢刊
2005 年 10 月，第三期，1-28

Temporal Aggregation of a Strong PGARCH(1,1) Process

Meng-Feng Yen*

Department of Finance, Chaoyang University of Technology
Tel: 04-23323000 ext 4601
E-Mail: yenmf@mail.cyut.edu.tw

Abstract

Bollerslev's (1986) standard GARCH(1,1) model has been successful in the literature of volatility modelling and forecasting in the past two decades. Many of its extensions are contributed to examine the stylized features often observed with financial asset data. One of the distinct success is Bollerslev and Ghysels' (1996) periodic GARCH model, which takes into account periodic variation in the volatility of the underlying process. However, Drost and Nijman (1993) find that the conventional GARCH formulation works for only one sampling interval arbitrarily decided for the data in hand. This formulation does not apply to any other time intervals due to the assumption of an i.i.d. probability assumption for the underlying data. One of the problems caused by this will be that we cannot use ML method to estimate the GARCH model if the model is not for the original data set, but rather, for its temporally aggregated or dis-aggregated counterpart. Dorst and Nijman (1993) introduce the so-called weak GARCH formulation to tackle this problem and find this form of GARCH models apply to all sampling intervals for any given set of data. However, the ML method does not apply to the weak form of GARCH models since this formulation does not assume any probability distribution for the underlying standardised innovations. They thus propose a set of formulae to map the parameters of a weak GARCH(1,1) process sampled at one time interval to those of the same process but sampled at any other time interval.

However, there is hitherto no analytical results for a weak PGARCH process. It is the main purpose of the paper to investigate the relationship amongst the parameters of a weak GARCH process before and after temporal aggregation. Our simulation results tend to suggest that a two-stage PGARCH process will aggregate into a weak GARCH process. Some analytical results about the aggregated process are introduced too.

Key words: strong- and weak-GARCH, temporal aggregation, Monte Carlo simulation

* The Authors gratefully acknowledge the helpful comments of two anonymous referees.

1. Introduction

The standard GARCH(1,1) model introduced by Bollerslev (1986) has witnessed its success in the literature of volatility modelling and forecasting. Voluminous studies have been committed to this model's application in empirical circumstances. Many of its extensions involve modelling stylised features of the financial time series. A pronounced one of them is intraday volatility pattern, or called periodicity in volatility as a general term. Failure to accommodate this feature into the standard GARCH(1,1) model will absolutely lead to a model mis-specification problem. Bollerslev and Ghysels (1996, hereafter BG), amongst others, introduce their periodic GARCH model (hereafter PGARCH) to tackle this problem. However, Drost and Nijman (1993, hereafter DN) find that the formulation of the standard GARCH(1,1) model proposed by Bollerslev (1986) is applicable to only one sampling frequency of any data. To be specific, if we assume Bollerslev (1986)'s standard GARCH(1,1) model for a set of hourly data, the aggregated daily or weekly data will no longer follow this model. As a result, we can not use (quasi-) maximum likelihood method to estimate the standard GARCH(1,1) model for the aggregated daily or weekly data. DN suggest that the reason why Bollerslev's GARCH formulation does not apply to all time spans is the assumption of an i.i.d. probability distribution for the standardised errors. They therefore suggest a weak form of the standard GARCH model and a set of formulae mapping the parameters of the model sampled at any two different frequencies.

1.1 Review of the Drost and Nijman Aggregation Theory

Based on the continuous-time diffusion limit of the GARCH(1,1) process developed by Nelson (1990), DN (1993) propose a theoretical framework for the temporal aggregation of weak GARCH processes. Assuming that there is an underlying GARCH(1,1) diffusion, they derive a set of formulae for parameter mapping between any two discrete GARCH(1,1) models sampled at different frequencies. For example, if one has a series of daily data and desires to know what a weekly GARCH(1,1) model looks like, a natural approach to achieving this is to simply aggregate the daily data into weekly data and estimate them directly. Alternatively, one may first estimate the daily data to get the daily GARCH(1,1) estimates. Substitution of the daily estimates into the DN aggregation formulae gives us the weekly GARCH(1,1) parameters.

Before introducing the DN temporal aggregation theory, it is necessary to understand the three types of GARCH(*p*,*q*) processes, one of which underlies the theory. The definitions of the three GARCH(*p*,*q*) processes are described as follows: Given that

$$B(L)h_t = \phi + \{A(L)-1\}\varepsilon_t^2 , \qquad (1a)$$

where

$A(L) = 1 + \sum_{i=1}^{q}\alpha_i L^i$, $B(L) = 1 - \sum_{i=1}^{p}\beta_i L^i$,

and $\sum_{i=1}^{p}\beta_i + \sum_{i=1}^{q}\alpha_i < 1$, (1b)

ε_t is said to be

1. a *strong GARCH(p,q)* process if $z_t = \varepsilon_t /(h_t)^{1/2}$ is i.i.d. with zero mean and unit variance;
2. a *semi-strong GARCH(p,q)* process if the conditional expectation of ε_t and ε_t^2 upon the lags of ε_t are equal to, respectively, 0 and h_t; and
3. a *weak GARCH(p,q)* process if the best linear predictor of ε_t and ε_t^2 in terms of 1, ε_{t-1}, ε_{t-2},…, ε_{t-1}^2, ε_{t-2}^2,…, are equal, respectively, to 0 and h_t.

Note that both strong and semi-strong GARCH processes also satisfy the definition of a weak GARCH process, whereas a strong GARCH process is also a semi-strong GARCH process. Of the three definitions above, only the weak GARCH process is closed under temporal aggregation. In other words, a weak GARCH process remains a weak GARCH process under temporal aggregation. On the contrary, (semi-) strong GARCH processes will no longer be (semi-) strong GARCH processes, but rather weak GARCH processes, under temporal aggregation. An implication of this is that the conventional treatment of volatility modelling by the (semi-) strong GARCH models is not appropriate, as these models will be valid at only one frequency. Unfortunately, the weak GARCH type of models is dealing with the best linear projection of the underlying process, rather than its conditional expectation that has been regarded as the subject under investigation in the literature of GARCH type of models. By definition, however, the best linear projection of any variable upon a given information set is equal in size to its conditional expectation upon the same information set. The weak GARCH class of models is thus still useful for modelling and forecasting conditional heteroscedasticity.

Within the framework of weak GARCH processes, DN derive a set of formulae for mapping the parameters of a GARCH(1,1) model sampled at any temporal frequency onto the parameters of another GARCH(1,1) model sampled at another temporal frequency. To illustrate, let $h_t = \phi^H + \alpha_1^H \varepsilon_{t-1}^2 + \beta_1^H h_{t-1}$ denote the best linear projection equation and k^H the unconditional kurtosis of the high-frequency weak GARCH(1,1) process ε_t, and let $h_\lambda = \phi_{(m)}^L + \alpha_{1(m)}^L \varepsilon_{\lambda-1}^2 + \beta_{1(m)}^L h_{\lambda-1}$ denote the best linear projection equation and $k_{(m)}^L$ the unconditional kurtosis of the aggregated low-frequency weak GARCH(1,1) process ε_λ, m being the number of high-frequency intervals that each low-frequency interval contains. Then,

$$\phi_{(m)}^L = f_1(m, \phi^H, \alpha_1^H, \beta_1^H), \tag{2a}$$
$$\beta_{1(m)}^L = f_2(m, \alpha_1^H, \beta_1^H, k^H), \tag{2b}$$
$$\alpha_{1(m)}^L = f_3(m, \alpha_1^H, \beta_1^H, \beta_{(m)}^L), \tag{2c}$$
$$k_{(m)}^L = f_4(m, \alpha_1^H, \beta_1^H, k^H), \tag{2d}$$

Refer to DN (1993) for the detailed formulae.

It is worth noting that k^H, the unconditional kurtosis of the high-frequency weak GARCH(1,1) process ε_t, appears explicitly in (2b) and (2d), and implicitly in (2c). The value of k^H is therefore required for the set of formulae to work, which does not apply to the temporal aggregation of ARMA models[1]. To ensure non-negativity for k^H, moreover, the sum of α_1^H and β_1^H must fall inside the unit circle and $1-(\alpha_1^H+\beta_1^H)^2$ must be larger than $(k_\xi^H - 1)(\alpha_1^H)^2$, where k_ξ^H denotes the kurtosis of the high-frequency standardised strong GARCH(1,1) process. Also note that the aggregation frequency m is also allowed to be smaller than 1, such a case constituting the dis-aggregation case. In other words, the implied parameters constitute a higher-frequency GARCH(1,1) model. Given a weak GARCH(1,1) model sampled at any frequency, therefore, one can derive any other weak GARCH(1,1) model sampled at another frequency without having to estimate it.

1.2 Intended contribution of the paper

Having reviewed DN's temporal aggregation theory for the standard GARCH model, it is one our main concerns in this paper to find out if there is any relationship between the parameters for a PGARCH(1,1) process before and after temporal aggregation. In particular, we will examine, via Monte Carlo simulations, what would happen if the standard GARCH(1,1) formulation is

[1] But the temporal aggregation of ARMA models requires the existence of second moments. For the results of the temporal aggregation of ARMA type of models, see e.g. Palm and Nijman (1984) and Nijman and Palm (1990a, b).

replaced by BG's PGARCH(1,1) model. Intuitively, if we aggregate a set of data given by a PGARCH(1,1) model into relatively lower-frequency intervals, the aggregation interval being equal to the lower-frequency interval in length, the volatility of the aggregated data should be free of periodicity. In other words, it is very likely that the aggregated data be characterised by the standard GARCH(1,1) formulation. The main thrust of this paper is therefore to uncover the volatility mechanism of a PGARCH process upon temporal aggregation. We adopt BG's PGARCH(1,1) model with a simple two-stage periodic variation in either the intercept or the alpha parameter. The remainder of the paper is organised as follows. Section 2 explains the simulation framework and section 3 discusses the simulation results. Section 4 concludes.

2. Monte Carlo Simulation Framework

Similar to the simulation framework in BG[2], the basic GARCH(1,1) model is characterised by Parameterisation 1 in Table 1 below. Parameterisation 1 is then varied to become Parameterisations 2 and 3 in Table 1 so as to show the shift in intercept (ϕ^H) or parameter α_1^H across the two stages of each periodic volatility cycle. Parameterisations 1 and 2 in Table 1 are used to mark the change in intercept (ϕ^H) of the GARCH(1,1) model, whereas Parameterisations 1 and 3 in Table 1 are employed to specify the shift in parameter α_1^H across the two stages of each volatility cycle.

Table 1: Parameterisations of the GARCH(1,1) model for the DGPs

		Parameterisations		
		1	2	3
ϕ^H		0.05	0.01	0.05
α_1^H		0.15	0.15	0.05
β_1^H		0.7	0.7	0.7
$\alpha_1^H + \beta_1^H$		0.85	0.85	0.75
$(\sigma^H)^2$		0.3333	0.0667	0.2
k_ξ^H	$\xi_t \sim t_5$	25	25	25
	$\xi_t \sim N(0,1)$	3	3	3

Notes: 1. k_ξ^H denotes the unconditional kurtosis of the standardised innovation ξ_t .

2. $(\sigma^H)^2$ denotes the unconditional variance of ε_t , implied by ϕ^H , α_1^H , and β_1^H . Namely, $(\sigma^H)^2 = \dfrac{\phi^H}{(1-\alpha_1^H-\beta_1^H)}$.

The actual models used in the DGPs for high-frequency observations are therefore given as follows.

2.1 DGPs for high-frequency data

[2] In BG''s (1996) simulations for a periodic change in parameter α_1^H of their PGARCH(1,1) model, they let α_{11}^H = 0.4666 and α_{12}^H = 0.0727, whilst ϕ^H and β_1^H are fixed at, respectively, 0.05 and 0.7.

High-Frequency observations *(in-sample size: 4960, out-of-sample size: 40)*

Conditional mean (zero mean):

$$y_t = \varepsilon_t{}^3 = \xi_t \sqrt{h_t}\ , \qquad \text{for } t = 1 \text{ to } 5100 \qquad (3)$$

where $\xi_t \sim$ *i.i.d.* $N(0,1)$[4] or t_5[5,6], h_t denotes the conditional variance of the innovation ε_t and follows a PGARCH(1,1) model given by (4a) and (4b) below.

DGP 1: *PGARCH(1,1) of two-stage periodicity in intercept*

$$h_t = \phi^H_{s(t)} + 0.15\varepsilon^2_{t-1} + 0.7h_{t-1} \qquad \text{for } t = 1 \text{ to } 5000, \text{ where} \qquad (4a)$$

$s(t)$ = 1 for odd t and 2 for even t, and

$\phi^H_1 = 0.05$

$\phi^H_2 = 0.01$.

DGP 2: *PGARCH(1,1) of two-stage periodicity in alpha*

$$h_t = 0.05 + \alpha^H_{1s(t)}\varepsilon^2_{t-1} + 0.7h_{t-1} \qquad \text{for } t = 1 \text{ to } 5000, \text{ where} \qquad (4b)$$

$s(t)$ = 1 for odd t and 2 for even t, and

$\alpha^H_{11} = 0.15$

$\alpha^H_{12} = 0.05$.

The parameters in (4a) and (4b) are selected in according to the common fact that the estimate of the lagged conditional variance dominates the estimate of the squared lagged return innovations for daily or intra-daily data. Moreover, these parameters satisfy the requirements for DN's aggregation formulae to work for both stages.

Given the high-frequency observations, we aggregate them into relatively low-frequency observations, the aggregation interval being equal to two consecutive high-frequency intervals in length to avoid any aliasing problem.

Low-Frequency Observations *(in-sample size: 2480, out-of-sample size: 20)*

$$y_\lambda = \sum y_t = \sum_{i=1}^{2} y_{i+2(\lambda-1)} = \sum_{i=1}^{2} \varepsilon_{i+2(\lambda-1)} \equiv \varepsilon_\lambda \qquad \text{for } \lambda = 1 \text{ to } 2500, \text{ and} \qquad (5)$$

$$h_\lambda = \sum h_t = \sum_{i=1}^{2} h_{i+2(\lambda-1)} \qquad \text{for } \lambda = 1 \text{ to } 2500, \qquad (6)$$

where subscripts λ and t refer, respectively, to the low-frequency time scale and the high-frequency time scale.

---

[3] Academic research has also been committed to periodicity in the conditional mean of financial asset returns. However, significant periodic patterns in the conditional mean of any particular financial asset would imply arbitrage opportunities, which would disappear as soon as they are perceived via intensive trading on that asset. In other words, a periodic pattern in the conditional mean would be very short-lived. We thus do not include this possibility in the mean equation here.

[4] A Gaussian disturbance is used to examine whether model mis-specification or fat-tailedness in the pre-filtered data is the major contributor to the size of the fourth moment of the filtered data.

[5] Similar to the definitions of the three types of GARCH processes described in Section 1, assuming an *i.i.d.* distribution for a PGARCH(1,1) DGP will cast it into the category of strong PGARCH processes.

[6] In order to take care of the fat-tails of the financial asset return distribution often observed in empirical studies, we also employ a *Student*'s t_5 as the driving disturbances.

According to DN (1993), strong GARCH processes will become weak GARCH processes under temporal aggregation in normal context. It is of interest to know whether a high-frequency GARCH process, characterised by a periodic pattern in the parameters, will also become a weak GARCH one upon temporal aggregation. We speculate that the answer is positive since the assumption of i.i.d. for the high-frequency strong GARCH(1,1) process might not hold under temporal aggregation.

Taking into account the choice of high-frequency driving disturbances among $N(0,1)$ and t_5 doubles the cases that are examined. Table 2 below summarises the settings of all four experiments.

Table 2: Specifications of the experiments

Case No	DGP No	Driving Disturbances	Case No	DGP No	Driving Disturbances
Case 1	DGP 1	$N(0,1)$	Case 3	DGP 1	t_5
Case 2	DGP 2	$N(0,1)$	Case 4	DGP 2	t_5

2.2 Model Specification and Forecasting Procedures

The standard GARCH(1,1) model assuming constant parameters is fitted to both the high-frequency and the aggregated low-frequency observations. It is specified as follows.

strong standard GARCH(1,1) for high-frequency observations:

$$y_t = \varepsilon_t = \xi_t \sigma_t , \tag{7a}$$

where $\xi_t \sim i.i.d.$ standardised t_{v_1} or $N(0,1)$,

$$\sigma_t^2 = \phi^H + \alpha_1^H \varepsilon_{t-1}^2 + \beta_1^H \sigma_{t-1}^2 \text{, and} \tag{7b}$$

superscript ^{H} stands for *high-frequency*.

strong standard GARCH(1,1) for aggregated observations:

For ease of reference, let the superscript ^{L} stand for *low-frequency*. Substitute ^{L} for superscript ^{H} in (7a) and (7b), and rewrite them as

$$y_\lambda = \varepsilon_\lambda = \xi_\lambda \sigma_\lambda {}^{7}, \tag{8a}$$

where $\xi_\lambda \sim i.i.d.$ standardised t_{v_2} or $N(0,1)$, and

$$\sigma_\lambda^2 = \phi_{(2)}^L + \alpha_{1(2)}^L \varepsilon_{\lambda-1}^2 + \beta_{1(2)}^L \sigma_{\lambda-1}^2 , \tag{8b}$$

where the parenthesised 2 in the subscripts denotes the number of high-frequency periods within each aggregation interval.

There are two points worth noting: i) in the case of t_5 disturbances, the two numbers of degrees of freedom v_1 (for the *HF* GARCH(1,1)) and v_2 (for the *LF* GARCH(1,1)) are also part of the parameters under estimation by the ML method; ii) both the *HF* and *LF* GARCH(1,1) models are strong GARCH models since they both require the assumption of *i.i.d.* for the return innovations.

[7] Having also conducted a separate set of experiments allowing for a constant term in (8a), however, we do not find the constant's estimate to be significantly different from zero. No subsequent results are different from those from assuming zero constant term either.

implied Aggregated Weak GARCH(1,1):

Thanks to the DN aggregation theory, the aggregated weak GARCH(1,1) model is given by

$$\zeta_{\lambda}^{2} = \phi_{(2)}^{DN} + \alpha_{1(2)}^{DN}\varepsilon_{\lambda-1}^{2} + \beta_{1(2)}^{DN}\zeta_{\lambda-1}^{2}, \quad (9)$$

where ζ_{λ}^{2} denotes the linear projection of the squared low-frequency innovation, $\varepsilon_{\lambda}^{2}$, on the Hilbert space spanned by $\{1,\ \varepsilon_{\lambda-1}, \varepsilon_{\lambda-2}, \ldots,\ \varepsilon_{\lambda-1}^{2},\ \varepsilon_{\lambda-2}^{2}, \ldots\}$.

2.3 Reported Summary Statistics

In Table 3 below, summary statistics will be reported including the ML estimates of the *HF* and *LF* GARCH(1,1) model parameters, the degrees of freedom of the driving disturbances (for the t_5 cases), the parameters of the DN implied low-frequency GARCH(1,1) model, the unconditional variance implied by the model parameter estimates, and the unconditional kurtosis.

The data are generated under each DGP to ensure that 5000 high-frequency (pre-aggregated) observations are available for each replication. The impact of initial values is allowed for by allowing the DGP to run for 200 observations before sampling. 10000[8] replications, i.e. N = 10000, are completed for each simulation.

The mis-specified strong standard GARCH(1,1) filter is estimated by the quasi-maximum likelihood[9] method using the BFGS algorithm, with the average of each of the parameters across the two stages in the PGARCH(1,1) DGP (for high-frequency estimation) and its corresponding DN implied parameters (for low-frequency estimation) as the initial values. WinRATS version 5.03 is used to perform the simulations and calculations on 4 PC's of Intel PIII 733K (and faster) CPUs.

3. Results

This section discusses our simulation results for the four cases in Table 2 above. But to help understand the impacts of volatility periodicity on the ML estimates of the mis-specified strong GARCH(1,1) model, we start with examination of the model estimates under standard no-periodicity conditions. In particular, we want to know how well the ML method estimates the parameters in such a context. Section 3.1 below reports and discusses the averaged biases of the ML estimates of the strong standard GARCH(1,1) filter for both high-frequency and aggregated (low-frequency) observations from their true values given no patterns of volatility periodicity in the DGP.

3.1 Biases of the ML Parameter Estimates of the Well-Specified Strong Standard GARCH(1,1) Filters under No-Periodicity Standard Circumstances

---

[8] We try 1000, 5000, and 10,000 replications for a few of the experiments, and suggest that 10,000 should be the minimum value to generate consistent results across different seed numbers for the *t* cases. Despite efficient codes and fast machines, however, the steps involved (data generation, ML estimation of the strong standard GARCH(1,1) filter for both the *HF* and *LF* observations, calculation of the aggregated weak GARCH(1,1) model, forecasting volatility and evaluation of them) for a sample size of 5000 takes a non-trivial amount of processing time given this number of replications.

[9] The likelihood function for a t_v distribution is formulated as follows:

$t-logl = ln\Gamma\frac{v+1}{2} - ln\Gamma\frac{v}{2} - \frac{ln(v-2)}{2} - \frac{lnh_t}{2} - \frac{v+1}{2}\cdot ln(1+\frac{\varepsilon_t^2}{(v-2)\cdot h_t})$, where $ln\Gamma(x)$ denotes the natural logarithm of $\Gamma(x)$, and v the number of degrees of freedom under estimation.

We use the strong standard GARCH(1,1) filter to estimate the in-sample 4960 high-frequency observations produced by a strong standard GARCH(1,1) DGP which is purged of periodic variation in its parameters. The same strong standard GARCH(1,1) filter is then estimated on the aggregated 2480 observations, the aggregation interval being two high-frequency observation periods in length. These estimations are replicated 10000 times for each of Parameterisations 1 to 3 under both $N(0,1)$ and t_5 driving disturbances. Taken together, there are six sets of Monte Carlo simulations that lead to the results summarised in Tables A.1 and A.2 in Appendix A, which present the true values and the ML estimates of the strong standard GARCH(1,1) filter for the high-frequency and the low-frequency observations respectively.

Before detailed discussion of results in Tables A.1 and A.2 in Appendix A, it is interesting to note that the average biases of the parameter estimates are generally smaller under t_5 disturbances than under $N(0,1)$ disturbances for both the high-frequency and aggregated low-frequency observations.

Discussions of Table A.1 in Appendix A: high-frequency estimates

It is obvious from Table A.1 that, despite significant average biases of the reported high-frequency estimates, they are quite close to their true values with both sets of driving disturbances. Very few exceptions appear on the estimates of beta and the unconditional kurtosis for the t_5 cases:

(a) β_1^H tends to be much more underestimated under Parameterisation 3 than under Parameterisation 1, irrespective of the driving disturbances. Since the only difference between Parameterisations 1 and 3 is the size of alpha, α_1^H being 0.15 under Parameterisation 1 and 0.05 under Parameterisation 1, it appears to suggest that beta's estimate is inclined to be more underestimated when the true value of alpha decreases. Since alpha is correctly estimated the apparent underestimation of beta under Parameterisation 3 causes the level of integratedness, $\alpha_1^H + \beta_1^H$, to be underestimated under Parameterisation 3 to an extent larger than under Parameterisation 1.

(b) The average non-parametric sample estimate for the unconditional kurtosis of the standardized high-frequency innovations, $\overline{\hat{k}_\xi^H}$, tends to be significantly underestimated under t_5 disturbances across all parameterisations. We doubt the possibility that estimation biases of the parameters of the standard *HF* GARCH(1,1) filter lead to the downward bias of $\overline{\hat{k}_\xi^H}$ in the t_5 cases. To see why it is not likely for the biases of the parameter estimates to be the reason, we compare the biases of the parameter estimates between under both disturbances in Table A.1. In general, the biases of $\overline{\hat{\phi}^H}$, $\overline{\hat{\alpha}_1^H}$, and $\overline{\hat{\beta}_1^H}$ in the case of t_5 disturbances are slightly smaller than their counterparts in the case of $N(0,1)$ disturbances. Under both disturbances, moreover, the biases are all practically trivial, although some of them are statistically significant. This finding rules out the possibility that estimation biases of the parameters cause the under-estimation of k_ξ^H under t_5 disturbances. As a matter of fact, the strong standard GARCH(1,1) filter for the high-frequency observations is correctly estimated under both disturbances. It is therefore likely that the unknown statistical property of the non-parametric sample estimate of unconditional kurtosis, i.e. $\hat{k}_\xi^H = [\frac{1}{T}\sum_{t=1}^{T}(\hat{\xi}_t)^4] / [\frac{1}{T}\sum_{t=1}^{T}(\hat{\xi}_t - \overline{\hat{\xi}_t})^2]^2$, is responsible for the under-estimation of k_ξ^H in the t_5 cases even though the strong standard GARCH(1,1) filter for the high-frequency observations is correctly specified and estimated. In other words, it should be part of the unknown statistical properties of

$\hat{k}_{\xi}^{H} = [\frac{1}{T}\sum_{t=1}^{T}(\hat{\xi}_t)^4\,] / [\frac{1}{T}\sum_{t=1}^{T}(\hat{\xi}_t - \overline{\hat{\xi}_t})^2]^2$ that the larger the fourth moment of the driving disturbances, the more is its non-parametric sample estimate smaller than its true value. However, verification of the above argument relies on the statistical property of $[\frac{1}{T}\sum_{t=1}^{T}(\hat{\xi}_t)^4\,] / [\frac{1}{T}\sum_{t=1}^{T}(\hat{\xi}_t - \overline{\hat{\xi}_t})^2]^2$ which is not available.

However, if we look at the estimate of the number of degrees of freedom for the high-frequency standardised innovations in the cases of t_5 disturbances, i.e. $\overline{\hat{v}_1}$, it is very close to its true value, 5, across all parameterisations despite the statistically significant but practically trivial positive biases. Given that the unconditional kurtosis of a standardised t distribution with v degrees of freedom is equal to $\frac{3(v-2)}{v-4}$, $v = 5$ corresponds to a kurtosis of 9. Therefore, $\overline{\hat{v}_1} \approx 5$ in Table A.1 implies $\overline{\hat{k}_{\xi}^{H}} \approx 9$. For example, $\overline{\hat{v}_1}$ is 5.04 under all three parameterisations in Table A.1, which implies $\overline{\hat{k}_{\xi}^{H}} = \frac{3(5.04-2)}{(5.04-4)} = 8.77$, much closer to its true value, 9, than is the non-parametric sample estimate of k_{ξ}^{H}.

In addition to the two points above, there are some other aspects in Table A.1 worth noting:

(i) The level of integratedness, $\alpha_1^H + \beta_1^H$, is well estimated under Parameterisation 1 and 2 under both disturbances. Under Parameterisation 3, although $\alpha_1^H + \beta_1^H$ is significantly underestimated, the downward biases under both types of disturbances are only as small as three hundredths of the true value in size. In general, therefore, the ML method is able to correctly estimate the level of integratedness of a strong GARCH(1,1) process. This result will be used as a benchmark, against which the distortion of the estimate of $\alpha_1^H + \beta_1^H$ in the context of volatility periodicity will be compared.

(ii) The asymptotic innovation variance implied by the parameter estimates, $\overline{(\hat{\sigma}^H)^2} = \overline{\frac{\hat{\phi}^H}{(1-\hat{\alpha}_1^H - \hat{\beta}_1^H)}}$, is found to be almost equal to its true value across all three parameterisations and the two disturbances. This result signals one of the advantages of using the ML method in GARCH model estimation. But note that this might not be the case when the true t-likelihood is replaced by the Gaussian likelihood under t_5 DGPs.

(iii) Finally, the fourth moment of the driving disturbances of the DGP does not seem to affect the fact that the high-frequency parameters of the GARCH(1,1) filter are accurately estimated by the ML method. In fact, the extent to which $\overline{\hat{\phi}^H}$, $\overline{\hat{\alpha}_1^H}$, $\overline{\hat{\beta}_1^H}$, $\overline{\hat{\alpha}_1^H + \hat{\beta}_1^H}$, and $\overline{(\hat{\sigma}^H)^2}$ are close to their true values is virtually the same across both driving disturbances, as reported in Table A.1. This result might be due to the use of the actual likelihood for both disturbances in the ML method. However, this might not be the case if we have applied the quasi-likelihood method to the t_5 cases. In other words, if we assume a Gaussian likelihood for the t_5 cases, the parameter estimates might not be as accurate as those under the use of the correct t likelihood for the t_5 cases.

Discussions of Table A.2 in Appendix A: low-frequency estimates

(i) The results in Table A.2 tend to suggest that, despite the significant average biases, most of the low-frequency parameter estimates are close to their true values. These accurate estimates justify the usefulness of ML method in estimating the GARCH parameters. However, there appear to be some exceptions arising under Parameterisation 3 for both disturbances, where $\overline{\hat{\phi}^L}$ is obviously upwardly biased from its true value and $\overline{\hat{\beta}_1^L}$ (and thus $\overline{\hat{\alpha}_1^L}+\overline{\hat{\beta}_1^L}$) downwardly biased from its true value. These biases could be due to the use of relatively smaller α_1^H under Parameterisation 3, 0.05, compared to 0.15 under Parameterisation 1 and 2. The theoretical value of α_1^L is thus smaller under Parameterisation 3, i.e. 0.038 under $N(0,1)$ and 0.057 under t_5. These two values are dwarfed, respectively, by 0.126 under $N(0,1)$, and 0.173 under t_5 for both Parameterisation 1 and 2. Since this fact that a larger downward bias of the beta estimate relates to the use of a smaller alpha in the GARCH(1,1) DGP is also observed in Table A.1 for the high-frequency estimates, we speculate that a smaller alpha in the GARCH(1,1) DGP will render the ML estimation less reliable. One possible explanation for this suggests that it is more difficult for the ML to distinguish the long-term variation form the short-term variation in the conditional innovation variances when the true alpha of the GARCH(1,1) DGP appears to be smaller.

(ii) Turning to the obvious upward bias of $\overline{\hat{\phi}^L}$ from its true values under Parameterisation 3 for both disturbances, this result should be discussed along with the estimate of the asymptotic low-frequency innovation variance, i.e. $\overline{(\hat{\sigma}^L)^2}$. It can be seen from Table A.2 that $\overline{(\hat{\sigma}^L)^2}$ is almost equal to its true value irrespective of the parameterisation and disturbance distribution. The accurate estimation of $(\sigma^L)^2$ justifies the upward bias of $\overline{\hat{\phi}^L}$ under Parameterisation 3, which offsets the downward bias of $\overline{\hat{\beta}_1^L}$. It seems to say that the ML method is always able to capture the level of the asymptotic innovation variance of a strong GARCH(1,1) process despite biased parameter estimates. The correct estimation of the asymptotic innovation variance by the ML is also reflected by the fact that the ratio of $\overline{(\hat{\sigma}^L)^2}$ to $\overline{(\hat{\sigma}^H)^2}$ being about 2 to 1 across all parameterisations with both disturbances corresponds exactly to the ratio of $(\sigma^L)^2$ to $(\sigma^H)^2$, and the aggregation frequency $m = 2$.

(iii) It is also interesting to note that the estimate for the level of integratedness of the *LF* GARCH(1,1) model, $\overline{\hat{\alpha}_1^L}+\overline{\hat{\beta}_1^L}$, is obviously much more underestimated relative to those for the *HF* GARCH(1,1) model. This could be caused by the problem of model mis-specification. In particular, the aggregated low-frequency observations are no longer a strong GARCH(1,1) process, but rather a weak GARCH(1,1) process. It is thus not appropriate to estimate the low-frequency observations by a strong GARCH model. However, there are hitherto no approaches to estimating a weak GARCH model. One can only use the ML method to give the 'best sample-based' parameterisation of the linear projection of a weak GARCH process. Under such a condition of model mis-specification for the low-frequency observations, it is not surprising to see the larger biases of the parameters of the *LF* strong GARCH(1,1) model than those of the *HF* GARCH(1,1) model.

(iv) Finally, it is worth noting that the value of $\overline{\hat{v}_2}$ seems to be in inverse proportion to the size of the true α_1^H. In particular, $\overline{\hat{v}_2}$ is 5.254 under both Parameterisations 1 and 2, where α_1^H is set equal to 0.15. Under Parameterisation 3, where α_1^H = 0.05, the value of $\overline{\hat{v}_2}$ goes up to 6.475. Despite the lack of theoretical value for v_2, $\hat{v}_2$ given by the strong standard GARCH(1,1) filter might be close to the unknown true v_2, since the same filter successfully estimate its high-frequency counterpart v_1. Given the formula for the unconditional kurtosis of a standardised t distribution with v degrees of freedom above, the smaller the true alpha of a high-frequency GARCH(1,1) process, other things equal, the smaller will be the fourth moment of the aggregated process standardised by the aggregated GARCH(1,1) model.

Having analysed the estimation results of the standard GARCH(1,1) model in the no-periodicity benchmark condition, the next section will discuss these same issues but in the context of model mis-specification. In other words, we will explore the effects of periodicity in the parameters of the GARCH(1,1) DGP on the estimation of both *HF* and *LF* strong standard GARCH(1,1) models.

3.2 Effects of Volatility Periodicity on Model Estimation of the Mis-Specified GARCH(1,1) Filters

To see how the volatility periodicity affects the ML parameter estimates of the mis-specified strong standard GARCH(1,1) models, we devise two types of volatility periodicities through DGPs 1 and 2. The details of both DGPs and the GARCH models used for the conditional variance estimation have been explained in Section 2. Discussions of the simulation results reported in Table A.3 in Appendix A will be categorised according to DGPs 1 and 2.

Since these two DGPs are based upon a two-stage PGARCH(1,1) model, it might be informative to examine the true unconditional variance of each stage of a two-stage PGARCH(1,1) model in (4a) and (4b). In particular, let $h_t = \phi^H_{s(t)} + \alpha^H_{1s(t)}\varepsilon^2_{t-1} + \beta^H_{1s(t)}h_{t-1}$ denote a high-frequency two-stage PGARCH(1,1) model, where $s(t)$ = 1 for odd t and 2 for even t. We have shown in Appendix B that the high-frequency unconditional innovation variance is given by

$$h^H_{odd} = \frac{\phi^H_1 + \phi^H_2(\alpha^H_{11} + \beta^H_{11})}{1-(\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})} \text{ for odd } t\text{, and} \quad (10a)$$

$$h^H_{even} = \frac{\phi^H_2 + \phi^H_1(\alpha^H_{12} + \beta^H_{12})}{1-(\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})} \text{ for even } t. \quad (10b)$$

If we aggregate the high-frequency PGARCH(1,1) process given by $h_t = \phi^H_{s(t)} + \alpha^H_{1s(t)}\varepsilon^2_{t-1} + \beta^H_{1s(t)}h_{t-1}$ into a low-frequency process, the aggregation interval being two high-frequency observation periods in length. Under the assumption of *i.i.d.* high-frequency driving disturbances, the unconditional variance for the aggregated innovations, ε_λ, is simply the sum of h^H_{odd} and h^H_{even}. That is,

$$h_\lambda = h^H_{odd} + h^H_{even} = \frac{\phi^H_1(1+\alpha^H_{12} + \beta^H_{12}) + \phi^H_2(1+\alpha^H_{11} + \beta^H_{11})}{1-(\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})}. \quad (11)$$

In passing, note that the aggregated low-frequency process, ε_λ, is covariance stationary as long as $0 \le (\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12}) < 1$. We refer to BG (1996) for proof. An implication of this argument is that either $\alpha^H_{11} + \beta^H_{11}$ (the level of integratedness for stage one), or $\alpha^H_{12} + \beta^H_{12}$ (the level of integratedness for stage two) can be larger than 1 as long as their product falls between 0 and 1.

Despite the lack of analytical results[10] for the parameters of a PGARCH(1,1) process under temporal aggregation, the formula for h_λ in (11) tends to suggest that:

(a) *the aggregated process is a weak GARCH(1,1) process*, (12a)

(b) following (a), *the aggregated intercept* $\phi^L_{(2)}$ *is equal to* $\phi^H_1(1+\alpha^H_{12} + \beta^H_{12}) + \phi^H_2(1+\alpha^H_{11} + \beta^H_{11})$, (12b)

and

(c) following (a),

[10] The DN aggregation theory does not apply to a periodic-GARCH (or PGARCH) process, albeit it is possible to extend the theory to a PGARCH version.

the level of integratedness $\alpha^{L}_{1(2)} + \beta^{L}_{1(2)}$ *of the aggregated low-frequency weak GARCH(1,1) process is equal to* $(\alpha^{H}_{11} + \beta^{H}_{11})(\alpha^{H}_{12} + \beta^{H}_{12})$. (12c)

Argument (12a) above might be a straightforward result of the fact that we set the length of the aggregation interval equal to the length of each periodic cycle of the high-frequency PGARCH(1,1) DGPs. That is, the aggregated observations do not show any periodicity in the parameters of the GARCH specification governing their conditional innovation variances. We will examine the validity of arguments (12b) and (12c) above via our simulation results for DGPs 1 and 2 as reported in Table A.3 below.

Discussions of Table A.3 in Appendix A:

Results of the cases for DGPs 1 and 2 in Table A.3 tend to suggest that the Monte Carlo standard deviations of the high-frequency parameter estimates and the implied low-frequency parameters are obviously larger than those of the low-frequency parameter estimates. For example, the standard deviation for $\overline{\hat{\phi}^{H}}$ is 0.005, and 0.016 for $\overline{\phi^{DN}_{(2)}}$, whereas the standard deviation for $\overline{\hat{\phi}^{L}_{(2)}}$ is 0.035 in case 1. This finding might be explained by the number of observations used for the model estimation. In particular, 4960 observations are used for the in-sample estimation of the *HF* standard GARCH(1,1) filter, contrasting with the 2480 observations used for the in-sample estimation of the *LF* standard GARCH(1,1) filter. Intuitively, the more data are estimated by ML in each replication, the smaller will be the variations of the model parameter estimates across replications. Since the parameters of the aggregated low-frequency weak GARCH(1,1) model are implied by the parameter estimates of the *HF* strong GARCH(1,1) model which are based on the 4960 observations, it is not surprising to see the smaller standard deviation of $\overline{\phi^{DN}_{(2)}}$ than of $\overline{\hat{\phi}^{L}_{(2)}}$. Similar results are found in cases 2, 3, and 4, and for the estimates of alpha, beta, number of degrees of freedom, unconditional innovation variance implied by the parameter estimates.

a. Impacts on the Average ML Estimated Intercepts of the Strong Standard GARCH(1,1) Model for both* HF *and* LF *Observations

(i) $\overline{\hat{\phi}^{H}}$: In cases 1 and 3, where DGP 1 in (4a) is employed to generate a two-stage periodicity in the intercept ϕ^{H}, the value of $\overline{\hat{\phi}^{H}}$ of the mis-specified *HF* GARCH(1,1) model is found to be 0.3, a value equal to the average of the true ϕ^{H} of the two stages, 0.05 and 0.01. This result is indicative that the ML method tends to take equal account of the high intercept (and thus the high unconditional innovation variance) in stage 1 and the low intercept (and thus the low unconditional innovation variance) in stage 2, when maximising the likelihood. Since this fact is observed in cases of both 1 (N(0,1) disturbances) and 3 (t_5 disturbances), it suggests that the size of the fourth moment of the driving disturbances does not seem to play a role in the ML estimation of the intercept of the mis-specified *HF* GARCH(1,1) model.

Turning to cases 2 and 4, where DGP 2 in (4b) is used to generate two-stage periodicity in the parameter alpha α^{H}_{1}, the estimated $\overline{\hat{\phi}^{H}}$ is found to be 0.052 in both cases. This estimate is very close to its true values across the two stages, i.e. $\phi^{H}_{1} = \phi^{H}_{2} = \phi^{H} = 0.05$. It appears to suggest that, provided that the intercept is not mis-specified in the standard GARCH(1,1) model, the ML method provides accurate intercept estimate in that mis-specification in the alpha parameter does not cause to bias of the intercept estimate.

Also, since the PGARCH(1,1) DGPs in (4a) and (4b) consist of Parameterisations 1 and 2 or of Parameterisations 1 and 3 in Table 1, it might be interesting to compare the values of $\overline{\hat{\phi}^H}$ in Panel A of Table A.3 to its counterparts in Table A.1 in Appendix A. For example, DGP 1 used in cases 1 and 3 consists of Parameterisations 1 (for stage 1) and 2 (for stage 2) in Table 1 above, we expect to observe the intercept estimate, $\overline{\hat{\phi}^H}$, of the mis-specified GARCH(1,1) filter to be the average of $\overline{\hat{\phi}^H}$ under Parameterisation 1, 0.051, and $\overline{\hat{\phi}^H}$ under Parameterisation 2, 0.0102. Indeed, $\overline{\hat{\phi}^H}$ is 0.03 in case 1, virtually equal to the average of 0.051 and 0.0102. Similar results can be observed with cases 2, 3, and 4. In line with the argument above, it tends to indicate that the ML method values equally the sizes of the true intercept of each stage of the PGARCH(1,1) DGP when estimating the mis-specified standard GARCH(1,1) filter.

(ii) $\overline{\hat{\phi}^L_{(2)}}$: We have argued above that the aggregated observations might follow a standard weak GARCH(1,1) process. In this context, the standard *LF* strong GARCH(1,1) model might not be regarded mis-specified although the ML method is not tailored for estimating a weak GARCH process. However, DN have documented that the ML estimates of a weak GARCH process are not biased from their true values to a significant extent. As a result, the values of $\overline{\hat{\phi}^L_{(2)}}$, $\overline{\hat{\alpha}^L_{1(2)}}$, and $\overline{\hat{\beta}^L_{1(2)}}$ are expected to be close to their true values.

In cases 1 and 3 using DGP 1 in (4a), where $\phi_1^H = 0.05$, $\phi_2^H = 0.01$, $\alpha_{11}^H = \alpha_{12}^H = 0.15$, $\beta_{11}^H = \beta_{12}^H = 0.7$, the formula in (11) gives

$$h^L_{(2)} = \frac{\phi_1^H(1+\alpha_{12}^H+\beta_{12}^H)+\phi_2^H(1+\alpha_{11}^H+\beta_{11}^H)}{1-(\alpha_{11}^H+\beta_{11}^H)(\alpha_{12}^H+\beta_{12}^H)} = \frac{0.05(1+0.15+0.7)+0.01(1+0.15+0.7)}{1-(0.15+0.7)(0.15+0.7)}$$

$$= \frac{0.111}{1-0.7225} = 0.4.$$

The vicinity of the numerator, 0.111, to the estimated $\overline{\hat{\phi}^L_{(2)}}$, 0.117, in cases 1 and 3 reported in Panel B of Table A.3 justifies our the argument (12b). That is, the true $\phi^L_{(2)}$ of the aggregated weak GARCH(1,1) process might well be equal to $\phi_1^H(1+\alpha_{12}^H+\beta_{12}^H)+\phi_2^H(1+\alpha_{11}^H+\beta_{11}^H)$.

To verify the argument in (12b) in the context of cases 2 and 4 using DGP 2 in (4b), where $\phi_1^H = \phi_2^H = 0.05$, $\alpha_{11}^H = 0.15$, $\alpha_{12}^H = 0.05$, and $\beta_{11}^H = \beta_{12}^H = 0.7$. Formula (11) therefore gives h_λ $= \frac{0.05(1+0.05+0.7)+0.05(1+0.15+0.7)}{1-(0.15+0.7)(0.05+0.7)} = \frac{0.18}{1-0.6375} = 0.49655$. Referring to Panel B of Table A.3, the numerator 0.18 above is also close to the estimated $\overline{\hat{\phi}^L_{(2)}}$, 0.195 (case 3) and 0.188 (case 4). As a result, the argument in (12b) is more than likely to be true.

(iii) $\overline{\phi^{DN}_{(2)}}$: The average aggregated intercept implied by $\overline{\hat{\phi}^H}$ of the mis-specified standard *HF* strong GARCH(1,1) model is seen to be close to $\overline{\hat{\phi}^L_{(2)}}$ in all four cases discussed above, i.e. cases 1, 2, 3, and 4: see Panel B of Table A.3 in Appendix A. Table 3 below compares the values of $\overline{\phi^{DN}_{(2)}}$ and $\overline{\hat{\phi}^L_{(2)}}$ to those of $\phi_1^H(1+\alpha_{12}^H+\beta_{12}^H)+\phi_2^H(1+\alpha_{11}^H+\beta_{11}^H)$, denoted $\phi^L_{(2)}$ hereafter, in the four cases:

Table 3: Comparison of the values of $\overline{\hat{\phi}^L_{(2)}}$ and $\overline{\phi^{DN}_{(2)}}$ to $\phi^L_{(2)}$

Case No	(DGP No, Disturbance)	$\overline{\hat{\phi}^L_{(2)}}$	$\overline{\phi^{DN}_{(2)}}$	$\phi^L_{(2)} = \phi_1^H(1+\alpha_{12}^H+\beta_{12}^H)+ \phi_2^H(1+\alpha_{11}^H+\beta_{11}^H)$

1	(DGP 1, $N(0,1)$)	0.117 (0.035)	0.112 (0.016)	0.111
2	(DGP 2, $N(0,1)$)	0.195 (0.082)	0.186 (0.036)	0.18
3	(DGP 1, t_5)	0.117 (0.027)	0.112 (0.016)	0.111
4	(DGP 2, t_5)	0.188 (0.059)	0.185 (0.035)	0.18

Obviously, the values of $\overline{\phi_{(2)}^{DN}}$ in all four cases above tend to be close to the values of $\phi_{(2)}^{L}$ more than are the values of $\overline{\hat{\phi}_{(2)}^{L}}$. The implication of this result will be discussed when we move to the results of the estimated level of integratedness of the *HF* GARCH(1,1) filter.

b. Impacts on the Average ML Estimated Dynamics of the Strong Standard GARCH(1,1) Model for both* HF *and* LF *Observations

(i) $\overline{\hat{\alpha}_1^H}$ (reported in Panel A) and $\overline{\hat{\beta}_1^H}$ (reported in Panel B): By similar argument to the discussions of $\overline{\hat{\phi}_1^H}$, moreover, we expect the values of $\overline{\hat{\alpha}_1^H}$ (or $\overline{\hat{\beta}_1^H}$) to be equal to the average of the two $\overline{\hat{\alpha}_1^H}$ (or $\overline{\hat{\beta}_1^H}$) under Parameterisation 1 and 2 (for cases 1 and 3) or 1 and 3 (for cases 2 and 4) in Table A.1 in appendix A. In particular, the value of $\overline{\hat{\alpha}_1^H}$ is 0.149 in case 1, which is almost the average of the values of $\overline{\hat{\alpha}_1^H}$, 0.15 under Parameterisation 1 and 0.15 under Parameterisation 2. In cases 2, 3, and 4, this is also true.

Similar results are observed with $\overline{\hat{\beta}_1^H}$. For instance, $\overline{\hat{\beta}_1^H}$ is 0.692 in case 4 (under DGP 2, made up of Parameterisation 1 and 3 with t_5 disturbances) as reported in Panel A of Table A.3, which is very close to the average of 0.698 under Parameterisation 1 and 0.681 under Parameterisation 3 in Table A.1. The values of $\overline{\hat{\beta}_1^H}$ in cases 1, 2, and 3 are also equally weighted mixtures of its values under Parameterisation 1 and 2 or 1 and 3, depending on which case is discussed.

The values of $\overline{\hat{\alpha}_1^H}$ and $\overline{\hat{\beta}_1^H}$ can also be discussed from another perspective by comparing them to their true values. In particular, in cases 1 and 3 using DGP 1, where $\phi_1^H = 0.05$, $\phi_2^H = 0.01$, $\alpha_{11}^H = \alpha_{12}^H = 0.15$, $\beta_{11}^H = \beta_{12}^H = 0.7$, the value of $\overline{\hat{\alpha}_1^H}$ is 0.149 (in case 1) and 0.15 (in case 3), which is almost equal to or equal to the true α_1^H ($= \alpha_{11}^H = \alpha_{12}^H = 0.15$). Moreover, the value of $\overline{\hat{\beta}_1^H}$ is 0.698 (in case 1) and 0.699 (in case 3), both being almost equal to the true β_1^H ($=\beta_{11}^H = \beta_{12}^H = 0.7$). Taken together, periodicity in the intercept of a PGARCH(1,1) process does not seem to deviate the dynamics estimates of the mis-specified strong standard GARCH(1,1) model from their true values. This result might well be due to the nature of the model mis-specification, that is: the strong standard GARCH(1,1) filter used to estimate a PGARCH(1,1) process of a two-stage periodicity in the intercept fails to correctly specify the intercept of the PGARCH(1,1) process only, but not its dynamics. It is known that the intercept decides only the level of the unconditional variance of a GARCH process, whereas the dynamics parameters characterise how the conditional variance of a GARCH process changes through time. As a result, it is not surprising to observe that the dynamics estimates of the standard GARCH(1,1) filter is the same as their true values despite that the filter does not capture the variation in the intercept of a PGARCH(1,1) DGP.

Turning to cases 2 and 4, where DGP 2 ($\phi_1^H = \phi_2^H = 0.05$, $\alpha_{11}^H = 0.15$, $\alpha_{12}^H = 0.05$, and $\beta_{11}^H = \beta_{12}^H = 0.7$) is used, the value of $\overline{\hat{\alpha}_1^H}$ is 0.1 (in case 2) and 0.098 (in case 4), which is equal to the average of α_{11}^H and α_{12}^H. This result tends to suggest that the ML method values equally the sizes of α_{11}^H and α_{12}^H when maximising the likelihood of the mis-specified strong standard GARCH(1,1) innovations. Moreover, the value of $\overline{\hat{\beta}_1^H}$ is 0.69 (in case 2) and 0.692 (in case 4), which are both close to their true value β_1^H (=0.7). It indicates that periodicity in the parameter alpha of a PGARCH(1,1) DGP does not cause a deviation in the beta estimate of the mis-specified GARCH(1,1) filter from its true value.

Another point worth noting is that the fourth moment of the driving disturbances of the underlying PGARCH(1,1) DGP does not seem to affect the ML estimation of the dynamics of the mis-specified standard GARCH(1,1) filter. To understand this, we can look at the values of $\overline{\hat{\alpha}_1^H}$ and $\overline{\hat{\beta}_1^H}$ in cases 1 (N(0,1) disturbances) and 3 (t_5 disturbances), for example. Both $\overline{\hat{\alpha}_1^H}$ and $\overline{\hat{\beta}_1^H}$ are almost unchanged across the two cases. This is also observed in cases 2 (N(0,1) disturbances) and 4 (t_5 disturbances).

(ii) $\overline{\hat{\alpha}_{1(2)}^L}$: $\overline{\hat{\alpha}_{1(2)}^L}$ is 0.124 in case 1 (N(0,1)), smaller than the high-frequency true α_1^H, 0.15, whereas $\overline{\hat{\alpha}_{1(2)}^L}$ is 0.161, larger than 0.15 in case 3 (t_5). Similarly, $\overline{\hat{\alpha}_{1(2)}^L}$ is 0.085 in case 2 (N(0,1)), smaller than the average of α_{11}^H=0.15 and α_{12}^H=0.05, 0.1, whereas $\overline{\hat{\alpha}_{1(2)}^L}$ is 0.11 in case 4 (t_5), larger than 0.1. The value of $\overline{\hat{\alpha}_{1(2)}^L}$ is obviously dependent on the fourth moment of the driving disturbances, which coincides with the DN aggregation formulae for the standard weak GARCH(1,1) process. It indicates that any development of the analytical results for the temporal aggregation of a weak PGARCH(1,1) process should also take into account the unconditional kurtosis of the driving disturbances.

(iii) $\overline{\hat{\alpha}_{1(2)}^L}$ vs. $\overline{\alpha_{1(2)}^{DN}}$ and $\overline{\hat{\beta}_{1(2)}^L}$ vs. $\overline{\beta_{1(2)}^{DN}}$: From Table A.3, the values of $\overline{\hat{\alpha}_{1(2)}^L}$ and $\overline{\hat{\beta}_{1(2)}^L}$ are close to the values of $\overline{\alpha_{1(2)}^{DN}}$ and $\overline{\beta_{1(2)}^{DN}}$ respectively under both DGPs 1 and 2 (cases 1, 2, 3, and 4). The differences between $\overline{\hat{\alpha}_{1(2)}^L}$ and $\overline{\alpha_{1(2)}^{DN}}$, and between $\overline{\hat{\beta}_{1(2)}^L}$ and $\overline{\beta_{1(2)}^{DN}}$ are not larger than those observed in the standard no-periodicity experiments discussed in Section 6.3.1: see discussions of Table A.2. Given the argument in (12a) that the aggregated observations are characterised by a standard weak GARCH(1,1) process, these differences should be attributed only to the errors caused by the application of the ML method to estimating a weak GARCH(1,1) process.

c. *Impacts on the Average ML Estimated levels of Integratedness of the Strong Standard GARCH(1,1) Model for both* HF *and* LF *Observations:*

(i) $\overline{\hat{\alpha}_1^H}+\overline{\hat{\beta}_1^H}$: the value of $\overline{\hat{\alpha}_1^H}+\overline{\hat{\beta}_1^H}$ is expected to be equal to the average of the values of $\overline{\hat{\alpha}_1^H}+\overline{\hat{\beta}_1^H}$ under Parameterisations 1 and 2 (for cases 1 and 3) or Parameterisations 1 and 3 (for cases 2 and 4). In fact, $\overline{\hat{\alpha}_1^H}+\overline{\hat{\beta}_1^H}$ is 0.790 in case 2 reported in Panel A of Table A.3, whereas $\overline{\hat{\alpha}_1^H}+\overline{\hat{\beta}_1^H}$ is 0.847 under Parameterisation 1 and 0.726 under Parameterisation 3 in Table A.1 with the N(0,1) disturbances. 0.79 is roughly the average of 0.847 and 0.726, which echoes our expectation of the value of $\overline{\hat{\alpha}_1^H}+\overline{\hat{\beta}_1^H}$ in case 2. For cases 1, 3, and 4, the same results are observed. It tends to suggest that periodicity in the intercept or parameter alpha of the underlying two-stage PGARCH(1,1) DGP does not cause the ML estimates of the levels of integratedness of the mis-specified strong standard GARCH(1,1) filter to deviate from the average of their true values of the two stages.

(ii) $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$: To discuss the values of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ in the four cases, we need to refer to the argument in (12c), that is, the true level of integratedness $\alpha^L_{1(2)} + \beta^L_{1(2)}$ of the aggregated low-frequency weak GARCH(1,1) process is equal to $(\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})$. To facilitate our discussions, we compare the values of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ and $\overline{\alpha^{DN}_{1(2)}} + \overline{\beta^{DN}_{1(2)}}$ to the values of $(\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})$ (termed $\alpha^L_{1(2)} + \beta^L_{1(2)}$, hereafter) in Table 4 below.

Table 4: Comparison of the values of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ ***and*** $\overline{\alpha^{DN}_{1(2)}} + \overline{\beta^{DN}_{1(2)}}$ ***to*** $\alpha^L_{1(2)} + \beta^L_{1(2)}$

Case No	**(DGP No, Disturbance)**	$\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$	$\overline{\alpha^{DN}_{1(2)}} + \overline{\beta^{DN}_{1(2)}}$	$\alpha^L_{1(2)} + \beta^L_{1(2)} = (\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})$
1	(DGP 1, $N(0,1)$)	0.707 (0.089)	0.719 (0.042)	0.7225
2	(DGP 2, $N(0,1)$)	0.606 (0.167)	0.626 (0.072)	0.6375
3	(DGP 1, t_5)	0.705 (0.069)	0.721 (0.042)	0.7225
4	(DGP 2, t_5)	0.619 (0.121)	0.627 (0.070)	0.6375

Note: The sample standard deviations for N=10000 estimates are reported in parentheses.

Despite the fact that the values of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ are slightly smaller than $\alpha^L_{1(2)} + \beta^L_{1(2)}$ throughout all four cases, the differences might be explained by the expedient use of the ML method to estimate the aggregated weak GARCH(1,1) process. Taking account of the effects of the ML method applied to the estimation of a weak GARCH(1,1) process, the values of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ are indeed very close to $\alpha^L_{1(2)} + \beta^L_{1(2)}$ in all four cases. Is this a coincidence? The answer is probably no if the argument in (12a) that the aggregate of a two-stage PGARCH(1,1) process turns into a weak GARCH(1,1) process is true. Given this argument, using the strong standard GARCH(1,1) model to filter the aggregated process is appropriate and that is why the values of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ of the strong standard GARCH(1,1) model are close to the true values of the dynamics of the aggregated process, i.e. $\alpha^L_{1(2)} + \beta^L_{1(2)} = (\alpha^H_{11} + \beta^H_{11})(\alpha^H_{12} + \beta^H_{12})$, in all four cases. The numerically trivial biases of $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$ from $\alpha^L_{1(2)} + \beta^L_{1(2)}$ might come from the problem caused by applying the ML method to the parameter estimation of a weak GARCH(1,1) process.

Moreover, the average low-frequency level of integratedness $\overline{\alpha^{DN}_{1(2)}} + \overline{\beta^{DN}_{1(2)}}$, implied by the parameter estimates of the mis-specified *HF* strong GARCH(1,1) filter, is also found to be close to $\alpha^L_{1(2)} + \beta^L_{1(2)}$ across the four cases. If we consider this result together with the finding that the values of both $\overline{\hat{\phi}^L_{(2)}}$ and $\overline{\phi^{DN}_{(2)}}$ are close to that of $\phi^L_{(2)}$ in all four cases, argument (12a) can be further strengthened by the following argument:

a two-stage strong PGARCH(1,1) process will turn into a weak GARCH(1,1) process upon temporal aggregation, the aggregation interval being two original (or high-frequency) observation periods in length, and this aggregated weak GARCH(1,1) process is simply the one obtained by the aggregation of the ML estimated mis-specified GARCH(1,1) model filtering the two-stage PGARCH(1,1) process.

d. Impacts on the Average Estimates of the Numbers of Degrees of Freedom for the Standardised Innovations of the Strong Standard GARCH(1,1) Model for both* HF *and* LF *Observations:

Note that these discussions only involve cases with t_5 driving disturbances since the number of degrees of freedom is also a parameter under estimation in these cases. The values of $\overline{\hat{v}_1}$ and $\overline{\hat{v}_2}$ are therefore not available in cases 1 and 2 of Panel A of Table A.3, where the $N(0,1)$ driving disturbances are used.

(i) $\overline{\hat{v}_1}$: The value of $\overline{\hat{v}_1}$ is 5.011 in case 3 (see Panel B of Table A.3), whereas it is 5.04 under both Parameterisations 1 and 2 (see Table A.1). Also, the value of $\overline{\hat{v}_1}$ in case 4 is 4.996 (see Panel B of Table A.3), whereas it is 5.04 under both Parameterisations 1 and 3 (see Table A.1). Both results suggest that periodicity in the intercept and alpha of the PGARCH(1,1)-t_5 DGP does not lead to a biased estimate of the number of degrees of freedom of the mis-specified standard GARCH(1,1)-t filter.

(ii) $\overline{\hat{v}_2}$: Unfortunately, analytical results for the true number of degrees of freedom of the standardised aggregated innovations, v_2 are not available. We thus cannot compare $\overline{\hat{v}_2}$ to its true value v_2. However, we still can compare the values of $\overline{\hat{v}_2}$ under DGP 1 (case 3) and DGP 2 (case 4) to their benchmark values under the standard GARCH(1,1) DGP reported in Table 6.6. In particular, $\overline{\hat{v}_2}$ is 5.136 under DGP 1 in case 3, a value close to its counterparts, 5.254 under both Parameterisations 1 and 2, which together form DGP 1. The periodicity in the intercept under DGP 1 does not seem to cause much deviation in the value of $\overline{\hat{v}_2}$ for the *LF* GARCH(1,1) standardised innovations relative to its benchmarks of the two stages of the periodic cycle. The fixed dynamics parameters across the two stages under DGP 1 might be responsible for this result. On the other hand, $\overline{\hat{v}_2}$ is 6.188 under DGP 2 in case 4, a value slightly larger than the average of $\overline{\hat{v}_2}$= 5.254 under Parameterisation 1 and $\overline{\hat{v}_2}$= 6.475 under Parameterisation 3 of the standard GARCH(1,1) DGP. We have found in Section 6.3.1 that the number of degrees of freedom for the aggregated low-frequency observations is in inverse proportion to the size of the true alpha of the high-frequency GARCH(1,1) DGP. The value of $\overline{\hat{v}_2}$ falling between its benchmarks of the two stages thus might be a result of the alternating stages of large α_{11}^H = 0.15 and small α_{12}^H = 0.05, under DGP 2.

e. Impacts on the Average Unconditional Innovation Variances Implied by the Parameter Estimates of the Strong Standard GARCH(1,1) Model for both* HF *and* LF *Observations

In Appendix B, we have shown the derivation of the formulae for the unconditional innovation variance of each stage of the periodic cycle of a two-stage PGARCH(1,1) process in (10a), (10b), and (11). The discussion of this part will be based on these formulae.

(i) $\overline{(\hat{\sigma}^H)^2}$: To facilitate the discussion of the values of $\overline{(\hat{\sigma}^H)^2}$ in the context of volatility periodicity, we compare the values of $\overline{(\hat{\sigma}^H)^2}$ to the values of h_{odd}^H, given by (10a), and h_{even}^H, given by (10b) in Table 5 below.

Table 5: Comparison of the values of $\overline{(\hat{\sigma}^H)^2}$ to the values of h_{odd}^H and h_{even}^H

Case No	(DGP No, Disturbance)	$\overline{(\hat{\sigma}^H)^2}$	h^H_{odd}	h^H_{even}
1	(DGP 1, $N(0,1)$)	0.200 (0.009)	0.2108	0.1892
2	(DGP 2, $N(0,1)$)	0.248 (0.008)	0.2552	0.2414
3	(DGP 1, t_5)	0.201 (0.015)	0.2108	0.1892
4	(DGP 2, t_5)	0.249 (0.013)	0.2552	0.2414

Note: The sample standard deviations for N=10000 estimates are reported in parentheses.

Interestingly, the value of $\overline{(\hat{\sigma}^H)^2}$ is equal to the average of the values of h^H_{odd} and h^H_{even} in all four cases. It follows from DN's aggregation formulae that the unconditional variance of the aggregated innovations, $(\sigma^L_{(m)})^2$, of a weak GARCH(1,1) process is equal to the number of high-frequency periods within each aggregation interval, m, times the unconditional variance of the high-frequency innovations, $(\sigma^H)^2$. In the context of an aggregation frequency of 2, the theory simply suggests that, if we aggregated the high-frequency innovations filtered by the mis-specified GARCH(1,1) model into their low-frequency counterparts, we would expect to see the unconditional variance of the aggregated innovations equal to $2\overline{(\hat{\sigma}^H)^2}$. Coincidently, $2\overline{(\hat{\sigma}^H)^2}$ is just equal to the sum of h^H_{odd} and h^H_{even} according to the results in Table 5 above. Moreover, we have shown in (11) that the unconditional variance of the aggregated weak GARCH(1,1) innovations, $h^L_{(2)}$, is equal to $h^H_{odd} + h^H_{even} = \dfrac{\phi^H_1(1+\alpha^H_{12}+\beta^H_{12})+\phi^H_2(1+\alpha^H_{11}+\beta^H_{11})}{1-(\alpha^H_{11}+\beta^H_{11})(\alpha^H_{12}+\beta^H_{12})}$. Consequently, $2\overline{(\hat{\sigma}^H)^2}$ should also be equal to $h^L_{(2)}$. An implication of this result is that a PGARCH(1,1) process and a GARCH(1,1) process with parameters from the ML estimates of a strong standard GARCH(1,1) model filtering the PGARCH(1,1) process will aggregate into the same low-frequency weak GARCH(1,1) process under an aggregation frequency of 2. We suspect that this might also be the case for the integer aggregation frequencies larger than 2.

(ii) $\overline{(\hat{\sigma}^L_{(2)})^2}$ and $\overline{(\sigma^{DN}_{(2)})^2}$: To strengthen our point above, we compare the unconditional variance of the aggregated innovations implied by the *LF* GARCH(1,1) parameter estimates, $\overline{(\hat{\sigma}^L_{(2)})^2}$, to its counterpart implied by the parameters of the calculated aggregated GARCH(1,1) model and to $h^L_{(2)}$ from (11) in Table 6 below.

Table 6: Comparison of the values of $\overline{(\hat{\sigma}^H)^2}$ to the values of h^H_{odd} and h^H_{even}

Case No	(DGP No, Disturbance)	$\overline{(\hat{\sigma}^L_{(2)})^2}$	$\overline{(\sigma^{DN}_{(2)})^2}$	$h^L_{(2)} = h^H_{odd} + h^H_{even}$
1	(DGP 1, $N(0,1)$)	0.400 (0.020)	0.400 (0.017)	0.4
2	(DGP 2, $N(0,1)$)	0.496 (0.019)	0.496 (0.015)	0.4966
3	(DGP 1, t_5)	0.399 (0.033)	0.402 (0.031)	0.4
4	(DGP 2, t_5)	0.493 (0.029)	0.497 (0.027)	0.4966

Note: The sample standard deviations for N=10000 estimates are reported in parentheses.

Obviously, the values of $\overline{(\hat{\sigma}^L_{(2)})^2}$, $\overline{(\sigma^{DN}_{(2)})^2}$, and $h^L_{(2)}$ are extremely close to one another in all four cases. The fact that $\overline{(\hat{\sigma}^L_{(2)})^2} \approx \overline{(\sigma^{DN}_{(2)})^2}$ along with $\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}} \approx \overline{\alpha^{DN}_{1(2)}} + \overline{\beta^{DN}_{1(2)}}$ and $\overline{\hat{\phi}^L_{(2)}} \approx \overline{\phi^{DN}_{(2)}}$ provides support for our argument in (12a) that the aggregated observations are governed by a weak GARCH(1,1) process which does not show any periodicity in the parameters. The rationale behind this is that the aggregated low-frequency model implied by the ML estimates of the mis-specified *HF* GARCH(1,1) model via the DN formulae is a weak GARCH(1,1) free of periodicity. Therefore, if we use a GARCH(1,1) model, though 'strong' in definition, to estimate the aggregated low-frequency observations, and we end up with parameter estimates and the resulting unconditional innovation variance which are similar in size to those of the aforementioned aggregated weak GARCH(1,1) model, it is more than likely that the aggregated observations are free of periodicity, i.e. they are no longer governed by a PGARCH(1,1) process, but rather by a GARCH(1,1) process.

Following the argument above, moreover, the strong standard GARCH(1,1) filter for the low-frequency observations is not mis-specified in its parameter settings, though still mis-specified in assuming *i.i.d.* innovations[11]. The fact that $\overline{(\hat{\sigma}^L_{(2)})^2} \approx h^L_{(2)}$ in Table 6 above thus tends to certify the correctness of the formula for $h^L_{(2)}$ in (11).

f. Impacts on the Average Sample Non-Parametric Estimates for the Unconditional Kurtoses of the High-Frequency Innovations Standardised by the Strong Standard GARCH(1,1) Filter, and of the Low-frequency Innovations Standardised by the Strong Standard GARCH(1,1) Filter and by the Aggregated Weak GARCH(1,1) Model

(i) $\overline{\hat{k}^H_\xi}$: Comparison of the values of $\overline{\hat{k}^H_\xi}$ in Panel A of Table A.3 to the values of $\overline{\hat{k}^H_\xi}$ under Parameterisations 1 to 3 in Table A.1 is made in Table 7 below.

Table 7: Comparison of the values of $\overline{\hat{k}^H_\xi}$ in cases 1, 2, 3, and 4 of Table A.3 to their counterparts under Parameterisations 1 to 3 in Table A.1

PGARCH(1,1) DGPs			$\overline{\hat{k}^H_\xi}$ under Standard GARCH(1,1) DGPs					
Case No	**(DGP No, Disturbance)**	$\overline{\hat{k}^H_\xi}$	**Parameterisation 1**		**Parameterisation 2**		**Parameterisation 3**	
			$\xi_t \sim N(0,1)$	$\xi_t \sim t_5$	$\xi_t \sim N(0,1)$	$\xi_t \sim t_5$	$\xi_t \sim N(0,1)$	$\xi_t \sim t_5$
1	(DGP 1, $N(0,1)$)	3.012 (0.071)	2.998 (0.070)		2.998 (0.070)			
2	(DGP 2, $N(0,1)$)	3.017 (0.072)	2.998 (0.070)				2.998 (0.070)	
3	(DGP 1, t_5)	7.819 (4.514)		7.798 (4.593)		7.799 (4.592)		
4	(DGP 2, t_5)	7.861 (4.709)		7.798 (4.593)				7.797 (4.598)

Note: The sample standard deviations for N = 10000 estimates are reported in parentheses

Results in Table 7 above tend to suggest that the periodicity in the intercept or alpha of the two-stage strong PGARCH(1,1) DGPs does not seem to affect the estimation of the true k^H_ξ . For example, in case 1 where DGP 1 and the $N(0,1)$ driving disturbances are employed, the average unconditional kurtosis estimate of the high-frequency innovations standardised by the mis-specified strong GARCH(1,1) filter, $\overline{\hat{k}^H_\xi}$, is 3.012. In the standard context of no periodicity,

[11] If the aggregated low-frequency observations do follow a standard GARCH(1,1) process, it must be a weak GARCH(1,1) model given the DN aggregation theory. Therefore the strong GARCH(1,1) model used to estimate the low-frequency observations under the ML method is doomed to be mis-specified since it assumes that the innovations are *i.i.d.*

on the other hand, the values of $\overline{\hat{k}_{\xi}^{H}}$ under Parameterisation 1 and 2, which together constitute DGP 1, are both 2.998. 3.012 and 2.998 are very close to each other and are almost equal to the true value, 3. The same result is observed with the other three cases.

However, there is another point worth noting. That is, the fourth moment of the driving disturbances used in the strong PGARCH(1,1) DGP seems to have impacts on the accuracy of $\overline{\hat{k}_{\xi}^{H}}$ and the Monte Carlo standard deviation of $\overline{\hat{k}_{\xi}^{H}}$. The values of $\overline{\hat{k}_{\xi}^{H}}$ reported in Table 7 under both the strong PGARCH(1,1) DGPs and the no-periodicity strong standard GARCH(1,1) DGPs are close to their true values, 3, in the cases of $N(0,1)$ driving disturbances. Under t_5 disturbances, on the contrary, the sizes of $\overline{\hat{k}_{\xi}^{H}}$ are obviously smaller than their true value, $k_{\xi}^{H} = 9$. Refer to part (b) of the discussion of Table A.1 for explanation of this result.

Also, the Monte Carlo standard deviations of $\overline{\hat{k}_{\xi}^{H}}$ are trivially small in the $N(0,1)$ cases, whereas their counterparts in the t_5 cases are large relative to the values of $\overline{\hat{k}_{\xi}^{H}}$ themselves. This tends to suggest that large fourth moments of the driving disturbances will make the Monte Carlo results of $\overline{\hat{k}_{\xi}^{H}}$ less reliable.

(ii) $\overline{\hat{k}_{\xi_{(2)}}^{L}}$ and $\overline{\hat{k}_{\xi_{(2)}}^{DN}}$: As reported in Panel B of Table A.3, it might be surprising to observe that the values of $\overline{\hat{k}_{\xi_{(2)}}^{L}}$ and $\overline{\hat{k}_{\xi_{(2)}}^{DN}}$ are very close to each other in all four cases. Following the same argument in point (ii) of the discussions in part e, this result tends to verify our speculation: argument (a) in (12a).

4 Concluding remarks

We have shown the formulae for the unconditional innovation variance of each of the two stages of a PGARCH(1,1) DGP. According to the formulae in (10a) and (10b), the two-stage periodic cycle in the intercept or alpha parameter of a PGARCH(1,1) DGP implies the two-stage variation in the unconditional innovation variance. Moreover, a straightforward implication of these formulae is the formula for the true unconditional variance of the aggregated innovations ε_λ, i.e. h_λ in (11).

The formula for h_λ in (11) tends to suggest two points. First of all, whilst the conditional variance of the high-frequency observations generated by the two-stage strong PGARCH(1,1) DGP demonstrate periodicity in the intercept or alpha parameter, the aggregated low-frequency observations follow a non-periodic weak GARCH(1,1) process. Second, following the argument above, the true intercept of the aggregated weak GARCH(1,1) process is equal to ϕ_1^H (1+ α_{12}^H + β_{12}^H)+ ϕ_2^H (1+ α_{11}^H + β_{11}^H), and its true level of integratedness is specified by ($\alpha_{11}^H + \beta_{11}^H$)($\alpha_{12}^H + \beta_{12}^H$). The first point is extended, moreover, by our finding that the aggregated weak GARCH(1,1) process is identical to the one obtained by temporal aggregation, in the DN sense, of the ML estimated mis-specified strong GARCH(1,1) model filtering the two-stage PGARCH(1,1) process. All these arguments above are strongly supported by our simulation results.

Interestingly, the periodicity in the intercept and alpha parameter of our two-stage or five-stage PGARCH(1,1) DGPs does not seem to have any special impacts on the model estimation of the mis-specified strong standard GARCH(1,1) filter for the high-frequency observations. The intercept estimate of the mis-specified *HF* GARCH(1,1) filter is simply equal to the average of the true intercepts of the two stages under DGP 1 or of the five stages under DGP 3. Similarly, the alpha estimate is equal to the average of the true alpha's of the two stages under DGP 2 or of the five stages under DGP 4. Moreover, the periodicity does not have impacts at all on the strong standard GARCH(1,1) filter for the aggregated low-frequency observations, a result of the first point above. Finally, these findings seem to be robust to the fourth moment of the driving disturbances.

Appendix A

Table A.1: GARCH(1,1) parameterisations for the high-frequency DGP and their ML estimates

High-Frequency Estimation	Driving Disturbances: $\xi_t^H \sim N(0,1)$ Parameterisation: 1	2	3	$\xi_t^H \sim t_5$ Parameterisation: 1	2	3
ϕ^H	0.05	0.01	0.05	0.05	0.01	0.05
$\overline{\hat{\phi}^H}$	0.051 (0.008) [0.001**]	0.0102 (0.002) [0.0002**]	0.055 (0.025) [0.005**]	0.051 (0.008) [0.001**]	0.0101 (0.002) [0.0001**]	0.054 (0.021) [0.004**]
α_1^H	0.15	0.15	0.05	0.15	0.15	0.05
$\overline{\hat{\alpha}_1^H}$	0.15 (0.016) [0]	0.15 (0.016) [0]	0.051 (0.014) [0.001**]	0.15 (0.019) [0]	0.15 (0.019) [0]	0.051 (0.015) [0.001**]
β_1^H	0.7	0.7	0.7	0.7	0.7	0.7
$\overline{\hat{\beta}_1^H}$	0.697 (0.034) [-0.003**]	0.697 (0.034) [-0.003**]	0.675 (0.132) [-0.025**]	0.698 (0.035) [-0.002**]	0.698 (0.035) [-0.002**]	0.681 (0.115) [-0.019**]
$\alpha_1^H + \beta_1^H$	0.85	0.85	0.75	0.85	0.85	0.75
$\overline{\hat{\alpha}_1^H} + \overline{\hat{\beta}_1^H}$	0.847 (0.025) [-0.003**]	0.847 (0.025) [-0.003**]	0.726 (0.125) [-0.024**]	0.848 (0.025) [-0.002**]	0.848 (0.025) [-0.002**]	0.732 (0.107) [-0.018**]
v_1				5	5	5
$\overline{\hat{v}_1}$				5.04 (0.36) [0.04**]	5.04 (0.36) [0.04**]	5.04 (0.36) [0.04**]
$(\sigma^H)^2$	0.33	0.067	0.2	0.33	0.0666	0.2
$\overline{(\hat{\sigma}^H)^2}$	0.33 (0.014) [0]	0.067 (0.003) [0]	0.2 (0.005) [0]	0.335 (0.025) [0.005**]	0.0669 (0.005) [0.0003**]	0.2 (0.008) [0]
k_ξ^H	3	3	3	9	9	9
$\overline{\hat{k}_\xi^H}$	2.998 (0.070) [-0.002**]	2.998 (0.070) [-0.002**]	2.998 (0.070) [-0.002**]	7.798 (4.593) [-1.202**]	7.799 (4.592) [-1.201**]	7.797 (4.598) [-1.203**]

Notes: 1. ϕ^H, α_1^H, and β_1^H refer to the parameters of the strong standard GARCH(1,1) filter for high-frequency observations. v_1 denotes the number of degrees of freedom of the *Student*'s *t* distribution assumed for the high-frequency return innovations. $(\sigma^H)^2$ denotes the unconditional innovation variance, i.e. $(\sigma^H)^2 = \dfrac{\phi^H}{(1-\alpha_1^H - \beta_1^H)}$, and k_ξ^H denotes the unconditional kurtosis of the standardised high-frequency innovation ξ_t.

2. A circumflexed symbol, say, $\hat{\phi}^H$, denotes the estimate for the true value in ϕ^H, whilst the bar over $\hat{\phi}^H$, i.e. $\overline{\hat{\phi}^H}$ denotes the mean of all $\hat{\phi}^H$ across the 10000 Monte Carlo simulations. Namely, $\overline{\hat{\phi}^H} = \dfrac{\sum_{i=1,\ldots,N} \hat{\phi}_i^H}{N}$, where $N = 10000$.

3. Parenthesised figures denote the sample standard deviations for the estimates from 10000 Monte Carlo replications, whilst figures in square brackets denote the average biases, i.e. the deviation of the average estimates from their true values, i.e. $\overline{\hat{\phi}^H}$ - ϕ^H , for example.
4. Significance tests are performed on the figures in the square brackets to see whether the average biases of the estimates are significantly different from zero. The test statistics, $\frac{\bar{x}-0}{StD(x)/\sqrt{N}}$, where x denotes the bias of the estimates and N refers to the number of Monte Carlo simulations, follows a standard normal distribution according to the central limit theorem. The ** sign shows significance at the 1% significance level, whereas * indicates significance at the 5% significance level. Figures in the shaded brackets are not significantly different from zero at the 5% significance level.

Table A.2: Parameterisations of the aggregated weak GARCH(1,1) model implied by the corresponding high-frequency parameterisations in Table A.1 and their ML estimates

Low-Frequency Estimation	**Driving Disturbances:**					
	$\xi_t \sim N(0,1)$			$\xi_t \sim t_5$		
	Parameterisation:			**Parameterisation:**		
	1	**2**	**3**	**1**	**2**	**3**
$\phi^L_{(2)}$	0.185	0.037	0.175	0.185	0.037	0.175
$\overline{\hat{\phi}^L_{(2)}}$	0.195 (0.057) [0.01**]	0.039 (0.011) [0.002**]	0.215 (0.147) [0.04**]	0.194 (0.044) [0.009**]	0.039 (0.009) [0.002**]	0.194 (0.107) [0.019**]
$\alpha^L_{1(2)}$	0.126	0.126	0.038	0.173	0.173	0.057
$\overline{\hat{\alpha}^L_{1(2)}}$	0.127 (0.027) [0.001**]	0.127 (0.027) [0.001**]	0.039 (0.021) [0.001**]	0.163 (0.033) [-0.01**]	0.163 (0.033) [-0.01**]	0.053 (0.023) [-0.004**]
$\beta^L_{1(2)}$	0.597	0.597	0.525	0.55	0.55	0.506
$\overline{\hat{\beta}^L_{1(2)}}$	0.581 (0.101) [-0.016**]	0.581 (0.101) [-0.016**]	0.422 (0.372) [-0.103**]	0.543 (0.083) [-0.007**]	0.543 (0.083) [-0.007**]	0.459 (0.279) [-0.047**]
$\alpha^L_{1(2)} + \beta^L_{1(2)}$	0.723	0.723	0.563	0.723	0.723	0.563
$\overline{\hat{\alpha}^L_{1(2)}} + \overline{\hat{\beta}^L_{1(2)}}$	0.708 (0.088) [-0.015**]	0.708 (0.088) [-0.015**]	0.461 (0.367) [-0.102**]	0.706 (0.068) [-0.017**]	0.706 (0.068) [-0.017**]	0.512 (0.272) [-0.051**]
$\overline{\hat{v}_2}$				5.254 (0.574)	5.254 (0.574)	6.475 (0.869)
$(\sigma^L_{(2)})^2$	0.67	0.13333	0.4	0.67	0.13	0.4
$\overline{(\hat{\sigma}^L_{(2)})^2}$	0.67 (0.033) [0]	0.1333 (0.007) [-0.00003**]	0.4 (0.013) [0]	0.663 (0.054) [-0.007**]	0.133 (0.011) [-0.003**]	0.398 (0.019) [-0.002**]

Notes: 1. Footnotes to Table A.1 also apply here, the superscript ^{L} being substituted for ^{H} to denote that the notation used here refers to the calculated and estimated parameters of the *aggregated* (*low-frequency*) weak GARCH(1,1) model. Parenthesised 2 in the subscripts denote the number of high-frequency observations within each aggregation interval.
2. The true values of the parameters are from the DN aggregation formulae.
3. As the theoretical value in the number of degrees of freedom for the standardised aggregated low-frequency innovations, i.e. v_2 , is not available, the implied unconditional kurtosis of the standardised aggregated low-frequency innovation, i.e. k^L_ξ , is not available either. We thus exclude both their average estimates from Table A.2.

Table A.3: The Average ML Estimates of the Strong Standard HF and LF GARCH(1,1) Filters under DGPs 1 and 2

Panel A: high-frequency estimates

Case No	**(DGP No, Driving Disturbance)**	$\phi^H_{s(t)}$	$\overline{\hat{\phi}^H}$	$\alpha^H_{1s(t)}$	$\overline{\hat{\alpha}^H_1}$	$\beta^H_{1s(t)}$	$\overline{\hat{\beta}^H_1}$	$\alpha^H_{1s(t)}+\beta^H_{1s(t)}$	$\overline{\hat{\alpha}^H_1}+\overline{\hat{\beta}^H_1}$	$(\sigma^H_{s(t)})^2$	$\overline{(\hat{\sigma}^H)^2}$	k^H_ξ	$\overline{\hat{k}^H_\xi}$
1	(DGP 1, $N(0,1)$)	ϕ^H_1 =0.05 ϕ^H_2 =0.01	0.030 (0.005)	$\alpha^H_{11}=\alpha^H_{12}$ =0.15	0.149 (0.016)	$\beta^H_{11}=\beta^H_{12}$ =0.7	0.698 (0.033)	$\alpha^H_{11}+\beta^H_{11}$ =0.85 $\alpha^H_{12}+\beta^H_{12}$ =0.85	0.848 (0.025)	$(\sigma^H_1)^2$ =0.2108 $(\sigma^H_2)^2$ =0.1892	0.2 (0.009)	3	3.012 (0.071)
2	(DGP 2, $N(0,1)$)	$\phi^H_1=\phi^H_2$ =0.05	0.052 (0.012)	α^H_{11} =0.15 α^H_{12} =0.05	0.100 (0.016)	$\beta^H_{11}=\beta^H_{12}$ =0.7	0.690 (0.056)	$\alpha^H_{11}+\beta^H_{11}$ =0.85 $\alpha^H_{12}+\beta^H_{12}$ =0.75	0.790 (0.047)	$(\sigma^H_1)^2$ =0.2552 $(\sigma^H_2)^2$ =0.2414	0.248 (0.008)	3	3.017 (0.072)

Case No	**(DGP No, Driving Disturbance)**	$\phi^H_{s(t)}$	$\overline{\hat{\phi}^H}$	$\alpha^H_{1s(t)}$	$\overline{\hat{\alpha}^H_1}$	$\beta^H_{1s(t)}$	$\overline{\hat{\beta}^H_1}$	$\alpha^H_{1s(t)}+\beta^H_{1s(t)}$	$\overline{\hat{\alpha}^H_1}$ + $\overline{\hat{\beta}^H_1}$	v_1	$\overline{\hat{v}_1}$	$(\sigma^H_{s(t)})^2$	$\overline{(\hat{\sigma}^H)^2}$	k^H_ξ	$\overline{\hat{k}^H_\xi}$
3	(DGP 1, t_5)	ϕ^H_1 =0.05 ϕ^H_2 =0.01	0.030 (0.005)	$\alpha^H_{11}=\alpha^H_{12}$ =0.15	0.150 (0.019)	$\beta^H_{11}=\beta^H_{12}$ =0.7	0.699 (0.035)	$\alpha^H_{11}+\beta^H_{11}$ =0.85 $\alpha^H_{12}+\beta^H_{12}$ =0.85	0.849 (0.025)	5	5.01 (0.36)	$(\sigma^H_1)^2$ =0.2108 $(\sigma^H_2)^2$ =0.1892	0.201 (0.015)	9	7.82 (4.51)
4	(DGP 2, t_5)	$\phi^H_1=\phi^H_2$ =0.05	0.052 (0.011)	α^H_{11} =0.15 α^H_{12} =0.05	0.098 (0.018)	$\beta^H_{11}=\beta^H_{12}$ =0.7	0.692 (0.055)	$\alpha^H_{11}+\beta^H_{11}$ =0.85 $\alpha^H_{12}+\beta^H_{12}$ =0.75	0.791 (0.045)	5	5.00 (0.36)	$(\sigma^H_1)^2$ =0.2552 $(\sigma^H_2)^2$ =0.2414	0.249 (0.013)	9	7.86 (4.71)

Panel B: low-frequency estimates

Case No	(DGP No, Driving Disturbance)	$\overline{\hat{\phi}^{L}_{(m)}}$	$\overline{\phi^{DN}_{(m)}}$	$\overline{\hat{\alpha}^{L}_{1(m)}}$	$\overline{\alpha^{DN}_{1(m)}}$	$\overline{\hat{\beta}^{L}_{1(m)}}$	$\overline{\beta^{DN}_{1(m)}}$	$\overline{\hat{\alpha}^{L}_{1(m)}} + \overline{\hat{\beta}^{L}_{1(m)}}$	$\overline{\alpha^{DN}_{1(m)}} + \overline{\beta^{DN}_{1(m)}}$	$\overline{(\hat{\sigma}^{L}_{(m)})^2}$	$\overline{(\sigma^{DN}_{(m)})^2}$
1	(DGP 1, $N(0,1)$)	0.117 (0.035)	0.112 (0.016)	0.124 (0.027)	0.125 (0.012)	0.583 (0.103)	0.594 (0.042)	0.707 (0.089)	0.719 (0.042)	0.4 (0.020)	0.4 (0.017)
2	(DGP 2, $N(0,1)$)	0.195 (0.082)	0.186 (0.036)	0.085 (0.025)	0.078 (0.01)	0.521 (0.177)	0.548 (0.073)	0.606 (0.167)	0.626 (0.072)	0.496 (0.019)	0.496 (0.015)

Case No	(DGP No, Driving Disturbance)	$\overline{\hat{\phi}^{L}_{(m)}}$	$\overline{\phi^{DN}_{(m)}}$	$\overline{\hat{\alpha}^{L}_{1(m)}}$	$\overline{\alpha^{DN}_{1(m)}}$	$\overline{\hat{\beta}^{L}_{1(m)}}$	$\overline{\beta^{DN}_{1(m)}}$	$\overline{\hat{\alpha}^{L}_{1(m)}} + \overline{\hat{\beta}^{L}_{1(m)}}$	$\overline{\alpha^{DN}_{1(m)}} + \overline{\beta^{DN}_{1(m)}}$	$\overline{\hat{v}_2}$	$\overline{(\hat{\sigma}^{L}_{(m)})^2}$	$\overline{(\sigma^{DN}_{(m)})^2}$	$\overline{\hat{k}^{L}_{\xi_{(m)}}}$
3	(DGP 1, t_5)	0.117 (0.027)	0.112 (0.016)	0.161 (0.033)	0.158 (0.018)	0.544 (0.084)	0.563 (0.044)	0.705 (0.069)	0.721 (0.042)	5.136 (0.548)	0.399 (0.033)	0.402 (0.031)	7.90 (10.74)
4	(DGP 2, t_5)	0.188 (0.059)	0.185 (0.035)	0.110 (0.029)	0.103 (0.016)	0.509 (0.134)	0.524 (0.073)	0.619 (0.121)	0.627 (0.070)	6.188 (0.796)	0.493 (0.029)	0.497 (0.027)	6.32 (7.21)

Notes: 1. Statistics with superscript ^{DN} are of the implied aggregated weak GARCH(1,1) model, calculated by the DN aggregation formulae applied on the mis-specified standard strong *HF* GARCH(1,1) filter. 2. Statistics with superscript ^{L} are related to the strong standard GARCH(1,1) filter estimated for the aggregated observations. Note that $\hat{v}_2$ and $\hat{k}^{L}_{\xi_{(m)}}$ denote, respectively, the estimated number of degrees of freedom and unconditional kurtosis for the *t*-distributed standardised aggregated innovations of the *LF* GARCH(1,1) filter. However, there are no formulae for the analytical results of v_2 and $k^{L}_{\xi_{(m)}}$. 3. The sample standard deviations for $N = 10000$ estimates are reported in parentheses.

Appendix B

—Derivations of the Formulae for the Unconditional Innovation Variances of a Two-Stage PGARCH(1,1) Model

A general two-stage PGARCH(1,1) model is given by

$$h_t = \phi_{s(t)} + \alpha_{1s(t)}\varepsilon_{t-1}^2 + \beta_{1s(t)}h_{t-1}, \quad \text{(B.1)}$$

where *s(t)* = 1 for odd *t* and 2 for even *t*.

It follows from (B.1) that $h_2 = \phi_2 + \alpha_{12}\varepsilon_1^2 + \beta_{12}h_1$. (B.2)

Moving the time tag one period forward and substituting h_2 by (B.2) leads to

$$h_3 = \phi_1 + \alpha_{11}\varepsilon_2^2 + \beta_{11}h_2 = \phi_1 + \alpha_{11}\varepsilon_2^2 + \beta_{11}(\phi_2 + \alpha_{12}\varepsilon_1^2 + \beta_{12}h_1)$$

$$= \phi_1 + \beta_{11}\phi_2 + \alpha_{11}\varepsilon_2^2 + \alpha_{12}\beta_{11}\varepsilon_1^2 + \beta_{11}\beta_{12}h_1. \quad \text{(B.3)}$$

Similarly, we can get

$$h_4 = \phi_2 + \alpha_{12}\varepsilon_3^2 + \beta_{12}h_3 = \phi_2 + \alpha_{12}\varepsilon_3^2 + \beta_{12}(\phi_1 + \beta_{11}\phi_2 + \alpha_{11}\varepsilon_2^2 + \alpha_{12}\beta_{11}\varepsilon_1^2 + \beta_{11}\beta_{12}h_1)$$

$$= \phi_2 + \beta_{12}\phi_1 + \beta_{11}\beta_{12}\phi_2 + \alpha_{12}\varepsilon_3^2 + \alpha_{11}\beta_{12}\varepsilon_2^2 + \alpha_{12}\beta_{11}\beta_{12}\varepsilon_1^2 + \beta_{11}\beta_{12}^2h_1, \quad \text{(B.4)}$$

$$h_5 = \phi_1 + \alpha_{11}\varepsilon_4^2 + \beta_{11}h_4$$

$$= \phi_1 + \alpha_{11}\varepsilon_4^2 + \beta_{11}(\phi_2 + \beta_{12}\phi_1 + \beta_{11}\beta_{12}\phi_2 + \alpha_{12}\varepsilon_3^2 + \alpha_{11}\beta_{12}\varepsilon_2^2 + \alpha_{12}\beta_{11}\beta_{12}\varepsilon_1^2 + \beta_{11}\beta_{12}^2h_1)$$

$$= \phi_1 + \beta_{11}\phi_2 + \beta_{11}\beta_{12}\phi_1 + \beta_{11}^2\beta_{12}\phi_2 + \alpha_{11}\varepsilon_4^2 + \alpha_{12}\beta_{11}\varepsilon_3^2 + \alpha_{11}\beta_{11}\beta_{12}\varepsilon_2^2 + \alpha_{12}\beta_{11}^2\beta_{12}\varepsilon_1^2 + \beta_{11}^2\beta_{12}^2h_1, \quad \text{(B.5)}$$

$$h_6 = \phi_2 + \alpha_{12}\varepsilon_5^2 + \beta_{12}h_5$$

$$= \phi_2 + \alpha_{12}\varepsilon_5^2 + \beta_{12}(\phi_1 + \beta_{11}\phi_2 + \beta_{11}\beta_{12}\phi_1 + \beta_{11}^2\beta_{12}\phi_2 + \alpha_{11}\varepsilon_4^2 + \alpha_{12}\beta_{11}\varepsilon_3^2 + \alpha_{11}\beta_{11}\beta_{12}\varepsilon_2^2 + \alpha_{12}\beta_{11}^2\beta_{12}\varepsilon_1^2 + \beta_{11}^2\beta_{12}^2h_1)$$

$$= \phi_1(\beta_{12} + \beta_{11}\beta_{12}^2) + \phi_2(1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2) + \alpha_{12}\varepsilon_5^2 + \alpha_{11}\beta_{12}\varepsilon_4^2 + \alpha_{12}\beta_{11}\beta_{12}\varepsilon_3^2 + \alpha_{11}\beta_{11}\beta_{12}^2\varepsilon_2^2 + \alpha_{12}\beta_{11}^2\beta_{12}^2\varepsilon_1^2 + \beta_{11}^2\beta_{12}^3h_1, \quad \text{(B.6)}$$

$$h_7 = \phi_1 + \alpha_{11}\varepsilon_6^2 + \beta_{11}h_6$$

$$= \phi_1 + \alpha_{11}\varepsilon_6^2 + \beta_{11}[\phi_1(\beta_{12} + \beta_{11}\beta_{12}^2) + \phi_2(1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2) + \alpha_{12}\varepsilon_5^2 + \alpha_{11}\beta_{12}\varepsilon_4^2 + \alpha_{12}\beta_{11}\beta_{12}\varepsilon_3^2 + \alpha_{11}\beta_{11}\beta_{12}^2\varepsilon_2^2 + \alpha_{12}\beta_{11}^2\beta_{12}^2\varepsilon_1^2 + \beta_{11}^2\beta_{12}^3h_1]$$

$$= \phi_1(1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2) + \phi_2(\beta_{11} + \beta_{11}^2\beta_{12} + \beta_{11}^3\beta_{12}^2) + \alpha_{11}\varepsilon_6^2 + \alpha_{12}\beta_{11}\varepsilon_5^2 + \alpha_{11}\beta_{11}\beta_{12}\varepsilon_4^2 + \alpha_{12}\beta_{11}^2\beta_{12}\varepsilon_3^2 + \alpha_{11}\beta_{11}^2\beta_{12}^2\varepsilon_2^2 + \alpha_{12}\beta_{11}^3\beta_{12}^2\varepsilon_1^2 + \beta_{11}^3\beta_{12}^3h_1 \quad \text{(B.7)}$$

Let h_{odd} denote the unconditional variance of ε_t for odd *t*, and h_{even} the unconditional variance of ε_t for even *t*. The form of h_{odd} can be induced from (B.3), (B.5), and (B.7) as follows:

$$h_{odd} = \phi_1 (1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + ...) + \phi_2\beta_{11} (1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + \ldots) + \alpha_{11} (1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + ...) h_{even} + \alpha_{12}\beta_{11} (1+ \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + \ldots) h_{odd} . \quad \text{(B.8)}$$

If $0 \le \beta_{11}\beta_{12} < 1$, (B.8) can be rewritten as

$$h_{odd} = \frac{\phi_1}{1-\beta_{11}\beta_{12}} + \frac{\phi_2\beta_{11}}{1-\beta_{11}\beta_{12}} + \frac{\alpha_{11}}{1-\beta_{11}\beta_{12}} h_{even} + \frac{\alpha_{12}\beta_{11}}{1-\beta_{11}\beta_{12}} h_{odd} . \quad \text{(B.9)}$$

Moving the h_{odd} term in the RHS of (B.9) to the LHS gives

$$(1 - \frac{\alpha_{12}\beta_{11}}{1-\beta_{11}\beta_{12}}) h_{odd} = \frac{\phi_1 + \phi_2\beta_{11}}{1-\beta_{11}\beta_{12}} + \frac{\alpha_{11}}{1-\beta_{11}\beta_{12}} h_{even} . \quad \text{(B.10)}$$

Multiplying $1-\beta_{11}\beta_{12}$ through (B.10) gives

$$(1-\beta_{11}\beta_{12} - \alpha_{12}\beta_{11}) h_{odd} = \phi_1 + \phi_2\beta_{11} + \alpha_{11} h_{even} \quad \text{(B.11)}$$

By similar arguments,

$$h_{even} = \phi_1\beta_{12} (1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + \ldots) + \phi_2 (1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + \ldots) + \alpha_{12} (1+ \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + \ldots) h_{odd} + \alpha_{11}\beta_{12} (1 + \beta_{11}\beta_{12} + \beta_{11}^2\beta_{12}^2 + \ldots) h_{even}$$

$$\Rightarrow (1 - \frac{\alpha_{11}\beta_{12}}{1-\beta_{11}\beta_{12}}) h_{even} = \frac{\phi_1\beta_{12} + \phi_2}{1-\beta_{11}\beta_{12}} + \frac{\alpha_{12}}{1-\beta_{11}\beta_{12}} h_{odd}$$

$$\Rightarrow (1-\beta_{11}\beta_{12} - \alpha_{11}\beta_{12}) h_{even} = \phi_1\beta_{12} + \phi_2 + \alpha_{12} h_{odd} \quad \text{(B.12)}$$

From (B.11), $h_{even} = -\dfrac{\phi_1 + \phi_2\beta_{11}}{\alpha_{11}} + \dfrac{1-\beta_{11}\beta_{12} - \alpha_{12}\beta_{11}}{\alpha_{11}} h_{odd}$ (B.13)

From (B.12), $h_{even} = \dfrac{\phi_1\beta_{12} + \phi_2}{1-\beta_{11}\beta_{12} - \alpha_{11}\beta_{12}} + \dfrac{\alpha_{12}}{1-\beta_{11}\beta_{12} - \alpha_{11}\beta_{12}} h_{odd}$, (B.14)

Equating the RHS of equations (B.13) and (B.14) above gives

$$-\frac{\phi_1 + \phi_2\beta_{11}}{\alpha_{11}} + \frac{1-\beta_{11}\beta_{12} - \alpha_{12}\beta_{11}}{\alpha_{11}} h_{odd} = \frac{\phi_1\beta_{12} + \phi_2}{1-\beta_{11}\beta_{12} - \alpha_{11}\beta_{12}} + \frac{\alpha_{12}}{1-\beta_{11}\beta_{12} - \alpha_{11}\beta_{12}} h_{odd} .$$

$$\therefore h_{odd} = \frac{\phi_1 + \phi_2(\alpha_{11} + \beta_{11})}{1 - (\alpha_{11} + \beta_{11})(\alpha_{12} + \beta_{12})} \tag{B.15}$$

Substituting equation (B.15) into equation (B.13) gives

$$h_{even} = \frac{\phi_2 + \phi_1(\alpha_{12} + \beta_{12})}{1 - (\alpha_{11} + \beta_{11})(\alpha_{12} + \beta_{12})}\,. \tag{B.16}$$

References

Bollerslev, T. (1986) "Generalised Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, **31**, 307-327.

Bollerslev, T. and E. Ghysels (1996) "Periodic Autoregressive Conditional Heteroscedasticity," *Journal of Business and Economic Statistics*, 14(2), 139-151.

Drost, F.C. and T.E. Nijman (1993) "Temporal Aggregation of GARCH Processes," *Econometrica*, 61, 909-927.

Nelson, D.B. (1990a) "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, **45**, 7-38

Nijman, T.E. and F.C. Palm (1990a) "Parameter Identification in ARMA-Processes in the Presence of Regular but Incomplete Sampling," *Journal of Time Series Analysis*, **11**, 239-248.

Nijman, T.E. and F.C. Palm (1990b) "Predictive Accuracy Gain from Disaggregation Sampling in ARIMA Models," *Journal of Business and Economic Statistics*, **8**, 405-415.

Palm, F.C. and T.E. Nijman (1984) "Missing Observations in the Dynamic Regression Model," *Econometrica*, **52**, 1415-1435.