Estimating and Forecasting

Long-Horizon Dollar Return Skewness*

Kevin Aretz† Jiayu Jin‡ Yifan Li§

Abstract

We develop a parametric estimator of the physical skewness of an asset's discrete ("dollar") return over long horizons from the assumption that the asset's value can be modelled using a stochastic process from the affine stochastic volatility (ASV) model class. Taking compounding and leverage effects into account, we demonstrate that our estimator is close to unbiased and efficient, setting it apart from other recent estimators. In a further contrast to those other estimators, it also lends itself naturally to forecasting skewness. Applying our estimator to representative stock indexes, we show that the skewness of long-horizon dollar returns is far less extreme than suggested in the current literature.

JEL classification: C13, C18, C22, C58.

Keywords: Skewness estimation; discrete returns; leverage effect; compounding effect; affine stochastic volatility models.

---

*We are indebted to Michael Brennan, Hening Liu as well as the participants at the 2020 Michael Brennan Day (Manchester), the Econometric Webinar at Konstanz, the 14$^{\text{th}}$ International Conference on Computational & Financial Econometrics (CFE 2020), the 6$^{\text{th}}$ Konstanz-Lancaster-Manchester-Freiburg (KoLaMaFr) Workshop on Financial Econometrics, and the 7$^{\text{th}}$ International Young Finance Scholars' Conference (IYFS) for valuable suggestions and advice.

†Corresponding author, Accounting and Finance Group, Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK. e-mail: kevin.aretz@manchester.ac.uk, website: www.kevin-aretz.com.

‡Accounting and Finance Group, Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK. e-mail: jiayu.jin-3@postgrad.manchester.ac.uk.

§Accounting and Finance Group, Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK. e-mail: yifan.li@manchester.ac.uk.

1 Introduction

While a huge theoretical literature dating back to the 1960s studies how the physical skewness of an asset's discrete ("dollar") return affects investor behavior and asset prices, only few empirical studies convincingly test the main predictions of that literature.[1] The reason for this gap is that it is challenging to empirically estimate skewness with realistic amounts of data, especially over the long return horizons the theoretical literature focuses on.[2] Only recently, a handful of studies including Fama and French (2018) and Farago and Hjalmarsson (2019) have started addressing the estimation issue head on, proposing new estimators relying on fewer data. Yet, as we will demonstrate, those estimators often deliver biased estimates under realistic assumptions, especially over long horizons.

In our paper, we develop a new estimator of the skewness of an asset's return over an arbitrary horizon under the assumption that the asset's value can be modelled using a stochastic process from the affine stochastic volatility (ASV) model class (see Duffie et al. (2000, 2003)). To do so, we first explain how to calculate the conditional and unconditional version of the skewness of the asset's return using the corresponding moment generating functions (MGF) of the stochastic process. Since the MGFs depend on the parameter values of the stochastic process, we can then obtain estimates of the two skewness versions from explicit estimates of the parameters. Crucially, using the most recent estimate of the conditional MGF, we are easily able to calculate a natural and internally consistent forecast of skewness over the desired future horizon.

We choose the Heston (1993) stochastic process as an example to implement

[1] Since we exclusively look into the *physical* skewness of an asset's *dollar* return, we will, for simplicity, refer to that skewness as the "skewness of an asset's return" from here on, unless stated otherwise. Given that only the dollar (and not the log) return measures the profitability of real-world investments (i.e., only the dollar return is "investable"), its skewness is a more relevant statistic for real-world investors than the skewness of the log return.

[2] To illustrate, even if monthly returns were independently and identically (i.i.d) normal, standard estimators such as the sample skewness would need about 600 observations (approximately 50 years of data) to estimate skewness with a standard error of 0.10. Needless to say, most assets, as, for example, most single stocks, do not exist for such long time periods.

our new methodology, the advantages being that the process is simple yet flexible, popular, and has an MGF available in closed form. Also, we are able to consistently estimate the parameters of the process with a novel GMM estimator. Using a Monte Carlo simulation exercise, we then confirm that our skewness estimator delivers close to unbiased and efficient estimates of the skewness of returns with an up to five-year horizon – even when we use no more than ten years of data in our parameter estimation. Under our base case parameter setting for the Heston process, the true *unconditional* skewness is, for example, 0.044, 0.097, 0.594, 1.406, and 2.120 at the weekly, monthly, annual, three-year, and five-year return horizon, respectively. In striking agreement, the mean estimate of our estimator is 0.041, 0.093, 0.575, 1.286, and 1.890, with an impressive mean squared error (MSE) of 0.001, 0.004, 0.018, 0.047 and 0.112, all respectively.

Interestingly, Fama and French's (2018) and Farago and Hjalmarsson's (2019) new estimators perform much worse in most simulations, with them overshooting skewness over all horizons and them failing to capture possible declines in skewness over shorter horizons. This occurs because their estimators miss out on one of two forces driving skewness: the "compounding" and the "leverage" effect. The compounding effect implies that compounding up short-horizon dollar returns produces right skewness in the long-horizon dollar return, with the effect being amplified by short-horizon return volatility (Bessembinder (2018)). The leverage effect implies that a negative dependence between return and volatility lowers skewness, especially over short horizons (Neuberger and Payne (2020)). While Fama and French's (2018) and Farago and Hjalmarsson's (2019) new estimators capture the compounding effect, their assumption that returns are i.i.d. makes them miss out on the leverage effect. Our estimator, in contrast, incorporates both the compounding effect as well as the leverage effect.

Notwithstanding, our estimator relies crucially on the assumption that an ASV model can accurately describe asset values. To illustrate that small deviations from that assumption do not completely invalidate the estimator, we next simulate asset values from a stochastic process similar to but even more realistic than the Heston process, namely, the multi-factor Heston process. Using our estimator based on the Heston process on those simulated data, we show that the estimator continues to strongly outperform the others.

We finally apply our skewness estimator to real-world stock indexes, an asset class for which the existence of a leverage effect has been extensively shown in prior research. Our evidence reveals that the leverage effect dominates the unconditional skewness of those indexes over shorter horizons, whereas the compounding effect dominates that skewness over longer horizons. To be more precise, we discover that the unconditional skewness of weekly returns is always larger than that of monthly returns and sometimes even larger than that of annual returns, suggesting that the "pull-down effect" caused by leverage can be strong. Notwithstanding, we also find that unconditional skewness always increases with the return horizon over sufficiently long horizons. We also learn that the unconditional skewness of long-horizon dollar returns is much lower than suggested in Bessembinder (2018) and Farago and Hjalmarsson (2019), mainly owing to the leverage effect. Given that, we do not share Farago and Hjalmarsson's (2019) sentiment that long-horizon skewness is too extreme to be useful in practical applications.

Our work contributes to a small but emerging literature on how to accurately estimate and forecast the skewness of returns over long horizons using limited amounts of data. As is well appreciated, standard estimators do not succeed in doing so.[3] Thus, Fama and French (2018) propose conducting a simple bootstrap on short-horizon returns, compounding up the draws to create a "bootstrap long-horizon return." Repeating that step multiple times, they compute the skewness of long-horizon returns from the artifical returns. Assuming that short-horizon returns are i.i.d., Farago and Hjalmarsson (2019) derive a closed-form

[3] In addition, Li (2020) shows that standard estimators deliver biased estimates for the skewness of stock returns in the presence of leverage and volatility feedback effects.

solution expressing the skewness of long-horizon returns as a function of short-horizon return moments. Approximating the variance and skewness operator, Neuberger (2012) and Neuberger and Payne (2020) derive a closed-form solution expressing the skewness of *log* long-horizon returns as a function of short-horizon return moments but without assuming i.i.d. returns. We add to those studies by proposing an alternative parametric estimator of the skewness of long-horizon returns producing more unbiased and efficient estimates. In contrast to the other estimators, our estimator also easily yields a forecast of skewness.

Our work further feeds into a well-established literature on how the skewness of returns affects financial decision-making and thus asset prices. Arditti (1967) and Scott and Horvath (1980) establish that von-Neumann-Morgenstern (NM) investors prefer more positively skewed returns. Simkowitz and Beedles (1978) and Conine Jr and Tamarkin (1981) highlight that a preference for skewness can lead the same investors to hold optimal underdiversified portfolios. Assuming NM preferences in combination with monetary separation, Rubinstein (1973), Kraus and Litzenberger (1976), and Harvey and Siddique (2000) document that co-skewness (i.e., systematic skewness) prices assets. Using non-NM preferences or violations of monetary separation, Brunnermeier et al. (2007), Mitton and Vorkink (2007), and Barberis and Huang (2008) reveal that idiosyncratic skewness may also price assets. Importantly, while the theories above all study the skewness of long-horizon returns, empirical studies typically rely on estimates of the skewness of short-horizon returns to test them, likely due to the difficulties in estimating long-horizon skewness with limited data.[4] We add to those studies by making a first step toward testing those theories using more meaningful proxies.

We finally also contribute to the literature on estimating the parameters of

[4]Empirical studies investigating the effect of skewness on stock and option returns include, for example, Boyer et al. (2010), Boyer and Vorkink (2014), Conrad et al. (2014), and Amaya et al. (2015). Those studies are distinct from others concentrating on the effect of risk-neutral skewness, as, for example, Bali and Murray (2013), Conrad et al. (2013), and Stilger et al. (2017). Since systematic risk lowers risk-neutral skewness relative to physical skewness, the results from the two types of studies should not be directly compared.

continuous stochastic volatility models. Previous approaches include generalized method of moments (GMM; e.g., Hansen and Scheinkman (1995) and Pan (2002)), maximum likelihood (e.g., Bakshi et al. (2006), Bates (2006), and Aït-Sahalia et al. (2007)), simulated methods of moments (e.g., Gallant and Tauchen (1998), Duffie et al. (2000), and Chernov et al. (2003)), Markov Chain Monte Carlo (e.g., Eraker et al. (2003) and Li et al. (2008)), and the empirical characteristic function method (e.g., Singleton (2001) and Jiang and Knight (2002)). Unlike existing GMM estimators, our GMM estimator is constructed by matching the theoretical central moments and cross-moments of the dollar return with its sample counterparts, substituting small time interval approximations for analytical expressions of the theoretical moments.

The paper proceeds as follows. In Section 2, we introduce our theoretical framework and derive our general skewness estimator. In Section 3, we propose a Heston-type estimator based on our framework and show how to implement it in practice. We further discuss alternative estimators proposed in the recent literature. Section 4 offers a simulation exercise studying the unbiasedness and efficiency of our estimator and the others. In Section 5, we apply our Heston-type estimator to real-world data. Section 6 sums up and concludes.

2 A General Parametric Skewness Estimator

In this section, we introduce our new parametric estimator of the skewness of an asset's dollar return over some arbitrary horizon. We start with defining the class of stochastic processes which can be used to model an asset's value in our setup and then outline how to calculate conditional and unconditional return skewness under those processes in the general case. We next narrow down our discussion to three popular stochastic processes, deriving their specific expressions for the two types of skewness. We finally turn to the topic of estimating conditional and unconditional skewness from a stochastic process in the ASV class.

2.1 Affine Stochastic Volatility (ASV) Models

In our derivations, we assume that the value of an asset can be described using a stochastic process from the ASV model class, as formalized in Duffie et al. (2000, 2003), Chernov et al. (2003), Bates (2006), and Cheridito et al. (2007). In line with those studies, we define such processes as follows:

Assumption 1. *On a canonical filtered space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$, *let* $(X_t, V_t)_{t\geq 0}$ *denote a time-homogenous stochastically continuous Markov process taking values in the state space* $D = \mathbb{R} \times \mathbb{R}^d_+$ *for some* $d \in \mathbb{N}$, *where the scalar process* X_t *is understood as the observed log-value process and the d-dimensional vector-valued process* V_t *is the latent state process associated with* X_t. *We assume that:*

1. *There exist unique functions* $\phi(u, w, h) \in \mathbb{C}$ *and* $\psi(u, w, h) \in \mathbb{C}^d$ *such that the joint conditional moment generating function (MGF)* $M_t(u, w, h; V_t)$ *defined in Equation (2.1) has an exponential-linear form:*

$$M_t(u, w, h; V_t) \equiv \mathrm{E}[e^{r_{t,h}u+\langle V_{t+h},w\rangle}|X_t, V_t] = e^{\mu hu+\phi(u,w,h)+\langle\psi(u,w,h),V_t\rangle}, \tag{2.1}$$

   for all $(u, w, h) \in \mathbb{C} \times \mathbb{C}^d \times \mathbb{R}_+$ *and* $t \geq 0$ *if* $\mathrm{E}[.]$ *exists, and where* $r_{t,h} \equiv X_{t+h} - X_t$ *is the log-return of* X_t *over horizon* h *and* μ *the drift of* X_t.[5]

2. *As* $t \to \infty$, V_t *converges in law to a unique invariant limit distribution* $L(v)$ *with MGF given by* $G(w) = \mathrm{E}[e^{\langle w,V_t\rangle}]$.

3. *For all* $t, h \in (0, \infty)$, $M_t(3, 0, h; V_t) < \infty$ *almost surely.*

Condition (1) follows Duffie et al. (2003) in defining an ASV model which is homogenous of order one in X_t. Assuming the existence of the limit distribution

[5] To be more precise, the MGF functions of the ASV models in this paper refer to the analytical continuation of the corresponding characteristic functions to the complex plane.

$L(v)$, Condition (2) ensures that a stationary regime for the Markov system exists and thus that the unconditional moments of X_t are well defined (see Keller-Ressel and Steiner (2008) and Keller-Ressel (2011) for more details). In turn, the third moment of X_t does not explode over any finite horizon h, implying the skewness of the dollar return over that horizon exists (see Andersen and Piterbarg (2007), Keller-Ressel (2011), and Keller-Ressel et al. (2015) for details).

Intuitively, $M_t(u, w, h; V_t)$ can be viewed as the MGF for $(r_{t,h}, V_{t+h})$ conditional on the latent state vector V_t. As we discuss below, $M_t(u, w, h; V_t)$ is inconvenient to use empirically since V_t is unobservable in practice. To avoid that issue, we could focus on the unconditional MGF, computed by interpreting V_t as a random variable with (limiting) distribution $L(v)$ and then taking the expectation of the conditional MGF with respect to that variable:

$$M(u, w, h) \equiv \mathrm{E}[M_t(u, w, h; V_t)] = e^{\mu h u + \phi(u,w,h)} G(\psi(u, w, h)). \tag{2.2}$$

Doing so, we can, however, only estimate unconditional skewness, not conditional skewness, as we will understand more fully in the next section. To keep our notation simple, we from here on suppress the MGF parameter w whenever it is equal to zero (i.e., $M_t(u, h) \equiv M_t(u, 0, h)$ and $M(u, h) \equiv M(u, 0, h)$).

2.2 Calculating True Skewness from ASV Models

2.2.1 The General Case

We next derive general closed-form or quasi-closed-form solutions for the skewness of an asset's dollar return over some horizon under the assumptions that the value of the asset can be modelled using a stochastic process from the ASV class and that the process parameters are known. To do so, we define the asset's gross dollar return measured from time t and with horizon h as $R_{t,h} \equiv e^{r_{t,h}} = e^{X_{t+h} - X_t}$, which is the ratio of the asset's value at time $t+h$ to its value at time t.

We define the unconditional skewness of the dollar return as:

$$\text{Skew}[R_{t,h}] \equiv \frac{\text{E}[(R_{t,h} - \text{E}[R_{t,h}])^3]}{\text{V}[R_{t,h}]^{3/2}}, \tag{2.3}$$

which can be interpreted as our best estimate of skewness in the absence of any information on V_t. Conversely, we define the conditional skewness as:

$$\text{Skew}[R_{t,h}|\mathcal{G}_t] \equiv \frac{\text{E}[(R_{t,h} - \text{E}[R_{t,h}|\mathcal{G}_t])^3|\mathcal{G}_t]}{\text{V}[R_{t,h}|\mathcal{G}_t]^{3/2}}, \tag{2.4}$$

where $\mathcal{G} = (\mathcal{G}_t)_{t\geq 0} \subset \mathcal{F}$ denotes the filtration generated by X_t. Equation (2.4) can be interpreted as our best estimate of skewness conditional on the observations available at time t. Notice that when X_t is continuously observed over time we can perfectly estimate the latent state process V_t, so that $\text{Skew}[R_{t,h}|\mathcal{G}_t]$ is equal to $\text{Skew}[R_{t,h}|V_t]$, the full-information conditional skewness.

Using Assumption (1) in Section 2.1, we are able to rewrite the unconditional skewness of the dollar return in Equation (2.3) as:

$$\text{Skew}[R_{t,h}] = \frac{M(3,h) - 3M(1,h)M(2,h) + 2M(1,h)^3}{[M(2,h) - M(1,h)^2]^{3/2}}, \tag{2.5}$$

and the conditional skewness of the dollar return in Equation (2.4) as:

$$\text{Skew}[R_{t,h}|\mathcal{G}_t] = \frac{M_t(3,h) - 3M_t(1,h)M_t(2,h) + 2M_t(1,h)^3}{[M_t(2,h) - M_t(1,h)^2]^{3/2}}. \tag{2.6}$$

2.2.2 Some Specific Examples

We now show the specific solutions for the unconditional and conditional skewness in case that the asset's value can be described using geometric Brownian motion (GBM), the Heston process, and the multifactor-Heston process.

Geometric Brownian motion. Suppose $(X_t, V_t)_{t \geq 0}$ satisfies:

$$X_t = X_0 + (\mu - \frac{1}{2}\sigma^2)t + \sqrt{V_t} dW_t \quad \text{and} \quad V_t = \sigma^2, \tag{2.7}$$

where μ and σ are parameters, and W_t is a Brownian motion. Since V_t is constant, the conditional and unconditional MGFs coincide and equal:

$$M_t(u, h) = M(u, h) = e^{(\mu h - \frac{1}{2}\sigma^2 h)u + \frac{1}{2}\sigma^2 h u^2}. \tag{2.8}$$

In turn, the conditional and unconditional skewness coincide and equal:

$$\text{Skew}[R_{t,h}|\mathcal{G}_t] = \text{Skew}[R_{t,h}] = \frac{e^{3h\sigma^2} - 3e^{h\sigma^2} + 2}{(e^{h\sigma^2} - 1)^{3/2}}, \tag{2.9}$$

which is derived from plugging Equation (2.8) into (2.5) and simplifying. To wit, Equation (2.9) is similar to Farago and Hjalmarsson's (2019) estimator for dollar return skewness, which also relies on the assumption that the short-horizon dollar return is i.i.d. but does not require it to be lognormally distributed.

The Heston model. To enable an asset's volatility to evolve stochastically over time and to allow the asset's value and volatility to be correlated, Heston (1993) assumes that $(X_t, V_t)_{t \geq 0}$ satisfies:

$$\begin{aligned} dX_t &= (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t} dW_t, \\ dV_t &= \kappa(\alpha - V_t)dt + \xi\sqrt{V_t} dB_t, \end{aligned} \tag{2.10}$$

where κ, α, and ξ are the mean reversion, the long-run variance, and the volatility-of-volatility parameter respectively, and W_t and B_t are Brownian motions with $[W, B]_t = \rho t$. Setting $\rho < 0$ creates a negative correlation between the asset's value and volatility, consistent with Black's (1976) leverage effect. To ensure that V_t is positive, we require $\kappa > 0$, $\alpha > 0$, and $\xi > 0$ as well as $2\kappa\alpha > \xi^2$.

Bates (2006), Andersen (2008), and Rollin et al. (2009) document that the conditional MGF of the Heston stochastic process is:

$$M_t(u, w, h; V_t) = e^{\mu h u + \phi(u,w,h) + \psi(u,w,h)V_t}, \tag{2.11}$$

where:

$$\begin{aligned}
\phi(u, w, h) &= \kappa\theta\left(\alpha_+(u)h + \frac{\alpha_-(u) - \alpha_+(u)}{P(u)} \ln\left(\frac{Q(u, w) - e^{P(u)h}}{Q(u, w) - 1}\right)\right), \\
\psi(u, w, h) &= \frac{Q(u, w)\alpha_+(u) - \alpha_-(u)e^{P(u)t}}{Q(u, w) - e^{P(u)t}}, \quad P(u) = \sqrt{(\kappa - \xi\rho u)^2 + \xi^2(u - u^2)}, \\
\alpha_\pm &= (\kappa - u\rho\xi \pm P(u))/\xi^2, \quad \text{and} \quad Q(u, w) = \frac{\alpha_-(u) - w}{\alpha_+(u) - w}.
\end{aligned} \tag{2.12}$$

Since $(V_t)_{t\geq 0}$ obeys a Gamma distribution in the Heston process in the time limit, the function $G(w)$ in Equation (2.2) takes the form:

$$G(w) \equiv \mathrm{E}[e^{wV_t}] = \left(1 - \frac{\xi^2}{2\kappa}w\right)^{-2\kappa\alpha/\xi^2}, \tag{2.13}$$

and the unconditional MGF is equal to:

$$M(u, w, h) = e^{\mu u h + \phi(u,w,h)}G(\psi(u, w, h)). \tag{2.14}$$

The multi-factor Heston model. To allow for a stochastic volatility model with even more flexibility, Christoffersen et al. (2009) propose the multi-factor Heston model, which generalizes the Heston model by modelling variance using $K > 1$ stochastic variables. In particular, $(X_t, V_t)_{t\geq 0}$ now satisfies:

$$\begin{aligned}
dX_t &= (\mu - \frac{1}{2}\sum_{k=1}^{K} V_t^{(k)})dt + \sum_{k=1}^{K}\sqrt{V_t^{(k)}}dW_t^{(k)}, \\
V_t^{(k)} &= \kappa^{(k)}(\alpha^{(k)} - V_t^{(k)})dt + \xi^{(k)}\sqrt{V_t^{(k)}}dB_t^{(k)}, \quad k \in [1, 2, \ldots, K],
\end{aligned} \tag{2.15}$$

where $[W^{(k)}, B^{(k)}]_t = \rho^{(k)} t$, $[W^{(k)}, W^{(k')}]_t = [B^{(k)}, W^{(k')}]_t = [B^{(k)}, B^{(k')}]_t = 0$ for all $k \neq k'$, and the parameters μ, $\kappa^{(k)}$, $\alpha^{(k)}$, $\xi^{(k)}$, and $\rho^{(k)}$ are defined as before. Since each $V_t^{(k)}$ evolves independently of the others, the conditional MGF of the multi-factor Heston process takes the form:

$$M_t(u, w, h; \theta, V_t) = e^{\mu h u + \sum_{k=1}^{K} \phi^{(k)}(u,w,h) + \sum_{k=1}^{K} \psi^{(k)}(u,w,h) V_t^{(k)}}, \tag{2.16}$$

where $\phi^{(k)}(u, w, h)$ and $\psi^{(k)}(u, w, h)$ are as in Equation (2.12) assuming $\kappa = \kappa^{(k)}$, $\alpha = \alpha^{(k)}$, $\xi = \xi^{(k)}$, and $\rho = \rho^{(k)}$. We further have that:

$$G(w) = \prod_{k=1}^{K} G^{(k)}(w), \tag{2.17}$$

where $G^{(k)}(w)$ is as in Equation (2.13) under the same assumption as above.

Other ASV models. Turning to other ASV models, Theorem 2.7 in Duffie et al. (2003) suggests that $\phi(u, w, h)$ and $\psi(u, w, h)$ can always be obtained as the solution of a system of general Riccati equations. Moreover, $G(w)$ can be derived as a by-product of that solution, as, for example, done in Keller-Ressel (2011). While the Ricatti-equation system does not always have a closed-form solution, it is always possible to solve it using numerical methods. As a result, we can always calculate dollar-return skewness from Equations (2.5) and (2.6) if the asset's value obeys an ASV-class stochastic process — even if we looked into processes more complicated and sophisticated than those explicitly discussed in this section (such as affine jump-diffusion processes as in e.g. Eraker et al. (2003) and Bates (2006)).[6]

[6] While only few studies derive closed-form solutions for the MGFs of general stochastic processes from the ASV class, the reason may be a lack of applications. We note, for example, that we only require $\phi(u, 0, h)$ and $\psi(u, 0, h)$ for the purpose of option pricing, with these functions being much easier to compute than their general (i.e., $w \neq 0$) counterparts. Pan (2002), Bates (2006), and Bates (2012) show how to derive MGFs for ASV-class stochastic processes with $K = 1$, while Bates (2019) discusses how to derive those MGFs for models with $K > 1$.

2.3 Estimating Skewness from ASV Models

Since we can calculate the skewness of an asset's dollar return assuming that the asset's value obeys an ASV-class stochastic process and the process parameters are known, we next show that we can consistently and efficiently estimate and forecast that skewness simply by estimating the process parameters. Assuming that we work with daily or intra-day price data, we let $(X_t)_{t\geq 0}$ be equidistantly observed on some finite interval $[0, T]$. To be more precise, denoting some fixed time interval by Δ (e.g., $\Delta = 1/252$ for one trading day), we observe $(X_{i\Delta})_{i=1:N}$, where $N = T/\Delta$ is the sample size. Since Δ is fixed, we, however, write X_i rather than $X_{i\Delta}$ to keep our notation as simple as possible. We use $\mathcal{G}_i = \sigma(X_j; j \in \{0, \dots, i\})$ to denote the filtration generated by $(X_j)_{j=1:i}$, with $i \leq N$.

To be able to estimate the parameters of ASV-class stochastic processes and to convert those into dollar return skewness estimates and forecasts, we impose the following additional assumptions on top of those in Assumption (1):

Assumption 2. *Consider an ASV model generating $(X_t, V_t)_{t\geq 0}$ defined as in Assumption 1. We further assume about that model that:*

1. *The model is uniquely determined by some k-dimensional parameter vector $\theta \in \Theta$, where the associated parameter space Θ is a compact subspace of $\mathbb{R}^k$. For some realization of $(X_t, V_t)_{t\geq 0}$, we denote the true data generating parameter vector by θ_0, taking values from the interior of Θ.*

2. *There exists an N-consistent estimator of θ_0 adapted to $\mathcal{G}_N$, which we denote by $\hat{\theta}$, such that $\hat{\theta} \xrightarrow{p} \theta_0$ as $N \to \infty$.*

3. *The unconditional and conditional moments $\mathrm{E}[R_{t,h}^k]$ and $\mathrm{E}[R_{t,h}^k|\mathcal{G}_t]$ are uniformly continuous in θ for all $h > 0$ and $k = 1, 2, 3$.*

Conditions (1) and (2) ensure that a consistent estimator of the parameter vector θ_0 can be constructed based on equidistantly observed data. Given that

Bates (2006) shows that numerous estimation techniques (including, e.g., GMM, the efficient method of moments (EMM), and the empirical characteristic function method) consistently estimate the parameters of a wide class of ASV models, the two conditions are not restrictive. Condition (3) ensures the exponential moments of an ASV model are smooth functions of the parameter vector, which is also satisfied by the vast majority of ASV models. The final condition guarantees that plugging the consistent estimates of an ASV-class stochastic process into our skewness formulas (2.5) and (2.6) yields consistent skewness estimates.

Under Assumptions (1) and (2), we can estimate the unconditional skewness of an asset's dollar return over time $i\Delta$ to $i\Delta + h$, $\text{Skew}[R_{i,h}; \hat{\theta}]$, using:

$$\text{Skew}[R_{i,h}; \hat{\theta}] = \frac{M(3, h; \hat{\theta}) - 3M(1, h; \hat{\theta})M(2, h; \hat{\theta}) + 2M(1, h; \hat{\theta})^3}{[M(2, h; \hat{\theta}) - M(1, h; \hat{\theta})^2]^{3/2}}. \tag{2.18}$$

Since we are unable to exactly infer V_i to calculate the value of the full-information conditional MGF, $M_i(u, w, h; V_i, \hat{\theta})$, with equidistantly observed data, we devise an optimal estimate $\bar{M}_i(u, w, h; \hat{\theta}) \equiv \text{E}[M_i(u, w, h; V_i, \hat{\theta})|\mathcal{G}_i]$ instead. The observed conditional skewness, $\text{Skew}[R_{i,h}; \hat{\theta}|\mathcal{G}_i]$, is then computed using:

$$\text{Skew}[R_{i,h}; \hat{\theta}|\mathcal{G}_i] = \frac{\bar{M}_i(3, h; \hat{\theta}) - 3\bar{M}_i(1, h; \hat{\theta})\bar{M}_i(2, h; \hat{\theta}) + 2\bar{M}_i(1, h; \hat{\theta})^3}{[\bar{M}_i(2, h; \hat{\theta}) - \bar{M}_i(1, h; \hat{\theta})^2]^{3/2}}. \tag{2.19}$$

While the optimal estimate $\bar{M}_i(u, w, h; \hat{\theta})$ is generally unavailable in closed-form, the following proposition from Bates (2006) allows us to iteratively calculate it:

Proposition 1. *Under Assumptions (1) and (2), assume that V_0 is a random draw from $L(v)$ and X_0 is a deterministic initial log-price. Let $G_i(w|\mathcal{G}_i) = \text{E}[e^{\langle w, V_i \rangle}|\mathcal{G}_i]$ denote the MGF of V_i conditioning on $\mathcal{G}_i$. It then holds that:*

$$\bar{M}_i(u, w, h; \hat{\theta}) = e^{\mu hu + \phi(u,w,h)} G_i(\psi(u, w, h)|\mathcal{G}_i), \tag{2.20}$$

$$G_{i+1}(w|\mathcal{G}_{i+1}) = \frac{\int_{0-i\infty}^{0+i\infty} \bar{M}_i(u, w, h; \hat{\theta}) e^{-ur_{i,1}} du}{\int_{0-i\infty}^{0+i\infty} \bar{M}_i(u, 0, h; \hat{\theta}) e^{-ur_{i,1}} du}, \tag{2.21}$$

where $G_0(w|\mathcal{G}_0) = G(w)$.

In practice, we always start with Equation (2.20), calculating $\bar{M}_0(u, w, h; \hat{\theta}) = e^{\mu hu + \phi(u,w,h)} G(w)$. We next plug $\bar{M}_0(u, w, h; \hat{\theta})$ into Equation (2.21), numerically evaluating the integrals and obtaining $G_1(w|\mathcal{G}_1)$. We then plug $G_1(w|\mathcal{G}_1)$ back into Equation (2.20), allowing us to compute the optimal expected MGF at time Δ, $\bar{M}_1(u, w, h; \hat{\theta})$. We continue in that manner until we have the optimal expected MGFs for all dates in our data (i.e., for all $i \in \{1, 2, \ldots, N\}$).

We note that a direct implementation of Proposition (1) suffers from the curse of dimensionality due to the iterative nature of the numerical integration. To avoid that problem, we follow Bates (2006) and approximate $G_i(w|\mathcal{G}_i)$ by the MGF of a Gamma distribution with parameters estimated numerically at every iteration. Despite the complexity, the approach is numerically stable, and filtration errors die out with an increasing number of observations.

Following from condition (3) of Assumption (2), the continuous mapping theorem implies that $\text{Skew}[R_{i,h}; \hat{\theta}]$ and $\text{Skew}[R_{i,h}; \hat{\theta}|\mathcal{G}_i]$ are consistent estimators of, respectively, $\text{Skew}[R_{i,h}; \theta_0]$ and $\text{Skew}[R_{i,h}; \theta_0|\mathcal{G}_i]$, as stated in the next proposition:

Proposition 2. *Under Assumptions (1) and (2), it holds as* $N \to \infty$ *that:*

$$\begin{aligned} \text{Skew}[R_{i,h}; \hat{\theta}] &\xrightarrow{p} \text{Skew}[R_{i,h}; \theta_0], \\ \text{Skew}[R_{i,h}; \hat{\theta}|\mathcal{G}_i] &\xrightarrow{p} \text{Skew}[R_{i,h}; \theta_0|\mathcal{G}_i]. \end{aligned} \tag{2.22}$$

We can also establish that our skewness estimates are asymptotically normal if $\hat{\theta}$ is so, too. Since establishing that $\hat{\theta}$ is asymptotically normal, however, requires additional assumptions, we omit those derivations for the sake of brevity.

3 A Heston-Type Estimator Based On Our Setup and Other Competing Estimators

In this section, we study a Heston-type estimator of the skewness of long-horizon returns based on the setup in Section 2, illustrating how we can use a novel GMM approach to calculate that estimator in practice. We also introduce estimators of the skewness of long-horizon returns advocated in recent studies.

3.1 The Heston-Type Estimator

3.1.1 Using GMM to Calculate the Heston-Type Estimator

We devise a Heston-type estimator of the conditional (unconditional) skewness of long-horizon dollar returns by plugging the conditional (unconditional) MGFs of the Heston stochastic process shown in Equation (2.11) ((2.14)) into skewness formula (2.6) ((2.5)) and evaluating the formula at estimates of the parameter values of the process, $\theta \equiv [\mu, \kappa, \alpha, \xi, \rho]$. To obtain the parameter estimates, we propose a novel GMM approach matching the theoretical central moments and cross-moments of $R_{i,\Delta}$ with their sample counterparts. To that end, denote the centered version of $R_{i,\Delta}$ as $\tilde{R}_{i,\Delta} \equiv R_{i,\Delta} - M(1,\Delta)$. Next denote the centered MGF and cross-MGF of $R_{i,\Delta}$ as, respectively, $\tilde{M}(k,\Delta) \equiv \mathrm{E}[\tilde{R}^k_{i,\Delta}]$ and $\tilde{C}(m,n,\Delta,j) \equiv \mathrm{Cov}[\tilde{R}^m_{i,\Delta}, \tilde{R}^n_{i+j,\Delta}]$, where j indicates the number of lags. We can then write the system of moment conditions used in our GMM approach as:

$$\mathrm{E}[h(R_{i,\Delta};\theta)] = 0, \tag{3.1}$$

where $h(R_{i,\Delta};\theta)$ is given by the following vector:

$$h(R_{i,\Delta};\theta) = \begin{pmatrix} R_{i,\Delta} - M(1,\Delta;\theta) \\ \tilde{R}^2_{i,\Delta} - \tilde{M}(2,\Delta;\theta) \\ \tilde{R}^3_{i,\Delta} - \tilde{M}(3,\Delta;\theta) \\ \tilde{R}^4_{i,\Delta} - \tilde{M}(4,\Delta;\theta) \\ \tilde{R}^1_{i,\Delta}\tilde{R}^2_{i+1,\Delta} - \tilde{C}(1,2,\Delta,1;\theta) \\ \vdots \\ \tilde{R}^1_{i,\Delta}\tilde{R}^2_{i+J,\Delta} - \tilde{C}(1,2,\Delta,J;\theta) \\ \tilde{R}^2_{i,\Delta}\tilde{R}^2_{i+1,\Delta} - \tilde{C}(2,2,\Delta,1;\theta) \\ \vdots \\ \tilde{R}^2_{i,\Delta}\tilde{R}^2_{i+J,\Delta} - \tilde{C}(2,2,\Delta,J;\theta) \end{pmatrix}_{(4+2J)\times 1}, \tag{3.2}$$

and J represents the longest lag between returns.

Expanding the definitions of $\tilde{M}(k,\Delta;\theta)$ and $\tilde{C}(m,n,\Delta,j;\theta)$ in vector (3.2), we show in the appendix that those centered MGFs and cross-MGFs can be written as linear functions of non-centered MGFs and cross-MGFs:

$$\begin{aligned}
\tilde{M}(2,\Delta) &= M(2,\Delta) - M^2(1,\Delta), && (3.3)\\
\tilde{M}(3,\Delta) &= M(3,\Delta) - 3M(1,\Delta)M(2,\Delta) + 2M^3(1,\Delta), && (3.4)\\
\tilde{M}(4,\Delta) &= M(4,\Delta) - 4M(1,\Delta)M(3,\Delta) + 6M^2(1,\Delta)M(2,\Delta) \\
&- 3M^4(1,\Delta), && (3.5)\\
\tilde{C}(1,2,\Delta,j) &= C(1,2,\Delta,j) - M(1,\Delta)M(2,\Delta), && (3.6)\\
\tilde{C}(2,2,\Delta,j) &= C(2,2,\Delta,j) - 2M(1,\Delta)C(1,2,\Delta,j) - M^2(2,\Delta) \\
&+ 2M^2(1,\Delta)M(2,\Delta), && (3.7)
\end{aligned}$$

where the non-centered cross-MGFs $C(m,n,\Delta,j)$ can be calculated from:

$$C(m,n,\Delta,j) = e^{\mu\Delta n+\phi(n,0,\Delta)+\phi(0,\psi(n,0,\Delta),j)} M(m,\psi(0,\psi(n,0,\Delta),j\Delta),\Delta).^{7} \tag{3.8}$$

Although our moment conditions include all first four central moments in the first four rows, they only include two cross-moments, namely those involving $\tilde{R}_{i,\Delta}\tilde{R}^2_{i+j,\Delta}$ and $\tilde{R}^2_{i,\Delta}\tilde{R}^2_{i+j,\Delta}$, starting from the fifth row. While the cross-moments involving $\tilde{R}_{i,\Delta}\tilde{R}^2_{i+j,\Delta}$ are meant to capture leverage effects, the ones involving $\tilde{R}^2_{i,\Delta}\tilde{R}^2_{i+j,\Delta}$ are meant to capture volatility clustering effects, two well-documented stylized facts of asset returns. We do not include other cross-moments since most of them are likely to be close to zero, making it hard to identify parameters from them. The cross-moments involving $\tilde{R}_{i,\Delta}\tilde{R}_{i+j,\Delta}$ will, for example, always approximately hold in a weak-form efficient market since the history of dollar returns does not predict future dollar returns in such a market.

As the system of moment conditions (3.1) is overidentified even when $J = 1$, we obtain parameter estimates by minimizing a quadratic form of the moment conditions. To that end, let $\boldsymbol{W}_{(4+2J)\times(4+2J)}$ denote a positive definite weighting matrix. We then define the GMM estimator of the parameter vector θ as:

$$\hat{\theta}_{GMM} = \underset{\theta}{\operatorname{argmin}}\, \bar{h}(R_{i,\Delta};\theta)'\boldsymbol{W}_{(4+2J)\times(4+2J)}\bar{h}(R_{i,\Delta};\theta), \tag{3.9}$$

where $\bar{h}(R_{i,\Delta};\theta) = \frac{1}{N}\sum_{i=1}^{N} h(R_{i,\Delta};\theta)$. Under standard regularity conditions (as discussed in, e.g., Chapter 3 in Hall (2005)), $\hat{\theta}_{GMM}$ is a consistent and asymptotically normal estimator of θ_0. Fixing the horizon h, the delta method suggests that our conditional and unconditional skewness estimators are also consistent and asymptotically normal estimators of their true counterparts.

[7] Derivation can be found in Appendix A.2.

3.1.2 Practical Implementation

We use a two-step approach with $J = 100$ to estimate the parameter vector θ in practice. In particular, we start with employing the identity matrix $\boldsymbol{I}_{204\times204}$ as weighting matrix $\boldsymbol{W}_{(4+2J)\times(4+2J)}$ in Equation (3.9) to obtain the preliminary parameter estimate vector $\hat{\theta}_1$. Relying on the vector $\hat{\theta}_1$, we compute the Newey and West (1987) variance-covariance matrix of $h(R_{i,\Delta};\hat{\theta}_1)$, often labelled the spectral density matrix. We finally use the inverted spectral density matrix as weighting matrix $\boldsymbol{W}_{(4+2J)\times(4+2J)}$ to obtain our final parameter estimate vector $\hat{\theta}_2$. In addition to the Feller condition (i.e. $2\kappa\alpha > \xi^2$), we always impose the constraints that $\kappa \in (0, 10]$, $\alpha \in (0, 1]$, $\xi \in (0, 1.5]$, and $\rho \in [-1, 0]$ to ensure reasonable estimates. We also impose $\rho \leq -\sqrt{30}/6$ or $\kappa/\xi \geq 6\rho + \sqrt{30}$ to ensure $M_t(6, \Delta; V_t) < \infty$ almost surely, implying a well defined variance of skewness. [8]

To increase computational efficiency and improve parameter identification, we further do not use the MGFs and cross-MGFs in Equations (3.3) to (3.7) in our GMM estimations but instead their short-time-increment power-expansion approximations. We show those approximations in Proposition 3.

Proposition 3. *As $\Delta \to 0$, it is the case that:*

1. $M(1, \Delta) = 1 + \mu\Delta + \frac{\mu^2}{2}\Delta^2 + o(\Delta^2)$.
2. $\tilde{M}(2, \Delta) = \alpha\Delta + \alpha(\frac{\xi^2}{4\kappa} + \frac{\alpha}{2} + 2\mu + \xi\rho)\Delta^2 + o(\Delta^2)$.
3. $\tilde{M}(3, \Delta) = 3\alpha(\alpha + \frac{\xi^2}{2\kappa} + \frac{\xi\rho}{2})\Delta^2 + o(\Delta^2)$.
4. $\tilde{M}(4, \Delta) = 3\alpha(\alpha + \frac{\xi^2}{2\kappa})\Delta^2 + o(\Delta^2)$.
5. $\tilde{C}(1, 2, \Delta, j) = \alpha\xi\rho e^{-\kappa(j-1)\Delta}\Delta^2 + o(\Delta^2)$.
6. $\tilde{C}(2, 2, \Delta, j) = \frac{\alpha\xi^2}{2\kappa}e^{-\kappa(j-1)\Delta}\Delta^2 + o(\Delta^2)$.

[8]Further explanation is provided in Appendix A.3.

Interestingly, the approximations in Proposition (3) often allow for new insights into the relations between the moments and the parameters. The expression for $\tilde{C}(1, 2, \Delta, j)$, the covariance between lagged return and current squared return, for example, suggests that $\rho < 0$ is a necessary condition for the leverage effect (i.e., for higher returns to predict a lower volatility, and vice versa).[9]

3.2 Fama and French's (2018) Bootstrap Estimator

Fama and French (2018) suggest a simple bootstrap to estimate the skewness of long-horizon returns. They start from a time-series of short-horizon (e.g., daily or weekly) returns. They next draw with replacement and equal probabilities a number of returns from that time-series sufficient for those returns to cover a period equal to the desired long horizon (e.g., one year). They then compound the drawn returns to generate a "bootstrap long-horizon return." Repeating the prior two steps multiple times, they generate a sample of bootstrap long-horizon returns. They finally compute the skewness of the long-horizon return by applying the sample skewness to the sample of bootstrap long-horizon returns.

While Fama and French's (2018) estimator is original, their simple bootstrap abstracts from dependence in returns, implicitly assuming that returns are i.i.d. over time. Given the i.i.d. assumption, an estimate obtained from their estimator is simultaneously a conditional and unconditional estimate.

3.3 Farago and Hjalmarsson's (2019) Estimator

Assuming that the short-horizon return, $R_{i,\Delta}$, is i.i.d., Farago and Hjalmarsson (2019) demonstrate that the skewness of the long-horizon return, which is the

[9] We generate the results in Proposition 3 using Wolfram Mathematica, offering our programming codes in the supplement material for this paper.

compounded up short-horizon return $\prod_{i=1}^{d} R_{i,\Delta}$, can be computed from:

$$\text{Skew}\left[\prod_{i=1}^{d} R_{i,\Delta}\right] = \frac{\theta_3^d - 3\theta_2^d + 2}{(\theta_2^d - 1)^{3/2}}, \tag{3.10}$$

where $\theta_2 = \frac{\text{Var}[R_{i,\Delta}]}{\text{E}[R_{i,\Delta}]^2} + 1$, $\theta_3 = -2 + 3\theta_2 + \text{Skew}[R_{i,\Delta}](\theta_2 - 1)^{3/2}$ and d denotes the number of short-horizon periods within the long-horizon. As before, the i.i.d. assumption implies that the estimator (3.10) simultaneously gives a conditional and unconditional estimate. In essence, Farago and Hjalmarsson's (2019) estimator is the closed-form equivalent of Fama and French's (2018) estimator.

3.4 Sample Skewness Estimator

The sample skewness estimator is defined as:

$$\text{Skew}[R_{i,\Delta}] = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(R_{i,\Delta} - \frac{1}{N}\sum_{i=1}^{N} R_{i,\Delta}\right)^3}{\left(\frac{1}{N}\sum_{i=1}^{N}\left(R_{i,\Delta} - \frac{1}{N}\sum_{i=1}^{N} R_{i,\Delta}\right)^2\right)^{3/2}}, \tag{3.11}$$

where N is the number of returns $R_{i,\Delta}$ in our sample period. To ensure that we have enough observations when we apply the sample skewness estimator to long-horizon returns, we consistently compute it on overlapping returns.

4 Simulation Exercise

In this section, we conduct a simulation exercise to investigate the performance of the Heston-type estimator outlined in Section 3 and to compare it with the other estimators advocated in the recent literature. We start with describing the data generating process, computing true skewness from that process, and studying how well we can estimate the process parameters using the novel GMM method developed by us. We next compare the unbiasedness and efficiency of the Heston-type estimator and the others in case asset values obey the Heston stochastic process. We finally repeat the former comparisons in case asset values obey a more realistic and flexible process, the double-Heston process.

4.1 Data Generating Process, Implied True Skewness, and Heston Process Parameter Estimation

We follow Broadie and Kaya (2006) and Andersen (2008) in simulating 10,000 sample paths with a ten-year length from the Heston-class processes (i.e. Equations (2.10) and (2.15)). To achieve that, we assume that the initial price for each path, P_0, is 50. Moreover, we draw the initial variance of each factor for each path, $V_0^{(k)}$, from their asymptotic Gamma distributions, $\Gamma(2\kappa^{(k)}\alpha^{(k)}/(\xi^{(k)})^2, 2\kappa^{(k)}/(\xi^{(k)})^2)$. We set the discretization step equal to one day (i.e., $\Delta = 1/252$), implying that μ, $\alpha^{(k)}$, and $\xi^{(k)}$ are all stated per annum and that $N = 2,520$. We finally employ the following four-step approach to consecutively simulate observations for each sample time $i \in \{1, \dots, N\}$:

1. We follow Cox et al. (2005) in simulating $V_i^{(k)}$ based on $V_{i-1}^{(k)}$ using:

$$V_i^{(k)} = \frac{\left(\xi^{(k)}\right)^2 \left(1 - e^{-\kappa^{(k)}\Delta}\right)}{4\kappa^{(k)}} \chi^2_{d,\lambda}, \tag{4.1}$$

   where $d = 4\kappa^{(k)}\alpha^{(k)}/(\xi^{(k)})^2$ and $\lambda = \frac{4\kappa^{(k)}e^{-\kappa^{(k)}\Delta}}{(\xi^{(k)})^2(1-e^{-\kappa^{(k)}\Delta})} V_{i-1}^{(k)}$ are, respectively, the degrees of freedom and the noncentrality parameter of the noncentral chi-square variable $\chi^2_{d,\lambda}$.

2. We simulate a standard Gaussian random variable $Z^{(k)}$ for each factor.

3. We follow Andersen (2008) in updating P_i by approximating the solution

to the Heston-class stochastic differential equation for the asset value as:

$$P_i = P_{i-1} \exp\left[\mu\Delta + \sum_{k=1}^{K} f\left(V^{(k)}\right)\right], \tag{4.2}$$

where

$$\begin{aligned} f\left(V^{(k)}\right) &= K_0^{(k)} + K_1^{(k)} V_{i-1}^{(k)} + K_2^{(k)} V_i^{(k)} + \sqrt{K_3^{(k)}(V_{i-1}^{(k)} + V_i^{(k)})} \cdot Z^{(k)}, \\ K_0^{(k)} &= -\frac{\rho^{(k)}\kappa^{(k)}\alpha^{(k)}}{\xi^{(k)}}\Delta, \quad K_1^{(k)} = \frac{1}{2}\Delta\left(\frac{\kappa^{(k)}\rho^{(k)}}{\xi^{(k)}} - \frac{1}{2}\right) - \frac{\rho^{(k)}}{\xi^{(k)}}, \\ K_2^{(k)} &= \frac{1}{2}\Delta\left(\frac{\kappa^{(k)}\rho^{(k)}}{\xi^{(k)}} - \frac{1}{2}\right) + \frac{\rho^{(k)}}{\xi^{(k)}}, \quad K_3^{(k)} = \frac{1}{2}\Delta\left(1 - \left(\rho^{(k)}\right)^2\right). \end{aligned} \tag{4.3}$$

4. We calculate the daily dollar return $R_{i,\Delta} = P_i/P_{i-1}$.

This four-step approach mitigates Euler discretization errors, reducing their impacts on our simulation exercise.

We select $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, and $\rho = -0.50$ as basecase parameters in our Heston simulations. To investigate how variations in those parameter values affect our skewness estimation outcomes, we also consider the alternative values $\kappa = 1$ or 5, $\alpha = 0.25$, $\xi = 0.10$ or 0.50 and $\rho = -0.90$ or 0.00, in each case, however, varying only one of the parameters from its basecase value.

To understand how the Heston process parameters affect the unconditional skewness of the return and its relation with the return horizon, Figure 1 plots that skewness evaluated at the basecase parameter values and its variations over horizons extending up to five years. The figure vividly shows that, under most parameter value choices, unconditional skewness monotonically rises with the return horizon. The only exception is the case in which the asset return and volatility share a correlation (ρ) of –0.90. In that case, unconditional skewness initially falls and only later rises with the return horizon. The lesson to be learned is thus that, when the asset return-volatility correlation is sufficiently negative, the leverage effect can dominate the compounding effect at shorter horizons, whilst it is always dominated by the compounding effect at longer horizons.

[**Insert Figure 1 here**]

The figure further shows that, at each return horizon, unconditional skewness rises with long-run variance α and the asset return-volatility correlation ρ but falls with the volatility of variance ξ. The positive relation with long-run variance is due to an increase in that variance boosting the compounding effect, whereas the positive relation with the asset return-volatility correlation is due to an increase in that correlation diminishing the leverage effect. Interesingly, the strength of mean reversion κ has an ambiguous effect on unconditional skewness, with the effect being negative at short horizons but positive at long ones.

Given that the performance of our Heston-type estimator must hinge crucially on how well we are able to estimate the μ, κ, α, ξ, and ρ parameters, Table 1 offers descriptive statistics obtained from applying the GMM estimator developed in Section 3 on each of our 10,000 simulated sample paths. The descriptive statistics include the true parameter values, the mean estimates, the median estimates, the standard deviation of the estimates, and the mean squared error (MSE) under both the basecase parameter values and its variations. The table confirms that the GMM estimator usually yields mean and median estimates close to the true values and small standard deviations and MSEs, except for the speed of mean reversion κ. The higher standard deviations and MSEs for κ presumably come from the likelihood surface being flat in κ, making it hard to infer that parameter independent of the estimation method applied (Atiya and Wall (2009)).

[**Insert Table 1 here**]

4.2 Unconditional Skewness Estimates in a Heston World

Table 2 contrasts the unbiasedness and efficiency of our Heston-type estimator and the others for the unconditional skewness of returns in case asset values obey the Heston stochastic process calibrated using the basecase parameter values or its variations. Given either parameter value set, the table gives the return horizon ("Horizon"), the true unconditional return skewness over a horizon ("True"), and the mean estimate ("Mean"), the absolute bias scaled by the true value ("%|Bias|"), and the mean squared error ("MSE") computed over the 10,000 sample paths for each estimator and horizon. The estimators are our Heston-type ("Ours"), the Fama-French (2018) bootstrap ("FF"), the Farago-Hjalmarsson (2019) closed-form ("FH"), and the sample skewness ("Conv.") estimator.

[Insert Table 2 here]

The table shows that our estimator strongly outperforms the other estimators except when the volatility of variance is close to zero ($\xi = 0.10$) or the asset return-volatility correlation is zero ($\rho = 0$). To be more specific, in the absence of the above two cases, our estimator always yields the smallest mean absolute percent error and the smallest MSE over all return horizons, with the small MSEs being particular noteworthy since the MSE reflects both bias and standard error of an estimate. The outperformance of our estimator is economically large. Considering the basecase and the one-year horizon, our estimator, for example, yields a mean absolute percent error at least 20 percentage points smaller than the others and an MSE of no more than 25% of those of the others. The worse performance of our estimator when $\xi = 0.10$ or $\rho = 0$ is attributable to asset returns being close to i.i.d. in those cases, implying that Fama and French's (2018) and Farago and Hjalmarsson's (2019) estimators gain statistical power from imposing reasonable restrictions. Even in those cases, the performance of our estimator is however, fortunately, not too far off from those of the other estimators.

To graphically compare the performance of the estimators, Figure 2 plots the evolution of true unconditional skewness and the mean estimates obtained from the estimators under the case $\rho = -0.9$ over return horizons up to 20 years. The results from the figure align with those from Table 2. In particular, the figure vividly shows that our estimator (dotted green line) yields a mean estimate consistently closest to true unconditional skewness (solid blue line). In contrast, the mean estimates from Fama and French's (2018; big red dots) and Farago and Hjalmarsson's (2019; violet broken line) estimator consistently overshoot true unconditional skewness, largely owing to their assumption that returns are i.i.d. and them thus abstracting from the leverage effect.[10] Notably, the extent of the overshooting rises with the horizon. Finally, the sample skewness (broken yellow line) is a decent estimator at short horizons. Yet, as we extend the horizon, it yields increasingly downward biased estimates, largely owing to the sample containing increasingly fewer independent return observations.[11]

[**Insert Figure 2 here**]

4.3 Conditional Skewness Estimates in a Heston World

Table 3 looks into the unbiasedness and efficiency of our Heston-type estimator for the *conditional* skewness of returns at the end of each simulated path in case asset values obey the Heston process calibrated using the same parameter value sets as before. While the table gives the same statistics as Table 2, for the sake of brevity, it omits the other (non-Heston-type) estimators since those produce conditional estimates identical to their corresponding unconditional estimates due to them assuming i.i.d. returns. In addition, the table further omits return

[10] We also find it interesting to note that Fama and French's (2018) and Farago and Hjalmarsson's (2019) estimators yield close to identical mean estimates up to ten year horizons, in line with our prior observation that Farago and Hjalmarsson's (2019) estimator is essentially the closed-form equivalent of Fama and French's (2018) bootstrap estimator.

[11] Notice that we cannot use the sample skewness to calculate the skewness of returns stretching over periods longer than ten years with ten years of sample data.

horizons shorter than one year. While we again estimate the process parameters underlying the skewness estimates from the entire ten years of data of a sample path, to speed up our computations, we use only the final two years of data from a path to compute the optimal estimates of the conditional MGFs, $\bar{M}_i(u, w, h; \hat{\theta})$, according to the iterative procedure outlined in Proposition 1.

We stress that the conditional estimates in this subsection are likely more relevant to a real-world investor than the unconditional estimates in the prior subsection. The reason is that such an investor is in most cases more interested in forecasting skewness over some concrete period (starting, e.g., from the current date) rather than over an equally long arbitrary period.

[**Insert Table 3 here**]

The table suggests that our Heston-type estimator also performs extremely well in estimating the conditional skewness of returns over long horizons. While true conditional skewness is now not a single number but varies over the 10,000 sample paths for each return horizon and parameter value set according to the value of the state variable (i.e., variance) at the end of a path, it is on average close to unconditional skewness (compare "True" columns in Tables 2 and 3). More crucially, our Heston-type estimates are generally close to conditional skewness, as demonstrated by the mean absolute percent biases and MSEs. Looking at the basecase and the five-year horizon, the mean absolute percent bias is, for example, only 10%, while the MSE is only 0.101. The one exception is the case in which long-run variance is high ($\alpha = 0.25$), in which our estimator finds it hard to capture the rapid rise of conditional skewness with the return horizon.

4.4 Unconditional Estimates in a Double-Heston World

Although our Heston-type estimator outperforms the others in case asset values obey the Heston process, there could be concern that the fact that our estimator assumes exactly the right asset value process grants it an unfair advantage over the other estimators. To wit, while it seems reasonable to argue that the Heston process goes some way toward modelling asset values in a more realistic way relative to i.i.d. processes, it is certainly not the case that it captures all stylized facts of asset values. To see how deviations between the process assumed by our estimator and the true process affect the performance of our estimator, Table 4 repeats our unbiasedness and efficiency tests for the unconditional estimates of the Heston-type estimator and the others in Table 2, this time, however, using a double-Heston process to simulate the 10,000 sample paths of data.

In contrast to the Heston process, the double-Heston process offers more flexibility in modelling the term structure of volatility (see Christoffersen et al. (2009)). Moreover, choosing a high value for the κ_1 mean reversion parameter in Equation (2.15) and a low value for the κ_2 parameter, the process is better able to fit the often different relations between short-term returns and variance and between long-term returns and variance in the data. To distinguish the double-Heston process in our simulations as much as possible from a Heston process, we however not only allow for variations in the mean reversion parameters but also in all others, setting the parameter value vector of the process, $\{\mu, \kappa_1, \alpha_1, \xi_1, \rho_1, \kappa_2, \alpha_2, \xi_2, \rho_2\}$, equal to $\{0.10, 1, 0.01, 0.10, -0.90, 5, 0.09, 0.50, -0.60\}$.

[**Insert Table 4 here**]

Using the same design as Table 2, Table 4 shows that our double-Heston process yields a slightly lower true unconditional skewness relative to our basecase Heston process over short return horizons, but a close to identical over longer horizons. More crucially, the table further reveals that our Heston-type estimator continues to strongly outperform the others in estimating unconditional skewness in terms of mean absolute percent bias and MSE. Looking at the five-year horizon, our estimator, for example, produces a mean absolute percent bias of 9% and an

MSE of 0.084, while the other estimators produce an mean absolute percent bias of at least 34% and an MSE of at least 0.654. All in all, we thus conclude that realistic deviations between the asset value process assumed by our estimator and the true process only marginally affect the performance of our estimator.

5 Empirical Application

We finally employ our Heston-type estimator as well as the others to estimate the unconditional skewness of the returns of the S&P 500, FTSE 100, and Nikkei 225 indexes over various return horizons, studying their evolutions and relations with the return horizon over time. To obtain the Heston-type estimates, we first estimate the Heston process parameters using ten-year rolling windows of daily returns and then plug the estimates into Equation (2.5) combined with Equation (2.14). We use the same windows of daily data to calculate the other estimators as described in Sections 3.2 to 3.4. The sample periods for the S&P 500, FTSE 100, and Nikkei 225 indexes are, respectively, 1950/01/03 to 2020/09/30, 1986/01/02 to 2020/09/30, and 1965/01/05 to 2020/09/30.

In Figure 3, we plot the unconditional skewness estimates obtained from the Heston-type estimator over time, with Panel A focussing on the S&P 500, Panel B on the FTSE 100, and Panel C on the Nikkei 225 index. In case of each index, the evolutions of the skewness estimates for different return horizons are highly correlated over time. Moreover, the weekly-return skewness estimates exceed the monthly-return skewness estimates in the vast majority of cases — and in some cases even the annual-return skewness estimates, suggesting the existence of a strong and time-varying leverage effect in stock indexes. The figure finally suggests that the skewness of long-horizon dollar returns is likely to be much lower than suggested in the recent work of Bessembinder (2018) and Farago and Hjalmarsson (2019), with even our estimates of the skewness of five-year returns never exceeding a value of three for any of the stock indexes.

[**Insert Figure 3 here**]

Figure 4 contrasts the unconditional monthly, annual, and five-year return skewness estimates from our Heston-type estimator (solid red line) with those from Fama and French's (2018; dotted green line), Farago and Hjalmarsson's (2019; broken violet line), and the sample skewness (broken blue line) estimators for each index. Raising confidence in our work, plots on the far left show that the skewness estimates from our estimator and the sample skewness are well aligned over short return horizons. Moving to the longer return horizons, the estimates from those two estimators continue to be aligned on average, with the sample skewness estimates however becoming much too volatile to be useful in practice, consistent with the simulation evidence in Table 2. In contrast, the estimates from Fama and French's (2018) and Farago and Hjalmarsson's (2019) estimators are more stable and correlate less strongly with the estimates from the two former estimators. Also consistent with our simulation evidence, they tend to yield higher skewness values than the Heston-type and sample skewness estimators.

[**Insert Figure 4 here**]

6 Concluding Remarks

In this paper, we derive a novel parametric estimator of the skewness of dollar returns from the assumption that asset values can be modelled using a stochastic process from the ASV model class. Using the Heston process as example, we run a simulation exercise comparing the unbiasedness and efficiency of our estimator with those of other existing estimators, namely Fama and French's (2018) boostrap estimator, Farago and Hjalmarsson's (2019) closed-form estimator as well as the sample skewness estimator. Our evidence suggests that our estimator strongly outperforms the others when asset values obey the Heston process or even some more complicated process, with it yielding the smallest mean absolute relative bias and MSE in the vast majority of cases. We finally apply our Heston-type estimator to real-world data on stock indexes, showing that there is an important time-varying leverage effect in that asset class and refuting the idea that the skewness of long-horizon returns is usually too high to be useful in practice.

References

Aït-Sahalia, Y., Kimmel, R., et al. (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413–452.

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, 118(1):135–167.

Andersen, L. B. (2008). Simple and efficient simulation of the Heston stochastic volatility model. *The Journal of Computational Finance*, 11(3):1–42.

Andersen, L. B. and Piterbarg, V. V. (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50.

Arditti, F. D. (1967). Risk and the required return on equity. *The Journal of Finance*, 22(1):19–36.

Atiya, A. F. and Wall, S. (2009). An analytic approximation of the likelihood function for the heston model volatility estimation problem. *Quantitative Finance*, 9(3):289–296.

Bakshi, G., Ju, N., and Ou-Yang, H. (2006). Estimation of continuous-time models with an application to equity volatility dynamics. *Journal of Financial Economics*, 82(1):227–249.

Bali, T. G. and Murray, S. (2013). Does risk-neutral skewness predict the cross section of equity option portfolio returns? *Journal of Financial and Quantitative Analysis*, pages 1145–1171.

Barberis, N. and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066–2100.

Bates, D. S. (2006). Maximum likelihood estimation of latent affine processes. *The Review of Financial Studies*, 19(3):909–965.

Bates, D. S. (2012). Us stock market crash risk, 1926–2010. *Journal of Financial Economics*, 105(2):229–259.

Bates, D. S. (2019). How crashes develop: Intradaily volatility and crash evolution. *The Journal of Finance*, 74(1):193–238.

Bessembinder, H. (2018). Do stocks outperform treasury bills? *Journal of Financial Economics*, 129(3):440–457.

Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the Business and Economics Section of the American Statistical Association*, pages 177–181.

Boyer, B., Mitton, T., and Vorkink, K. (2010). Expected idiosyncratic skewness. *The Review of Financial Studies*, 23(1):169–202.

Boyer, B. H. and Vorkink, K. (2014). Stock options as lotteries. *The Journal of Finance*, 69(4):1485–1527.

Broadie, M. and Kaya, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2):217–231.

Brunnermeier, M. K., Gollier, C., and Parker, J. A. (2007). Optimal beliefs, asset prices, and the preference for skewed returns. *American Economic Review*, 97(2):159–165.

Cheridito, P., Filipović, D., and Kimmel, R. L. (2007). Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83(1):123–170.

Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257.

Christoffersen, P., Heston, S., and Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, 55(12):1914–1932.

Conine Jr, T. E. and Tamarkin, M. J. (1981). On diversification given asymmetry in returns. *The Journal of Finance*, 36(5):1143–1155.

Conrad, J., Dittmar, R. F., and Ghysels, E. (2013). Ex ante skewness and expected stock returns. *The Journal of Finance*, 68(1):85–124.

Conrad, J., Kapadia, N., and Xing, Y. (2014). Death and jackpot: Why do individual investors hold overpriced stocks? *Journal of Financial Economics*, 113(3):455–475.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (2005). A theory of the term structure of interest rates. In *Theory of Valuation*, pages 129–164. World Scientific.

Duffie, D., Filipović, D., Schachermayer, W., et al. (2003). Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376.

Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, 58(3):1269–1300.

Fama, E. F. and French, K. R. (2018). Long-horizon returns. *The Review of Asset Pricing Studies*, 8(2):232–252.

Farago, A. and Hjalmarsson, E. (2019). Compound returns. In *Proceedings of Paris December 2019 Finance Meeting EUROFIDAI-ESSEC*.

Gallant, A. R. and Tauchen, G. (1998). Reprojecting partially observed systems with application to interest rate diffusions. *Journal of the American Statistical Association*, 93(441):10–24.

Hall, A. (2005). *Generalized Method of Moments*. Advanced texts in econometrics. Oxford University Press.

Hansen, L. and Scheinkman, J. (1995). Back to the future: Generating moment implications for continuous-time markov processes. *Econometrica*, 63(4):767–804.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3):1263–1295.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343.

Jiang, G. J. and Knight, J. L. (2002). Estimation of continuous-time processes via the empirical characteristic function. *Journal of Business & Economic Statistics*, 20(2):198–212.

Keller-Ressel, M. (2011). Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance*, 21(1):73–98.

Keller-Ressel, M., Mayerhofer, E., et al. (2015). Exponential moments of affine processes. *The Annals of Applied Probability*, 25(2):714–752.

Keller-Ressel, M. and Steiner, T. (2008). Yield curve shapes and the asymptotic short rate distribution in affine one-factor models. *Finance and Stochastics*, 12(2):149–172.

Kraus, A. and Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance*, 31(4):1085–1100.

Li, H., Wells, M. T., and Yu, C. L. (2008). A bayesian analysis of return dynamics with lévy jumps. *The Review of Financial Studies*, 21(5):2345–2378.

Li, Y. (2020). Nearly unbiased estimation of sample skewness. *Economics Letters*, page 109174.

Mitton, T. and Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies*, 20(4):1255–1288.

Neuberger, A. (2012). Realized skewness. *The Review of Financial Studies*, 25(11):3423–3455.

Neuberger, A. and Payne, R. (2020). The skewness of the stock market over long horizons. *The Review of Financial Studies*.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–08.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50.

Rollin, S. d. B., Ferreiro-Castilla, A., and Utzet, F. (2009). A new look at the heston characteristic function. *Stochastic Processes and Their Applications*.

Rubinstein, M. E. (1973). The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis*, pages 61–69.

Scott, R. C. and Horvath, P. A. (1980). On the direction of preference for moments of higher order than the variance. *The Journal of Finance*, 35(4):915–919.

Simkowitz, M. A. and Beedles, W. L. (1978). Diversification in a three-moment world. *Journal of Financial and Quantitative Analysis*, 13(5):927–941.

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102(1):111–141.

Stilger, P. S., Kostakis, A., and Poon, S.-H. (2017). What does risk-neutral skewness tell us about future stock returns? *Management Science*, 63(6):1814–1834.

Table 1: Parameter Estimates under Various Heston Specifications

This table reports some descriptive statistics of the estimated distributions of Heston parameters under various Heston specifications. *SD* and *MSE* are standard deviation and mean square error separately. Parameters μ, κ, α, ξ and ρ respectively indicate drift, mean-reversion, long-run variance, volatility-of-volatility and correlation between Brownian motions in a Heston model. All estimations are based on 10,000 replications of 10-year daily dollar returns generated by the corresponding Heston process, and all estimates are obtained from a simple two-step GMM estimator, where we match the theoretical central moments and cross-moments of daily returns with their sample counterparts.

Base Case						Small ξ					
Parameter	True	Mean	Median	SD	MSE	Parameter	True	Mean	Median	SD	MSE
μ	0.1	0.090	0.092	0.106	0.011	μ	0.1	0.090	0.090	0.104	0.011
κ	3	3.964	3.650	1.971	4.813	κ	3	4.824	4.132	3.487	15.484
α	0.09	0.072	0.071	0.008	0.000	α	0.09	0.074	0.074	0.003	0.000
ξ	0.3	0.252	0.243	0.082	0.009	ξ	0.1	0.095	0.066	0.086	0.007
ρ	-0.5	-0.473	-0.467	0.192	0.038	ρ	-0.5	-0.677	-0.886	0.365	0.165
Small κ						Large ξ					
Parameter	True	Mean	Median	SD	MSE	Parameter	True	Mean	Median	SD	MSE
μ	0.1	0.081	0.086	0.116	0.014	μ	0.1	0.087	0.091	0.108	0.012
κ	1	1.984	1.775	1.117	2.215	κ	3	4.082	3.862	1.656	3.915
α	0.09	0.068	0.064	0.021	0.001	α	0.09	0.068	0.067	0.012	0.001
ξ	0.3	0.285	0.270	0.097	0.010	ξ	0.5	0.424	0.415	0.099	0.016
ρ	-0.5	-0.411	-0.391	0.229	0.060	ρ	-0.5	-0.442	-0.434	0.168	0.032
Large κ						Small ρ					
Parameter	True	Mean	Median	SD	MSE	Parameter	True	Mean	Median	SD	MSE
μ	0.1	0.091	0.093	0.104	0.011	μ	0.1	0.091	0.095	0.106	0.011
κ	5	5.840	5.599	2.486	6.884	κ	3	4.124	3.867	1.661	4.021
α	0.09	0.073	0.073	0.005	0.000	α	0.09	0.072	0.071	0.008	0.000
ξ	0.3	0.236	0.226	0.088	0.012	ξ	0.3	0.269	0.264	0.061	0.005
ρ	-0.5	-0.512	-0.495	0.213	0.045	ρ	-0.9	-0.775	-0.788	0.145	0.037
Large α						Large ρ					
Parameter	True	Mean	Median	SD	MSE	Parameter	True	Mean	Median	SD	MSE
μ	0.1	0.063	0.063	0.175	0.032	μ	0.1	0.085	0.084	0.105	0.011
κ	3	4.196	3.582	2.745	8.963	κ	3	4.288	4.084	2.256	6.748
α	0.25	0.204	0.204	0.014	0.002	α	0.09	0.072	0.071	0.008	0.000
ξ	0.3	0.234	0.219	0.116	0.018	ξ	0.3	0.271	0.250	0.126	0.017
ρ	-0.5	-0.588	-0.569	0.273	0.082	ρ	0	-0.096	-0.019	0.146	0.030

Table 2: Skewness Estimates under Various Heston Specifications

This table displays the results of unconditional skewness estimation under various Heston specifications, with horizon ranging from one week to five years. Our base case parameters are: $\mu = 0.1$, $\kappa = 3$, $\alpha = 0.09$, $\xi = 0.3$, and $\rho = -0.5$, respectively indicating drift, mean-reversion, long-run variance, volatility-of-volatility and correlation between Brownian motions in a Heston model. For experiments left, we only change the value of the stated one. *True* unconditional skewness are obtained by plugging specifications into the closed-form skewness of the Heston model, while all estimations are based on 10,000 replications of 10-year daily dollar returns generated by the corresponding Heston process. *FF*, *FH* and *Conv.* respectively represent Fama and French's (2018) estimator, Farago and Hjalmarsson's (2019) estimator, and the sample skewness estimator. *%|Bias|* refers to the absolute relative bias in percentage, and *MSE* is the mean square error.

Base case						
Horizon	True		Ours	FF	FH	Conv.
1W	0.044	Mean	0.041	0.111	0.111	0.050
		%|Bias|	7%	155%	156%	14%
		MSE	0.001	0.006	0.006	0.012
1M	0.097	Mean	0.093	0.253	0.253	0.106
		%|Bias|	4%	161%	161%	10%
		MSE	0.004	0.025	0.025	0.042
1Y	0.594	Mean	0.575	0.944	0.946	0.438
		%|Bias|	3%	59%	59%	26%
		MSE	0.018	0.128	0.129	0.217
3Y	1.406	Mean	1.286	1.832	1.840	0.588
		%|Bias|	9%	30%	31%	58%
		MSE	0.047	0.219	0.207	0.906
5Y	2.120	Mean	1.890	2.672	2.691	0.567
		%|Bias|	11%	26%	27%	73%
		MSE	0.112	0.482	0.384	2.707

$\kappa = 1$						
Horizon	True		Ours	FF	FH	Conv.
1W	0.085	Mean	0.060	0.118	0.119	0.091
		%|Bias|	29%	40%	40%	8%
		MSE	0.003	0.004	0.003	0.024
1M	0.173	Mean	0.125	0.255	0.254	0.173
		%|Bias|	28%	47%	47%	0%
		MSE	0.013	0.009	0.009	0.090
1Y	0.608	Mean	0.532	0.938	0.939	0.391
		%|Bias|	13%	54%	54%	36%
		MSE	0.081	0.136	0.134	0.326
3Y	1.197	Mean	1.135	1.830	1.837	0.522
		%|Bias|	5%	53%	53%	56%
		MSE	0.172	0.568	0.559	0.710
5Y	1.753	Mean	1.666	2.688	2.714	0.562
		%|Bias|	5%	53%	55%	68%
		MSE	0.320	1.554	1.444	1.771

$\alpha = 0.25$						
Horizon	True		Ours	FF	FH	Conv.
1W	0.161	Mean	0.144	0.202	0.202	0.162
		%|Bias|	10%	25%	25%	0%
		MSE	0.001	0.003	0.002	0.008
1M	0.336	Mean	0.304	0.433	0.433	0.326
		%|Bias|	9%	29%	29%	3%
		MSE	0.004	0.011	0.010	0.035
1Y	1.410	Mean	1.294	1.743	1.748	0.916
		%|Bias|	8%	24%	24%	35%
		MSE	0.032	0.133	0.121	0.494
3Y	3.464	Mean	2.993	4.276	4.358	1.013
		%|Bias|	14%	23%	26%	71%
		MSE	0.346	2.141	0.910	6.320
5Y	6.453	Mean	5.211	7.782	8.718	0.951
		%|Bias|	19%	21%	35%	85%
		MSE	2.169	18.035	6.162	30.651

$\kappa = 5$						
Horizon	True		Ours	FF	FH	Conv.
1W	0.036	Mean	0.039	0.109	0.109	0.040
		%|Bias|	6%	201%	200%	9%
		MSE	0.001	0.007	0.006	0.010
1M	0.090	Mean	0.095	0.252	0.252	0.095
		%|Bias|	5%	179%	179%	5%
		MSE	0.003	0.027	0.026	0.035
1Y	0.677	Mean	0.638	0.944	0.945	0.478
		%|Bias|	6%	39%	40%	29%
		MSE	0.011	0.075	0.073	0.225
3Y	1.538	Mean	1.384	1.833	1.838	0.600
		%|Bias|	10%	19%	20%	61%
		MSE	0.039	0.115	0.098	1.116
5Y	2.297	Mean	2.016	2.673	2.686	0.557
		%|Bias|	12%	16%	17%	76%
		MSE	0.106	0.286	0.174	3.319

		$\xi = 0.1$				
Horizon	True		Ours	FF	FH	Conv.
1W	0.095	Mean	0.085	0.121	0.121	0.095
		%|Bias|	10%	28%	28%	0%
		MSE	0.001	0.002	0.001	0.007
1M	0.197	Mean	0.181	0.257	0.258	0.193
		%|Bias|	9%	30%	31%	2%
		MSE	0.002	0.004	0.004	0.027
1Y	0.810	Mean	0.750	0.948	0.949	0.555
		%|Bias|	7%	17%	17%	32%
		MSE	0.009	0.022	0.020	0.261
3Y	1.668	Mean	1.485	1.838	1.843	0.617
		%|Bias|	11%	10%	10%	63%
		MSE	0.047	0.051	0.034	1.339
5Y	2.459	Mean	2.131	2.679	2.691	0.568
		%|Bias|	13%	9%	9%	77%
		MSE	0.138	0.188	0.063	3.862

		$\rho = -0.9$				
Horizon	True		Ours	FF	FH	Conv.
1W	-0.040	Mean	-0.030	0.094	0.095	-0.028
		%|Bias|	24%	336%	338%	29%
		MSE	0.001	0.020	0.019	0.011
1M	-0.066	Mean	-0.041	0.244	0.245	-0.041
		%|Bias|	38%	471%	472%	38%
		MSE	0.005	0.097	0.097	0.035
1Y	0.279	Mean	0.359	0.941	0.943	0.271
		%|Bias|	29%	237%	238%	3%
		MSE	0.026	0.444	0.444	0.154
3Y	1.063	Mean	1.077	1.832	1.838	0.550
		%|Bias|	1%	72%	73%	48%
		MSE	0.032	0.629	0.619	0.501
5Y	1.704	Mean	1.653	2.665	2.689	0.572
		%|Bias|	3%	56%	58%	66%
		MSE	0.054	1.095	1.027	1.611

		$\xi = 0.5$				
Horizon	True		Ours	FF	FH	Conv.
1W	0.011	Mean	0.003	0.104	0.105	0.019
		%|Bias|	70%	858%	861%	78%
		MSE	0.003	0.011	0.011	0.026
1M	0.032	Mean	0.022	0.249	0.249	0.042
		%|Bias|	32%	685%	686%	32%
		MSE	0.011	0.049	0.048	0.085
1Y	0.445	Mean	0.442	0.942	0.943	0.324
		%|Bias|	1%	112%	112%	27%
		MSE	0.040	0.258	0.256	0.230
3Y	1.225	Mean	1.129	1.833	1.838	0.542
		%|Bias|	8%	50%	50%	56%
		MSE	0.068	0.440	0.425	0.717
5Y	1.889	Mean	1.690	2.680	2.694	0.569
		%|Bias|	11%	42%	43%	70%
		MSE	0.138	0.907	0.806	2.081

		$\rho = 0$				
Horizon	True		Ours	FF	FH	Conv.
1W	0.148	Mean	0.113	0.131	0.131	0.145
		%|Bias|	23%	11%	11%	2%
		MSE	0.002	0.002	0.001	0.014
1M	0.302	Mean	0.232	0.263	0.263	0.284
		%|Bias|	23%	13%	13%	6%
		MSE	0.007	0.003	0.002	0.051
1Y	1.037	Mean	0.831	0.909	0.950	0.623
		%|Bias|	20%	12%	8%	40%
		MSE	0.052	0.013	0.011	0.406
3Y	1.946	Mean	1.574	1.838	1.845	0.619
		%|Bias|	19%	6%	5%	68%
		MSE	0.162	0.050	0.028	1.998
5Y	2.819	Mean	2.236	2.673	2.698	0.567
		%|Bias|	21%	5%	4%	80%
		MSE	0.393	0.206	0.072	5.369

Table 3: Skewness Forecasts under Various Heston Specifications

This table summarizes the results of conditional skewness forecast under various Heston specifications, with horizon ranging from one year to ten years. Our base case parameters are: $\mu = 0.1$, $\kappa = 3$, $\alpha = 0.09$, $\xi = 0.3$, and $\rho = -0.5$, respectively representing drift, mean-reversion, long-run variance, volatility-of-volatility and correlation between Brownian motions in a Heston model. For experiments left, we only change the value of the stated one. All forecasts are based on 10,000 replications of the last two years' daily dollar returns of the 10-year sample generated by the corresponding Heston process. $\%|Bias|$ indicates the absolute relative bias in percentage, and MSE is the mean square error.

Base Case					$\xi = 0.1$				
Horizon	True	Mean	%|Bias|	MSE	Horizon	True	Mean	%|Bias|	MSE
1Y	0.582	0.590	1%	0.018	1Y	0.809	0.759	6%	0.008
3Y	1.402	1.301	7%	0.042	3Y	1.668	1.497	10%	0.042
5Y	2.118	1.905	10%	0.101	5Y	2.459	2.146	13%	0.126
10Y	4.155	3.516	15%	0.652	10Y	4.874	3.961	19%	0.980
$\kappa = 1$					$\xi = 0.5$				
Horizon	True	Mean	%|Bias|	MSE	Horizon	True	Mean	%|Bias|	MSE
1Y	0.414	0.520	26%	0.090	1Y	0.410	0.454	11%	0.043
3Y	1.100	1.153	5%	0.163	3Y	1.216	1.146	6%	0.062
5Y	1.689	1.693	0%	0.292	5Y	1.884	1.707	9%	0.126
10Y	3.222	3.117	3%	1.435	10Y	3.686	3.138	15%	0.686
$\kappa = 5$					$\rho = -0.9$				
Horizon	True	Mean	%|Bias|	MSE	Horizon	True	Mean	%|Bias|	MSE
1Y	0.675	0.648	4%	0.010	1Y	0.266	0.369	39%	0.032
3Y	1.537	1.392	9%	0.036	3Y	1.059	1.086	3%	0.032
5Y	2.297	2.024	12%	0.100	5Y	1.702	1.662	2%	0.051
10Y	4.538	3.749	17%	0.742	10Y	3.372	3.116	8%	0.248
$\alpha = 0.25$					$\rho = 0$				
Horizon	True	Mean	%|Bias|	MSE	Horizon	True	Mean	%|Bias|	MSE
1Y	1.403	1.316	6%	0.028	1Y	1.024	0.847	17%	0.041
3Y	3.461	3.027	13%	0.315	3Y	1.942	1.592	18%	0.145
5Y	6.450	5.264	18%	2.034	5Y	2.816	2.256	20%	0.366
10Y	27.486	18.293	33%	109.843	10Y	5.651	4.153	27%	2.568

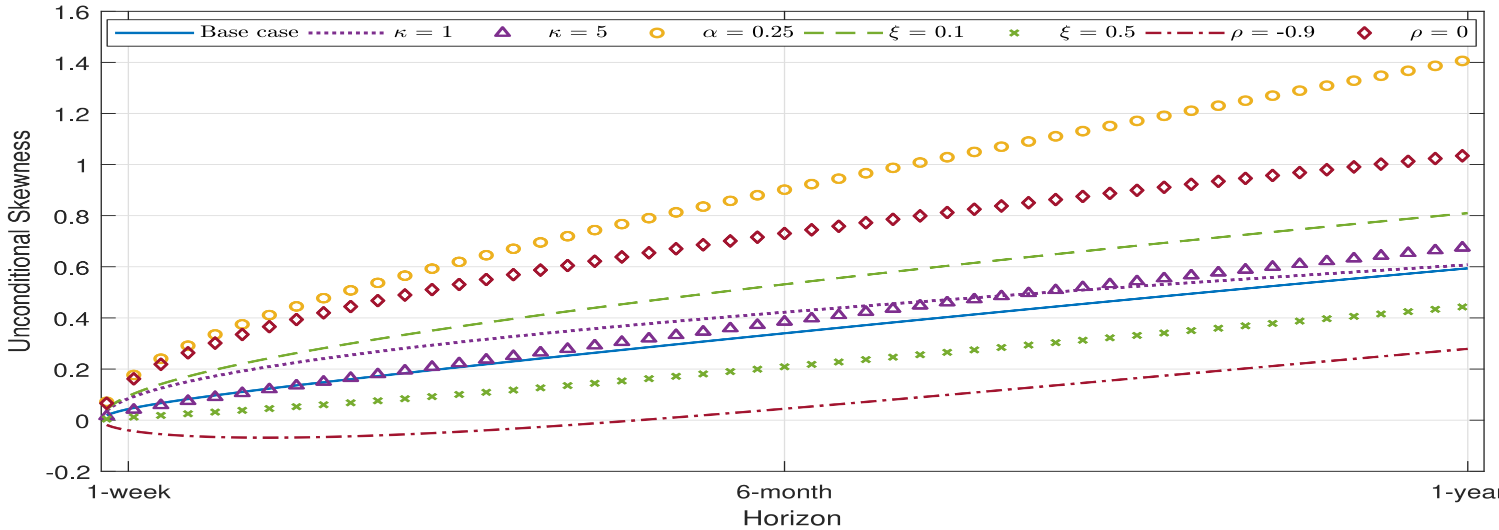
Table 4: Skewness Estimates with Double-Heston Data

This table shows the true unconditional skewness of a double-Heston process, with horizon ranging from one week to five years. Alongside, there are statistics of estimates produced by each estimator. The double-Heston parameter vector $\{\mu, \kappa_1, \alpha_1, \xi_1, \rho_1, \kappa_2, \alpha_2, \xi_2, \rho_2\}$ is set to $\{0.1, 1, 0.01, 0.1, -0.9, 5, 0.09, 0.5, -0.6\}$, where μ, κ_i, α_i, ξ_i and ρ_i respectively indicate drift, mean-reversion, long-run variance, volatility-of-volatility and correlation between corresponding Brownian motions of the price and the volatility processes. *True* unconditional skewness are obtained by plugging specifications into the closed-form skewness of the double-Heston model, while all estimations are based on 10,000 replications of 10-year daily dollar returns generated by this double-Heston process. *Ours*, *FF*, *FH* and *Conv.* respectively represent our parametric estimator derived from the Heston assumption, Fama and French's (2018) estimator, Farago and Hjalmarsson's (2019) estimator, and the sample skewness estimator. $\%|Bias|$ refers to the absolute relative bias in percentage, and *MSE* is the mean square error.

Horizon	True		Ours	FF	FH	Conv.
1W	-0.018	Mean	-0.009	0.104	0.104	-0.011
		%|Bias|	48%	679%	680%	40%
		MSE	0.002	0.017	0.016	0.014
1M	-0.012	Mean	0.008	0.260	0.260	0.001
		%|Bias|	161%	2214%	2215%	109%
		MSE	0.006	0.075	0.075	0.045
1Y	0.527	Mean	0.537	0.998	1.000	0.401
		%|Bias|	2%	89%	90%	24%
		MSE	0.017	0.228	0.227	0.194
3Y	1.406	Mean	1.321	1.967	1.974	0.603
		%|Bias|	6%	40%	40%	57%
		MSE	0.033	0.361	0.343	0.888
5Y	2.167	Mean	1.975	2.911	2.932	0.606
		%|Bias|	9%	34%	35%	72%
		MSE	0.084	0.808	0.654	2.750

Figure 1: Relation Between Unconditional Skewness and Return Horizon in a Heston World

This figure plots the true unconditional skewness obtained from eight Heston parameter vectors against horizon up to five years. Our base case parameters are: $\mu = 0.1$, $\kappa = 3$, $\alpha = 0.09$, $\xi = 0.3$, and $\rho = -0.5$, respectively indicating drift, mean-reversion, long-run variance, volatility-of-volatility and correlation between Brownian motions in a Heston model. For experiments left, we only change the value of the stated one.



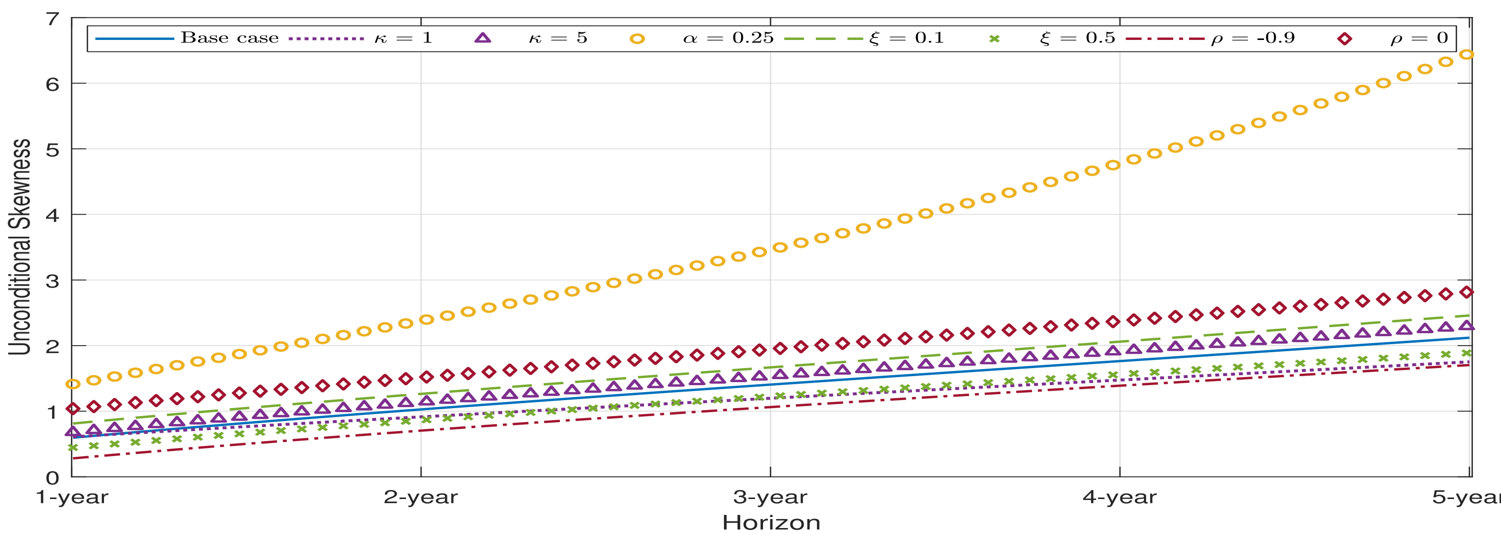


Figure 2: Unconditional Skewness Estimates of a Heston Process

This figure depicts the relation between true unconditional skewness and return horizon of a Heston (1993) process, alongside of which are curves of mean estimates produced by different estimators. The parameter vector $\{\mu, \kappa, \alpha, \xi, \rho\}$ is set to $\{0.1, 3, 0.09, 0.3, -0.9\}$, where μ, κ, α, ξ and ρ respectively represent drift, mean reversion, long-run variance, volatility-of-volatility and correlation between Brownian motions. All estimations are based on 10,000 replications of 10-year daily dollar returns generated by this Heston process.

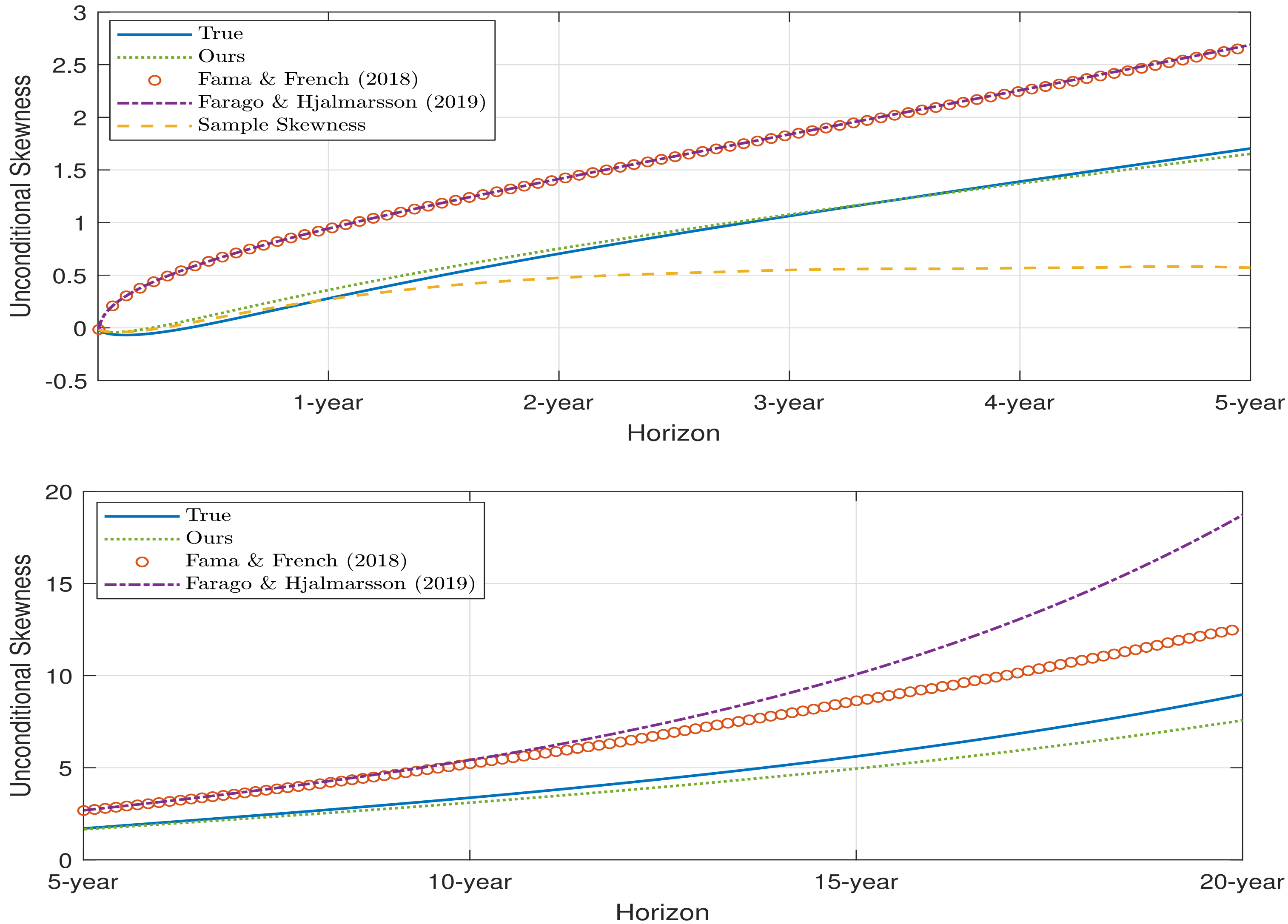
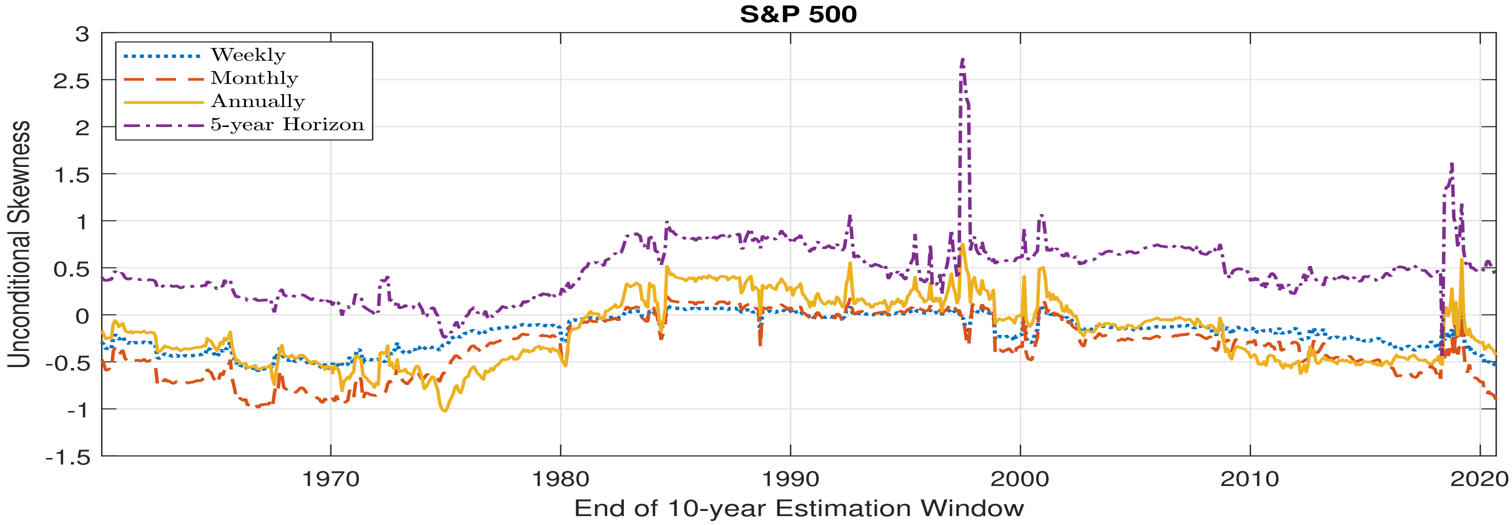
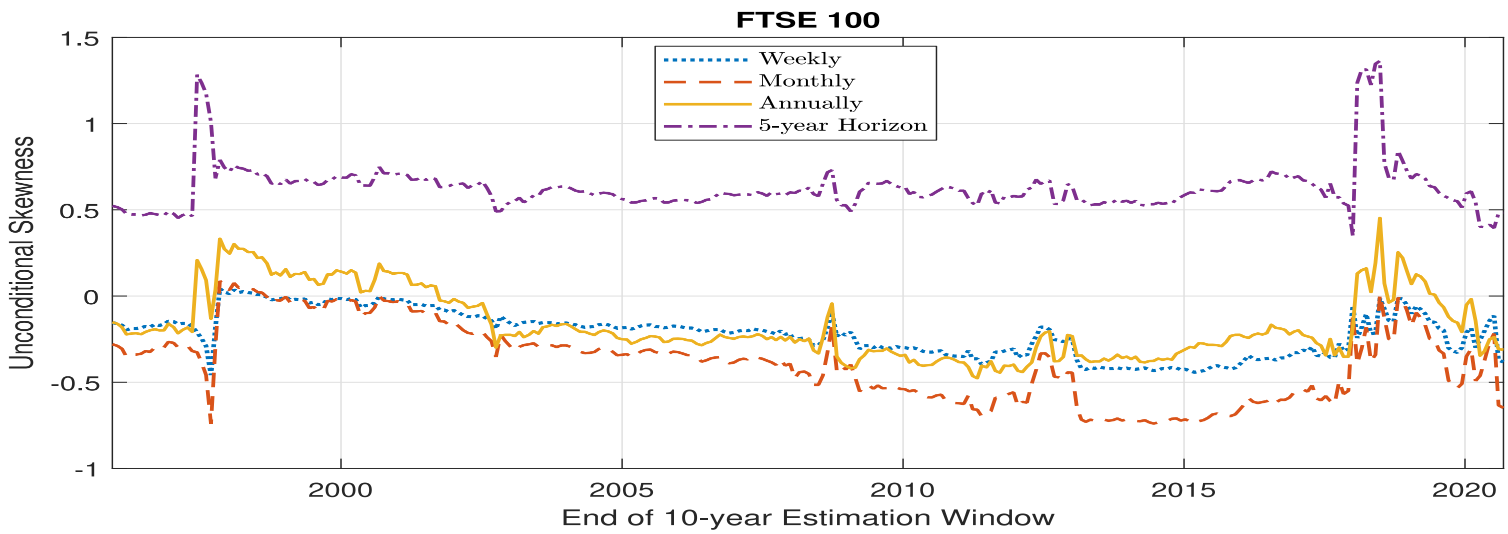


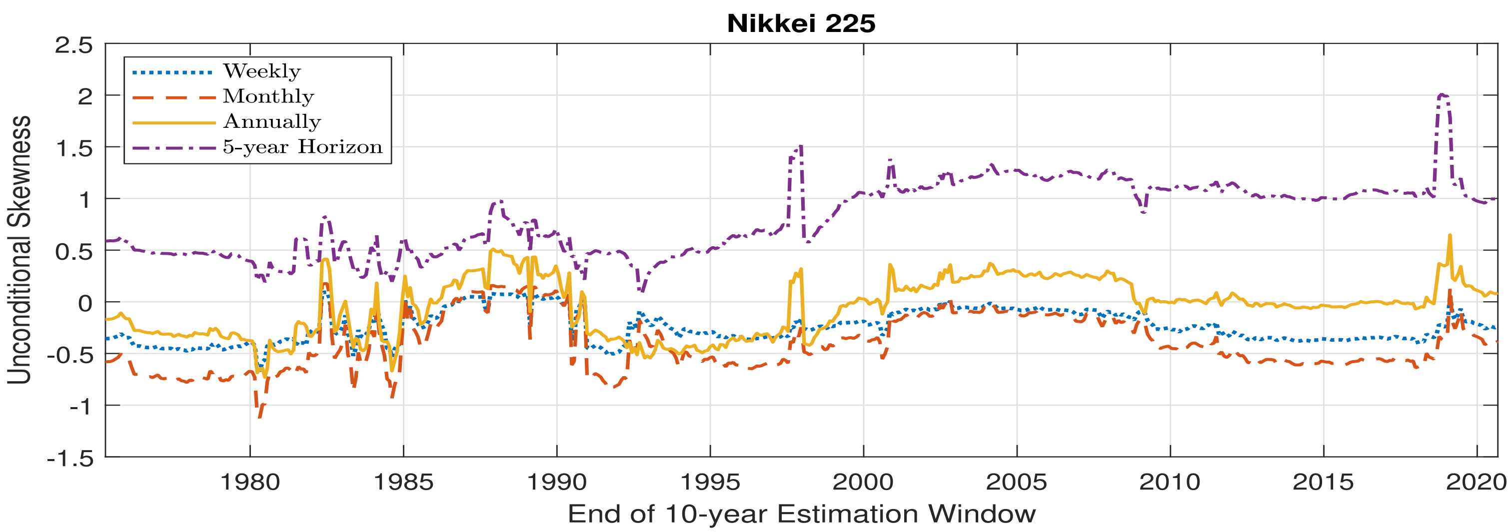
Figure 3: The Estimated Trends of Unconditional Skewness for Representative Stock Indexes

This figure plots our unconditional skewness estimates at various horizons for S&P500, FTSE100 and Nikkei225. We download their daily prices for the periods 1950/01/03 – 2020/09/30, 1986/01/02 – 2020/09/30, and 1965/01/05 – 2020/09/30 respectively. We set the length of estimation window to ten years, and we roll one month over every time we shift the window.





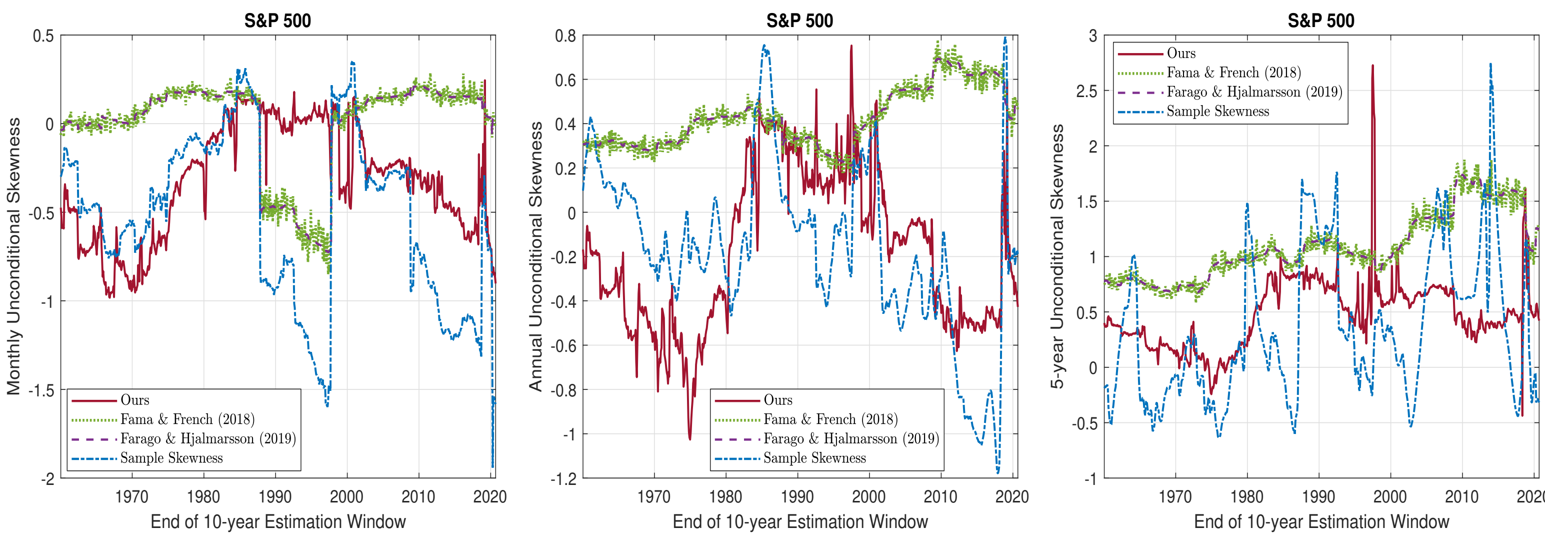
FTSE 100
Weekly
Monthly
Annually
5-year Horizon
Unconditional Skewness
1.5
1
0.5
0
-0.5
-1
2000
2005
2010
2015
2020
End of 10-year Estimation Window

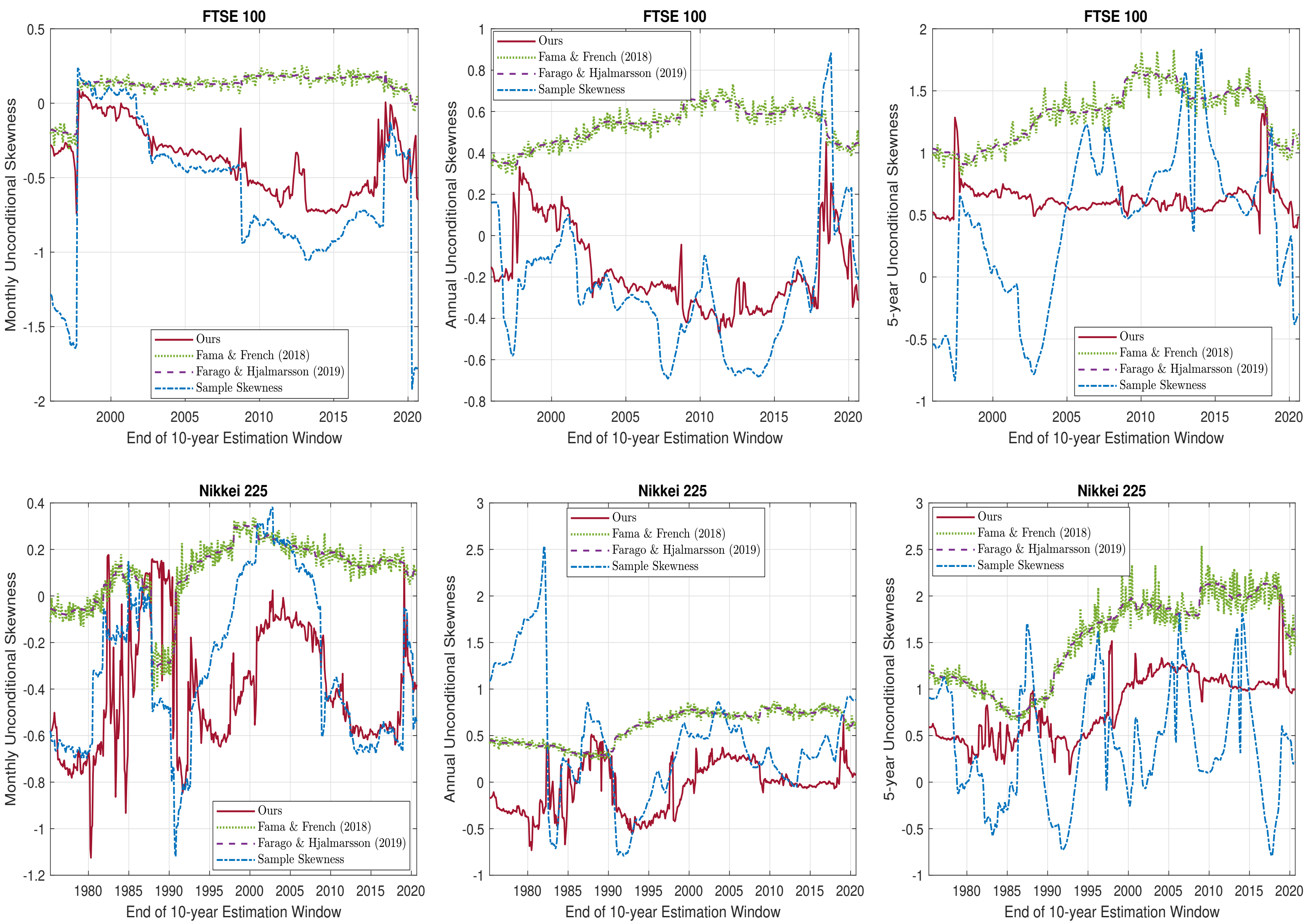


Nikkei 225
Weekly
Monthly
Annually
5-year Horizon
Unconditional Skewness
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
1980
1985
1990
1995
2000
2005
2010
2015
2020
End of 10-year Estimation Window

Figure 4: Unconditional Skewness Estimates by Different Estimators for Representative Stock Indexes

This figure compares the monthly, annual and five-year unconditional skewness estimates produced by different estimators for S&P500, FTSE100 and Nikkei225. We download their daily prices for the periods 1950/01/03 – 2020/09/30, 1986/01/02 – 2020/09/30, and 1965/01/05 – 2020/09/30 respectively. We set the length of estimation window to ten years, and we roll one month over every time we shift the window.




FTSE 100
Monthly Unconditional Skewness
End of 10-year Estimation Window
Ours
Fama & French (2018)
Farago & Hjalmarsson (2019)
Sample Skewness
FTSE 100
Annual Unconditional Skewness
End of 10-year Estimation Window
Ours
Fama & French (2018)
Farago & Hjalmarsson (2019)
Sample Skewness
FTSE 100
5-year Unconditional Skewness
End of 10-year Estimation Window
Ours
Fama & French (2018)
Farago & Hjalmarsson (2019)
Sample Skewness
Nikkei 225
Monthly Unconditional Skewness
End of 10-year Estimation Window
Ours
Fama & French (2018)
Farago & Hjalmarsson (2019)
Sample Skewness
Nikkei 225
Annual Unconditional Skewness
End of 10-year Estimation Window
Ours
Fama & French (2018)
Farago & Hjalmarsson (2019)
Sample Skewness
Nikkei 225
5-year Unconditional Skewness
End of 10-year Estimation Window
Ours
Fama & French (2018)
Farago & Hjalmarsson (2019)
Sample Skewness

Appendix A Proofs

A.1 Expressions for Centred MGFs and Cross-MGFs

To show that centered MGFs and cross-MGFs are simply linear combinations of non-centered MGFs and cross-MGFs, we use the moments involved in our GMM estimator as examples (see Equation (3.2)) and derive their expressions. Recall that $M(k, \Delta) = \mathrm{E}[R_{i,\Delta}^k]$ and $\tilde{M}(k, \Delta) \equiv \mathrm{E}[\tilde{R}_{i,\Delta}^k]$. Let us start from $\tilde{M}(2, \Delta)$.

$$\begin{aligned}
\tilde{M}(2, \Delta) &\equiv \mathrm{E}\left[\tilde{R}_{i,\Delta}^2\right] \\
&= \mathrm{E}\left[(R_{i,\Delta} - M(1, \Delta))^2\right] \\
&= \mathrm{E}\left[R_{i,\Delta}^2 - 2R_{i,\Delta}M(1, \Delta) + M^2(1, \Delta)\right] \\
&= M(2, \Delta) - 2M^2(1, \Delta) + M^2(1, \Delta) \\
&= M(2, \Delta) - M^2(1, \Delta)
\end{aligned} \tag{A.1}$$

Next, we derive $\tilde{M}(3, \Delta)$.

$$\begin{aligned}
\tilde{M}(3, \Delta) &\equiv \mathrm{E}\left[\tilde{R}_{i,\Delta}^3\right] \\
&= \mathrm{E}\left[(R_{i,\Delta} - M(1, \Delta))^3\right] \\
&= \mathrm{E}\left[R_{i,\Delta}^3 - 3R_{i,\Delta}^2M(1, \Delta) + 3R_{i,\Delta}M^2(1, \Delta) - M^3(1, \Delta)\right] \\
&= M(3, \Delta) - 3M(1, \Delta)M(2, \Delta) + 3M^3(1, \Delta) - M^3(1, \Delta) \\
&= M(3, \Delta) - 3M(1, \Delta)M(2, \Delta) + 2M^3(1, \Delta)
\end{aligned} \tag{A.2}$$

Now, we move on to expanding $\tilde{M}(4, \Delta)$.

$$
\begin{aligned}
\tilde{M}(4,\Delta) &\equiv \mathrm{E}\left[\tilde{R}_{i,\Delta}^4\right] \\
&= \mathrm{E}\left[\left(R_{i,\Delta} - M(1,\Delta)\right)^4\right] \\
&= \mathrm{E}\left[R_{i,\Delta}^4 - 4R_{i,\Delta}^3 M(1,\Delta) + 6R_{i,\Delta}^2 M^2(1,\Delta) - 4R_{i,\Delta} M^3(1,\Delta) + M^4(1,\Delta)\right] \\
&= M(4,\Delta) - 4M(1,\Delta)M(3,\Delta) + 6M^2(1,\Delta)M(2,\Delta) - 4M^4(1,\Delta) + M^4(1,\Delta) \\
&= M(4,\Delta) - 4M(1,\Delta)M(3,\Delta) + 6M^2(1,\Delta)M(2,\Delta) - 3M^4(1,\Delta)
\end{aligned}
\tag{A.3}
$$

For centered cross-MGFs $\tilde{C}(m,n,\Delta,j) \equiv \mathrm{Cov}[\tilde{R}_{i,\Delta}^m, \tilde{R}_{i+j,\Delta}^n]$, we assume a weak-form efficient market, where $\mathrm{Cov}[R_{i,\Delta}^m, R_{i+j,\Delta}] = 0$. In other words, $\mathrm{E}[R_{i,\Delta}^m R_{i+j,\Delta}] = \mathrm{E}[R_{i,\Delta}^m]\,\mathrm{E}[R_{i+j,\Delta}]$. Then,

$$
\begin{aligned}
\tilde{C}(1,2,\Delta,j) &\equiv \mathrm{Cov}\left[\tilde{R}_{i,\Delta}, \tilde{R}_{i+j,\Delta}^2\right] \\
&= \mathrm{E}\left[\tilde{R}_{i,\Delta}\tilde{R}_{i+j,\Delta}^2\right] - \mathrm{E}\left[\tilde{R}_{i,\Delta}\right]\mathrm{E}\left[\tilde{R}_{i+j,\Delta}^2\right] \\
&= \mathrm{E}\left[\left(R_{i,\Delta} - M(1,\Delta)\right)\left(R_{i+j,\Delta} - M(1,\Delta)\right)^2\right] - \mathrm{E}\left[R_{i,\Delta} - M(1,\Delta)\right]\mathrm{E}\left[\left(R_{i+j,\Delta} - M(1,\Delta)\right)^2\right] \\
&= \mathrm{E}\left[R_{i,\Delta}R_{i+j,\Delta}^2 - 2R_{i,\Delta}R_{i+j,\Delta}M(1,\Delta) + R_{i,\Delta}M^2(1,\Delta)\right. \\
&\quad \left. - R_{i+j,\Delta}^2 M(1,\Delta) + 2R_{i+j,\Delta}M^2(1,\Delta) - M^3(1,\Delta)\right] \\
&= C(1,2,\Delta,j) - 2M^3(1,\Delta) + M^3(1,\Delta) - M(1,\Delta)M(2,\Delta) + 2M^3(1,\Delta) - M^3(1,\Delta) \\
&= C(1,2,\Delta,j) - M(1,\Delta)M(2,\Delta)
\end{aligned}
\tag{A.4}
$$

Finally, we expand $\tilde{C}(2,2,\Delta,j)$.

$$
\begin{aligned}
\tilde{C}(2,2,\Delta,j) &\equiv \operatorname{Cov}\left[\tilde{R}^2_{i,\Delta}, \tilde{R}^2_{i+j,\Delta}\right] \\
&= \mathrm{E}\left[\tilde{R}^2_{i,\Delta}\tilde{R}^2_{i+j,\Delta}\right] - \mathrm{E}\left[\tilde{R}^2_{i,\Delta}\right]\mathrm{E}\left[\tilde{R}^2_{i+j,\Delta}\right] \\
&= \mathrm{E}\left[(R_{i,\Delta} - M(1,\Delta))^2 (R_{i+j,\Delta} - M(1,\Delta))^2\right] - \mathrm{E}\left[(R_{i,\Delta} - M(1,\Delta))^2\right]\mathrm{E}\left[(R_{i+j,\Delta} - M(1,\Delta))^2\right] \\
&= \mathrm{E}\left[R^2_{i,\Delta}R^2_{i+j,\Delta} - 2R^2_{i,\Delta}R_{i+j,\Delta}M(1,\Delta) + R^2_{i,\Delta}M^2(1,\Delta) - 2R_{i,\Delta}R^2_{i+j,\Delta}M(1,\Delta)\right. \\
&\quad \left. + 4R_{i,\Delta}R_{i+j,\Delta}M^2(1,\Delta) - 2R_{i,\Delta}M^3(1,\Delta) + R^2_{i+j,\Delta}M^2(1,\Delta) - 2R_{i+j,\Delta}M^3(1,\Delta) + M^4(1,\Delta)\right] \\
&\quad - \mathrm{E}\left[R^2_{i,\Delta} - 2R_{i,\Delta}M(1,\Delta) + M^2(1,\Delta)\right]\mathrm{E}\left[R^2_{i+j,\Delta} - 2R_{i+j,\Delta}M(1,\Delta) + M^2(1,\Delta)\right] \\
&= C(2,2,\Delta,j) - 2M^2(1,\Delta)M(2,\Delta) + M^2(1,\Delta)M(2,\Delta) - 2M(1,\Delta)C(1,2,\Delta,j) \\
&\quad + 4M^4(1,\Delta) - 2M^4(1,\Delta) + M^2(1,\Delta)M(2,\Delta) - 2M^4(1,\Delta) + M^4(1,\Delta) \\
&\quad - \left(M(2,\Delta) - 2M^2(1,\Delta) + M^2(1,\Delta)\right)^2 \\
&= C(2,2,\Delta,j) - 2M(1,\Delta)C(1,2,\Delta,j) - M^2(2,\Delta) + 2M^2(1,\Delta)M(2,\Delta)
\end{aligned}
\tag{A.5}
$$

A.2 Derivation for the Non-Centered Cross-MGF

Let us denote the non-centered cross-MGF as $C(m,n,\Delta,j) = \mathrm{E}\left[R^m_{i,\Delta}R^n_{i+j,\Delta}\right]$. The following formula holds whenever the moments exist.

$$
\begin{aligned}
\mathrm{E}\left[R^m_{i,\Delta}R^n_{i+j,\Delta}\right] &= \mathrm{E}\left[R^m_{i,\Delta}\,\mathrm{E}\left[R^n_{i+j,\Delta}|\mathcal{F}_{i+j}\right]\right] \\
&= \mathrm{E}\left[R^m_{i,\Delta}M_{i+j}(n,0,\Delta)\right] \\
&= \mathrm{E}\left[R^m_{i,\Delta}e^{\mu\Delta n + \phi(n,0,\Delta) + \langle\psi(n,0,\Delta),V_{i+j}\rangle}\right] \\
&= e^{\mu\Delta n + \phi(n,0,\Delta)}\,\mathrm{E}\left[R^m_{i,\Delta}\,\mathrm{E}\left[e^{\langle\psi(n,0,\Delta),V_{i+j}\rangle}|\mathcal{F}_i\right]\right] \\
&= e^{\mu\Delta n + \phi(n,0,\Delta)}\,\mathrm{E}\left[R^m_{i,\Delta}M_i\left(0,\psi(n,0,\Delta),j\Delta\right)\right] \\
&= e^{\mu\Delta n + \phi(n,0,\Delta)}\,\mathrm{E}\left[R^m_{i,\Delta}e^{\phi(0,\psi(n,0,\Delta),j\Delta) + \langle\psi(0,\psi(n,0,\Delta),j\Delta),V_i\rangle}\right] \\
&= e^{\mu\Delta n + \phi(n,0,\Delta) + \phi(0,\psi(n,0,\Delta),j\Delta)}\,\mathrm{E}\left[e^{mr_{i,\Delta} + \langle\psi(0,\psi(n,0,\Delta),j\Delta),V_i\rangle}\right] \\
&= e^{\mu\Delta n + \phi(n,0,\Delta) + \phi(0,\psi(n,0,\Delta),j)}M(m,\psi(0,\psi(n,0,\Delta),j\Delta),\Delta)
\end{aligned}
\tag{A.6}
$$

A.3 Parameter Constraints Required to Ensure $M_t(6, \Delta; V_t) < \infty$

According to Rollin et al. (2009) Proposition 3.1.1, $M_t(k, \Delta; V_t) < \infty$ holds for every $\rho \in [-1, 0]$ on the interval $[k_-, k_+]$ with:

$$k_\pm = \frac{1 - 2\rho\frac{\kappa}{\xi} \pm \sqrt{4(\frac{\kappa}{\xi})^2 - 4\rho\frac{\kappa}{\xi} + 1}}{2(1-\rho^2)}. \tag{A.7}$$

If we assume $k_+ \geq 6$ and let $x = \kappa/\xi > 0$, this is equivalent of determining the range of x and ρ such that:

$$\frac{1 - 2\rho x + \sqrt{4x^2 - 4\rho x + 1}}{2(1-\rho^2)} \geq 6, \tag{A.8}$$

which is equivalent to:

$$\sqrt{4x^2 - 4\rho x + 1} \geq 11 - 12\rho^2 + 2\rho x. \tag{A.9}$$

We now divide into two scenarios.

Scenario 1. If $11 - 12\rho^2 + 2\rho x \leq 0$, then Equation (A.9) holds since the left hand side is always non-negative. This implies that $x \geq 6\rho - \frac{11}{2\rho}$ is a sufficient condition for Equation (A.9) to hold. We therefore see that if $6\rho - \frac{11}{2\rho} \leq 0$, which implies that $\rho \leq -\sqrt{11/12}$, then Equation (A.9) holds for all $x > 0$.

Scenario 2. We now assume $11 - 12\rho^2 + 2\rho x \geq 0$, or $x \leq 6\rho - \frac{11}{2\rho}$. Squaring both sides of Equation (A.9) and rearranging yields:

$$(1-\rho^2)x^2 - 12\rho(1-\rho^2)x - 36\rho^4 + 66\rho^2 - 30 \geq 0. \tag{A.10}$$

This is quadratic in x with at most one positive real root x_+. One therefore sees that Equation (A.9) holds under Scenario 2 when $6\rho - \frac{11}{2\rho} \geq x \geq x_+$. Solving the above equation w.r.t. x yields:

$$x_\pm = 6\rho \pm \sqrt{30}. \tag{A.11}$$

We now discuss two additional conditions: (1) $x_+ \leq 6\rho - \frac{11}{2\rho}$. If this does not hold, then scenario 2 is trivial as it is subsumed into Scenario 1. (2) $x_+ \geq 0$. If this does not hold, then effectively Equation A.9 holds for all $x > 0$. Condition (1) is equivalent to the following:

$$\rho \geq -\frac{11\sqrt{30}}{60} \approx -1.004, \tag{A.12}$$

which indicates that condition (1) always holds. Condition (2) is equivalent to the condition that $\rho \geq -\sqrt{30}/6 \approx -0.913$. We therefore sees that if $\rho \leq -\sqrt{30}/6$, then Equation (A.9) holds for any $x > 0$. As this includes scenario 1, scenario 1 is obsolete.