UNDERSTANDING TEMPORAL AGGREGATION EFFECTS ON KURTOSIS IN FINANCIAL INDICES

By

Offer Lieberman and Peter C.B. Phillips

October 2, 2018

RESEARCH INSTITUTE FOR ECONOMETRICS

DISCUSSION PAPER NO. 1-18

DEPARTMENT OF ECONOMICS
BAR-ILAN UNIVERSITY
RAMAT-GAN 5290002, ISRAEL

http://econ.biu.ac.il/en/node/2473

Understanding Temporal Aggregation Effects on Kurtosis in Financial Indices*

Offer Lieberman† and Peter C. B. Phillips‡

June 17, 2018

Abstract

Indices of financial returns typically display sample kurtosis that declines towards the Gaussian value 3 as the sampling interval increases. This paper uses stochastic unit root (STUR) and continuous time analysis to explain the phenomenon. Limit theory for the sample kurtosis reveals that STUR specifications provide two sources of excess kurtosis, both of which decline with the sampling interval. Limiting kurtosis is shown to be random and is a functional of the limiting price process. Using a continuous time version of the model under no-drift, local drift, and drift inclusions, we suggest a new continuous time kurtosis measure for financial returns that assists in reconciling these models with the empirical kurtosis characteristics of returns. Simulations are reported and applications to several financial indices demonstrate the usefulness of this approach.

Key words and phrases: Autoregression; Diffusion; Kurtosis; Stochastic unit root; Time-varying coefficients.

JEL Classification: C22

1 Introduction

Asset pricing models with roots in the vicinity of unity that correspond to near martingale generating mechanisms have attracted considerable attention in financial theory,

*We thank Tim Ginker for his research assistance.

†Bar-Ilan University. Support from Israel Science Foundation grant No. 1182-17 is gratefully acknowledged. Correspondence to: Department of Economics and Research Institute for Econometrics (RIE), Bar-Ilan University, Ramat Gan 52900, Israel. E-mail: offer.lieberman@gmail.com

‡Yale University, University of Auckland, Southampton University, and Singapore Management University. Research support is acknowledged from the Kelly Foundation at the University of Auckland. Email: peter.phillips@yale.edu

predictive regression analyses, and empirical applications. Given the well-established stylized features of heavy-tailedness, high peakedness, and higher moment conditional dependence that are displayed by asset returns, plausible models also need to generate non-Gaussian behavior and accommodate conditional heterogeneity. One class of model that is capable of producing these characteristics while retaing near martingale behavior is a nonlinear time dependent autoregression with a root that is local to unity or stochastically local to unity.

A secondary stylized fact of financial asset returns is that their sample kurtosis typically declines towards 3 as the sampling interval increases. This paper explores whether variants of stochastic unit root (STUR) models are capable of mimicking this additional characteristic. We use discrete time STUR models together with continuous time analogues of these models and of the usual kurtosis measures to assist in explaining this additional stylized fact of empirical asset return data.

To fix ideas, we consider the following local stochastic unit root (LSTUR) model (Lieberman and Phillips, 2017c, henceforth LP)

$$\begin{aligned} Y_1 &= \mu + \varepsilon_1, \\ Y_t &= \mu + \beta_{nt} Y_{t-1} + \varepsilon_t, \ t = 2, ..., n, \\ \beta_{nt} &= \exp\left(\frac{c}{n} + a\frac{u_t}{\sqrt{n}}\right), \end{aligned} \tag{1}$$

where a and c are localizing coefficient parameters and $\mu = \mu_n$ is a drift parameter that may be zero, non-zero, or local to zero. Conditions on the STUR driver variable u_t and the error ε_t are given later in Assumption 1. For brevity, we write the time varying autoregressive coefficient as $\beta_{nt} = \beta_t$ in what follows. This autoregressive coefficient is a stochastically time varying parameter that fluctuates with u_t and allows for additional departures from unity by means of a conventional local-to-unit-root (LUR) specification involving the fixed localizing coefficient c. The model is therefore 'hybrid' in the sense that β_t includes both a deterministic localizing component and a stochastic component, thus bringing together into one model two main streams of literature on autoregressions with near unit roots, viz., LUR and STUR formulations. For the background literature on these specifications see, among others, Chan and Wei (1987), Phillips (1987) and Bykhovskaya and Phillips (2018, 2019) for the former stream, and Leybourne, McCabe and Mills (1996), Leybourne, McCabe and Tremayne (1996), Granger and Swanson (1997), McCabe and Smith, (1998), Yoon (2006), Lieberman (2012) and Lieberman and Phillips (2014, 2017a, 2017b) for the latter. The hybrid model that combines these elements was applied by LP in explaining the spread between an index of investment grade rated corporate debt and the spot Treasury curve as a function of the return on the S&P500 index.

Limit theory for the $\sqrt{n}$-normalized process as well as for the nonlinear least squares estimators (NLLS) of a, c and of β_t were established by LP in the $\mu = 0$ case and were shown to be functionals of a nonlinear diffusion process that satisfies a nonlinear stochastic differential equation corresponding to a structural model of option pricing that has been considered in the continuous time mathematical finance literature (Föllmer and Schweizer, 1993) and in some recent continuous time econometric work (Tao et al. 2018). The results were shown to generalize the theory already known in the special cases of LUR (Phillips 1987) and STUR (Lieberman and Phillips 2017a).

In this paper we show that the sample kurtosis of temporally aggregated returns based on the LSTUR model converges to a random variable which exceeds the Gaussian value 3 and decreases according to the level of aggregation. This result is consistent with much financial return data and provides a model-based explanation for the empirical phenomena. To assist in the analysis, we introduce new measures of kurtosis that are based on continuous time versions of the model and investigate their limiting forms for various configurations of base model, allowing for zero drift, local drift and dominant drift cases. A further contribution is the asymptotic analysis of a fitted misspecified fixed-coefficient autoregression and its associated kurtosis measures.

The plan for the rest of the paper is as follows. Notation and assumptions are given in Section 2. Limit theory for the sample kurtosis of temporally aggregated return data for the $\mu = 0$ case is established in Section 3 and for $\mu \neq 0$ in Section 4. In Section 5 we analyze the effects of misspecification of an LSTUR model by a simple AR(1) model and in Section 6 we introduce measures of kurtosis based on continuous time versions of the model. Simulations are provided to explore numerical support for the limit theory and the theoretical results on kurtosis in Section 7. An empirical application is given in Section 8. Section 9 concludes and proofs are placed in the Appendix.

2 Notation and Assumptions

The following assumption is used in developing asymptotic theory of the LSTUR model and estimated kurtosis coefficients for temporally aggregated data. The results that follow no doubt hold with some modification under far weaker conditions, particularly concerning temporal dependence as implied by the limit theory in Lieberman and Phillips (2017b & c). Some generality is sacrificed in what follows in order to deliver simpler formulae without compromising the validity of the main findings of the paper.

Assumption 1. $u_t \sim_{iid} \left(0, \mathbb{E}u_t^2 = \sigma_u^2, \mathbb{E}u_t^4 = \mu_{4,u}\right), \varepsilon_t \sim_{iid} \left(0, \mathbb{E}\varepsilon_t^2 = \sigma_\varepsilon^2, \mathbb{E}\varepsilon_t^4 = \mu_{4,\varepsilon}\right),$ *both are symmetrically distributed about zero, and* u_t *is independent of* ε_s *for all* t, s.

Then partial sums of $w_t = (u_t, \varepsilon_t)'$ satisfy the invariance principle

$$n^{-1/2} \sum_{t=1}^{\lfloor n \cdot \rfloor} w_t \Rightarrow B(\cdot) \equiv \mathrm{BM}(\Sigma), \;\; \Sigma = \mathrm{diag}\left(\sigma_u^2, \sigma_\varepsilon^2\right), \tag{2}$$

where $\lfloor \cdot \rfloor$ is the floor function and $B = (B_u, B_\varepsilon)'$ is vector Brownian motion. By Lemma

1 of LP, when $\mu = 0$,

$$\frac{Y_{t=\lfloor nr \rfloor}}{\sqrt{n}} \Rightarrow G_{a,c}(r) := e^{rc+a'B_u(r)} \int_0^r e^{-pc-a'B_u(p)} dB_\varepsilon(p). \tag{3}$$

It is convenient to set $G^*(r) = \sigma_\varepsilon^{-1} G_{a,c}(r) = e^{rc+a'B_u(r)} \int_0^r e^{-pc-a'B_u(p)} dW(p)$ where dependence of G^* on (a, c) is suppressed for notational simplicity and where $W(r)$ is standard Brownian motion.

Sample statistics are often calculated using temporally aggregated data, such as

$$Y_1^m = Y_m, Y_2^m = Y_{2m}, ..., Y_{n/m}^m = Y_n,$$

where m is an aggregation parameter. For example, $m = 5$ for weekly financial data when the original observations are daily. For simplicity in what follows and with no loss of generality we assume that n/m is integer valued.The properties of the model and the limit theory depend on whether or not $\mu = 0$. These cases are therefore analyzed separately.

3 The Case $\mu = 0$

If Y_t is a price process, then the return series of temporally aggregated data created from $\{Y_t^m\}_{t=1}^{n/m}$, is given by

$$\begin{aligned}
\Delta_t^m &= Y_t^m - Y_{t-1}^m \\
&= (Y_{tm} - Y_{tm-1}) + (Y_{tm-1} - Y_{tm-2}) + \cdots + \left(Y_{tm-(m-1)} - Y_{tm-m}\right) \\
&= ((\beta_{tm} - 1) Y_{tm-1} + \varepsilon_{tm}) + \left(\left(\beta_{tm-1} - 1\right) Y_{tm-2} + \varepsilon_{tm-1}\right) \\
&+ \cdots + \left(\left(\beta_{tm-(m-1)} - 1\right) Y_{tm-m} + \varepsilon_{tm-(m-1)}\right) \\
&= \sum_{s=-(m-1)}^{0} \left(\left(\beta_{tm+s} - 1\right) Y_{tm+s-1} + \varepsilon_{tm+s}\right), \ t = 2, ..., n/m. \qquad (4)
\end{aligned}$$

Let $b = (a\sigma_u)^2$ and denote the standardized fourth moments of u and ε by $\rho_{4,u} = \mu_{4,u}/\sigma_u^4$ and $\rho_{4,\varepsilon} = \mu_{4,\varepsilon}/\sigma_\varepsilon^4$. The limit distribution of the sample kurtosis of the m-aggregated data is given in the following result.

Theorem 1 *Under Assumption 1, for the model (1) with $\mu = 0$, as $n \to \infty$ with m*

fixed

$$\gamma_n^m = \frac{\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^4}{\left(\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^2\right)^2} \Rightarrow \gamma^m$$

$$:= 3 + \frac{3b^2\left(\int_0^1 G^{*4}(r)\,dr - \left(\int_0^1 G^{*2}(r)\,dr\right)^2\right)}{\left(1 + b\int_0^1 G^{*2}(r)\,dr\right)^2} + \frac{1}{m}\frac{\left(\rho_{4,\varepsilon} - 3\right) + \left(\rho_{4,u} - 3\right) b^2 \int_0^1 G^{*4}(r)\,dr}{\left(1 + b\int_0^1 G^{*2}(r)\,dr\right)^2}. \tag{5}$$

Evidently from the limit expression (5), γ^m falls as m increase, matching the observed behavior in the kurtosis measures of much financial data. When $b = 0$ (i.e., either $a = 0$ or $\sigma_u = 0$) the limit form γ^m reduces to

$$\gamma^m = 3 + \frac{1}{m}\left(\rho_{4,\varepsilon} - 3\right), \tag{6}$$

which is the result given by Lau and Wingender (1989, eq'n (10)) in the iid case. Otherwise, the limit kurtosis (5) is a random variable. If ε_t is Gaussian

$$\gamma_n^m \Rightarrow 3 + \frac{3b^2\left(\int_0^1 G^{*4}(r)\,dr - \left(\int_0^1 G^{*2}(r)\,dr\right)^2\right)}{\left(1 + b\int_0^1 G^{*2}(r)\,dr\right)^2} + \frac{\left(\rho_{4,u} - 3\right) b^2 \int_0^1 G^{*4}(r)\,dr}{m\left(1 + b\int_0^1 G^{*2}(r)\,dr\right)^2},$$

which still depends on m. Thus, the LSTUR model has the property that the sample kurtosis declines with m whether the error process is Gaussian or otherwise.

Also, irrespective of whether $\rho_{4,\varepsilon} \geq 3$ and since $\int_0^1 G^{*4}(r)\,dr > \left(\int_0^1 G^{*2}(r)\,dr\right)^2$ *a.s.*, which follows as in Phillips and Hansen (1990, lemmas A2 and A3), the model is consistent with $\gamma^m > 3$ whenever $b \neq 0$, which is in line with observed financial index data. In other words, higher kurtosis in the observed process in the LSTUR model is not dependent on Gaussian errors and kurtosis declines as temporal aggregation rises. If the data generating mechanism (1) is STUR rather than LSTUR (i.e., $c = 0$), the result (5) changes only by the form of the limiting G process, with the corresponding limiting STUR process $G_a(r)$ replacing that of the LSTUR process $G_{a,c}(r)$. Thus, these findings apply to both STUR and LSTUR generating mechanisms.

4 The Case $\mu \neq 0$

Under Assumption 1 when $\mu \neq 0$, simple derivations following those for the driftless case show that

$$\frac{Y_t}{n} \Rightarrow H_{a,c}(r) := \mu e^{rc + aB_u(r)} \int_0^r e^{-pc - aB_u(p)} dp. \tag{7}$$

Aggregating in this case leads to

$$\begin{aligned}
\Delta_t^m &= Y_t^m - Y_{t-1}^m \\
&= (Y_{tm} - Y_{tm-1}) + (Y_{tm-1} - Y_{tm-2}) + \cdots + \left(Y_{tm-(m-1)} - Y_{tm-m}\right) \\
&= (\mu + (\beta_{tm} - 1) Y_{tm-1} + \varepsilon_{tm}) + \left(\mu + \left(\beta_{tm-1} - 1\right) Y_{tm-2} + \varepsilon_{tm-1}\right) \\
&+ \cdots + \left(\mu + \left(\beta_{tm-(m-1)} - 1\right) Y_{tm-m} + \varepsilon_{tm-(m-1)}\right) \\
&= m\mu + \sum_{s=-(m-1)}^{0} \left(\left(\beta_{tm+s} - 1\right) Y_{tm+s-1} + \varepsilon_{tm+s}\right),
\end{aligned} \tag{8}$$

for $t = 2, ..., n/m$. Let $H(r) = \mu H^*(r) := H_{a,c}(r)$, where for brevity we omit the parameter dependencies in $H_{a,c}(r)$.

Theorem 2 *Under Assumption 1 for the model (1) with $\mu \neq 0$,*

$$\gamma_n^{m,\mu} = \frac{\frac{m}{n}\sum_{t=2}^{n/m} \left(\Delta_t^m\right)^4}{\left(\frac{m}{n}\sum_{t=2}^{n/m} \left(\Delta_t^m\right)^2\right)^2} \Rightarrow \gamma^{m,\mu} := \frac{\left(\frac{1}{m}\left(\rho_{4,u} - 3\right) + 3\right) \int_0^1 H^{*4}(r)\, dr}{\left(\int_0^1 H^{*2}(r)\, dr\right)^2}. \tag{9}$$

As $n \to \infty$ followed by $a \to 0$ (so that lower order terms are eliminated), the limit (9) becomes

$$\frac{\left(\frac{1}{m}\left(\rho_{4,u} - 3\right) + 3\right) \int_0^1 (\mu r)^4}{\left(\int_0^1 (\mu r)^2\right)^2} = \frac{\left(\frac{1}{m}\left(\rho_{4,u} - 3\right) + 3\right) \int_0^1 r^4}{\left(\int_0^1 r^2\right)^2} = \frac{9\left(\frac{1}{m}\left(\rho_{4,u} - 3\right) + 3\right)}{5},$$

which equals 5.4 when the STUR variable u_t is Gaussian.

5 Effects of Misspecification

The use of a simple fitted AR(1) regression involves misspecifying the LSTUR model as a fixed coefficient autoregression. As we have seen in Sections 3 and 4, the local and stochastically local to unity specification leads to analytic formulae for the excess kurtosis and in Section 8 these formulae will be shown to match closely direct computations of kurtosis in the observed data. Similar close correspondence will be found in the case of fitted values from a simple AR (1) regression. The explanation for this phenomenon is that the misspecification in the fitted autoregression involves an error of $O_p(n^{-1})$ and this error is sufficiently small to ensure the fitted AR kurtosis is asymptotically equivalent to that of the data. Importantly, however, the AR(1) model does not explain the source of the excess kurtosis and only provides an analytic asymptotic formula for the kurtosis that is based on the underlying LSTUR model.

We denote the least squares estimators of β in a fitted AR(1) model under $\mu = 0$ and $\mu \neq 0$ by $\hat{\beta} = \sum_{t=2}^{n} Y_t Y_{t-1} / \sum_{t=2}^{n} Y_{t-1}^2$ and $\hat{\beta}^{\mu} = \sum_{t=2}^{n} \tilde{Y}_t \tilde{Y}_{t-1} / \sum_{t=2}^{n} \tilde{Y}_{t-1}^2$, respectively, where $\tilde{Y}_t := Y_t - \bar{Y}$ and $\bar{Y} := n^{-1} \sum_{t=1}^{n} Y_t$. By $\hat{\mu} = \bar{Y} - \hat{\beta}\bar{Y}_{-1}$ we denote the least squares estimator of μ in the fitted AR(1) model under $\mu \neq 0$, where $\bar{Y}_{-1} = n^{-1} \sum_{t=1}^{n-1} Y_t$. Finally, we let $\hat{Y}_t = \hat{\beta} Y_{t-1}$ in the case $\mu = 0$, and $\hat{Y}_t^{\mu} = \hat{\mu} + \hat{\beta}^{\mu} Y_{t-1}$ in the case $\mu \neq 0$, with the associated differences $\hat{\Delta}_t = \hat{Y}_t - \hat{Y}_{t-1}$, $\hat{\Delta}_t^m = \hat{Y}_t^m - \hat{Y}_{t-1}^m$, $\hat{\Delta}_t^{\mu} = \hat{Y}_t^{\mu} - \hat{Y}_{t-1}^{\mu}$ and $\hat{\Delta}_t^{m,\mu} = \hat{Y}_t^{m,\mu} - \hat{Y}_{t-1}^{m,\mu}$.

Theorem 3 *Under Assumption 1, for the model (1), fitted data from a misspecified AR(1) model have kurtosis coefficients and limiting kurtosis as follows:*
(i) When $\mu = 0$, the sample kurtosis which is based on a fitted AR(1) model that does not include an intercept is

$$\gamma_{AR,n}^{m} := \frac{\frac{m}{n} \sum_{t=2}^{n/m} \left(\hat{\Delta}_t^m \right)^4}{\left(\frac{m}{n} \sum_{t=2}^{n/m} \left(\hat{\Delta}_t^m \right)^2 \right)^2} = \frac{\hat{\beta}^4 \frac{m}{n} \sum_{t=2}^{n/m} \left(\Delta_t^m \right)^4}{\hat{\beta}^4 \left(\frac{m}{n} \sum_{t=2}^{n/m} \left(\Delta_t^m \right)^2 \right)^2} \Rightarrow \gamma^m,$$

where γ^m is given in (5).
(ii) When $\mu \neq 0$, the sample kurtosis which is based on a fitted AR(1) model that includes an intercept is

$$\gamma_{AR,n}^{m,\mu} := \frac{\frac{m}{n} \sum_{t=2}^{n/m} \left(\hat{\Delta}_t^{m,\mu} \right)^4}{\left(\frac{m}{n} \sum_{t=2}^{n/m} \left(\hat{\Delta}_t^{m,\mu} \right)^2 \right)^2} \Rightarrow \gamma^{m,\mu},$$

where $\gamma^{m,\mu}$ is given by (9).
(iii) When $\mu \neq 0$, the least squares estimator of μ in a fitted AR(1) satisfies

$$\hat{\mu} \Rightarrow \mu + a \int_0^1 G(s)\, dB_u(s) - \frac{b}{2} \int_0^1 G(s)\, ds.$$

These results show that the constant coefficient AR fitted kurtosis simply reproduces the empirical sample kurtosis in models with and without drift, provided an intercept is fitted in modeling the data in the case with drift. In the latter case, the result holds in spite of the fact that μ is inconsistently estimated, as shown in part (iii) of Theorem 3. Part (iii) of the theorem will be shown to be particularly useful in Section below becauses it anticipates an important practical distinction between the empirical fitting of an AR(1) model and an LSTUR model.

6 Continuous Time Measures of Kurtosis

6.1 The case of zero drift

When $\mu = 0$, an instantaneous kurtosis measure for the process increments $dG(r)$ at r can be obtained using the stochastic differential equation representation

$$dG(r) = aG(r)\,dB_u(r) + dB_\varepsilon(r) + \left(c + \frac{b}{2}\right) G(r)\,dr. \tag{10}$$

As indicated in the result that follows, we may define instantaneous kurtosis as in (11) in terms of the conditional moments of the increment process $dG(r)$ in (10).

Theorem 4 *For the process (10),*

$$\kappa_{b,c}(r) := \frac{\mathbb{E}\left(\mathbb{E}\left[(dG(r))^4 \,|\mathcal{F}_r\right]\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(dG(r))^2 \,|\mathcal{F}_r\right]\right)\right\}^2} \tag{11}$$

$$= 3\left\{1 + \frac{b^2\left[\mathbb{E}\left(G(r)^4\right) - \left(\mathbb{E}\left(G(r)^2\right)\right)^2\right]}{b^2\left(\mathbb{E}\left(G(r)^2\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2\mathbb{E}\left(G(r)^2\right)}\right\} + o(1). \tag{12}$$

The second term in braces in (12) shows the excess kurtosis in the process increments arising from the non-Gaussianity of $G(r)$. As $b \to 0$ evidently $\kappa_{b,c} \to 3$, as expected since $G(r) \to J_c(r) = \int_0^r e^{(r-p)c} dB_\varepsilon(p)$ which is Gaussian in this case. But when $c \to 0$, $G(r) \to G_a(r) = e^{aB_u(r)} \int_0^r e^{-a'B_u(p)} dB_\varepsilon(p)$, which is non-Gaussian and then $\kappa_{b,0} > 3$. For large b, after some calculation, we have

$$\kappa_{b,0}(r) = \frac{3\mathbb{E}\left(G(r)^4\right)}{\left(\mathbb{E}\left(G(r)^2\right)\right)^2} + o(1) \sim \frac{9}{6}e^{4br},$$

and the kurtosis of the process increments $dG(r)$ grows exponentially with b irrespective of the fixed value of c.

6.2 Local to zero drift

We next consider the case in which the limit of the standardized discrete time model $Y_t/\sqrt{n}$ has a discrete time drift comparable in magnitude to the stochastic term,which occurs when the discrete model has drift local to zero of the form $\mu{=}\alpha/\sqrt{n}$. The stochastic differential equation for G(r) in this case is

$$dG(r) = aG(r)\,dB_u(r) + dB_\varepsilon(r) + \left[\alpha + \left(c + \frac{b}{2}\right) G(r)\right] dr, \tag{13}$$

whose solution is given by Föllmer and Schweizer (1993; theorem 3.5). With initial condition $G(0) = 0$, this solution has the following form

$$G(r) = \alpha e^{rc+aB_u(r)} \int_0^r e^{-pc-aB_u(p)} dp + e^{rc+aB_u(r)} \int_0^r e^{-pc-aB_u(p)} dB_\varepsilon(p). \qquad (14)$$

Define $X = c + b/2$ and $D(r) = bG^2(r) + \sigma_\varepsilon^2$.

Theorem 5 *For the process (13),*

$$\begin{aligned} \kappa_{b,c,\alpha} &= \frac{\mathbb{E}\left(\mathbb{E}\left[(dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r))^4 |\mathcal{F}_r\right]\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r))^2 |\mathcal{F}_r\right]\right)\right\}^2} \\ &= 3 + \frac{3b^2 Var(G^2(r)) + 6\mathbb{E}\left[\{b\{G^2(r) - \mathbb{E}(G^2(r))\}\} X^2 \{G(r) - \mathbb{E}(G(r))\}^2\right](dr)}{\{\mathbb{E}(D(r))\}^2 + 2\mathbb{E}(D(r)) X^2 Var(G(r))(dr)} + O\left((dr)^2\right). \end{aligned}$$

To first order it again follows that $\kappa_{b,c,\alpha} > 3$ provided $b \neq 0$, and as $b \to 0$ evidently $\kappa_{b,c,a} \to 3$.

6.3 Dominating drift term

When the drift is fixed rather than local to zero in the discrete time model (1) the resulting trend in the time series dominates asymptotically. In this case the limit process corresponding to $Y_{t=\lfloor nr \rfloor}/n$ is

$$G(r) = \mu e^{rc+aB_u(r)} \int_0^r e^{-pc-aB_u(p)} dp,$$

in place of (14). Then, $\mathbb{E}(dG(r)) = \{\mu + X\mathbb{E}G(r)\} dr = \left\{\mu + \left(e^{Xr} - 1\right)\right\} dr$,

$$\begin{aligned} dG(r) &= cG(r)dr + \mu e^{rc+aB_u(r)} \int_0^r e^{-pc-aB_u(p)} dp a dB_u(r) \\ &\quad + \frac{1}{2}\mu e^{rc+aB_u(r)} \int_0^r e^{-pc-aB_u(p)} dp a^2 \sigma_u^2 dr + \mu dr \\ &= \{XG(r) + \mu\} dr + aG(r) dB_u(r), \end{aligned}$$

and $\mathbb{E}(dG(r)|\mathcal{F}_r) = \{XG(r) + \mu\} dr$, so that $dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r) = aG(r)dB_u(r)$. It follows that the instantaneous kurtosis at r is given by

$$\begin{aligned} \kappa_{b,c,\mu}(r) &= \frac{\mathbb{E}\left(\mathbb{E}\left[(dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r))^4 |\mathcal{F}_r\right]\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r))^2 |\mathcal{F}_r\right]\right)\right\}^2} \\ &= \frac{\mathbb{E}\left(a^4 G(r)^4 \mathbb{E}\left[(dB_u(r))^4 |\mathcal{F}_r\right]\right)}{\left\{\mathbb{E}\left(a^2 G(r)^2 \mathbb{E}\left[dB_u(r)^2 |\mathcal{F}_r\right]\right)\right\}^2} = 3\frac{\mathbb{E}\left[G(r)^4\right]}{\left\{\mathbb{E}\left[G(r)^2\right]\right\}^2}. \end{aligned}$$

Observe that when $X = 0$ and $a \to 0$, $G(r) \to \mu r$ is deterministic and $\kappa_{b,c,\mu} \to 3$.

By contrast, we may consider the kurtosis of the averaged relative increments $\frac{dG(r)-\mathbb{E}(dG(r)|\mathcal{F}_r)}{\sqrt{dr}}$ over an interval such as $[0,q]$, defined as

$$\begin{aligned}
\bar{\kappa}_{b,c,\mu}(q) &= \frac{\mathbb{E}\left(\frac{1}{q}\int_0^q \mathbb{E}\left[\left(\frac{dG(r)-\mathbb{E}(dG(r)|\mathcal{F}_r)}{\sqrt{dr}}\right)^4 |\mathcal{F}_r\right] dr\right)}{\left\{\mathbb{E}\left(\frac{1}{q}\int_0^q \mathbb{E}\left[\left(\frac{dG(r)-\mathbb{E}(dG(r)|\mathcal{F}_r)}{\sqrt{dr}}\right)^2 |\mathcal{F}_r\right] dr\right)\right\}^2} \\
&= 3\frac{\mathbb{E}\left(\frac{1}{q}\int_0^q a^4\sigma_u^4 G(r)^4 dr\right)}{\left\{\mathbb{E}\left(\frac{1}{q}\int_0^q a^2\sigma_u^2 G(r)^2 dr\right)\right\}^2} = 3\frac{\frac{1}{q}\int_0^q \mathbb{E}\left[G(r)^4\right] dr}{\left(\frac{1}{q}\int_0^q \mathbb{E}\left[G(r)^2\right] dr\right)^2}
\end{aligned}$$

In this case when $X = 0$ and $a \to 0$, $G(r) \to \mu r$ and

$$\bar{\kappa}_{b,c,\mu}(q) \to \frac{\mu^4 \frac{1}{q}\int_0^q r^4 dr}{\mu^4\left(\frac{1}{q}\int_0^q r^2 dr\right)^2} = 3\frac{\frac{1}{5}\mu^4 q^4}{\left(\frac{1}{3}q^2\right)^2} = \frac{27}{5} = 5.4,$$

which can also be obtained by a direct calculation from the discrete time model in this special case.

7 Simulations

The purpose of this section is to corroborate some of the analytical results. To this end we have simulated a driftless STUR model with parameter settings $a = 0.7$, $u_t \overset{iid}{\sim} t(6)$ (so that $\sigma_u^2 = 1.5$ and $\rho_{4,u} = 6$), and $\varepsilon_t \overset{iid}{\sim} N(0,1)$, as well as a drifting STUR model with parameter $\mu = 2.5$. In addition, an LSTUR model with $c = -b$ was simulated with the same settings for both drift and driftless cases.

For each model we compared the mean and standard deviation of the sample kurtosis with those of the asymptotic formulae given in (5) and (9), corresponding to the zero mean and drift cases. We also constructed PP-plots of the finite sample distributions of the sample kurtosis against the limit distributions in each case. The simulation design comprised temporal average settings with $m = 1, 3, 5$, sample sizes $n = 9,000$, $21,000$, $48,000$, $99,000$, $240,000$, 500 integral points and $5,000$ replications. The results are summarized in Tables 1-4 and Figures 1-4.[1]

Inspection of the figures reveals that the PP-plots move closer to the 45 degree line as the sample size increases, irrespective of whether the generating mechanism contains a drift or otherwise. This movement is also clear in the tables. When the model contains a drift (and less so in the driftless case), there is a marked shift near the origination point

[1] For brevity only four figures are displayed. Other cases deliver very similar conclusions.

in the PP-plots which reflects the fact that the finite sample distribution of the sample kurtosis has a non-zero probability of being less than 3, whereas for the asymptotic distribution this probability is zero. The shift at zero becomes smaller as n increases, but even with $240,000$ observations the discontinuity does not entirely vanish. Overall, the simulations corroborate the analytic findings on the kurtosis patterns associated with these non-Gaussian models.

8 An Empirical Application

This section explores how well the kurtosis theory developed above is supported by observed return data on various indices. In particular, we demonstrate that the sample kurtosis of the observed data match those which are based on the fitted LSTUR-estimates and decline with the frequency of the data, just as in direct calculations from the data. Using Theorem 3(iii), through the drift parameter estimates, evidence will be given in favour of the LSTUR specification and against the traditional fixed coefficient autoregression. The data employed in these comparisons is now briefly described.

8.1 Data

The data for the empirical application comprises Exchange Traded Fund (ETF) and Exchange Traded Note (ETN) data obtained from the following sources. ETF closing prices data were retrieved from Yahoo Finance, covering the period from January 2010 to December 2017, giving a total of 2013, 417 and 95 observations for the daily, weekly and monthly frequencies, respectively. The US equity ETF data includes SPDR S&P 500 ETF (SPY) , SPDR S&P 600 Small Cap ETF (SLY) and SPDR S&P MidCap 400 ETF (MDY). IShares MSCI Emerging Markets ETF (EEM) and iShares Global 100 ETF (IOO) were also used, the first seeking to track the investment results of the SCI Emerging Markets and the second - the S&P Global 100 indices.

The bond ETF data includes iShares Core US Aggregate Bond ETF (AGG), iShares 1-3 Year Treasury Bond ETF (SHY), iShares iBoxx investment grade corporate bond ETF (LQD) and SPDR Bloomberg Barclays high yield bond ETF (JNK) closing price series.

In addition, we have used the ETFs and ETNs replicating the prices of commodities. These include SPDR Gold Shares (GLD), iShares Silver Trust (SLV), iPath S&P GSCI crude oil ETN (OIL) and Rogers International Agriculture commodity total return index ETN (RJA).

Finally, we have also used the Currency Shares Euro ETF (FXE), which tracks changes in value of the Euro relative to the US dollar.

8.2 Results

LSTUR and simple AR(1) models with and without drift were estimated for the empirical data according to whether they exhibited trend. The LSTUR model was estimated under the restriction $c + b = 0$. For simplicity, for each realization of the LSTUR variates, u_t was generated as $u_t \overset{iid}{\sim} N(0,1)$ with 2000 replications. Average estimates over the replications are reported in Table 5 for a model with fitted drift and in Table 6 for a model without drift.

For the model fitted with a drift, the kurtosis estimates from the observed return data, from the returns based on the fitted LSTUR-estimates, and from the returns based on a fitted AR(1) model, are all very close to each other and all decline with the frequency of the data. The kurtosis estimates for the model fitted without a drift exhibit a similar pattern, decreasing as the frequency increases, as expected.

The findings also corroborate Theorem 3, which shows that the kurtosis estimate based on a fitted AR(1) model is first order equivalent to that based on an LSTUR model. Importantly, the LSTUR-based drift parameter estimates are very close to the actual return means in all cases, whereas those based on the AR(1) model are evidently biased upwards, again corresponding to the asymptotic results. Indeed, this feature of the empirical results matches the finding in part (iii) of Theorem 3, which shows that when LSTUR is the generating mechanism and the fitted model is misspecified as an AR(1), the drift parameter estimate from the misspecified autoregression is inconsistent, which is further evidence of the usefulness the local stochastic unit root specification over traditional fixed coefficient autoregression.

9 Conclusions

Asset price data are typically well described as martingales and their returns as martingale differences, thereby capturing the prominent feature of near-unpredictability. A secondary feature of great practical importance is their characteristic peaked and heavy-tailed distribution, with high kurtosis that steadily declines towards the Gaussian value of 3 with increasing temporal aggregation of the returns. The present results reveal that all of these features are captured by stochastic unit root models. The nonlinear stochastic nature of these models induces non-Gaussian behavior even when the first two moments correspond closely to a simple process like a Gaussian random walk. Most notably, the asymptotic behavior of sample kurtosis measures from these variants of STUR models and their continuous time analogues and associated kurtosis measures all mimick the defining characteristics of observed financial returns. Several empirical applications to a variety of financial indices confirm these useful capabilities.

References

Bykhovskaya, A. and P. C. B. Phillips (2019) "Point Optimal Testing with Roots that are Functionally Local to Unity", *Journal of Econometrics* (forthcoming).

Bykhovskaya, A. and P. C. B. Phillips (2018)"Boundary Limit Theory for Functional Local to Unity Regression", *Journal of Time Series Analysis,* 39, 523-562.

Chan, N. H. and C. Z. Wei (1987). "Asymptotic inference for nearly nonstationary AR(1) processes," *Annals of Statistics* 15, 1050–1063.

Föllmer, H. and M. Schweizer (1993). "A microeconomic approach to diffusion models for stock prices," *Mathematical Finance* 3(1), 1–23.

Granger, C. W. and N. R. Swanson (1997). "An introduction to stochastic unit-root processes," *Journal of Econometrics 80(1), 35–62.*

Ibragimov, R. and P. C. B. Phillips (2008). "Regression asymptotics using martingale convergence methods," *Econometric Theory* 24, 888–947.

Lau, H. S. and J. R. Wingender (1989). "The analytics of the intervaling effect on skewness and kurtosis of stock returns," *The Financial Review* 24(2) 215–233.

Leybourne, S. J., McCabe, B. P., and T. C. Mills (1996). "Randomized unit root processes for modelling and forecasting financial time series: theory and applications," *Journal of Forecasting* 15(3), 253-270.

Leybourne, S. J., McCabe, B. P. and A. R. Tremayne (1996). "Can economic time series be differenced to stationarity?," *Journal of Business and Economic Statistics* 14(4), 435-446.

Lieberman, O. (2012). "A similarity-based approach to time-varying coefficient nonstationary autoregression," *Journal of Time Series Analysis* 33, 484–502.

Lieberman, O. and P. C. B. Phillips (2014). "Norming rates and limit theory for some time-varying coefficient autoregressions," *Journal of Time Series Analysis* 35, 592–623.

Lieberman, O. and P. C. B. Phillips (2017a). "A multivariate stochastic unit root model with an application to derivative pricing," *Journal of Econometrics,* 196, 99-110.

Lieberman, O. and P. C. B. Phillips (2017b). "IV and GMM Inference in Endogenous Stochastic Unit Root Models". *Econometric Theory,* doi.org/10.1017/S0266466617000330.

Lieberman, O. and P. C. B. Phillips (2017c). "Hybrid stochastic local unit roots". Cowles Foundation Discussion Paper 2113.

McCabe, B. P. and R. J. Smith (1998). "The power of some tests for difference stationarity under local heteroscedastic integration," *Journal of the American Statistical Association* 93(442), 751-761.

Phillips, P. C. B. (1987). "Towards a unified asymptotic theory for autoregression," *Biometrika* 74, 535–547.

Phillips, P. C. B. (2014). "On confidence intervals for autoregressive roots and predictive regression," *Econometrica* 82(3), 1177-1195.

Phillips, P. C. B. and B. E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies* 57, 99–125.

Tao, Y., P. C. B. Phillips and J. Yu (2017). "Random coefficient continuous systems: testing for extreme sample path behavior", Working paper, Singapore Management University.

Yoon, G. (2006). "A note on some properties of STUR processes," *Oxford Bulletin of Economics and Statistics* 68(2), 253-260.

Table 1. STUR with a drift

n		Δ^1	γ^1	Δ^3	γ^3	Δ^5	γ^5
9000	Mean	12.40	12.84	8.37	8.56	7.57	7.64
	Std. dev.	7.34	3.58	3.42	2.39	3.03	2.09
21000	Mean	12.73	12.93	8.50	8.62	7.61	7.71
	Std. dev.	10.62	3.67	4.66	2.45	2.62	2.18
48000	Mean	12.64	12.85	8.44	8.57	7.66	7.45
	Std. dev.	7.87	3.50	3.20	2.33	2.39	2.20
99000	Mean	12.79	12.84	8.53	8.56	7.57	7.66
	Std. dev.	6.42	3.70	3.05	2.46	2.28	2.14
240000	Mean	12.74	12.81	8.51	8.54	7.72	7.72
	Std. dev.	4.58	3.61	2.58	2.41	2.20	2.17

Notes: calculated with 5000 replications and 500 integral points, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $\mu = 2.5$, $a = 0.7$.

Table 2. STUR without a drift

n		Δ^1	γ^1	Δ^3	γ^3	Δ^5	γ^5
9000	Mean	4.15	4.18	3.66	3.67	3.57	3.58
	Std. dev.	2.47	1.93	1.31	1.14	1.07	1.01
21000	Mean	4.2	4.23	3.68	3.7	3.56	3.6
	Std. dev.	2.36	1.97	1.2	1.16	1.05	1.06
48000	Mean	4.18	4.21	3.68	3.69	3.57	3.59
	Std. dev.	2.13	2.01	1.17	1.19	0.97	1
99000	Mean	4.2	4.21	3.68	3.69	3.58	3.58
	Std. dev.	2.34	2.01	1.26	1.19	0.95	0.99
240000	Mean	4.18	4.22	3.68	3.7	4.18	4.19
	Std. dev.	2.14	1.94	1.21	1.15	3.57	3.58

Notes: calculated with 5000 replications and 500 integral points, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $a = 0.7$.

Table 3. LSTUR with a drift

n		Δ^1	γ^1	Δ^3	γ^3	Δ^5	γ^5
9000	Mean	11.55	11.58	7.68	7.72	6.87	6.96
	Std. dev.	12.60	2.93	4.67	1.95	3.29	1.74
21000	Mean	11.34	11.61	7.62	7.74	6.80	6.96
	Std. dev.	7.21	2.96	2.95	1.97	2.02	1.83
48000	Mean	11.48	11.62	7.69	7.75	6.92	6.98
	Std. dev.	5.89	2.97	3.13	1.98	2.03	1.79
99000	Mean	11.52	11.52	7.70	7.68	6.98	6.97
	Std. dev.	5.15	2.84	2.47	1.89	2.92	1.78
240000	Mean	11.52	11.63	7.69	7.75	6.91	6.92
	Std. dev.	4.32	2.99	2.26	1.99	1.77	1.75

Notes: calculated with 5000 replications and 500 integral points, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $\mu = 2.5$, $a = 0.7$, $c = -b$.

Table 4. LSTUR without a drift

n		Δ^1	γ^1	Δ^3	γ^3	Δ^5	γ^5
9000	Mean	3.7	3.69	3.39	3.39	3.32	3.34
	Std. dev.	1.62	1.179	0.84	0.68	0.68	0.62
21000	Mean	3.71	3.71	3.39	3.4	3.39	3.38
	Std. dev.	1.92	1.19	0.88	0.7	0.74	0.69
48000	Mean	3.71	3.71	3.4	3.41	3.33	3.32
	Std. dev.	2.08	1.26	0.81	0.74	0.61	0.57
99000	Mean	3.7	3.71	3.4	3.41	3.35	3.35
	Std. dev.	1.3	1.29	0.75	0.76	0.69	0.63
240000	Mean	3.69	3.69	3.39	3.39	3.34	3.34
	Std. dev.	1.34	1.18	0.74	0.69	0.64	0.63

Notes: calculated with 5000 replications and 500 integral points, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $a = 0.7$, $c = -b$.

Table 5. Estimates for the model with a drift

Ticker	Frequency	$\bar{\Delta y}$	$\hat{\mu}_{n,LSTUR}$	$\hat{\mu}_{n,OLS}$	$\hat{\beta}_{n,OLS}$	$\hat{a}_{n,LSTUR}$	$\kappa_{\Delta y}$	$\kappa_{\Delta \hat{y},LSTUR}$	$\kappa_{\Delta \hat{y},OLS}$
SPY	daily	0.0004	0.0004	0.0031	0.9995	-0.0029	7.622	7.594	7.620
SPY	weekly	0.0021	0.0021	0.0124	0.9980	0.0023	5.051	5.035	5.042
SPY	monthly	0.0096	0.0095	0.0560	0.9910	0.0008	3.478	3.442	3.442
SLY	daily	0.0005	0.0005	0.0063	0.9987	0.0019	6.574	6.569	6.573
SLY	weekly	0.0022	0.0022	0.0280	0.9940	0.0029	4.914	4.899	4.904
SLY	monthly	0.0101	0.0096	0.1230	0.9750	0.0045	3.330	3.329	3.333
MDY	daily	0.0005	0.0005	0.006	0.9990	-0.0034	7.916	7.886	7.915
MDY	weekly	0.0023	0.0023	0.0250	0.9960	-0.0030	5.772	5.744	5.760
MDY	monthly	0.0105	0.0100	0.1080	0.9820	-0.0050	3.640	3.588	3.612
IOO	daily	0.0002	0.0002	0.0093	0.998	-0.0015	7.822	7.815	7.818
IOO	weekly	0.0010	0.0010	0.0360	0.9920	-0.0120	5.028	4.904	5.017
IOO	monthly	0.0051	0.0049	0.1440	0.967	-0.0043	3.369	3.331	3.336
AGG	daily	0.0000	0.0000	0.0390	0.9920	0.0002	4.805	4.803	4.803
AGG	weekly	0.0001	0.0001	0.1780	0.9620	0.0005	4.341	4.335	4.336
AGG	monthly	0.0005	0.0005	0.5900	0.8740	-0.0027	3.656	3.573	3.618
LQD	daily	0.0001	0.0001	0.0230	0.9950	0.0017	5.006	4.982	5.004
LQD	weekly	0.0004	0.0004	0.1130	0.9760	-0.0001	4.882	4.880	4.880
LQD	monthly	0.0015	0.0012	0.4410	0.9080	-0.0038	3.218	3.184	3.201
OIL	daily	-0.0007	-0.0007	0.0003	0.9990	0.0110	7.504	7.480	7.500
OIL	weekly	-0.0033	-0.0033	0.0004	0.9990	0.0027	3.923	3.925	3.926
OIL	monthly	-0.0134	-0.0130	0.0002	0.9950	-0.0360	3.368	3.327	3.360

Table 5 Continued

Ticker	Frequency	$\bar{\Delta y}$	$\hat{\mu}_{n,LSTUR}$	$\hat{\mu}_{n,OLS}$	$\hat{\beta}_{n,OLS}$	$\hat{a}_{n,LSTUR}$	$\kappa_{\Delta y}$	$\kappa_{\Delta\hat{y},LSTUR}$	$\kappa_{\Delta\hat{y},OLS}$
SLV	daily	0.0000	−0.0001	0.0054	0.9980	0.0120	9.711	9.656	9.707
SLV	weekly	−0.0001	0.0002	0.0270	0.9910	0.0370	13.391	12.720	13.397
SLV	monthly	0.0001	0.0000	0.1340	0.9960	0.0036	4.040	4.003	4.004
RJA	daily	−0.0001	−0.0001	0.0016	0.9990	−0.0028	6.700	6.690	6.700
RJA	weekly	−0.0007	−0.0007	0.007	0.9960	−0.0009	4.895	4.887	4.887
RJA	monthly	−0.002	−0.002	0.051	0.9740	−0.0024	4.452	4.407	4.407
FXE	daily	−0.0001	−0.0001	0.0100	0.9980	−0.0017	4.483	4.477	4.483
FXE	weekly	−0.0005	−0.0005	0.0500	0.9900	−0.0013	3.451	3.452	3.453
FXE	monthly	−0.0019	−0.0018	0.2010	0.9580	0.0012	3.617	3.581	3.582

Notes: u_t is simulated as IID $N(0,1)$; $\hat{\mu}_{n,LSTUR}$ and $\hat{a}_{n,LSTUR}$ are the means of the MLEs of LSTUR in 2000 replications; $\hat{\mu}_{n,OLS}$ and $\hat{\beta}_{n,OLS}$ are the OLS intercept and slope estimates; $\kappa_{\Delta\hat{y},LSTUR}$ is the mean kurtosis of $\Delta\hat{y}_{LSTUR}$ in 2000 replications; $\kappa_{\Delta\hat{y},OLS}$ is the kurtosis of $\Delta\hat{y}_{OLS}$.

Table 6. Estimates for the model without a drift

Ticker	Frequency	$\bar{\Delta y}$	$\hat{a}_{n,LSTUR}$	$\kappa_{\Delta y}$	$\kappa_{\Delta\hat{y},LSTUR}$	$\kappa_{\Delta\hat{y},OLS}$
EEM	daily	0.0001	−0.0020	6.232	6.227	6.230
EEM	weekly	0.0003	−0.0110	4.668	4.628	4.660
EEM	monthly	0.0022	−0.0140	4.552	4.556	4.513
SHY	daily	0.0000	−0.0010	5.149	5.144	5.147
SHY	weekly	0.0000	0.0000	4.350	4.344	4.345
SHY	monthly	0.0000	0.0005	3.685	3.680	3.713
JNK	daily	0.0000	0.0006	11.632	11.624	11.627
JNK	weekly	−0.0001	0.0030	6.770	6.729	6.754
JNK	monthly	−0.0006	0.0050	4.775	4.667	4.735
GLD	daily	0.0001	−0.0004	8.483	8.481	8.481
GLD	weekly	0.0003	0.0030	4.040	4.040	4.042
GLD	monthly	0.0016	−0.0030	2.817	2.795	2.795

Notes: u_t is simulated as IID $N(0,1)$; $\hat{a}_{n,LSTUR}$ is the mean of the MLE of a in LSTUR in 2000 replications; $\kappa_{\Delta\hat{y},LSTUR}$ is the mean kurtosis of $\Delta\hat{y}_{LSTUR}$ in 2000 replications; $\kappa_{\Delta\hat{y},OLS}$ is the kurtosis of $\Delta\hat{y}_{OLS}$; $\hat{\beta}_{n,OLS} \simeq 1$ for all indices (the deviations from unity are $\leq 10^{-4}$ in all cases).

Proofs

Proof of Theorem 1. Using (4)

$$\begin{aligned}\bar{\Delta}^m &= \frac{m}{n}\sum_{t=2}^{n/m}\Delta_t^m = \frac{m}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\left(\left(\beta_{tm+s}-1\right)Y_{tm+s-1}+\varepsilon_{tm+s}\right) \\ &= \frac{m}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\frac{au_{tm+s}}{\sqrt{n}}Y_{tm+s-1}+o_p\left(1\right)\Rightarrow 0,\end{aligned}$$

so that

$$\begin{aligned}\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^2 = \frac{m}{n}\sum_{t=2}^{n/m}&\left(\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^2+\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2\right) \\ &+\frac{2m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right). \qquad (15)\end{aligned}$$

The first term in (15) is

$$\begin{aligned}\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^2 &= \frac{m}{n}\left\{\left(\varepsilon_{m+1}+\cdots+\varepsilon_{2m}\right)^2+\left(\varepsilon_{2m+1}+\cdots+\varepsilon_{3m}\right)^2+\right. \\ &\quad \left. ...+\left(\varepsilon_{n-(m-1)}+\cdots+\varepsilon_n\right)^2\right\} \\ &= \frac{m}{n}\sum_{t=m+1}^{n}\varepsilon_t^2+o_p\left(1\right)\to_p m\sigma_\varepsilon^2.\end{aligned}$$

The second term in (15) is

$$
\begin{aligned}
& \frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2 \\
= \quad & \frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\frac{c}{n}+\frac{au_{tm+s}}{\sqrt{n}}+\frac{\left(au_{tm+s}\right)^2}{2n}\right)Y_{tm+s-1}\right)^2+o_p\left(1\right) \\
= \quad & \frac{m}{n}\frac{a^2}{n}\left\{\left(u_{m+1}Y_m+\cdots+u_{2m}Y_{2m-1}\right)^2\right. \\
& \left.+\left(u_{2m+1}Y_{2m}+\cdots+u_{3m}Y_{3m-1}\right)^2+...+\left(u_{n-(m-1)}Y_{n-m}+\cdots+u_nY_{n-1}\right)^2\right\}+o_p\left(1\right) \\
\Rightarrow \quad & mb\int_0^1 G^2\left(r\right)dr.
\end{aligned}
$$

The third term in (15) is

$$
\begin{aligned}
& \frac{2m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right) \\
= \quad & \frac{2ma}{n^{3/2}}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}u_{tm+s}Y_{tm+s-1}\right)+o_p\left(1\right)\Rightarrow 0.
\end{aligned}
$$

Hence, $\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^2 \Rightarrow m\sigma_\varepsilon^2\left(1+\frac{b}{\sigma_\varepsilon^2}\int_0^1 G^2\left(r\right)dr\right)$. Now

$$
G\left(r\right)=e^{aB_u(r)}\int_0^r e^{-aB_u(p)}dB_\varepsilon\left(p\right)=\sigma_\varepsilon e^{aB_u(r)}\int_0^r e^{-aB_u(p)}dW\left(p\right)=\sigma_\varepsilon G^*\left(r\right), \text{ say}
$$

and so $\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^2 \Rightarrow m\sigma_\varepsilon^2\left(1+b\int_0^1 G^{*2}(r)\,dr\right)$. Continuing,

$$\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^4=\frac{m}{n}\sum_{t=2}^{n/m}\left\{\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^4+4\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^3\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)\right.$$
$$+6\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^2\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2$$
$$\left.+4\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^3+\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^4\right\}, \tag{16}$$

which leads to

$$\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^4=\frac{m}{n}\left(\sum_{t=m+1}^{n}\varepsilon_t^4+\sum_{t=2}^{n/m}\sum_{\substack{s,s'=-(m-1)\\ s\neq s'}}^{0}\binom{4!}{2!2!}\varepsilon_{tm+s}^2\varepsilon_{tm+s'}^2\right)+o_p(1).$$

The first term in the last equation converges in probability to $m\mu_{4,\varepsilon}$. For each $t = 2,...,n/m$, there are $m(m-1)/2$ cross products $\varepsilon_{tm+s}^2\varepsilon_{tm+s'}^2$, $s \neq s'$, and so

$$6\frac{m}{n}\sum_{t=2}^{n/m}\sum_{\substack{s,s'=-(m-1)\\ s\neq s'}}^{0}\varepsilon_{tm+s}^2\varepsilon_{tm+s'}^2 \to_p 3m(m-1)\sigma_\varepsilon^4.$$

Therefore

$$\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^4 \to_p m\mu_{4,\varepsilon}+3m(m-1)\sigma_\varepsilon^4=m^2\sigma_\varepsilon^4\left(3+\frac{1}{m}\left(\frac{\mu_{4,\varepsilon}}{\sigma_\varepsilon^4}-3\right)\right)$$
$$=m^2\sigma_\varepsilon^4\left(3+\frac{1}{m}\left(\rho_{4,\varepsilon}-3\right)\right).$$

In the $m=1$ case, note that $\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^4 \rightarrow_p \mu_{4,\varepsilon}=\sigma_\varepsilon^4\rho_{4,\varepsilon}$. Further,

$$\begin{aligned}
&\frac{6m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^2\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2\\
&=\frac{6m}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}^2\sum_{s'=-(m-1)}^{0}\frac{a^2u_{tm+s'}^2}{n}Y_{tm+s'-1}^2+o_p\left(1\right)\\
&=\frac{6m}{n}\sum_{t=m+1}^{n}\left(m\sigma_\varepsilon^2\right)\left(\frac{a^2\sigma_u^2}{n}Y_t^2\right)+o_p\left(1\right)\Rightarrow 6m^2\sigma_\varepsilon^2 b\int_0^1 G^2\left(r\right)dr. \qquad (17)
\end{aligned}$$

When $m=2$ we have

$$\begin{aligned}
&\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^4\\
=\;&\frac{m}{n}\left(\left(\frac{au_{m+1}}{\sqrt{n}}Y_m+\frac{au_{2m}}{\sqrt{n}}Y_{m+1}\right)^4+\cdots+\left(\frac{au_{n-1}}{\sqrt{n}}Y_{n-2}+\frac{au_n}{\sqrt{n}}Y_{n-1}\right)^4\right)+o_p\left(1\right)\\
=\;&\frac{ma^4}{n^3}\left[\left(\left(u_3+u_4\right)Y_2\right)^4+\left(\left(u_5+u_6\right)Y_4\right)^4+\cdots\right]+o_p\left(1\right)\\
=\;&\frac{ma^4}{n^3}\left[\left(\left(u_3^4+u_4^4+6u_3^2u_4^2\right)Y_2^4\right)+\left(\left(u_5^4+u_6^4+6u_5^2u_6^2\right)Y_4^4\right)+\cdots\right]+o_p\left(1\right)\\
\Rightarrow\;& ma^4\mu_{4,a}\int_0^1 G^4+6b^2\int_0^1 G^4.
\end{aligned}$$

More generally,

$$\begin{aligned}
&\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^4\Rightarrow ma^4\mu_{4,a}\int_0^1 G^4+3m\left(m-1\right)b^2\int_0^1 G^4\\
=\;&\left(3+\frac{1}{m}\left(\frac{\mu_{4,u}}{\sigma_u^4}-3\right)\right)m^2\sigma_\varepsilon^4 b^2\int_0^1 G^{*4}\left(r\right)dr=\left(3+\frac{1}{m}\left(\rho_{4,u}-3\right)\right)m^2\sigma_\varepsilon^4 b^2\int_0^1 G^{*4}\left(r\right)dr.
\end{aligned}$$

All other terms in (16) are negligible. The sample kurtosis of the m-period return is thus

$$\begin{aligned}
\gamma_n &= \frac{\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^4}{\left(\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^2\right)^2} \\
&\Rightarrow \frac{m^2\sigma_\varepsilon^4\left(3+\frac{1}{m}\left(\rho_{4,\varepsilon}-3\right)\right)+6m^2\sigma_\varepsilon^2 b\int_0^1 G^2\left(r\right)dr+\left(3+\frac{1}{m}\left(\rho_{4,u}-3\right)\right)m^2\sigma_\varepsilon^4 b^2\int_0^1 G^{*4}\left(r\right)dr}{m^2\sigma_\varepsilon^4\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2} \\
&= \frac{m^2\sigma_\varepsilon^4\left(3+\frac{1}{m}\left(\rho_{4,\varepsilon}-3\right)\right)+6m^2\sigma_\varepsilon^4 b\int_0^1 G^{*2}\left(r\right)dr+\left(3+\frac{1}{m}\left(\rho_{4,u}-3\right)\right)m^2\sigma_\varepsilon^4 b^2\int_0^1 G^{*4}\left(r\right)dr}{m^2\sigma_\varepsilon^4\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2} \\
&= \frac{\left(3+\frac{1}{m}\left(\rho_{4,\varepsilon}-3\right)\right)+6b\int_0^1 G^{*2}\left(r\right)dr+\left(3+\frac{1}{m}\left(\rho_{4,u}-3\right)\right)b^2\int_0^1 G^{*4}\left(r\right)dr}{\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2} \\
&= \frac{3+6b\int_0^1 G^{*2}\left(r\right)dr+3b^2\int_0^1 G^{*4}\left(r\right)dr+\frac{1}{m}\left(\rho_{4,\varepsilon}-3\right)+\frac{1}{m}\left(\rho_{4,u}-3\right)b^2\int_0^1 G^{*4}\left(r\right)dr}{\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2} \\
&= \frac{3\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2+3b^2\left(\int_0^1 G^{*4}\left(r\right)dr-\left(\int_0^1 G^{*2}\left(r\right)dr\right)^2\right)}{\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2} \\
&+\frac{\frac{1}{m}\left(\rho_{4,\varepsilon}-3\right)+\frac{1}{m}\left(\rho_{4,u}-3\right)b^2\int_0^1 G^{*4}\left(r\right)dr}{\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2} \\
&= 3+\frac{3b^2\left(\int G^{*4}\left(r\right)dr-\left(\int_0^1 G^{*2}\left(r\right)dr\right)^2\right)+\frac{1}{m}\left\{\left(\rho_{4,\varepsilon}-3\right)+\left(\rho_{4,u}-3\right)b^2\int_0^1 G^{*4}\left(r\right)dr\right\}}{\left(1+b\int_0^1 G^{*2}\left(r\right)dr\right)^2}.
\end{aligned} \tag{18}$$

■

Proof of Theorem 2. In the case $\mu \neq 0$, using (8) and letting $r_s = \lfloor tm + s \rfloor / \sqrt{n}$,

$$\begin{aligned}
\bar{\Delta}^m = \frac{m}{n}\sum_{t=2}^{n/m} \Delta_t^m &= \frac{m}{n}\sum_{t=2}^{n/m}\left(m\mu + \sum_{s=-(m-1)}^{0} \left(\left(\beta_{tm+s} - 1\right) Y_{tm+s-1} + \varepsilon_{tm+s} \right) \right) \\
&= m\mu + \frac{m}{n}\sum_{t=2}^{n/m} \sum_{s=-(m-1)}^{0} \left(\frac{c}{n} + \frac{a u_{tm+s}}{\sqrt{n}} + \frac{1}{2}\left(\frac{a u_{tm+s}}{\sqrt{n}}\right)^2 \right) Y_{tm+s-1} + o_p(1) \\
&\Rightarrow m\mu + c\sum_{s=-(m-1)}^{0} \int_0^1 H(r_s)\, dr + a\sum_{s=-(m-1)}^{0} \int_0^1 H(r_s)\, dB_u(r_s) + \frac{b}{2}\sum_{s=-(m-1)}^{0} \int_0^1 H(r_s)\, dr \\
&= m\mu + L, \text{ say.}
\end{aligned}$$

We then have

$$\begin{aligned}
&\frac{m}{n}\sum_{t=2}^{n/m} \left(\Delta_t^m - \bar{\Delta}^m\right)^2 \\
&= \frac{m}{n}\sum_{t=2}^{n/m}\left(m\mu + \sum_{s=-(m-1)}^{0} \left(\left(\beta_{tm+s} - 1\right) Y_{tm+s-1} + \varepsilon_{tm+s} \right) - m\mu - L \right)^2 + o_p(1) \\
&= \frac{m}{n}\sum_{t=2}^{n/m}\left(\left(\sum_{s=-(m-1)}^{0} \varepsilon_{tm+s} \right)^2 + \left(\sum_{s=-(m-1)}^{0} \left(\beta_{tm+s} - 1\right) Y_{tm+s-1} \right)^2 \right) \\
&+ \frac{2m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0} \varepsilon_{tm+s} \right)\left(\sum_{s=-(m-1)}^{0} \left(\beta_{tm+s} - 1\right) Y_{tm+s-1} \right) \\
&+ L^2 - 2L\frac{m}{n}\sum_{t=2}^{n/m}\left(\left(\sum_{s=-(m-1)}^{0} \varepsilon_{tm+s} \right) + \left(\sum_{s=-(m-1)}^{0} \left(\beta_{tm+s} - 1\right) Y_{tm+s-1} \right) \right) + o_p(1).
\end{aligned} \tag{19}$$

The first term in (19) is

$$\begin{aligned}
\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0} \varepsilon_{tm+s} \right)^2 &= \frac{m}{n}\left(\left(\varepsilon_{m+1} + \cdots + \varepsilon_{2m}\right)^2 + \left(\varepsilon_{2m+1} + \cdots + \varepsilon_{3m}\right)^2 \right. \\
&\quad \left. + \left(\varepsilon_{n-(m-1)} + \cdots + \varepsilon_n\right)^2 \right) \\
&= \frac{m}{n}\sum_{t=m+1}^{n} \varepsilon_t^2 + o_p(1) \to_p m\sigma_\varepsilon^2.
\end{aligned}$$

As $Y_t = O_p(n)$, the second term in (19) is

$$\begin{aligned}
&\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2 \\
&= \frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\frac{c}{n}+\frac{au_{tm+s}}{\sqrt{n}}+\frac{1}{2}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2\right)Y_{tm+s-1}\right)^2 \\
&= \frac{m}{n}\sum_{t=2}^{n/m}\left(\frac{c^2}{n^2}\left(\sum_{s=-(m-1)}^{0}Y_{tm+s-1}\right)^2+\left(\sum_{s=-(m-1)}^{0}\frac{au_{tm+s}}{\sqrt{n}}Y_{tm+s-1}\right)^2\right. \\
&\left.+\sum_{s=-(m-1)}^{0}\frac{au_{tm+s}}{\sqrt{n}}\sum_{s'=-(m-1)}^{0}\left(\frac{au_{tm+s'}}{\sqrt{n}}\right)^2Y_{tm+s-1}Y_{tm+s'-1}\right) \\
&+\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\frac{1}{2}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2Y_{tm+s-1}\right)^2. \qquad (20)
\end{aligned}$$

The first term in (20) satisfies

$$\frac{m}{n}\sum_{t=2}^{n/m}\frac{c^2}{n^2}\left(\sum_{s=-(m-1)}^{0}Y_{tm+s-1}\right)^2 \Rightarrow c^2 m\int_0^1 H(r)\,dr,$$

and the second term satisfies

$$\frac{1}{n}\frac{m}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\mathbb{E}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2 Y_{tm+s-1}^2 = \frac{mb}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\frac{Y_{tm+s-1}^2}{n^2} \Rightarrow mb\int_0^1 H^2(r)\,dr.$$

Similarly, the third term in (20) yields

$$\frac{1}{n}\frac{m}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\frac{au_{tm+s}}{\sqrt{n}}\sum_{s'=-(m-1)}^{0}\left(\frac{au_{tm'+s}}{\sqrt{n}}\right)^2 Y_{tm+s-1}Y_{tm+s'-1} \Rightarrow 0.$$

Consider the case $m=2$. The fourth term in (20) yields

$$
\begin{aligned}
&\frac{1}{n}\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\frac{1}{2}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2 Y_{tm+s-1}\right)^2 \\
=\;& \frac{m}{4n^2}\left(\left(\left(\frac{au_{m+1}}{\sqrt{n}}\right)^2 Y_m+\left(\frac{au_{2m}}{\sqrt{n}}\right)^2 Y_{m+1}\right)^2+\cdots\right. \\
&\left.+\left(\left(\frac{au_{n-1}}{\sqrt{n}}\right)^2 Y_{n-2}+\left(\frac{au_{n}}{\sqrt{n}}\right)^2 Y_{n-1}\right)^2\right) \\
=\;& \frac{ma^4}{4n^4}\left[\left(\left(u_3^2+u_4^2\right)Y_2\right)^2+\left(\left(u_5^2+u_6^2\right)Y_4\right)^2+\cdots\right]+o_p(1)\Rightarrow 0,
\end{aligned}
$$

and more generally

$$
\frac{1}{n}\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\frac{1}{2}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2 Y_{tm+s-1}\right)^2 \Rightarrow 0.
$$

It follows that

$$
\frac{m}{n^2}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2 \Rightarrow mb\int_0^1 H^2(r)\,dr.
$$

The third term in (15) is

$$
\begin{aligned}
&\frac{2m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right) \\
=\;& \frac{2ma}{n^{3/2}}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}u_{tm+s}Y_{tm+s-1}\right)+O_p(1)=O_p(1).
\end{aligned}
$$

The fourth and fifth terms in (15) are $O_p(1)$ and the last term is

$$
-2L\frac{m}{n}\sum_{t=2}^{n/m}\sum_{s=-(m-1)}^{0}\left(\frac{c}{n}+\frac{au_{tm+s}}{\sqrt{n}}+\frac{1}{2}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2\right)Y_{tm+s-1}=O_p(1).
$$

Hence,

$$\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m-\bar{\Delta}^m\right)^2=\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2+O_p\left(1\right)=O_p\left(n\right)$$

and we deduce that

$$\frac{m}{n^2}\sum_{t=2}^{n/m}\left(\Delta_t^m-\bar{\Delta}^m\right)^2\Rightarrow mb\int_0^1 H^2\left(r\right)dr.$$

Continuing, we find that

$$\begin{aligned}
&\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m-\bar{\Delta}^m\right)^4\\
&=\frac{m}{n}\sum_{t=2}^{n/m}\left(\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^4+4\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^3\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)\right.\\
&+6\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)^2\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^2\\
&+4\left(\sum_{s=-(m-1)}^{0}\varepsilon_{tm+s}\right)\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^3\\
&\left.+\left(\sum_{s=-(m-1)}^{0}\left(\beta_{tm+s}-1\right)Y_{tm+s-1}\right)^4\right)+O_p\left(1\right),
\end{aligned}$$

and the $O_p\left(1\right)$ involves functions of L. The leading term in the last equation is

$$\frac{m}{n}\sum_{t=2}^{n/m}\left(\sum_{s=-(m-1)}^{0}\left(\frac{c}{n}+\frac{au_{tm+s}}{\sqrt{n}}+\frac{1}{2}\left(\frac{au_{tm+s}}{\sqrt{n}}\right)^2\right)Y_{tm+s-1}\right)^4=O_p\left(n^2\right).$$

Consider the case $m = 2$. We have

$$\begin{aligned}
& \frac{m}{n^3} \sum_{t=2}^{n/m} \left(\sum_{s=-(m-1)}^{0} \left(\beta_{tm+s} - 1 \right) Y_{tm+s-1} \right)^4 \\
= \; & \frac{m}{n^3} \left(\left[\left(\frac{au_{m+1}}{\sqrt{n}} + \frac{\left(au_{m+1}\right)^2}{2n} \right) Y_m + \left(\frac{au_{2m}}{\sqrt{n}} + \frac{\left(au_{2m}\right)^2}{2n} \right) Y_{m+1} \right]^4 \right. \\
& \left. + \cdots + \left[\left(\frac{au_{n-1}}{\sqrt{n}} + \frac{\left(au_{n-1}\right)^2}{2n} \right) Y_{n-2} + \left(\frac{au_{n}}{\sqrt{n}} + \frac{\left(au_{n}\right)^2}{2n} \right) Y_{n-1} \right]^4 \right) + o_p(1) \\
\Rightarrow \; & ma^4 \mu_{4,a} \int_0^1 H^4(r)\, dr + 3m(m-1) b^2 \int_0^1 H^4(r)\, dr.
\end{aligned}$$

The sample kurtosis of the m-period return is given by

$$\begin{aligned}
\gamma = \; & \frac{\frac{m}{n} \sum_{t=2}^{n/m} \left(\Delta_t^m\right)^4}{\left(\frac{m}{n} \sum_{t=2}^{n/m} \left(\Delta_t^m\right)^2 \right)^2} = \frac{\frac{1}{n^2} \frac{m}{n} \sum_{t=2}^{n/m} \left(\Delta_t^m\right)^4}{\left(\frac{1}{n} \frac{m}{n} \sum_{t=2}^{n/m} \left(\Delta_t^m\right)^2 \right)^2} \\
= \; & \frac{\frac{1}{n^2} \frac{m}{n} \sum_{t=2}^{n/m} \left(\sum_{s=-(m-1)}^{0} \left(\frac{au_{tm+s}}{\sqrt{n}} + \frac{1}{2} \left(\frac{au_{tm+s}}{\sqrt{n}} \right)^2 \right) Y_{tm+s-1} \right)^4}{\left(\frac{1}{n} \frac{m}{n} \sum_{t=2}^{n/m} \left(\sum_{s=-(m-1)}^{0} \left(\beta_{tm+s} - 1 \right) Y_{tm+s-1} \right)^2 \right)^2} + o_p(1) \\
\Rightarrow \; & \frac{ma^4 \mu_{4,u} \int_0^1 H^4(r)\, dr + 3m(m-1) b^2 \int_0^1 H^4(r)\, dr}{(mb)^2 \left(\int_0^1 H^2(r)\, dr \right)^2} = \frac{mb^2 \frac{\mu_{4,u}}{\sigma_u^4} \int_0^1 H^4(r)\, dr + 3m(m-1) b^2 \int_0^1 H^4(r)}{(mb)^2 \left(\int_0^1 H^2(r)\, dr \right)^2} \\
= \; & \frac{\left(mb^2 \left(\rho_{4,u} - 3 \right) + 3m^2 b^2 \right) \int_0^1 H^4(r)}{(mb)^2 \left(\int_0^1 H^2(r)\, dr \right)^2} = \frac{\left(\frac{1}{m} \left(\rho_{4,u} - 3 \right) + 3 \right) \int_0^1 H^4(r)}{\left(\int_0^1 H^2(r)\, dr \right)^2}.
\end{aligned}$$

Note that as $H(r) = \mu H^*(r)$, with $H^*(r) = e^{rc + aB_u(r)} \int_0^r e^{-rp - aB_u(p)} dp$, we have

$$\gamma^{m,\mu} = \frac{\left(\frac{1}{m} \left(\rho_{4,u} - 3 \right) + 3 \right) \int_0^1 H^4(r)\, dr}{\left(\int_0^1 H^2(r)\, dr \right)^2} = \frac{\left(\frac{1}{m} \left(\rho_{4,u} - 3 \right) + 3 \right) \int_0^1 H^{*4}(r)\, dr}{\left(\int_0^1 H^{*2}(r)\, dr \right)^2},$$

which is independent of μ. ■

Proof of Theorem 3. In the $\mu = 0$ case,

$$\hat{\beta} = \frac{\sum_{t=1}^{n} Y_t Y_{t-1}}{\sum_{t=1}^{n} Y_{t-1}^2} = \frac{\sum_{t=1}^{n} \beta_t Y_{t-1}^2 + \sum_{t=1}^{n} Y_{t-1}\varepsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2}.$$

Since β_t has the approximate form

$$\begin{aligned}\beta_t = e^{\frac{c}{n}+a\frac{u_t}{\sqrt{n}}} &= e^{\frac{c}{n}}\left\{1 + a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{u_t^2}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\} \\ &= e^{\frac{c}{n}}\left\{1 + \frac{b}{2n} + a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{\left(u_t^2-\sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\},\end{aligned} \tag{21}$$

we have

$$\begin{aligned}\hat{\beta} &= \frac{e^{\frac{c}{n}}\sum_{t=1}^{n}\left\{e^{\frac{b}{2n}} + a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{\left(u_t^2-\sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\}Y_{t-1}^2 + \sum_{t=1}^{n} Y_{t-1}\varepsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2} \\ &= e^{\frac{c+b/2}{n}} + \frac{e^{\frac{c}{n}}\sum_{t=1}^{n}\left\{a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{\left(u_t^2-\sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\}Y_{t-1}^2 + \sum_{t=1}^{n} Y_{t-1}\varepsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2}.\end{aligned}$$

Then

$$\hat{\beta} - e^{\frac{c+b/2}{n}} = \frac{\sum_{t=1}^{n}\left\{a\frac{u_t}{\sqrt{n}} + a^2\frac{\left(u_t^2-\sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\}Y_{t-1}^2 + \sum_{t=1}^{n} Y_{t-1}\varepsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2}$$

and

$$\begin{aligned} n\left(\hat{\beta} - e^{\frac{c+b/2}{n}}\right) &= \frac{\frac{1}{n}\sum_{t=1}^{n}\left\{a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{\left(u_t^2-\sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\}Y_{t-1}^2 + \frac{1}{n}\sum_{t=1}^{n} Y_{t-1}\varepsilon_t}{\frac{1}{n^2}\sum_{t=1}^{n} Y_{t-1}^2} \\ &= \frac{\sum_{t=1}^{n}\left\{a\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2\frac{u_t}{\sqrt{n}} + \frac{a^2}{2\sqrt{n}}\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2\frac{\left(u_t^2-\sigma_u^2\right)}{\sqrt{n}} + O_p\left(\frac{1}{n^{3/2}}\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2\right)\right\} + \sum_{t=1}^{n}\frac{Y_{t-1}}{\sqrt{n}}\frac{\varepsilon_t}{\sqrt{n}}}{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2} \\ &= \frac{a\sum_{t=1}^{n}\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2\frac{u_t}{\sqrt{n}} + \sum_{t=1}^{n}\frac{Y_{t-1}}{\sqrt{n}}\frac{\varepsilon_t}{\sqrt{n}}}{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2} + O_p\left(\frac{1}{\sqrt{n}}\right) \\ &\Rightarrow \frac{a\int_0^1 G(r)^2\,dB_u(r) + \int_0^1 G(r)\,dB_\varepsilon(r)}{\int_0^1 G(r)^2\,dr},\end{aligned}$$

by virtue of the limit theory in Ibragimov and Phillips (2008) in the present case of independence between u_t and ε_s. Next, consider the kurtosis measure based on the simple fitted AR regression $\hat{Y}_t = \hat{\beta} Y_{t-1}$, for which

$$\hat{\Delta}_t := \hat{Y}_t - \hat{Y}_{t-1} = \hat{\beta}\left(Y_{t-1} - Y_{t-2}\right) = \left\{e^{\frac{c+b/2}{n}} + O_p\left(\frac{1}{n}\right)\right\}\left(Y_{t-1} - Y_{t-2}\right)$$

and upon temporal aggregation,

$$\begin{aligned} \hat{\Delta}_t^m \quad : \quad &= \hat{Y}_t^m - \hat{Y}_{t-1}^m = \left(\hat{Y}_{tm} - \hat{Y}_{tm-1}\right) + \left(\hat{Y}_{tm-1} - \hat{Y}_{tm-2}\right) + \ldots + \left(\hat{Y}_{tm-(m-1)} - \hat{Y}_{tm-m}\right) \\ &= \hat{\beta}\left\{\left(Y_{tm} - Y_{tm-1}\right) + \left(Y_{tm-1} - Y_{tm-2}\right) + \ldots + \left(Y_{tm-(m-1)} - Y_{tm-m}\right)\right\} \\ &= \hat{\beta}\Delta_t^m = \left\{e^{\frac{c+b/2}{n}} + O_p\left(\frac{1}{n}\right)\right\}\Delta_t^m. \end{aligned}$$

The asymptotic behavior of the kurtosis measure based on the fitted, misspecified constant parameter AR is therefore given by

$$\gamma_{AR,n}^m := \frac{\frac{m}{n}\sum_{t=2}^{n/m}\left(\hat{\Delta}_t^m\right)^4}{\left(\frac{m}{n}\sum_{t=2}^{n/m}\left(\hat{\Delta}_t^m\right)^2\right)^2} = \frac{\hat{\beta}^4 \frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^4}{\hat{\beta}^4\left(\frac{m}{n}\sum_{t=2}^{n/m}\left(\Delta_t^m\right)^2\right)^2} \Rightarrow \gamma^m,$$

and part (i) of the theorem is completed.

To establish parts (ii) and (iii), in the $\mu \neq 0$ case the least squares regression under the misspecification that $\beta_t = \beta$ is constant gives

$$\hat{\beta}^\mu = \sum_{t=1}^n Y_t \tilde{Y}_{t-1} / \sum_{t=1}^n \tilde{Y}_{t-1}^2 = \frac{\sum_{t=1}^n \beta_{nt}\tilde{Y}_{t-1}^2 + \sum_{t=1}^n \tilde{Y}_{t-1}\varepsilon_t}{\sum_{t=1}^n \tilde{Y}_{t-1}^2},$$

where $\tilde{Y}_{t-1} = Y_t - \bar{Y}$ is defined prior to Theorem 3. Since β_t has the approximate form (21), we deduce that

$$\begin{aligned} \hat{\beta}^\mu \quad &= \quad \frac{e^{\frac{c}{n}}\sum_{t=1}^n\left\{e^{\frac{b}{2n}} + a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{\left(u_t^2 - \sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\}\tilde{Y}_{t-1}^2 + \sum_{t=1}^n \tilde{Y}_{t-1}\varepsilon_t}{\sum_{t=1}^n \tilde{Y}_{t-1}^2} \\ &= \quad e^{\frac{c+b/2}{n}} + \frac{\frac{e^{\frac{c}{n}}}{n^3}\sum_{t=1}^n\left\{a\frac{u_t}{\sqrt{n}} + \frac{a^2}{2}\frac{\left(u_t^2 - \sigma_u^2\right)}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right\}\tilde{Y}_{t-1}^2 + \frac{1}{n^3}\sum_{t=1}^n \tilde{Y}_{t-1}\varepsilon_t}{\frac{1}{n^3}\sum_{t=1}^n \tilde{Y}_{t-1}^2}. \end{aligned}$$

Hence,

$$
\begin{aligned}
n\left(\hat{\beta}^{\mu}-e^{\frac{c+b/2}{n}}\right) &= \frac{\frac{e^{\frac{c}{n}}}{n^2}\sum_{t=1}^{n}\left\{a\frac{u_t}{\sqrt{n}}+\frac{a^2}{2}\frac{\left(u_t^2-\sigma_u^2\right)}{n}+O_p\left(\frac{1}{n^{3/2}}\right)\right\}\tilde{Y}_{t-1}^2+\frac{1}{n^2}\sum_{t=1}^{n}\tilde{Y}_{t-1}\varepsilon_t}{\frac{1}{n^3}\sum_{t=1}^{n}\tilde{Y}_{t-1}^2} \\
&\Rightarrow \frac{a\int_0^1 \tilde{H}(s)^2\, dB_u(s)}{\int_0^1 \tilde{H}(s)^2\, ds},
\end{aligned}
$$

where $\tilde{H}(r) := H(r) - \int_0^1 H(s)\, ds$, and $H(r)$ is given in (7).

Next, consider the estimate of the intercept in the regression $\hat{Y}_t = \hat{\mu} + \hat{\beta}^{\mu} Y_{t-1}$, where

$$
\hat{\mu} = \bar{Y} - \hat{\beta}^{\mu}\bar{Y}_{-1} = \mu + \frac{1}{n}\sum_{t=1}^{n}\beta_t Y_{t-1} + \bar{\varepsilon} - \hat{\beta}^{\mu}\bar{Y}_{-1} = \mu + \frac{1}{n}\sum_{t=1}^{n}\left(\beta_t - \hat{\beta}^{\mu}\right) Y_{t-1} + \bar{\varepsilon}
$$

$$
\begin{aligned}
&= \mu + \frac{1}{n}\sum_{t=1}^{n}\left(e^{\frac{c}{n}+a\frac{u_t}{\sqrt{n}}} - e^{\frac{c+b/2}{n}} + O_p\left(\frac{1}{n}\right)\right) Y_{t-1} + \bar{\varepsilon} \\
&= \mu + \frac{e^{\frac{c}{n}}}{n}\sum_{t=1}^{n}\left(a\frac{u_t}{\sqrt{n}} - \frac{b}{2n} + O_p\left(\frac{1}{n}\right)\right) Y_{t-1} + o_p(1) \\
&= \mu + ae^{\frac{c}{n}}\sum_{t=1}^{n}\frac{Y_{t-1}}{n}\frac{u_t}{\sqrt{n}} - \frac{e^{\frac{c}{n}}b}{2n}\sum_{t=1}^{n}\frac{Y_{t-1}}{n} + o_p(1) \\
&\Rightarrow \mu + a\int_0^1 H(s)\, dB_u(s) - \frac{b}{2}\int_0^1 H(s)\, ds,
\end{aligned}
$$

showing that the fitted intercept estimator is inconsistent.

Now, consider the kurtosis measure based on the simple fitted AR regression $\hat{Y}_t = \hat{\mu} + \hat{\beta}^{\mu} Y_{t-1}$ so that $\tilde{Y}_t = \hat{\beta}^{\mu}\tilde{Y}_{t-1}$. Then

$$
\hat{\Delta}_t^{\mu} := \hat{Y}_t^{\mu} - \hat{Y}_{t-1}^{\mu} = \hat{\beta}^{\mu}\left(Y_{t-1} - Y_{t-2}\right) = \left\{e^{\frac{c+b/2}{n}} + O_p\left(\frac{1}{n}\right)\right\}\left(Y_{t-1} - Y_{t-2}\right)
$$

and upon temporal aggregation,

$$
\begin{aligned}
\hat{\Delta}_t^{m,\mu} \quad &: \quad = \hat{Y}_t^{m,\mu} - \hat{Y}_{t-1}^{m,\mu} = \left(\hat{Y}_{tm}^{\mu} - \hat{Y}_{tm-1}\right) + \left(\hat{Y}_{tm-1} - \hat{Y}_{tm-2}\right) + \ldots + \left(\hat{Y}_{tm-(m-1)} - \hat{Y}_{tm-m}\right) \\
&= \hat{\beta}^{\mu}\left\{\left(Y_{tm} - Y_{tm-1}\right) + \left(Y_{tm-1} - Y_{tm-2}\right) + \ldots + \left(Y_{tm-(m-1)} - Y_{tm-m}\right)\right\} \\
&= \hat{\beta}^{\mu}\Delta_t^m = \left\{e^{\frac{c+b/2}{n}} + O_p\left(\frac{1}{n}\right)\right\}\Delta_t^m.
\end{aligned}
$$

The asymptotic behavior of the kurtosis measure based on the fitted, misspecified constant parameter AR is then

$$\gamma_{AR,n}^{m,\mu} := \frac{\frac{m}{n}\sum_{t=2}^{n/m}\left(\hat{\Delta}_t^{m,\mu}\right)^4}{\left(\frac{m}{n}\sum_{t=2}^{n/m}\left(\hat{\Delta}_t^{m,\mu}\right)^2\right)^2} \Rightarrow \gamma^{m,\mu},$$

just as in (9). So, the constant coefficient AR fitted kurtosis measure again reproduces the actual data kurtosis in spite of the fact that the intercept is inconsistently estimated.■

Proof of Theorem 4. Defining the filtration $\mathcal{F}_r = \sigma\{(B_u(s), B_\varepsilon(s)), 0 \leq s \leq r\}$ and letting $X = c + \frac{b}{2}$, we have

$$\begin{aligned}
&\mathbb{E}\left[(dG(r))^4 |\mathcal{F}_r\right] = \mathbb{E}\left[aG(r)\,dB_u(r) + dB_\varepsilon(r) + XG(r)\,dr|\mathcal{F}_r\right]^4 \\
&= \mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^4 |\mathcal{F}_r\right] \\
&+ 4X\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^3 G(r) |\mathcal{F}_r\right] dr \\
&+ 6X^2\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^2 G(r)^2 |\mathcal{F}_r\right] (dr)^2 \\
&+ 4X^3\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))\, G(r)^3 |\mathcal{F}_r\right] (dr)^3 \\
&+ X^4\mathbb{E}\left[G(r)^4 |\mathcal{F}_r\right] (dr)^4
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^4 |\mathcal{F}_r\right] \\
&+ 6X^2\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^2 G(r)^2 |\mathcal{F}_r\right] (dr)^2 + O\left((dr)^4\right) \\
&= \left[3b^2G(r)^4 + 6b\sigma_\varepsilon^2 G(r)^2 + 3\sigma_\varepsilon^4\right] (dr)^2 \\
&+ 6X^2\left[bG(r)^4 + G(r)^2\sigma_\varepsilon^2\right] (dr)^3 + O\left((dr)^4\right),
\end{aligned} \tag{22}$$

since $\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^3 G(r) |\mathcal{F}_r\right] dr = 0$. Similarly

$$\begin{aligned}
&\mathbb{E}\left[(dG(r))^2 |\mathcal{F}_r\right] = \mathbb{E}\left[aG(r)\,dB_u(r) + dB_\varepsilon(r) + XG(r)\,dr|\mathcal{F}_r\right]^2 \\
= \;& \mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^2 |\mathcal{F}_r\right] + X^2\mathbb{E}\left[G(r)^2 |\mathcal{F}_r\right] (dr)^2 \\
= \;& \mathbb{E}\left[\left(a^2\sigma_u^2 G(r)^2 + \sigma_\varepsilon^2\right) |\mathcal{F}_r\right] dr + X^2\mathbb{E}\left[G(r)^2 |\mathcal{F}_r\right] (dr)^2 \\
= \;& \left[bG(r)^2 + \sigma_\varepsilon^2\right] dr + X^2 G(r)^2 (dr)^2.
\end{aligned} \tag{23}$$

We define the instantaneous kurtosis measure

$$\kappa_{b,c}(r) = \frac{\mathbb{E}\left(\mathbb{E}\left[(dG(r))^4 |\mathcal{F}_r\right]\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(dG(r))^2 |\mathcal{F}_r\right]\right)\right\}^2}$$

and using (22) - (23) we obtain

$$\kappa_{b,c}(r) = \frac{\mathbb{E}\left(\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^4 \,|\mathcal{F}_r\right]\right) + o\left((dr)^2\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^2 \,|\mathcal{F}_r\right]\right) + O\left((dr)^2\right)\right\}^2}$$

$$\begin{aligned}
&= \frac{3b^2\mathbb{E}\left(G(r)^4\right) + 6b\sigma_\varepsilon^2\mathbb{E}\left(G(r)^2\right) + 3\sigma_\varepsilon^4}{\left(b\mathbb{E}\left(G(r)^2\right) + \sigma_\varepsilon^2\right)^2} + o(1) \\
&= \frac{3b^2\mathbb{E}\left(G(r)^4\right) + 6b\sigma_\varepsilon^2\mathbb{E}\left(G(r)^2\right) + 3\sigma_\varepsilon^4}{b^2\left(\mathbb{E}\left(G(r)^2\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2\mathbb{E}\left(G(r)^2\right)} + o(1) \\
&= 3 + \frac{3b^2\left[\mathbb{E}\left(G(r)^4\right) - \left(\mathbb{E}\left(G(r)^2\right)\right)^2\right]}{b^2\left(\mathbb{E}\left(G(r)^2\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2\mathbb{E}\left(G(r)^2\right)} + o(1),
\end{aligned}$$

as required. ■

Proof of Theorem 5. Under independence of B_u and B_ε,

$$\mathbb{E}G(r) = \alpha\int_0^r e^{(r-p)c}\mathbb{E}e^{a(B_u(r)-B_u(p))}dp = \alpha\int_0^r e^{X(r-p)}dp = \alpha\frac{e^{Xr}-1}{X}$$

and

$$\mathbb{E}(dG(r)) = \{\alpha + X\mathbb{E}G(r)\}\,dr = \left\{\alpha + \alpha\left(e^{Xr}-1\right)\right\}dr = \alpha e^{Xr}dr.$$

We have

$$\begin{aligned}
\Lambda(r) &\equiv dG(r) - \mathbb{E}(dG(r)) \\
&= aG(r)\,dB_u(r) + dB_\varepsilon(r) + (\alpha + XG(r))\,dr - (\alpha + X\mathbb{E}(G(r)))\,dr \\
&= aG(r)\,dB_u(r) + dB_\varepsilon(r) + X(G(r)) - \mathbb{E}(G(r))\,dr.
\end{aligned}$$

It follows that the instantaneous kurtosis of the returns in this case is

$$\begin{aligned}
\kappa_{b,c,\alpha} &= \frac{\mathbb{E}\left(\mathbb{E}\left[(dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r))^4 \,|\mathcal{F}_r\right]\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(dG(r) - \mathbb{E}(dG(r)|\mathcal{F}_r))^2 \,|\mathcal{F}_r\right]\right)\right\}^2} \\
&= \frac{\mathbb{E}\left(\mathbb{E}\left[(aG(r)\,dB_u(r) + dB_\varepsilon(r))^4 \,|\mathcal{F}_r\right]\right) + o\left((dr)^2\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[(aGdB_u(r) + dB_\varepsilon(r))^2 \,|\mathcal{F}_r\right]\right) + o\left((dr)^2\right)\right\}^2}.
\end{aligned}$$

We have

$$\begin{aligned}
\mathbb{E}\left(\Lambda^2(r)\right) &= \left(b\mathbb{E}\left(G^2(r)\right) + \sigma_\varepsilon^2\right)dr + \mathbb{E}\left[[X\{G(r) - \mathbb{E}(G(r))\}]^2\right](dr)^2 \\
&= \mathbb{E}(D(r))\,dr + X^2Var(G(r))\,(dr)^2.
\end{aligned}$$

Furthermore,

$$\left(\mathbb{E}\left(\Lambda^{2}\left(r\right)\right)\right)^{2}=\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}\left(dr\right)^{2}+2\mathbb{E}\left(D\left(r\right)\right)X^{2}Var\left(G\left(r\right)\right)\left(dr\right)^{3}+o\left(dr^{3}\right),$$

and

$$\begin{aligned}
& \mathbb{E}\left(\Lambda^{4}\left(r\right)\right) \\
= & \mathbb{E}\left[aG\left(r\right)dB_{u}\left(r\right)+dB_{\varepsilon}\left(r\right)\right]^{4} \\
& +6\mathbb{E}\left\{\left[aG\left(r\right)dB_{u}\left(r\right)+dB_{\varepsilon}\left(r\right)\right]^{2}X^{2}\left\{G\left(r\right)-\mathbb{E}\left(G\left(r\right)\right)\right\}\right\}\left(dr\right)^{3}+o\left(dr^{3}\right) \\
= & \left\{3b^{2}\mathbb{E}\left(G^{4}\left(r\right)\right)+3\sigma_{\varepsilon}^{4}+6b\sigma_{\varepsilon}^{2}\mathbb{E}\left(G^{2}\left(r\right)\right)\right\}\left(dr\right)^{2} \\
& +6\mathbb{E}\left[\left\{aG\left(r\right)dB_{u}\left(r\right)+dB_{\varepsilon}\left(r\right)\right\}^{2}X^{2}\left\{G\left(r\right)-\mathbb{E}\left(G\left(r\right)\right)\right\}^{2}\right]\left(dr\right)+o\left(dr^{3}\right) \\
= & 3b^{2}Var\left(G^{2}\left(r\right)\right)\left(dr\right)^{2} \\
& +3\left\{b^{2}\left(\mathbb{E}\left(G^{2}\left(r\right)\right)\right)^{2}+\sigma_{\varepsilon}^{4}+2b\sigma_{\varepsilon}^{2}\mathbb{E}\left(G^{2}\left(r\right)\right)\right\}\left(dr\right)^{2} \\
& +6\mathbb{E}\left[\left\{b\left\{G^{2}\left(r\right)-\mathbb{E}\left(G^{2}\left(r\right)\right)\right\}+b\mathbb{E}\left(G^{2}\left(r\right)\right)+\sigma_{\varepsilon}^{2}\right\}X^{2}\left\{G\left(r\right)-\mathbb{E}\left(G\left(r\right)\right)\right\}^{2}\right]\left(dr\right)^{3} \\
& +o\left(dr^{3}\right) \\
= & 3b^{2}Var\left(G^{2}\left(r\right)\right)\left(dr\right)^{2}+3\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}\left(dr\right)^{2} \\
& +6\mathbb{E}\left[\left\{b\left\{G^{2}\left(r\right)-\mathbb{E}\left(G^{2}\left(r\right)\right)\right\}\right\}X^{2}\left\{G\left(r\right)-\mathbb{E}\left(G\left(r\right)\right)\right\}^{2}\right]\left(dr\right)^{3} \\
& +6\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}X^{2}Var\left(G\left(r\right)\right)\left(dr\right)^{3}+o\left(dr^{3}\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\kappa_{b,c,m} &= \frac{\mathbb{E}\left(\mathbb{E}\left[\left(dG\left(r\right)-\mathbb{E}\left(dG(r)|\mathcal{F}_{r}\right)\right)^{4}|\mathcal{F}_{r}\right]\right)}{\left\{\mathbb{E}\left(\mathbb{E}\left[\left(dG\left(r\right)-\mathbb{E}\left(dG(r)|\mathcal{F}_{r}\right)\right)^{2}|\mathcal{F}_{r}\right]\right)\right\}^{2}} \\
&= \frac{3\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}\left(dr\right)^{2}+6\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}X^{2}Var\left(G\left(r\right)\right)\left(dr\right)^{3}}{\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}\left(dr\right)^{2}+2\mathbb{E}\left(D\left(r\right)\right)X^{2}Var\left(G\left(r\right)\right)\left(dr\right)^{3}+o\left(dr^{3}\right)} \\
&\quad+\frac{3b^{2}Var\left(G^{2}\left(r\right)\right)\left(dr\right)^{2}}{\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}\left(dr\right)^{2}+2\mathbb{E}\left(D\left(r\right)\right)X^{2}Var\left(G\left(r\right)\right)\left(dr\right)^{3}+o\left(dr^{3}\right)} \\
&\quad+\frac{6b\mathbb{E}\left[\left\{G^{2}\left(r\right)-\mathbb{E}\left(G^{2}\left(r\right)\right)\right\}X^{2}\left\{G\left(r\right)-\mathbb{E}\left(G\left(r\right)\right)\right\}^{2}\right]\left(dr\right)^{3}+o\left(dr^{3}\right)}{\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}\left(dr\right)^{2}+2\mathbb{E}\left(D\left(r\right)\right)X^{2}Var\left(G\left(r\right)\right)\left(dr\right)^{3}+o\left(dr^{3}\right)} \\
&= 3+\frac{3b^{2}Var\left(G^{2}\left(r\right)\right)+6\mathbb{E}\left[\left\{b\left\{G^{2}\left(r\right)-\mathbb{E}\left(G^{2}\left(r\right)\right)\right\}\right\}X^{2}\left\{G\left(r\right)-\mathbb{E}\left(G\left(r\right)\right)\right\}^{2}\right]\left(dr\right)}{\left\{\mathbb{E}\left(D\left(r\right)\right)\right\}^{2}+2\mathbb{E}\left(D\left(r\right)\right)X^{2}Var\left(G\left(r\right)\right)\left(dr\right)+O\left(dr^{2}\right)} \\
&\quad+O\left(\left(dr\right)^{2}\right).
\end{aligned}$$

■

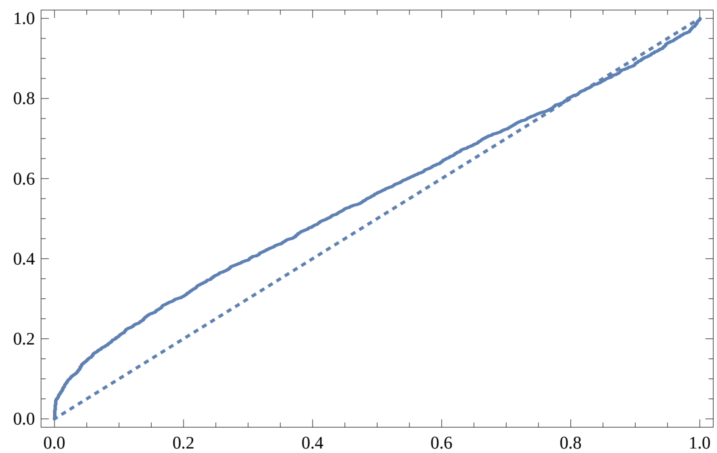

Figure 1: LSTUR with a drift: PP Plot of Δ^5 against its asymptotic distribution, $n = 9000$, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $\mu = 2.5$, $a = 0.7$, $c = -b$.

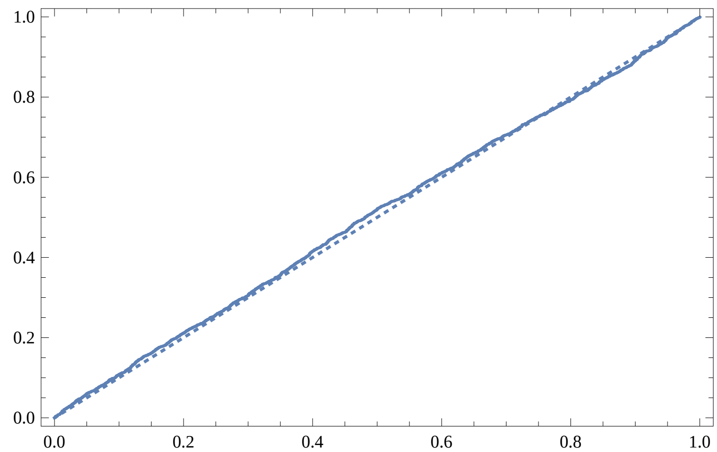

Figure 2: LSTUR with a drift: PP Plot of Δ^5 against its asymptotic distribution, $n = 240000$, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $\mu = 2.5$, $a = 0.7$, $c = -b$.

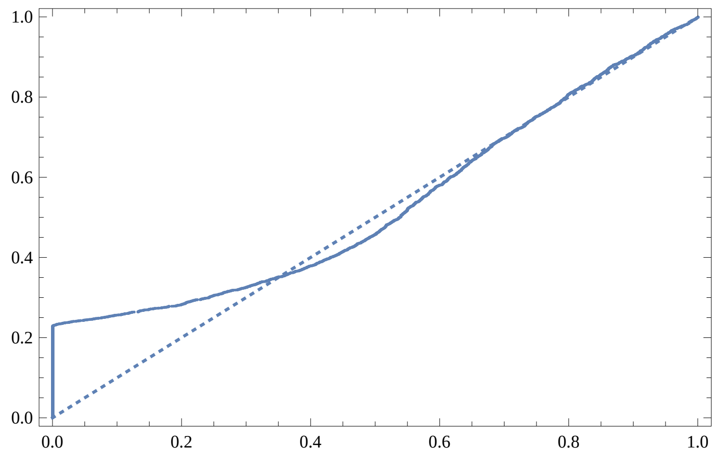

Figure 3: LSTUR without a drift: PP Plot of Δ^5 against its asymptotic distribution, $n = 9000$, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $a = 0.7$, $c = -b$.

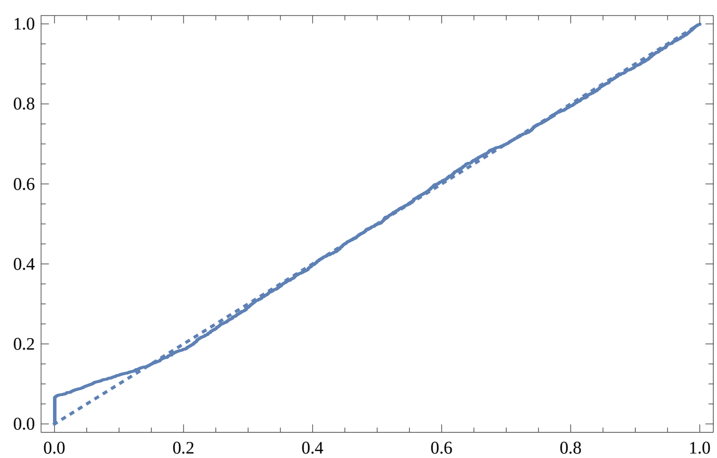

Figure 4: LSTUR without a drift: PP Plot of Δ^5 against its asymptotic distribution, $n = 240000$, $u_t \overset{iid}{\sim} t(6)$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $a = 0.7$, $c = -b$.