Three Essays on the Skewness of Discrete Stock Returns

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities

2024

Jiayu Jin

Alliance Manchester Business School
Accounting and Finance Division

Contents

Word count: 34,543

List of Figures

List of Tables

Abstract

In the first chapter, we develop an estimator of the physical skewness of an asset's discrete return over any time horizon based on the assumption that the asset's price follows a stochastic process from the affine stochastic volatility (ASV) model class. Conceptually, our estimator improves upon others by (i) focusing on discrete returns; (ii) allowing us to capture compounding and leverage effects; yielding (iii) horizon-consistent (iv) unconditional and conditional (i.e., in-sample and out-of-sample) estimates; and (v) not requiring ad-hoc conditioning variables. Our simulation exercise shows that our estimator is highly precise even when the true data-generating process partially deviates from that assumed by the estimator, and that it comfortably beats others advocated in recent studies. Using options data, we further show that our estimator best captures time-series variations in the risk-neutral conditional skewness of the S&P 500 index.

In the second chapter, we apply our newly developed skewness estimator to U.S. single-stock data. We first show that our estimates easily beat those of other estimators from the literature. We next identify those stocks more likely to exhibit a U-shaped skewness-return horizon relation or an explosive skewness. Moreover, we show that while historical and forward-looking skewness align over long return horizons, historical skewness is a weak predictor of forward-looking skewness over short horizons, with the spread conditioned by various firm fundamentals and the economic state. We next reveal how our estimates relate to firm fundamentals and firm-fundamental-based skewness estimators. We finally estimate the term structure of forward-looking skewness premiums orthogonal to firm characteristics.

In the third chapter, we empirically show that retail investors' preference for skewed assets meaningfully affects the skewness premiums of single stocks. Measuring the skewness premium as the difference between forward-looking skewness obtained from our newly developed skewness estimator and risk-neutral skewness obtained from options data, we report that the skewness premium of the average single stock significantly drops from the start of the COVID-19 pandemic. Exploiting the profound rise in retail investing over the same period, we establish that these drops in skewness premiums occur only for those stocks which observed the largest increases in their retail holdings.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I am immensely grateful to my brilliant PhD supervisors, Kevin Aretz and Yifan Li, for their endless support and invaluable guidance, both academically and mentally, throughout my doctoral journey. Doing a PhD is perhaps the most challenging decision I have ever made, but lucky enough, I was accepted by two incredibly responsible and thoughtful supervisors, which felt like winning the lottery. Without their help, I would not have achieved what I have accomplished.

I am also thankful to many colleagues, including but not limited to Hening Liu, Maria Marchica, Chi-Yang Tsou, Emiel Jerphanion, Olga Kolokolova, and Alex Taylor, for their valuable academic and personal advice, as well as encouragement, at different stages of my PhD. My unforgettable PhD experience is, unsurprisingly, also attributable to my supportive PhD friends, particularly Shifan Yu, Kevin Schneider, Mengqian Chen, Mo Hao, Tianzong Wang, Nan Hu, and many others.

I want to thank each and every member of my warm family for their ongoing concern and strong emotional support, particularly my parents, who have been my firmest support in every aspect.

Finally, many thanks to Alex Kostakis and Olivier Scaillet for being the internal and external examiners for this thesis.

Chapter 1

Estimating and Forecasting Skewness Using Affine Stochastic Volatility Models

1.1 Introduction

While a huge theoretical literature dating back to the 1960s looks into the origins and implications of asset skewness, we still massively struggle to empirically measure and predict that skewness with reasonable precision, especially over long return horizons.[1] Although the recent estimators of, for example, Fama and French (2018), Neuberger and Payne (2021), and Farago and Hjalmarsson (2023) have made significant progress along the precision dimension, they suffer from several serious shortcomings. Most crucially, they either (i) do not focus on discrete ("investable") returns; (ii) fail to adequately capture both compounding and leverage effects; (iii) do not yield horizon-consistent estimates; (iv) only measure unconditional (i.e., in-sample) but not conditional (i.e., out-of-sample) skewness; and/or (v) rely on ad-hoc conditioning variables.

[1] To aid with intuition, if returns were independently and identically (i.i.d) normal, the sample skewness would require about 600 observations (50 years of data) to estimate the true skewness with a standard error of 0.10.

In our paper, we propose a flexible, precise, and computationally fast parametric skewness estimator addressing the shortcomings of existing estimators. Spurred by an emerging literature in economics extracting forward-looking volatility estimates from stochastic volatility models (see Fernández-Villaverde et al. (2011), Chan (2017), Shin and Zhong (2020), etc.), the key assumption of our estimator is that we can use a stochastic process from the affine stochastic volatility (ASV) model class to describe the evolution of an asset's price over time. Under that assumption, the conditional and unconditional moment-generating functions (MGFs) of the stochastic process over any time horizon, and thus the corresponding skewnesses, are explicit or implicit functions of the stochastic process parameters. The upshot then is that we can produce horizon-consistent unconditional and conditional (i.e., in-sample and out-of-sample) skewness estimates accounting for compounding, leverage, and other effects in asset prices simply by calibrating short-horizon (e.g., daily) asset returns to some realistic ASV class process. Heeding that strategy, the main challenge becomes to accurately and efficiently estimate the process parameters.

We use Heston's (1993) stochastic volatility process as an illustrative example to implement our estimator. Our rationale for choosing this process is that it is simple, flexible, and popular, accounts for both compounding and leverage effects, and yields a closed-form MGF. Moreover, we show that we can use a simple novel generalized method of moments (GMM) estimator to consistently and precisely estimate the parameters of that process. Relying on a comprehensive Monte Carlo simulation exercise in which Heston's (1993) process is also the true data-generating process, we demonstrate that our estimator delivers highly precise estimates of the skewness of discrete returns over up to five-year horizons — even when we employ no more than ten years of daily data in the parameter estimation. While the mean true conditional (i.e., out-of-sample) skewnesses of the daily, weekly, monthly, quarterly, annual, three-year, and five-year return in the base case are, for example, 0.005, 0.015, 0.051, 0.154, 0.576, 1.400, and 2.117, the mean (out-of-sample) estimates of our estimator are 0.012 (mean squared error (MSE): 0.001), 0.029 (0.005), 0.075 (0.017), 0.179 (0.029), 0.559 (0.040), 1.286 (0.074), and 1.910 (0.145), all respectively. Our comparative statics report that the estimator performs similarly under a wide set of alternative plausible stochastic process parameterizations. In the same vein, the estimator shows a comparable performance in capturing unconditional skewness (i.e., in-sample analyses) over the same set of horizons.

As stressed, the impressive performance of our estimator in the simulation exercise must ultimately derive from our novel GMM estimator being able to accurately estimate the parameters of Heston's (1993) process. In agreement, the GMM estimator does indeed do an excellent job in estimating the asset price drift (μ), the long-run variance (α), the volatility of variance (ξ), and the asset price-asset volatility correlation (ρ) but struggles with the mean-reversion-in-variance speed (κ). While the true values of μ, κ, α, ξ, and ρ are, for example, 0.100, 3.000, 0.090, 0.300, and 0.500 in the base case, the average GMM estimates are 0.108 (MSE: 0.011), 3.727 (4.994), 0.077 (0.000), 0.306 (0.019), and 0.506 (0.049), respectively. Our comparative statics show that the GMM estimator performs similarly well under the other parameterizations.

We next benchmark our estimator against the sample skewness and the recent alternative estimators of Fama and French (2018), Neuberger and Payne (2021), and Farago and Hjalmarsson (2023). While Fama and French (2018) and Farago and Hjalmarsson (2023) derive a bootstrap and closed-form estimator based on the assumption that asset returns are i.i.d., Neuberger and Payne (2021) derive a closed-form estimator not requiring that assumption by approximating return moments. Strikingly, the other estimators are all markedly less precise than ours. In line with Farago and Hjalmarsson (2023), we find that the sample skewness is on target over short horizons but becomes increasingly downward biased over longer horizons. Conversely, Fama and French's (2018) and Farago and Hjalmarsson's (2023) mean estimates are consistently upward biased since they fail to pick up the negative impact of the leverage effect on skewness. Finally, Neuberger and

Payne's (2021) mean estimate is also consistently downward biased since it approximates away the positive impact of compounding up returns to longer horizons on skewness (see Bessembinder (2018)). Perhaps even more noteworthily, the MSE of our estimator is usually at least three times smaller than those of the other estimators.

A concern with our estimator is that deviations between the stochastic process assumed by the estimator to describe the evolution of an asset's price over time and the true process may lead the performance of our estimator to collapse. To mitigate that concern, we next assume that the true process is either the double-Heston (see Christoffersen et al. (2009)) or the Heston-jump (see Bates (1996)) process. We then repeat our Monte Carlo simulation exercise, applying the Heston (1993) estimator to data simulated from those processes. Our evidence shows that the deviations from the true process only marginally affect the performance of our estimator, with our mean estimates remaining close to their true values. While the mean true conditional (i.e., out-of-sample) skewnesses of the daily, weekly, monthly, quarterly, annual, three-year, and five-year return are, for example, 0.030, 0.059, 0.069, 0.027, 0.515, 1.402, and 2.165 in our double-Heston simulation, our mean (out-of-sample) estimates are 0.017, 0.032, 0.029, 0.060, 0.497, 1.303, and 1.978, respectively. More importantly, our estimator continues to yield the lowest MSE regardless of the return horizon.

We finally evaluate whether our estimator captures time-series variations in the (risk-neutral) conditional skewness of the S&P 500 as embedded in SPX option prices. Running regressions of that skewness on the conditional (i.e., out-of-sample) estimates obtained from our estimator and its competitors, we find that our estimator is far superior to the others in capturing those time-series variations, with its adjusted R-squareds at least around twice the magnitude of those of the others. For example, while our estimator captures 42% of the variations in the conditional risk-neutral skewness of the one-month S&P 500 return, the sample skewness captures only 17% and the other estimators less than 4%. Noteworthily, Fama and French's (2018) and Farago and Hjalmarsson's (2023) estimators even predict those variations with the wrong (i.e., a negative) sign.

Our work adds to a small emerging literature on how to accurately estimate and forecast the skewness of long-horizon asset returns using limited amounts of data. As is well-known, standard estimators such as the sample skewness yield highly imprecise estimates in that situation (see, e.g., Lau et al. (1989)). Also, their estimates are biased in the presence of leverage and volatility feedback effects (Li (2020)). To increase precision, Fama and French (2018) propose creating artificial long-horizon returns by compounding up short-horizon returns obtained from a simple bootstrap and then applying the sample skewness to the artificial long-horizon returns, implicitly assuming i.i.d. returns. While Farago and Hjalmarsson (2023) develop an equivalent closed-form estimator, Aretz and Arisoy (2023) extend Fama and French's (2018) to a block bootstrap to account for return dependencies. Neuberger and Payne (2021) approximate an asset's variance and skewness operator to derive a closed-form estimator not requiring i.i.d. returns. Noteworthily, all these estimators capture unconditional (i.e., in-sample) but not conditional (i.e., out-of-sample) skewness.[2] We contribute by developing a precise parametric estimator that (i) is able to account for return dependencies; (ii) does not rely on approximations (i.e., is exact); and (iii) simultaneously yields unconditional and conditional (i.e., in-sample and out-of-sample) estimates.

Our work further relates to studies estimating the parameters of continuous-time stochastic volatility models. Previous approaches include GMM (e.g., Hansen and Scheinkman (1995) and Pan (2002)); maximum likelihood (e.g., Bakshi et al. (2006), Bates (2006), and Aït-Sahalia and Kimmel (2007)); simulated method 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

[2]An exception is Boyer et al. (2010) who use cross-sectional OLS regressions to forecast the skewness coefficient of daily discrete stock returns computed over the past 60 months using lagged exogenous variables (including that same skewness coefficient, daily stock return volatility, momentum, share turnover, size dummies, industry dummies, and an exchange dummy).

(2001) and Jiang and Knight (2002)). Unlike existing GMM estimators, we construct ours by matching the theoretical central moments and cross-moments of discrete returns with their sample counterparts, substituting small-time-interval approximations for the exact analytical moment expressions. A key advantage of us using the approximations is that it greatly reduces estimation time, making it more feasible to apply our estimator to large amounts of data.

We finally also contribute to an emerging literature in economics extracting macroeconomic and financial uncertainty proxies from latent stochastic volatility processes. To be more specific, Fernández-Villaverde et al. (2011), Fernández-Villaverde et al. (2015), Chan (2017), and Mumtaz and Musso (2021) employ that strategy to measure (forward-looking) interest rate, fiscal policy, inflation, and regional uncertainty. Similarly, Alessandri and Mumtaz (2019) and Shin and Zhong (2020) use it to measure financial uncertainty. In the same spirit, we extract skewness (i.e., the second and third moments) from stochastic volatility processes.

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 (1993) estimator based on our setup and show how to implement it in practice. Section 4 offers a simulation exercise studying the unbiasedness and stability of our Heston (1993) estimator as well as those competing ones advocated in the recent literature. In Section 5, we evaluate the reliability of our Heston (1993) estimator using real-world data. Section 6 sums up and concludes.

1.2 A General Parametric Skewness Estimator

In this section, we introduce our novel parametric estimator of the skewness of an asset's discrete return over any arbitrary horizon. We start with defining the subset of affine stochastic volatility (ASV) processes which we can use to model an asset's price dynamics under our estimator. We then outline how to calculate unconditional and conditional discrete-return skewness for those processes. We next narrow down our discussion to four specific processes within our subset of the ASV class, deriving expressions for the two types of skewness. We finally describe how to estimate conditional and unconditional discrete-return skewness under our estimator.

1.2.1 Affine Stochastic Volatility Processes

To derive our estimator, we assume that we can use a stochastic process from a subset of the ASV class to describe an asset's price dynamics. Initially formalized in Dai and Singleton (2000) and Duffie et al. (2003), the ASV class has been widely used due to its rich dynamics and well-known analytical properties (see e.g., Duffie et al. (2000), Chernov et al. (2003), Cheridito et al. (2007), Anderson et al. (2010), Mencía and Sentana (2013), etc.). More specifically, the subset of ASV stochastic processes studied by us satisfies the following assumption:

Assumption 1. *On a canonical filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$, let $(X_t, \boldsymbol{Y_t})_{t \geq 0}$ 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 log-price process, and the d-dimensional vector-valued process $\boldsymbol{Y_t}$ is the latent state process associated with X_t. We use **bold** font to distinguish vectors from scalars. $\boldsymbol{Y_t}$ can be further decomposed into $(\boldsymbol{V_t}, \boldsymbol{N_t})'$, where $\boldsymbol{V_t}$ is a d_1-dimensional process collecting the continuous components of $\boldsymbol{Y_t}$, and $\boldsymbol{N_t}$ is a d_2-dimensional pure jump process with $d_2 = d - d_1$. We further assume that:*

1. *There exist unique functions $\phi(u, \boldsymbol{w}, h) \in \mathbb{C}$ and $\boldsymbol{\psi}(u, \boldsymbol{w}, h) \in \mathbb{C}^{d_1}$ such that the joint conditional moment generating function (MGF) $M_t(u, \boldsymbol{w}, h; \boldsymbol{Y_t})$ has an exponential-linear form:*

$$M_t(u, \boldsymbol{w}, h; \boldsymbol{Y_t}) := \mathrm{E}[e^{r_{t,h}u + \langle \boldsymbol{Y_{t+h}}, \boldsymbol{w} \rangle} | X_t, \boldsymbol{Y_t}] = e^{\mu h u + \phi(u,\boldsymbol{w},h) + \langle \boldsymbol{\psi}(u,\boldsymbol{w},h), \boldsymbol{V_t} \rangle + \langle \boldsymbol{w_2}, \boldsymbol{N_t} \rangle}, \tag{1.2.1}$$

for all $(u, \boldsymbol{w}, h) \in \mathbb{C} \times \mathbb{C}^d \times \mathbb{R}_+$ *and* $t \geq 0$ *if* $\mathrm{E}[\cdot]$ *exists, and where* $r_{t,h} := X_{t+h} - X_t$ *is the log-return of* X_t *over horizon* h, μ *is the drift of* X_t, *and* $\boldsymbol{w_2}$ *is the subvector of* $\boldsymbol{w} = (\boldsymbol{w_1}, \boldsymbol{w_2})'$ *that corresponds to* $\boldsymbol{N_t}$.[3]

2. *As* $t \to \infty$, $\boldsymbol{V_t}$ *converges in law to a unique invariant limit distribution with MGF given by* $G(\boldsymbol{w_1}) = \mathrm{E}[e^{\langle \boldsymbol{w_1}, \boldsymbol{V_t} \rangle}]$.

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

In Assumption 1, we usually interpret $\boldsymbol{V_t}$ as the components of spot variance and $\boldsymbol{N_t}$ as a Poisson counter capturing the jump occurrence in X_t. For example, the Heston (1993) and the double-Heston process of Christoffersen et al. (2009) model, respectively, $\boldsymbol{V_t}$ as a single and two independent Cox-Ingerson-Ross (CIR; 1985) processes, with N_t a trivial zero process. Assuming the same dynamics for V_t as Heston (1993), the Heston-jump process of Bates (1996) allows N_t to be nonzero. The specific functional forms of $\phi(u, \boldsymbol{w}, h)$ and $\boldsymbol{\psi}(u, \boldsymbol{w}, h)$ depend on the choice of processes governing $\boldsymbol{Y_t}$. We will offer concrete examples in Section 1.2.2.2 below.

The conditions in Assumption 1 restrict the type of ASV stochastic processes which we can assume under our estimator. In line with Duffie et al. (2003), condition 1 defines an ASV process which is homogenous of degree one in X_t. Condition 2 implies that $\boldsymbol{V_t}$ is a stationary process, so that a stationary regime for $(X_t, \boldsymbol{V_t})$ exists (see Keller-Ressel and Steiner (2008) and Keller-Ressel (2011)) and the unconditional moments of X_t are defined. Nevertheless, it precludes models with non-stationary spot variance.[4] Condition 3 states that the third moment of X_t does not explode over any finite horizon h, guaranteeing the existence

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

[4]If the volatility process is not stationary, there is no steady-state distribution for the volatility process. This simply means that one cannot even define the "unconditional" skewness, since the return process itself does not have a time-invariant moment. Non-stationary volatility process includes stochastic processes that are not stationary in nature (e.g. a Brownian motion, or any jump process with stationary jump increments).

of discrete-return skewness over any arbitrary horizon. This assumption is critical as the exponential moments of an ASV stochastic process do not necessarily exist over an arbitrary horizon unless certain parameter constraints are satisfied (see Andersen and Piterbarg (2007), Keller-Ressel (2011), and Keller-Ressel and Mayerhofer (2015) for an extensive discussion).

The joint MGF $M_t(u, \boldsymbol{w}, h; \boldsymbol{Y_t})$ conditional on $\boldsymbol{Y_t}$ plays a critical role for our estimator, reflecting the joint exponential moments of $(X_t, \boldsymbol{Y_t})$ based on the current values of the state variables. Since $\boldsymbol{Y_t}$ is unobservable in practice, we cannot evaluate $M_t(u, \boldsymbol{w}, h; \boldsymbol{Y_t})$ directly. To avoid this issue, we first ignore the $\boldsymbol{N_t}$ component in $\boldsymbol{Y_t}$ by setting $\boldsymbol{w_2} = \boldsymbol{0}$. Defining $\boldsymbol{w^-_{d\times 1}} := (\boldsymbol{w_1}, \boldsymbol{0})'$, the joint conditional MGF of $(X_t, \boldsymbol{V_t})$ then takes on the form:

$$M_t(u, \boldsymbol{w}^-, h; \boldsymbol{Y_t}) = e^{\mu h u + \phi(u, \boldsymbol{w}^-, h) + \langle \boldsymbol{\psi}(u, \boldsymbol{w}^-, h), \boldsymbol{V_t} \rangle}, \tag{1.2.2}$$

which is independent of $\boldsymbol{N_t}$ but still requires knowledge about the latent process $\boldsymbol{V_t}$ to be computable. To derive the corresponding unconditional joint MGF of $(X_t, \boldsymbol{V_t})$, we next rely on Assumption 1 and integrate out $\boldsymbol{V_t}$ based on its invariant limit distribution:

$$M(u, \boldsymbol{w}^-, h) := \mathrm{E}[M_t(u, \boldsymbol{w}^-, h; \boldsymbol{V_t})] = e^{\mu h u + \phi(u, \boldsymbol{w}^-, h)} G(\boldsymbol{\psi}(u, \boldsymbol{w}^-, h)), \tag{1.2.3}$$

which depends only on the unconditional distribution of $\boldsymbol{V_t}$ through the $G(\boldsymbol{w_1})$ function and can be computed without knowledge of $\boldsymbol{Y_t}$. Since $M_t(u, \boldsymbol{w}^-, h; \boldsymbol{Y_t})$ and $M(u, \boldsymbol{w}^-, h)$ are linear combinations of exponential moments of $r_{t,h}$ and $\boldsymbol{V_{t+h}}$, we then calculate the, respectively, conditional and unconditional moments necessary to derive discrete-return skewness from:

$$M_t(u, \boldsymbol{0}, h; \boldsymbol{V_t}) = \mathrm{E}[e^{r_{t,h} u} | \boldsymbol{V_t}] \qquad \mathrm{E}[R^u_{t,h} | \boldsymbol{V_t}] \tag{1.2.4}$$

$$M(u, \boldsymbol{0}, h) = \mathrm{E}[e^{r_{t,h} u}] \qquad \mathrm{E}[R^u_{t,h}], \tag{1.2.5}$$

where $R_{t,h} := e^{r_{t,h}} = e^{X_{t+h} - X_t}$ is the discrete return. To simplify our notation, we suppress the

MGF parameter $\boldsymbol{w}$ if it is zero (i.e., $M_t(u,h;\boldsymbol{V_t}) \equiv M_t(u,\mathbf{0},h;\boldsymbol{V_t})$ and $M(u,h) \equiv M(u,\mathbf{0},h)$). The MGFs $M_t(u,h;\boldsymbol{V_t})$ and $M(u,h)$ serve as the building blocks for the derivation of our ASV-model-implied skewness estimator, to be introduced in the next subsection.

1.2.2 ASV Model Implied Discrete-Return Skewness

1.2.2.1 The General Case

We now derive the closed-form or quasi-closed-form expressions for the skewness of discrete returns over any arbitrary horizon h implied by the subset of ASV stochastic processes defined in Assumption 1. Recall that the (standardized) conditional skewness of $R_{t,h}$ is defined as:

$$\text{Skew}[R_{t,h}|\mathcal{F}_t] := \frac{\text{E}[(R_{t,h} - \text{E}[R_{t,h}|\mathcal{F}_t])^3|\mathcal{F}_t]}{\text{E}[(R_{t,h} - \text{E}[R_{t,h}|\mathcal{F}_t])^2|\mathcal{F}_t]^{3/2}}, \tag{1.2.6}$$

whereas the corresponding unconditional skewness is defined as:

$$\text{Skew}[R_{t,h}] := \frac{\text{E}[(R_{t,h} - \text{E}[R_{t,h}])^3]}{\text{E}[(R_{t,h} - \text{E}[R_{t,h}])^2]^{3/2}}. \tag{1.2.7}$$

Following from Section 1.2.1, the skewness of $R_{t,h}$ conditional on $\boldsymbol{Y_t}$ can be written in terms of the conditional moments $M_t(u,h;\boldsymbol{V_t})$ in Eq. (1.2.4):

$$\text{Skew}[R_{t,h}|\mathcal{G}_t] = \frac{M_t(3,h;\boldsymbol{V_t}) - 3M_t(1,h;\boldsymbol{V_t})M_t(2,h;\boldsymbol{V_t}) + 2M_t(1,h;\boldsymbol{V_t})^3}{[(M_t(2,h;\boldsymbol{V_t}) - M_t(1,h;\boldsymbol{V_t}))^2]^{3/2}}, \tag{1.2.8}$$

where $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0} \subseteq \mathcal{F}$ and $\mathcal{G}_t = \sigma(X_t, t \geq 0)$ denotes the internal filtration of log-prices up to time t. We can interpret Eq. (1.2.8) as our best skewness estimate of $R_{t,h}$ conditional on the full price trajectory until time t. Conversely, the unconditional skewness of $R_{t,h}$

can be written in terms of the unconditional moments $M(u,h)$ Eq. (1.2.5):

$$\text{Skew}[R_{t,h}] = \frac{M(3,h) - 3M(1,h)M(2,h) + 2M(1,h)^3}{[M(2,h) - M(1,h)^2]^{3/2}}, \tag{1.2.9}$$

which can be interpreted as our best estimate in the absence of information on $\boldsymbol{Y_t}$.

We stress that Eq. (1.2.8) crucially depends on the assumption that we are able to continuously observe X_t over time. In that case, we can perfectly estimate the latent process $\boldsymbol{V_t}$, so that we can compute $\text{Skew}[R_{t,h}|\boldsymbol{V_t}]$, and it follows from the Markov property of the ASV class defined in Section 1.2.1 that $\text{Skew}[R_{t,h}|\mathcal{G}_t] = \text{Skew}[R_{t,h}|\boldsymbol{V_t}]$. Unfortunately, as the complete trajectory of X_t is usually unavailable, we must view $\text{Skew}[R_{t,h}|\mathcal{G}_t]$ as an infeasible *holy grail* for predicting skewness in practice. What we usually have at our disposal is a discrete equidistant sample of X_t. Let us denote that sample by $(X_{i\Delta})_{i=0:N}$, where Δ is the fixed time interval between two consecutive prices and N is the number of observations (e.g., $\Delta = 1/252$ and $N = 252$ indicate a sample of daily log-prices over a year). To ease our notation, let us further write X_i rather than $X_{i\Delta}$ to differentiate discrete from continuous-time observations. Denote by $\mathcal{G}_{i\Delta} = \sigma(X_n, n \in \{0, 1, \ldots, i\})$ the information set generated by the discrete observations $(X_n)_{n=0:i}$, with $\mathcal{G}_{i\Delta} \subseteq \mathcal{G}_{i\Delta}$. We then take the $\mathcal{G}_{i\Delta}$-conditional expectation of the full-path conditional MGF $M_{i\Delta}(u, \boldsymbol{w}^{-}, h; \boldsymbol{V_{i\Delta}})$ to reflect the fact that we are conditioning on a coarser information set:

$$M_{i\Delta}(u, \boldsymbol{w}^{-}, h) := \mathrm{E}[M_{i\Delta}(u, \boldsymbol{w}^{-}, h; \boldsymbol{V_{i\Delta}})|\mathcal{G}_{i\Delta}]. \tag{1.2.10}$$

While $M_{i\Delta}(u, \boldsymbol{w}^{-}, h)$ is generally unavailable in closed form for ASV-class processes, the following proposition provides a method to iteratively evaluate $M_{i\Delta}(u, \boldsymbol{w}^{-}, h)$ and thus $M_{i\Delta}(u, h)$:

Proposition 1. *Under Assumption 1, assume that $\boldsymbol{V_0}$ is a random draw from its in-*

variant limit distribution, and X_0 is a deterministic initial log-price. Let $G_i(\boldsymbol{w_1}|\mathcal{G}_i) :=$ $\mathrm{E}[e^{\langle \boldsymbol{w_1}, \boldsymbol{V_{i\Delta}} \rangle}|\mathcal{G}_i]$ denote the MGF of $\boldsymbol{V_{i\Delta}}$ conditioning on $\mathcal{G}_i$. It must then hold that:

$$M_i(u, \boldsymbol{w}^-, h) = e^{\mu hu + \phi(u, \boldsymbol{w}^-, h)} G_i(\boldsymbol{\psi}(u, \boldsymbol{w}^-, h)|\mathcal{G}_i), \tag{1.2.11}$$

$$G_{i+1}(\boldsymbol{w_1}|\mathcal{G}_{i+1}) = \frac{\int_{0\ \ i1}^{0+i1} M_i(u, \boldsymbol{w}^-, h) e^{\ ur_{i,1}} du}{\int_{0\ \ i1}^{0+i1} M_i(u, h) e^{\ ur_{i,1}} du}, \tag{1.2.12}$$

where $G_0(\boldsymbol{w_1}|\mathcal{G}_0) = G(\boldsymbol{w_1})$, and $r_{i,1} = X_{i+1} \quad X_i$.

Proof. See Proposition 1 of Bates (2006) and his discussion in Section 1.2. □

In practice, we always start with Eq. (1.2.11) and calculate:

$$M_0(u, \boldsymbol{w}^-, h) = e^{\mu hu + \phi(u, \boldsymbol{w}^-, h)} G(\boldsymbol{\psi}(u, \boldsymbol{w}^-, h)). \tag{1.2.13}$$

We next plug $M_0(u, \boldsymbol{w}^-, h)$ into Eq. (1.2.12), numerically evaluating the integrals and obtaining $G_1(\boldsymbol{w_1}|\mathcal{G}_1)$. We then compute $G_1(\boldsymbol{w_1}|\mathcal{G}_1)$ at $\boldsymbol{\psi}(u, \boldsymbol{w}^-, h)$ and plug it back into Eq. (1.2.11), allowing us to compute the optimal expected MGF at time Δ, $M_1(u, \boldsymbol{w}^-, h)$. We iterate these steps until index i to obtain $M_i(u, \boldsymbol{w}^-, h)$. However, as stressed by Bates (2006), a direct implementation of Proposition 1 suffers from the curse of dimensionality due to the iterative nature of the numerical integration. We thus follow the steps in Section 1.3 of Bates (2006) and use a moment-matching technique to obtain a numerically stable approximation of $M_i(u, \boldsymbol{w}^-, h)$. We offer more details about the moment-matching technique in Appendix 1.A.

Relying on $M_i(u, \boldsymbol{w}^-, h)$, we define the observed-path conditional skewness as:

$$\mathrm{Skew}[R_{i\ ,h}|\mathcal{G}_i] = \frac{M_i(3,h) \quad 3M_i(1,h)M_i(2,h) + 2M_i(1,h)^3}{[M_i(2,h) \quad M_i(1,h)^2]^{3/2}}, \tag{1.2.14}$$

which we interpret as our best estimate conditional on observed prices.

1.2.2.2 Some Specific Examples

We next present the conditional and unconditional MGFs of some popular ASV processes, including the geometric Brownian motion, Heston, Heston-jump, and multifactor-Heston processes. We can then directly calculate the conditional and unconditional skewness of discrete returns from Eqs. (1.2.8) and (1.2.9), the relevant MGFs, and parameter values.

Geometric Brownian motion. Suppose $(X_t, \boldsymbol{Y_t})_{t \geq 0}$ satisfies $N_t \equiv 0$, $V_t = \sigma^2$, and:

$$dX_t = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t, \tag{1.2.15}$$

where σ is the volatility parameter, and W_t is a Brownian motion. Since V_t is constant, the conditional and unconditional MGFs coincide:

$$M_t(u, h; V_t) = M(u, h) = e^{(\mu h - \frac{1}{2}\sigma^2 h)u + \frac{1}{2}\sigma^2 h u^2}, \tag{1.2.16}$$

as do the (full and observed-path) conditional and unconditional skewness:

$$\text{Skew}[R_{t,h}|\mathcal{G}_t] = \text{Skew}[R_{t,h}] = \frac{e^{3h\sigma^2} - 3e^{h\sigma^2} + 2}{(e^{h\sigma^2} - 1)^{3/2}}, \tag{1.2.17}$$

which we obtain from plugging Eq. (1.2.16) into Eq. (1.2.9) and simplifying. Eq. (1.2.17) is similar to Farago and Hjalmarsson's (2023) estimator, which also assumes i.i.d. short-horizon discrete returns but does not require them to be lognormally distributed.

The Heston process. To let an asset's volatility evolve stochastically and to allow its price and volatility to be correlated, Heston (1993) assumes that $(X_t, \boldsymbol{Y_t})_{t \geq 0}$ satisfies $N_t \equiv 0$ and:

$$\begin{aligned} dX_t &= (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t, \\ dV_t &= \kappa(\alpha - V_t)dt + \xi\sqrt{V_t}dB_t, \end{aligned} \tag{1.2.18}$$

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 $[W, B]_t = \rho t$. 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 Del Baño Rollin et al. (2010) all derive the conditional MGF of the Heston process. Relying on Bates's (2006) expression, we have:

$$M_t(u, w_1, h; V_t) = e^{\mu h u + \phi_H(u,w_1,h) + \psi_H(u,w_1,h)V_t}, \tag{1.2.19}$$

where:

$$\begin{aligned} \phi_H(u, w_1, h) &= \frac{\kappa\alpha h}{\xi^2}D(u) - \frac{2\kappa\alpha}{\xi^2}\ln\{(1 + \frac{e^{I(u)h} - 1}{2I(u)}D(u))(1 - w_1 A(u,h))\}, \\ \psi_H(u, w_1, h) &= \frac{(u^2 - u)A(u,h)}{\xi^2} + \frac{w_1 B(u,h)}{1 - w_1 A(u,h)}, \quad A(u,h) = \frac{\xi^2}{I(u)C(u,h) + \kappa - u\xi\rho}, \\ B(u,h) &= \frac{C(u,h)^2 - 1}{(C(u,h) + \frac{\kappa - u\xi\rho}{I(u)})^2}, \quad C(u,h) = \frac{e^{I(u)h} + 1}{e^{I(u)h} - 1}, \quad D(u) = \kappa - u\xi\rho + I(u), \\ \text{and} \quad I(u) &= \sqrt{(\kappa - u\xi\rho)^2 - \xi^2(u^2 - u)} \end{aligned} \tag{1.2.20}$$

As the invariant limit distribution of $(V_t)_{t \geq 0}$ is Gamma, the function $G(w_1)$ in Eq. (1.2.3) is:

$$G(w_1) = (1 - \frac{\xi^2}{2\kappa}w_1)^{-2\kappa\alpha/\xi^2}, \tag{1.2.21}$$

and, as a consequence, the unconditional MGF is:

$$M(u, w_1, h) = e^{\mu u h + \phi_H(u, w_1, h)} G(\psi_H(u, w_1, h)). \tag{1.2.22}$$

The Heston-jump process. To allow for stochastic volatility and discontinuity, Bates (1996) adds a jump component to the price dynamic of the Heston process, with $(X_t, \boldsymbol{Y_t})_{t \geq 0}$ satisfying:

$$\begin{aligned} dX_t &= (\mu - \lambda\overline{\gamma} - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t + \gamma dN_t, \\ dV_t &= \kappa(\alpha - V_t)dt + \xi\sqrt{V_t}dB_t, \end{aligned} \tag{1.2.23}$$

where $\gamma \sim N(\mu_j, \sigma_j^2)$ is a random Gaussian jump, $\overline{\gamma} := \mathrm{E}[e^{\gamma} - 1] = e^{\mu_j + \frac{1}{2}\sigma_j^2} - 1$ is the expected jump size, and λ is the jump intensity. Bates (2006) expresses the conditional MGF of this process as:

$$M_t(u, \boldsymbol{w}, h; \boldsymbol{Y_t}) = e^{\mu h u + \phi_{HJ}(u, \boldsymbol{w}, h) + \psi_H(u, w_1, h)V_t + w_2 N_t}, \tag{1.2.24}$$

where:

$$\phi_{HJ}(u, \boldsymbol{w}, h) = \phi_H(u, w_1, h) + \lambda h(e^{u\mu_j + \frac{1}{2}u^2\sigma_j^2 + w_2} - u\overline{\gamma} - 1). \tag{1.2.25}$$

As a result, the unconditional MGF is:

$$M(u, \boldsymbol{w}, h) = e^{\mu h u + \phi_{HJ}(u, \boldsymbol{w}, h)} G(\psi_H(u, w_1, h)) e^{\lambda(e^{w_2} - 1)}.[5] \tag{1.2.26}$$

The multi-factor Heston process. To allow for a stochastic volatility model with multiple volatility factors, Christoffersen et al. (2009) propose the multi-factor Heston process, which generalizes the Heston process by modelling variance using $d_1 > 1$ independent CIR

[5] We could allow N_t to have a volatility-dependent intensity. Yet, as the resulting stochastic process is much more complicated than the described one, we refer the interested reader to Appendix A.2 of Bates (2006).

processes. $(X_t, \boldsymbol{Y_t})_{t \geq 0}$ now satisfies $N_t \equiv 0$, $\boldsymbol{V_t} = (V_t^{(k)})'_{k=1:d_1}$, and:

$$
\begin{aligned}
dX_t &= (\mu - \frac{1}{2}\sum_{k=1}^{d_1} V_t^{(k)})dt + \sum_{k=1}^{d_1} \sqrt{V_t^{(k)}} dW_t^{(k)}, \\
V_t^{(k)} &= \kappa^{(k)}(\alpha^{(k)} - V_t^{(k)})dt + \xi^{(k)} \sqrt{V_t^{(k)}} dB_t^{(k)}, \quad k \in [1, 2, \dots, d_1],
\end{aligned}
\tag{1.2.27}
$$

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 as in the Heston process. Since each $V_t^{(k)}$ evolves independently, the conditional MGF of the multi-factor Heston process takes the form:

$$
M_t(u, \boldsymbol{w_1}, h; \boldsymbol{V_t}) = e^{\mu h u + \sum_{k=1}^{d_1} \phi_H^{(k)}(u, w_1^{(k)}, h) + \sum_{k=1}^{d_1} \psi_H^{(k)}(u, w_1^{(k)}, h) V_t^{(k)}}, \tag{1.2.28}
$$

where $\boldsymbol{w_1} = (w_1^{(k)})'_{k=1:d_1}$, $\phi_H^{(k)}(u, w_1^{(k)}, h)$ and $\psi_H^{(k)}(u, w_1^{(k)}, h)$ are as in Eq. (1.2.20) using $\kappa = \kappa^{(k)}$, $\alpha = \alpha^{(k)}$, $\xi = \xi^{(k)}$, and $\rho = \rho^{(k)}$. The unconditional MGF thus has the below form:

$$
M(u, \boldsymbol{w_1}, h) = e^{\mu h u + \sum_{k=1}^{d_1} \phi_H^{(k)}(u, w_1^{(k)}, h)} \prod_{k=1}^{d_1} G^{(k)}(\psi_H^{(k)}(u, w_1^{(k)}, h)), \tag{1.2.29}
$$

where $G^{(k)}(w_1^{(k)})$ is as in Eq. (1.2.21) under the same assumption as above.

Other ASV processes and beyond. When Assumption 1 holds, Theorem 2.7 in Duffie et al. (2003) shows that $\phi(\cdot)$ and $\boldsymbol{\psi}(\cdot)$ can always be obtained as the solution of a system of general Riccati equations. Also, $G(\cdot)$ 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 numerically. Given that, we can always calculate discrete-return skewness from Eqs. (1.2.8) and (1.2.9) if the asset's price obeys Assumption 1, even when we look into more complicated and sophisticated processes than those explicitly considered.[6]

[6] While only a 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 one only needs $\phi(\cdot)$

In essence, the ASV class imposes a parametric structure on the asset price in continuous-time, allowing us to compute conditional and unconditional skewness from the corresponding MGFs. While it is possible to specify a parametric model for returns outside the ASV class (e.g., the discrete-time stochastic volatility model in Taylor (1994) or the GARCH-type models in Engle (1982), Bollerslev (1986), Nelson (1991), etc.), to our best knowledge, only the ASV class features time-scalable MGFs. The upshot is that making inferences about skewness over arbitrary horizons becomes challenging for models outside the ASV class, as it requires simulating the long-horizon return distribution for each target horizon and conditioning variable. We thus do not consider this alternative and focus on ASV stochastic processes in our work.

1.2.3 Estimating Skewness from ASV Stochastic Processes

Assuming that an asset's price dynamics follow a stochastic process from the ASV class defined in Assumption 1, Section 1.2.2 shows that the conditional and unconditional skewness of the asset's discrete return over any arbitrary horizon are (explicit or implicit) functions of the corresponding MGFs of the process and, in turn, of its parameters. Based on those insights, we next show that we can thus estimate the two types of skewness simply by estimating those parameters. To do so, let us assume that we start from a sample of daily log-prices $(X_i)_{i=0:N}$. To estimate the parameters of ASV-class stochastic processes and to convert them into skewness estimates, we make the following assumptions in addition to those already stated:

Assumption 2. *Consider a stochastic process from the ASV model class in Assumption 1 generating $(X_t, \boldsymbol{Y_t})_{t \geq 0}$. We further assume about that stochastic process that:*

and $\boldsymbol{\psi}( )$ with $u, h \neq 0$ for the purpose of option pricing, which are much easier to compute than their (i.e., $\boldsymbol{w} \neq 0$) general counterparts. Pan (2002), Bates (2006), and Bates (2012) show how to derive MGFs for ASV processes with $d_1 = 1$, while Bates (2019) discusses how to derive those MGFs for processes with $d_1 > 1$.

1. *The process is uniquely determined by some K-dimensional parameter vector $\boldsymbol{\theta} \in \Theta$, where the space Θ is a compact subspace of $\mathbb{R}^K$. For some realization of $(X_t, \boldsymbol{Y_t})_{t \geq 0}$, we denote the true data-generating parameter vector by $\boldsymbol{\theta_0}$, taking values from the interior of Θ.*

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

3. *The conditional and unconditional MGFs $M_t(u, h; \boldsymbol{V_t})$ and $M(u, h)$ are almost everywhere continuous functions of $\boldsymbol{\theta}$ on Θ for all $h > 0$ and $u = 1, 2, 3$.*

Condition 1 ensures that the parameter vector $\boldsymbol{\theta_0}$ associated with the ASV stochastic process is identified, while condition 2 assumes the existence of an N-consistent estimator of $\boldsymbol{\theta_0}$ based on equidistantly observed data. As Bates (2006) shows that many estimation techniques (as, e.g., the generalized method of moments (GMM), the efficient method of moments, and the empirical characteristic function-based methods) consistently estimate the parameters of a wide class of ASV processes, the two conditions are not restrictive. Condition 3 ensures that the exponential moments of the process are smooth functions of the parameter vector, which is also satisfied by the vast majority of ASV processes. The final condition guarantees that plugging the consistent parameter estimates into Eqs. (1.2.8) and (1.2.9) yields consistent skewness estimates.

Under Assumptions 1 and 2, we can estimate the unconditional skewness of an asset's discrete return over time $i\Delta$ to $i\Delta + h$, $\text{Skew}[R_{i\Delta,h}; \hat{\boldsymbol{\theta}}]$, and the observed-path conditional skewness, $\text{Skew}[R_{i\Delta,h}; \hat{\boldsymbol{\theta}} | \mathcal{G}_{i\Delta}]$, by plugging the corresponding MGFs evaluated at the fitted parameter vector $\hat{\boldsymbol{\theta}}$ into Eqs. (1.2.9) and (1.2.14), respectively. Following from condition 3 of Assumption 2, the consistency of our skewness estimates immediately follows:

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

$$\begin{aligned} \text{Skew}[R_{i\,,h}; \hat{\boldsymbol{\theta}}] &\xrightarrow{p} \text{Skew}\,[R_{i\,,h}; \boldsymbol{\theta_0}]\,, \\ \text{Skew}[R_{i\,,h}; \hat{\boldsymbol{\theta}}|\mathcal{G}_{i}\,] &\xrightarrow{p} \text{Skew}[R_{i\,,h}; \boldsymbol{\theta_0}|\mathcal{G}_{i}\,]. \end{aligned} \tag{1.2.30}$$

Proof. Based on condition 3 of Assumption 2, the consistency of the skewness estimates follows from the continuous mapping theorem, and the proof is complete. □

We could in principle use the delta method to establish the asymptotic normality of the skewness estimates if $\hat{\boldsymbol{\theta}}$ is asymptotically normal. Notwithstanding, Yu (2012) shows that the finite sample distribution of the ASV estimates is typically nowhere close to normal, so that it is preferable to use bootstrap-based inferences to conduct statistical tests of the parameter estimates.

1.3 A Heston Process Estimator of Skewness

In this section, we offer more details on how to use a version of our parametric estimator assuming that asset price dynamics are governed by Heston's (1993) stochastic volatility process to estimate the conditional and unconditional skewness of discrete returns over arbitrary horizons (recall the second example in Section 1.2.2.2). To do so, we first devise a novel GMM approach to estimate the parameters of that stochastic process. Next, we propose approximations of the moments used in that approach to speed up our numerical solution technique and to make it more likely that the solution technique converges to a global (and not a local) minimum.

1.3.1 A New GMM Estimator for the Heston Process Parameters

We use a GMM system based on the centered MGFs and cross-MGFs of Heston's (1993) stochastic process to estimate the parameters of that process, $\boldsymbol{\theta} = (\mu, \kappa, \alpha, \xi, \rho)$. Toward that goal, we start with a series of daily log-price $(X_i)_{i=1:N}$ and use that to compute daily discrete returns $R_{i,\Delta} := e^{X_{i+1}-X_i}$. We next match the theoretical central moments and cross-moments of $R_{i,\Delta}$ with their sample counterparts. To be more rigorous, let us denote the centered version of $R_{i,\Delta}$ as $\tilde{R}_{i,\Delta} := R_{i,\Delta} - M(1,\Delta)$ and the centered MGF and cross-MGF of $R_{i,\Delta}$ as $\tilde{M}(k,\Delta) := \mathrm{E}[\tilde{R}_{i,\Delta}^k]$ and $\tilde{C}(m_1, m_2, \Delta, j) := \mathrm{Cov}[\tilde{R}_{i,\Delta}^{m_1}, \tilde{R}_{i+j,\Delta}^{m_2}]$, respectively, where $j = 1:J$ indicates the number of lags. In the following lemma, we establish that we are able to evaluate the cross-MGF function of the Heston (1993) process from its unconditional MGF function:

Lemma 1. *The uncentred cross-MGF function of the Heston (1993) process takes the form:*

$$\begin{aligned} C(m_1, m_2, \Delta, j) = \exp(\mu\Delta m_2 + \phi_H(m_2, 0, \Delta) + \phi_H(0, \psi_H(m_2, 0, \Delta), (j-1)\Delta)) \\ \times M(m_1, \psi_H(0, \psi_H(m_2, 0, \Delta), (j-1)\Delta), \Delta). \end{aligned} \tag{1.3.1}$$

Proof. See Appendix 1.B.1 for the step-by-step derivation. □

The expressions of $\tilde{M}(k,\Delta)_{k\in\{2,3,4\}}$ and $\tilde{C}(m,2,\Delta,j)_{m\in\{1,2\}}$ are offered in Appendix 1.B.2.

Using the centered MGFs and cross-MGFs, we consider the following system of unconditional moment conditions to be used in our GMM approach:

$$\mathrm{E}[\boldsymbol{h}(R_{i,\Delta};\boldsymbol{\theta})] := \mathrm{E}\begin{pmatrix} \tilde{R}_{i,\Delta} \\ (\tilde{R}_{i,\Delta}^k - \tilde{M}(k,\Delta))'_{k=2:4} \\ (\tilde{R}_{i,\Delta}^1 \tilde{R}_{i+j,\Delta}^2 - \tilde{C}(1,2,\Delta,j))'_{j=1:J} \\ (\tilde{R}_{i,\Delta}^2 \tilde{R}_{i+j,\Delta}^2 - \tilde{C}(2,2,\Delta,j))'_{j=1:J} \end{pmatrix} = \mathbf{0}_{(4+2J)\times 1}. \tag{1.3.2}$$

We choose the moment conditions in line with intuition to match the interpretation of the parameters. First, the four central moments allow us to identify the deviations from normality implied by Heston's (1993) process, as captured by μ, α and ξ. Second, the cross-moments $\tilde{C}(1,2,\Delta,j)$ reflect the decay of cross-covariances between past return and future variance, as captured by ρ and κ. Similarly, the cross-moments $\tilde{C}(2,2,\Delta,j)$ reflect the decay of autocovariances in squared demeaned returns, capturing the volatility clustering effect embedded in the stochastic process and allowing us identify κ and ξ. Critically, moment conditions of the form $\tilde{C}(n,1,\Delta,j)$ do not carry incremental information about the parameters over and above the n-th central moment as the Markov property of the Heston process implies that:

$$\mathrm{E}[\tilde{R}_{i-j,\Delta}^{n} \tilde{R}_{i,\Delta}] = \mathrm{E}[\tilde{R}_{i-j,\Delta}^{n}]\,\mathrm{E}[\tilde{R}_{i,\Delta}] = 0, \quad \forall j \geq 1. \tag{1.3.3}$$

Also, higher-order moment conditions are of a much smaller scale than the moment conditions used by us here when Δ is small. As a result, we do not consider those.

As the system of moment conditions (1.3.2) is overidentified even when $J = 1$, we use the GMM estimator to obtain parameter estimates. To that end, let $\boldsymbol{W}_{(4+2J)\times(4+2J)}$ denote a positive definite weighting matrix. The GMM estimator of the parameter vector $\boldsymbol{\theta}$ is defined as:

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{GMM}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, \bar{\boldsymbol{h}}(R_{i,\Delta};\boldsymbol{\theta})' \boldsymbol{W}_{(4+2J)\times(4+2J)} \bar{\boldsymbol{h}}(R_{i,\Delta};\boldsymbol{\theta}), \tag{1.3.4}$$

where $\bar{\boldsymbol{h}}(R_{i,\Delta};\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{h}(R_{i,\Delta};\boldsymbol{\theta})$ is the sample mean of the moment conditions. Under standard regularity conditions which hold for the Heston (1993) process (as discussed in, e.g., Chapter 3 of Hall (2005)), $\hat{\boldsymbol{\theta}}_{\boldsymbol{GMM}}$ is a consistent estimator of $\boldsymbol{\theta_0}$. As a result, keeping h fixed, Proposition 2 implies that the skewness estimates derived from $\hat{\boldsymbol{\theta}}_{\boldsymbol{GMM}}$ are also consistent.

In our simulation exercise and empirical implementation of our GMM estimator, we use a two-step approach with $J = 100$ lags to obtain $\hat{\boldsymbol{\theta}}_{\boldsymbol{GMM}}$. To be specific, we start with the identity matrix $\boldsymbol{I}_{204 \times 204}$ as weighting matrix in Eq. (1.3.4) to obtain a first-stage parameter vector estimate $\hat{\boldsymbol{\theta}}_{\mathbf{1}}$. Relying on $\hat{\boldsymbol{\theta}}_{\mathbf{1}}$, we compute the Newey and West (1987) variance-covariance matrix of $\boldsymbol{h}(R_{i,\cdot}; \hat{\boldsymbol{\theta}}_{\mathbf{1}})$, often labelled the spectral density matrix.[7] We then use the inverted spectral density matrix as weighting matrix to obtain our final parameter vector estimate $\hat{\boldsymbol{\theta}}_{\mathbf{2}}$. Regarding the choice of J, we choose it large enough to capture the decay in cross-covariances, but not so large to induce excessive noise into our estimation and/or to make the estimation too computationally burdensome. Our later simulation exercise suggests that $J = 100$ strikes a reasonable balance between accuracy and speed, motivating us to use that J value in our main paper.

1.3.2 Approximating the GMM Moment Conditions

To aid the numerical performance of our GMM estimator, we next develop small-time approximations of the MGFs and cross-MGFs used in System (1.3.2), inspired by our choice of $\Delta = 1/252$ being (relatively) small. Doing so is beneficial since solving the (non-approximated) optimization problem in Eq. (1.3.4) is not a trivial task. First, the problem is highly non-linear due to the complicated form of the MGFs, leading standard gradient-based algorithms to typically only converge to local minima depending on the choice of initial values. Moreover, the evaluation of the theoretical cross-MGFs is extremely time-consuming due to their recursive structure. Thus, it can take a significant amount of time to obtain even the first step estimate $\hat{\boldsymbol{\theta}}_{\mathbf{1}}$.

We present the approximated MGFs and cross-MGFs in Proposition 3:

Proposition 3. *The MGFs and cross-MGFs below can be expanded as follows as $\Delta \to 0$:*

1. $M(1, \Delta) = 1 + \mu\Delta + \frac{\mu^2}{2}\Delta^2 + o\left(\Delta^2\right)$.

[7] We use the Bartlett kernel with an automatic bandwidth choice $\lfloor 4(N/100)^{2/9} \rfloor$ to calculate that matrix.

2. $\tilde{M}(2,\Delta) = \alpha\Delta + \alpha\left(\frac{\xi^2}{4\kappa} + \frac{\alpha}{2} + 2\mu + \xi\rho\right)\Delta^2 + o\left(\Delta^2\right)$.

3. $\tilde{M}(3,\Delta) = 3\alpha\left(\alpha + \frac{\xi^2}{2\kappa} + \frac{\xi\rho}{2}\right)\Delta^2 + o\left(\Delta^2\right)$.

4. $\tilde{M}(4,\Delta) = 3\alpha\left(\alpha + \frac{\xi^2}{2\kappa}\right)\Delta^2 + o\left(\Delta^2\right)$.

5. $\tilde{C}(1,2,\Delta,j) = \alpha\xi\rho e^{-\kappa(j-1)}\Delta^2 + o\left(\Delta^2\right)$.

6. $\tilde{C}(2,2,\Delta,j) = \frac{\alpha\xi^2}{2\kappa} e^{-\kappa(j-1)}\Delta^2 + o\left(\Delta^2\right)$.

Proof. See the supplementary material of this paper. □

Not only are the approximated MGFs and cross-MGFs new to the literature but they also offer several interesting insights into the small-time behavior of the Heston (1993) stochastic process. First, it is obvious that we cannot identify all parameters of the process from the first four central moments alone because those moments only feature ξ^2/κ and $\xi\rho$ (and not those parameters on their own). Yet, also considering the cross-MGFs, the separate identification of all parameters becomes feasible. Next, the sign of $\tilde{C}(1,2,\Delta,j)$ is dictated by ρ, the parameter directly conditioning the strength of the leverage effect. Also, the $e^{-\kappa j}$ term appears in both $\tilde{C}(1,2,\Delta,j)$ and $\tilde{C}(2,2,\Delta,j)$ and controls the speed of decay of these cross-MGFs and, in turn, the degree of volatility clustering. Finally, higher-order moments above the fourth (as, e.g., $\tilde{M}(k,\Delta)$ for $k > 4$) have a numerically negligible order of $o(\Delta^2)$ in comparison to the first four central moments, providing further support for our choice of moment conditions.

The simple formulas for the approximated MGFs and cross-MGFs greatly ease the functional form of the optimization problem in Eq. (1.3.4), providing vast gains in computational speed and stability compared to the original problem. To further improve the speed and reliability of our numerical solution technique, we also rely on several other settings, including user-supplied gradients, randomized starting points, and parameter constraints ensuring that Assumption 1 holds. We offer more details about those other settings in Appendix 1.C.

1.4 Simulation Exercise

In this section, we run a simulation exercise to evaluate the absolute and relative performance of our Heston (1993) estimator from Section 1.3. We start by describing how we simulate asset prices from Heston-type stochastic processes (i.e., Heston (1993), Bates (1996), and Christoffersen et al. (2009)). We then briefly introduce recent competing estimators. Concentrating on the Heston (1993) process, we next show how the process parameters affect the true unconditional discrete-return skewness across horizons. After that, we examine how well our GMM estimator estimates the Heston (1993) process parameters and compare the preciseness of our Heston (1993) estimator with those of the competitors. We finally repeat the comparisons allowing the data-generating processes to deviate from the Heston (1993) process.

1.4.1 Data-Generating Processes

For each single simulation, we follow Broadie and Kaya (2006) and Andersen (2008) in generating 10,000 asset-price paths from a Heston-type process over a ten-year period. In doing so, we assume that the initial price, P_0, is 50, and we draw the initial variance of each factor, $V_0^{(k)}$, from its asymptotic Gamma distribution, $\Gamma(\frac{2\kappa^{(k)}\alpha^{(k)}}{(\xi^{(k)})^2}, \frac{2\kappa^{(k)}}{(\xi^{(k)})^2})$.[8] Relying on a daily sampling frequency, we set the discretization step to one day (i.e., $\Delta = \frac{1}{252}$), implying that μ, $\alpha^{(k)}$, $\xi^{(k)}$, and λ are stated per annum and that $N = 2,520$. We use the following approach to consecutively simulate observations for each time $i \in \{1, \ldots, N\}$ within a sample path:

[8]While there is only one variance factor in the Heston (1993) or the Heston-jump process, namely V, there are multiple such factors in the multi-factor Heston process.

1. We follow Cox et al. (1985) in simulating $V_i^{(k)}$ based on $V_{i-1}^{(k)}$ using:

$$V_i^{(k)} = \frac{\left(\xi^{(k)}\right)^2 \left(1 - e^{-\kappa^{(k)}}\right)}{4\kappa^{(k)}} \chi^2_{df,nc}, \tag{1.4.1}$$

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

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

3. In the absence of jumps, we follow Andersen (2008) in updating P_i by approximating the solutions to the stochastic differential equations for the asset value as:

$$P_i = P_{i-1} \exp\left[\mu\Delta + \sum_{k=1}^{d_1} f\left(V^{(k)}\right)\right], \tag{1.4.2}$$

where:

$$\begin{aligned} f\left(V^{(k)}\right) &= K_0^{(k)} + K_1^{(k)} V_{i-1}^{(k)} + K_2^{(k)} V_i^{(k)} + \sqrt{K_3^{(k)}\left(V_{i-1}^{(k)} + V_i^{(k)}\right)}\; Z^{(k)}, \\ K_0^{(k)} &= \frac{\rho^{(k)}\kappa^{(k)}\alpha^{(k)}}{\xi^{(k)}}\Delta, \quad K_1^{(k)} = \frac{1}{2}\Delta\left(\frac{\kappa^{(k)}\rho^{(k)}}{\xi^{(k)}} - \frac{1}{2}\right) - \frac{\rho^{(k)}}{\xi^{(k)}}, \\ K_2^{(k)} &= \frac{1}{2}\Delta\left(\frac{\kappa^{(k)}\rho^{(k)}}{\xi^{(k)}} - \frac{1}{2}\right) + \frac{\rho^{(k)}}{\xi^{(k)}}, \quad K_3^{(k)} = \frac{1}{2}\Delta\left(1 - \left(\rho^{(k)}\right)^2\right). \end{aligned} \tag{1.4.3}$$

4. If jumps can occur, we further simulate a Gaussian random variable $\gamma \sim N(\mu_j, \sigma_j^2)$ and a binomial random variable $Y \sim B(1, \lambda\Delta)$, and modify P_i updating scheme to:

$$P_i = P_{i-1} \exp\left[\left(\mu - \lambda\bar{\gamma}\right)\Delta + \sum_{k=1}^{d_1} f\left(V^{(k)}\right) + \gamma Y\right]. \tag{1.4.4}$$

5. We calculate the daily discrete return $R_{i,} = P_i/P_{i-1}$.

A key advantage of this approach is that it mitigates Euler discretization errors.

1.4.2 Competing Estimators

We next introduce a number of alternative skewness estimators.

1.4.2.1 Sample Skewness Estimator

The sample skewness of the discrete return $R_{i,h}$ is:

$$\text{Skew}[R_{i,h}] = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(R_{i,h} \quad \frac{1}{N}\sum_{i=1}^{N} R_{i,h}\right)^{3}}{\left(\frac{1}{N}\sum_{i=1}^{N}\left(R_{i,h} \quad \frac{1}{N}\sum_{i=1}^{N} R_{i,h}\right)^{2}\right)^{3/2}}, \tag{1.4.5}$$

where N is the number of observations. To ensure we have enough observations, even when looking into long-horizon returns, we consistently apply the estimator to overlapping returns for horizons longer than daily.

1.4.2.2 Fama and French's (2018) Bootstrap Estimator

Fama and French (2018) suggest to rely on a simple bootstrap to estimate the unconditional skewness of long-horizon discrete returns. In particular, they propose to create a large number of bootstrap samples by drawing with replacement and equal probabilities observations from a time-series of short-horizon (e.g., daily) returns. Using each bootstrap sample, they next compound up the short-horizon returns to the desired long horizon. They finally apply the sample skewness estimator to the bootstrapped long-horizon returns to compute the skewness estimate. While Fama and French's (2018) estimator is intuitive, their simple bootstrap abstracts from return dependence, implicitly assuming that returns are i.i.d. over time. To allow for some return dependence, Aretz and Arisoy (2023) propose to replace the simple bootstrap used to create the bootstrapped long-horizon returns with a block bootstrap. However, considering that the optimal block length suggested in Politis and White (2004) is designed for long-run variance not skewness, we decide to stick to Fama and French's (2018) estimator.

1.4.2.3 Neuberger and Payne's (2021) Closed-Form Estimator

Assuming that an asset's price is a strongly stationary martingale, Neuberger and Payne (2021) derive a closed-form estimator for the unconditional skewness of the log return by approximating the moments of short-horizon log returns.[9] They show that this unconditional skewness can be expressed as the sum of the unconditional short-horizon skewness and the scaled covariance between past returns and current short-horizon variance:

$$\mathrm{Skew}[\ln R_{i,h}] = \left(\mathrm{Skew}[\ln R_{i,}\;] + 3\frac{\mathrm{Cov}[y^{(1)}, x^{(2E)}(\ln R_{i,}\;)]}{\mathrm{Var}[\ln R_{i,}\;]^{3/2}}\right) / \sqrt{\frac{h}{\Delta}}, \tag{1.4.6}$$

where:

$$\begin{aligned} &\mathrm{Skew}[\ln R_{i,}\;] = 6((R_{i,}\; + 1)\ln R_{i,}\quad 2(R_{i,}\quad 1))/\,\mathrm{Var}[\ln R_{i,}\;]^{3/2}, \\ &\mathrm{Var}[\ln R_{i,}\;] = 2(R_{i,}\quad 1 \quad \ln R_{i,}\;), \quad y_i^{(1)} = \sum_{h^0=1}^{h/\quad 1} (R_{i\quad 1,h^0}\quad 1)/\sqrt{\frac{h}{\Delta}}, \text{ and} \\ &x^{(2E)}(\ln R_{i,}\;) = 2(R_{i,}\;\ln R_{i,}\quad R_{i,}\; + 1). \end{aligned} \tag{1.4.7}$$

1.4.2.4 Farago and Hjalmarsson's (2023) Closed-Form Estimator

Farago and Hjalmarsson (2023) derive the closed-form equivalent of Fama and French's (2018) skewness estimator. In particular, assuming that the short-horizon return, $R_{i,}$, is i.i.d., they show that the unconditional skewness of the long-horizon discrete return,

[9] Their earlier working paper shows that they actually approximate the moments of the discrete not log return, but that accidentally and unexpectedly, they obtain an estimator which more closely reflects the skewness of log rather than discrete returns.

$R_{i,h} = \prod_{i=1}^{h/} R_{i,}$, can be analytically computed from short-horizon return moments:

$$\text{Skew}[R_{i,h}] = \frac{\theta_3^{h/} \quad 3\theta_2^{h/} \quad + 2}{(\theta_2^{h/} \quad 1)^{3/2}}, \tag{1.4.8}$$

where $\theta_2 = \text{Var}[R_{i,} \ ]/\,\text{E}[R_{i,} \ ]^2 + 1$ and $\theta_3 = \quad 2 + 3\theta_2 + \text{Skew}[R_{i,} \ ](\theta_2 \quad 1)^{3/2}$.

1.4.3 The Heston World

We next study the performance of our Heston (1993) estimator and its competitors in a world in which asset prices obey the Heston (1993) process. Toward that goal, we first look into the true skewness of the discrete return over various horizons in such a world. We next evaluate how well our novel GMM approach from Section 1.3 estimates the process parameters in that world. Finally, we contrast the mean-squared errors (MSEs) generated by our estimator and the others in capturing unconditional and conditional skewness (i.e., in-sample and out-of-sample analyses) over various horizons.

Having scanned a large number of papers on estimating or calibrating the Heston (1993) model (e.g., Aït-Sahalia and Kimmel (2007), Mariani et al. (2008), Mrázek et al. (2016), etc.), we select $\mu = 0.1$, $\kappa = 3$, $\alpha = 0.09$, $\xi = 0.3$, and $\rho = \quad 0.5$, a moderate set of values, as basecase parameters for 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, $\alpha = 0.02$ or 0.25, $\xi = 0.1$ or 0.5 and $\rho = \quad 0.9$ or 0, in each case, however, varying only one of the parameters from its basecase value. These values almost map out the entire reasonable range of each parameter documented in the literature.

1.4.3.1 The True Skewness of the Discrete Return

We first investigate the true skewness of the discrete return in the Heston (1993) world. To that end, Figure 1.1 plots the unconditional version of that skewness calculated over horizons up to five years under our basecase parameter choice and the alterations. In line with Bessembinder's (2018) argument that the compounding up of short-horizon returns induces right skewness in long-horizon returns, the figure demonstrates that the unconditional skewness rises with the return horizon, at least when studying sufficiently long horizons exceeding six months.

Notwithstanding, the actual shape of the unconditional skewness-return horizon relation is strongly modulated by the process parameters. To wit, while unconditional skewness rises with long-run variance α and the asset return-volatility correlation ρ, it falls with the volatility of variance ξ. The positive relation with long-run variance arises since a greater variance boosts the compounding effect, while the positive relation with the asset return-volatility correlation arises since a higher correlation diminishes the leverage effect (i.e., the tendency of low returns to be followed by high variance), lowering skewness. Interestingly, mean reversion κ exerts an ambiguous effect on unconditional skewness. In agreement, the figure further reveals that when the compounding effect is sufficiently weak (e.g., $\alpha = 0.02$) and/or the leverage effect is sufficiently strong (e.g., $\rho = \quad 0.9$), unconditional skewness can initially drop with the return horizon over short horizons up to 3-4 months before eventually starting to rise over longer horizons.

1.4.3.2 Heston Process Parameter Estimation

As the performance of our Heston (1993) estimator critically hinges on how well our novel GMM approach is able to estimate the process parameters, Table 1.1 gives the mean parameter estimates and MSEs (in parentheses) obtained from applying that approach to the 10,000 simulated Heston (1993) sample paths under our basecase parameter choice

(Panel A) as well as its variations (Panel B). The table confirms that our estimator usually yields mean estimates close to the corresponding true values. Moreover, it reveals that the MSEs are generally small. The only exception occurs for the mean reversion parameter κ whose estimates tend to be off-target. The poor performance of our estimator for κ aligns with Atiya and Wall's (2009) conclusion that the flatness of the likelihood surface in κ renders it hard for any estimator to precisely capture that parameter. Nevertheless, our magnitude of bias roughly match the level documented in Yu (2012).

1.4.3.3 Unbiasedness and Efficiency of the Estimators

We next contrast the unbiasedness and efficiency of our Heston (1993) estimator and its competitors in estimating the unconditional and conditional skewness of discrete returns (i.e., in-sample and out-of-sample analyses) over various horizons in a Heston (1993) world. Starting with unconditional skewness, Table 1.2 (i.e., in-sample analyses) offers the true skewness ("True"), the mean estimates of our Heston (1993; "HE"), the sample skewness ("SS"), Fama and French's (2018; "FF"), Neuberger and Payne's (2021; "NP"), and Farago and Hjalmarsson's (2023; "FH") estimators, and the corresponding MSEs (in parentheses) over a daily, weekly, monthly, quarterly, annual, three-year, and five-year horizon. We compute both the means and MSEs from 10,000 sample paths simulated under the basecase parameters.

The table suggests that our estimator almost always yields the closest-to-unbiased estimates with the smallest MSEs, with its outperformance often being economically meaningful. Looking into the annual horizon, our estimator, for example, yields an absolute bias or MSE of no more than 25% of those of the others. While the sample skewness ("SS") also yields close-to-unbiased estimates over short horizons, it becomes increasingly downward biased over longer horizons (owing to the fact that the sample contains increasingly fewer independent observations), and its MSEs are at least seven times those of our estimator. Next, Fama and French's (2018; "FF") and Farago and Hjalmarsson's (2023; "FH") estimators tend to overshoot true unconditional skewness over each horizon since their i.i.d. return assumption leads them to miss out on the negative influence of the leverage effect on skewness. Finally, Neuberger and Payne's (2021; "NP") estimator tends to greatly undershoot skewness, especially over long horizons, since their estimator approximates away the positive effect of compounding up short-horizon returns on skewness.

In Table 1.3 (i.e., in-sample analyses), we give comparative statics for our unconditional-skewness-estimation evaluations in Table 1.2, separately replacing the mean reversion speed κ (Panel A), the long-run variance α (Panel B), the volatility of variance ξ (Panel C) or the asset price-variance correlation ρ (Panel D) in the data-generating process with a low (columns (1) to (5)) or high ((6) to (10)) value. The table reveals that our estimator consistently outperforms the others in terms of unbiasedness and MSE except when the asset price-volatility correlation is zero ($\rho = 0$). The worse performance of our estimator in the case of $\rho = 0$ is hardly surprising since, in this scenario, both Fama and French's (2018; "FF") and Farago and Hjalmarsson's (2023; "FH") estimators impose the correct restriction of no return dependence, helping them to gain statistical power. Yet, even in this case, our estimator shows an only marginally worse performance than the former two.

We next switch to evaluating how well the five estimators capture the conditional skewness of the discrete return over the above horizons (i.e., out-of-sample analyses). Doing so is plausibly more relevant than studying their ability to capture unconditional skewness since real-world market participants are likely more interested in forecasting skewness over some future period rather than measuring it over some historical period. Toward that goal, we now calculate the true skewness as the average of the conditional skewness based on the spot variance at the end of a sample path taken over the 10,000 paths. Also, while we again estimate the process parameters from applying our GMM approach to an entire sample path, we use only the final two years of that path to compute the optimal estimate of the conditional MGF, $M_i \ (u,h;\hat{\theta})$, according to the iterative procedure outlined in Proposition 1, to speed up our computations. We finally stress that since the competing estimators are only able to yield unconditional skewness estimates, we assess the ability of those estimates to capture conditional skewness in their case.

Relying on the same design as Table 1.2, Table 1.4 (i.e., out-of-sample analyses) gives the average true skewness ("True"), the mean estimates of the five estimators, and their corresponding MSEs (in parentheses) over the same periods as above. The table shows that conditional skewness is, on average, slightly lower than unconditional skewness (contrast the "True" columns in Tables 1.2 and 1.4). More importantly, it further suggests that our estimator consistently outperforms the others in forecasting skewness, producing the closest-to-unbiased estimates as well as the smallest MSEs. Looking at the annual horizon again, our estimator yields an absolute bias of no more than 13% and an MSE of no more than 28% of the others. While the margin of outperformance of our estimator has slightly deteriorated along the MSE dimension relative to the unconditional-skewness case (i.e., in-sample analyses), the margin of outperformance has greatly improved along the unbiasedness dimension. Switching to comparative statics, Table 1.5 demonstrates that our estimator also consistently outperforms its competitors under the alternative process-parameters also used in Table 1.3.

Figures 1.2 and 1.3 sum up the performance of the estimators in measuring unconditional and conditional skewness (i.e., in-sample and out-of-sample analyses), respectively, under the basecase and the alternative parameter value sets, plotting the MSEs of the estimators over the monthly horizon. The figures vividly show that, except in the few aforementioned cases, our estimator greatly outperforms the others.

1.4.4 The Double-Heston and the Heston-Jump Worlds

While Section 1.4.3 shows that our Heston (1993) estimator comfortably beats its competitors in a Heston (1993) world, we could ratiocinate that the exact alignment between the stochastic process assumed in our data-generating process and in the estimator grants an unfair advantage to our estimator. To address that concern, we now apply our Heston (1993) estimator and its competitors to data-generating processes more complicated than the Heston (1993) process. Our goal is to show that even in situations in which the Heston (1993) process only goes some way toward accurately modelling asset prices but does not fully capture all stylized facts of them, we can meaningfully use our Heston (1993) to measure and forecast skewness. To facilitate that analysis, we now use the Christoffersen et al. (2009) double-Heston and the Heston-jump process as alternative data-generating processes in our skewness estimate evaluations (see Section 1.2.2.2 for the exact mathematical definitions of the alternative stochastic processes).

Using a panel design identical to the design of Table 1.2, Table 1.6 repeats our unbiasedness and efficiency tests for the unconditional (Panel A; i.e., in-sample analyses) and conditional (Panel B; i.e., out-of-sample analyses) skewness estimates produced by the estimators under the assumption that asset prices follow the double-Heston process. In contrast to the Heston (1993) process, the double-Heston process offers more flexibility in modelling the term structure of volatility. In addition, choosing a high κ_1 and a low κ_2 value in Eq. (1.2.27), the process can better capture the often different effects of short-term and long-term returns on variance. To differentiate the double-Heston process used in our simulation as much as possible from a Heston (1993) process, we however allow for variations in all parameters, and not only the mean reversion parameters. In particular, we choose $\mu = 0.10$, $\kappa_1 = 1.00$, $\alpha_1 = 0.01$, $\xi_1 = 0.10$, $\rho_1 = \quad 0.90$, $\kappa_2 = 5.00$, $\alpha_2 = 0.09$, $\xi_2 = 0.50$, and $\rho_2 = \quad 0.60$.

The table suggests that our double-Heston process yields a slightly lower true unconditional skewness relative to our basecase Heston (1993) process over short return horizons but a close to identical over longer horizons. More crucially, it also demonstrates that our Heston (1993) estimator continues to strongly outperform its competitors in estimating and forecasting skewness along the unbiasedness and MSE dimensions. Looking at the annual horizon, our estimator, for example, generates a mean unconditional and conditional estimate only 0.035 and 0.017 away from the corresponding true value and an MSE of 0.031 and 0.036, all respectively. In contrast, the other estimators produce a mean estimate at least 0.128 and 0.116 away from that same true value and an MSE of at least 0.197 and 0.199, again all respectively.

To also investigate how our Heston (1993) estimator fares under mixed jump-diffusion data-generating processes, we next repeat our evaluation exercise in Table 1.6 under the assumption that asset prices obey the Heston-jump process. To do so, we choose $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, $\rho =$ .50, $\lambda = 1.70$, $\overline{\gamma} =$ 0.0301, and $\delta = 0.0078$, in line with the jump parameter values from Chernov et al. (2003). Using the same design as Table 1.6, Table 1.7 demonstrates that our Heston-jump process produces a slightly lower true unconditional skewness at super-short horizons (e.g., daily and weekly) relative to our basecase Heston process but else a slightly higher skewness. These changes align with the intuition that, as jump size and intensity rise, the downward (upward) effect of jumps on unconditional skewness over super-short (longer) horizons amplify. More crucially, the table again confirms that our Heston (1993) estimator consistently outperforms its competitors along both the unbiasedness and MSE dimensions.[10]

Taken together, this section shows that realistic deviations between the asset price process assumed by our estimator and the true process only marginally affect the performance of our estimator, with our estimator still comfortably beating its competitors.

[10] In Appendix 1.D, we reveal how we can improve our super-short horizon mean estimates by generalizing our Heston (1993) estimator to a Heston-jump estimator.

1.5 Empirical Application

We finally offer some evidence on how well our Heston (1993) estimator and its competitors are able to fit the conditional skewness of the S&P 500 in real-world data. To do so, let $R^s_{i,h}$ be the "standardized" S&P 500 return over the h horizon (i.e., the index return net of its expectation scaled by its volatility). Denoting the equivalent martingale measure by Q, we have:

$$\mathrm{E}^{\mathsf{Q}}[(R^s_{i,h})^3] = \mathrm{E}\left[m_h (R^s_{i,h})^3\right] = \frac{1}{R_{f,h}} \mathrm{E}[(R^s_{i,h})^3] + \mathrm{Cov}(m_h, (R^s_{i,h})^3), \tag{1.5.1}$$

where m_h is the stochastic discount factor over horizon h and $R_{f,h}$ the risk-free rate of return over that horizon. Assuming the risk-free rate $R_{f,h}$ and the risk premium $\mathrm{Cov}(m_h, (R^s_{i,h})^3)$ are constant over time, risk-neutral skewness $\mathrm{E}^{\mathsf{Q}}[(R^s_{i,h})^3]$ is linear in physical skewness $\mathrm{E}[(R^s_{i,h})^3]$, implying that a perfect estimate of conditional (i.e., out-of-sample) physical skewness should also perfectly forecast conditional risk-neutral skewness. Under mild violations of those assumptions, the same perfect estimate should still explain conditional risk-neutral skewness with a high accuracy.

We rely on Carr and Madan's (2001) formula to infer $\mathrm{E}^{\mathsf{Q}}[(R^s_{i,h})^3]$ from options data:

$$g(S_T) = g(S_t) + g'(S_t)(S_T - S_t) + \int_0^{\infty} g''(K_s) O_t(K_s) dK_s, \tag{1.5.2}$$

where $g(\cdot)$ is a real-valued twice-differentiable payoff function, S_t is the S&P 500 value at time t, and $O_t(K_s)$ is the price of an out-of-the-money (OTM) option written on the S&P 500 with strike price K_s at that time. Let $g_u(S_T) := (S_T/S_t)^u = R^u_{t,T-t}$, where $u \in \{1, 2, 3\}$. Given $\mathrm{E}^{\mathsf{Q}}_t[S_T] = F_t$, the forward price of the S&P 500 at time T, the expectations of $g_u(S_T)$

under Q are:

$$
\begin{aligned}
\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}] &= \frac{F_t}{S_t}, \\
\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}^2] &= \quad 1 + \frac{2F_t}{S_t} + \frac{2}{S_t^2}\int_0^{1} O_t(K_s)dK_s, \\
\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}^3] &= \quad 2 + \frac{3F_t}{S_t} + \frac{6}{S_t^3}\int_0^{1} K_s O_t(K_s)dK_s.
\end{aligned}
\tag{1.5.3}
$$

Consequently, the risk-neutral skewness equals

$$
\text{RN-Skew}[R_{t,T\ \ t}] = \frac{\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}^3] \quad 3\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}]\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}^2] + 2\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}]^3}{(\mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}^2] \quad \mathrm{E}_t^{\mathsf{Q}}[R_{t,T\ \ t}]^2)^{3/2}}. \tag{1.5.4}
$$

To facilitate our analysis, we use standard (i.e., settled on the third Friday of each month) SPX OTM options over the period from start-1996 to end-2022. After imposing conventional filters used in the literature,[11] we only keep those OTM options with one month and three months left to maturity to ensure a high liquidity. The upshot is that we exclusively look into the skewness of the one and three-month discrete return in this section. We finally employ the trapezoidal rule to approximate the two integrals shown in Eq. (1.5.3).

In Table 1.8, we present the results from time-series regressions of the conditional risk-neutral skewness of the discrete return over the one (Panel A) or three (Panel B) months to maturity on our Heston (1993) estimates and its competitors. To align the dependent and independent variables, we apply our estimator and its competitors to daily data over the ten year period ending at the start of the return period used to compute the dependent variable. Remarkably, the table suggests that our estimator is far superior to the others in capturing variations in risk-neutral skewness, with it consistently yielding the most significant slope coefficients (t-statistic > 6) and highest R-squareds (around

[11] Our filters are: i) an option's price must lie within its arbitrage-free boundaries; ii) its bid-ask spread must be positive; iii) its volume and open interest must be positive; and iv) its price must exceed 0.125.

40%). While the sample skewness is also able to capture some fraction of that variation, its R-squareds are only about half of ours. Conversely, Fama and French's (2018) and Farago and Hjalmarsson's (2023) estimators hardly capture variation in risk-neutral skewness, with them consistently producing close-to-zero R-squareds.

1.6 Conclusion

In this paper, we propose a novel parametric estimator of the skewness of discrete returns based on the assumption that an asset's price can be described using a stochastic process from the ASV model class. Relying on the Heston (1993) process as example, we run a simulation exercise to compare the unbiasedness and efficiency of our estimator with those of other existing estimators, such as the sample skewness, Fama and French's (2018) bootstrap estimator, and Neuberger and Payne's (2021) and Farago and Hjalmarsson's (2023) closed-form estimators. Our evidence suggests that our estimator strongly outperforms the others under the assumption that the asset price obeys the Heston (1993) process or even some slightly more complicated processes, with the estimator yielding the smallest absolute bias and MSE in the vast majority of cases. We finally apply our Heston (1993) estimator and the others to capture the risk-neutral conditional skewness of the one- and three-month S&P 500 return as embedded in option prices, showing that our estimates best mimic the time-series variations in that conditional skewness.

Table 1.1: Heston Process Parameter Estimation

This table presents the true values of the Heston (1993) process used to simulate sample paths of the daily discrete returns as well as the mean estimates ("Est.") and mean squared errors (MSEs; in parentheses) obtained from separately applying our GMM estimator introduced in Section 1.3 to 10,000 sample paths obtained from that process. Each of the 10,000 sample paths features ten years of daily discrete returns.

Drift (μ)		Mean Reversion Variance Speed (κ)		Long-Run Variance (α)		Volatility of Variance (ξ)		Price-Variance Correlation (ρ)	
True	Est.	True	Est.	True	Est.	True	Est.	True	Est.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Basecase									
0.100	0.108	3.000	3.727	0.090	0.077	0.300	0.306	0.500	0.506
	(0.011)		(4.994)		(0.000)		(0.019)		(0.049)
Panel B: Comparative Statics									
0.100	0.099	1.000	2.002	0.090	0.072	0.300	0.332	0.500	0.474
	(0.011)		(3.449)		(0.001)		(0.027)		(0.070)
0.100	0.109	5.000	5.501	0.090	0.079	0.300	0.307	0.500	0.505
	(0.010)		(6.499)		(0.000)		(0.021)		(0.049)
0.100	0.103	3.000	3.639	0.020	0.016	0.300	0.267	0.500	0.504
	(0.002)		(3.901)		(0.000)		(0.009)		(0.047)
0.100	0.096	3.000	4.094	0.250	0.221	0.300	0.333	0.500	0.521
	(0.029)		(8.174)		(0.001)		(0.036)		(0.064)
0.100	0.100	3.000	4.703	0.090	0.081	0.100	0.173	0.500	0.508
	(0.010)		(15.018)		(0.000)		(0.032)		(0.124)
0.100	0.108	3.000	3.658	0.090	0.072	0.500	0.474	0.500	0.499
	(0.011)		(3.780)		(0.000)		(0.026)		(0.045)
0.100	0.118	3.000	3.549	0.090	0.077	0.300	0.317	0.900	0.788
	(0.010)		(2.382)		(0.000)		(0.008)		(0.044)
0.100	0.092	3.000	4.313	0.090	0.077	0.300	0.338	0.000	0.060
	(0.010)		(9.086)		(0.000)		(0.049)		(0.063)

Table 1.2: The Unconditional Performance of the Heston Skewness Estimator: Basecase

This table presents the true unconditional skewness of the discrete return over various horizons (column (1)) as well as the mean estimates of that skewness plus their mean-squared errors (MSEs; in parentheses) obtained from separately applying our Heston (1993) estimator ("HE"; (2)) or its competitors ((3) to (6)) to 10,000 sample paths obtained from our basecase Heston (1993) process. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The basecase parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, and $\rho =$ 0.50.

Horizon	True	Heston Estimator (HE)	Sample Skewness (SS)	Fama-French (FF)	Neuberger-Payne (NP)	Farago-Hjalmarsson (FH)
	(1)	(2)	(3)	(4)	(5)	(6)
Daily	0.019	0.016	0.022	0.023	0.022	0.022
		(0.000)	(0.004)	(0.005)	(0.009)	(0.004)
Weekly	0.044	0.037	0.050	0.111	0.069	0.111
		(0.002)	(0.012)	(0.006)	(0.018)	(0.005)
Monthly	0.097	0.087	0.106	0.253	0.177	0.253
		(0.006)	(0.042)	(0.025)	(0.082)	(0.025)
Quarterly	0.200	0.187	0.197	0.451	0.274	0.451
		(0.014)	(0.096)	(0.064)	(0.238)	(0.063)
Annual	0.594	0.556	0.438	0.944	0.316	0.946
		(0.031)	(0.217)	(0.128)	(0.855)	(0.127)
3-Year	1.406	1.278	0.588	1.832	0.223	1.840
		(0.074)	(0.906)	(0.219)	(2.694)	(0.207)
5-Year	2.120	1.900	0.567	2.672	0.156	2.691
		(0.149)	(2.707)	(0.482)	(5.226)	(0.384)

Table 1.3: The Unconditional Performance of the Heston Skewness Estimator: Comparative Statics

This table presents the true unconditional skewness of the discrete return over various horizons (columns (1) and (7)) as well as the mean estimates of that skewness plus their mean-squared errors (MSEs; in parentheses) obtained from separately applying our Heston (1993) estimator ("HE"; (2) and (8)) or its competitors ((3) to (6) and (9) to (12)) to 10,000 sample paths obtained from our comparative static Heston (1993) processes. In Panels A to D, we vary the variance mean reversion speed κ, the long-run variance α, the volatility of variance ξ, and the price-variance correlation ρ from its basecase values, respectively, setting them to a lower (columns (1) to (6)) or higher ((7) to (12)) value. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The basecase parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, and $\rho =$ 0.50. We report the alternative values in the panel headings.

Horizon	True	HE	SS	FF	NP	FH	True	HE	SS	FF	NP	FH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Variance Mean Reversion Speed												
	$\kappa = 1$						$\kappa = 5$					
Daily	0.038	0.027	0.041	0.041	0.033	0.041	0.015	0.015	0.017	0.017	0.019	0.017
		(0.001)	(0.008)	(0.010)	(0.015)	(0.008)		(0.000)	(0.003)	(0.004)	(0.008)	(0.003)
Weekly	0.084	0.060	0.091	0.118	0.061	0.118	0.036	0.036	0.039	0.109	0.071	0.109
		(0.005)	(0.024)	(0.004)	(0.032)	(0.003)		(0.001)	(0.009)	(0.007)	(0.016)	(0.006)
Monthly	0.173	0.125	0.173	0.254	0.171	0.254	0.090	0.089	0.095	0.252	0.170	0.252
		(0.018)	(0.090)	(0.009)	(0.136)	(0.009)		(0.005)	(0.035)	(0.027)	(0.073)	(0.026)
Quarterly	0.298	0.225	0.263	0.448	0.295	0.448	0.214	0.210	0.203	0.449	0.242	0.450
		(0.048)	(0.201)	(0.028)	(0.388)	(0.027)		(0.009)	(0.079)	(0.057)	(0.216)	(0.056)
Annual	0.608	0.519	0.391	0.938	0.462	0.939	0.677	0.634	0.477	0.944	0.228	0.945
		(0.124)	(0.326)	(0.136)	(1.236)	(0.134)		(0.017)	(0.224)	(0.075)	(0.832)	(0.073)
3-Year	1.197	1.110	0.522	1.830	0.450	1.837	1.538	1.408	0.600	1.833	0.145	1.838
		(0.264)	(0.710)	(0.568)	(2.885)	(0.559)		(0.042)	(1.116)	(0.115)	(2.849)	(0.098)
5-Year	1.753	1.641	0.562	2.688	0.334	2.714	2.297	2.074	0.557	2.673	0.101	2.686
		(0.486)	(1.771)	(1.553)	(4.548)	(1.444)		(0.094)	(3.319)	(0.286)	(5.774)	(0.174)
Panel B: Long-Run Variance												
	$\alpha = 0.02$						$\alpha = 0.25$					
Daily	0.053	0.055	0.045	0.045	0.050	0.045	0.072	0.067	0.073	0.073	0.011	0.073
		(0.002)	(0.013)	(0.017)	(0.033)	(0.013)		(0.000)	(0.003)	(0.004)	(0.011)	(0.003)
Weekly	0.116	0.120	0.099	0.027	0.145	0.027	0.161	0.150	0.162	0.202	0.044	0.202
		(0.007)	(0.041)	(0.024)	(0.020)	(0.023)		(0.001)	(0.008)	(0.003)	(0.046)	(0.002)
Monthly	0.218	0.221	0.176	0.105	0.369	0.106	0.335	0.315	0.326	0.433	0.107	0.433
		(0.026)	(0.129)	(0.106)	(0.057)	(0.106)		(0.004)	(0.035)	(0.011)	(0.200)	(0.010)
Quarterly	0.299	0.293	0.214	0.202	0.568	0.202	0.604	0.573	0.544	0.774	0.165	0.775
		(0.056)	(0.227)	(0.252)	(0.141)	(0.251)		(0.009)	(0.103)	(0.032)	(0.597)	(0.030)
Annual	0.176	0.178	0.053	0.421	0.641	0.421	1.410	1.333	0.916	1.743	0.191	1.748
		(0.085)	(0.255)	(0.360)	(0.350)	(0.359)		(0.035)	(0.494)	(0.132)	(2.573)	(0.121)
3-Year	0.257	0.192	0.201	0.752	0.450	0.753	3.464	3.138	1.013	4.276	0.134	4.358
		(0.080)	(0.219)	(0.255)	(0.659)	(0.254)		(0.308)	(6.320)	(2.141)	(12.961)	(0.910)
5-Year	0.546	0.444	0.266	0.997	0.316	0.998	6.453	5.595	0.951	7.782	0.095	8.718
		(0.078)	(0.399)	(0.222)	(0.910)	(0.220)		(1.801)	(30.651)	(18.035)	(42.906)	(6.162)

(continued on next page)

Horizon	True	HE	SS	FF	NP	FH	True	HE	SS	FF	NP	FH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel C: Volatility of Variance												
	$\xi = 0.10$						$\xi = 0.50$					
Daily	0.042	0.040	0.043	0.043	0.049	0.043	0.004	-0.002	0.008	0.008	0.005	0.008
		(0.000)	(0.003)	(0.003)	(0.006)	(0.003)		(0.001)	(0.008)	(0.011)	(0.015)	(0.008)
Weekly	0.094	0.089	0.095	0.121	0.005	0.121	0.011	-0.002	0.019	0.104	0.134	0.105
		(0.001)	(0.007)	(0.002)	(0.013)	(0.001)		(0.005)	(0.026)	(0.011)	(0.033)	(0.011)
Monthly	0.197	0.189	0.193	0.257	0.051	0.258	0.032	0.009	0.042	0.249	0.300	0.249
		(0.002)	(0.026)	(0.004)	(0.065)	(0.004)		(0.017)	(0.085)	(0.049)	(0.129)	(0.048)
Quarterly	0.357	0.346	0.327	0.453	0.088	0.454	0.096	0.067	0.092	0.447	0.454	0.448
		(0.004)	(0.074)	(0.010)	(0.202)	(0.010)		(0.038)	(0.156)	(0.126)	(0.338)	(0.126)
Annual	0.810	0.777	0.555	0.948	0.107	0.949	0.445	0.390	0.324	0.942	0.510	0.943
		(0.009)	(0.261)	(0.022)	(0.846)	(0.020)		(0.073)	(0.230)	(0.258)	(0.981)	(0.256)
3-Year	1.668	1.549	0.617	1.838	0.077	1.843	1.225	1.073	0.542	1.833	0.354	1.838
		(0.036)	(1.339)	(0.051)	(3.054)	(0.034)		(0.130)	(0.717)	(0.440)	(2.586)	(0.425)
5-Year	2.459	2.240	0.568	2.679	0.054	2.691	1.889	1.638	0.569	2.679	0.253	2.694
		(0.098)	(3.862)	(0.188)	(6.326)	(0.063)		(0.230)	(2.081)	(0.907)	(4.695)	(0.806)
Panel D: Price-Variance Correlation												
	$\rho = 0.90$						$\rho = 0.00$					
Daily	0.019	0.019	0.015	0.015	0.015	0.015	0.066	0.061	0.066	0.065	0.063	0.066
		(0.000)	(0.004)	(0.005)	(0.009)	(0.004)		(0.001)	(0.004)	(0.005)	(0.009)	(0.004)
Weekly	0.040	0.039	0.028	0.094	0.150	0.095	0.148	0.135	0.145	0.131	0.028	0.131
		(0.001)	(0.011)	(0.019)	(0.017)	(0.019)		(0.003)	(0.013)	(0.002)	(0.020)	(0.001)
Monthly	0.066	0.062	0.041	0.244	0.330	0.245	0.302	0.273	0.284	0.263	0.013	0.263
		(0.005)	(0.035)	(0.097)	(0.077)	(0.097)		(0.010)	(0.051)	(0.003)	(0.091)	(0.002)
Quarterly	0.049	0.035	0.010	0.445	0.497	0.446	0.519	0.463	0.447	0.457	0.007	0.457
		(0.013)	(0.072)	(0.245)	(0.215)	(0.245)		(0.022)	(0.136)	(0.005)	(0.276)	(0.005)
Annual	0.279	0.288	0.271	0.941	0.563	0.943	1.037	0.914	0.623	0.949	0.004	0.950
		(0.031)	(0.154)	(0.444)	(0.742)	(0.444)		(0.051)	(0.406)	(0.013)	(1.095)	(0.011)
3-Year	1.063	1.011	0.550	1.832	0.394	1.838	1.946	1.704	0.619	1.837	0.003	1.845
		(0.054)	(0.501)	(0.628)	(2.165)	(0.619)		(0.133)	(1.998)	(0.050)	(3.818)	(0.028)
5-Year	1.704	1.596	0.572	2.665	0.275	2.688	2.819	2.431	0.567	2.673	0.001	2.698
		(0.091)	(1.610)	(1.095)	(3.959)	(1.027)		(0.303)	(5.369)	(0.206)	(7.989)	(0.072)

Table 1.4: The Conditional Performance of the Heston Skewness Estimator: Basecase

This table presents the mean true conditional skewness of the discrete return over various horizons (column (1)) as well as the mean estimates of that skewness plus their mean-squared errors (MSEs; in parentheses) obtained from separately applying our Heston (1993) estimator ("HE"; (2)) or its competitors ((3) to (6)) to 10,000 sample paths obtained from our basecase Heston (1993) process. We measure conditional skewness (true or estimated) at the end of each sample path. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The basecase parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, and $\rho =$ 0.50.

Horizon	True	Heston Estimator (HE)	Sample Skewness (SS)	Fama-French (FF)	Neuberger-Payne (NP)	Farago-Hjalmarsson (FH)
	(1)	(2)	(3)	(4)	(5)	(6)
Daily	0.005	0.012	0.022	0.023	0.022	0.022
		(0.001)	(0.005)	(0.006)	(0.010)	(0.005)
Weekly	0.015	0.029	0.050	0.111	0.069	0.111
		(0.005)	(0.015)	(0.013)	(0.015)	(0.012)
Monthly	0.051	0.075	0.106	0.253	0.177	0.253
		(0.017)	(0.052)	(0.048)	(0.066)	(0.047)
Quarterly	0.154	0.179	0.197	0.451	0.274	0.451
		(0.029)	(0.107)	(0.097)	(0.205)	(0.097)
Annual	0.576	0.559	0.438	0.944	0.316	0.946
		(0.040)	(0.217)	(0.145)	(0.828)	(0.143)
3-Year	1.400	1.286	0.588	1.832	0.223	1.840
		(0.074)	(0.901)	(0.225)	(2.680)	(0.213)
5-Year	2.117	1.910	0.567	2.672	0.156	2.691
		(0.145)	(2.702)	(0.485)	(5.214)	(0.387)

Table 1.5: The Conditional Performance of the Heston Skewness Estimator: Comparative Statics

This table presents the mean true conditional skewness of the discrete return over various horizons (columns (1) and (7)) as well as the mean estimates of that skewness plus their mean-squared errors (MSEs; in parentheses) obtained from separately applying our Heston (1993) estimator ("HE"; (2) and (8)) or its competitors ((3) to (6) and (9) to (12)) to 10,000 sample paths obtained from our comparative static Heston (1993) processes. We measure conditional skewness (true or estimated) at the end of each sample path. In Panels A to D, we vary the variance mean reversion speed κ, the long-run variance α, the volatility of variance ξ, and the price-variance correlation ρ from its basecase values, respectively, setting them to a lower (columns (1) to (6)) or higher ((7) to (12)) value. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The basecase parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, and $\rho =$ 0.50. We report the alternative values in the panel headings.

Horizon	True	HE	SS	FF	NP	FH	True	HE	SS	FF	NP	FH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Panel A: Variance Mean Reversion Speed						
	$\kappa = 1$						$\kappa = 5$					
Daily	0.006	0.007	0.041	0.041	0.033	0.041	0.007	0.014	0.017	0.017	0.019	0.017
		(0.005)	(0.012)	(0.014)	(0.020)	(0.012)		(0.001)	(0.004)	(0.005)	(0.008)	(0.004)
Weekly	0.010	0.017	0.091	0.118	0.061	0.118	0.020	0.033	0.039	0.109	0.071	0.109
		(0.022)	(0.044)	(0.028)	(0.025)	(0.027)		(0.003)	(0.011)	(0.010)	(0.014)	(0.010)
Monthly	0.000	0.051	0.173	0.254	0.171	0.254	0.068	0.088	0.095	0.252	0.170	0.252
		(0.070)	(0.149)	(0.094)	(0.078)	(0.093)		(0.009)	(0.038)	(0.038)	(0.065)	(0.037)
Quarterly	0.062	0.136	0.263	0.448	0.295	0.448	0.197	0.212	0.203	0.449	0.242	0.450
		(0.138)	(0.293)	(0.199)	(0.218)	(0.199)		(0.013)	(0.082)	(0.068)	(0.204)	(0.067)
Annual	0.387	0.452	0.391	0.938	0.462	0.939	0.673	0.640	0.477	0.944	0.228	0.945
		(0.222)	(0.336)	(0.372)	(0.877)	(0.371)		(0.018)	(0.224)	(0.078)	(0.825)	(0.076)
3-Year	1.086	1.079	0.522	1.830	0.450	1.837	1.537	1.414	0.600	1.833	0.145	1.838
		(0.327)	(0.627)	(0.727)	(2.588)	(0.719)		(0.041)	(1.115)	(0.116)	(2.846)	(0.099)
5-Year	1.680	1.621	0.562	2.688	0.334	2.714	2.297	2.081	0.557	2.673	0.101	2.686
		(0.529)	(1.662)	(1.673)	(4.311)	(1.563)		(0.092)	(3.318)	(0.286)	(5.772)	(0.175)
						Panel B: Long-Run Variance						
	$\alpha = 0.02$						$\alpha = 0.25$					
Daily	0.115	0.081	0.045	0.045	0.050	0.045	0.065	0.066	0.073	0.073	0.011	0.073
		(0.017)	(0.028)	(0.032)	(0.061)	(0.028)		(0.001)	(0.003)	(0.004)	(0.010)	(0.003)
Weekly	0.230	0.169	0.099	0.027	0.145	0.027	0.146	0.149	0.162	0.202	0.044	0.202
		(0.053)	(0.084)	(0.096)	(0.053)	(0.095)		(0.003)	(0.009)	(0.005)	(0.041)	(0.005)
Monthly	0.366	0.284	0.176	0.105	0.369	0.106	0.310	0.314	0.326	0.433	0.107	0.433
		(0.102)	(0.205)	(0.261)	(0.074)	(0.261)		(0.009)	(0.038)	(0.020)	(0.182)	(0.019)
Quarterly	0.416	0.340	0.214	0.202	0.568	0.202	0.575	0.576	0.544	0.774	0.165	0.775
		(0.121)	(0.295)	(0.412)	(0.122)	(0.411)		(0.018)	(0.106)	(0.048)	(0.561)	(0.046)
Annual	0.211	0.190	0.053	0.421	0.641	0.421	1.395	1.344	0.916	1.743	0.191	1.748
		(0.107)	(0.275)	(0.410)	(0.330)	(0.410)		(0.042)	(0.485)	(0.148)	(2.534)	(0.136)
3-Year	0.249	0.193	0.201	0.752	0.450	0.753	3.456	3.158	1.013	4.276	0.134	4.358
		(0.085)	(0.222)	(0.264)	(0.652)	(0.263)		(0.304)	(6.293)	(2.150)	(12.915)	(0.926)
5-Year	0.542	0.447	0.266	0.997	0.316	0.998	6.444	5.628	0.951	7.782	0.095	8.718
		(0.080)	(0.400)	(0.225)	(0.908)	(0.224)		(1.758)	(30.587)	(18.039)	(42.815)	(6.192)

(continued on next page)

Horizon	True	HE	SS	FF	NP	FH	True	HE	SS	FF	NP	FH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel C: Volatility of Variance												
	$\xi = 0.10$						$\xi = 0.50$					
Daily	0.041	0.039	0.043	0.043	0.049	0.043	0.041	0.020	0.008	0.008	0.005	0.008
		(0.000)	(0.003)	(0.003)	(0.006)	(0.003)		(0.007)	(0.014)	(0.016)	(0.020)	(0.014)
Weekly	0.092	0.089	0.095	0.121	0.005	0.121	0.081	0.040	0.019	0.104	0.134	0.105
		(0.001)	(0.007)	(0.002)	(0.013)	(0.001)		(0.027)	(0.049)	(0.050)	(0.029)	(0.049)
Monthly	0.193	0.189	0.193	0.257	0.051	0.258	0.105	0.045	0.042	0.249	0.300	0.249
		(0.003)	(0.027)	(0.005)	(0.063)	(0.005)		(0.068)	(0.134)	(0.154)	(0.086)	(0.153)
Quarterly	0.352	0.348	0.327	0.453	0.088	0.454	0.032	0.020	0.092	0.447	0.454	0.448
		(0.005)	(0.075)	(0.012)	(0.198)	(0.011)		(0.095)	(0.203)	(0.259)	(0.244)	(0.259)
Annual	0.808	0.782	0.555	0.948	0.107	0.949	0.399	0.380	0.324	0.942	0.510	0.943
		(0.010)	(0.260)	(0.023)	(0.843)	(0.021)		(0.100)	(0.235)	(0.316)	(0.909)	(0.315)
3-Year	1.667	1.556	0.617	1.838	0.077	1.843	1.212	1.080	0.542	1.833	0.354	1.838
		(0.035)	(1.338)	(0.051)	(3.052)	(0.034)		(0.137)	(0.708)	(0.457)	(2.553)	(0.443)
5-Year	2.459	2.248	0.568	2.679	0.054	2.691	1.881	1.649	0.569	2.679	0.253	2.694
		(0.094)	(3.861)	(0.188)	(6.324)	(0.064)		(0.230)	(2.071)	(0.917)	(4.671)	(0.817)
Panel D: Price-Variance Correlation												
	$\rho =$ 0.90						$\rho = 0.00$					
Daily	0.035	0.027	0.015	0.015	0.015	0.015	0.056	0.057	0.066	0.065	0.063	0.066
		(0.002)	(0.005)	(0.006)	(0.011)	(0.005)		(0.001)	(0.005)	(0.006)	(0.009)	(0.005)
Weekly	0.073	0.055	0.028	0.094	0.150	0.095	0.125	0.128	0.145	0.131	0.028	0.131
		(0.011)	(0.017)	(0.034)	(0.016)	(0.033)		(0.003)	(0.014)	(0.002)	(0.016)	(0.002)
Monthly	0.117	0.084	0.041	0.244	0.330	0.245	0.263	0.264	0.284	0.263	0.013	0.263
		(0.030)	(0.052)	(0.142)	(0.064)	(0.142)		(0.011)	(0.053)	(0.003)	(0.073)	(0.002)
Quarterly	0.095	0.051	0.010	0.445	0.497	0.446	0.475	0.458	0.447	0.457	0.007	0.457
		(0.043)	(0.091)	(0.306)	(0.189)	(0.306)		(0.025)	(0.134)	(0.006)	(0.237)	(0.005)
Annual	0.264	0.290	0.271	0.941	0.563	0.943	1.015	0.918	0.623	0.949	0.004	0.950
		(0.044)	(0.160)	(0.468)	(0.722)	(0.468)		(0.053)	(0.393)	(0.014)	(1.056)	(0.011)
3-Year	1.059	1.020	0.550	1.832	0.394	1.838	1.939	1.713	0.619	1.837	0.003	1.845
		(0.056)	(0.500)	(0.635)	(2.156)	(0.626)		(0.133)	(1.983)	(0.050)	(3.796)	(0.028)
5-Year	1.701	1.606	0.572	2.665	0.275	2.688	2.814	2.443	0.567	2.673	0.001	2.698
		(0.090)	(1.611)	(1.099)	(3.953)	(1.031)		(0.301)	(5.352)	(0.205)	(7.967)	(0.070)

Table 1.6: The Performance of the Heston Skewness Estimator Under a Double-Heston Process

This table presents the true unconditional (Panel A) and mean conditional (Panel B) skewness of the discrete return over various horizons (column (1)) as well as the mean estimates of those skewnesses plus their mean-squared errors (MSEs; in parentheses) obtained from separately applying our Heston (1993) estimator ("HE"; (2)) or its competitors ((3) to (6)) to 10,000 sample paths obtained from a double-Heston process. We measure conditional skewness (true or estimated) at the end of each sample path. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The parameter values are $\mu = 0.10$, $\kappa_1 = 1.00$, $\alpha_1 = 0.01$, $\xi_1 = 0.10$, $\rho_1 =$ 0.90, $\kappa_2 = 5.00$, $\alpha_2 = 0.09$, $\xi_2 = 0.50$, and $\rho_2 =$ 0.60.

Horizon	True	Heston Estimator (HE)	Sample Skewness (SS)	Fama-French (FF)	Neuberger-Payne (NP)	Farago-Hjalmarsson (FH)
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Unconditional Skewness			
Daily	0.009	0.010	0.006	0.005	0.016	0.006
		(0.001)	(0.005)	(0.006)	(0.010)	(0.005)
Weekly	0.018	0.018	0.011	0.104	0.147	0.104
		(0.003)	(0.014)	(0.017)	(0.023)	(0.016)
Monthly	0.012	0.012	0.001	0.260	0.310	0.260
		(0.009)	(0.046)	(0.075)	(0.098)	(0.075)
Quarterly	0.070	0.069	0.080	0.471	0.432	0.472
		(0.017)	(0.091)	(0.163)	(0.269)	(0.163)
Annual	0.527	0.493	0.399	0.999	0.410	1.001
		(0.031)	(0.197)	(0.229)	(0.906)	(0.227)
3-Year	1.406	1.296	0.601	1.973	0.269	1.976
		(0.057)	(0.890)	(0.370)	(2.844)	(0.345)
5-Year	2.167	1.971	0.595	2.913	0.187	2.937
		(0.116)	(2.793)	(0.803)	(5.587)	(0.660)
			Panel B: Conditional Skewness			
Daily	0.030	0.017	0.006	0.005	0.016	0.006
		(0.003)	(0.007)	(0.008)	(0.012)	(0.007)
Weekly	0.059	0.032	0.011	0.104	0.147	0.104
		(0.011)	(0.021)	(0.033)	(0.020)	(0.033)
Monthly	0.069	0.029	0.001	0.260	0.310	0.260
		(0.028)	(0.061)	(0.119)	(0.079)	(0.119)
Quarterly	0.027	0.060	0.080	0.471	0.432	0.472
		(0.036)	(0.104)	(0.208)	(0.238)	(0.208)
Annual	0.515	0.497	0.399	0.999	0.410	1.001
		(0.036)	(0.199)	(0.244)	(0.887)	(0.242)
3-Year	1.402	1.303	0.601	1.973	0.269	1.976
		(0.057)	(0.887)	(0.374)	(2.833)	(0.350)
5-Year	2.165	1.978	0.595	2.913	0.187	2.937
		(0.114)	(2.788)	(0.806)	(5.577)	(0.663)

Table 1.7: The Performance of the Heston Skewness Estimator Under a Heston-Jump Process

This table presents the true unconditional (Panel A) and mean conditional (Panel B) skewness of the discrete return over various horizons (column (1)) as well as the mean estimates of those skewnesses plus their mean-squared errors (MSEs; in parentheses) obtained from separately applying our Heston (1993) estimator ("HE"; (2)) or its competitors ((3) to (6)) to 10,000 sample paths obtained from a Heston-jump process. We measure conditional skewness (true or estimated) at the end of each sample path. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, $\rho =$ 0.50, $\lambda = 1.70$, $\overline{\gamma} =$ 0.0301, and $\delta = 0.0078$.

Horizon	True	Heston Estimator (HE)	Sample Skewness (SS)	Fama-French (FF)	Neuberger-Payne (NP)	Farago-Hjalmarsson (FH)
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Unconditional Skewness			
Daily	0.010	0.013	0.007	0.007	0.009	0.007
		(0.001)	(0.004)	(0.005)	(0.009)	(0.004)
Weekly	0.033	0.031	0.038	0.099	0.082	0.099
		(0.002)	(0.012)	(0.006)	(0.019)	(0.005)
Monthly	0.097	0.076	0.104	0.248	0.178	0.248
		(0.006)	(0.041)	(0.024)	(0.083)	(0.023)
Quarterly	0.207	0.175	0.201	0.450	0.271	0.451
		(0.014)	(0.094)	(0.061)	(0.240)	(0.060)
Annual	0.609	0.552	0.449	0.951	0.310	0.952
		(0.030)	(0.222)	(0.122)	(0.870)	(0.121)
3-Year	1.432	1.289	0.585	1.853	0.216	1.859
		(0.072)	(0.949)	(0.215)	(2.755)	(0.201)
5-Year	2.160	1.923	0.568	2.702	0.150	2.725
		(0.146)	(2.843)	(0.465)	(5.381)	(0.377)
			Panel B: Conditional Skewness			
Daily	0.035	0.010	0.007	0.007	0.009	0.007
		(0.007)	(0.008)	(0.009)	(0.013)	(0.008)
Weekly	0.001	0.025	0.038	0.099	0.082	0.099
		(0.008)	(0.016)	(0.014)	(0.016)	(0.014)
Monthly	0.050	0.068	0.104	0.248	0.178	0.248
		(0.018)	(0.051)	(0.046)	(0.066)	(0.046)
Quarterly	0.162	0.170	0.201	0.450	0.271	0.451
		(0.028)	(0.104)	(0.093)	(0.208)	(0.092)
Annual	0.591	0.556	0.449	0.951	0.310	0.952
		(0.036)	(0.221)	(0.138)	(0.842)	(0.137)
3-Year	1.426	1.297	0.585	1.853	0.216	1.859
		(0.069)	(0.943)	(0.220)	(2.739)	(0.206)
5-Year	2.156	1.932	0.568	2.702	0.150	2.725
		(0.138)	(2.835)	(0.469)	(5.366)	(0.380)

Table 1.8: The Ability of the Skewness Estimators to Capture Real-World Risk-Neutral Skewness

This table presents the results from time-series regressions of the conditional risk-neutral skewness of the one-month (Panel A) and three-month (Panel B) discrete S&P 500 return extracted from options on the conditional skewness estimates obtained from our Heston (1993) estimator and the unconditional estimates obtained from its competitors. We use the ten-years of daily data ending with the date on which we back out risk-neutral skewness from options prices to calculate each skewness estimate. The competing estimators are the sample skewness, the Fama-French (2018) bootstrap estimator, the Neuberger-Payne (2021) closed-form estimator, and the Farago-Hjalmarsson (2023) closed-form estimator. Plain numbers are coefficient estimates, while those in square brackets are Newey-West (1987) t-statistics with a twelve-month lag length. We report the adjusted R-squared of each regression at the bottom of each panel.

			Panel A: Monthly Horizon			
Heston Estimator	1.471					1.325
	[6.96]					[5.74]
Sample Skewness		0.682				1.315
		[3.21]				[2.04]
Fama-French			0.694			0.208
			[2.47]			[0.32]
Neuberger-Payne				0.223		1.112
				[1.19]		[2.05]
Farago-Hjalmarsson					0.729	1.124
					[2.61]	[1.04]
Constant	0.554	1.162	1.612	1.310	1.610	1.491
	[3.58]	[7.51]	[11.71]	[4.50]	[11.73]	[2.66]
Adj. R^2	0.422	0.172	0.023	0.035	0.025	0.488
			Panel B: Quarterly Horizon			
Heston Estimator	1.061					0.802
	[7.15]					[5.17]
Sample Skewness		0.596				0.943
		[3.79]				[3.10]
Fama-French			0.548			0.575
			[2.75]			[0.83]
Neuberger-Payne				0.213		0.605
				[1.37]		[2.44]
Farago-Hjalmarsson					0.558	0.558
					[2.99]	[0.69]
Constant	0.557	0.913	1.314	1.019	1.314	0.998
	[4.92]	[8.07]	[12.45]	[4.30]	[12.45]	[4.05]
Adj. R^2	0.375	0.220	0.024	0.054	0.026	0.469

Figure 1.1: Heston Process's Unconditional Skewness-Return Horizon Relation

This figure plots the true unconditional skewness of the discrete return under the Heston (1993) process against the return horizon. Our base case parameters are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, and $\rho = \quad 0.50$. For each alternative set, we change one of the process parameters to the value shown in the legend.

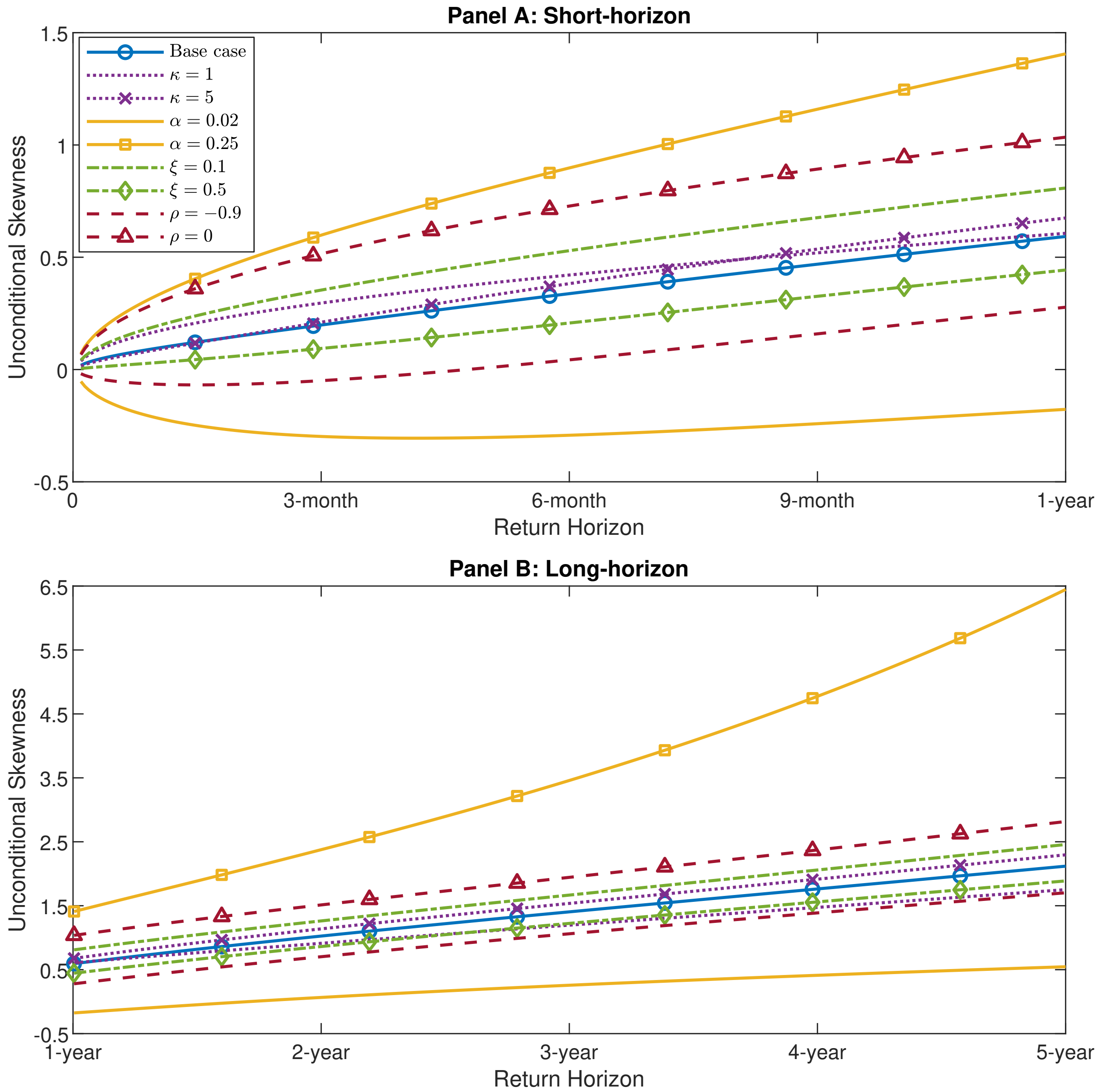


Figure 1.2: The Mean Squared Errors of the Monthly Unconditional Skewness Estimates

This figure plots the mean-squared errors of our Heston (1993) estimator (“HE”) and its competitors obtained from estimating the true unconditional skewness of the monthly discrete return from 10,000 sample paths generated from our basecase and comparative statics Heston (1993) processes. The competing estimators are the sample skewness (“SS”), the Fama-French (2018) bootstrap estimator (“FF”), the Neuberger-Payne (2021) closed-form estimator (“NP”), and the Farago-Hjalmarsson (2023) closed-form estimator (“FH”). Each of the 10,000 sample paths features ten years of daily discrete returns. The basecase parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, $\rho =$ 0.50. In the alternative sets, we vary a single parameter from its basecase value, indicating the value of that parameter in the panel heading.

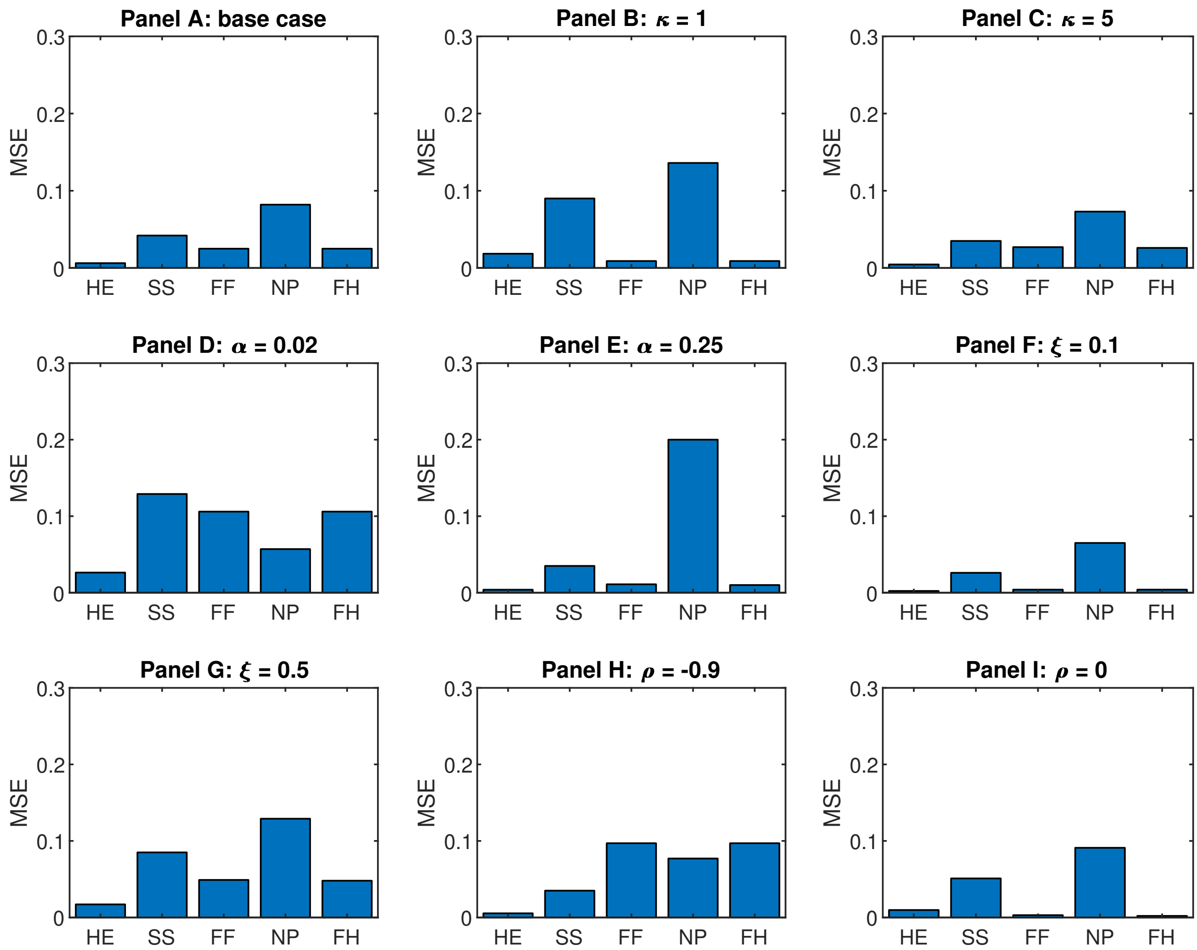
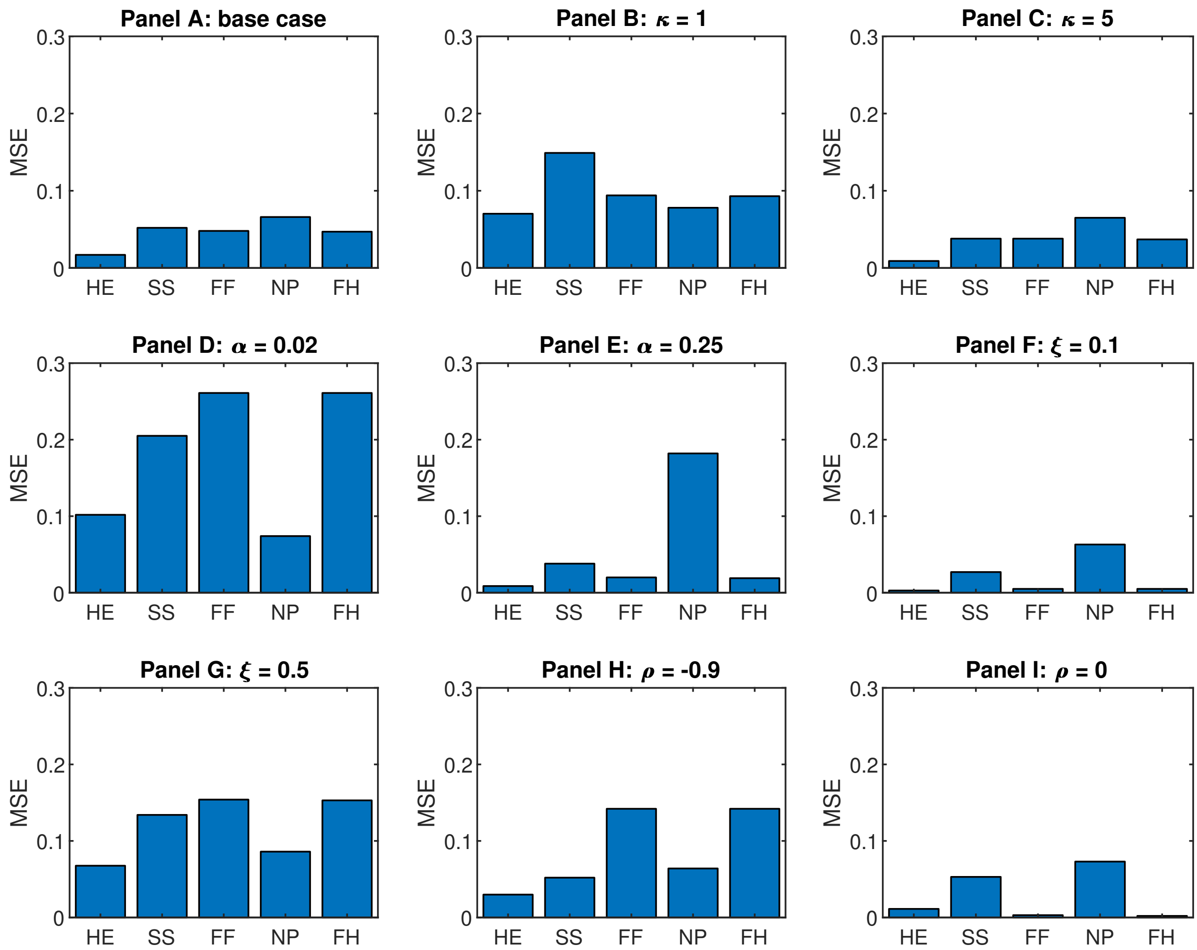


Figure 1.3: The Mean Squared Errors of the Monthly Conditional Skewness Estimates

This figure plots the mean-squared errors of our Heston (1993) estimator ("HE") and its competitors obtained from estimating the true conditional skewness of the monthly discrete return from 10,000 sample paths generated from our basecase and comparative statics Heston (1993) processes. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap estimator ("FF"), the Neuberger-Payne (2021) closed-form estimator ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimator ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The basecase parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, $\rho =$ 0.50. In the alternative sets, we vary a single parameter from its basecase value, indicating the value of that parameter in the panel heading.



Appendix 1.A Implementing Proposition 1

In this section, we explain how we approximate $M_i \ (u, \boldsymbol{w}^{-}, h)$ to avoid the curse of dimensionality often plaguing iterative numerical integrations. As stated in Section 1.2.2.1, we always start from calculating $M_0 \ (u, \boldsymbol{w}^{-}, h)$ to implement Proposition 1. Using that estimate, we then compute $G_1(\boldsymbol{w}_1 | \mathcal{G}_1 \ )$. We next approximate $G_1(\boldsymbol{w}_1 | \mathcal{G}_1 \ )$ by the MGF of a distribution of the same type as the limit distribution. Take the Heston (1993) process as example. In that process, V_t follows a CIR process with limit distribution $\Gamma(\frac{2\kappa\alpha}{\xi^2}, \frac{\xi^2}{2\kappa})$. To approximate $G_1(w_1 | \mathcal{G}_1 \ )$ by a gamma distribution, we first pin down its scale parameter (sc) by numerically solving:

$$\frac{\ln(1 \quad sc \quad w_{11} \quad i)}{\ln(1 \quad sc \quad w_{12} \quad i)} = \frac{\ln(G_1(w_{11} | \mathcal{G}_1 \ ))}{\ln(G_1(w_{12} | \mathcal{G}_1 \ ))}, \tag{1.A.1}$$

where w_{11} and w_{12} are two arbitrary values of w_1, and we can rely on the scale parameter of the limit distribution (i.e., $\frac{\xi^2}{2\kappa}$) as initial guess. Once we have obtained sc, we can then compute the shape parameter (sh) of this gamma distribution from:

$$sh = \quad \frac{\ln(G_1(w_1 | \mathcal{G}_1 \ ))}{\ln(1 \quad sc \quad w_1 \quad i)} \tag{1.A.2}$$

Using the ratios of log MGFs, we can separately estimate the two parameters of the gamma distribution, increasing approximation precision. Evaluating the MGF of $\Gamma(sh, sc)$ at $\psi(u, w1, h)$, we replace $G_1(w_1 | \mathcal{G}_1 \ )$ in Eq. (1.2.11) with $(1 \quad sc \quad w_1) \quad ^{sh}$. We use that same approximation for each $G_i(w_1 | \mathcal{G}_i \ )$ with $i \in \{1, 2, \ldots, N\}$ until we obtain the desired $M_i \ (u, w_1, h)$.

If the numerical integration of $M_i \ (u, \boldsymbol{w}^{-}, h)$ fails, we calculate its inverse Fourier transform from the Fourier Cosine transform of Fang and Oosterlee (2009).

Appendix 1.B Derivations of Moment Expressions

1.B.1 Proof of Lemma 1

Let us denote the non-centered cross-MGF as $C(m_1, m_2, \Delta, j) := \mathrm{E}[R_{i,}^{m_1} R_{i+j,}^{m_2}\ ]$. Using the definition of the MGFs, we can then show that:

$$
\begin{aligned}
C(m_1, m_2, \Delta, j) &:= \mathrm{E}[R_{i,}^{m_1} R_{i+j,}^{m_2}\ ] \\
&= \mathrm{E}[R_{i,}^{m_1} \mathrm{E}[R_{i+j,}^{m_2}\ jG_{(i+j)}\ ]] \\
&= e^{\mu\ \ m_2 + \phi(m_2, 0,\ \ )} \mathrm{E}[R_{i,}^{m_1} e^{\psi(m_2, 0,\ \ ) V_{(i+j)\Delta}}] \\
&= e^{\mu\ \ m_2 + \phi(m_2, 0,\ \ )} \mathrm{E}[R_{i,}^{m_1} \mathrm{E}[e^{\psi(m_2, 0,\ \ ) V_{(i+j)\Delta}} jG_{(i+1)}\ ]] \\
&= e^{\mu\ \ m_2 + \phi(m_2, 0,\ \ ) + \phi(0, \psi(m_2, 0,\ \ ), (j\ \ 1)\ \ )} \mathrm{E}[R_{i,}^{m_1} e^{\psi(0, \psi(m_2, 0,\ \ ), (j\ \ 1)\ \ ) V_{(i+1)\Delta}}] \\
&= e^{\mu\ \ m_2 + \phi(m_2, 0,\ \ ) + \phi(0, \psi(m_2, 0,\ \ ), (j\ \ 1)\ \ )} M(m_1, \psi(0, \psi(m_2, 0, \Delta), (j\ \ \ 1)\Delta), \Delta),
\end{aligned}
\tag{1.B.1}
$$

which holds whenever the moments exist.

1.B.2 Centred MGFs and Cross-MGFs for the Heston Process

To illustrate that the centered MGFs and cross-MGFs are simply linear combinations of the non-centered MGFs and cross-MGFs, we use the moments from our GMM estimator as example (see Eq. (1.3.2)) and derive their expressions. Recall that $M(k, \Delta)\ \ \ \mathrm{E}[R_{i,}^{k}\ ]$ and $\tilde{M}(k, \Delta) := \mathrm{E}[\tilde{R}_{i,}^{k}\ ]$. Starting from $\tilde{M}(2, \Delta)$, we can then show that:

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

We next derive $\tilde{M}(3,\Delta)$:

$$\begin{aligned}
\tilde{M}(3,\Delta) &:= \mathrm{E}[\tilde{R}^3_{i,} \ ] \\
&= \mathrm{E}[(R_{i,} \quad M(1,\Delta))^3] \\
&= \mathrm{E}[R^3_{i,} \quad 3R^2_{i,} \ M(1,\Delta) + 3R_{i,} \ M^2(1,\Delta) \quad M^3(1,\Delta)] \\
&= M(3,\Delta) \quad 3M(1,\Delta)M(2,\Delta) + 2M^3(1,\Delta)
\end{aligned} \tag{1.B.3}$$

We now expand $\tilde{M}(4,\Delta)$:

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

Further recalling that $\tilde{C}(m_1, m_2, \Delta, j) := \mathrm{Cov}[\tilde{R}^{m_1}_{i,}, \tilde{R}^{m_2}_{i+j,} \ ]$, we have:

$$\begin{aligned}
\tilde{C}(1,2,\Delta,j) &:= \mathrm{Cov}[\tilde{R}_{i,} \ , \tilde{R}^2_{i+j,} \ ] \\
&= \mathrm{E}[\tilde{R}_{i,} \ \tilde{R}^2_{i+j,} \ ] \quad \mathrm{E}[\tilde{R}_{i,} \ ]\,\mathrm{E}[\tilde{R}^2_{i+j,} \ ] \\
&= \mathrm{E}[(R_{i,} \quad M(1,\Delta))(R_{i+j,} \quad M(1,\Delta))^2] \\
&\qquad \mathrm{E}[R_{i,} \quad M(1,\Delta)]\,\mathrm{E}[(R_{i+j,} \quad M(1,\Delta))^2] \\
&= \mathrm{E}[R_{i,} \ R^2_{i+j,} \quad 2R_{i,} \ R_{i+j,} \ M(1,\Delta) + R_{i,} \ M^2(1,\Delta) \\
&\qquad R^2_{i+j,} \ M(1,\Delta) + 2R_{i+j,} \ M^2(1,\Delta) \quad M^3(1,\Delta)] \\
&= C(1,2,\Delta,j) \quad 2M^3(1,\Delta) + M^3(1,\Delta) \\
&\qquad M(1,\Delta)M(2,\Delta) + 2M^3(1,\Delta) \quad M^3(1,\Delta) \\
&= C(1,2,\Delta,j) \quad M(1,\Delta)M(2,\Delta)
\end{aligned} \tag{1.B.5}$$

We finally expand $\tilde{C}(2,2,\Delta,j)$.

$$
\begin{aligned}
\tilde{C}(2,2,\Delta,j) \quad & \mathrm{Cov}[\tilde{R}^2_{i,} \;, \tilde{R}^2_{i+j,} \;] \\
&= \mathrm{E}[\tilde{R}^2_{i,} \; \tilde{R}^2_{i+j,} \;] \quad \mathrm{E}[\tilde{R}^2_{i,} \;]\,\mathrm{E}[\tilde{R}^2_{i+j,} \;] \\
&= \mathrm{E}[(R_{i,} \quad M(1,\Delta))^2\,(R_{i+j,} \quad M(1,\Delta))^2] \\
&\qquad \mathrm{E}[(R_{i,} \quad M(1,\Delta))^2]\,\mathrm{E}[(R_{i+j,} \quad M(1,\Delta))^2] \\
&= \mathrm{E}[R^2_{i,} \; R^2_{i+j,} \quad 2R^2_{i,} \; R_{i+j,} \; M(1,\Delta) + R^2_{i,} \; M^2(1,\Delta) \\
&\qquad 2R_{i,} \; R^2_{i+j,} \; M(1,\Delta) + 4R_{i,} \; R_{i+j,} \; M^2(1,\Delta) \quad 2R_{i,} \; M^3(1,\Delta) \\
&\quad + R^2_{i+j,} \; M^2(1,\Delta) \quad 2R_{i+j,} \; M^3(1,\Delta) + M^4(1,\Delta)] \\
&\qquad \mathrm{E}[R^2_{i,} \quad 2R_{i,} \; M(1,\Delta) + M^2(1,\Delta)] \\
&\quad \; \mathrm{E}[R^2_{i+j,} \quad 2R_{i+j,} \; M(1,\Delta) + M^2(1,\Delta)] \\
&= C(2,2,\Delta,j) \quad 2M^2(1,\Delta)M(2,\Delta) + M^2(1,\Delta)M(2,\Delta) \\
&\qquad 2M(1,\Delta)C(1,2,\Delta,j) + 4M^4(1,\Delta) \quad 2M^4(1,\Delta) \\
&\quad + M^2(1,\Delta)M(2,\Delta) \quad 2M^4(1,\Delta) + M^4(1,\Delta) \\
&\qquad (M(2,\Delta) \quad 2M^2(1,\Delta) + M^2(1,\Delta))^2 \\
&= C(2,2,\Delta,j) \quad 2M(1,\Delta)C(1,2,\Delta,j) \quad M^2(2,\Delta) + 2M^2(1,\Delta)M(2,\Delta)
\end{aligned}
\tag{1.B.6}
$$

Appendix 1.C Implementation of Our GMM Estimator

In this section, we provide further details on the numerical implementation of our GMM estimator in Section 1.3. We use `fmincon` in MATLAB with the sequential quadratic programming (SQP) algorithm and user-supplied analytical gradient to solve the optimization problem stated in Eq. (1.3.4). The key advantage of using analytical gradients is that it boosts the convergence speed of the numerical algorithm. Rewriting Eq. (1.3.2) as:

$$
\begin{aligned}
h_1(\boldsymbol{\theta}) &= \bar{R} \quad 1 \quad \mu\Delta \quad \frac{\mu^2}{2}\Delta^2 \\
h_2(\boldsymbol{\theta}) &= \overline{\tilde{R}^2} \quad \alpha\Delta \quad \alpha(\frac{\xi^2}{4\kappa} + \frac{\alpha}{2} + 2\mu + \xi\rho)\Delta^2 \\
h_3(\boldsymbol{\theta}) &= \overline{\tilde{R}^3} \quad 3\alpha(\alpha + \frac{\xi^2}{2\kappa} + \frac{\xi\rho}{2})\Delta^2 \\
h_4(\boldsymbol{\theta}) &= \overline{\tilde{R}^4} \quad 3\alpha(\alpha + \frac{\xi^2}{2\kappa})\Delta^2 \\
h_{12,j}(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{i=j+1}^{T} \tilde{R}_{i,} \ (\tilde{R}^2_{i+j,} \qquad \overline{\tilde{R}^2}) \quad \alpha\xi\rho e^{\ \kappa(j \ \ 1)} \ \Delta^2 \\
h_{22,j}(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{i=j+1}^{T} (\tilde{R}^2_{i,} \qquad \overline{\tilde{R}^2})(\tilde{R}^2_{i+j,} \qquad \overline{\tilde{R}^2}) \quad \frac{\alpha\xi^2}{2\kappa} e^{\ \kappa(j \ \ 1)} \ \Delta^2,
\end{aligned}
\tag{1.C.1}
$$

where $\bar{R} \quad \frac{1}{T}\sum_{i=1}^{T} R_{i,}$, and $\overline{\tilde{R}^k} \quad \frac{1}{T}\sum_{i=1}^{T} \tilde{R}^k_{i,}$, with:

$$
\frac{\partial \tilde{R}^k_{i,}}{\partial \mu} = \quad k\tilde{R}^{k \ \ 1}_{i,} e^{\mu} \ \ \Delta, \qquad \frac{\partial \overline{\tilde{R}^k}}{\partial \mu} = \quad k\overline{\tilde{R}^{k \ \ 1}} e^{\mu} \ \ \Delta,
\tag{1.C.2}
$$

we then derive the analytical gradients with respect to:

(i) μ:

$$
\begin{aligned}
\frac{\partial h_1(\boldsymbol{\theta})}{\partial \mu} &= \quad \Delta \quad \mu\Delta^2, \quad \frac{\partial h_2(\boldsymbol{\theta})}{\partial \mu} = \quad 2(\bar{R} \quad e^{\mu} \ \ )e^{\mu} \ \ \Delta \quad 2\alpha\Delta^2, \\
\frac{\partial h_3(\boldsymbol{\theta})}{\partial \mu} &= \quad 3\overline{\tilde{R}^2} e^{\mu} \ \ \Delta, \quad \frac{\partial h_4(\boldsymbol{\theta})}{\partial \mu} = \quad 4\overline{\tilde{R}^3} e^{\mu} \ \ \Delta, \\
\frac{\partial h_{12,j}(\boldsymbol{\theta})}{\partial \mu} &= \quad \frac{1}{T}\sum_{i=j+1}^{T} (\tilde{R}^2_{i+j,} \qquad \overline{\tilde{R}^2} + 2\tilde{R}_{i,} \ (\tilde{R}_{i,} \qquad \bar{R} + e^{\mu} \ \ ))e^{\mu} \ \ \Delta, \\
\frac{\partial h_{22,j}(\boldsymbol{\theta})}{\partial \mu} &= \quad \frac{2}{T}\sum_{i=j+1}^{T} ((\tilde{R}^2_{i,} \qquad \overline{\tilde{R}^2})(\tilde{R}_{i+j,} \qquad \bar{R} + e^{\mu} \ \ ) \\
&\quad + (\tilde{R}^2_{i+j,} \qquad \overline{\tilde{R}^2})(\tilde{R}_{i,} \qquad \bar{R} + e^{\mu} \ \ ))e^{\mu} \ \ \Delta;
\end{aligned}
\tag{1.C.3}
$$

(ii) κ:

$$
\begin{aligned}
\frac{\partial h_2(\boldsymbol{\theta})}{\partial \kappa} &= \frac{\alpha\xi^2}{4\kappa^2}\Delta^2, \ \frac{\partial h_3(\boldsymbol{\theta})}{\partial \kappa} = \frac{\partial h_4(\boldsymbol{\theta})}{\partial \kappa} = \frac{3\alpha\xi^2}{2\kappa^2}\Delta^2, \\
\frac{\partial h_{12,j}(\boldsymbol{\theta})}{\partial \kappa} &= \alpha\xi\rho(j-1)e^{-\kappa(j-1)\Delta}\Delta^3, \\
\frac{\partial h_{22,j}(\boldsymbol{\theta})}{\partial \kappa} &= \frac{\alpha\xi^2}{2\kappa}e^{-\kappa(j-1)\Delta}\Delta^2\left(\frac{1}{\kappa} + (j-1)\Delta\right);
\end{aligned}
\tag{1.C.4}
$$

(iii) α:

$$
\begin{aligned}
\frac{\partial h_2(\boldsymbol{\theta})}{\partial \alpha} &= \Delta - \left(\frac{\xi^2}{4\kappa} + \alpha + 2\mu + \xi\rho\right)\Delta^2, \\
\frac{\partial h_3(\boldsymbol{\theta})}{\partial \alpha} &= 3\left(2\alpha + \frac{\xi^2}{2\kappa} + \frac{\xi\rho}{2}\right)\Delta^2, \ \frac{\partial h_4(\boldsymbol{\theta})}{\partial \alpha} = 3\left(2\alpha + \frac{\xi^2}{2\kappa}\right)\Delta^2, \\
\frac{\partial h_{12,j}(\boldsymbol{\theta})}{\partial \alpha} &= \xi\rho e^{-\kappa(j-1)\Delta}\Delta^2, \ \frac{\partial h_{22,j}(\boldsymbol{\theta})}{\partial \alpha} = \frac{\xi^2}{2\kappa}e^{-\kappa(j-1)\Delta}\Delta^2;
\end{aligned}
\tag{1.C.5}
$$

(iv) ξ:

$$
\begin{aligned}
\frac{\partial h_2(\boldsymbol{\theta})}{\partial \xi} &= \alpha\left(\frac{\xi}{2\kappa} + \rho\right)\Delta^2, \ \frac{\partial h_3(\boldsymbol{\theta})}{\partial \xi} = 3\alpha\left(\frac{\xi}{\kappa} + \frac{\rho}{2}\right)\Delta^2, \ \frac{\partial h_4(\boldsymbol{\theta})}{\partial \xi} = 3\frac{\alpha\xi}{\kappa}\Delta^2, \\
\frac{\partial h_{12,j}(\boldsymbol{\theta})}{\partial \xi} &= \alpha\rho e^{-\kappa(j-1)\Delta}\Delta^2, \ \frac{\partial h_{22,j}(\boldsymbol{\theta})}{\partial \xi} = \frac{\alpha\xi}{\kappa}e^{-\kappa(j-1)\Delta}\Delta^2;
\end{aligned}
\tag{1.C.6}
$$

(v) ρ:

$$
\frac{\partial h_2(\boldsymbol{\theta})}{\partial \rho} = \alpha\xi\Delta^2, \ \frac{\partial h_3(\boldsymbol{\theta})}{\partial \rho} = \frac{3\alpha\xi}{2}\Delta^2, \ \frac{\partial h_{12,j}(\boldsymbol{\theta})}{\partial \rho} = \alpha\xi e^{-\kappa(j-1)\Delta}\Delta^2. \tag{1.C.7}
$$

To make the optimization less dependent on the initial parameter value guesses, we generate ten sets of random initial guesses and use the one with the minimum error as input for the second step. 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 $\rho \in [-1, 1]$ to ensure that our parameter estimates fall within reasonable ranges. We also impose $\kappa > \xi\rho$ and $\sqrt{(\kappa - 3\xi\rho)^2 - 6\xi^2} + \kappa + 3\xi\rho > 0$ to guarantee the existence of skewness (see Section 6.4 of Keller-Ressel (2011)).

Appendix 1.D A Heston-Jump Estimator

In this section, we describe how we generalize our Heston (1993) estimator to a Heston-jump estimator (see Bates (1996)). We also investigate how much better the Heston-jump estimator performs relative to our main estimator in a Heston-jump world.

To obtain a Heston-jump estimator, we use the exact same GMM framework as in Section 1.3 but alter some implementation details. The most important change is the short-time-increment power-expansion approximations used in our GMM estimation. We find that jump with constant intensity does not affect the mean or the two central cross-moments of returns but changes $\tilde{M}(2,\Delta)$, $\tilde{M}(3,\Delta)$, and $\tilde{M}(4,\Delta)$ to:

$$\begin{aligned}
\tilde{M}(2,\Delta) &= (\alpha + \lambda I(2,0))\Delta + \left(\alpha\left(2\mu + \frac{\alpha}{2} + \frac{\xi^2}{4\kappa} + \xi\rho\right)\right. \\
&\quad + \frac{\lambda^2}{2} I^2(2,0) + (2\mu\lambda + \alpha\lambda) I(2,0)\Delta^2 + o\left(\Delta^2\right). \\
\tilde{M}(3,\Delta) &= \lambda(I(3,0) - 3I(2,0))\Delta + \left(3\alpha\left(\alpha + \frac{\xi^2}{2\kappa} + \frac{\xi\rho}{2}\right)\right. \\
&\quad + \frac{\lambda^2}{2}\left(I^2(3,0) - 3I^2(2,0)\right) + 3\mu\lambda(I(3,0) - 3I(2,0)) \\
&\quad \left. + 3\alpha\lambda(I(3,0) - I(2,0))\right)\Delta^2 + o\left(\Delta^2\right). \\
\tilde{M}(4,\Delta) &= \lambda(I(4,0) - 4I(3,0) + 6I(2,0))\Delta + \left(3\alpha\left(\alpha + \frac{\xi^2}{2\kappa}\right)\right. \\
&\quad + \frac{\lambda^2}{2}\left(I^2(4,0) - 4I^2(3,0) + 6I^2(2,0)\right) + 4\mu\lambda(I(4,0) - 4I(3,0) + 6I(2,0)) \\
&\quad \left. + 6\alpha\lambda(I(4,0) - 2I(3,0) + I(2,0))\right)\Delta^2 + o\left(\Delta^2\right).
\end{aligned} \tag{1.D.1}$$

Regarding the parameter constraints, we use the same boundaries for the Heston parameters while require $\lambda \in (0, 6]$, $\bar{\gamma} \in [-0.25, 0.25]$, and $\delta \in (0, 1]$ to get reasonable estimates. Due to mathematical complexity, we are unable to update constraints that guarantee skewness existence, thus we keep it unchanged.

Table 1.D.1 copies the entire Table 1.7 and adds an extra column to reflect the performance of our Heston-jump estimator. The new column shows that correctly specifying the data-generating process does improve mean estimates at very short-horizons like daily. For instance, the absolute bias for daily unconditional skewness shrinks from 0.023 to 0.007. But in terms of MSEs, the precision of the Heston-jump estimator is not much different from the Heston (1993) estimator.

Table 1.D.1: The Performance of the Heston-Jump Skewness Estimator Under a Heston-Jump Process

This table presents the true unconditional (Panel A) and mean conditional (Panel B) skewness of the discrete return over various horizons (column (1)) as well as the mean estimates of those skewnesses plus their mean-squared errors (MSEs; in brackets) obtained from separately applying our Heston (1993) and Heston-jump estimator ("HE" and "HJE"; (2) and (7)) or their competitors ((3) to (6)) to 10,000 sample paths obtained from a Heston-jump process. We measure conditional skewness (true or estimated) at the end of each sample path. We look into the daily, weekly, monthly, quarterly, annual, three-year, and five-year horizons. The competing estimators are the sample skewness ("SS"), the Fama-French (2018) bootstrap ("FF"), the Neuberger-Payne (2021) closed-form ("NP"), and the Farago-Hjalmarsson (2023) closed-form estimators ("FH"). Each of the 10,000 sample paths features ten years of daily discrete returns. The parameter values are $\mu = 0.10$, $\kappa = 3.00$, $\alpha = 0.09$, $\xi = 0.30$, $\rho =$ 0.50, $\lambda = 1.70$, $\overline{\gamma} =$ 0.0301, and $\delta = 0.0078$.

Horizon	True	Heston Estimator (HE)	Sample Skewness (SS)	Fama-French (FF)	Neuberger-Payne (NP)	Farago-Hjalmarsson (FH)	Heston-Jump Estimator (HJE)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Unconditional Skewness			
Daily	0.010	0.013	0.007	0.007	0.009	0.007	0.003
		(0.001)	(0.004)	(0.005)	(0.009)	(0.004)	(0.002)
Weekly	0.033	0.031	0.038	0.099	0.082	0.099	0.027
		(0.002)	(0.012)	(0.006)	(0.019)	(0.005)	(0.002)
Monthly	0.097	0.076	0.104	0.248	0.178	0.248	0.079
		(0.006)	(0.041)	(0.024)	(0.083)	(0.023)	(0.005)
Quarterly	0.207	0.175	0.201	0.450	0.271	0.451	0.181
		(0.014)	(0.094)	(0.061)	(0.240)	(0.060)	(0.012)
Annual	0.609	0.552	0.449	0.951	0.310	0.952	0.556
		(0.030)	(0.222)	(0.122)	(0.870)	(0.121)	(0.029)
3-Year	1.432	1.289	0.585	1.853	0.216	1.859	1.292
		(0.072)	(0.949)	(0.215)	(2.755)	(0.201)	(0.070)
5-Year	2.160	1.923	0.568	2.702	0.150	2.725	1.926
		(0.146)	(2.843)	(0.465)	(5.381)	(0.377)	(0.143)
				Panel B: Conditional Skewness			
Daily	0.035	0.010	0.007	0.007	0.009	0.007	0.006
		(0.007)	(0.008)	(0.009)	(0.013)	(0.008)	(0.007)
Weekly	0.001	0.025	0.038	0.099	0.082	0.099	0.021
		(0.008)	(0.016)	(0.014)	(0.016)	(0.014)	(0.008)
Monthly	0.050	0.068	0.104	0.248	0.178	0.248	0.071
		(0.018)	(0.051)	(0.046)	(0.066)	(0.046)	(0.017)
Quarterly	0.162	0.170	0.201	0.450	0.271	0.451	0.177
		(0.028)	(0.104)	(0.093)	(0.208)	(0.092)	(0.027)
Annual	0.591	0.556	0.449	0.951	0.310	0.952	0.562
		(0.036)	(0.221)	(0.138)	(0.842)	(0.137)	(0.035)
3-Year	1.426	1.297	0.585	1.853	0.216	1.859	1.301
		(0.069)	(0.943)	(0.220)	(2.739)	(0.206)	(0.067)
5-Year	2.156	1.932	0.568	2.702	0.150	2.725	1.937
		(0.138)	(2.835)	(0.469)	(5.366)	(0.380)	(0.134)

Chapter 2

The Skewness of Discrete Single-Stock Returns

2.1 Introduction

The skewness of an asset's discrete returns has always been of great interest to both academics and practitioners. While the initial academic literature focuses on whether investors like assets with skewed returns (Arditti (1967)), the effects of skewness on portfolio diversification (Conine and Tamarkin (1981)), and the pricing of co-skewness (Rubinstein (1973)), more recent studies have turned to the pricing of idiosyncratic skewness (Mitton and Vorkink (2007) and Barberis and Huang (2008)) and the skewness of asset returns over long horizons (Fama and French (2018) and Farago and Hjalmarsson (2023)). Yet, despite our comprehensive understanding of the theoretical properties and implications of skewness, we know far less about its empirical properties, especially over long return horizons, mostly because it is challenging to estimate an asset's discrete-return skewness with satisfactory precision (see, e.g., Lau et al. (1989)).

In this paper, we aim to provide deeper insights into the empirical properties of the skewness of discrete single-stock returns over various short and long horizons. To achieve that goal, we rely on the precise and computationally fast skewness estimator newly developed by Aretz et al. (2024). In a nutshell, these authors assume that the evolution of an asset's price can be approximated by an affine stochastic volatility (ASV) process. This allows them to estimate the parameters of that process, and to calculate the historical (or: forward-looking) skewness[1] of the asset from the fitted unconditional (or: conditional) moment-generating function (MGF) of the process. A significant advantage of their approach is that, by imposing a parametric assumption about the continuous-time stock price process, their estimator requires far less data to produce precise skewness estimates even over long return horizons.

Equipped with novel estimates, we reassess important empirical conclusions about the skewness of discrete single-stock returns from the literature. As a crucial first step, we show that our estimates easily outperform others in forecasting future daily, weekly, and monthly sample skewness. Spurred by Bessembinder's (2018), Fama and French's (2018), and Farago and Hjalmarsson's (2023) claim that the skewness of discrete single-stock returns increases monotonically with the return horizon, we then reexamine this claim, discovering large systematic exceptions to it. Motivated by studies which use historical estimates of the skewness of short-horizon returns to determine the pricing of skewness in stock markets (see, e.g., Boyer et al. (2010) and Amaya et al. (2015)), we next also assess the relations between historical and forward-looking skewness and between skewness over different return horizons. We find that while historical and forward-looking skewness align closely over long-return horizons, they can, for example, significantly differ over shorter and/or mismatched horizons and in economic expansions. Motivated by studies modelling skewness using firm fundamentals (e.g., Chen et al. (2001), Boyer et al. (2010), and Aretz and Arisoy (2023)), we also evaluate the relations between skewness and a comprehensive set of firm fundamentals,

[1]Throughout this chapter, we use "historical" and "forward-looking" to refer to in-sample and out-of-sample skewness estimates, respectively.

finding that only three out of 13 produce significant relations across our horizons, and that their overall explanatory power over long horizons is low. We finally reexamine the stock pricing of forward-looking skewness orthogonalized with respect to firm fundamentals known to price stocks over various horizons, reporting that its pricing power varies across horizons.

We begin by showing that our historical and forward-looking skewness estimates outperform the alternative sample skewness, Fama and French's (2018), Neuberger and Payne's (2021), and Farago and Hjalmarsson's (2023) estimates in forecasting daily, weekly, and monthly sample skewness (i.e., out-of-sample analyses). To that end, we conduct cross-sectional Mincer and Zarnowitz (1969) regressions of each future sample skewness computed from the three years of data subsequent to the end of month t on estimates obtained from each skewness estimator computed from data up to that month end. Our evidence suggests that our estimates are always most strongly related to the future sample skewnesses and always yield the highest mean adjusted R^2. Looking into monthly future sample skewness, our forward-looking estimator, for example, yields a slope coefficient of 0.35 (t-statistic: 16.08) and a mean adjusted R^2 of 4.22%. In comparison, none of the four alternative estimators produces a slope coefficient with a t-statistic greatly exceeding ten and a mean adjusted R^2 above 3.57%.

We then investigate how the skewness of discrete single-stock returns relates to the return horizon. Consistent with the claims in other studies, we report that the historical skewness of the median single stock is indeed monotonically positively related to the return horizon, rising from 0.12 at the daily horizon to 4.64 at the five-year horizon. Interestingly, however, lower-percentile stocks do not exhibit a monotonic pattern, with, for example, the historical skewness of the fifth-percentile stock initially dropping from 0.03 at the daily horizon to 0.09 at the monthly horizon before then rising to 1.40 at the five-year horizon. Crucially, we show that the variations in the skewness-time horizon patterns are systematic. Using a logit regression, we demonstrate that $MarketBeta$, $Size$, $BookToMarket$, $Profitability$, and $Leverage$ positively affect the likelihood of a stock exhibiting a U-shaped pattern, while $AssetGrowth$, $HistoricalVolatility$, $AssetTangibility$, $SalesGrowth$, and the $NBER$ recession indicator negatively impact this likelihood. We finally report that the skewness of a small number of stocks *explodes* with the return horizon, with, for example, about five percent of all stocks having a five-year historical or forward-looking skewness in excess of 1,000. The tendency of skewness to explode with the return horizon is stronger for stocks with a high $Momentum$, $AssetGrowth$, and $HistoricalVolatility$ as well as stocks with a low $MarketBeta$, $Size$, $BookToMarket$, $Profitability$, and $Leverage$.

Turning to the relations between historical and forward-looking skewness measured over potentially mismatched return horizons, we use Mincer and Zarnowitz (1969) regressions to show that while historical skewness is an imperfect and biased predictor of forward-looking skewness at the daily horizon (mean adjusted R^2: 63.10%), its precision improves and its bias vanishes with increases in the return horizon, with the two skewnesses virtually identical over the five-year horizon (mean adjusted R^2: 99.68%). Noteworthily, the absolute spread in the two skewnesses becomes larger in economic expansions, and not recessions. In addition, increases in $MarketBeta$, $Momentum$, $Profitability$, $HistoricalVolatility$, and $CompanyAge$, or decreases in $Size$, $ShareTurnover$, $AssetTangibility$, $Leverage$, and $SalesGrowth$ widen the spread. Allowing for mismatches in the return horizon, we establish that historical skewness can become an even less perfect and more biased predictor of forward-looking skewness. In fact, the larger the mismatch, the worse the performance is. For example, the mean adjusted R^2 obtained from a regression of five-year forward-looking skewness on daily historical skewness is only about 1.22%.

We further reevaluate the relations between the skewness of discrete single-stock returns and firm fundamentals conjectured in earlier studies to reflect skewness or skewness estimates mostly constructed from those same fundamentals. Using cross-sectional regressions, we find support for Bessembinder's (2018) hypothesis that $HistoricalVolatility$ is a primary driver of cross-sectional variations in stock skewness, especially over longer return horizons. We also corroborate Aretz and Arisoy's (2023) hypothesis that a stock's *AssetGrowth* further explain those same variations. In contrast, *Size*, *BookToMarket*, *Momentum*, *Profitability*, *AssetTangibility*, and *SalesGrowth* only explain meaningful fractions of those variations over short or moderately long horizons (up to the quarterly or annual), while *Leverage* and *ShareIssuance* never capture any meaningful fractions. Remarkably, the explanatory power of *ShareTurnover* and *CompanyAge* is primarily concentrated in capturing forward-looking rather than historical skewness. Turning to the skewness estimates mostly constructed from the firm fundamentals, we reveal that only Boyer et al.'s (2010) OLS-based and Aretz and Arisoy's (2023) quantile-regression-based estimates are moderately related to our estimates, but not Bali et al.'s (2011) maximum-return and Conrad et al.'s (2014) logit-based ones.

We finally take another look into the term structure of the stock skewness premium. To guard against the possibility that the premium could be driven by skewness being determined by firm fundamentals known to price stocks, such as stock volatility, we run partial least squares regressions of our forward-looking skewness estimates on various fundamentals, to isolate the residual skewness components. Using these residual components as pricing factors, our portfolio sorts suggest that both simple and risk-adjusted skewness premiums are significantly negative from daily to annual, but not at the three- and five-year return horizons. Our Fama-MacBeth (FM; 1973) regressions demonstrate that the skewness premium first drops and then rises with the return horizon, reaching its trough at around nine months. Moreover, the skewness premiums are significant at the 95% confidence level up to return horizons of about 30 months.

Our work feeds into studies estimating and forecasting the skewness of discrete returns. While the primary focus of these studies may not lie on exploring the empirical properties of skewness, they still enhance our understanding from two perspectives. First, many studies offer evidence on how skewness relates to the return horizon. Applying the sample skewness estimator to individual stocks, Bessembinder (2018), for example, finds that skewness monotonically increases with the return horizon. Assuming independently and identically distributed (i.i.d.) returns, the skewness estimators of Fama and French (2018) and Farago and Hjalmarsson (2023) reach the same conclusion. Neuberger and Payne (2021), however, argue that return dependency plays a crucial role in how skewness evolves over the return horizon. Allowing for return dependency, Aretz et al.'s (2024) estimator suggests that skewness may initially decrease before eventually increasing. Second, many studies (including, e.g., Boyer et al. (2010) and Aretz and Arisoy (2023)) employ various firm characteristics to model forward-looking skewness but often fall short of stating the rationales behind their choices of characteristics. We add to those studies by offering fresh evidence on (i) the skewness-return horizon relation; (ii) the tendency of skewness to explode; (iii) the spread between historical and forward-looking skewness over potentially mismatched horizons; and (iv) those firm characteristics most strongly related to skewness.

Our work also contributes to studies establishing how the skewness of an asset's discrete future return affects its expected return. Scott and Horvath (1980) show that von-Neumann-Morgenstern (NM) investors like positive skewness. Since diversification can reduce portfolio skewness, Simkowitz and Beedles (1978) and Conine and Tamarkin (1981) claim that these preferences can possibly explain why investors hold underdiversified portfolios. Further assuming monetary separation, Rubinstein (1973) establishes that only an asset's co-skewness (but not its total skewness) explains its expected return. To establish a pricing role for total skewness, recent studies look into settings violating monetary separation. For example, Mitton and Vorkink (2007) assume that investors have heterogeneous skewness preferences, while Brunnermeier et al. (2007) and Barberis and Huang (2008) use non-NM preferences. While several empirical studies offer evidence that skewness proxies negatively price assets (see Aretz and Arisoy (2023)), neither they nor the theoretical studies clarify over which return horizon skewness should be most strongly priced. We add to those studies by estimating the entire term structure of the forward-looking orthogonalized skewness premium from the daily to the five-year horizon.

We proceed as follows. Section 2 introduces Aretz et al.'s (2024) skewness estimator, while Section 3 briefly reviews competing estimators. Section 4 summarizes our data sources. In Section 5, we first confirm the reliability of our estimates and then explore the empirical properties of historical and forward-looking skewness across various return horizons. Section 6 links our skewness estimates to firm characteristics argued by the literature to reflect skewness and to skewness estimates mostly constructed from those. In Section 7, we investigate the pricing power of skewness orthogonalized with respect to firm fundamentals known to price stocks across return horizons. Section 8 sums up and concludes.

2.2 A New Parametric Heston-Process-Based Skewness Estimator

In this section, we give an overview of the precise and computationally fast skewness estimator newly developed by Aretz et al. (2024) conditional on the assumption that we can use Heston's (1993) stochastic process to approximate the evolution of an asset's price. We start with a brief technical introduction to the estimator. We next discuss how we can use a generalized method of moments (GMM) approach to estimate the parameters of Heston's (1993) stochastic process. We refer the reader to Aretz et al. (2024) for the statistical properties of the estimator.

2.2.1 An Introduction to the Estimator

Aretz et al. (2024) develop a skewness estimator from the assumption that we can use a stochastic process from the ASV class to describe an asset's price path. Given its simplicity, popularity, and the fact that it yields closed-form conditional and unconditional MGFs, we rely on Heston's (1993) stochastic process to model single-stock prices. We can write that process as:

$$
\begin{aligned}
dX_t &= (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t, \\
dV_t &= \kappa(\alpha - V_t)dt + \xi\sqrt{V_t}dB_t,
\end{aligned}
\tag{2.2.1}
$$

where $t \geq 0$, X_t is the natural logarithm of the stock price, V_t is the associated underlying stock variance, μ, κ, α, and ξ are the drift, the mean reversion, the long-run variance, and the volatility-of-volatility parameter,respectively, and W_t and B_t are Brownian motions with a correlation ρ. As shown in Section 2.2.2 of Aretz et al. (2024), the joint conditional and unconditional MGF of the Heston (1993) process can be written as:

$$
\begin{aligned}
M_t(u,w,h;V_t) &:= \mathrm{E}[e^{ur_{t,h}+wV_{t+h}}|X_t,V_t] = e^{\mu hu+\phi(u,w,h)+\psi(u,w,h)V_t}, \\
M(u,w,h) &:= \mathrm{E}[M_t(u,w,h;V_t)] = e^{\mu uh+\phi(u,w,h)}(1-\frac{\xi^2}{2\kappa}\psi(u,w,h))^{-2\kappa\alpha/\xi^2},
\end{aligned}
\tag{2.2.2}
$$

where $(u,w,h) \in \mathbb{C} \times \mathbb{C} \times \mathbb{R}_+$, $r_{t,h} := X_{t+h} - X_t$,

$$
\begin{aligned}
\phi(u,w,h) &= \frac{\kappa\alpha h}{\xi^2}D(u) - \frac{2\kappa\alpha}{\xi^2}\ln\{(1+\frac{e^{I(u)h}-1}{2I(u)}D(u))(1-wA(u,h))\}, \\
\psi(u,w,h) &= \frac{(u^2-u)A(u,h)}{\xi^2} + \frac{wB(u,h)}{1-wA(u,h)}, \quad A(u,h) = \frac{\xi^2}{I(u)C(u,h)+\kappa-u\xi\rho}, \\
B(u,h) &= \frac{C(u,h)^2-1}{(C(u,h)+\frac{\kappa-u\xi\rho}{I(u)})^2}, \quad C(u,h) = \frac{e^{I(u)h}+1}{e^{I(u)h}-1}, \quad D(u) = \kappa - u\xi\rho + I(u), \\
\text{and} \quad I(u) &= \sqrt{(\kappa-u\xi\rho)^2-\xi^2(u^2-u)}
\end{aligned}
\tag{2.2.3}
$$

Defining the discrete return $R_{t,h} := e^{r_{t,h}}$ and using $M(u,0,h) = \mathrm{E}[e^{r_{t,h}u}] = \mathrm{E}[R^u_{t,h}]$, we can easily express the historical (i.e., unconditional) skewness of $R_{t,h}$ as:

$$\text{Skew}[R_{t,h}] = \frac{M(3,0,h) - 3M(1,0,h)M(2,0,h) + 2M(1,0,h)^3}{[M(2,0,h) - M(1,0,h)^2]^{3/2}}. \tag{2.2.4}$$

It is then tempting to calculate the forward-looking (i.e., conditional) skewness of $R_{t,h}$ by simply replacing $M(u,w,h)$ with $M_t(u,w,h;V_t)$ in Eq. (2.2.4). However, as noted by Aretz et al. (2024), since X_t is not continuously observable, we are unable to exactly identify V_t, and thus it is infeasible to calculate $M_t(u,w,h;V_t)$ in practice. To overcome this issue, Aretz et al. (2024) use a proposition from Bates (2006) to devise the optimal estimate $M_{i\Delta}(u,w,h) := \text{E}[M_{i\Delta}(u,w,h;V_{i\Delta})|\mathcal{G}_{i\Delta}]$, where Δ is an equidistant time interval (e.g., a day), i is the sequence of observations, and $\mathcal{G}_{i\Delta}$ is the filtration generated by the observed log stock price X up to time $i\Delta$. To be more specific, let $M_0(u,w,h) = M(u,w,h)$ and use it as the initial input (i.e., $i = 0$) in:

$$G_{i+1}(w|\mathcal{G}_{i+1}) = \frac{\int_{0-i\infty}^{0+i\infty} M_{i}(u,w,h)e^{-ur_{i,1}}du}{\int_{0-i\infty}^{0+i\infty} M_{i}(u,0,h)e^{-ur_{i,1}}du}. \tag{2.2.5}$$

Numerically solving the integrals, we next compute $M_1(u,w,h)$ from:

$$M_{i+1}(u,w,h) = e^{\mu hu + \phi(u,w,h)}G_{i+1}(\psi(u,w,h)|\mathcal{G}_{i+1}). \tag{2.2.6}$$

Iterating on Eqs. (2.2.5) and (2.2.6) until the desired i, we finally replace $M(u,w,h)$ in Eq. (2.2.4) by $M_{i\Delta}(u,w,h)$ to obtain a feasible forward-looking skewness estimate.

Compared to skewness estimators from prior studies, the parametric estimator of Aretz et al. (2024) has several important advantages. First, it is the first estimator to simultaneously yield consistent historical and forward-looking estimates over arbitrary-long return horizons. Second, it does not rely on adhoc exogenous variables used to model skewness. Third, it is flexible enough to account for important return dependencies, such as Black's (1976) leverage effect and volatility clustering. Moreover, while we could have used a Heston (1993) process with independent jumps, the so-called Bates (1996) process, we notice that jumps are rare at low frequencies and that it is markedly more challenging to estimate the parameters of the Bates (1996) process. Given that, we believe that our choice of the Heston (1993) process is a good compromise.

2.2.2 Estimating the Heston Process Parameters

The key challenge in deriving skewness estimates from Aretz et al.'s (2024) estimator in Section 2.2.1 is to estimate the parameters of the Heston (1993) process. To do so, we use the two-step GMM estimator developed in Section 3.1 of Aretz et al. (2024). The basic idea is to contrast the first four central moments and the first two central cross-moments of daily discrete returns, $R_{i,\Delta=1/252}$, with their theoretical counterparts. Defining the centered return, the centered MGF, and the centered cross-MGF as $\tilde{R}_{i,\Delta} := R_{i,\Delta} - M(1,0,\Delta)$, $\tilde{M}(k,\Delta) := \mathrm{E}[\tilde{R}_{i,\Delta}^k]$ and $\tilde{C}(m_1,m_2,\Delta,j) := \mathrm{Cov}[\tilde{R}_{i,\Delta}^{m_1}\tilde{R}_{i+j,\Delta}^{m_2}]$, respectively, we thus use the following GMM moment conditions:

$$\mathrm{E}\begin{pmatrix} \tilde{R}_{i,\Delta} \\ (\tilde{R}_{i,\Delta}^k - \tilde{M}(k,\Delta))'_{k=2:4} \\ (\tilde{R}_{i,\Delta}^1 \tilde{R}_{i+j,\Delta}^2 - \tilde{C}(1,2,\Delta,j))'_{j=1:J} \\ (\tilde{R}_{i,\Delta}^2 \tilde{R}_{i+j,\Delta}^2 - \tilde{C}(2,2,\Delta,j))'_{j=1:J} \end{pmatrix} = 0_{(4+2J)\times 1}, \tag{2.2.7}$$

where $j = 1:J$ is the number of lags, with a maximum of $J = 100$. We notice that the two cross-moments, $\tilde{C}(1,2,\Delta,j)$ and $\tilde{C}(2,2,\Delta,j)$, in Eq. (2.2.7) capture the leverage and volatility clustering effects in stock returns, respectively. While Aretz et al. (2024) derive the theoretical expressions for the centered MGFs and cross-MGFs, we follow them in using the approximated MGFs and cross-MGFs in their Proposition 3, to enhance parameter identification and computational efficiency. As parameter constraints, we first impose the Feller condition (i.e., $2\kappa\alpha > \xi^2$) to ensure a positive variance. Next, we require $\sqrt{(\kappa - 4\xi\rho)^2 - 12\xi^2} + \kappa + 4\xi\rho > 0$ and $\kappa > \xi\rho$ to guarantee the existence of the fourth moment (see Keller-Ressel (2011)). Also, we impose $\mu \in [-1, 1]$, $\kappa \in (0, 10]$, $\alpha \in (0, 1]$, $\xi \in (0, 1.5]$, and $\rho \in [-1, 1]$ to obtain reasonable estimates for the stock drift, the variance mean-reversion speed, the long-run variance, the volatility of variance, and the value-variance correlation, respectively.

We estimate the historical and forward-looking single-stock skewness over rolling windows containing ten years of daily data and ending at the end of sample month t. We exclude a stock if it has less than five years of consecutive daily returns up to the end of the window. We require a long window with consecutive returns since we would else not be able to calculate the cross-moments in the GMM system in Eq. (2.2.7).

2.3 The Competing Skewness Estimators

In this section, we briefly review eight alternative skewness estimators from the literature. We first discuss four direct estimators, including the sample skewness, Fama and French's (2018) i.i.d.-return simple bootstrap estimator, Neuberger and Payne's (2021) approximated closed-form estimator, and Farago and Hjalmarsson's (2023) i.i.d.-return closed-form estimator. We then discuss more indirect estimators, mostly formed from firm fundamental, including Boyer et al.'s (2010) least-squares forecast, Bali et al.'s (2011) maximum return, Conrad et al.'s (2014) logit-model forecast, and Aretz and Arisoy's (2023) quantile-regression forecast.

2.3.1 Direct Skewness Estimators

2.3.1.1 Sample Skewness

The sample skewness of an asset's discrete return over horizon h is:

$$\mathrm{Skew}[R_{i,h}] = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(R_{i,h} \quad \frac{1}{N}\sum_{i=1}^{N}R_{i,h}\right)^{3}}{\left(\frac{1}{N}\sum_{i=1}^{N}\left(R_{i,h} \quad \frac{1}{N}\sum_{i=1}^{N}R_{i,h}\right)^{2}\right)^{3/2}}, \tag{2.3.1}$$

where $R_{i,h}$ is the discrete return from time ih to $ih+h$ and N is the number of observations.

2.3.1.2 Fama and French's (2018) bootstrapped Skewness

Assuming that short-horizon returns are i.i.d., Fama and French (2018) propose a simple bootstrap estimator. To estimate the skewness of returns over horizon h, they first randomly sample a number of shorter-horizon returns over equidistant interval Δ from some estimation window longer than horizon h with equal probabilities and replacement. They next compound the sampled returns to create an artificial return over horizon h. Repeating this process a large number of times, they finally apply the sample skewness to the artificial returns to estimate the skewness of returns over horizon h. While Aretz and Arisoy (2023) show how to use a block bootstrap to extend Fama and French's (2018) estimator to allow for dependence in returns, we exclusively focus on the simple-bootstrap estimator in our work.

2.3.1.3 Neuberger and Payne's (2021) Closed-Form Skewness

Assuming that an asset's price follows a strongly stationary martingale process and relying on approximations of the unconditional moments, Neuberger and Payne (2021) derive a closed-form solution for the skewness of log returns over horizon h, expressed as:

$$\text{Skew}[\ln R_{i,h}] = \left(\text{Skew}[\ln R_{i,}\ \ ] + 3\frac{\text{Cov}[y^{(1)}, x^{(2E)}(\ln R_{i,}\ \ )]}{\text{Var}[\ln R_{i,}\ \ ]^{3/2}} \right) / \sqrt{\frac{h}{\Delta}}, \tag{2.3.2}$$

where:

$$\begin{aligned} &\text{Skew}[\ln R_{i,}\ \ ] = 6((R_{i,}\ \ + 1)\ln R_{i,}\ \ \ \ 2(R_{i,}\ \ \ \ 1))/\,\text{Var}[\ln R_{i,}\ \ ]^{3/2}, \\ &\text{Var}[\ln R_{i,}\ \ ] = 2(R_{i,}\ \ \ \ 1\ \ \ \ln R_{i,}\ \ ), \quad y_i^{(1)} = \sum_{h^0=1}^{h/\ \ \ 1} (R_{i\ \ 1,h^0}\ \ \ \ 1)/\sqrt{\frac{h}{\Delta}}, \\ &x^{(2E)}(\ln R_{i,}\ \ ) = 2(R_{i,}\ \ \ln R_{i,}\ \ \ \ R_{i,}\ \ + 1), \end{aligned} \tag{2.3.3}$$

and, as in Section 2.3.1.2, Δ is an equidistant interval shorter than h (e.g., a day).

2.3.1.4 Farago and Hjalmarsson's (2023) Closed-Form Skewness

Assuming that an asset's return over equidistant interval Δ is i.i.d., Farago and Hjalmarsson (2023) illustrate that the skewness of returns over the longer horizon h is:

$$\text{Skew}\left[R_{i,h}\right] = \text{Skew}\left[\prod_{i=1}^{N} R_{i,}\ \ \right] = \frac{\theta_3^N\ \ \ 3\theta_2^N + 2}{(\theta_2^N\ \ \ 1)^{3/2}}, \tag{2.3.4}$$

where $\theta_2 = \frac{\text{Var}[R_{i,\Delta}]}{\text{E}[R_{i,\Delta}]^2} + 1$, $\theta_3 = \ \ 2 + 3\theta_2 + \text{Skew}[R_{i,}\ \ ](\theta_2\ \ \ 1)^{3/2}$ and N is the number of returns over interval Δ spanning the h period. We note that this estimator is essentially the closed-form equivalent of Fama and French's (2018) simple bootstrap estimator.

2.3.2 Indirect Skewness Estimators

2.3.2.1 Boyer et al.'s (2010) Least-Squares Forecast

Boyer et al. (2010) compute an estimate of the expected skewness of the daily discrete return of asset a from the following cross-sectional regression:

$$Skew_{a,t} = \alpha + \beta_{1,t}\ Skew_{a,t-60} + \beta_{2,t}\ Vol_{a,t-60} + \boldsymbol{\gamma}_t'\ \boldsymbol{X}_{a,t-60}^{BMV} + \epsilon_{a,t}, \tag{2.3.5}$$

where $Skew_{a,t}$ and $Vol_{a,t}$ are the sample skewness and standard deviation, respectively, of the daily discrete return of asset a, estimated from the first day of month $t-60+1$ up until the end of month t. $\boldsymbol{X}^{BMV}$ is a vector of firm characteristics, including momentum, share turnover, small and medium-sized firm dummies, industry dummies, and an exchange dummy. To estimate the expected skewness at the end of month t, they first measure all independent variables for each asset at both month $t-60$ and month t. They then regress month t sample skewness on month $t-60$ independent variables to obtain the coefficients. Finally, they combine the coefficient estimates with the month t independent variables to produce the estimate.

2.3.2.2 Bali et al's (2011) Maximum Return

Bali et al. (2011) use an asset's maximum daily return over the prior month to evaluate the asset's likelihood of yielding a lottery-like payoff. While the estimator is not meant to directly reflect skewness, the general consensus in the literature is that it does so indirectly.

2.3.2.3 Conrad et al.'s (2014) Logit-Model Forecast

Conrad et al. (2014) employ the following logit model to predict the probability of an asset's one-year-ahead log return exceeding 100% (which equals a discrete return of 170%):

$$P_{t-1}(Jackpot_{a,t,t+12} = 1) = \frac{\exp(\alpha + \boldsymbol{\beta}'\boldsymbol{X}^{CKX}_{a,t-1})}{1 + \exp(\alpha + \boldsymbol{\beta}'\boldsymbol{X}^{CKX}_{a,t-1})}, \qquad (2.3.6)$$

where $Jackpot_{a,t,t+12}$ is a dummy variable equal to one if an asset's return from start July of year $t-1$ to end June of year t exceeds the above threshold and else zero. The independent variable vector $\boldsymbol{X}^{CKX}$ includes the momentum return, the sample volatility and skewness of the daily log return computed over the past three months, share turnover detrended over the last five years, size, firm age, asset tangibility, and sales growth. They first estimate the coefficients of the model using only June observations over the past 20 years. They subsequently construct out-of-sample estimates of the probabilities for the next twelve months.

2.3.2.4 Aretz and Arisoy's (2023) Quantile-Regression Forecast

Motivated by Ghysels et al. (2016), Aretz and Arisoy (2023) use panel data quantile regressions to construct a skewness estimator for the annual discrete return. They first rely on monthly data over the prior 20 years to fit the first, fifth, tenth, 25th, 50th, 75th, 90th, 95th, and 99th quantiles of the annual return with independent variables measured until the start of the return horizon. The independent variables include firm age, asset tangibility, sales growth, the book-to-market ratio, share issuances, asset growth, profitability, and those used by Boyer et al. (2010) (although they replace the size dummies with a continuous variable). Having estimated the coefficients, they combine these with the independent variables at the end of the estimation window to produce forecasts of the quantiles of annual returns.

To convert the forecasts into a skewness estimate, Aretz and Arisoy (2023) next posit that the discrete return distribution of an asset is uniform between two consecutive quantiles and ignore density mass outside the top-to-bottom quantile range. They then compute the first three conditional moments of an asset's annual returns from:

$$\mathrm{E}[R_{i,h} | q_{\tau_{(j-1)}} \leq R_{i,h} < q_{\tau_{(j)}}] = \frac{q_{\tau_{(j-1)}} + q_{\tau_{(j)}}}{2}, \tag{2.3.7}$$

$$\mathrm{E}[R^2_{i,h} | q_{\tau_{(j-1)}} \leq R_{i,h} < q_{\tau_{(j)}}] = \frac{q^2_{\tau_{(j-1)}} + q_{\tau_{(j-1)}} q_{\tau_{(j)}} + q^2_{\tau_{(j)}}}{3}, \tag{2.3.8}$$

$$\mathrm{E}[R^3_{i,h} | q_{\tau_{(j-1)}} \leq R_{i,h} < q_{\tau_{(j)}}] = \frac{q^3_{\tau_{(j-1)}} + q^2_{\tau_{(j-1)}} q_{\tau_{(j)}} + q_{\tau_{(j-1)}} q^2_{\tau_{(j)}} + q^3_{\tau_{(j)}}}{4}, \tag{2.3.9}$$

where $q_{\tau_{(j)}}$ is the τth quantile forecast of $R_{i,h}$, with $j = \{1, 2, ..., J\}$ indexing the quantile forecasts in ascending order, and J is the number of estimated quantiles. They finally estimate the corresponding unconditional moments from the law of total probability:

$$\mathrm{E}[R^n_{i,h}] = \sum_{j=2}^{J} \frac{F^{-1}(q_{\tau_{(j)}}) - F^{-1}(q_{\tau_{(j-1)}})}{F^{-1}(q_{\tau_{(J)}}) - F^{-1}(q_{\tau_{(1)}})} \mathrm{E}[R^n_{i,h} | q_{\tau_{(j-1)}} \leq R_{i,h} < q_{\tau_{(j)}}], \tag{2.3.10}$$

where $n \in \{1, 2, 3\}$. They finally plug $\mathrm{E}[R^n_{i,h}]$ into the (expanded) definition of skewness to obtain a skewness estimate for the asset's annual returns:

$$Skew = \frac{\mathrm{E}[R^3_{i,h}] - 3\,\mathrm{E}[R_{i,h}](\mathrm{E}[R^2_{i,h}] - \mathrm{E}[R_{i,h}]^2) - \mathrm{E}[R_{i,h}]^3}{(\mathrm{E}[R^2_{i,h}] - \mathrm{E}[R_{i,h}]^2)^{3/2}}. \tag{2.3.11}$$

2.4 Data Sources

We collect stock data from CRSP, accounting data from Compustat, and data on the Fama-French (FF) benchmark factors, FF industry portfolios, and the risk-free rate of return from Kenneth French's website.[2] We focus on common stocks (share codes: 10 and 11) traded on the NYSE, AMEX, and Nasdaq, dropping stocks from highly regulated (SIC

[2]The URL address for Kenneth French's website is:
<http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>.

codes: 4000-4999) and financial (6000-6799) industries. To mitigate microstructure biases, we exclude stocks worth less than $5 at the end of the skewness estimation window. To match the stock and accounting data, we use the accounting data from the fiscal year in calendar year t 1 over the period from June of calendar year t to May of calendar year $t + 1$. Our sample period is June 1973 to December 2020.

2.5 Properties of Our Single-Stock Skewness Estimates

In this section, we take an initial look at the historical and forward-looking single-stock skewness estimates obtained from Aretz et al.'s (2024) estimator. We start by offering descriptive statistics for those estimates over return horizons from one day to five years and validating that our estimates outperform those obtained from the alternative direct estimators in Section 2.3.1 in capturing future sample skewness. We next study how the estimates relate to the return horizon and which types of stocks are more likely to produce an explosive skewness. We finally contrast the historical and forward-looking skewness estimates over potentially mismeasured return horizons and identify factors driving differences between those two types of skewness.

2.5.1 Descriptive Statistics and the Validity of Estimates

In Table 2.1, we offer descriptive statistics for the historical (Panel A) and forward-looking (Panel B) single-stock skewness estimates calculated over the daily, weekly, monthly, quarterly, annual, three-year, and five-year return horizons. We compute the descriptive statistics first by cross-section and then average over our sample period. The table shows that both types of skewness are positive on average and tend to rise with the return horizon. Looking into the historical estimates, Panel A, for example, reports that the median stock yields a skewness of 0.12, 0.53, 1.56, and 4.64 over the daily, monthly, annual, and five-year horizons, respectively. Contrasting the historical and forward-looking estimates, the panels suggest that historical skewness is, on median, slightly more positive than forward-looking skewness over all horizons. The difference in the medians between historical and forward-looking skewness is, for example, 0.06 (0.12–0.06) over the daily horizon, which is identical to that over the five-year horizon (4.64–4.58).

Notwithstanding, the table also suggests that skewness does not always rise with the return horizon. To wit, both the historical and forward-looking skewness of the 5th-percentile stock initially drop from the daily to the monthly return horizon, before then rising from the monthly to the five-year horizon. Considering the forward-looking estimates in Panel B, we, for example, see that the skewness of the 5th-percentile stock initially falls from 0.10 over the daily to 0.25 over the monthly horizon, before then rising to 1.36 over the five-year horizon. The table also reveals that both the historical and forward-looking skewness of more than five percent of stocks ***explode*** with the return horizon, with values surpassing 1,000 over the five-year horizon.

We next verify that our skewness estimates are more accurate than those obtained from the alternative direct estimators in Section 2.3.1. Doing so is critical since if they were not, it would make little sense to look into their empirical properties. To facilitate our analysis, we run cross-sectional regressions of future sample skewness computed from daily data over the three years after the end of month t on either our historical or forward-looking estimates, sample skewness, or Fama and French's (2018), Neuberger and Payne's (2021), and Farago and Hjalmarsson's (2023) estimates computed from daily data until the end of that same month (i.e., out-of-sample analyses). We calculate the future sample skewness estimates over only the daily, weekly, and monthly return horizons since Farago and Hjalmarsson (2023) show that those estimates become increasingly biased over longer horizons (see their Figure 2). The numbers of observations available for computing future sample skewness over those horizons are (approximately) 750, 150, and 36, respectively.

Table 2.2 gives the predictive performance of our estimates and its competitors, with Panels A, B, and C focusing on the daily, weekly, and monthly future sample skewness, respectively. While the plain numbers in the table are time-series averages, those in parentheses are Newey-West (1987) t-statistics calculated with a 36-month lag length. The table demonstrates that our estimates outperform the others, with them always yielding the most significant slope coefficients and the highest mean R-squareds. Looking into the daily horizon, the t-statistic of the slope coefficient of our historical estimates ("HH") and its mean adjusted R-squared are, for example, at least about two and one-and-a-half times those of the other estimators, respectively. Remarkably, however, our historical estimates consistently slightly outperform our forward-looking estimates ("HF"), presumably because the three-year windows which we use to calculate future sample skewness yield closer-to-unconditional rather than conditional skewness estimates.

2.5.2 The Skewness-Return Horizon Relation and Explosive Skewness

Our evidence in Table 2.1 raises the interesting questions of whether we can identify (i) systematic variations in the shape of the skewness-return horizon relation and/or (ii) stocks more likely to produce an explosive skewness. To address the first question, Table 2.3 reports the results from a panel logit regression of a dummy variable equal to one if a stock's skewness at the end of month t initially drops and only later rises (i.e., forms a U-shaped pattern) with the return horizon and else zero, on a comprehensive set of predictor variables measured at that time. While the first three columns look into historical skewness, the next three consider forward-looking skewness. The predictors are $MarketBeta$, $Size$, $BookToMarket$, $Momentum$, $AssetGrowth$, $Profitability$, $HistoricalVolatility$, $ShareTurnover$, $AssetTangibility$, $CompanyAge$, $Leverage$, $SalesGrowth$, *ShareIssuance*, and an $NBER$ recession indicator. See Table 2.A.1 in the appendix for the predictors' definitions. Plain numbers are estimates; those in parentheses are z-scores. To better gauge economic significance, the table also reports the effect on the log odds ratio from raising an exogenous variable from its sample mean to one-standard-deviation above that mean. As Aretz et al. (2024) explain, while a more negative stock return-volatility correlation contributes to creating a U-shaped skewness-return horizon relation, a higher stock volatility works against it by boosting the compounding effect (Bessembinder (2018)).

Table 2.3 shows that $MarketBeta$, $Size$, $BookToMarket$, $Profitability$, and $Leverage$ positively affect the likelihood of a stock exhibiting a U-shaped historical skewness-return horizon relation, while $AssetGrowth$, $HistoricalVolatility$, $AssetTangibility$, $SalesGrowth$, and the $NBER$ recession indicator negatively impact that likelihood. The positive $Leverage$ effect aligns with Black's (1976) original hypothesis that a higher financial leverage creates a stronger leverage effect (i.e., a more negative stock price-volatility relation), thus making a U-shaped skewness-return horizon relation more likely. The negative $HistoricalVolatility$ effect further supports the conjecture that the compounding effect works against finding a U-shaped relation between skewness and the return horizon. The negative effect of the $NBER$ recession indicator may be partially explained by the fact that stock volatility is countercyclical (i.e., rises during recessions). Looking into forward-looking skewness, we reach similar conclusions. Despite that, *CompanyAge* and *ShareTurnover* now reduce the likelihood of a U-shaped relation, while the effect of $Momentum$ switches from negative to positive. Importantly, the change in the odds ratio indicates that, among all predictors, the likelihood of observing a U-shaped pattern is most sensitive to changes in size. In particular, a one-standard-deviation increase in size raises the log odds ratio by a whooping 160.30 (212.60) for historical (forward-looking) skewness.

Turning to the second question, Table 2.4 presents the results from a panel logit regression similar to that of Table 2.3, with the only difference being that the dependent dummy variable now equals one if a stock's five-year (either historical or forward-looking) skewness at the end of month t exceeds 1,000 and else zero. The table shows that stocks with a high $Momentum$, $AssetGrowth$, and $HistoricalVolatility$ and a low $MarketBeta$, $Size$, $BookToMarket$, $Profitability$, and $Leverage$ are more likely to exhibit an explosive skewness. As in Table 2.3, the effects of $HistoricalVolatility$ and the $NBER$ recession indicator have the same sign. Rather unsurprisingly, the tendency to produce an explosive skewness is particularly sensitive to changes in $HistoricalVolatility$, with a one-standard-deviation increase in it raising the log odds ratio by more than 350 for both historical and forward-looking skewness. The sensitivities of that tendency to changes in $Size$ or $BookToMarket$, are also high, with the effects on the log odds ratio being at least three and two times the magnitude of those of the others, respectively.

2.5.3 Historical versus Forward-Looking Skewness

We next examine how well our historical skewness estimates explain our forward-looking skewness estimates. We do so since many studies (as, e.g., Bali et al. (2011) and Amaya et al. (2015)) use historical skewness measures as proxies for forward-looking skewness. To that end, Table 2.5 offers the results from cross-sectional regressions of each of our forward-looking skewness estimates over the seven return horizons on the corresponding historical skewness estimates. Plain numbers are time-series averages, while those in parentheses are Newey-West (1987) t-statistics.

The first column of Table 2.5 suggests that daily historical skewness is a biased and imperfect estimate of daily forward-looking skewness. More specifically, its mean constant estimate (i.e., bias) is 0.06 (t-statistic: 7.56), and its mean adjusted R-squared is 63.10%. Interestingly, however, the bias in the historical estimate vanishes and its precision moves toward 100% with increases in the return horizon. The last three columns of the table, for example, reveal that the mean constant estimate is no longer significant from the annual horizon onward, and that the mean adjusted R-squared rises from 95.76% at the annual to 99.68% at the five-year horizon. The reason may be that the dependence on state variable, asset volatility, weakens as the horizon expands, causing the historical and forward-looking estimates to align. The upshot is that although we can likely use long-horizon historical skewness measures to proxy for long-horizon forward-looking skewness, we should not do the same over shorter horizons.

In Table 2.6, we study the determinants of the absolute difference between the two types of skewness over shorter horizons (i.e., daily, weekly, monthly, and quarterly). To do so, we run panel regressions of those differences on the same set of predictors used in Tables 2.3 and 2.4 and firm fixed effects. The table suggests that an increase in $MarketBeta$, $Momentum$, $Profitability$, $HistoricalVolatility$, and $CompanyAge$ widens the absolute difference across horizons, while an increase in $Size$, $ShareTurnover$, $AssetTangibility$, $Leverage$, and $SalesGrowth$ reduces it. The table also indicates that this difference depends on the economic state, with it being larger during economic expansions. Looking into the monthly horizon, we, for example, find that the absolute difference is 0.02 (t-statistic: 14.27) lower in NBER recession periods.

In Table 2.7, we repeat the cross-sectional regressions of forward-looking skewness on historical skewness from Table 2.5, this time, however, also allowing for a return horizon mismatch between the skewness variables. Our reason for doing so is that some studies, as, for example, Boyer et al. (2010), implicitly use skewness proxies over short horizons, like the daily, to make inferences about longer-horizon skewness. In line with intuition, the table demonstrates that a greater mismatch in the return horizons across the two types of skewness further deteriorates the ability of historical skewness to capture forward-looking skewness. While the mean adjusted R-squared obtained from the regression of annual forward-looking skewness on quarterly historical skewness is still 67.91%, that mean adjusted R-squared obtained from the regression of five-year forward-looking skewness on daily historical skewness is a mere 1.22%. The implication is that it is not advisable to use short-horizon historical skewness to capture long-horizon forward-looking skewness.

2.6 The Relations of Our Skewness Estimates with the Literature

In this section, we assess how our single-stock historical and forward-looking skewness estimates relate to firm fundamentals conjectured to reflect skewness in the literature. We also compare our estimates with those from the indirect skewness estimators, mostly formed from the fundamentals.

2.6.1 Our Skewness Estimates and Firm Fundamentals

While prior studies (as, e.g., Boyer et al. (2010), Conrad et al. (2014), and Aretz and Arisoy (2023)) use various firm fundamentals (including, e.g., $Size$, $BookToMarket$, $Momentum$, $AssetGrowth$, $Profitability$, $HistoricalVolatility$, $ShareTurnover$, *AssetTangibility*, $CompanyAge$, $SalesGrowth$, and $ShareIssuance$) to model or predict skewness, only $Momentum$, $HistoricalVolatility$ and $ShareTurnover$ are motivated by theoretical work (see Chen et al. (2001), Bessembinder (2018), and Hong and Stein (2003)). Given that, we now evaluate how our historical and forward-looking skewness estimates are associated with these firm fundamentals.

Table 2.8 provides the results from cross-sectional regressions of our historical (Panel A) and forward-looking (Panel B) skewness estimates over seven return horizons on the firm fundamentals. Plain numbers are time-series means and those in parentheses are their Newey and West (1987) t-statistics computed using a 120-month lag length. The table shows that only three of the 13 firm fundamentals, $MarketBeta$, $AssetGrowth$, and $HistoricalVolatility$, significantly explain our skewness estimates over all return horizons. In contrast, $Size$, $BookToMarket$, $Momentum$, $Profitability$, $AssetTangibility$, and $SalesGrowth$ only reflect skewness over short or moderately long horizons, while $Leverage$ and $ShareIssuance$ never capture it. Interestingly, the explanatory power of $ShareTurnover$ and $CompanyAge$ is primarily concentrated in our forward-looking (rather than historical) skewness estimates. Overall, our selected firm fundamentals better reflect historical skewness than forward-looking skewness, except over long horizons over which capturing either becomes challenging. While the mean adjusted R-squared of the regression modelling daily historical skewness is, for example, 42.90%, the mean adjusted R-squared of the regression modelling daily forward-looking skewness is a markedely lower 24.93%.

2.6.2 Comparison with Indirect Skewness Estimators

We next compare our historical and forward-looking skewness estimates with those obtained from the indirect skewness estimators in Section 2.3.2. To do so, we run cross-sectional regressions of our estimates over each of our return horizons on the (return-horizon-independent) estimates from the indirect estimators. To only study the cross-sectional correlation between them, we use rank-variable transformations as dependent and independent variable in these regressions.

Table 2.9 presents the regression results, with Panels A and B focusing on our historical and forward-looking skewness estimates, respectively. Plain numbers are mean estimates, while those in parentheses are Newey-West (1987) t-statistics. The table suggests that only Boyer et al.'s (2010) and Aretz and Arisoy's (2023) estimates moderately align with ours. To be specific, Panel A, for example, shows that the estimates from Boyer et al.'s

(2010) and Aretz and Arisoy's (2023) estimators exhibit correlations (i.e., coefficients) of around 0.60 and 0.40 with our historical estimates, respectively, with the R-squared ranging from 31% to 35% for the former and 27% to 31% for the latter. In contrast, the connection between Bali et al.'s (2011) estimates and ours is much weaker, with its R-squared never exceeding 17%. Strikingly, Conrad et al.'s (2014) estimates almost reflect nothing in common with ours, with both the coefficient and R-squared consistently close to zero. Comparing the explanatory power of these indirect skewness estimators across Panels A and B, we find it interesting that those designed to capture future skewness (i.e., Boyer et al.'s (2010), Conrad et al.'s (2014), and Aretz and Arisoy's (2023) estimators) connect more strongly with our historical estimates, while Bali et al.'s (2011) backward-looking proxy performs better in explaining our forward-looking skewness estimates.

2.7 The Term Structure of the Skewness Premium

In this section, we assess the term structure of the forward-looking skewness premium. In doing so, we first employ partial least squares (PLS) regressions of our forward-looking skewness estimates on our set of firm fundamentals to isolate the components of skewness not reflecting the pricing of other stock return determinants. We next use univariate portfolio sorts and Fama-MacBeth (FM; 1973) regressions to derive the term structure of the orthogonalized skewness premium.

2.7.1 Orthogonalizing Our Skewness Estimates

We use PLS regressions to orthogonalize our forward-looking skewness estimates with respect to the set of firm fundamentals used in our other tests, mitigating concerns that the pricing of our skewness estimates is attributable to these fundamentals. Addressing that concern head on is critical since Table 2.8 suggests that the firm fundamentals often explain a significant fraction of the variations in our forward-looking skewness estimates, with them, for example, explaining 46.96% of those variations over the annual horizon. In turn, Hou et al. (2015), Green et al. (2017), and many others offer evidence that virtually all of our firm fundamentals price stock returns. Our reason for using PLS regressions to orthogonalize our skewness estimates is that they handle multicollinearity more effectively than ordinary least squares regressions. They do so by projecting both the response and predictor variables into a new space in which the projected predictors are orthogonal to one another, leading the regression estimates to be more robust. We set the number of PLS components to 13 (matching the number of firm fundamentals) and refer to the PLS residual from our regressions as "orthogonalized skewness" from here on.

2.7.2 The Pricing of Skewness Over Different Return Horizons

We start with portfolio sorts to see how our orthogonalized skewness estimates are priced over different return horizons. At the end of each sample month t 1, we thus sort our sample stocks into ten portfolios based on the decile breakpoints of our orthogonalized skewness over each of the horizons. We value-weight the portfolios and hold them over month t. We then create a spread portfolio by taking a long position in the top portfolio and a short position in the bottom portfolio. To adjust for risk, we regress excess portfolio returns on the CAPM and the Fama-French three-factor model (FF3) and report the intercepts (CAPM or FF3 alpha). We finally calculate Newey-West (1987) t-statistics with a twelve-month lag length for the mean returns and alphas of the spread portfolios.

Table 2.10 reports our portfolio sort results. While the plain numbers are mean excess returns or alphas, those in paretheses are t-statistics. The table suggests that the mean portfolio returns drop significantly over the portfolios formed from the daily, weekly, monthly, quarterly, and annual skewness variables, but not over those formed from the three and five-year skewness variables. Notwithstanding, the declines are mostly driven by the initial and last three portfolios, but not by the four middle portfolios. In accordance, only the daily to annual skewness spread portfolios produce significantly negative mean returns and alphas. For example, the mean return, the CAPM alpha, and the FF3 alpha of the daily skewness spread portfolio are 0.21%, 0.22%, and 0.20% (t-statistics: 2.35, 2.50, and 2.37), all respectively. These results imply that investors are primarily concerned with short-term (e.g., daily) to medium-term (e.g., annual) discrete-return skewness, with their preferences varying across return horizons.

In Table 2.11, we complement our portfolio sorts with FM regressions, where we project the single-stock excess return over month t on each of the skewness variables as well as $MarketBeta$, $Size$, $BookToMarket$, $Momentum$, $AssetGrowth$, and $Profitability$. As always, plain numbers are estimates, while those in parentheses are Newey and West (1987) t-statistics with a twelve-month lag length. Due to the limited number of return horizons used in the portfolio sorts, it is hard to visualize how the skewness premium evolves over the return horizon. As a result, we increase the number of horizons to 60 in the FM regressions, starting with the monthly horizon and going to the 60-month horizon in one-month increments. Our evidence suggests that skewness is significantly negatively priced up to the 30-month horizon. Plotting our regression outcomes, Figure 2.1 shows that the skewness premium becomes more negative up to the nine-month horizon, before becoming more positive again over longer horizons.

2.8 Conclusion

We offer a comprehensive study of the skewness of discrete single-stock returns. Using the new skewness estimator of Aretz et al. (2024), we first confirm that our estimates easily beat those obtained from other recent competing estimators. We next reveal that skewness does not always rise with the return horizon but can sometimes form a U-shaped with that horizon, and that about five percent of stocks have a five-year skewness in excess of 1,000. Crucially, we identify several well-known firm characteristics able to predict either a U-shaped pattern or an explosive skewness with some precision. We then demonstrate that the spread between historical and forward-looking skewness can be large over shorter horizons, especially in economic expansions and when there is a mismatch in the return horizons. We also find that more than half of all the predictor variables used in prior studies modelling skewness fail to consistently produce a strong association with skewness. To assess the pricing power of skewness, we first orthogonalize our forward-looking skewness estimates with respect to the firm fundamentals well-known to price stocks. Using the orthogonalized estimates in asset pricing tests, we find that the skewness premium is significantly negative up to about two-and-a-half years. In particular, we establish that the skewness premium initially becomes more negative with the return horizon before eventually rising up toward zero, reaching its trough over the close-to-nine-month horizon.

Table 2.1: Descriptive Statistics

This table reports descriptive statistics including the average number of observations ($\overline{\text{Obs}}$), the mean, the standard deviation (StD), the fifth percentile (Pct5), the first quartile (Q1), the median, the third quartile (Q3), and the 95th percentile (Pct95). While Panel A focuses on our historical skewness estimates for daily, weekly, monthly, quarterly, annual, three-year, and five-year discrete returns, Panel B concentrates on our forward-looking skewness estimates. For each stock at month t, we estimate its historical and forward-looking skewness by applying Aretz et al.'s (2024) skewness estimator to daily discrete returns over a ten-year estimation window ending with month t. The sample period spans from June 1973 to December 2020.

	$\overline{\text{Obs}}$	Mean	StD	Pct5	Q1	Median	Q3	Pct95
				Panel A: Historical Skewness				
Daily	1,622	0.12	0.09	0.03	0.05	0.12	0.18	0.28
Weekly	1,622	0.26	0.21	0.06	0.12	0.26	0.39	0.62
Monthly	1,622	0.53	0.40	0.09	0.25	0.53	0.78	1.23
Quarterly	1,622	0.90	0.63	0.04	0.43	0.86	1.28	2.06
Annual	1,622	1.99	1.91	0.33	0.89	1.56	2.46	5.04
Three-year	1,622	32.57	317.39	0.95	1.71	2.94	5.72	49.76
Five-year	1,622	9.74 10^3	1.74 10^5	1.40	2.46	4.64	12.62	1.17 10^3
				Panel B: Forward-Looking Skewness				
Daily	1,622	0.07	0.13	0.10	0.01	0.06	0.13	0.27
Weekly	1,622	0.17	0.25	0.19	0.02	0.15	0.30	0.59
Monthly	1,622	0.36	0.42	0.25	0.09	0.33	0.61	1.13
Quarterly	1,622	0.71	0.63	0.17	0.27	0.64	1.07	1.92
Annual	1,622	1.84	1.83	0.25	0.80	1.41	2.29	4.86
Three-year	1,622	30.34	296.21	0.89	1.66	2.86	5.57	47.33
Five-year	1,622	9.06 10^3	1.64 10^5	1.36	2.42	4.58	12.33	1.12 10^3

Table 2.2: The Validity of Estimates

This table presents the results from the cross-sectional regression:

$$\textit{Future Discrete-Return Skewness}_{t,t+36} = \alpha + \beta \quad \textit{Skewness Estimate}_t.$$

We compute the future discrete-return skewness of a stock at month t as the sample skewness of non-overlapping returns within the three-year window immediately following month t. The examined skewness estimates include Aretz et al.'s (2024) historical ("HH") and forward-looking ("HF"), the sample skewness ("SS"), Fama and French's (2018; "FF"), Neuberger and Payne's (2021; "NP"), and Farago and Hjalmarsson's (2023; "FH") estimates. See Section 2.2 and 2.3 for implementation details. We only evaluate daily, weekly, and monthly estimates due to insufficient data for computing future skewness. Our sample month t covers the period from June 1973 to December 2020. Coefficients, Constants, and adjusted R-squareds are time-series averages. T-statistics in parentheses are calculated using Newey-West (1987) standard errors with a 36-month lag length.

	Heston Historical (HH)	Heston Forward (HF)	Sample Skewness (SS)	Fama-French (FF)	Neuberger-Payne (NP)	Farago-Hjalmarsson (FH)
			Panel A: Daily			
Coefficient	2.41	1.73	0.21	0.19	0.19	0.21
	(17.11)	(14.48)	(8.28)	(8.28)	(6.35)	(8.28)
Constant	0.11	0.25	0.27	0.28	0.31	0.27
	(2.31)	(6.33)	(6.52)	(6.77)	(7.38)	(6.52)
Adj. R^2	3.85 %	3.21 %	2.58 %	2.35 %	1.77 %	2.58 %
			Panel B: Weekly			
Coefficient	0.97	0.82	0.24	0.38	0.24	0.39
	(19.99)	(17.12)	(8.98)	(8.82)	(10.71)	(8.66)
Constant	0.10	0.21	0.20	0.20	0.32	0.19
	(2.18)	(4.66)	(4.96)	(4.73)	(7.18)	(4.60)
Adj. R^2	5.98 %	5.31 %	5.15 %	3.79 %	2.78 %	3.93 %
			Panel C: Monthly			
Coefficient	0.35	0.35	0.16	0.41	0.20	0.42
	(16.00)	(16.08)	(10.31)	(9.05)	(14.61)	(8.66)
Constant	0.11	0.16	0.19	0.08	0.31	0.08
	(2.23)	(3.42)	(4.49)	(1.71)	(7.16)	(1.56)
Adj. R^2	4.24 %	4.22 %	3.01 %	3.45 %	2.23 %	3.57 %

Table 2.3: The Skewness-Return Horizon Relation

This table displays the results from the panel logit regression:

$$D^1_{s,t} = \frac{\exp(\alpha + \beta \quad F_{s,t})}{1 + \exp(\alpha + \beta \quad F_{s,t})},$$

where $D^1 = 1$ if the skewness of a stock s at the end of month t initially drops and only later rises with the return horizon, and zero otherwise. F represents a vector of predictor variables (see Table 2.A.1 for descriptions). Both historical and forward-looking skewness estimates for the stock s at the end of month t are obtained from applying Aretz et al.'s (2024) estimator to a ten-year estimation window of daily discrete returns ending with month t. The sample period spans from June 1973 to December 2020. Z-scores and percent (%) changes in odds ratio are computed using cluster-robust standard errors, clustered by stock.

	Panel A: Historical Skewness			Panel B: Forward-Looking Skewness		
	Coefficient	Z-score	% change in odds ratio for a 1σ change in X	Coefficient	Z-score	% change in odds ratio for a 1σ change in X
MarketBeta	0.23	(5.04)	14.00	0.24	(9.72)	14.60
Size	0.47	(22.75)	160.30	0.56	(40.85)	212.60
BookToMarket	0.22	(5.28)	20.30	0.22	(8.91)	20.20
Momentum	0.87	(22.39)	28.70	0.08	(4.58)	3.30
AssetGrowth	0.58	(7.44)	12.00	0.72	(14.70)	14.50
Pro tability	0.43	(4.10)	13.30	0.53	(8.40)	16.60
HistoricalVolatility	4.82	(20.58)	59.40	0.91	(8.23)	15.60
ShareTurnover	0.01	(2.18)	6.10	0.02	(10.61)	17.80
AssetTangibility	0.77	(9.39)	24.50	0.60	(10.73)	19.60
CompanyAge	0.00	(0.23)	0.70	0.01	(5.62)	11.00
Leverage	0.82	(5.10)	14.80	0.45	(4.54)	7.80
SalesGrowth	0.30	(6.04)	8.30	0.31	(9.12)	8.50
ShareIssuance	0.06	(0.36)	0.70	0.07	(0.65)	0.70
NBER	0.37	(11.05)	11.00	1.08	(36.42)	28.90
Constant	6.46	(22.41)		7.83	(41.55)	
Pseudo R^2	21.29 %			17.93 %		

Table 2.4: The Explosive Skewness Potential

This table provides the results from the panel logit regression:

$$D^2_{s,t} = \frac{\exp(\alpha + \beta \quad F_{s,t})}{1 + \exp(\alpha + \beta \quad F_{s,t})},$$

where $D^2 = 1$ if the five-year skewness of a stock s at the end of month t exceeds 1,000, and zero otherwise. F represents a vector of predictor variables (see Table 2.A.1 for descriptions). Both historical and forward-looking skewness estimates for the stock s at the end of month t are obtained from applying Aretz et al.'s (2024) estimator to a ten-year estimation window of daily discrete returns ending with month t. The sample period spans from June 1973 to December 2020. Z-scores and percent (%) changes in odds ratio are computed using cluster-robust standard errors, clustered by stock.

	Panel A: Historical Skewness			Panel B: Forward-Looking Skewness		
	Coefficient	Z-score	% change in odds ratio for a 1σ change in X	Coefficient	Z-score	% change in odds ratio for a 1σ change in X
MarketBeta	0.12	(2.81)	6.70	0.12	(2.67)	6.50
Size	0.29	(10.70)	45.10	0.29	(10.65)	45.20
BookToMarket	0.43	(11.36)	30.20	0.44	(11.38)	30.40
Momentum	0.20	(5.72)	8.20	0.21	(6.00)	8.70
AssetGrowth	0.51	(7.14)	11.90	0.52	(7.18)	12.00
Pro tability	0.16	(2.54)	4.40	0.16	(2.57)	4.50
HistoricalVolatility	8.03	(43.04)	350.30	8.05	(42.68)	352.20
ShareTurnover	0.00	(1.58)	1.80	0.00	(1.99)	2.50
AssetTangibility	0.28	(2.13)	9.90	0.27	(2.03)	9.50
CompanyAge	0.00	(0.97)	6.60	0.00	(1.04)	7.10
Leverage	0.89	(4.28)	13.90	0.90	(4.29)	14.00
SalesGrowth	0.02	(0.55)	0.60	0.02	(0.45)	0.50
ShareIssuance	0.02	(0.18)	0.20	0.00	(0.02)	0.00
NBER	0.41	(6.72)	13.90	0.43	(7.06)	14.60
Constant	5.10	(14.87)		5.17	(14.94)	
Pseudo R^2	44.28 %			44.59 %		

Table 2.5: Forward-Looking versus Historical Skewness of the Same Return Horizon

This table offers the results from the cross-sectional regression:

$$Forward\text{-}LookingSkew_t = \alpha + \beta \quad HistoricalSkew_t.$$

We obtain both historical ($HistoricalSkew$) and forward-looking ($Forward\text{-}LookingSkew$) skewness estimates for a stock at month t by applying Aretz et al.'s (2024) estimator to daily discrete returns within a ten-year estimation window ending with month t. Our sample period ranges from June 1973 to December 2020. Coefficients, constants, and adjusted R-squareds are time-series averages. T-statistics in parentheses are calculated using Newey-West (1987) standard errors with a 120-month lag length.

	Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
Coefficient	1.10	1.03	0.95	0.92	0.94	0.93	0.93
	(25.54)	(34.22)	(45.94)	(50.62)	(139.02)	(161.05)	(125.88)
Constant	0.06	0.11	0.14	0.12	0.03	0.02	130.22
	(7.56)	(6.85)	(5.30)	(3.57)	(0.73)	(0.11)	(1.65)
Adj. R^2	63.10%	74.56%	81.73%	86.05%	95.76%	99.47%	99.68%

Table 2.6: Variables Explaining the Difference between Forward-Looking and Historical Skewness

This table shows the results from the panel regression:

$$\text{Abs}(Forward\text{-}LookingSkew_{s,t} \quad HistoricalSkew_{s,t}) = \alpha + \beta \quad F_{s,t},$$

where historical ($HistoricalSkew$) and forward-looking ($Forward\text{-}LookingSkew$) skewness estimates for the stock s at month t are both obtained from applying Aretz et al.'s (2024) estimator to a ten-year estimation window of daily discrete returns ending with month t. F represents a vector of predictor variables (see Table 2.A.1 for descriptions). Our sample period ranges from June 1973 to December 2020. T-statistics in parentheses are calculated using bootstrapped standard errors.

	Daily	Weekly	Monthly	Quarterly
MarketBeta	0.00	0.01	0.02	0.02
	(8.78)	(12.49)	(11.36)	(10.85)
Size	0.00	0.01	0.01	0.02
	(5.51)	(6.59)	(6.36)	(6.49)
BookToMarket	0.00	0.00	0.00	0.00
	(1.56)	(2.07)	(1.57)	(0.14)
Momentum	0.01	0.02	0.02	0.02
	(20.99)	(19.70)	(18.26)	(13.74)
AssetGrowth	0.00	0.00	0.01	0.01
	(1.23)	(2.53)	(2.13)	(1.47)
Pro tability	0.01	0.01	0.02	0.02
	(5.46)	(6.89)	(5.67)	(4.95)
HistoricalVolatility	0.00	0.03	0.12	0.26
	(0.01)	(5.07)	(11.06)	(19.79)
ShareTurnover	0.00	0.00	0.00	0.00
	(7.34)	(10.94)	(7.89)	(3.81)
AssetTangibility	0.01	0.03	0.04	0.04
	(7.12)	(5.82)	(6.83)	(4.60)
CompanyAge	0.00	0.00	0.00	0.01
	(19.25)	(25.59)	(23.28)	(20.23)
Leverage	0.01	0.03	0.05	0.07
	(4.10)	(5.76)	(6.52)	(6.63)
SalesGrowth	0.00	0.00	0.00	0.01
	(2.78)	(3.32)	(3.48)	(3.47)
ShareIssuance	0.00	0.00	0.01	0.01
	(0.37)	(1.01)	(1.77)	(1.46)
NBER	0.01	0.01	0.02	0.00
	(18.65)	(22.97)	(14.27)	(2.29)
Firm Fixed Effects	Yes	Yes	Yes	Yes
R^2	1.26%	3.61%	5.23%	5.99%

Table 2.7: Forward-Looking versus Historical Skewness of Unequal Horizons

This table summarizes the results from the cross-sectional regression:

$$Forward\text{-}LookingSkew_t\ (LH) = \alpha + \beta \quad HistoricalSkew_t\ (SH),$$

where LH and SH stand for long horizons (i.e., annual, three-year, and five-year) and short horizons (i.e., daily, weekly, monthly, and quarterly), respectively. For each stock at month t, ranging from June 1973 to December 2020, we apply Aretz et al.'s (2024) estimator to daily discrete returns within a ten-year estimation window ending with month t to obtain both historical ($HistoricalSkew$) and forward-looking ($Forward\text{-}LookingSkew$) skewness estimates. Coefficients, constants, and adjusted R-squareds are time-series averages. T-statistics in parentheses are calculated using Newey-West (1987) standard errors with a 120-month lag length.

	Daily Historical	Weekly Historical	Monthly Historical	Quarterly Historical
Panel A: Annual Forward-Looking				
Coefficient	13.77	6.34	3.42	2.34
	(8.52)	(8.68)	(9.35)	(11.50)
Constant	0.22	0.18	0.02	0.29
	(2.37)	(1.89)	(0.23)	(2.49)
Adj. R^2	50.15%	51.32%	56.03%	67.91%
Panel B: Three-year Forward-Looking				
Coefficient	602.96	280.97	160.07	127.32
	(2.89)	(2.93)	(3.09)	(3.50)
Constant	39.62	42.52	54.72	87.69
	(2.47)	(2.54)	(2.81)	(3.32)
Adj. R^2	3.57%	3.76%	4.63%	7.75%
Panel C: Five-year Forward-Looking				
Coefficient	2.24 10^5	1.04 10^5	5.96 10^4	4.85 10^4
	(2.37)	(2.40)	(2.51)	(2.78)
Constant	1.68 10^4	1.79 10^4	2.26 10^4	3.59 10^4
	(2.24)	(2.28)	(2.45)	(2.78)
Adj. R^2	1.22%	1.30%	1.66%	3.10%

Table 2.8: Our Skewness Estimates and Firm Fundamentals

This table reports the results from the cross-sectional regression:

$$Forward\text{-}LookingSkew_t \text{ or } HistoricalSkew_t = \alpha + \beta \quad F_t.$$

We obtain both historical ($HistoricalSkew$) and forward-looking ($Forward\text{-}LookingSkew$) skewness estimates for a stock at month t by applying Aretz et al.'s (2024) estimator to daily discrete returns within a ten-year estimation window ending with month t. F represents a vector of predictor variables (see Table 2.A.1 for descriptions). The sample period spans from June 1973 to December 2020. Coefficients, constants, and adjusted R-squareds are time-series averages. T-statistics in parentheses are calculated using Newey-West (1987) standard errors with a 120-month lag length.

	Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
			Panel A: Historical Skewness				
MarketBeta	0.01	0.03	0.06	0.10	0.36	23.08	5.18 10^3
	(4.22)	(4.26)	(4.42)	(4.83)	(6.12)	(4.12)	(3.44)
Size	0.02	0.05	0.10	0.15	0.20	3.39	1.12 10^3
	(8.17)	(8.30)	(8.83)	(10.32)	(14.45)	(1.33)	(0.93)
BookToMarket	0.01	0.03	0.05	0.09	0.22	6.77	1.58 10^3
	(5.52)	(5.51)	(5.52)	(5.65)	(5.72)	(1.94)	(1.03)
Momentum	0.02	0.05	0.09	0.14	0.27	4.68	260.65
	(5.75)	(5.80)	(5.99)	(6.39)	(6.96)	(1.98)	(0.24)
AssetGrowth	0.03	0.06	0.11	0.18	0.43	16.76	6.39 10^3
	(15.89)	(16.03)	(16.37)	(15.82)	(6.83)	(2.72)	(2.22)
Pro tability	0.01	0.02	0.04	0.07	0.16	0.13	836.37
	(4.01)	(3.92)	(3.60)	(3.01)	(2.13)	(0.03)	(0.34)
HistoricalVolatility	0.15	0.34	0.76	1.59	6.55	343.46	9.59 10^4
	(4.10)	(4.37)	(5.49)	(8.93)	(14.63)	(3.86)	(2.61)
ShareTurnover	0.00	0.00	0.00	0.00	0.01	0.75	93.04
	(0.15)	(0.18)	(0.34)	(0.76)	(3.59)	(3.71)	(0.83)
AssetTangibility	0.01	0.03	0.05	0.06	0.03	5.47	5.28 10^3
	(3.27)	(3.22)	(2.97)	(2.27)	(0.45)	(0.74)	(1.60)
CompanyAge	0.00	0.00	0.00	0.00	0.01	0.32	26.78
	(1.63)	(1.64)	(1.67)	(1.87)	(2.99)	(2.11)	(0.62)
Leverage	0.01	0.02	0.05	0.07	0.17	7.39	2.97 10^3
	(1.07)	(1.07)	(1.08)	(1.10)	(1.51)	(0.75)	(0.58)
SalesGrowth	0.01	0.01	0.03	0.04	0.09	1.63	14.24
	(3.00)	(2.99)	(2.99)	(2.97)	(2.39)	(0.42)	(0.01)
ShareIssuance	0.00	0.01	0.02	0.02	0.06	10.60	7.78 10^3
	(0.55)	(0.53)	(0.46)	(0.34)	(0.74)	(1.10)	(1.54)
Constant	0.36	0.77	1.44	2.02	1.64	155.12	4.28 10^4
	(7.65)	(7.79)	(8.37)	(9.98)	(5.87)	(2.41)	(1.62)
Adj. R^2	42.90%	43.59%	46.20%	51.63%	48.77%	8.23%	3.61%

(continued on next page)

	Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
			Panel B: Forward-Looking Skewness				
MarketBeta	0.02	0.03	0.06	0.10	0.35	22.21	5.05 10^3
	(3.56)	(3.81)	(4.05)	(4.49)	(6.62)	(4.03)	(3.38)
Size	0.03	0.06	0.11	0.15	0.21	3.06	1.05 10^3
	(9.83)	(11.61)	(13.89)	(16.58)	(17.76)	(1.33)	(0.94)
BookToMarket	0.01	0.03	0.05	0.08	0.22	6.00	1.31 10^3
	(5.73)	(5.85)	(5.59)	(5.59)	(5.99)	(1.88)	(0.91)
Momentum	0.01	0.02	0.04	0.07	0.20	4.81	10.34
	(2.08)	(2.23)	(2.64)	(3.39)	(5.19)	(2.24)	(0.01)
AssetGrowth	0.03	0.06	0.10	0.16	0.38	14.17	5.51 10^3
	(10.92)	(10.82)	(10.26)	(10.30)	(6.65)	(2.63)	(2.05)
Pro tability	0.02	0.04	0.07	0.11	0.23	0.79	809.80
	(6.28)	(6.50)	(6.32)	(5.68)	(3.43)	(0.19)	(0.36)
HistoricalVolatility	0.07	0.18	0.48	1.16	5.82	319.45	8.99 10^4
	(1.69)	(2.74)	(5.09)	(10.91)	(14.65)	(3.82)	(2.59)
ShareTurnover	0.00	0.00	0.01	0.01	0.01	0.49	47.69
	(2.66)	(2.72)	(2.58)	(2.35)	(1.18)	(2.38)	(0.42)
AssetTangibility	0.02	0.04	0.07	0.08	0.05	5.28	5.02 10^3
	(5.56)	(5.04)	(4.14)	(2.88)	(0.72)	(0.79)	(1.65)
CompanyAge	0.00	0.00	0.00	0.00	0.01	0.31	27.94
	(5.57)	(5.52)	(5.03)	(4.30)	(3.48)	(2.11)	(0.71)
Leverage	0.02	0.04	0.07	0.10	0.21	5.81	2.42 10^3
	(1.47)	(1.46)	(1.37)	(1.32)	(1.56)	(0.66)	(0.53)
SalesGrowth	0.00	0.01	0.02	0.04	0.07	1.05	104.16
	(1.80)	(2.40)	(2.72)	(2.85)	(2.03)	(0.27)	(0.08)
ShareIssuance	0.01	0.01	0.02	0.03	0.07	12.11	9.05 10^3
	(1.22)	(1.09)	(0.99)	(0.83)	(0.95)	(1.30)	(1.73)
Constant	0.40	0.82	1.45	2.09	1.99	143.80	4.03 10^4
	(6.87)	(8.17)	(10.53)	(14.21)	(6.60)	(2.40)	(1.62)
Adj. R^2	24.93%	30.33%	36.21%	43.64%	46.94%	8.49%	3.70%

Table 2.9: Comparing Our Skewness Estimates with Indirect Skewness Measures

This table presents the results from the cross-sectional regression:

$$HistoricalSkew_t \text{ or } ForecastedSkew_t = \alpha + \beta \; IndirectSkewEst_t.$$

Historical ($HistoricalSkew$) and Forward-looking ($ForecastedSkew$) skewness estimates for each stock at month t, from June 1973 to December 2020, are obtained from applying Aretz et al.'s (2024) estimator to daily discrete returns within a ten-year estimation window ending with month t. $IndirectSkewEst$ are estimates produced by indirect skewness estimators introduced in Section 2.3. To be fair to indirect proxies, we generate percentile ranks for all estimates and use those as response and predictor variables. Coefficients and adjusted R-squareds are time-series averages. T-statistics in parentheses are calculated using Newey-West (1987) standard errors with a 120-month lag length.

		Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
Panel A: Historical Skewness								
	Coefficient	0.57	0.57	0.58	0.58	0.57	0.55	0.54
Boyer-Mitton-		(27.33)	(27.02)	(26.08)	(24.82)	(22.71)	(18.84)	(17.17)
Vorkink	Constant	0.21	0.21	0.21	0.21	0.21	0.22	0.23
(BMV)		(20.54)	(20.21)	(19.21)	(17.99)	(16.88)	(15.22)	(14.58)
	Adj. R^2	33.54%	33.68%	34.16%	34.74%	34.37%	32.50%	31.33%
	Coefficient	0.26	0.26	0.27	0.30	0.34	0.38	0.39
Bali-Cakici-		(10.98)	(11.26)	(12.33)	(14.61)	(17.82)	(18.22)	(18.16)
Whitelaw	Constant	0.37	0.37	0.36	0.35	0.33	0.31	0.30
(BCW)		(31.06)	(31.41)	(32.68)	(34.84)	(34.02)	(29.49)	(27.96)
	Adj. R^2	8.03%	8.16%	8.66%	9.82%	12.83%	15.57%	16.47%
	Coefficient	0.05	0.05	0.06	0.06	0.06	0.06	0.06
Conrad-		(2.88)	(2.93)	(2.99)	(3.01)	(3.10)	(2.98)	(3.12)
Kapadia-Xing	Constant	0.48	0.48	0.48	0.48	0.48	0.48	0.48
(CKX)		(46.67)	(46.72)	(46.26)	(44.94)	(44.29)	(46.82)	(49.60)
	Adj. R^2	0.13%	0.14%	0.17%	0.23%	0.27%	0.13%	0.10%
	Coefficient	0.40	0.40	0.41	0.45	0.44	0.42	0.41
Aretz-		(14.40)	(16.07)	(19.94)	(21.09)	(23.60)	(23.29)	(22.35)
Arisoy	Constant	0.32	0.32	0.32	0.30	0.30	0.31	0.32
(AA)		(24.27)	(23.94)	(22.94)	(21.72)	(19.65)	(17.99)	(17.35)
	Adj. R^2	29.79%	29.93%	30.41%	30.99%	30.62%	28.75%	27.58%

(continued on next page)

		Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
				Panel B: Forward-Looking Skewness				
Boyer-Mitton-Vorkink (BMV)	Coefficient	0.49	0.49	0.51	0.53	0.55	0.54	0.53
		(19.99)	(20.06)	(19.99)	(18.98)	(17.34)	(16.52)	(16.04)
	Constant	0.26	0.25	0.25	0.24	0.23	0.23	0.23
		(20.89)	(20.61)	(19.45)	(16.94)	(14.28)	(14.13)	(14.13)
	Adj. R^2	24.81%	25.25%	26.67%	29.11%	31.76%	31.10%	30.42%
Bali-Cakici-Whitelaw (BCW)	Coefficient	0.33	0.33	0.34	0.36	0.38	0.40	0.41
		(8.82)	(9.49)	(11.86)	(16.33)	(18.48)	(17.81)	(17.83)
	Constant	0.34	0.34	0.33	0.32	0.31	0.30	0.30
		(18.30)	(19.29)	(22.89)	(29.25)	(29.68)	(26.53)	(25.93)
	Adj. R^2	12.10%	12.26%	12.73%	13.64%	15.50%	16.96%	17.45%
Conrad-Kapadia-Xing (CKX)	Coefficient	0.04	0.04	0.04	0.04	0.05	0.05	0.06
		(3.62)	(3.60)	(3.79)	(3.72)	(3.29)	(3.22)	(3.26)
	Constant	0.49	0.49	0.49	0.49	0.48	0.48	0.48
		(75.83)	(76.71)	(80.36)	(78.56)	(61.68)	(54.61)	(52.71)
	Adj. R^2	0.06%	0.06%	0.07%	0.02%	0.11%	0.12%	0.13%
Aretz-Arisoy (AA)	Coefficient	0.32	0.32	0.34	0.40	0.42	0.41	0.40
		(16.26)	(9.11)	(13.85)	(15.25)	(18.23)	(20.97)	(21.22)
	Constant	0.37	0.36	0.36	0.33	0.32	0.32	0.32
		(24.62)	(24.34)	(23.18)	(20.67)	(17.05)	(16.90)	(16.90)
	Adj. R^2	21.06%	21.50%	22.92%	25.36%	28.01%	27.35%	26.67%

Table 2.10: Orthogonalized Skewness Premium: Portfolio Sorts

This table provides mean excess monthly returns (in %) and alphas (in %) for portfolios univariately sorted on orthogonalized skewness, which is measured by the residual from running partial least squares regressions of Aretz et al.'s (2024) forward-looking skewness estimates on a comprehensive set of firm fundamentals (see Table 2.A.1 for more details). At the end of each month between June 1973 and December 2020, we sort stocks into value-weighted decile portfolios based on their orthogonalized skewness of one of the horizons and hold them for one month. We then create a spread portfolio by longing the tenth portfolio and shorting the first. For risk adjustments, we adopt the CAPM and the Fama-French three-factor model (FF3) and record the intercept (alpha). T-statistics in parentheses are based on Newey-West (1987) standard errors with a twelve-month lag length.

	Sorting Variable (Orthogonalized Skewness)						
Decile	Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
1 (Low)	1.02	1.02	1.04	1.03	1.06	0.89	0.82
2	0.94	0.95	0.96	0.98	0.99	0.98	0.96
3	1.06	1.01	1.00	0.99	0.98	0.89	0.92
4	1.00	1.00	0.96	0.99	0.95	0.99	0.96
5	0.94	0.97	0.98	0.95	0.93	0.92	0.95
6	0.92	0.91	0.93	0.93	0.92	0.96	0.93
7	0.84	0.85	0.86	0.87	0.91	0.94	0.94
8	0.85	0.85	0.83	0.85	0.88	0.92	0.94
9	0.82	0.79	0.83	0.82	0.86	0.90	0.90
10 (High)	0.81	0.83	0.81	0.77	0.71	0.81	0.88
Spread Return	0.21	0.19	0.23	0.27	0.35	0.08	0.06
t-statistic	(2.35)	(2.22)	(2.67)	(2.93)	(3.27)	(0.54)	(0.35)
CAPM Alpha	0.22	0.20	0.24	0.26	0.30	0.06	0.20
t-statistic	(2.50)	(2.22)	(2.54)	(2.75)	(2.71)	(0.40)	(1.09)
FF3 Alpha	0.20	0.18	0.22	0.23	0.29	0.00	0.12
t-statistic	(2.37)	(2.12)	(2.43)	(2.56)	(2.83)	(0.01)	(0.84)

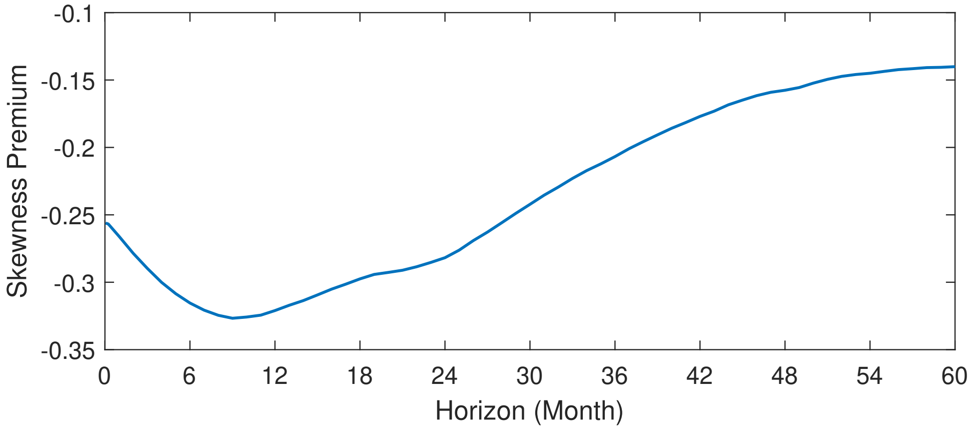
Table 2.11: Orthogonalized Skewness Premium: Fama-MacBeth Regressions

This table offers the results from Fama-MacBeth (1973) regressions of month $t+1$ single-stock excess returns on the percentile ranks of month t orthogonalized skewness, which is measured by the residual from running partial least squares regressions of Aretz et al.'s (2024) forward-looking skewness estimates on a comprehensive set of firm fundamentals (see Table 2.A.1 for more details). Control variables include market beta, size, book-to-market, momentum, asset growth, and profitability. The sample period spans from June 1973 to December 2020. Coefficients, constants, and adjusted R-squareds are time-series averages. T-statistics in parentheses are based on Newey-West (1987) standard errors with a twelve-month lag length.

	Daily	Weekly	Monthly	Quarterly	Annual	Three-year	Five-year
Orthogonalized Skewness	0.26	0.26	0.27	0.29	0.32	0.21	0.14
	(3.12)	(3.14)	(3.20)	(3.44)	(3.60)	(1.61)	(1.06)
MarketBeta	0.06	0.06	0.06	0.06	0.05	0.04	0.04
	(0.44)	(0.43)	(0.42)	(0.41)	(0.34)	(0.28)	(0.30)
Size	0.06	0.06	0.06	0.06	0.06	0.05	0.05
	(2.08)	(2.10)	(2.13)	(2.14)	(1.98)	(1.92)	(1.86)
BookToMarket	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	(1.70)	(1.67)	(1.67)	(1.67)	(1.71)	(1.70)	(1.70)
Momentum	0.72	0.72	0.72	0.72	0.72	0.72	0.74
	(3.86)	(3.86)	(3.87)	(3.87)	(3.84)	(3.91)	(4.03)
AssetGrowth	0.46	0.47	0.47	0.47	0.48	0.49	0.49
	(3.17)	(3.23)	(3.23)	(3.23)	(3.29)	(3.29)	(3.23)
Pro tability	0.55	0.54	0.55	0.55	0.55	0.53	0.52
	(3.75)	(3.73)	(3.73)	(3.73)	(3.75)	(3.55)	(3.44)
Constant	1.50	1.51	1.52	1.54	1.50	1.41	1.35
	(3.38)	(3.40)	(3.42)	(3.45)	(3.36)	(3.21)	(3.08)
Adj. R^2	5.27%	5.27%	5.27%	5.27%	5.27%	5.36%	5.39%

Figure 2.1: Orthogonalized Skewness Premium

This figure depicts the monthly premium pattern (in %) from Fama-MacBeth (1973) regressions of month $t + 1$ single-stock excess returns on the percentile ranks of month t orthogonalized skewness, which is measured by the residual from running partial least squares regressions of Aretz et al.'s (2024) forward-looking skewness estimates on a comprehensive set of firm fundamentals (see Table 2.A.1 for more details). Control variables include market beta, size, book-to-market, momentum, asset growth, and profitability. The sample period spans from June 1973 to December 2020.



Appendix 2.A Variable Definitions

Table 2.A.1: Variable Definitions

This table gives the definitions of the analysis variables used in this chapter. We indicate the updating frequency of these variables as "M" for monthly or "A" for annual. Variables updated annually are calculated using data from calendar year $t-1$ and are then used from June of calendar year t to May of calendar year $t+1$. CRSP and Compustat mnemonics are shown in parentheses.

Variable Name	Variable Definitions
MarketBeta (A)	Coefficient estimates from regressing daily excess stock returns on daily excess market returns over the past 12 months.
Size (A)	Log of market capitalization (abs(prc) × shrout).
BookToMarket (A)	The ratio of the book value of equity to its market capitalization (abs(prc) × shrout), where the book value of equity equals total assets (at) plus deferred taxes and investment tax credit (txditc) less total liabilities (lt) and preferred stock (pstkrv, pstkl, pstk, or zero).
Momentum (M)	Past twelve-month compounded return (ret) excluding the most recent month.
AssetGrowth (A)	Log of gross percentage change in total assets (at).
Profitability (A)	The ratio of sales (sale) less cost of goods sold (cogs), selling, general, and administrative expense (xsga), and interest expense (xint) to the book value of equity, which equals total assets (at) plus deferred taxes and investment tax credit (txditc) less total liabilities (lt) and preferred stock (pstkrv, pstkl, pstk, or zero).
HistoricalVolatility (M)	The standard deviation of daily discrete returns (ret) over the prior five years times $\sqrt{252}$.
ShareTurnover (M)	Past one-month average of the ratio of daily share volume (vol) to daily shares outstanding (shrout).
AssetTangibility (A)	The ratio of gross property, plant, and equipment (ppegt) to total assets (at).
CompanyAge (M)	Number of years since appearance in CRSP.
Leverage (A)	The ratio of total debt (dt) to total assets (at).
SalesGrowth (A)	Log of gross percent change in sales (sale).
ShareIssuance (A)	Log of gross percentage change in split-adjusted shares outstanding (csho × ajex).
NBER (M)	The NBER recession indicator, where a value of one indicates a recession.

Chapter 3

Retail Investing and the Skewness Premium: Evidence from the COVID-19 Pandemic

3.1 Introduction

There is a widely-held belief in academia and practice that individuals have a strong preference for skewness and are thus willing to "overpay" for skewed assets. Noteworthy manifestations of that belief are that many individuals are happy to buy lottery tickets with negative expected returns (Grote and Matheson (2006)) and that retail investors skew their portfolio holdings toward assets with skewed returns (Kumar (2009)). Spurred by this belief, Brunnermeier et al. (2007), Mitton and Vorkink (2007), and Barberis and Huang (2008) develop neoclassical or behavioral models evaluating the asset pricing implications of assuming utility or value functions implying a preference for skewness. Supporting their models' predictions, Boyer et al. (2010), Bali et al. (2011), Boyer and Vorkink (2014), Conrad et al. (2014), and others empirically show that idiosyncratic skewness negatively prices stocks and options. Despite that, we so far lack empirical evidence directly linking retail investing to skewness premiums.

In this paper, we exploit the marked rise in retail investing in the most popular stocks from before to after the start of the COVID-19 pandemic (see, e.g., Ortmann et al. (2020)[1]) to evaluate the direct influence of retail investing on the skewness premiums of single stocks. Toward that goal, we measure the skewness premium as the spread between forward-looking physical skewness obtained from Aretz et al.'s (2024) methodology and risk-neutral skewness obtained from Carr and Madan's (2001) methodology.[2] We first show that the skewness premium of the average single stock is mostly positive over the period before the pandemic (i.e., 2013 to 2019). Yet, consistent with the conjecture that individuals stuck at home massively entered stock markets as retail investors and started overpaying for skewness in 2020, we find a significant drop in that premium around that time. Further supporting that conjecture, the drop is almost entirely driven by those stocks which saw the largest increases in their retail holdings.

We start off with looking into the general empirical properties of the skewness premium in single stocks. Considering the pre-pandemic period, the skewness premium tends to be positive on average (except for the monthly horizon) and rises with the return horizon. While the skewness premium of the average single stock is, for example, 0.14 over the one-month horizon, it rises monotonically to 0.15 over the three- and to 0.37 over the six-month horizon. Those statistics, however, significantly drop from the start of the pandemic to the end of our sample period in December 2022. While the one-month-horizon skewness premium of the average single stock is, for example, 0.28 over that later period, it now rises monotonically to only 0.02 and 0.24 over the three- and six-month horizons,

[1]The popular press also noted the rise in retail investing. See, e.g., the CNBC article titled "A large chunk of the retail investing crowd started during the pandemic, Schwab survey shows" from April 8, 2021.

[2]This measure follows the skewness swap literature, not directly equivalent to the excess expected return of an asset with positively skewed returns.

respectively. The upshot is that investors appear to be willing to pay a much higher price for skewness since the onset of the pandemic.

We next explore whether changes in retail holdings are able to explain the decrease in the skewness premiums from before to after the start of the pandemic. In doing so, we exploit the wrinkle that retail investors only started investing into a subsample of (the more popular) stocks at that time. In particular, while the first quartile of percent retail holdings remains between 5% and 10% from before to after the start of the pandemic, the third quartile shoots up from about 25% to about 42%. Splitting our sample stocks into those with an average after-minus-before retail holdings change within the top 30% and those with a change within the bottom 30%, we find that those stocks with a top 30% change produce far more significant drops in the various-horizon skewness premiums than the bottom 30% stocks. We finally also report that past changes in retail holdings seem to explain the cross-section of skewness premiums, with stronger explanatory power in the post-pandemic period than in the pre-pandemic period.

Our work contributes to the literature studying skewness premiums in single stocks and stock indexes. One strand of this literature relies on conventional cross-sectional methods such as portfolio sorts and Fama-MacBeth (1973) regressions. In particular, Boyer et al. (2010), Bali et al. (2011), and Conrad et al. (2014) define the skewness premium as the mean spread return between assets with high and low (total or idiosyncratic) skewness and document that the such-defined premium is negative. Aretz and Arisoy (2023) further reveal that the such-defined premiums are mostly driven by the skewness of short- rather than long-horizon returns. Another strand relies on skewness swap trading strategies. To be specific, Kozhan et al. (2013), Pederzoli (2023), and Orłowski et al. (2024) define the skewness premium as the difference between physical and risk-neutral skewness and use the payoffs of the above strategies to show that the such-defined skewness premium is positive. Heeding studies in the second strand, we report that the skewness premiums of single stocks tend to rise with the return horizon, and that they are negatively linked to the amount of retail investors holding the underlying single stocks.

We also add to the emerging literature developing methodologies inferring the skewness premium from the difference between physical and risk-neutral skewness. Existing studies, such as Kozhan et al. (2013), Schneider and Trojani (2015), Paul Schneider (2019), Pederzoli (2023), and Orłowski et al. (2024), all measure the skewness premium as the average profit from skewness swap trading strategies in which investors buy the risk-neutral skewness of an asset's log return and receive its physical skewness. In contrast to them, our methodology concentrates on investable discrete (and not uninvestable log) returns.[3] To gain statistical precision, we further measure physical skewness not as the average payoff of the option replication strategy but rather using a direct estimate derived from Aretz et al.'s (2024) estimator. Despite those advantages, our skewness premium estimates do, in contrast to the others, not control for changes in the underlying asset's value (i.e., they are not delta-neutral). In the future, we aim to mitigate that concern by directly controlling for underlying asset risk in our tests.

Our work finally connects to studies investigating how asset premiums behave after the start of the pandemic. Berkman and Malloch (2023), for example, find sharply downward-sloping term structures of equity risk premiums across markets at that time. Heston and Todorov (2023) observe notable increases in variance risk premiums across multiple assets, with the rise in the S&P 500 premium being the most pronounced. Nieto and Rubio (2022) evaluate the performance of the three Fama-French (1993) factors, Carhart's (1997) momentum factor, and Asness et al.'s (2019) quality factor over the pandemic. They discover that only the last

[3] Our focus on discrete rather than log returns greatly influences the skewness premium estimates obtained in our empirical work. The reason is that while log single-stock returns are negatively skewed (Pederzoli (2023)), discrete single-stock returns are positively skewed (see Bessembinder (2018) and Fama and French (2018)). As a result, a log-return skewness swap is a strategy which mostly yields close-to-zero payoffs but occasionally hugely *negative* ones. In comparison, a discrete-return skewness swap is a strategy which mostly yields close-to-zero payoffs but occasionally hugely *positive* ones. The upshot is that the log-return skewness swap strategy is much riskier than the discrete-return skewness swap strategy, inducing it to produce more positive skewness premiums.

two factors outperform over that period. We add to these studies by documenting significant drops in the skewness premiums of a large number of single stocks at the start of pandemic.

We proceed as follows. In Section 2, we review a theory suggesting that a higher share of retail investors in a market lowers skewness premiums. In Section 3, we discuss our methodology to infer skewness premiums, compare it to other corresponding methodologies from the recent literature, and outline our data. In Section 4, we offer our main empirical evidence, showing that skewness premiums drop significantly after the start of the pandemic and linking these drops to increases in retail investing. Section 5 sums up and concludes.

3.2 Theoretical Framework

We first review the theoretical work of Mitton and Vorkink (2007) to formalize our intuition that a greater number of retail investors in an asset market raises (lowers) the prices (expected returns) of assets with skewed returns in that market. To wit, these authors consider a one-period model with both mean-variance (MV) and mean-variance-skewness (MVS) investors. They use the utility function $U^{MV}(W)$ to describe the preferences of the MV investors:

$$U^{MV}(W) = \mathrm{E}[W] \quad \frac{1}{2\tau}\,\mathrm{Var}[W], \tag{3.2.1}$$

where W is terminal wealth, E[ ] and Var[ ] are the expectation and variance operators, respectively, and $\tau > 0$ is the relative risk aversion parameters. Conversely, they use the utility function $U^{MVS}(W)$ to describe the preferences of the MVS investors:

$$U^{MVS}(W) = \mathrm{E}[W] \quad \frac{1}{2\tau}\,\mathrm{Var}[W] + \frac{1}{3\phi}\mathrm{Skew}[W], \tag{3.2.2}$$

where Skew[ ] is the skewness operator and $\phi > 0$ is the skewness preference parameter.

There are three risky assets in positive supply and a risk-free asset in net zero supply. The vector $\boldsymbol{R} = [R_1, R_2, R_3]'$ collects the returns of the risky assets, where the subscripts identify the three assets, and the matrix $\boldsymbol{V}$ is the variance-covariance matrix of the returns. To allow for skewed returns, let us define the matrixes $\boldsymbol{M}_1$, $\boldsymbol{M}_2$, and $\boldsymbol{M}_3$. The $\boldsymbol{M}_i$ matrix contains elements of the form: $M_{ijk} = E[(R_i - \bar{R}_i)(R_j - \bar{R}_j)(R_k - \bar{R}_k)]$, with $\bar{R}_x = E[R_x]$. Intuitively, the $i = j = k$ elements give the idiosyncratic skewness of asset i, the $i \neq j = k$ elements its co-skewness (i.e., the curvilinear interaction of asset i and j), and the $i \neq j \neq k$ its triple skewness (i.e., the triplicate product moment of the three assets). The risk-free rate of return is $r > 0$.

Assuming that the vector $\boldsymbol{X}_j = [x_{j1}, x_{j2}, x_{j3}]'$ contains the dollar investments of either the MV ($j = MV$) of the MVS ($j = MVS$) investors into the three assets, the terminal wealth of investors j can be written as: $W_j = W_{0,j}(1 + r) + \boldsymbol{X}_j'(\boldsymbol{R} - r\boldsymbol{1})$, where $W_{0,j}$ is the initial wealth of those investors and $\boldsymbol{1}$ a $[3 \times 1]$ vector of ones. Plugging the terminal wealth equation into Eq. (3.2.1), the utility function of the MV investors, taking expectations, and maximizing with respect to those investors' dollar investments, $\boldsymbol{X}_{MV}$, we obtain the solution:

$$\boldsymbol{X}_{MV} = \tau \boldsymbol{V}^{-1}(\bar{\boldsymbol{R}} - r). \tag{3.2.3}$$

Conversely, plugging the terminal wealth equation into Eq. (3.2.2), the utility function of the MVS investors, taking expectations, deriving the first partial derivative with respect to those investors' dollar investments, $\boldsymbol{X}_{MVS}$, and setting to zero, we obtain the solution:

$$(\bar{\boldsymbol{R}} - r) - \frac{1}{\tau}\boldsymbol{V}\boldsymbol{X}_{MVS} + \frac{1}{\phi}[(x_{MVS,1}\boldsymbol{M}_1 + x_{MVS,2}\boldsymbol{M}_2 + x_{MVS,3}\boldsymbol{M}_3)\boldsymbol{X}_{MVS}] = \boldsymbol{0}, \tag{3.2.4}$$

where $\boldsymbol{0}$ is a vector of zeros. Unfortunately, it is impossible to solve Eq. (3.2.4) for $\boldsymbol{X}_{MVS}$ in closed-form, so that we must numerically find the optimal value for $\boldsymbol{X}_{MVS}$.

Assuming that the risky assets are in unit supply and that only the second yields a skewed return, with its skewness entirely idiosyncratic (i.e., $M_{222} > 0$ but all the other elements of the $\boldsymbol{M}$ matrixes zero), Mitton and Vorkink (2007) solve the model by choosing reasonable parameter values and setting the demand for the risky assets equal to their supply. Starting with a market without MVS investors (by, e.g., letting ϕ increase to infinity), the standard MV-investor dollar demands in Eq. (3.2.3) solve the model and the idiosyncratic skewness of the second asset is thus irrelevant for pricing. Lowering ϕ and raising the preference for skewness, the MVS investors demand more of the second asset relative to the MV investors. We can easily understand that from Eq. (3.2.4) whose second element (i.e., the partial derivative with respect to the second asset's dollar investment) is positive when $\phi < \boldsymbol{1}$ and $\boldsymbol{X}_{MVS} = \boldsymbol{X}_{MV}$. Yet, since the demand for the second asset now exceeds its supply, that asset's expected return must drop, inducing the demands of both types of investors for it to drop and restoring equilibrium.

Taken together, the upshot is that, in markets featuring assets with skewed returns, we require investors with a preference for skewness (i.e., the MVS investors in the model) to lower (raise) the expected returns (prices) of those assets. In our empirical work, we argue that retail investors usually have a much stronger preference for skewness than instititutional investors, in line with the empirical evidence in Kumar (2009). As a result, as retail investors enter a market, the expected returns (prices) of skewed assets in that market should drop (rise).

3.3 A New Method to Compute the Skewness Premium

In this section, we propose a novel method for estimating the skewness premium, defined as the difference between forward-looking physical and risk-neutral skewness.[4] We first

[4] Premium under this definition is not directly equivalent to the excess expected return for an asset with positively skewed returns.

elaborate on how we estimate forward-looking physical skewness. We next explain how we compute risk-neutral skewness. We then compare our estimator with others from the recent literature. We finally describe the data sources used and the filters applied in our empirical analysis in Section 3.4.

3.3.1 Measuring Forward-Looking Physical Skewness

We adopt the precise and computationally fast estimator newly developed in Aretz et al. (2024) to estimate forward-looking physical skewness. Key advantages of that estimator are that it focuses on the skewness of (investable) discrete rather than (uninvestable) log returns and that it enables us to deduce the forward-looking density. Our application of that estimator starts from the assumption that we can approximate the evolution of a single-stock's price with Heston's (1993) stochastic volatility process:

$$\begin{aligned} dX_t &= (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t, \\ dV_t &= \kappa(\alpha - V_t)dt + \xi\sqrt{V_t}dB_t, \end{aligned} \tag{3.3.1}$$

where $t \geq 0$, X_t is the log stock price, V_t is the associated underlying variance, μ, κ, α, and ξ are the drift, mean reversion, long-run variance, and volatility-of-volatility parameters, respectively, and W_t and B_t are Brownian motions with a correlation of ρ. As shown in Section 2.2.2 of Aretz et al. (2024), the joint conditional MGF of Heston's (1993) process is:

$$M_t(u, w, h; V_t) := \mathrm{E}[e^{ur_{t,h} + wV_{t+h}} \mid X_t, V_t] = e^{\mu hu + \phi(u,w,h) + \psi(u,w,h)V_t}, \tag{3.3.2}$$

where $(u, w, h) \in \mathbb{C} \times \mathbb{C} \times \mathbb{R}_+$, $r_{t,h} := X_{t+h} - X_t$,

$$
\begin{aligned}
\phi(u, w, h) &= \frac{\kappa\alpha h}{\xi^2} D(u) - \frac{2\kappa\alpha}{\xi^2} \ln\{(1 + \frac{e^{I(u)h} - 1}{2I(u)} D(u))(1 - wA(u, h))\}, \\
\psi(u, w, h) &= \frac{(u^2 - u)A(u, h)}{\xi^2} + \frac{wB(u, h)}{1 - wA(u, h)}, \quad A(u, h) = \frac{\xi^2}{I(u)C(u, h) + \kappa - u\xi\rho}, \\
B(u, h) &= \frac{C(u, h)^2 - 1}{(C(u, h) + \frac{\kappa - u\xi\rho}{I(u)})^2}, \quad C(u, h) = \frac{e^{I(u)h} + 1}{e^{I(u)h} - 1}, \quad D(u) = \kappa - u\xi\rho + I(u), \\
\text{and} \quad I(u) &= \sqrt{(\kappa - u\xi\rho)^2 - \xi^2(u^2 - u)}
\end{aligned}
\tag{3.3.3}
$$

Let $R_{t,h} := e^{r_{t,h}}$ denote the discrete return. Since $M_t(u, 0, h; V_t) = \mathrm{E}[e^{r_{t,h}u} | X_t, V_t] \equiv \mathrm{E}[R_{t,h}^u | X_t, V_t]$, we can easily obtain the forward-looking physical skewness of $R_{t,h}$ from:

$$
\text{P-Skew}[R_{t,h} | V_t] = \frac{M_t(3, 0, h; V_t) - 3M_t(1, 0, h; V_t)M_t(2, 0, h; V_t) + 2M_t(1, 0, h; V_t)^3}{[(M_t(2, 0, h; V_t) - M_t(1, 0, h; V_t))^2]^{3/2}}.
\tag{3.3.4}
$$

An issue with the estimator in Eq. (3.3.4) is that X_t is not continuously observable, so that we are unable to exactly identify V_t and thus to compute $M_t(u, w, h; V_t)$. To circumvent that problem, we use a discretely observed equidistant sample $X_{t=i\Delta} \equiv X_i$, where $i \in \{0, 1, 2, \ldots\}$ and Δ is the fixed time interval between two consecutive prices (e.g., $\Delta = 1/252$ for a sample of daily returns). We then let $\mathcal{G}_{i\Delta} = \sigma(X_n, n \in \{0, 1, \ldots, i\})$ represent the information set generated by those observations. Relying on a proposition from Bates (2006), we can now devise the optimal estimate $M_{i\Delta}(u, w, h) := \mathrm{E}[M_{i\Delta}(u, w, h; V_{i\Delta}) | \mathcal{G}_{i\Delta}]$. To be concrete, we first calculate:

$$
M_{0}(u, w, h) = \mathrm{E}[M_t(u, w, h; V_t)] = e^{\mu u h + \phi(u,w,h)}(1 - \frac{\xi^2}{2\kappa}\psi(u, w, h))^{-2\kappa\alpha/\xi^2}
\tag{3.3.5}
$$

and then plug the result from that calculation into:

$$G_{i+1}(w|\mathsf{G}_{i+1}) = \frac{\int_{0\ \ i1}^{0+i1} M_i\ \ (u,w,h)e^{\ \ ur_{i,1}}du}{\int_{0\ \ i1}^{0+i1} M_i\ \ (u,0,h)e^{\ \ ur_{i,1}}du}. \tag{3.3.6}$$

Numerically solving the integrals, we compute $M_1\ \ (u,w,h)$ from $G_1(w|\mathsf{G}_1\ \ )$ using:

$$M_{i+1}(u,w,h) = e^{\mu hu+\phi(u,w,h)}G_{i+1}(\psi(u,w,h)|\mathsf{G}_{i+1}). \tag{3.3.7}$$

Iterating on Eqs. (3.3.6) and (3.3.7) until the desired time t, we finally estimate forward-looking physical skewness by replacing $M_t(u,w,h;V_t)$ in Eq. (3.3.4) with $M_i\ \ (u,w,h)$.

Based on the above insights, the main challenge in calculating forward-looking skewness from Aretz et al.'s (2024) estimator is then to estimate the five parameters of the Heston (1993) process, μ, κ, α, ξ, and ρ. We also follow Aretz et al. (2024) toward that goal. In particular, we rely on a novel GMM estimator whose moment conditions are:

$$\mathrm{E}\begin{pmatrix} \tilde{R}_{i,} \\ (\tilde{R}_{i,}^{k}\quad \tilde{M}(k,\Delta))'_{k=2:4} \\ (\tilde{R}_{i,}^{1}\ \ \tilde{R}_{i+j,}^{2}\quad \tilde{C}(1,2,\Delta,j))'_{j=1:J} \\ (\tilde{R}_{i,}^{2}\ \ \tilde{R}_{i+j,}^{2}\quad \tilde{C}(2,2,\Delta,j))'_{j=1:J} \end{pmatrix} = \mathbf{0}_{(4+2J)\ \ 1}, \tag{3.3.8}$$

where $\tilde{R}_{i,}\ \ := R_{i,}\quad M(1,0,\Delta), \tilde{M}(k,\Delta) := \mathrm{E}[\tilde{R}_{i,}^{k}\ \ ]$, and $\tilde{C}(m_1,m_2,\Delta,j) := \mathrm{Cov}[\tilde{R}_{i,}^{m_1}, \tilde{R}_{i+j,}^{m_2}\ \ ]$ are the centered return, the centered MGF, and the centered cross-MGF, respectively, and $j = 1:J$ is the number of lags, with a maximum of $J = 100$. We use the two cross-moments, $\tilde{C}(1,2,\Delta,j)$ and $\tilde{C}(2,2,\Delta,j)$, to capture the leverage and volatility clustering effects, respectively. In line with Aretz et al. (2024), we improve parameter identification and computational efficiency by substituting the theoretical moments $M(1,0,\Delta)$, $\tilde{M}(\ )$ and $\tilde{C}(\ )$ with their corresponding small-time approximations (see Proposition 3 in Aretz et al. (2024)), which offer sufficient precision for short return horizons (as, e.g., the daily).

3.3.2 Measuring Risk-Neutral Skewness

Carr and Madan (2001) show that any payoff at time T depending on the value of an underlying asset at the same time can be replicated using a portfolio consisting of the underlying asset and a continuum of European call and put options. To be specific, they establish that:

$$g(S_T) = g(S_t) + g'(S_t)(S_T - S_t) + \int_0^{\infty} g''(K_s) O_t(K_s) dK_s, \tag{3.3.9}$$

where $g(\cdot)$ is a twice-differentiable function, S_t is the price of an underlying asset at time t, and $O_t(K_s)$ is the price of an out-of-the-money (OTM) option written on that asset with strike price K_s and maturity time T at time t. Let us now denote $g_p(S_T) := (S_T/S_t)^p = R^p_{t,T-t}$, where $p \in \{1, 2, 3\}$. Intuitively, $g_p(S_T)$ gives the pth power of the discrete asset return from time t to T. Denoting by $\mathbb{Q}$ the risk-neutral measure at maturity time T, and recalling that $\mathrm{E}_t^{\mathbb{Q}}[S_T] = F_t$, the forward price of the asset at time t, it immediately follows from replication Eq. (3.3.9) that:

$$\begin{aligned}
\mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}] &= \frac{F_t}{S_t}, \\
\mathrm{E}_t^{\mathbb{Q}}[R^2_{t,T-t}] &= -1 + \frac{2F_t}{S_t} + \frac{2}{S_t^2} \int_0^{\infty} O_t(K_s) dK_s, \\
\mathrm{E}_t^{\mathbb{Q}}[R^3_{t,T-t}] &= -2 + \frac{3F_t}{S_t} + \frac{6}{S_t^3} \int_0^{\infty} K_s O_t(K_s) dK_s.
\end{aligned} \tag{3.3.10}$$

Since there are only a discrete (and not a continuum) number of options written on each single stock in the real world, we rely on a trapezoidal-rule-based approximation of Eq. (3.3.10) to calculate the risk-neutral expectations. We finally compute risk-neutral skewness as:

$$\text{RN-Skew}[R_{t,T-t}] = \frac{\mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}^3] - 3\mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}]\mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}^2] + 2\mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}]^3}{(\mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}^2] - \mathrm{E}_t^{\mathbb{Q}}[R_{t,T-t}]^2)^{3/2}}. \tag{3.3.11}$$

3.3.3 Comparing our Methodology with Others Used in the Literature

Just like us, prior studies rely on skewness swap trading strategies to estimate the skewness premium. As we already said, those strategies lead investors to buy the risk-neutral skewness of an asset's returns using a forward and option portfolio at initiation (fixed leg) and to receive the realized skewness of that return at maturity (floating leg). The studies then use the average payoff of those strategies as skewness premium estimate. Kozhan et al. (2013), for example, construct a daily-rebalanced delta-hedged skewness swap. To do so, they initially form a cubic log-return swap whose fixed leg is the value of a portfolio long OTM calls and short OTM puts and the floating leg is the approximated realized cubed log return. While Orłowski et al. (2024) rely on a similar rebalancing methodology, Schneider and Trojani (2015), Paul Schneider (2019) and Pederzoli (2023) rely on Bakshi et al.'s (2003) methodology to calculate risk-neutral expectation, allowing them to only rebalance the underlying asset. Pederzoli (2023), for example, defines a swap whose fixed leg is the value of a weighted option portfolio of OTM calls and puts replicating the cubed log return while the floating leg is the payoff of that portfolio. In addition, she uses dynamic trading in the underlying asset to eliminate underlying-asset risk.

Our skewness premium estimates differ in two key aspects. First, we look into investable and positively-skewed discrete returns, rather than into uninvestable and negatively-skewed log returns. An implication is that our skewness trading strategy usually yields close-to-zero but occasionally extremely ***positive*** payoffs. In comparison, the strategies used in the other studies usually yield close-to-zero but occasionally extremely ***negative*** payoffs. Given that, our skewness trading strategy is far less risky than those used in other studies, in line with us later finding much lower skewness premiums relative to those shown in the literature. Second, instead of using the average payoff of the skewness trading strategy to capture its expected payoff, we directly use an estimate of forward-looking skewness obtained from the skewness estimator of Aretz et al. (2024). Doing so is critical in our case since we only have 36 time-series observations over the pandemic period, rendering it difficult to use an average to estimate the expected payoff. On the downside, our skewness premium estimates do not control for changes in the underlying-asset's value (i.e., they are not delta-neutral). To mitigate that concern, we aim to more directly control for underlying asset risk in future versions of our paper.

3.3.4 Data

We obtain option and underlying-stock data from OptionMetrics, and institutional holdings data from Thomson Reuters. We exclusively focus on common stocks (shrcd: 10 or 11) traded on the NYSE, AMEX, or Nasdaq (exchcd: 1, 2, and 3). To estimate forward-looking physical skewness, we set the length of the estimation window to 10 years and employ a rolling window estimation approach, advancing one month at a time. We exclude stocks from a specific window if they have less than five years of consecutive daily returns up to the end of the estimation window to enhance precision. To compute risk-neutral skewness, we use volatility surface data to obtain options with a wide range of strike prices and maturity dates for each underlying stock. We then select times-to-maturity of 30, 60, 91, 122, 152, and 182 calendar days. To exclusively focus on standardized options, we retain only options maturing on the third Friday of each month. Also, we omit options with a price below \$0.125. To improve the accuracy of the numerical approximation of Eq. (3.3.10), we require the number of OTM options to be no less than 15. Since all options on single stocks are American-style, we exclude options on dividend-paying stocks, so that at least our sample call options are equivalent to their European counterparts. This also allows us to calculate $F_t = S_t e^{r_f(T-t)}$ according to standard no-arbitrage arguments, where r_f is the risk-free interest rate (i.e., the zero-coupon yield). Regarding institutional holdings data, we use it to infer patterns in retail holdings by subtracting the percentage of institutional holdings from one. Due to reporting frequency and minimum reporting size, this serves only as a rough proxy for retail holdings. Our sample period spans from January 2013 to December 2022.

3.4 Empirical Findings

In this section, we examine how the skewness premiums of single stocks change with the onset of the COVID-19 pandemic at the start of 2020 and whether those changes are conditioned by the change in retail investing in a stock at that time. We thus first report descriptive statistics and visualize the time-varying patterns of the skewness premiums and retail investing. We then conduct cross-sectional regressions to investigate the relation between the two variables.

3.4.1 Descriptive Statistics and Time-Varying Patterns

In Table 3.1, we offer descriptive statistics for the skewness premiums of single stocks over horizons from one month to half-a-year and for retail holdings, both before (Panel A) and after (Panel B) the start of the pandemic. We compute the descriptive statistics first by cross-section and then average over the two subsample periods. Focusing first on the skewness premiums before the pandemic, the table suggests that the skewness premium rises with the return horizon and tends to be positive except at the monthly horizon. Panel A, for example, shows that the skewness premium of the mean single stock is −0.14 over the monthly horizon and rises monotonically to 0.37 over the six-month horizon. Noteworthily, however, Panel B suggests that the premiums are markedly lower over the pandemic period.

The skewness premium of the mean single stock is now 0.28 over the monthly horizon, twice as negative as before, and 0.24 over the six-month horizon, about 65% of its previous level. Figure 3.1 plots the mean and median skewness premiums over our sample period. The figure reveals that the patterns are similar across horizons, with an increase shortly before the onset of the pandemic, followed by a continuous decline, although the skewness premiums for horizons longer than quarterly appear to stabilize in 2021.

We now turn to the retail holdings. Consistent with Jones et al.'s (2023) finding that the volume of retail trading rapidly grew in the U.S. from \$325 billion in 2019 to \$852 billion by mid-2020 and has remained high since then, the descriptive statistics in Table 3.1 show marked increases in retail holdings from before to after the start of the pandemic, albeit in an unbalanced manner (compare Panels A and B). To be more specific, while the first quartile of retail holdings remains almost unchanged, the third quartile exhibits an increase of 7%. Figure 3.2 depicts the median, first quartile, and third quartile of retail holdings over our sample period. While the first quartile remains almost flat over time, the median shows a mild increase and the third quartile a pronounced increase from 2020 to 2021. These results suggest that retail investors only flocked into a subsample of (the likely more popular) stocks at the start of the pandemic.

We next verify that the changes in skewness premiums and retail investing from before to after the start of the pandemic are statistically significant. We do so by conducting time-series regressions of cross-sectional sample means of either the skewness premiums over the alternative return horizons or retail holdings on a dummy variable equal to one during the post-pandemic period (January 2020 to December 2022) and else zero. Table 3.2 summarizes the results, with t-statistics calculated using Newey-West (1987) standard errors with a 60-month lag length. The table confirms that skewness premiums over all horizons decrease significantly after the onset of the pandemic, with the coefficients on the dummy variable being always significantly negative. In the one-month horizon skewness premium regression, the coefficient is, for example, 0.15 (t-statistic: 3.31). Conversely, retail holdings increase significantly. To be more specific, the retail holding regression yields a coefficient on the dummy variable of 0.03 (t-statistic: 10.06).

3.4.2 The Relation Between Skewness Premium and Retail Holdings

To investigate whether the changes in retail investing help explain the declines in the skewness premiums, we use a subsample spanning from 2017 to 2022 to run an initial experiment. We exclude from that subsample stocks with less than one year of data either before or after the start of the pandemic. We then calculate the difference in a stock's average retail holdings from before to after end-2019. We then plot the three-month moving average of the mean skewness premium over each horizon separately for stocks within the top 30% of average retail holding changes and those within the bottom 30%. Figure 3.3 shows that, regardless of the return horizon, the mean skewness premium for the 30% of stocks with the largest increase in average retail holdings drops significantly more than those of the 30% of stocks with the smallest increase.

We next more formally test whether changes in retail investing affect the skewness premiums over alternative horizons. To do so, Table 3.3 provides the results of cross-sectional regressions of the skewness premiums on lagged changes in retail holdings run over the full sample (Panel A), the pre-pandemic (Panel B), and the post-pandemic (Panel C) period. The plain numbers are the time-series average coefficients, constants, and adjusted R-squareds, while those in parentheses are Newey-West (1987) t-statistics. The table suggests that the increase in retail holdings consistently negatively predicts the skewness premium. Panel A, for example, shows that the semiannual horizon regression yields a retail holding change coefficien of 1.14 (t-statistic: 6.79). Notably, the mean adjusted R-squareds of the regressions are much higher over the post- compared to the pre-pandemic period. These findings suggest that retail investors may have a stronger preference for positive skewness, inducing them to pay more for assets with more skewed returns.

There may be concerns that the drop in the skewness premium after the start of the pandemic is due to unreliable estimates of forward-looking physical skewness, potentially resulting from misspecification of the underlying stock dynamics. We, therefore, separately plot the six-month moving average of the mean physical and risk-neutral components of our skewness premium in Figure 3.4. The figure shows that the two components generally align over time, with the gap between physical and risk-neutral skewness becoming narrower (wider) for horizons with mostly positive (negative) premiums.

3.5 Conclusion

In this paper, we examine how the skewness premiums of single stocks vary from before to after the outbreak of the COVID-19 pandemic and whether these variations relate to changes in retail investing. We measure a stock's skewness premium as the difference between the forward-looking physical skewness estimate obtained from Aretz et al.'s (2024) methodology and the risk-neutral skewness obtained from Carr and Madan's (2001) methodology. Using our novel estimates, we show that the skewness premium in single stocks tends to be positive except at the monthly horizon, and that its term structure from the one- to the six-month horizon is upward-sloping. We also observe significant declines in skewness premiums after the onset of the pandemic regardless of the horizon. We finally find that changes in retail holdings negatively predict the skewness premium, suggesting that retail investors may be highly attracted to lottery-like payoffs and, in turn, overpay for skewness.

Table 3.1: Descriptive Statistics

This table reports descriptive statistics, including the mean, the standard deviation (StD), the fifth percentile (Pct5), the first quartile (Q1), the median, the third quartile (Q3), and the 95th percentile (Pct95), for the skewness premiums of single stocks ("SP"), computed over horizons ranging from monthly to semiannual, and for retail holdings ("RH"), both before and after the start of the COVID-19 pandemic.

	Mean	StD	PCT5	Q1	Median	Q3	PCT95
Panel A: Pre-Pandemic (January 2013 to December 2019)							
1-Month SP	0.14	0.90	1.78	0.57	0.04	0.38	1.17
2-Month SP	0.05	0.78	1.25	0.34	0.08	0.49	1.24
3-Month SP	0.15	0.76	1.02	0.27	0.13	0.56	1.41
4-Month SP	0.24	0.77	0.86	0.21	0.18	0.62	1.58
5-Month SP	0.31	0.81	0.79	0.18	0.23	0.69	1.76
6-Month SP	0.37	0.87	0.77	0.16	0.26	0.76	1.94
RH	0.21	0.19	0.02	0.07	0.14	0.28	0.65
Panel B: Post-Pandemic (January 2020 to December 2022)							
1-Month SP	0.28	1.15	2.25	0.75	0.11	0.35	1.07
2-Month SP	0.09	0.85	1.59	0.52	0.02	0.43	1.14
3-Month SP	0.02	0.87	1.47	0.49	0.01	0.49	1.31
4-Month SP	0.10	0.88	1.25	0.42	0.07	0.59	1.54
5-Month SP	0.18	0.93	1.13	0.38	0.11	0.67	1.79
6-Month SP	0.24	1.01	1.12	0.37	0.13	0.74	2.02
RH	0.24	0.21	0.03	0.08	0.17	0.35	0.71

Table 3.2: The Significance of the Post-Pandemic Changes in Skewness Premiums and Retail Holdings

This table displays the results from the time-series regression:

$$Mean\ SkewnessPremium_t \text{ or } RetailHoldings_t = \delta_0 + \delta_1 \quad D_t^1.$$

For each sample month t, ranging from January 2013 to December 2022, we measure a stock's skewness premium as the difference between the forward-looking physical skewness estimate from Aretz et al.'s (2024) estimator and the risk-neutral skewness derived from Carr and Madan's (2001) formula (for details see Section 3.3) and its retail holdings as one minus the percentage of institutional holdings. We then compute their sample means and regress them on D^1, a dummy variable equal to one if the sample mean belongs to a post-pandemic month (i.e., January 2020 to December 2022). T-statistics are calculated using Newey-West (1987) standard errors with a 60-month lag length.

Response	D^1	
Variable	Coefficient	t-statistic
1-Month SP	0.15	3.31
2-Month SP	0.14	4.08
3-Month SP	0.17	4.02
4-Month SP	0.14	2.94
5-Month SP	0.13	2.39
6-Month SP	0.12	2.09
RH	0.03	10.06

Table 3.3: The Relation Between Skewness Premium and the Change in Retail Holdings

This table presents the results from the cross-sectional regression:

$$SkewPremium_t = \delta_0 + \delta_1 \quad ChangeInRetailHoldings_{t-1}.$$

For each sample month t, we measure a stock's skewness premium as the difference between the forward-looking physical skewness estimate from Aretz et al.'s (2024) estimator and the risk-neutral skewness derived from Carr and Madan's (2001) formula (see Section 3.3 for more details), while the retail holdings is measured as one minus the percentage of institutional holdings. Coefficients, constants, and adjusted R-squareds are time-series averages. T-statistics in parentheses are calculated using Newey-West (1987) standard errors with a twelve-month lag length.

	1-Month	2-Month	3-Month	4-Month	5-Month	6-Month
Panel A: Full Sample (January 2013 to December 2022)						
Coefficient	0.35	0.82	0.89	0.90	0.99	1.14
	(2.86)	(7.53)	(6.43)	(6.39)	(6.45)	(6.79)
Constant	0.18	0.01	0.10	0.19	0.27	0.33
	(5.46)	(0.21)	(2.37)	(4.77)	(6.32)	(7.40)
Adj. R^2	0.07%	0.22%	0.33%	0.38%	0.42%	0.45%
Panel B: Pre-Pandemic (January 2013 to December 2019)						
Coefficient	0.23	0.73	0.76	0.78	0.88	1.00
	(1.77)	(6.07)	(5.13)	(4.82)	(4.87)	(5.21)
Constant	0.14	0.05	0.15	0.23	0.30	0.36
	(5.05)	(1.47)	(3.55)	(4.99)	(5.85)	(6.42)
Adj. R^2	0.01%	0.16%	0.24%	0.32%	0.35%	0.39%
Panel C: Post-Pandemic (January 2020 to December 2022)						
Coefficient	0.64	1.05	1.19	1.18	1.27	1.45
	(3.54)	(5.85)	(4.90)	(4.79)	(4.61)	(4.68)
Constant	0.29	0.09	0.02	0.09	0.18	0.24
	(4.59)	(2.43)	(0.60)	(2.97)	(6.56)	(10.72)
Adj. R^2	0.20%	0.35%	0.53%	0.54%	0.57%	0.61%

Figure 3.1: Skewness Premiums Patterns

This figure depicts the evolving patterns of the cross-sectional mean and median skewness premiums over horizons ranging from monthly to semiannual, from January 2013 through December 2022. We define the skewness premium of a stock as the difference between the forward-looking physical skewness estimate from Aretz et al.'s (2024) estimator and the risk-neutral skewness derived from Carr and Madan's (2001) formula (for details see Section 3.3).

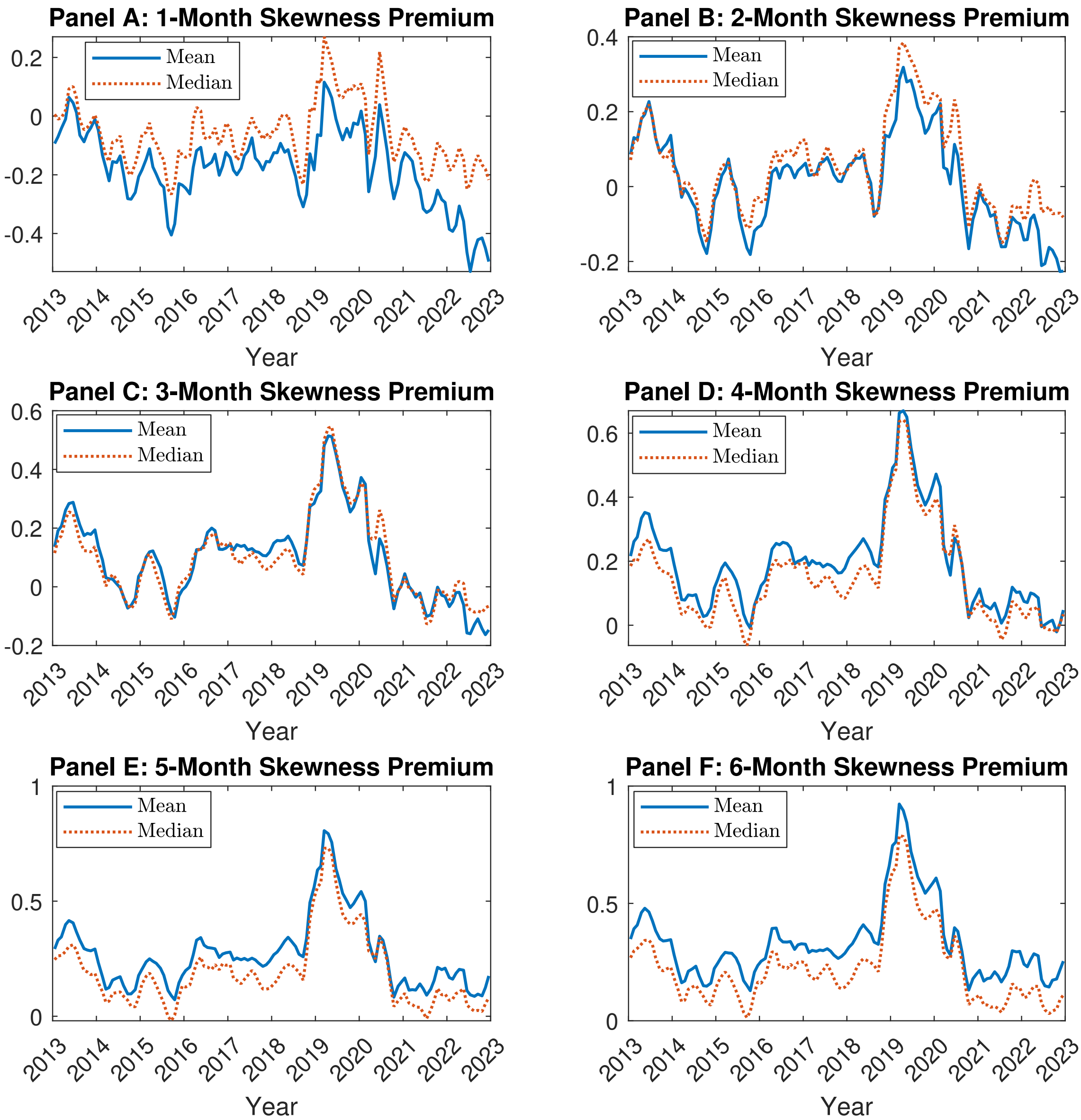


Figure 3.2: Retail Holdings Patterns

This figure plots the evolving patterns of the median, first quartile, and third quartile of retail holdings over the period from January 2013 to December 2022. Retail holdings are inferred by subtracting the percentage of institutional holdings from one.

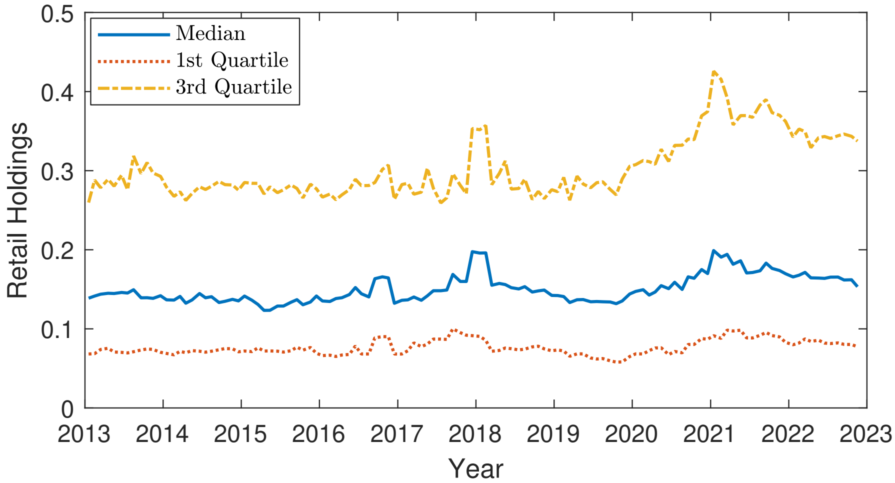


Figure 3.3: Skewness Premiums Over Time for Different Changing Levels of Retail Holdings

This figure shows how average skewness premium changes after the start of the COVID-19 pandemic for stocks within the top 30% in terms of change in average retail holdings, as well as for those within the bottom 30%, over time. For each sample month t, ranging from January 2017 to December 2022, we measure a stock's skewness premium as the difference between the forward-looking physical skewness estimate from Aretz et al.'s (2024) skewness estimator and the risk-neutral skewness derived from Carr and Madan's (2001) formula (see Section 3.3 for more details). The groups of stocks are splited by the difference in the average retail holdings before and after the end of 2019.

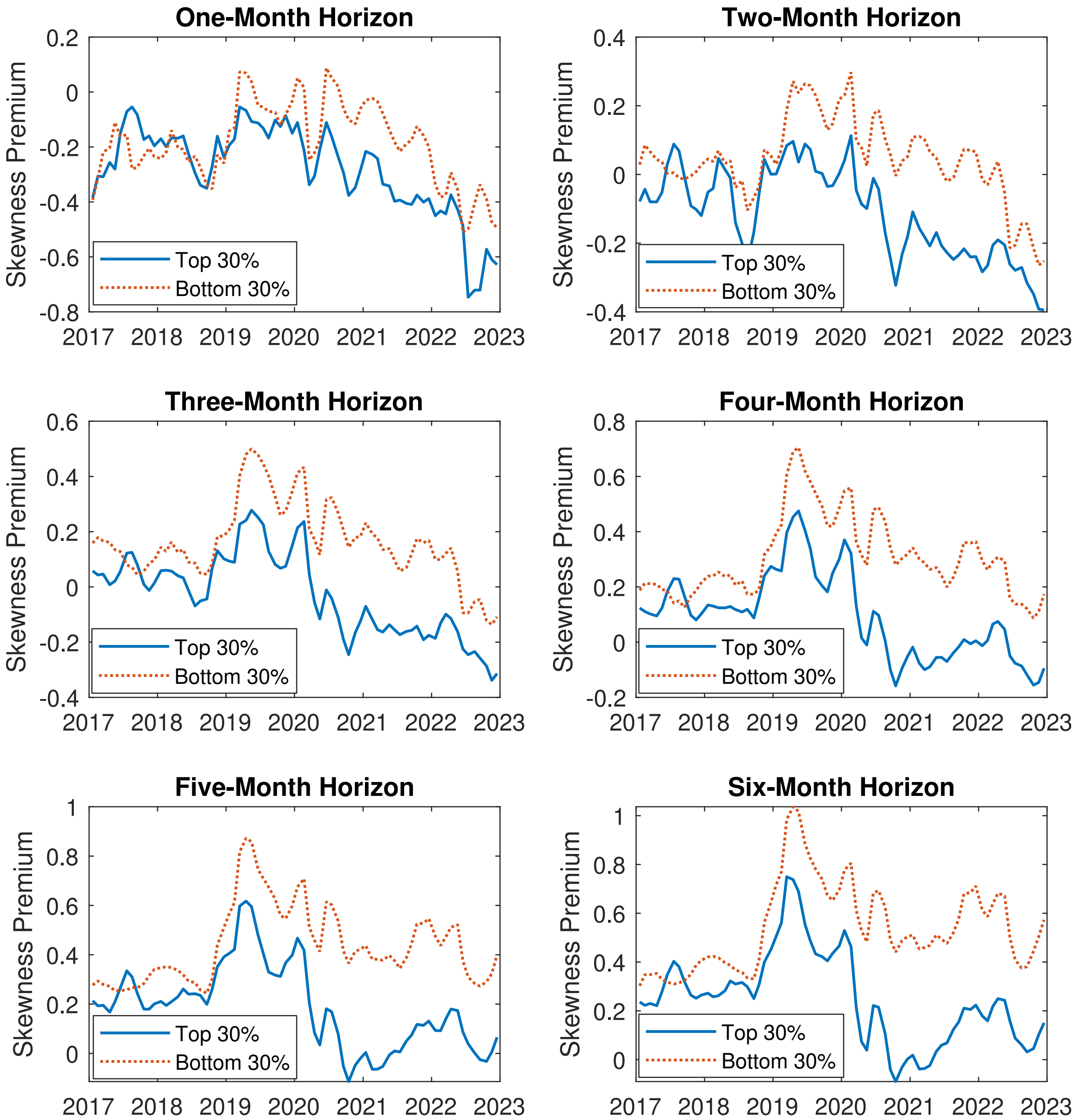
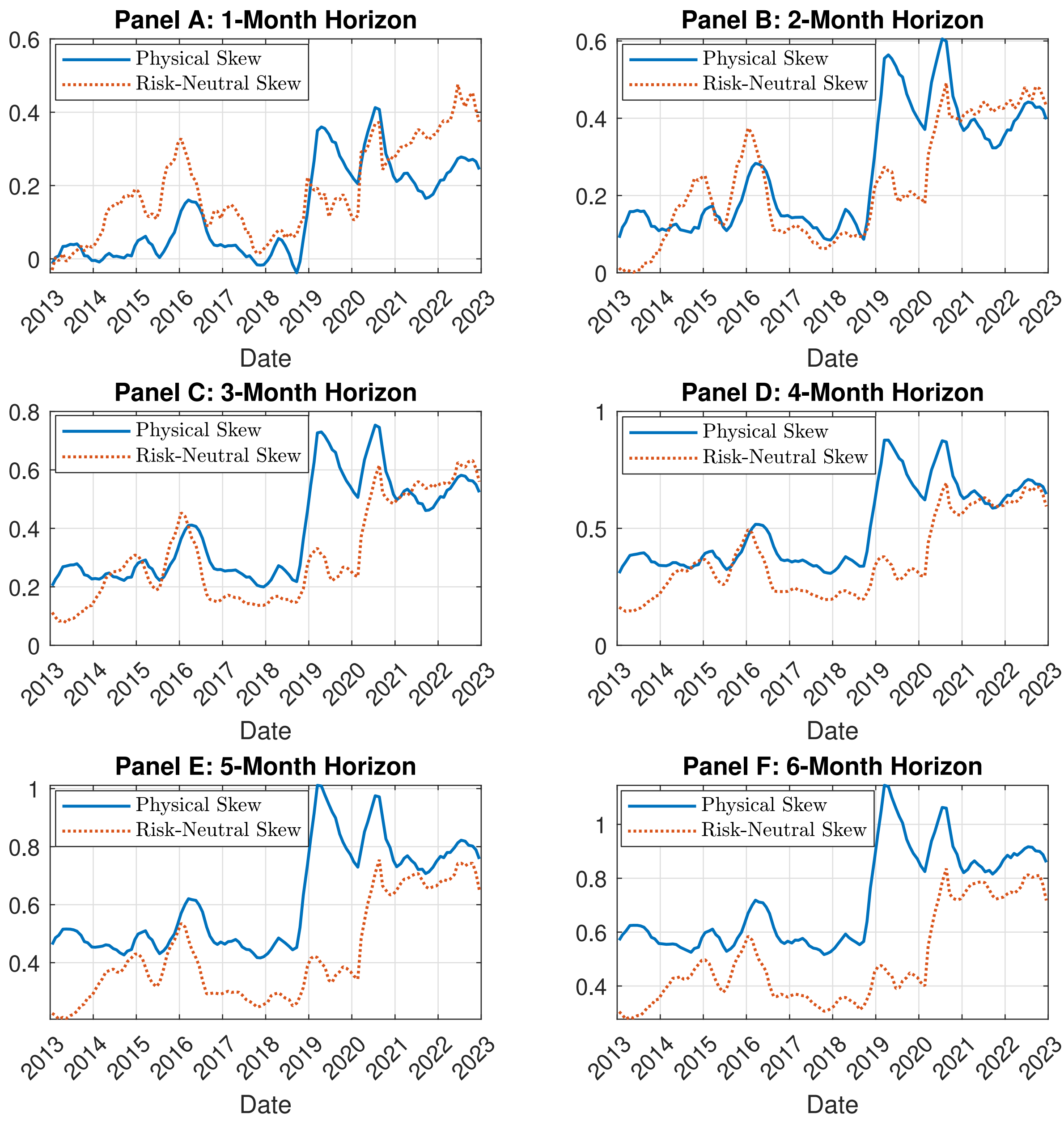


Figure 3.4: Physical and Risk-Neutral Skewness Over Time

This figure displays how the means of the physical and risk-neutral components of our skewness premium measure vary from January 2013 to December 2022. For each sample month t, we obtain a stock's forward-looking physical skewness estimate using Aretz et al.'s (2024) skewness estimator and derive the risk-neutral skewness from Carr and Madan's (2001) formula (see Section 3.3 for more details).



References

Aït-Sahalia, Yacine, and Robert Kimmel, 2007, Maximum likelihood estimation of stochastic volatility models, *Journal of Financial Economics* 83, 413–452.

Alessandri, Piergiorgio, and Haroon Mumtaz, 2019, Financial regimes and uncertainty shocks, *Journal of Monetary Economics* 101, 31–46.

Amaya, Diego, Peter Christoffersen, Kris Jacobs, and Aurelio Vasquez, 2015, Does realized skewness predict the cross-section of equity returns?, *Journal of Financial Economics* 118, 135–167.

Andersen, Leif B. G., 2008, Simple and efficient simulation of the Heston stochastic volatility model, *Journal of Computational Finance* 11, 1–42.

Andersen, Leif B. G., and Vladimir V. Piterbarg, 2007, Moment explosions in stochastic volatility models, *Finance and Stochastics* 11, 29–50.

Anderson, Bing, Peter J. Hammond, and Cyrus A. Ramezani, 2010, Affine models of the joint dynamics of exchange rates and interest rates, *Journal of Financial and Quantitative Analysis* 45, 1341–1365.

Arditti, Fred D., 1967, Risk and the required return on equity, *Journal of Finance* 22, 19–36.

Aretz, Kevin, and Y. Eser Arisoy, 2023, The pricing of skewness over different return horizons, *Journal of Banking & Finance* 148, 106713.

Aretz, Kevin, Jiayu Jin, and Yifan Li, 2024, Estimating and forecasting skewness with affine stochastic volatility models, *Working Paper* .

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen, 2019, Quality minus junk, *Review of Accounting Studies* 24, 34–112.

Atiya, Amir F., and Steve Wall, 2009, An analytic approximation of the likelihood function for the Heston model volatility estimation problem, *Quantitative Finance* 9, 289–296.

Bakshi, Gurdip, Nengjiu Ju, and Hui Ou-Yang, 2006, Estimation of continuous-time models with an application to equity volatility dynamics, *Journal of Financial Economics* 82, 227–249.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427–446.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review* 98, 2066–2100.

Bates, David S., 1996, Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options, *Review of Financial Studies* 9, 69–107.

Bates, David S., 2006, Maximum likelihood estimation of latent affine processes, *Review of Financial Studies* 19, 909–965.

Bates, David S., 2012, US stock market crash risk, 1926–2010, *Journal of Financial Economics* 105, 229–259.

Bates, David S., 2019, How crashes develop: Intradaily volatility and crash evolution, *Journal of Finance* 74, 193–238.

Berkman, Henk, and Hamish Malloch, 2023, Stock valuation during the COVID-19 pandemic: An explanation using option-based discount rates, *Journal of Banking & Finance* 147, 106386.

Bessembinder, Hendrik, 2018, Do stocks outperform treasury bills?, *Journal of Financial Economics* 129, 440–457.

Black, Fischer, 1976, Studies of stock market volatility changes, *Proceedings of the American Statistical Association, Business and Economic Statistics Section* 177–181.

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169–202.

Boyer, Brian H., and Keith Vorkink, 2014, Stock options as lotteries, *Journal of Finance* 69, 1485–1527.

Broadie, Mark, and Özgür Kaya, 2006, Exact simulation of stochastic volatility and other affine jump diffusion processes, *Operations Research* 54, 217–231.

Brunnermeier, Markus K., Christian Gollier, and Jonathan A. Parker, 2007, Optimal beliefs, asset prices, and the preference for skewed returns, *American Economic Review* 97, 159–165.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Carr, Peter, and Dilip Madan, 2001, Towards a theory of volatility trading, in *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, 458–476 (Cambridge University Press).

Chan, Joshua C. C., 2017, The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling, *Journal of Business & Economic Statistics* 35, 17–28.

Chen, Joseph, Harrison Hong, and Jeremy C Stein, 2001, Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices, *Journal of Financial Economics* 61, 345–381.

Cheridito, Patrick, Damir Filipović, and Robert L. Kimmel, 2007, Market price of risk specifications for affine models: Theory and evidence, *Journal of Financial Economics* 83, 123–170.

Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen, 2003, Alternative models for stock price dynamics, *Journal of Econometrics* 116, 225–257.

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2009, The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well, *Management Science* 55, 1914–1932.

Conine, Thomas E., and Maurry J. Tamarkin, 1981, On diversification given asymmetry in returns, *Journal of Finance* 36, 1143–1155.

Conrad, Jennifer, Nishad Kapadia, and Yuhang Xing, 2014, Death and jackpot: Why do individual investors hold overpriced stocks?, *Journal of Financial Economics* 113, 455–475.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Del Baño Rollin, Sebastian, Albert Ferreiro-Castilla, and Frederic Utzet, 2010, On the density of log-spot in the Heston volatility model, *Stochastic Processes and their Applications* 120, 2037–2063.

Duffie, Darrell, Damir Filipović, and Walter Schachermayer, 2003, Affine processes and applications in finance, *The Annals of Applied Probability* 13, 984–1053.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Engle, Robert F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987–1007.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in volatility and returns, *Journal of Finance* 58, 1269–1300.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2018, Long-horizon returns, *The Review of Asset Pricing Studies* 8, 232–252.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Fang, Fang, and Cornelis W. Oosterlee, 2009, A novel pricing method for European options based on Fourier-cosine series expansions, *SIAM Journal on Scienti c Computing* 31, 826–848.

Farago, Adam, and Erik Hjalmarsson, 2023, Long-horizon stock returns are positively skewed, *Review of Finance* 27, 495–538.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Keith Kuester, and Juan Rubio-Ramírez, 2015, Fiscal volatility shocks and economic activity, *American Economic Review* 105, 3352–3384.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez, and Martin Uribe, 2011, Risk matters: The real effects of volatility shocks, *American Economic Review* 101, 2530–2561.

Gallant, A. Ronald, and George Tauchen, 1998, Reprojecting partially observed systems with application to interest rate diffusions, *Journal of the American Statistical Association* 93, 10–24.

Ghysels, Eric, Alberto Plazzi, and Rossen Valkanov, 2016, Why invest in emerging markets? The role of conditional return asymmetry, *Journal of Finance* 71, 2145–2192.

Green, Jeremiah, John R. M. Hand, and X. Frank Zhang, 2017, The characteristics that provide independent information about average US monthly stock returns, *Review of Financial Studies* 30, 4389–4436.

Grote, Kent R., and Victor A. Matheson, 2006, In search of a fair bet in the lottery, *Eastern Economic Journal* 32, 673–684.

Hall, Alastair R., 2005, *Generalized Method of Moments* (Oxford University Press).

Hansen, Lars Peter, and José Alexandre Scheinkman, 1995, Back to the future: Generating moment implications for continuous-time Markov processes, *Econometrica* 63, 767–804.

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

Heston, Steven L., and Karamfil Todorov, 2023, Exploring the variance risk premium across assets, *Working paper* .

Hong, Harrison, and Jeremy C. Stein, 2003, Differences of opinion, short-sales constraints, and market crashes, *Review of Financial Studies* 16, 487–525.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

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

Jones, Charles M., Lin Tan, Xiaoyan Zhang, and Xinran Zhang, 2023, When price discovery and market quality are most needed: The role of retail investors during pandemic, *PBCSF-NIFR Research Paper Forthcoming* .

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

Keller-Ressel, Martin, and Eberhard Mayerhofer, 2015, Exponential moments of affine processes, *The Annals of Applied Probability* 25, 714–752.

Keller-Ressel, Martin, and Thomas Steiner, 2008, Yield curve shapes and the asymptotic short rate distribution in affine one-factor models, *Finance and Stochastics* 12, 149–172.

Kozhan, Roman, Anthony Neuberger, and Paul Schneider, 2013, The skew risk premium in the equity index market, *Review of Financial Studies* 26, 2174–2203.

Kumar, Alok, 2009, Who gambles in the stock market?, *Journal of Finance* 64, 1889–1933.

Lau, Hon-Shiang, John R. Wingender, and Amy Hing-Ling Lau, 1989, On estimating skewness in stock returns, *Management Science* 35, 1139–1142.

Li, Haitao, Martin T. Wells, and Cindy L. Yu, 2008, A Bayesian analysis of return dynamics with Lévy jumps, *Review of Financial Studies* 21, 2345–2378.

Li, Yifan, 2020, Nearly unbiased estimation of sample skewness, *Economics Letters* 192, 109174.

Mariani, Francesca, Graziella Pacelli, and Francesco Zirilli, 2008, Maximum likelihood estimation of the Heston stochastic volatility model using asset and option prices: An application of nonlinear filtering theory, *Optimization Letters* 2, 177–222.

Mencía, Javier, and Enrique Sentana, 2013, Valuation of VIX derivatives, *Journal of Financial Economics* 108, 367–391.

Mincer, Jacob A., and Victor Zarnowitz, 1969, The evaluation of economic forecasts, in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, 3–46 (National Bureau of Economic Research, Inc).

Mitton, Todd, and Keith Vorkink, 2007, Equilibrium underdiversification and the preference for skewness, *Review of Financial Studies* 20, 1255–1288.

Mrázek, Milan, Jan Posppíšil, and Tomáš Sobotka, 2016, On calibration of stochastic and fractional stochastic volatility models, *European Journal of Operational Research* 254, 1036–1046.

Mumtaz, Haroon, and Alberto Musso, 2021, The evolving impact of global, region-specific, and country-specific uncertainty, *Journal of Business & Economic Statistics* 39, 466–481.

Nelson, Daniel B., 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59, 347–70.

Neuberger, Anthony, and Richard Payne, 2021, The skewness of the stock market at long horizons, *Review of Financial Studies* 34, 1572–1616.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Nieto, Belén, and Gonzalo Rubio, 2022, The effects of the COVID-19 crisis on risk factors and option-implied expected market risk premia: An international perspective, *Journal of Risk and Financial Management* 15, 13.

Orłowski, Piotr, Paul Schneider, and Fabio Trojani, 2024, On the nature of (jump) skewness risk premia, *Management Science* 70, 1154–1174.

Ortmann, Regina, Matthias Pelster, and Sascha Tobias Wengerek, 2020, COVID-19 and investor behavior, *Finance Research Letters* 37, 101717.

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

Paul Schneider, Fabio Trojani, 2019, Divergence and the price of uncertainty, *Journal of Financial Econometrics* 17, 341–396.

Pederzoli, Paola, 2023, The crash risk in individual stocks embedded in skewness swap returns, *Swiss Finance Institute Research Paper* .

Politis, Dimitris N., and Halbert White, 2004, Automatic block-length selection for the dependent bootstrap, *Econometric Reviews* 23, 53–70.

Rubinstein, Mark E., 1973, The fundamental theorem of parameter-preference security valuation, *Journal of Financial and Quantitative Analysis* 8, 61–69.

Schneider, Paul, and Fabio Trojani, 2015, Fear trading, *Swiss Finance Institute Research Paper* .

Scott, Robert C., and Philip A. Horvath, 1980, On the direction of preference for moments of higher order than the variance, *Journal of Finance* 35, 915–919.

Shin, Minchul, and Molin Zhong, 2020, A new approach to identifying the real effects of uncertainty shocks, *Journal of Business & Economic Statistics* 38, 367–379.

Simkowitz, Michael A., and William L. Beedles, 1978, Diversification in a three-moment world, *Journal of Financial and Quantitative Analysis* 13, 927–941.

Singleton, Kenneth J., 2001, Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* 102, 111–141.

Taylor, Stephen J., 1994, Modeling stochastic volatility: A Review and comparative study, *Mathematical Finance* 4, 183–204.

Yu, Jun, 2012, Bias in the estimation of the mean reversion parameter in continuous time models, *Journal of Econometrics* 169, 114–122.