Cross-sectional Skewness*

Sangmin S. Oh† Jessica A. Wachter‡

June 13, 2021

Abstract

What distribution best characterizes the time series and cross section of individual stock returns? To answer this question, we estimate the degree of cross-sectional return skewness relative to a benchmark that nests many models considered in the literature. We find that cross-sectional skewness in monthly returns far exceeds what this benchmark model predicts. However, cross-sectional skewness in long-run returns in the data is substantially below what the model predicts. We show that fat-tailed idiosyncratic events appear to be necessary to explain skewness in the data.

Keywords: rare events, jumps, idiosyncratic volatility, power laws

*We thank Rui Albuquerque, Hendrik Bessembinder, John Campbell, Ian Dew-Becker, Priyank Gandhi, Marco Grotteria, Nishad Kapadia, Alessandro Rebucci, Yapai Zhang, and seminar participants at the 2019 JHU Finance Carey Conference, 2020 SFS Cavalcade, 2020 MFA Annual Meeting and at the Wharton School and Chicago Booth for helpful comments.

†Booth School of Business, The University of Chicago. Email: oh@chicagobooth.edu

‡The Wharton School, University of Pennsylvania. Email: jwachter@wharton.upenn.edu.

1 Introduction

What distribution best characterizes the time series and cross section of individual stock returns? Classic papers (Markowitz, 1952; Sharpe, 1964; Lintner, 1965) assumed the multivariate normal distribution. The literature quickly examined and rejected this assumption, finding instead that returns were positively skewed (Fama, 1965). Subsequent work assumed a multivariate lognormal distribution (Merton, 1973; Breeden, 1979), which implies positive skewness as the data would seem to require. Since then, many papers on the cross-section have taken the lognormal distribution as given.[1]

In this paper, we evaluate the degree of cross-sectional return skewness in the data relative to a benchmark lognormal model that accounts for the equity premium and equity volatility. We extend the model to allow for negative skewness in aggregate returns, a well-known finding in the data (Campbell and Hentschel, 1992; Duffee, 1995). We depart from previous work in that our focus is on skewness relative to the lognormal, as opposed to the normal model, and in that we measure skewness using the cross-section rather than a time series of an individual stock. When we calibrate the model to the CRSP universe on stock returns, we find that the degree of skewness in the data is far greater than that implied by lognormality. This is the case even when we consider that the skewness in the data might be a time series rather than a cross-sectional phenomenon. Namely, we consider the possibility that skewness results from "superstar months," rather than "superstar stocks." We find that, even when adjusting for superstar months, there are far too many superstar stocks to be explained by the lognormal model. We show that large, rare idiosyncratic jumps appear to be required to explain cross-sectional skewness.

Our focus on cross-sectional skewness is motivated for multiple reasons. The first reason is that asset pricing models generate predictions about the full distribution of stock returns, and in evaluating them there is no particular reason to stop at variance. Especially when one has several models that can explain, say mean and volatility, then implications

[1] A very partial list includes Berk et al. (1999), Cochrane et al. (2008), Campbell and Vuolteenaho (2004), Gomes et al. (2003), Martin (2013), Zhang (2005).

for higher-order moments will be useful in distinguishing among the competing theories. Furthermore, even if one were to focus on the first two moments, accounting for the presence of skewed shocks is important for their reliable estimation, a claim we later corroborate though a set of quantitative simulations.

Furthermore, there is a growing appreciation of cross-sectional skewness in macroeconomics and in finance, which is particularly important for understanding the source of aggregate shocks in the economy. A burgeoning research finds skewness in the cross-sectional distribution of employment growth (Ilut et al. (2018)), sales and investment (Crouzet and Mehrotra (2020)), labor earnings (Guvenen et al. (2015), Harmenberg and Sievertsen (2017)), and productivity (Salgado et al. (2019)). Also, an important growing discussion surrounds the micro-foundations of long-run cross-sectional skewness in wealth (Gomez and Gouin-Bonenfant (2020)). In interpreting these facts, however, misreading skewness will result in a lack of understanding of the relation between individual shocks and the macro-economy. As an example, a literature on credit constraints (e.g. Khan and Thomas (2013)) argues that heterogeneous exposure to these constraints is a primary source of cycles. The skewness in the size distribution of firms, on the other hand, may suggest that constraints on small firms matter less than those in an economy less skewed towards larger firms. Asset pricing moments such as skewness could then be a key source of discipline for this literature.

Our results shed light on other recent findings concerning skewness. First, our results relate to recent work by Bessembinder (2018), who shows that most stocks underperform Treasury bills most of the time. Perhaps surprisingly, we show that the underperformance of most stocks in the data does not pose a challenge to the lognormal model, while monthly and long-horizon skewness do. The lognormal model captures the underperformance of most stocks, most of the time, even while it fails to capture monthly cross-sectional skewness.

On the other hand, there is a sense in which the model predicts too much skewness. While long-term growth in market capitalizations and cumulative returns is highly skewed in the data, we show that it is even more skewed in the model. The model implies that, relatively quickly, one firm takes over the entire economy. Thus, while the distribution of firm sizes is highly skewed in the data (Axtell, 2001; Gabaix, 2009), as is the long-run distribution of returns (Bessembinder, 2018), the model implies a distribution that is even more skewed. To summarize: short-run individual returns in the data are far more positively skewed than the lognormal model would predict (while, at the same time, aggregate returns are negatively skewed), and long-run individual stock returns are less skewed than the lognormal model would predict.

Our work relates to the literature on co-skewness (Harvey and Siddique, 2000; Dittmar, 2002) and on idiosyncratic skewness (Bali et al., 2011; Boyer et al., 2010; Kapadia, 2006; Kapadia and Zekhnini, 2019) relative to the normal model. Kraus and Litzenberger (1976) show that non-increasing absolute risk aversion implies that positive co-skewness is negatively priced and that positive idiosyncratic skewness has a price of zero. As we show, however, the lognormal distribution features both types of skewness and, nonetheless, leads to a result that closely resembles the CAPM, even with constant relative risk aversion (CRRA). The assumption of CRRA utility together with lognormality, "cancel each other out," leading back to (something similar) to the CAPM. Thus we are concerned explicitly with skewness relative to the lognormal distribution.[2] We find that deviations from the lognormal benchmark are very large.

Our work also relates to a theoretical literature on skewness. Much of this work focuses on how institutional considerations (short-sale constraints, announcements, strategic disclosure) can qualitatively generate skewness relative to a normal distribution (Xu, 2007; Acharya et al., 2011; Albuquerque, 2012). Unlike this literature, our primary contribution is in documenting the very large degree of skewness relative to the lognormal distribution, and in providing a model of the stochastic discount factor and of cash flows that fits the data.

The paper proceeds as follows. Section 2 summarizes empirical findings on return skewness that motivate our study. Section 3 describes a general class of models that

[2]Our work complements a recent literature on ex ante skewness as measured in options data (Chang et al., 2013; Conrad et al., 2013).

provides a convenient lens through which to see these data. Section 4 takes this class of models to the data on return skewness and discusses the results. Section 5 illustrates the importance of accounting for skewness in the measurement of key moments. Section 6 concludes.

2 Facts about return skewness

We assume a time series of stock market return data $t = 1, \ldots, T$. There are a total of N assets (stocks), but at any point in time, only a subset of these assets have returns in the database. Let $\mathcal{J}_t \subset \{1, \ldots, N\}$ be the set of stocks available at time t. Similarly, let $\mathcal{T}_j \subset \{1, \ldots, T\}$ be the subset of time points for which stock j has available data. Assume that each stock j has consecutive return data from some $t_{0j} \geq 1$ to $T_j \leq T$. Define time-series and cross-sectional averages:

$$\bar{R}_j^{TS} \equiv \frac{1}{|\mathcal{T}_j|} \sum_{t \in \mathcal{T}_j} R_{jt} \quad (1)$$

$$\bar{R}^{CS} \equiv \frac{1}{\sum_{t=1}^{T} |\mathcal{J}_t|} \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_t} R_{jt} \quad (2)$$

$$\bar{R}_{cs,t} \equiv \frac{1}{|\mathcal{J}_t|} \sum_{j \in \mathcal{J}_t} R_{jt}, \quad (3)$$

where $|\mathcal{J}_t|$ denotes the number of elements in $\mathcal{J}_t$. Equation (1) is the time series mean of stock j, (2) the pooled (cross-sectional) mean of all stock returns, and (3) the equal-weighted average of the stock returns in the sample at time t. Now consider three types of skewness:

- Time-series skewness for stock j:

$$\gamma_j^{TS} \equiv \frac{\frac{1}{|\mathcal{T}_j|} \sum_{t \in \mathcal{T}_j} (R_{jt} - \bar{R}_j^{TS})^3}{\left[\frac{1}{|\mathcal{T}_j|-1} \sum_{t \in \mathcal{T}_j} (R_{jt} - \bar{R}_j^{TS})^2\right]^{3/2}} \quad (4)$$

We can also use (4) to define time-series skewness for the aggregate market.

- Cross-sectional skewness (measured using pooled returns)

$$\gamma^{CS} \equiv \frac{\frac{1}{\sum_{t=1}^{T}|\mathcal{J}_t|}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}(R_{jt}-\bar{R}^{CS})^3}{\left[\frac{1}{\sum_{t=1}^{T}|\mathcal{J}_t|-1}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}^{CS}\right)^2\right]^{3/2}}. \tag{5}$$

- Cross-sectional skewness at time t:

$$\gamma_{cs,t} \equiv \frac{\frac{1}{|\mathcal{J}_t|}\sum_{j\in\mathcal{J}_t}(R_{jt}-\bar{R}_{cs,t})^3}{\left[\frac{1}{|\mathcal{J}_t|-1}\sum_{j\in\mathcal{J}_t}(R_{jt}-\bar{R}_{cs,t})^2\right]^{3/2}}, \tag{6}$$

Beedles (1979), Harvey and Siddique (2000), Dittmar (2002), and Patton (2004) focus on the first of these, whereas Kapadia (2006) focuses on the third. Following Bessembinder (2018), we also consider two other measures of skewness, namely, the percent of returns greater than some fixed amount (say, the average Treasury bill rate), and the percent of total increase in value accounted for by the top 10 firms.

What could be the reasons for looking at these various measures of skewness? If skewness if a fixed quantity belonging to firm j, then (4) will be a consistent estimator of it. However, it may not be an efficient or unbiased estimator, particularly if skewness is very large; in that case, it might be useful to bring information from other assets to bear, as in (5).[3] One might object, however, that (5) does not only capture skewness in returns, but it also captures variation in idiosyncratic volatility. Indeed, Campbell et al. (2001) show significant time-series variation in idiosyncratic volatility. Presumably (6) is immune to this; however, like (4) it is likely to be missing important observations, and thus runs the risk of understating true skewness.[4]

To better understand the meaning of pooled skewness, it is also useful to decompose the expression for non-standardized skewness. Let us consider the non-standardized analogs of the previously defined skewness measures. Specifically, let $\bar{\gamma}_j^{TS}$ denote the

[3]Tsai and Wachter (2016) show that, in the presence of right-tailed shocks, the median skewness from simulated sample paths lies well below the population skewness.

[4]We corroborate this claim in Section 5 via simulations.

non-standardized time-series skewness for stock j, $\bar{\gamma}^{CS}$ the non-standardized pooled cross-sectional skewness, and $\bar{\gamma}_{cs,t}$ the non-standardized cross-sectional skewness at time t. Denote $\bar{\sigma}^2_{TS,j}$ and $\bar{\sigma}^2_{cs,t}$ as non-standardized time-series, pooled cross-sectional, and monthly cross-sectional variance. Appendix A shows that non-standardized pooled cross-sectional skewness can be expressed as

$$\begin{aligned}
\bar{\gamma}^{CS} &= \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_t} \left(R_{jt} - \bar{R}^{CS} \right)^3 \\
&= \sum_{t=1}^{T} \bar{\gamma}_{cs,t} + 3 \sum_{t=1}^{T} \left(\bar{R}_{cs,t} - \bar{R}^{CS} \right) \bar{\sigma}^2_{cs,t} + \sum_{t=1}^{T} |\mathcal{J}_t| \left(\bar{R}_{cs,t} - \bar{R}^{CS} \right)^3 \qquad (7)
\end{aligned}$$

The first term in (7) is the time-series average of monthly cross-sectional skewness. The second term is the time-series average of monthly cross-sectional variance, weighted by the deviation of each month's cross-sectional mean from the pooled mean. The final term captures the skewness of cross-sectional average returns. Together, these three components constitute pooled cross-sectional skewness.

An alternate decomposition is possible, this time in terms of time-series skewness:

$$\bar{\gamma}^{CS} = \sum_{j=1}^{J} \bar{\gamma}_j^{TS} + 3 \sum_{j=1}^{J} \left(\bar{R}_j^{TS} - \bar{R}^{CS} \right) \bar{\sigma}^2_{TS,j} + \sum_{j=1}^{J} |\mathcal{T}_j| \left(\bar{R}_j^{TS} - \bar{R}^{CS} \right)^3 . \qquad (8)$$

The first term in (8) is the cross-sectional average of time-series skewness. The second term is the cross-sectional average of time-series variance, weighted by the deviation of each stock's time-series mean from the pooled mean. The final term captures the skewness of time-series average returns. Both decompositions illustrate that pooled skewness encodes information not captured by time-series or monthly cross-sectional skewness. Importantly, it includes information on the cross-sectional and time-series variance as well as the skewness in average returns across time and within the cross-section.

As a first look at the data, Table 1 shows cross-sectional skewness γ^{CS} across various subsets of the CRSP universe.[5] Monthly skewness equals about 6. We also confirm the

[5]See Section 4 for further description of these subsets.

result of Bessembinder (2018) that most stock-month combinations deliver lower returns than the average 1-month Treasury bill. Table 2 reports corresponding statistics for multi-month returns for the CRSP common stocks from July 1926 to July 2017. Longer-horizon returns exhibit comparably high – if not higher – levels of pooled skewness as monthly returns. Furthermore, the percentage of returns positive and exceeding the T-bill rate is increasing in the return horizon, while the percentage of returns outperforming the market returns is decreasing in the return horizon.

Table 3 takes a deeper look at these results by examining statistics on γ^{TS} and on cross-sectional skewness at a point in time $\gamma_{cs,t}$. We look at both returns and log returns. Median time series skewness is not particularly large (0.9 even for the larger subset of firms), and is zero for log returns. Median cross-sectional skewness $\gamma_{cs,t}$ is higher, at 2.4. Interestingly, even cross-sectional skewness from pooled returns γ^{CS}, is negative when we look at log returns. Thus, if one were to look at time series skewness, or at log returns, one might conclude that a model with idiosyncratic lognormal shocks might adequately describe the cross-section.

3 Model

We build a model for the stochastic discount factor and for cash flows as a means of measuring the degree of skewness in the data. We introduce the most general form of our model and then consider special cases. A key component of our model is the compound Poisson process, which will allow us to tractably introduce rare events in a discrete-time framework.[6] Our model becomes the standard Poisson model in the continuous-time limit, and is closely-related to the double-jump model proposed by Kou (2002). While we use a reduced-form model for the stochastic discount factor, our main findings would not change if we used a preference-based model for the pricing kernel (Campbell and Cochrane (1999); Bansal and Yaron (2004)). For the moments of interest, these models produce

[6]See Drechsler and Yaron (2011) and Schmidt (2016). Alternative tractable ways of capturing skewness include centered Gamma shocks (Bekaert et al., 2019) and skew-normal shocks (Colacito et al., 2016).

very similar conclusions to the special case of our model in which shocks are lognormally distributed.

3.1 The Compound Poisson process

Let Q_t be a compound Poisson process parameterized by λ_t and random jump size ζ. Namely, λ_t is the expected number of jumps in the time period $(t, t+1]$. Agents in the model view jumps in $(t, t+1]$ as occurring at $t+1$. Then:

$$Q_{t+1} = \begin{cases} \sum_{k=1}^{\mathcal{N}_{t+1}-\mathcal{N}_t} \zeta_k & \text{if } \mathcal{N}_{t+1} - \mathcal{N}_t > 0 \\ 0 & \text{if } \mathcal{N}_{t+1} - \mathcal{N}_t = 0, \end{cases}$$

where $\mathcal{N}_t$ is a Poisson counting process and $\mathcal{N}_{t+1} - \mathcal{N}_t$ is the number of jumps in the time interval $(t, t+1]$. It follows that for $u \in \mathbb{R}$,

$$\log \mathbb{E}_t \left[e^{uQ_{t+1}} \right] = \lambda_t \left(\mathbb{E} \left[e^{u\zeta} \right] - 1 \right). \tag{9}$$

Note that the conditional expected value of Q_{t+1} equals

$$\mathbb{E}_t \left[Q_{t+1} \right] = \mathbb{E}_t \left[\zeta_1 + \cdots + \zeta_{\mathcal{N}_{t+1}-\mathcal{N}_t} \right] = \mathbb{E}_t \left[\mathcal{N}_{t+1} - \mathcal{N}_t \right] \mathbb{E}_t \left[\zeta \right],$$

so that

$$\mathbb{E}_t \left[Q_{t+1} \right] = \lambda_t \mathbb{E} \left[\zeta \right].$$

The conditional variance of Q_{t+1} equals

$$\mathrm{Var}_t \left[Q_{t+1} \right] = \lambda_t \left(\mathrm{Var} \left(\zeta \right) + \left(\mathbb{E} \left[\zeta \right] \right)^2 \right).$$

For small values, λ_t can be interpreted as the probability of a Poisson event over the next interval of time.

3.2 The stochastic discount factor

We assume all cash flows are subject to a common standard-normal shock $\epsilon_{c,t+1}$ and a compound Poisson shock Q_{t+1} parameterized as above. The latter captures negative skewness. Because these shocks are not diversifiable, they determine risk premia. The stochastic discount factor (SDF) will thus depend on these shocks.

As is well-known, prices are more volatile than dividends, and the lack of predictability in fundamentals strongly suggests that their volatility cannot be accounted for by rational forecasts of future dividends (see Campbell (2003)). We follow a substantial literature in allowing for a time-varying price of risk x_t in the stochastic discount factor. Let R_f denote the gross return on the short-term riskfree asset. We propose the following model for the SDF:

$$M_{t+1} = R_f^{-1} \exp\left\{ -\frac{1}{2}x_t^2 - \lambda_t \left(\mathbb{E}\left[e^{\zeta}\right] - 1 \right) - x_t \epsilon_{c,t+1} + Q_{t+1} \right\}, \tag{10}$$

where λ_t is the expected number of jumps as defined in the previous section, and ζ is the outcome of the jumps. Applying (9), along with properties of the lognormal distribution implies $\mathbb{E}_t[M_{t+1}] = R_f^{-1}$. We will assume the outcome ζ has support on the positive numbers. Then we can interpret Poisson events Q_{t+1} as increasing the SDF. As we explain in what follows, this is required by negatively skewed aggregate returns and decreasing marginal utility.

Note that x_t and λ_t are both time-varying. Let

$$x_{t+1} = (1-\varphi)\,\bar{x} + \varphi x_t + \sigma_x \epsilon_{x,t+1}, \tag{11}$$

for a standard normal $\epsilon_{x,t+1}$ that is independent of all other shocks. Lettau and Wachter (2007) argue these price-of-risk shocks should be uncorrelated with fundamental shocks $\epsilon_{c,t+1}$ to account for the cross-section. Let

$$\lambda_{t+1} = (1-\varphi_\lambda)\,\bar{\lambda} + \varphi_\lambda \lambda_t + Q_{t+1}^{\lambda},$$

where

$$Q_{t+1}^{\lambda} \sim \text{ Compound Poisson} \left(\nu, \zeta^{\lambda}\right),$$

and is independent of Q_{t+1} conditional on time-t information. We discuss the reason for this additional process in Section 3.6.

3.3 Dividend growth

Consider a cross-section of N stocks. Define D_{jt} as the dividend on stock j at time t. In addition to the common shocks $\epsilon_{c,t+1}$ and Q_{t+1}, dividends are subject to normal idiosyncratic shocks $\epsilon_{j,t+1}^{i}$ and idiosyncratic Poisson events $Q_{j,t+1}^{i}$, where

$$Q_{j,t+1}^{i} \sim \text{ Compound Poisson} \left(\lambda^{i}, \zeta_{j}^{i}\right).$$

These are independent of each other and of all other shocks, conditional on time-t information.[7] Log dividend growth $d_{jt} = \log D_{jt}$ is as follows:

$$\Delta d_{j,t+1} = \left(\mu_j - \frac{1}{2}(\sigma_j^i)^2 - \frac{1}{2}\beta_j^2\sigma_c^2\right) + \beta_j\sigma_c\epsilon_{c,t+1} + \sigma_j^i\epsilon_{j,t+1}^i - \beta_j^Q Q_{t+1} + Q_{j,t+1}^i. \tag{12}$$

Several comments about (12) are in order. First, absent occurences of the Poisson shocks, average dividend growth equals μ_j. Second, loadings on the common sources of risk differ across firms, with β_j the loading on the common normal-times risk, and β_j^Q the loading on the Poisson risk. We will show that the data require negative Poisson events in the aggregate. Therefore (while the formulas do not require that we sign β_j^Q) it is helpful to think of $\beta_j^Q > 0$ (note that $\beta_j^Q Q_{t+1}$ enters into (12) with a negative sign). While we do not derive (10) from a model of utility, reasonably assuming some link between aggregate shocks in (12) and the consumption or wealth of investors, together with decreasing marginal utility, explains why Q_{t+1} must enter into (10) with a positive sign. When there are adverse shocks to fundamentals, agents become poorer and their

[7]This is a "multiple-trees" model (Cochrane et al., 2008; Martin, 2013).

marginal utility rises.

On the other hand, the data will require that the idiosyncratic positive Poisson be positive. This is why $Q^i_{j,t+1}$ enters into (12) with a positive sign. In recent work, Salgado et al. (2019) document considerable skewness in firm-level employment and sales. These skewed shocks could be the origin of the skewed shocks in asset-level cash flows in (12).

3.4 Solution for prices

Let $P_{j,n,t}$ be the time-t price of an n-period dividend strip on asset j, namely the asset that pays the dividend for j n periods from now. First note that absence of arbitrage implies

$$\frac{P_{j,n,t}}{D_{j,t}} = \mathbb{E}_t \left[M_{t,t+n} \frac{D_{j,t+n}}{D_{j,t}} \right],$$

where $M_{t,t+n} = \prod_{k=1}^{n} M_{t+k}$. The Markov assumptions on dividend growth and the processes determining the SDF imply that the expectation depends only on x_t and λ_t. We therefore define a function

$$\Phi_j(x_t, \lambda_t, n) \equiv \frac{P_{j,n,t}}{D_{j,t}}. \tag{13}$$

Absence of arbitrage also implies prices must satisfiy

$$\mathbb{E}_t \left[M_{t+1} \left(\frac{P_{j,n-1,t+1}}{P_{j,n,t}} \right) \right] = 1, \tag{14}$$

and therefore,

$$\frac{P_{j,n,t}}{D_{j,t}} = \mathbb{E}_t \left[M_{t+1} \left(\frac{D_{j,t+1}}{D_{j,t}} \right) \left(\frac{P_{j,n-1,t+1}}{D_{j,t+1}} \right) \right], \tag{15}$$

with $P_{0,j,t}/D_{j,t} = 1$. Appendix B shows that

$$\Phi_j(x_t, \lambda_t, n) = \exp \left\{ a_j(n) + b_{xj}(n) x_t + b_{\lambda j}(n) \lambda_t \right\} \tag{16}$$

where

$$b_{xj}(n) = -\beta_j\sigma_c\frac{1-\varphi^n}{1-\varphi} \tag{17}$$

$$b_{\lambda j}(n) = \mathbb{E}\left[e^{-(\beta_j^Q-1)\zeta} - e^{\zeta}\right]\frac{1-\varphi_\lambda^n}{1-\varphi_\lambda}, \tag{18}$$

and where $a_j(n)$, given in Appendix B, does not play an important role in what follows. Note that $b_{xj}(n), b_{\lambda j}(n) < 0$ (for assets with positive exposure, i.e. $\beta_j, \beta_j^Q > 0$) and decreasing. These loadings capture the dependence of prices on time-varying risk premia. For small ζ, the expectations term in (18) is approximately $-\beta_j^Q\mathbb{E}\zeta$. Otherwise, $-\mathbb{E}\left[e^{-(\beta_j^Q-1)\zeta} - e^{\zeta}\right]$ is the rare-events analogue of the covariance $\beta_j\sigma_c$, and is increasing in β_j^Q.

The strip returns equals

$$R_{j,n,t+1} = \frac{P_{j,n-1,t+1}}{P_{jnt}} = \frac{\Phi_{n-1}^j\left(x_{t+1},\lambda_{t+1}\right)}{\Phi_{j,n}\left(x_t,\lambda_t\right)}\frac{D_{j,t+1}}{D_{jt}}. \tag{19}$$

Substituting (16) and the laws of motion for the x_t, λ_t, and $\Delta d_{j,t+1}$ into (19) implies

$$\begin{aligned}\log R_{j,n,t+1} = {} & \log R_f + \beta_j\sigma_c x_t + \mathbb{E}\left[e^{\zeta} - e^{-\left(\beta_j^Q-1\right)\zeta}\right]\lambda_t \\ & + b_{xj}(n-1)\sigma_x\epsilon_{x,t+1} + b_{\lambda j}(n-1)Q_{t+1}^{\lambda} + \beta_j\sigma_c\epsilon_{c,t+1} + \sigma_j^i\epsilon_{j,t+1}^i - \beta_j^Q Q_{t+1} + Q_{j,t+1}^i \\ & - \frac{1}{2}(\sigma_j^i)^2 - \frac{1}{2}\beta_j^2\sigma_c^2 - \frac{1}{2}\left(b_{xj}(n-1)\sigma_x\right)^2 \\ & - \nu\left(\mathbb{E}\left[e^{b_{\lambda j}(n-1)\zeta^{\lambda}}\right]-1\right) - \lambda_j^i\left(\mathbb{E}\left[e^{\zeta_j^i}\right]-1\right)\end{aligned} \tag{20}$$

Then the risk premium on the dividend strip for asset j equals

$$\log\mathbb{E}_t\left[R_{j,n,t+1}/R_f\right] = \beta_j\sigma_c x_t + \lambda_t\mathbb{E}\left[(e^{-\beta_j^Q\zeta}-1)(1-e^{\zeta})\right]. \tag{21}$$

Aggregate parameters σ_x, ν and the idiosyncratic parameters σ_{ij} and λ^i do not appear. Risk premia are a compensation for bearing the aggregate normal risk summarized by $\beta_j\sigma_c$, and the aggregate rare-event risk, summarized by the covariance $(e^{-\beta^Q\zeta}-1)(e^{\zeta}-1)$.

By assumption, shocks to x_t and to λ_t are unpriced. Idiosyncratic risk is also unpriced.

In the special case of no skewed shocks, the model takes the form of a consumption-CAPM in log returns. Under the further restriction of constant x_t, the model is precisely the consumption CAPM of Breeden (1979). We can recover this model from equilibrium using constant relative risk aversion. Thus, apparent skewness preference from the DARA (under the special case of CRRA utility) exactly "cancels out" the skewness present in log returns to produce a model closely resembling the CAPM, even though neither normality or mean-variance preferences hold.

3.5 Pricing long-lived assets

In order to estimate the model for stock returns, we need to derive prices of long-lived assets which we denote as S_{jt}. The price-dividend ratio on stock j, denoted as $\Psi_j(x_t, \lambda_t)$, is equal to the sum of the price-dividend ratios on the corresponding dividend strips:

$$\Psi_j(x_t, \lambda_t) \equiv \frac{S_{jt}}{D_{jt}} = \sum_{n=1}^{\infty} \Phi_j(x_t, \lambda_t, n). \tag{22}$$

By definition, the gross return on stock j equals:

$$R_{j,t+1} = \frac{\frac{S_{j,t+1}}{D_{j,t+1}} + 1}{\frac{S_{jt}}{D_{jt}}} \frac{D_{j,t+1}}{D_{jt}}. \tag{23}$$

Intuitiviely, stock j is a portfolio of strips. Therefore, its return must be a weighted average of the strip returns (19):

$$R_{j,t+1} = \sum_{n=1}^{\infty} w_j(x_t, \lambda_t, n) R_{j,n,t+1}, \tag{24}$$

with weights proportional to the market values:

$$w_j(x_t, \lambda_t, n) = \frac{\Phi_j(x_t, \lambda_t, n)}{\sum_{n=1}^{\infty} \Phi_j(x_t, \lambda_t, n)}, \tag{25}$$

(note that the term $D_{j,t}$ is equal across strips and therefore cancels out in computing the weights). Equation 24 can also be derived from first principles from (23) and (22).

As the time interval shrinks, the following approximation becomes more accurate:

$$\begin{aligned}\log R_{j,t+1} &= \log\left(\sum_{n=1}^{\infty} w_j(x_t,\lambda_t,n)R_{j,n,t+1}\right)\\ &\approx \text{constant as of time } t \ + \sum_{n=1}^{\infty} w_j(x_t,\lambda_t,n)\log R_{j,n,t+1}\end{aligned}$$

It therefore follows from (20) that

$$\begin{aligned}&\log R_{j,t+1} \approx \text{constant as of time } t\\ &\quad + \sum_{n=1}^{\infty} w_j(x_t,\lambda_t,n)b_{xj}(n-1)\sigma_x\epsilon_{x,t+1} + \sum_{n=1}^{\infty} w_j(x_t,\lambda_t,n)b_{\lambda j}(n-1)Q^{\lambda}_{t+1}+\\ &\quad + \beta_j\sigma_c\epsilon_{c,t+1} + \sigma^i_j\epsilon^i_{j,t+1} - \beta^Q_j Q_{t+1} + Q^i_{j,t+1}. \qquad (26)\end{aligned}$$

Equation (26) specifies how the return on stock j depends on the underlying sources of risk. Random variables $\epsilon_{ct}, \epsilon^i_{jt}, Q_t, Q^i_{jt}$ drive permanent shifts in dividends, and therefore affect all maturities equally. This is why they appear without any coefficients in (26). On the other hand, x_t and λ_t differentially affect the valuation of long versus short-term dividends. When x_t is high, for example, discount rates are high, and weight (25) shift towards short-maturity assets. This diminishes the effect both ϵ_{xt} and Q^{λ}_t have on R_{jt}.

Given that our purpose is to capture cross-sectional skewness, it is not essential that we capture time-variation in exposure to aggregate shocks arising from time-varying duration of long-lived assets. Therefore define weighted averages of the coefficients

$$\begin{aligned}b^*_{xj} &\approx \sum_{n=0}^{\infty} w_j(x_t,\lambda_t,n+1)b_{xj}(n)\\ b^*_{\lambda j} &\approx \sum_{n=0}^{\infty} w_j(x_t,\lambda_t,n+1)b_{\lambda j}(n),\end{aligned}$$

implying that (26) reduces to

$$\log R_{j,t+1} \approx \text{constant as of time } t \; + b^*_{xj}\sigma_x\epsilon_{x,t+1} + b^*_{\lambda j}Q^{\lambda}_{t+1} \\ + \beta_j\sigma_c\epsilon_{c,t+1} + \sigma^i_j\epsilon^i_{j,t+1} - \beta^Q_j Q_{t+1} + Q^i_{j,t+1} \quad (27)$$

The requirement that

$$\mathbb{E}[M_{t+1}R_{j,t+1}] = 1 \quad (28)$$

pins down the time-t constant in (27). Intuitively, the same economic forces governing risk premia on strips (21) govern risk premia on stocks. Moreover, variation in λ_t and x_t is unpriced. Therefore:

$$\log \mathbb{E}_t\left[R_{j,t+1}/R_f\right] = \beta_j\sigma_c x_t + \lambda_t\mathbb{E}\left[(e^{-\beta^Q_j\zeta} - 1)(1 - e^{\zeta})\right]. \quad (29)$$

Combining (27) with the requirement (29), together with properties of the lognormal and Poisson processes, gives the constant term in (27). Equivalently, this term can be derived from (28). See Appendix B for more details.

When calibrating the parameters, it is convenient to consider an asset that is not subject to idiosyncratic risk and with unit loadings on the shocks. This way, we can be sure that the common parameters are consistent, in at least an approximate sense, with the aggregate market. We will refer to this as the reference asset, and, give it the subscript m, because it represents the aggregate market in the calibration. The cash flows on this asset equal:

$$\Delta d_{t+1} = \left(\mu - \frac{1}{2}\sigma_c^2\right) + \sigma_c\epsilon_{c,t+1} - Q_{t+1}$$

We derive the price of this asset in a manner that is the same as the individual assets j.

The return on this asset equals:

$$
\begin{aligned}
\log R_{t+1} = \log R_f + \sigma_c x_t + \mathbb{E}_t \left[e^{\zeta} - 1 \right] \lambda_t \\
- \frac{1}{2}\sigma_c^2 - \frac{1}{2} \left(b_x^* \sigma_x \right)^2 - \nu \mathbb{E}_t \left[e^{b_\lambda^* \zeta^\lambda} - 1 \right] \\
+ b_x^* \sigma_x \epsilon_{x,t+1} + b_\lambda^* Q_{t+1}^\lambda + \sigma_c \epsilon_{c,t+1} - Q_{t+1} \quad (30)
\end{aligned}
$$

$$
\begin{aligned}
b_x^* &\approx -\sigma_c \sum_{n=1}^{\infty} w(x_t, \lambda_t, n) \frac{1 - \varphi^{n-1}}{1 - \varphi} \\
b_\lambda^* &\approx \mathbb{E}\left[1 - e^{\zeta} \right] \sum_{n=1}^{\infty} w(x_t, \lambda_t, n) \frac{1 - \varphi_\lambda^{n-1}}{1 - \varphi_\lambda},
\end{aligned}
$$

where we define w analogously to w_j.

3.6 Summary of the models

- *Consumption CAPM.* This model has dividend process (12) with the compound Poisson processes Q_t and Q_{jt}^i identically zero.[8] The model has a limiting version of SDF (10) with $\sigma_x = 0$, and $Q_t^\lambda = 0$. Thus x_t is constant at $\bar{x}$, which has the interpretation of risk aversion.

- Lettau and Wachter (2007) (henceforth LW) lognormal model. This model has dividend process (12) with the compound Poisson processes Q_t and Q_{jt}^i equal to zero. Moreover, in the SDF (10) $Q_t^\lambda = 0$. We call this the *Lognormal model.*

- LW model with rare dividend disasters. This model takes (12) and sets the idiosyncratic compound Poisson process Q_{jt}^i to zero. Moreover, in the SDF (10), $Q^\lambda = 0$.

- LW model with normal-times negative systematic skewness. This model takes (12) and sets Q_{jt}^i to zero. The SDF is given by (10). We call this the *Lognormal-N model.*

[8] That is, we set $\lambda_t = 0$ and $\lambda^i = 0$.

- LW model with negative market skewness and positive idiosyncratic skewness. This model uses (12) and (10) as given. We call this the *Lognormal-NP model*.

In what follows, we compare the lognormal model, the lognormal-N model, and the lognormal-NP model. We do not consider either the Consumption CAPM (because of its inability to match stock market volatility) or the model with skewed shocks to dividends only. The model with skewed shocks to dividends only seems, on its face, an appealing model. However, this model would only be able to match negative skewness based on negative skewness in dividend growth whereas, as we will see, negative aggregate skewness appears outside of disaster times.[9] Our model does require positive skewness to arise from positively skewed shocks to the dividend process itself; an observationally equivalent assumption (because our identification comes from returns) would be to have skewed shocks to moments of dividend growth.[10]

4 Model Estimation and Tests

4.1 Data

For the main part of our analysis, we focus on the CRSP subsample beginning in 1973, because this corresponds to inclusion of NASDAQ firms (Figure 1 shows the evolution of the number of firms over time). The data consist of monthly returns on ordinary common shares of stocks traded on all major exchanges, available on CRSP, from January 1973 to December 2016. Unless stated otherwise, we use holding period returns (i.e. with invested dividends). When computing multi-period returns, we follow entities using PERMNO. We use one-month Treasury bill returns from Kenneth French's website. We exclude all firms with fewer than 60 months of returns, to allow for plausible estimation of the parameters. We also consider a smaller subset of firms that are continuously part of the

[9]One micro-foundation for the lognormal-NP model is learning and recursive utility. See Wachter and Zhu (2019).

[10]Xu (2007) proposes such a model in the context of short-sale constraints and disagreement.

sample from 1973 to 2016. Unless otherwise stated, return statistics are monthly, and (when relevant) in percentage terms.

4.2 Simulated method of moments estimation

Because statistics such as cross-sectional and pooled skewness are not available in closed form, we must use simulations to estimate the model. We use two types of simulations. In the first, we focus on fictitious samples designed to replicate the firms continuously in existence (there are 404 such firms). In the second, we set the number of firms in our simulation equal to the median number of firms in existence in the sample at each point in time. Clearly, there are many more firms in the estimation than there are stocks in the simulation (14,786 versus 5,447). We thus use the following bootstrap procedure. At the start of each fictitious sample, we draw 5,447 stocks from the universe of 14,786 without replacement. To reflect the fact that some firm-level parameters are statistically unlikely—the firms to which they belong are only present on the exchanges for a small period of time— we assign different probability weights to different firms. That is, if $|\mathcal{T}_j|$ is the number of months that firm j is listed, we draw from the estimated parameters of firm j with probability $\frac{|\mathcal{T}_j|}{\sum_{j=1}^{14,786}|\mathcal{T}_j|}$. Thus, across fictitious samples, we should roughly capture the true distribution of firms in the cross section.

Of the two approaches to the simulation, the first is subject to survival bias. The second is subject to the criticism that we still may overweight firms that only exist for a short time (because once a firm has entered the simulation it does not leave). As we will see, however, the conclusions we reach are robust across these two strategies. This strongly suggests that more complicated ways of capturing the cross section of firms would lead to nearly identical results.

For each sample, we draw the aggregate shocks ϵ_c, ϵ_x, Q, and Q_λ. We then draw the appropriate number of sequences of firm-specific shocks, and use (27) as the asset return. We set the riskfree rate to a constant in the simulations and equal to the average rate on the 1-month Treasury bill.

4.2.1 Calibration within the simulation

The number of parameters is very large, and so to make the problem tractable, we calibrate a subset of parameters that are not the focus of the study. For aggregate parameters, we follow Lettau and Wachter (2007). We calibrate the unconditional mean of the price of risk $\bar{x}$ to generate a maximal monthly Sharpe ratio of 0.20.[11] We calibrate σ_c to annual dividend volatility. We calibrate φ to match the persistence of the price-dividend ratio. Given an annual persistence of 0.87, $\varphi = 0.988 = 0.87^{1/12}$. Finally, we set the size of the Poisson occurrence to be 15%, $\zeta = 0.15$ to match the consumption decline in the Great Depression (Barro, 2006).

Note that x_t is a latent process, and, given our reference asset (30), the volatility σ_x is not pinned down. Without loss of generality, we assume $b_x^* = -\sigma_c$. We also assume $b_{xj}^* = -\beta_j \sigma_c$, assuming that assets' exposure to ϵ_{t+1}^x is proportional to its exposure to ϵ_{t+1}^c. [12] These assumptions imply a one-factor model for normally distributed risk. We thus estimate β_j and β_j^Q from a CAPM regression of log market returns on log asset returns. We calibrate average dividend growth μ_j such that the expected dividend growth for each stock matches the average S&P 500 dividend growth.

For the lognormal model, we set idiosyncratic volatility parameters to match total volatility. That is, $\mathrm{Var}(R_j) = (\sigma_j^i)^2 + \beta_j^2 \mathrm{Var}(R_m)$. Table 4 gives our calibrated values.

4.2.2 Estimation of skewness parameters

The lognormal-N model requires four parameters to be estimated: parameters governing the disaster frequency process $\left(\bar{\lambda}, \varphi^\lambda\right)$ and those governing the distribution of shocks to disaster frequency $\left(\nu, \zeta^\lambda\right)$. For the lognormal-NP model, we introduce another scaling parameter, k, that governs the distribution of idiosyncratic Poisson shocks. Specifically, we assume that λ_j^is are drawn from a shifted power law distribution with minimum $\bar{\lambda}$ and

[11] Setting $\sqrt{e^{x^2} - 1} = 0.20$, implies $\bar{x} = 0.198$.

[12] For assets of fixed duration, this assumption is correct. However, assets with longer durations will have more exposure to x_t. Given that our purpose is skewness in individual stock returns rather than explaining average returns, accounting for this effect seems unlikely to make a difference.

mean $(k+1)\,\bar{\lambda}$ and that ζ_j^is are similarly drawn from a shifted power law distribution with minimum ζ and mean $(k+1)\,\zeta$.

We choose moments that are informative about the performance of the models in matching the data. We consider eight moments: the mean, standard deviation, and skewness of market returns; the mean, standard deviation, and skewness of pooled equity returns; the median of equity time-series skewness; and the median of monthly cross-sectional skewness. We use the identity matrix to determine the weightings of these moments. The reason for the choice of the identity matrix is that, given the mix of pooled and time-series moments, standard results concerning the joint variance matrix of the moments do not apply. For the same reason, we conduct significance tests not using asymptotic standard errors, but rather using simulations.

We initialize the states (x_0, λ_0) by drawing each from its unconditional stationary distribution. For x_t, the stationary distribution is normal with mean $\bar{x}$ and $\sigma_x^2/\left(1-\varphi^2\right)$. For λ_t, we construct the distribution numerically. Specifically, we iterate the law of motion for λ_t 500 times starting at 0 to obtain a stationary λ_t, and we conduct this exercise 500 times to obtain a stationary distribution. In each simulation, we then initialize λ_0 by randomly sampling from this distribution. The results of the estimation are summarized in Table 5.

Table 6 reports results for the aggregate market in the data and compares them to the reference asset (namely the asset with no idiosyncratic shocks). The table shows that, while all three models can match the equity premium and stock market volatility, the model with only lognormal shocks cannot match negative aggregate market skewness.

4.3 Cross-sectional skewness in pooled returns

Table 7 and Table 8 show our main results. We discuss only Table 7, as the results for Table 8 are similar. For all three models, we are interested in their ability to match moments of both level returns and log returns. As we show below, the models diverge in their ability to match skewness in level returns but not in log returns.

The data column of Table 7 shows the mean and and standard deviation of pooled stock returns, as well as the cross-sectional skewness. The table shows that all three models successfully match the first two moments of pooled stock returns. More surprisingly, both the lognormal and lognormal-N model come close to matching some moments relating to skewness. For example, both models predict that log returns are slightly negatively skewed, which turns out to be the case in the data. Thus focusing on log returns would lead one to erroneously conclude that the lognormal model generates sufficient skewness to explain the cross-section.

Both the lognormal and lognormal-N models can also account for the fact that over 50% of stock return observations are lower than the Treasury bill value (Bessembinder, 2018). To understand this result, consider, for simplicity, the lognormal model. How is it that a model with a positive price of risk can produce returns that are below Treasury bill returns most of the time? This requires us to evaluate

$$\Pr(R_j > R_f) = \Pr(\log R_M > \log R_f), \tag{31}$$

where the inequality holds because the log in a monotonic transformation. Consider (27), specialized to the case of no Poisson shocks:

$$\log R_{j,t+1} - \log R_f = \beta_j \sigma_c x_t - \frac{1}{2}(\sigma^i_j)^2 - \frac{1}{2}\beta_j^2\sigma_c^2 - \frac{1}{2}\left(b^*_{xj}\sigma_x\right)^2 \\ + b^*_{xj}\sigma_x\epsilon_{x,t+1} + \beta_j\sigma_c\epsilon_{c,t+1} + \sigma^i_j\epsilon^i_{j,t+1}.$$

Of the terms on the right-hand side, only $\beta_j\sigma_c x_t$ represents a risk premium. When the price of risk x_t is positive, as it would be all of time in a consumption-based model (and most of the time in any model with an equity premium), this term must be positive for an asset exposed to systematic risk. The log return, however, can easily fall below the riskfree rate because of the Jensen's inequality adjustments $-\frac{1}{2}(\sigma^i)^2_j - \frac{1}{2}\beta_j^2\sigma_c^2 - \frac{1}{2}\left(b^*_{xj}\sigma_x\right)^2$. It is the log return that matters when comparing the median return to the riskfree rate, given (31). The agent is willing to accept a lower return more than half of the time because of the upside potential represented by Jensen's inequality.

Finally, the lognormal and lognormal-N model both come close to matching the degree of time series skewness in the data. However, they are far from matching pooled cross-sectional skewness. The *maximal* value generated across simulations of these models less than two, whereas pooled cross-sectional skewness, in the data, is close to six. Only the lognormal-NP model, with the positive idiosyncratic Poisson events, can match the cross-sectional skewness in the data.

4.4 Cross-sectional skewness at a fixed point in time

So far we have reported that returns in the data are far more skewed than what the lognormal model would predict. One possible reason for this skewness is that, in pooling returns in the data, we have aggregated over many different idiosyncratic volatility regimes. If firm-level volatilities become more dispersed—if idiosyncratic volatility is higher at some points in time than in others (Campbell et al., 2001; Herskovic et al., 2016)—we might expect to find skewness in pooled returns. However, at any particular point in time, skewness in the population of returns would be much less. As we state in the introduction: skewness might arise from superstar months, not superstar stocks. As such, it could be consistent with a lognormal distribution.

To confront this concern, we compute skewness in the cross-section at each point in time. That is, we consider the measure (6). Figure 2 shows a histogram of these skewness observations. Consistent with time-varying idiosyncratic volatility, the majority of observations fall below the pooled statistic, with a large cluster close to zero. Even so, the data firmly reject the model. The dotted line in the figure shows the maximum skewness obtained in simulations in the model. The majority of data observations exceed the maximum value implied in the model simulations. Figure 3, which shows analogous results for the 404 firms, tells a similar story. Finally, Table 9 shows that average cross-sectional skewness, while below the pooled skewness, is far above what the model is capable of generating. We can therefore conclude that the lognormal model is not capable of generating the cross-sectional skewness observed in the data.[13] Thus, while skewness may partially arise from superstar months, it certainly also reflects superstar stocks.

4.5 Discussion of Model Fit

We also conduct a more formal test of the fit of the lognormal-N and lognormal-NP model to the data. Given our deviations from the traditional SMM approach, a direct test of over-identifying restrictions is ill-defined in our case.[14] As an alternative, we construct 50%, 90%, and 95% confidence ellipses for pairs of the relevant moments used in our estimation. This approach has the advantage of leveraging correlations across our moments that can only be obtained in simulations and not in the data.

Figures 4 through 6 compare the performance of lognormal-N and lognormal-NP models on the set of 404 firms. Figures 7 through 9 provide the analogous results for the set of 14,786 firms. Across the figures, the message is clear: the lognormal-NP model can jointly match or get much closer to replicating the high degree of pooled, time-series, and monthly cross-sectional skewness observed in the data, while lognormal and lognormal-N cannot.

4.6 Skewness in market capitalization

Axtell (2001) and Gabaix (2009) show that firm sizes follow a power law distribution. The skewness in this distribution implies that what might appear to be idiosyncratic shocks are not necessarily idiosyncratic in that they don't average out as the central limit

[13] In this analysis, we simulate from a lognormal model with constant idiosyncratic volatilities. One might argue then that we are not perhaps accounting for spikes in idiosyncratic volatility as documented by Herskovic et al. (2016). Because of the infrequency of these spikes, however, they cannot account for the inability to match cross-sectional skewness throughout the data sample.

[14] To see this point clearly, the test statistic for a test of over-identifying restriction in GMM / SMM is typically given as

$$\xi_T = \frac{TS}{1+S} \text{(minimized SMM objective function)} \sim \chi^2(q-p)$$

where p is the number of parameters, q is the number of moments, S is the number of simulations, and T is the number of observations in the data. Since we examine both pooled and medians of time-series and cross-sectional skewness, there is not a single T that is applicable to our setting.

theorem would imply (Gabaix, 2011). Furthermore, Bessembinder (2018) shows that the distribution in cumulative distributions is highly skewed. A related finding is extreme positive skewness in the wealth distribution (Gomez, 2019).

Motivated by these findings, we examine the implications of the model for the long-run distribution of firm sizes. All versions of the model deliver strong implications for the long-run distribution of firm sizes: it is even more skewed than in the data. Before reporting results from simulations, we discuss theoretical findings that show why this might be true.

Our first result is that the difference between growth rates of firm capitalizations is a stationary random variable, plus a random walk with drift. Note that, between any time t and $t+\tau$, the growth in the capitalization of firm j equals:

$$\log(S_{j,t+\tau}/S_{j,t}) = \log(\Psi_j(x_{t+\tau}, \lambda_{t+\tau})/\Psi_j(x_t, \lambda_t)) + \log(D_{j,t+\tau}/D_{j,t}) \tag{32}$$

The first term, $\log(\Psi_j(x_{t+\tau}, \lambda_{t+\tau})/\Psi_j(x_t, \lambda_t))$ is stationary, whereas $\log(D_{j,t+\tau}/D_{j,t})$ is a random walk with drift for $\tau = 1, 2, \ldots, \infty$.

It follows that

Lemma 1. *For any firms j, k with $j \neq k$, the difference between log growth in market capitalization is the sum of a stationary component and a random walk with drift.*

Namely, because the growth rate in one firm's capitalization in a random walk with drift, the idiosyncratic terms imply that the difference between the growth rates must be a random walk with drift.

The random walk is well-known to be non-stationary: its variance grows linearly with the horizon. This means that the log spread in capitalizations can wander anywhere on the real line. In practical terms, it will spend most of its time at very large positive or negative numbers.

Corollary 1. *There is a subsequence t_ℓ such that for every j, k, $j \neq k$, $\log S_{j,t_\ell}/S_{k,t_\ell}$ approaches negative or positive infinity as t_ℓ approaches infinity.*

The proof follows from the nonstationarity of the random walk (see (Feller, 1968, Chapter 14)). It is closely related to the result of Martin (2012) that cumulative returns either tend to zero or infinity.

Now consider ratios of the form

$$\frac{S_{jT}/S_{jt}}{\sum_{k=1}^{N}(S_{kT}/S_{kt})},$$

namely growth rates, as a percent of the growth rates of all firms in the economy.[15] Corollary 1 implies that these ratios to be either very close to 1, or very close to zero most of the time, provided that T is sufficiently large. One firm will take over the economy, where the others will have shares approaching zero.

To illustrate how quickly this might happen, we compute the ratios of market capitalization growth for the top ten firms in the model and compare them to the data. Because firm valuation ratios are stationary, for convenience we assume no changes in the price-dividend ratio (the stationary component of (32)), and compute shares in dividend levels. In Figures 10 and 11, we show the value of the top-ten share as a function of the number of simulation months in the economy, assuming equal shares at the start.

All three models generate skewness a top-ten share that exceeds that in the data for even a small number of simulation months.[16] One firm quickly takes over the entire economy, in effect making that firm "the market." This qualitative difference between the model and the data suggest the need to modify models of the cross-section to confront long-run implications. Finance theory, and data, require shocks to growth rates (not levels), implying non-stationary levels. Unanswered is the question of what economic force maintains what appears to be stationary levels, and what are its implications for pricing.

[15]Considering growth rates, as opposed to sizes, eliminates the dependence on initial sizes. Clearly the long-run distribution of share of growth rates, and share of sizes, will be the same.

[16]For illustrative purposes, we show the top-ten share for the 404 firms; the value for the larger group of firms is even lower. Replacing market capitalization by cumulative return, as in Bessembinder (2018) makes little difference in the analysis.

4.7 Evaluation of Model Assumption

In calibrating stock returns, we have used the approximation (27) to move from dividend strips to long-lived assets, We provide here an ex-post assessment of this approximation and show that it holds quite tightly.

Using the calibrated and estimated parameters for the lognormal-NP model, we simulate returns to equity strips of varying maturities for a single stock j. For simplicity, we set $\beta_j = \beta_j^Q = 0.8$ and set $\sigma_j = 0.078$, the median value in the calibration from the set of 404 firms. λ_j and ζ_j are also set to $\bar{\lambda}$ and ζ respectively. Importantly, both λ_t and x_t are allowed to vary across time since their degree of deviations from their long-term means are a first-order concern for the approximations.

Having simulated the returns on equity strips, we first compute the exact log equity return given by (24) where the sum is computed using equity strips with N different maturities. We then compute the approximate log equity return given by the right-hand side of (27) using the same number of strips. Repeating this exercise for $N = 1$ to 500 then allows us to examine the tightness of our approximation as a function of the number of strips used to compute the returns.

In the first panel of Figure 12, we plot the exact and approximate log equity return for a sample month with weights $w_{j,n}(x_t, \lambda_t)$ computed using equation (25). In the second panel, we repeat the exercise with equal weights, namely $\forall n : w_{j,n}(x_t, \lambda_t) = 1/N$. For both panels, the gap between the two lines represents the tightness of the approximation as a function of the number of equity strips used. At $N = 1$, we verify that the approximation is exact given there is only one element in the summation. For $N > 1$, the difference is negligible. The tightness of the approximation reveals that in our simulations, the weights and the ratio $\Phi_{j,n-1}(x_{t+1}, \lambda_{t+1})/\Phi_{j,n}(x_t, \lambda_t)$ – which constitute the summand in (26) – exhibit little variation across maturities. Overall, this exercise lends assurance to our approximation and provides additional robustness to our simulation results.

5 Skewness and the Measurement of Key Moments

Why is it important to get cross-sectional skewness right? Here we provide additional answers through stylized simulations based on the lognormal, lognormal-N, and lognoraml-NP models.

Specifically, we set up simplified versions of lognormal, lognormal-N, and lognormal-NP models and consider the returns of a single stock j. We assume that $\beta_j = \beta_j^Q = 0.8$ and set $\sigma_j = 0.078$, the median value in the calibration from the set of 404 firms. For lognormal-N and lognormal-NP, we scale σ_j so that the volatility of level returns is constant across the three models. We also assume that $\lambda_t = \bar{\lambda}$ and $x_t = \bar{x}$ following our original calibration. For each model, we generate returns according to (27), where (29) gives the mean return that is consistent with no arbitrage. We simulate 1000 sample with various lengths: a short sample of 10 observations, a realistic sample of 50 observations, and a long sample of 500 observations.

Table 10 reports the distribution of the sample mean. Note that all three models feature substantial skewness and volatility. Under all three null hypotheses, the mean is poorly estimated as indicated both by wide confidence intervals and by median values well below true values for realistic sample lengths. Table 11 reports estimates of the standard deviation. In constrast, under all three null hypotheses, the standard deviation is well-estimated, though the estimates do feature a downward bias. Finally, Table 12 reports estimates of skewness. Skewness is badly estimated for all three models with short sample lengths and well-estimated given sufficient data. However, for realistic samples, skewness is well-estimated for the lognormal and lognormal-N model, but badly estimated for the lognormal-NP model. That is, if one were living in a lognormal-NP world, one wouldn't know it from the data; even though true skewness is nearly double, the estimated skewness is not far from that in the lognormal model.

Figures 13 and 14 offer another view of the finite-sample properties of mean, volatility, and skewness. In Figure 13, we plot the time-series of the median estimates of expected return and volatility at longer horizons. Here we find that with enough time-series, the expanding-window estimates of both expected return and volatility converge to their true values. For the lognormal-NP model, which is the only model that delivers realistic magnitude of cross-sectional skewness, it takes more than 1,000 months for the volatility estimator to converge. Volatility estimates are within 10 basis points of the true value relatively quickly (which is why Table 11 shows the volatility as well-estimated). Nonetheless, the nature of the underlying shocks matters for the interpretation of volatility estimates.

Figure 14 contains the analogous results for median estimates of skewness at longer horizons. For sufficiently large samples, all models are correct on average despite a great deal of noise in the estimation. But for realistic samples, in which lognormal-NP appears to resemble the lognormal model, the median estimates of skewness substantially understate the true population skewness. Given the finite horizon that most empirical estimates face, the results therefore highlight the measurement error associated with estimating time-series skewness. On the other hand, pooled cross-sectional skewness, which encodes important information about time-series skewness as argued in Equation (8), allows the econometrician to overcome this problem by using all the information available across assets and through time.

6 Conclusion

We have considered the implications of three types of skewness – time series, cross-sectional (pooled), and cross-sectional at a fixed point in time – through the lens of a standard asset pricing model. Table 13 summarizes the models and the conclusions. All three models we consider are calibrated to match the equity premium and equity volatility. Interestingly, this is sufficient to match the low percentage of returns that exceed the riskfree rate, as well as skewness in long-horizon returns. However, non-normal positive shocks appear to be necessary to account for the high degree of skewness in the cross section.

Cross-sectional skewness has implications beyond simply measurement. For example, several recent models explain asset pricing facts through the mechanism of innovation

(Dou, 2017; Kogan et al., 2019; Garleanu and Panageas, 2018). Presumably, such innovation begins at the level of individual firms, and understanding the degree of cross-sectional skewness can help to calibrate this. Moreover, the pricing of volatility risk remains a topic of active debate (Dew-Becker et al., 2019). To the extent that much of idiosyncratic volatility is in fact upside risk, as measured by skewness, this suggests a mechanism by which volatility may contribute positively to investment opportunities. Finally, the skewed long-run distribution creates an intriguing theoretical problem for the micro-foundations of firms and the aggregate market. If firm dividends contain a permanent source of uncertainty, then some firms will inevitably grow and "take over" the economy. We leave these topics to future research.

Appendix

A Decomposition of Pooled Skewness

We first provide a decomposition in terms of monthly cross-sectional skewness:

$$
\begin{aligned}
&\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}^{CS}\right)^3\\
=&\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}_{cs,t}+\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)\right)^3\\
=&\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}_{cs,t}\right)^3+\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)^3+3\left(R_{jt}-\bar{R}_{cs,t}\right)^2\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)+3\left(R_{jt}-\bar{R}_{cs,t}\right)\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)^2\\
=&\sum_{t=1}^{T}\underbrace{\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}_{cs,t}\right)^3}_{\equiv\bar{\gamma}_{cs,t}}+\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)^3\\
&+3\sum_{t=1}^{T}\underbrace{\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}_{cs,t}\right)^2}_{\equiv\bar{\sigma}^2_{cs,t}}\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)+3\sum_{t=1}^{T}\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)^2\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}_{cs,t}\right)
\end{aligned}
$$

Since

$$\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}_{cs,t}\right)=\sum_{j\in\mathcal{J}_t}R_{jt}-\sum_{j\in\mathcal{J}_t}\bar{R}_{cs,t}=\left|\mathcal{J}_t\right|\bar{R}_{cs,t}-\left|\mathcal{J}_t\right|\bar{R}_{cs,t}=0$$

it follows that:

$$\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(R_{jt}-\bar{R}^{CS}\right)^3=\sum_{t=1}^{T}\bar{\gamma}_{cs,t}+\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_t}\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)^3+3\sum_{t=1}^{T}\left(\bar{R}_{cs,t}-\bar{R}^{CS}\right)\bar{\sigma}^2_{cs,t}$$

An alternate decomposition in terms of time-series skewness can be obtained using analogous steps, shown below:

$$
\begin{aligned}
&\sum_{j=1}^{J}\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}^{CS}\right)^3\\
=&\sum_{j=1}^{J}\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}_j^{TS}+\bar{R}_j^{TS}-\bar{R}^{CS}\right)^3\\
=&\sum_{j=1}^{J}\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}_j^{TS}\right)^3+\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)^3+3\left(R_{jt}-\bar{R}_j^{TS}\right)^2\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)+3\left(R_{jt}-\bar{R}_j^{TS}\right)\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)^2\\
=&\sum_{j=1}^{J}\underbrace{\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}_j^{TS}\right)^3}_{\equiv\bar{\gamma}_j^{TS}}+\sum_{j=1}^{J}\sum_{t\in\mathcal{T}_j}\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)^3\\
&+3\sum_{j=1}^{J}\underbrace{\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}_j^{TS}\right)^2}_{\equiv\bar{\sigma}^2_{TS,j}}\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)+3\sum_{j=1}^{J}\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)^2\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}_j^{TS}\right)
\end{aligned}
$$

By similar logic as above,

$$
\sum_{j=1}^{J}\sum_{t\in\mathcal{T}_j}\left(R_{jt}-\bar{R}^{CS}\right)^3=\sum_{j=1}^{J}\bar{\gamma}_j^{TS}+\sum_{j=1}^{J}\sum_{t\in\mathcal{T}_j}\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)^3+3\sum_{j=1}^{J}\left(\bar{R}_j^{TS}-\bar{R}^{CS}\right)\bar{\sigma}^2_{TS,j}
$$

B Model Solution

Substituting (13) and (16) into (15) implies:

$$
\begin{aligned}
& a_j(n) + b_{xj}(n)\,x_t + b_{\lambda j}(n)\,\lambda_t \\
& = \log \mathbb{E}_t \left[M_{t+1} \left(\frac{D_{j,t+1}}{D_{j,t}} \right) \exp\left\{ a_j(n-1) + b_{xj}(n-1)\,x_{t+1} + b_{\lambda j}(n-1)\,\lambda_{t+1} \right\} \right] \\
& = -r_f - \frac{1}{2}x_t^2 - \lambda_t \left(\mathbb{E}\left[e^{\zeta}\right] - 1 \right) + \mu_j - \frac{1}{2}(\sigma_j^i)^2 - \frac{1}{2}\beta_j^2\sigma_c^2 + a_j(n-1) \\
& \quad + \log \mathbb{E}_t \left[e^{(\beta_j\sigma_c - x_t)\epsilon_{c,t+1} + \sigma_j^i \epsilon_{j,t+1}^i + b_{xj}(n-1)x_{t+1} - \left(\beta_j^Q - 1\right)Q_{t+1} + Q_{j,t+1}^i + b_{\lambda j}(n-1)\left\{(1-\varphi_\lambda)\bar{\lambda} + \varphi_\lambda \lambda_t + Q_{\lambda,t+1}\right\}} \right] \\
& = -r_f - \frac{1}{2}x_t^2 - \lambda_t \left(\mathbb{E}\left[e^{\zeta}\right] - 1 \right) + \mu_j - \frac{1}{2}(\sigma_j^i)^2 - \frac{1}{2}\beta_j^2\sigma_c^2 + a_j(n-1) + b_{\lambda j}(n-1)\left\{(1-\varphi_\lambda)\,\bar{\lambda} + \varphi_\lambda\lambda_t\right\} \\
& \quad + \log \mathbb{E}_t \left[e^{(\beta_j\sigma_c - x_t)\epsilon_{c,t+1} + \sigma_j^i \epsilon_{j,t+1}^i + b_{xj}(n-1)x_{t+1} - \left(\beta_j^Q - 1\right)Q_{t+1} + b_{\lambda j}(n-1)Q_{\lambda,t+1} + Q_{j,t+1}^i} \right]
\end{aligned}
$$

Due to independence, the last term equals:

$$
\mathbb{E}_t \left[e^{(\beta_j\sigma_c - x_t)\epsilon_{c,t+1} + \sigma_j^i \epsilon_{j,t+1}^i + b_{xj}(n-1)x_{t+1}} \right] \mathbb{E}_t \left[e^{-\left(\beta_j^Q - 1\right)Q_{t+1}} \right] \mathbb{E}_t \left[e^{b_{\lambda j}(n-1)Q_{\lambda,t+1}} \right] \mathbb{E}_t \left[e^{Q_{j,t+1}^i} \right],
$$

where, using properties of the lognormal,

$$
\begin{aligned}
& \log \mathbb{E}_t \left[e^{(\beta_j\sigma_c - x_t)\epsilon_{c,t+1} + \sigma_j^i \epsilon_{j,t+1}^i + b_{xj}(n-1)x_{t+1}} \right] = \\
& \qquad b_{xj}(n-1)\left\{(1-\varphi)\,\bar{x} + \varphi x_t\right\} + \frac{1}{2}\left\{ \left(b_{xj}(n-1)\,\sigma_x\right)^2 + (\sigma_j^i)^2 + \left(\beta_j\sigma_c - x_t\right)^2 \right\} \qquad \text{(B.1)}
\end{aligned}
$$

and using properties of the compound Poisson,

$$
\begin{aligned}
\log \mathbb{E}\left[e^{-\left(\beta_j^Q - 1\right)Q_{t+1}} \right] &= \lambda_t \left(\mathbb{E}\left[e^{-\left(\beta_j^Q - 1\right)\zeta} \right] - 1 \right) \\
\log \mathbb{E}_t\left[e^{b_{\lambda j}(n-1)Q_{t+1}^{\lambda}} \right] &= \nu \left(\mathbb{E}\left[e^{b_{\lambda j}(n-1)\zeta^{\lambda}} \right] - 1 \right) \\
\log \mathbb{E}_t\left[e^{Q_{j,t+1}^i} \right] &= \lambda_j^i \left(\mathbb{E}\left[e^{\zeta_j^i} \right] - 1 \right).
\end{aligned}
$$

Combining the expressions:

$$
\begin{aligned}
&a_j\left(n\right)+b_{xj}\left(n\right)x_t+b_{\lambda j}\left(n\right)\lambda_t \\
&=-r_f+\mu_j+a_j\left(n-1\right)+b_{\lambda j}\left(n-1\right)\left(1-\varphi_\lambda\right)\bar{\lambda}+b_{xj}\left(n-1\right)\left(1-\varphi\right)\bar{x}+\frac{1}{2}\left(b_{xj}\left(n-1\right)\sigma_x\right)^2 \\
&\quad+\nu\left(\mathbb{E}\left[e^{b_{\lambda j}(n-1)\zeta^\lambda}\right]-1\right)+\lambda_j^i\left(\mathbb{E}\left[e^{\zeta_j^i}\right]-1\right) \\
&\quad+\left\{b_{xj}\left(n-1\right)\varphi-\beta_j\sigma_c\right\}x_t+\left\{b_{\lambda j}\left(n-1\right)\varphi_\lambda+\mathbb{E}\left[e^{-\left(\beta_j^Q-1\right)\zeta}\right]-\mathbb{E}\left[e^{\zeta}\right]\right\}\lambda_t
\end{aligned}
$$

Matching terms:

$$
\begin{aligned}
a_j\left(n\right)&=-r_f+\mu_j+a_j\left(n-1\right)+b_{\lambda j}\left(n-1\right)\left(1-\varphi_\lambda\right)\bar{\lambda} \\
&\quad+b_{xj}\left(n-1\right)\left(1-\varphi\right)\bar{x}+\frac{1}{2}\left(b_{xj}\left(n-1\right)\sigma_x\right)^2+\nu\left(\mathbb{E}\left[e^{b_{\lambda j}(n-1)\zeta^\lambda}\right]-1\right) \\
&\quad+\lambda_j^i\left(\mathbb{E}\left[e^{\zeta_j^i}\right]-1\right) \\
b_{xj}\left(n\right)&=b_{xj}\left(n-1\right)\varphi-\beta_j\sigma_c \\
b_{\lambda j}\left(n\right)&=b_{\lambda j}\left(n-1\right)\varphi_\lambda+\mathbb{E}\left[e^{-\left(\beta_j^Q-1\right)\zeta}\right]-\mathbb{E}\left[e^{\zeta}\right]
\end{aligned}
$$

Solving the recursion forward yields the formulas in the text.

References

Acharya, V. V., DeMarzo, P., and Kremer, I. (2011). Endogenous information flows and the clustering of announcements. *American Economic Review*, 101(7):2955–79.

Albuquerque, R. (2012). Skewness in stock returns: Reconciling the evidence on firm versus aggregate returns. *Review of Financial Studies*, 25(5):1630–1673.

Axtell, R. L. (2001). Zipf distribution of US firm sizes. *Science*, 293:1818–1820.

Bali, T. G., Cakici, N., and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99(2):427–446.

Bansal, R. and Yaron, A. (2004). Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3):823–866.

Bates, D. S. (1996). Testing option pricing models. In Maddala, G. S. and Rao, C. R., editors, *Handbook of Statistics*, volume 14, chapter 20, pages 567–611. Elsevier.

Beedles, W. L. (1979). Return, dispersion, and skewness: Synthesis and investment strategy. *Journal of Financial Research*, 2(1):71–80.

Bekaert, G., Engstrom, E., and Xu, N. (2019). The time variation in risk appetite and uncertainty. NBER Working Paper No. 25673.

Berk, J. B., Green, R. C., and Naik, V. (1999). Optimal investment, growth options, and security returns. *Journal of Finance*, 54(5):1553–1607.

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

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

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7:265–296.

Campbell, J. Y. (2003). Consumption-based asset pricing. In Constantinides, G., Harris, M., and Stulz, R., editors, *Handbook of the Economics of Finance, vol. 1b*, pages 803–887. Elsevier Science, North-Holland.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107:205–251.

Campbell, J. Y. and Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31(3):281–318.

Campbell, J. Y., Lettau, M., Malkiel, B. G., and Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance*, 56(1):1–43.

Campbell, J. Y. and Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review*, 94(5):1249–1275.

Chang, B. Y., Christoffersen, P., and Jacobs, K. (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, 107:46–68.

Cochrane, J. H., Longstaff, F. A., and Santa-Clara, P. (2008). Two trees. *Review of Financial Studies*, 21(1):347–385.

Colacito, R., Ghysels, E., Meng, J., and Siwasarit, W. (2016). Skewness in expected macro fundamentals and the predictability of equity returns: Evidence and theory. *The Review of Financial Studies*, 29(8):2069–2109.

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

Crouzet, N. and Mehrotra, N. R. (2020). Small and large firms over the business cycle. *American Economic Review*, 110(11):3549–3601.

Dew-Becker, I., Giglio, S., and Kelly, B. T. (2019). Hedging macroeconomic and financial uncertainty and volatility. Working paper, Northwestern University and Yale University.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance*, 57(1):369–403.

Dou, W. W. (2017). Embrace or fear uncertainty : Growth options, limited risk sharing, and asset prices. Working paper, University of Pennsylvania.

Drechsler, I. and Yaron, A. (2011). What's vol got to do with it. *Review of Financial Studies*, 24(1):1–45.

Duffee, G. R. (1995). Stock returns and volatility a firm-level analysis. *Journal of Financial Economics*, 37(3):399–420.

Fama, E. F. (1965). Portfolio analysis in a stable Paretian market. *Management science*, 11(3):404–419.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, volume I. John Wiley & Sons, New York, NY, 3rd edition.

Gabaix, X. (2009). Power laws in economics and finance. *Annual review of economics*, 1:255–293.

Gabaix, X. (2011). Disasterization: A simple way to fix the asset pricing properties of macroeconomic models. *American Economic Review*, 101(3):406–409.

Garleanu, N. and Panageas, S. (2018). Finance in a time of disruptive growth. Working paper, University of California at Berkeley and at Los Angeles.

Gomes, J., Kogan, L., and Zhang, L. (2003). Equilibrium cross section of returns. *Journal of Political Economy*, 111(4):693–732.

Gomez, M. (2019). Displacement and the rise in top wealth inequality. Working paper, Columbia University.

Gomez, M. and Gouin-Bonenfant, E. (2020). A Q-theory of inequality. Working paper, Columbia University.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2015). What do data on millions of U.S. workers reveal about life-cycle earnings risk? Working Paper 20913, National Bureau of Economic Research.

Harmenberg, K. and Sievertsen, H. H. (2017). The labor-market origins of cyclical income risk. Working paper, Copenhagen Business School and University of Bristol.

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

Herskovic, B., Kelly, B., Lustig, H., and Van Nieuwerburgh, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics*, 119(2):249–283.

Ilut, C., Kehrig, M., and Schneider, M. (2018). Slow to hire, quick to fire: Employment dynamics with asymmetric responses to news. *Journal of Political Economy*, 126(5):2011–2071.

Kapadia, N. (2006). The next Microsoft? skewness, idiosyncratic volatility, and expected returns. Working paper, Tulane University.

Kapadia, N. and Zekhnini, M. (2019). Do idiosyncratic jumps matter? *Journal of Financial Economics*, 131(3):666 – 692.

Khan, A. and Thomas, J. K. (2013). Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, 121(6):1055–1107.

Kogan, L., Papanikolaou, D., and Stoffman, N. (2019). Left behind: Creative destruction, inequality, and the stock market. *Journal of Political Economy*, 128(3):855–906.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101.

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

Lettau, M. and Wachter, J. A. (2007). Why is long-horizon equity less risky? a duration-based explanation of the value premium. *Journal of Finance*, 62:55–92.

Lintner, J. (1965). Security prices, risk and maximal gains from diversification. *Journal of Finance*, 20:587–615.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Martin, I. (2012). On the valuation of long-dated assets. *Journal of Political Economy*, 120(2):346–358.

Martin, I. (2013). The Lucas orchard. *Econometrica*, 81(1):55–111.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41:867–887.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, 2(1):130–168.

Salgado, S., Guvenen, F., and Bloom, N. (2019). Skewed business cycles. Working paper, University of Minnesota, University of Pennsylvania, and Stanford University.

Schmidt, L. (2016). Climbing and falling off the ladder: Asset pricing implications of labor market evenet risk. MIT Sloan Working Paper 5500-16.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–444.

Tsai, J. and Wachter, J. A. (2016). Rare booms and disasters in a multisector endowment economy. *Review of Financial Studies*, 29(5):1113–1169.

Wachter, J. A. and Zhu, Y. (2019). Learning with rare disasters. Working paper, University of Pennsylvania.

Xu, J. (2007). Price convexity and skewness. *The Journal of Finance*, 62(5):2521–2552.

Zhang, L. (2005). The value premium. *Journal of Finance*, 60:67–103.

Table 1: Statistics on Pooled Monthly Level Returns

	All CRSP 1926 – 2016	All CRSP 1945 – 2016	14,786 Select 1973 – 2016	404 Select 1973 – 2016
Mean (in %)	1.118	1.297	1.332	1.331
Median (in %)	0.000	0.000	0.000	0.936
Std. Dev (in %)	17.83	16.82	17.70	10.27
Skewness	6.335	5.987	5.986	1.321
% Positive	48.40	48.91	48.94	54.51
% $\geq$ 1-Month T-Bill	47.75	48.19	48.03	53.07
% $\geq$ VW Mkt Return	46.34	46.71	47.11	50.11
% $\geq$ EQ Mkt Return	45.83	46.19	47.01	49.26

Source: CRSP

Notes: The table reports selected statistics on pooled CRSP common stock monthly level returns for different time horizons and different universe of stocks. The first and second columns examine pooled monthly returns of all CRSP common stocks from July 1926 to December 2016 and November 1945 to December 2016, respectively. The third column concerns pooled monthly returns of all CRSP common stocks with at least 60 monthly returns from January 1973 to December 2016. The fourth column concerns pooled returns of all CRSP common stocks without missing data for monthly returns from January 1973 to December 2016.

Table 2: Statistics on Pooled Multi-Month Level Returns

Time Horizon (Months)	1	2	3	4	6	12
Mean (in %)	14.17	14.21	14.81	14.58	15.04	15.45
Median (in %)	0.000	2.564	4.444	4.782	5.373	6.061
Std. Dev (in %)	61.67	63.04	64.73	69.34	68.82	72.19
Skewness	6.349	6.258	5.333	13.36	10.08	7.075
% Positive	48.43	51.49	52.69	53.27	54.18	56.27
% $\geq$ 1-Month T-Bill	47.77	49.38	50.17	50.33	50.89	51.70
% $\geq$ VW Mkt Return	46.32	46.10	45.84	45.59	45.20	43.98
% $\geq$ EQ Mkt Return	45.83	45.41	44.82	44.47	43.48	41.55

Source: CRSP

Notes: The table reports selected statistics on pooled CRSP common stock monthly level returns of different horizons from July 1926 to July 2017. The n-month return is computed by compounding n monthly returns, starting from July 1926, in a non-overlapping manner. We exclude returns for which the compounding is infeasible due to missing observations. The reported mean, median, and standard deviation are annualized numbers.

Table 3: Skewness in Time-series, Cross-section, & Pooled Distribution of Returns

Skewness Type	Statistic	404 Firms		14,786 Firms	
		skew[R]	skew[$\log R$]	skew[R]	skew[$\log R$]
Time-series	Min.	-1.010	-4.018	-3.662	-17.20
	5th	-0.188	-1.128	-0.179	-1.461
	50th	0.362	-0.204	0.891	0.021
	95th	2.118	0.376	3.627	1.181
	Max.	7.583	1.640	16.36	6.330
	S&P500	-0.441	-0.707	-0.441	-0.707
Cross-section	Min.	-2.572	-10.09	-0.288	-4.146
	5th	-0.689	-1.612	0.393	-1.708
	50th	0.694	0.105	2.400	-0.085
	95th	3.351	1.673	8.486	1.520
	Max.	12.88	5.794	46.00	2.656
Pooled		1.321	-0.423	5.987	-0.240

Source: CRSP

Notes: The table reports the skewness in the time-series, the cross-section, and the pooled distribution of monthly returns for different time horizons and different universe of stocks. The first and second data columns examine monthly returns of all CRSP common stocks without missing data for monthly returns from January 1973 to December 2016. The third and fourth data columns concern monthly returns of all CRSP common stocks with at least 60 monthly returns from January 1973 to December 2016. For time-series skewness, we report the distribution across different assets as well as that of the S&P 500 during the same time period. For cross-sectional skewness, we report the distribution across different months for each relevant time period.

Table 4: Calibrated Parameters of the Model

Variable	Value
$\bar{x}$	0.198
β_j, β_j^Q	Estimated from CAPM Regression
φ	0.988
σ_c	0.0419
σ_x	0.0693
$\sigma_{i,j}$	Set to match $\sigma_j^2 = \sigma_{i,j}^2 + \beta_j^2 \sigma_M^2$
r_f	0.00395
μ_j	Set to match $\mathbb{E}[\Delta d_{j,t+1}]$ to average S&P500 dividend growth
ζ	0.15

Source: CRSP, Ken French's Website

Notes: The table reports the calibrated parameters of the model, which is simulated at a monthly frequency. $\bar{x}$ is chosen such that when x_t is at its long-run mean, the maximal Sharpe ratio is 0.20. β_j and β_j^Q are estimated from a CAPM regression of log returns, and the persistence (φ) is set to match the autocorrelation of the price-dividend ratio. σ_c and σ_x are chosen as in in Lettau and Wachter (2007). μ_js are calibrated to match $\mathbb{E}[\Delta d_{j,t+1}]$ to the average dividend growth of S&P500 in the data. Size of the aggregate jump (ζ) is set to 0.15.

Table 5: Results from Simulated Method of Moments (SMM) Estimation

Parameter		LNM-N	LNM-NP
	Estimation with 14,786 Firms		
$\bar{\lambda}$	Mean for the law of motion for λ_t	0.005	0.005
φ^{λ}	Persistence in the law of motion for λ_t	0.990	0.963
ν	Frequency of Compound Poisson shock Q^{λ}_{t+1}	0.004	0.004
ζ^{λ}	Size of Compound Poisson shock Q^{λ}_{t+1}	0.133	0.133
k	Scaling parameter	·	1.500
	Estimation with 404 Firms		
$\bar{\lambda}$	Mean for the law of motion for λ_t	0.005	0.0037
φ^{λ}	Persistence in the law of motion for λ_t	0.990	0.990
ν	Frequency of Compound Poisson shock Q^{λ}_{t+1}	0.0037	0.005
ζ^{λ}	Size of Compound Poisson shock Q^{λ}_{t+1}	0.200	0.134
k	Scaling parameter	·	1.000

Source: CRSP, Simulations

Notes: The table reports the results of the SMM estimation for the lognormal-N (LNM-N) and the lognormal-NP (LNM-NP) models, which chooses uncalibrated model parameters by matching the moments from a simulated panel of economy to the corresponding moments from the data. The first panel reports the parameters for the set of 14,786 firms, and the second panel reports the parameters for the set of 404 firms. The simulations are at a monthly frequency.

Table 6: Inference on Market Returns

	Empirical Value	Simulated Values					Simulation Population
		Min	5th	50th	95th	Max	
Panel A. Lognormal Model							
$E[R_m] - 1$	0.922	-3.028	-1.304	0.241	1.871	2.512	0.535
$\sigma[R_m]$	4.550	3.950	4.110	4.420	4.910	5.090	4.530
$\gamma^{TS}[R_m]$	-0.518	-0.281	-0.066	0.139	0.441	0.488	0.168
Panel B. Lognormal-N Model							
$E[R_m] - 1$	0.922	-2.530	-0.620	0.859	2.611	3.214	1.173
$\sigma[R_m]$	4.550	4.200	4.290	4.700	6.130	6.600	5.450
$\gamma^{TS}[R_m]$	-0.518	-1.524	-1.068	-0.198	0.348	0.561	-0.638
Panel C. Lognormal-NP Model							
$E[R_m] - 1$	0.922	-2.093	-0.145	1.221	2.891	3.201	1.163
$\sigma[R_m]$	4.550	4.220	4.370	4.930	6.480	9.280	6.530
$\gamma^{TS}[R_m]$	-0.518	-1.365	-1.036	-0.294	0.280	0.353	-0.878

Source: CRSP and simulations

Notes: We conduct 100 monthly simulations of the stock market for each type of model assuming 14,786 firms. Sampling distribution of each statistic is obtained from the simulations of the reference asset. The first column shows the statistic for the corresponding value in data where the moments are computed from the returns on Fama-French's market portfolio; the next five columns show the distribution of the statistic obtained from the simulations, and the last column illustrates the statistic for the pooled values of 50 simulations. $E[R] - 1$ and $\sigma[R]$ are reported in percentages. $\gamma^{TS}[R_m]$ denotes the time-series skewness of returns on the market.

Table 7: Inference on Pooled Monthly Returns (Simulation with 14,786 Firms)

	Empirical Value	Simulated Values: Min	5th	50th	95th	Max	Simulation Population
Panel A. Lognormal Model							
$E[R]-1$	1.332	-3.550	-1.578	0.230	2.119	2.890	0.912
$\sigma[R]$	17.70	15.89	16.26	16.63	17.10	17.33	16.78
$\gamma^{CS}[R]$	5.986	0.739	0.809	0.894	0.978	1.005	0.929
$\gamma^{CS}[\log R]$	-0.235	-0.319	-0.248	-0.177	-0.102	-0.079	-0.156
$\% \log R > \log R_f$	48.03	36.09	41.63	46.45	51.65	53.58	48.38
$\tilde{\gamma}^{TS}[R_j]$	0.875	0.397	0.407	0.424	0.450	0.458	·
Panel B. Lognormal-N Model							
$E[R]-1$	1.332	-2.977	-0.792	0.983	2.981	3.736	0.723
$\sigma[R]$	17.70	16.35	16.61	17.12	17.71	18.13	16.99
$\gamma^{CS}[R]$	5.986	0.712	0.824	0.942	1.063	1.083	0.909
$\gamma^{CS}[\log R]$	-0.235	-0.549	-0.343	-0.193	-0.092	-0.067	-0.209
$\% \log R > \log R_f$	48.03	37.94	43.85	48.62	54.18	55.63	47.98
$\tilde{\gamma}^{TS}[R_j]$	0.875	0.256	0.327	0.405	0.437	0.461	·
Panel C. Lognormal-NP Model							
$E[R]-1$	1.332	-2.453	-0.281	1.403	3.318	3.706	1.566
$\sigma[R]$	17.70	18.08	18.24	18.84	19.79	21.26	20.53
$\gamma^{CS}[R]$	5.986	2.499	2.975	5.325	22.82	55.59	81.02
$\gamma^{CS}[\log R]$	-0.235	-0.595	-0.164	0.003	0.100	0.132	-0.190
$\% \log R > \log R_f$	48.03	38.60	45.02	49.32	54.30	55.29	49.91
$\tilde{\gamma}^{TS}[R_j]$	0.875	0.325	0.460	0.537	0.571	0.576	·

Source: CRSP and simulations

Notes: We conduct 100 monthly simulations of the stock market for each type of model using the universe of 14,786 firms with at least 60 monthly returns from January 1973 to December 2016. Sampling distribution of each statistic is obtained from the simulations. The first column shows the statistic for the corresponding value in data; the next five columns show the distribution of the statistic obtained from the simulations, and the last column illustrates the statistic for the pooled values of 50 simulations. $E[R]-1$ and $\sigma[R]$ are reported in percentages. γ^{CS} denotes cross-sectional skewness across pooled returns, while $\gamma^{TS}[R_j]$ denotes the time-series skewness of returns for firm j. We also report the $\tilde{\gamma}^{TS}[R_j]$, the median value of $\gamma^{TS}[R_j]$ across the 14,786 firms.

Table 8: Inference on Pooled Monthly Returns (Simulation with 404 Firms)

	Empirical	Simulated Values					Simulation
	Value	Min	5th	50th	95th	Max	Population
Panel A. Lognormal Model							
$E[R]-1$	1.331	-2.367	-0.521	1.056	2.763	3.422	0.749
$\sigma[R]$	10.27	9.750	9.920	10.12	10.44	10.51	10.15
$\gamma^{CS}[R]$	1.321	0.273	0.379	0.476	0.579	0.620	0.464
$\gamma^{CS}[\log R]$	-0.432	-0.213	-0.127	-0.033	0.072	0.107	-0.045
$\% \log R > \log R_f$	54.51	36.83	44.33	50.71	57.72	60.24	49.53
$\tilde{\gamma}^{TS}[R_j]$	0.362	0.230	0.243	0.273	0.309	0.340	·
Panel B. Lognormal-N Model							
$E[R]-1$	1.331	-2.456	-0.403	1.249	2.969	3.991	2.554
$\sigma[R]$	10.27	9.890	10.16	10.78	11.76	13.44	10.88
$\gamma^{CS}[R]$	1.321	-0.126	0.122	0.352	0.497	0.602	0.431
$\gamma^{CS}[\log R]$	-0.432	-0.947	-0.571	-0.205	-0.019	0.064	-0.139
$\% \log R > \log R_f$	54.51	39.10	46.580	52.28	58.91	62.44	57.13
$\tilde{\gamma}^{TS}[R_j]$	0.362	-0.186	-0.045	0.148	0.270	0.283	·
Panel C. Lognormal-NP Model							
$E[R]-1$	1.331	-1.123	0.577	2.277	3.929	4.940	2.102
$\sigma[R]$	10.27	11.38	11.62	12.23	13.18	13.58	12.98
$\gamma^{CS}[R]$	1.321	0.999	1.202	1.613	2.089	4.539	1.782
$\gamma^{CS}[\log R]$	-0.432	-0.435	-0.273	0.036	0.282	0.322	-0.002
$\% \log R > \log R_f$	54.51	43.22	49.76	55.71	61.83	65.52	54.79
$\tilde{\gamma}^{TS}[R_j]$	0.362	-0.072	0.019	0.160	0.264	0.296	·

Source: CRSP and simulations

Notes: We conduct 100 monthly simulations of the stock market for each type of model using the universe of 404 firms with no missing returns from January 1973 to December 2016. Sampling distribution of each statistic is obtained from the simulations. The first column shows the statistic for the corresponding value in data; the next five columns show the distribution of the statistic obtained from the simulations, and the last column illustrates the statistic for the pooled values of 50 simulations. $E[R]-1$ and $\sigma[R]$ are reported in percentages. γ^{CS} denotes cross-sectional skewness across pooled returns, while $\gamma^{TS}[R_j]$ denotes the time-series skewness of returns for firm j. We also report the $\tilde{\gamma}^{TS}[R_j]$, the median value of $\gamma^{TS}[R_j]$ across the 404 firms.

Table 9: Inference on Monthly Cross-sectional Skew

	Empirical $\tilde{\gamma}_{cs}$	$\tilde{\gamma}_{cs}$ from Simulated Values					% of Months with Empirical γ_{cs} $\geq$ Max
		Min	5th	50th	95th	Max	
Panel A. Simulation with 14,786 Firms (1973.01 - 2016.12)							
Lognormal	2.400	0.735	0.796	0.862	0.911	0.942	85.42
Lognormal-N	2.400	0.794	0.887	0.945	1.005	1.052	83.90
Lognormal-NP	2.400	1.693	1.719	1.806	1.882	1.949	62.31
Panel B. Simulation with 404 Firms (1973.01 - 2016.12)							
Lognormal	0.694	0.207	0.286	0.381	0.453	0.495	59.47
Lognormal-N	0.694	0.237	0.316	0.397	0.478	0.531	57.39
Lognormal-NP	0.694	0.585	0.628	0.756	0.839	0.884	41.48

Source: CRSP and simulations

Notes: We conduct 100 monthly simulations of the stock market for the three models. The first column shows the median monthly cross-sectional skewness ($\tilde{\gamma}_{cs}$) for the corresponding universe of stocks and sample period. The next five columns illustrate the distribution of $\tilde{\gamma}_{cs}$ obtained from simulations. The final column reports the percentage of months in the sample period in which empirical γ_{cs} (obtained from the data) is greater than the maximum $\tilde{\gamma}_{cs}$ obtained from the simulations.

Table 10: Estimates of Expected Returns in Stylized Simulations

	5th	50th	95th	Mean	True $E[R_j]-1$
		$T=10$			
Lognormal	-3.050	0.941	5.507	1.062	1.063
Lognormal-N	-3.236	0.957	5.350	1.062	1.072
Lognormal-NP	-3.148	0.929	5.406	1.061	1.072
		$T=50$			
Lognormal	-0.877	1.008	3.015	1.060	1.063
Lognormal-N	-0.857	1.020	3.034	1.071	1.072
Lognormal-NP	-0.856	1.100	3.003	1.081	1.072
		$T=500$			
Lognormal	0.457	1.062	1.641	1.059	1.063
Lognormal-N	0.443	1.078	1.646	1.069	1.072
Lognormal-NP	0.457	1.067	1.692	1.070	1.072

Source: Simulations

Notes: We simulate 1,000 samples of length T under the null hypothesis of the lognormal, the lognormal with rare negative events (lognormal-N) and the lognormal with rare negative and positive events (lognormal-NP). The table shows the distribution of the sample mean as well as the true mean.

Table 11: Estimates of Standard Deviations in Stylized Simulations

	5th	50th	95th	Mean	True $\sigma[R_j]$
			$T = 10$		
Lognormal	5.022	8.260	11.714	8.301	8.595
Lognormal-N	5.054	8.227	11.728	8.302	8.595
Lognormal-NP	5.000	8.123	12.056	8.254	8.595
			$T = 50$		
Lognormal	7.037	8.534	10.106	8.569	8.595
Lognormal-N	7.040	8.514	10.098	8.567	8.595
Lognormal-NP	6.965	8.508	10.285	8.546	8.595
			$T = 500$		
Lognormal	8.145	8.584	9.050	8.587	8.595
Lognormal-N	8.136	8.580	9.043	8.585	8.595
Lognormal-NP	8.073	8.566	9.133	8.585	8.595

Source: Simulations

Notes: We simulate 1,000 samples of length T under the null hypothesis of the lognormal, the lognormal with rare negative events (lognormal-N) and the lognormal with rare negative and positive events (lognormal-NP). The table shows the distribution of the sample standard deviation as well as the true standard deviation.

Table 12: Estimates of Skewness in Stylized Simulations

	5th	50th	95th	Mean	True Skewness
			$T = 10$		
Lognormal	-0.921	0.129	1.311	0.165	0.255
Lognormal-N	-0.927	0.118	1.329	0.156	0.243
Lognormal-NP	-0.914	0.167	1.457	0.207	0.434
			$T = 50$		
Lognormal	-0.320	0.229	0.898	0.241	0.255
Lognormal-N	-0.339	0.217	0.887	0.233	0.243
Lognormal-NP	-0.287	0.294	1.210	0.358	0.434
			$T = 500$		
Lognormal	0.067	0.244	0.460	0.255	0.255
Lognormal-N	0.059	0.236	0.442	0.243	0.243
Lognormal-NP	0.155	0.411	0.776	0.427	0.434

Source: Simulations

Notes: We simulate 1,000 samples of length T under the null hypothesis of the lognormal, the lognormal with rare negative events (lognormal-N) and the lognormal with rare negative and positive events (lognormal-NP). The table shows the distribution of the sample skewness as well as the true skewness.

Table 13: Model Summary

Stylized Facts	Lognormal	Lognormal-N	Lognormal-NP
A. Monthly Returns (Stocks)			
Risk Premium	✓	✓	✓
Volatility	✓	✓	✓
% > Risk-free Rate	✓	✓	✓
TS Skewness (> 0) γ_j^{TS}	✓	✓	✓
Pooled Skewness γ^{CS}	×	×	✓
Monthly cross-sectional skewness $\gamma_{cs,t}$	×	×	✓
B. Monthly Returns (Market)			
Risk Premium	✓	✓	✓
Volatility	✓	✓	✓
TS Skewness for the Market (< 0) γ_M^{TS}	×	✓	✓
C. Long-run Returns			
Positive CS Skewness	✓	✓	✓
Stationary Distribution of Firm Size	×	×	×

Source: CRSP, Simulations

Notes: The table summarizes the performance of the three models in matching the observed stylized facts from data. Panel A is relevant to moments of pooled monthly returns across stocks, and Panel B pertains to analogous moments for the market returns. Panel C pertains to distributions of long-run returns and firm sizes across all stocks.

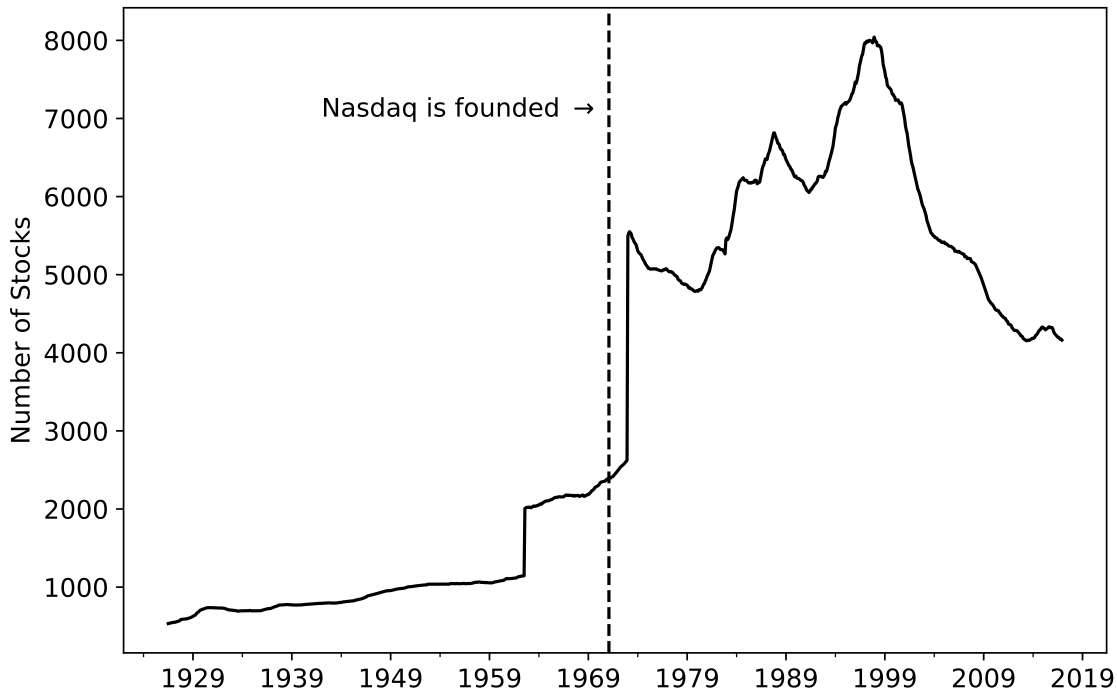


Figure 1: Historical Number of CRSP Common Stocks
On the first day of each month from July 1926 to December 2016, we count the number of unique common stocks in the cross-section, as available in CRSP. The jump on January 1973, from 2,623 to 5,494, roughly corresponds to the establishment of Nasdaq in February of 1971.

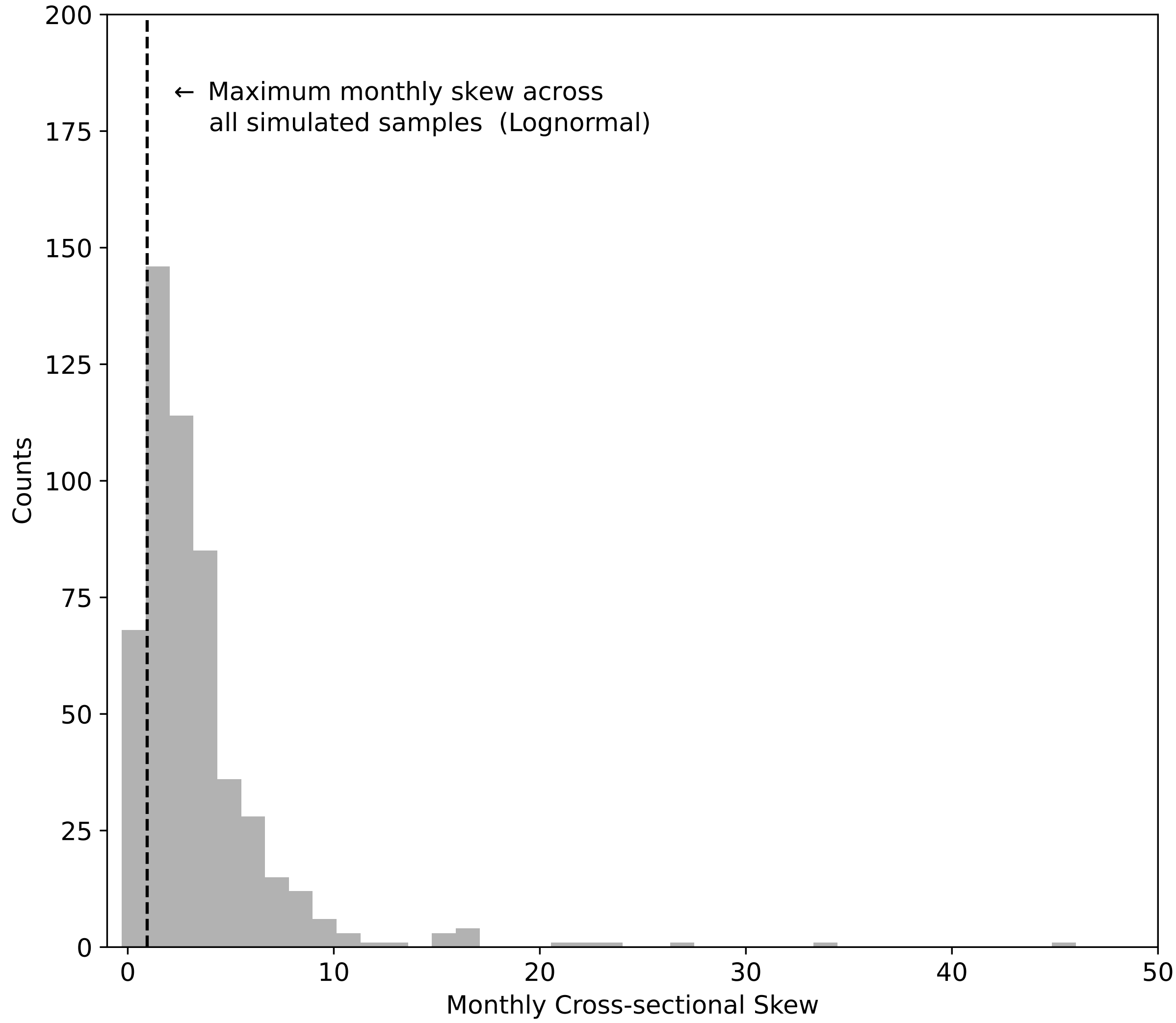


Figure 2: Distribution of Monthly Cross-sectional Skewness (14,786 Firms)
The figure illustrates the distribution of monthly cross-sectional skewness, defined as the skewness of monthly level returns for the cross-section of firms in each given month. The graph pertains to set of 14,786 firms with at least 60 monthly returns from January 1973 to December 2016. The vertical line on the graph represents the maximum of the average monthly cross-sectional skewness obtained from the 100 simulations.

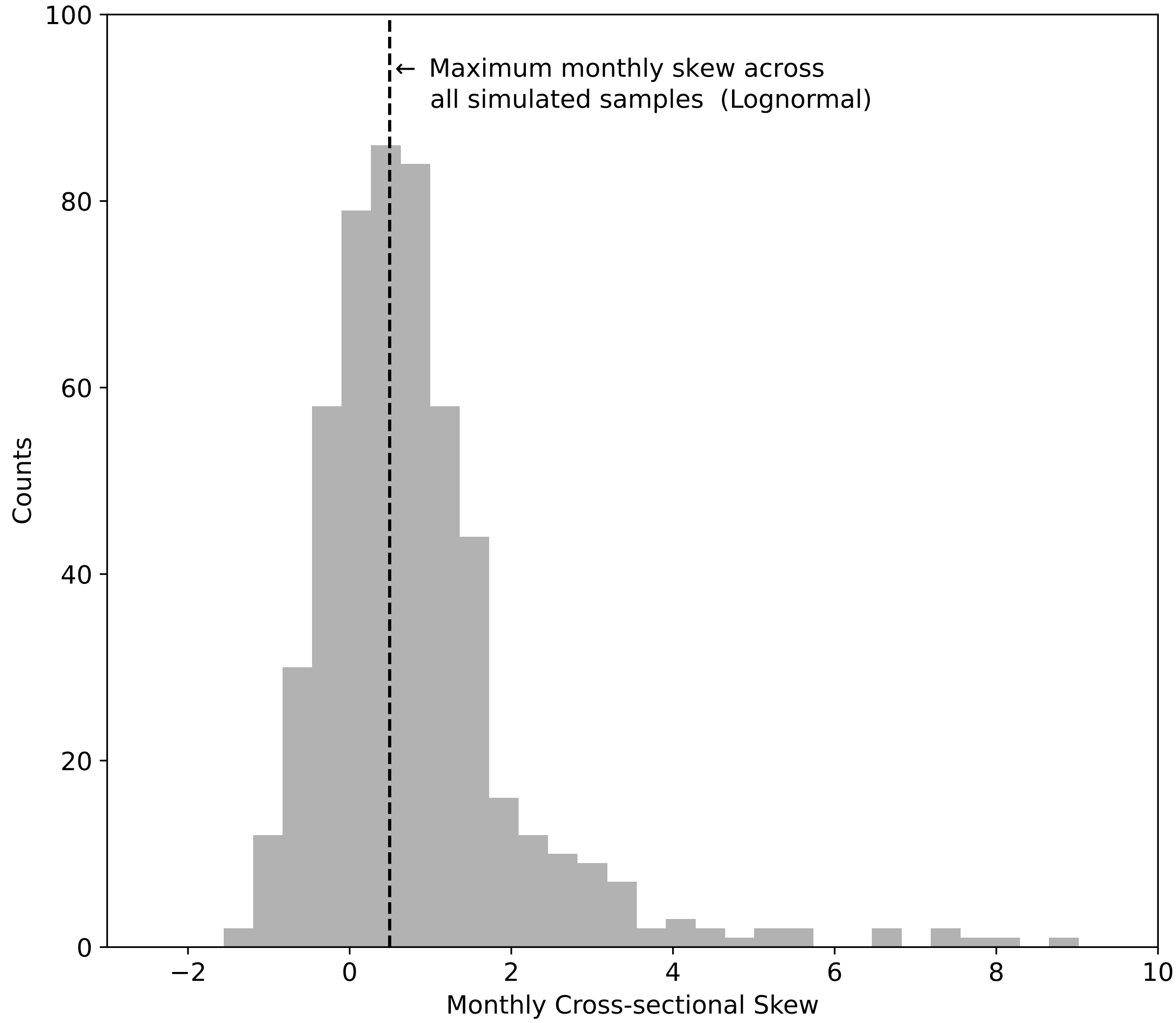


Figure 3: Distribution of Monthly Cross-sectional Skewness (404 Firms)
The figure illustrates the distribution of monthly cross-sectional skewness, defined as the skewness of monthly level returns for the cross-section of firms in each given month. The graph pertains to set of 404 firms without missing data for monthly returns from January 1973 to December 2016. The vertical line on the graph represents the maximum of the average monthly cross-sectional skewness obtained from the 100 simulations.

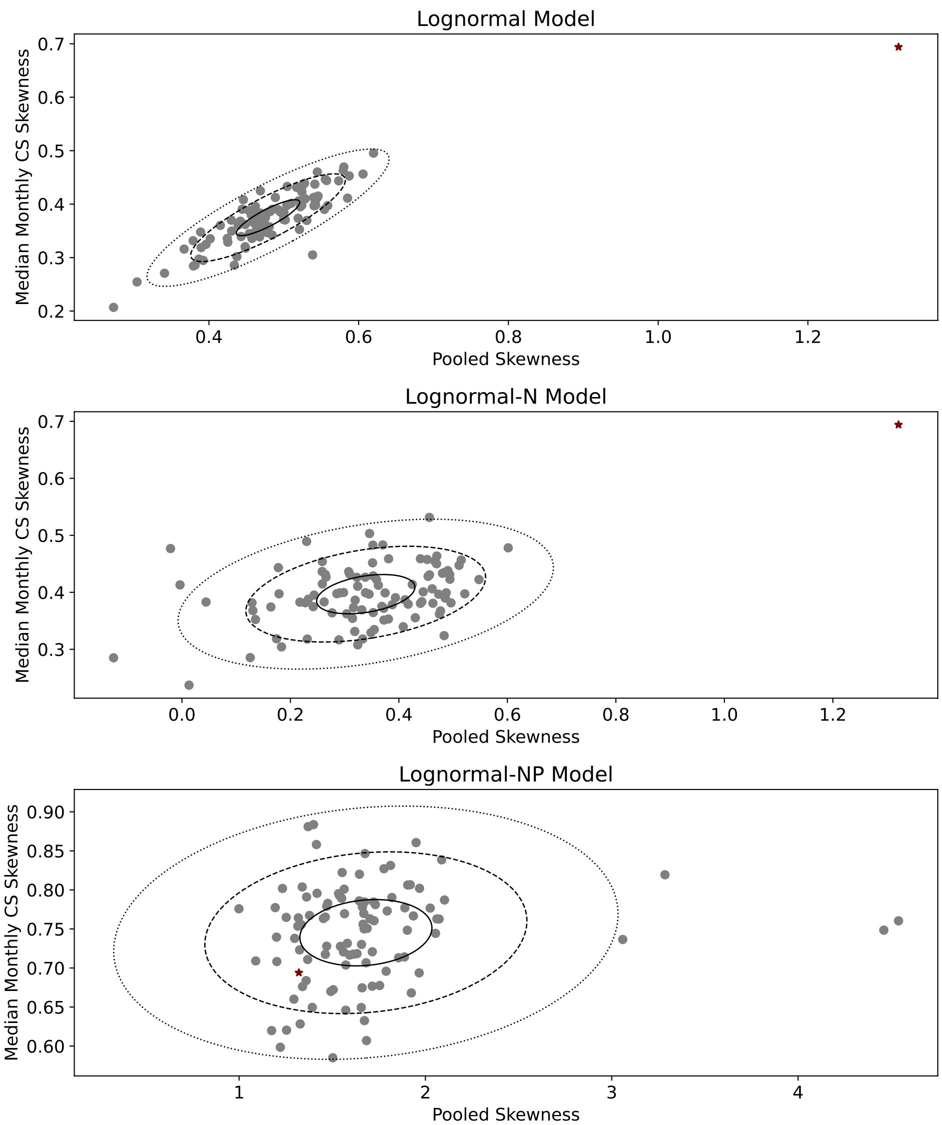


Figure 4: Joint Tests of Pooled Skewness and Monthly CS Skewness (404 Firms)
This figure plots the pooled cross-sectional skewness against the median monthly cross-sectional skewness from 100 simulations. Each gray circle represents the moment from one simulated sample, while the red star represents the data value. The ellipse with the solid border represents the 50% confidence level based on simulation samples, followed by 95% and 99% confidence ellipses represented by the dotted borders.

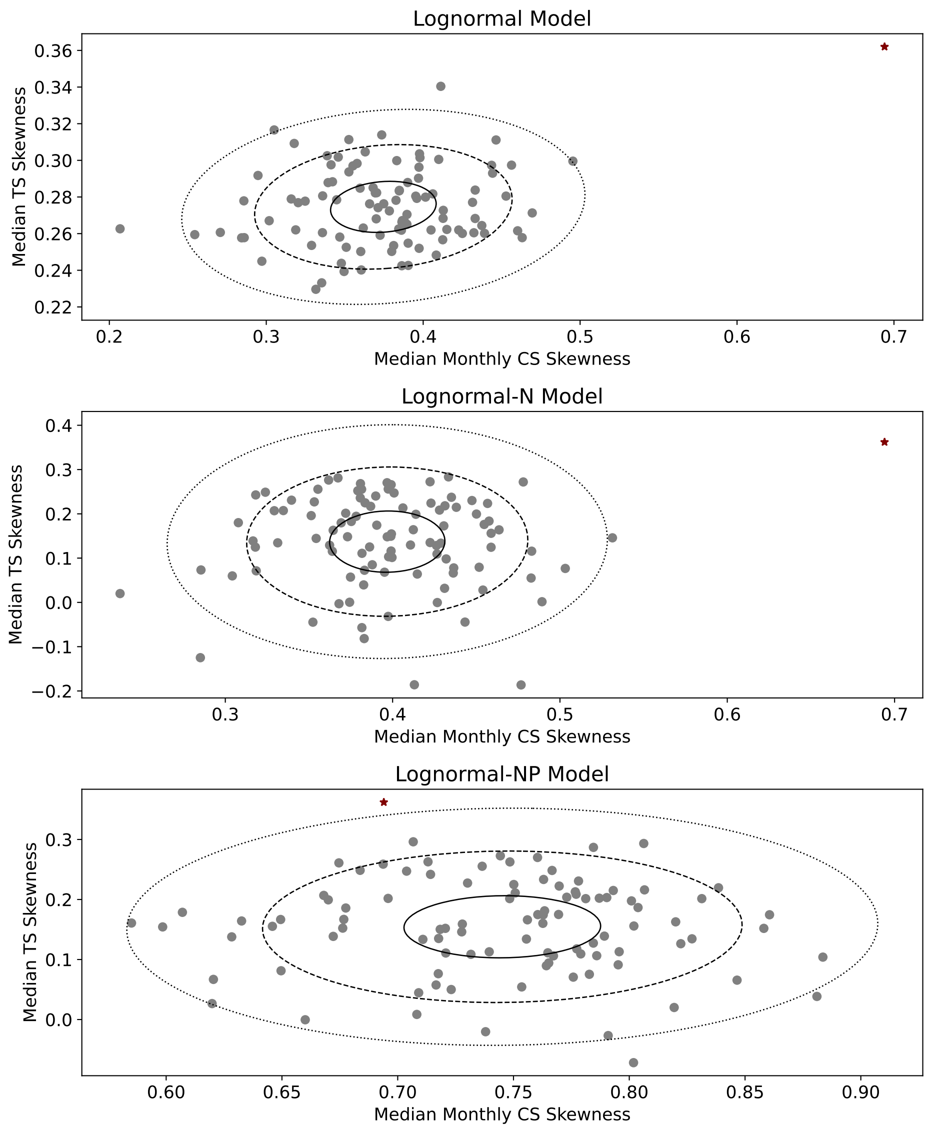


Figure 5: Joint Tests of Monthly CS Skewness and TS Skewness (404 Firms)
This figure plots the median monthly cross-sectional skewness against the median time-series skewness from 100 simulations. Each gray circle represents the moment from one simulated sample, while the red star represents the data value. The ellipse with the solid border represents the 50% confidence level based on simulation samples, followed by 95% and 99% confidence ellipses represented by the dotted borders.

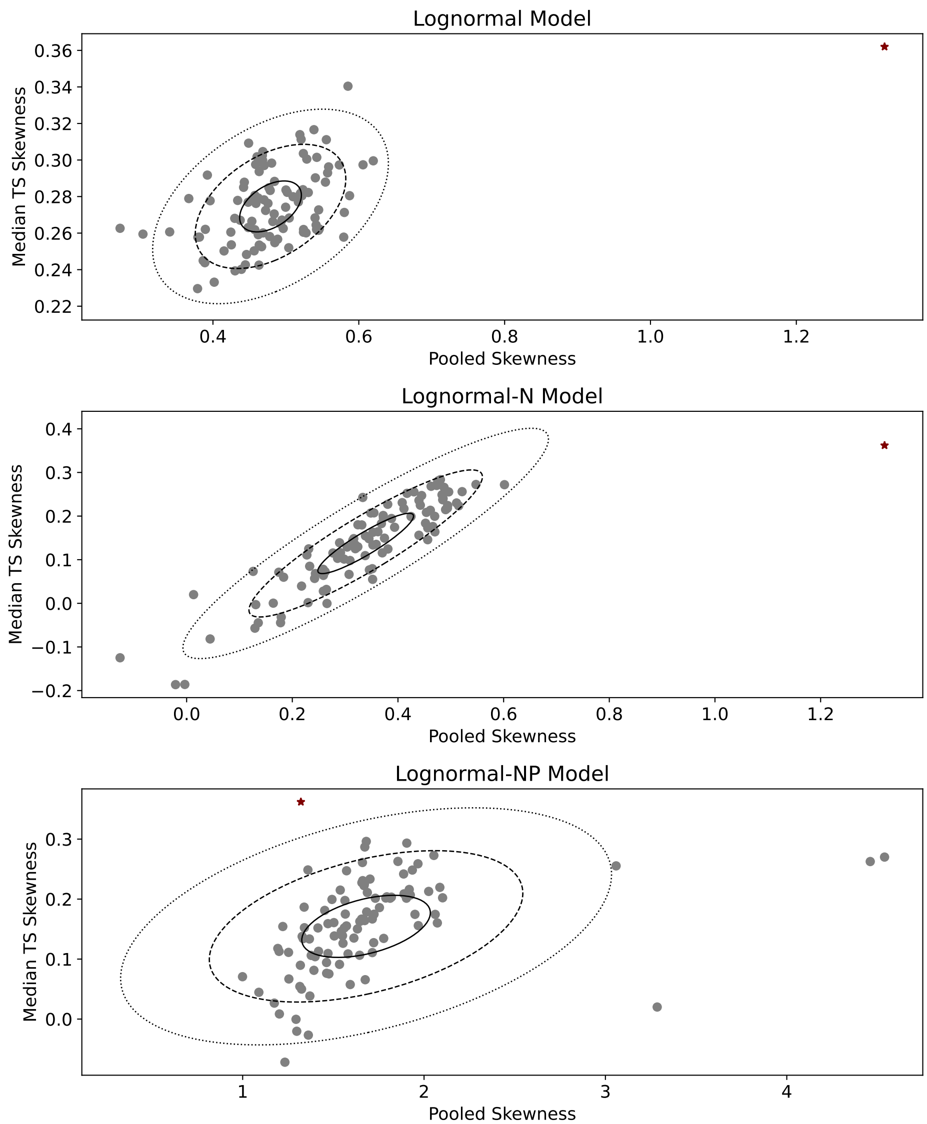


Figure 6: Joint Tests of Pooled and TS Skewness (404 Firms)
This figure plots the pooled cross-sectional skewness against the median time-series skewness from 100 simulations. Each gray circle represents the moment from one simulated sample, while the red star represents the data value. The ellipse with the solid border represents the 50% confidence level based on simulation samples, followed by 95% and 99% confidence ellipses represented by the dotted borders.

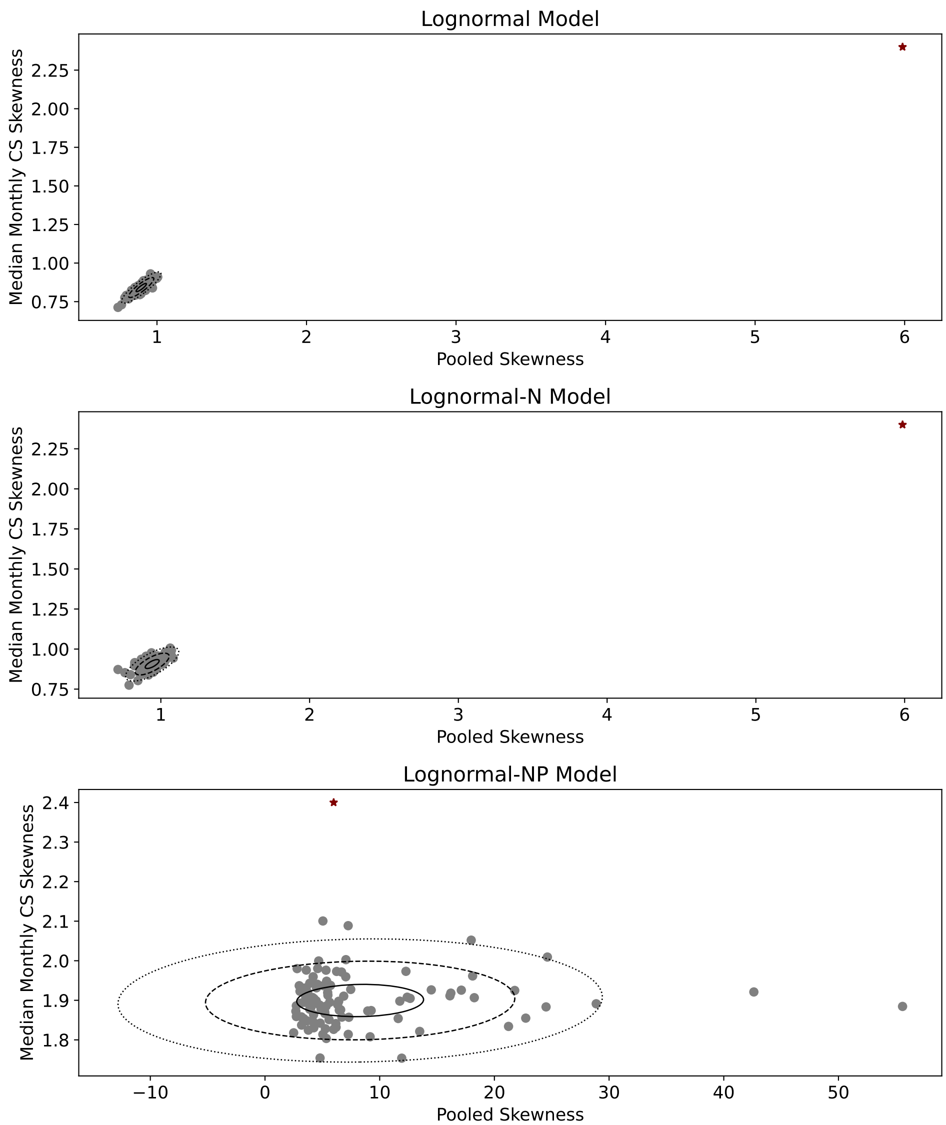


Figure 7: Joint Tests of Pooled Skewness and Monthly CS Skewness (14,786 Firms)

This figure we plots the pooled cross-sectional skewness against the median monthly cross-sectional skewness from 100 simulations. Each gray circle represents the moment from one simulated sample, while the red star represents the data value. The ellipse with the solid border represents the 50% confidence level based on simulation samples, followed by 95% and 99% confidence ellipses represented by the dotted borders.

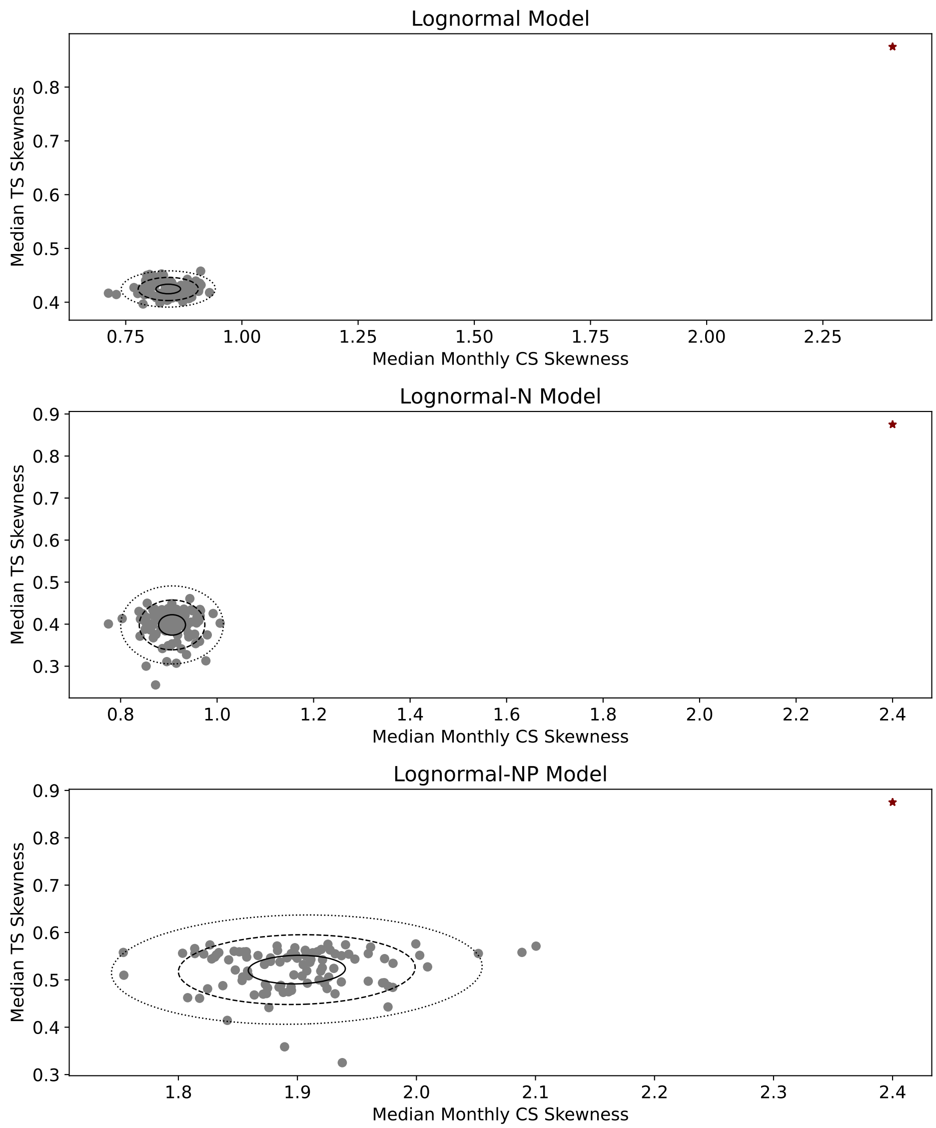


Figure 8: Joint Tests of Monthly CS Skewness and TS Skewness (14,786 Firms) This figure plots the median monthly cross-sectional skewness against the median time-series skewness from 100 simulations. Each gray circle represents the moment from one simulated sample, while the red star represents the data value. The ellipse with the solid border represents the 50% confidence level based on simulation samples, followed by 95% and 99% confidence ellipses represented by the dotted borders.

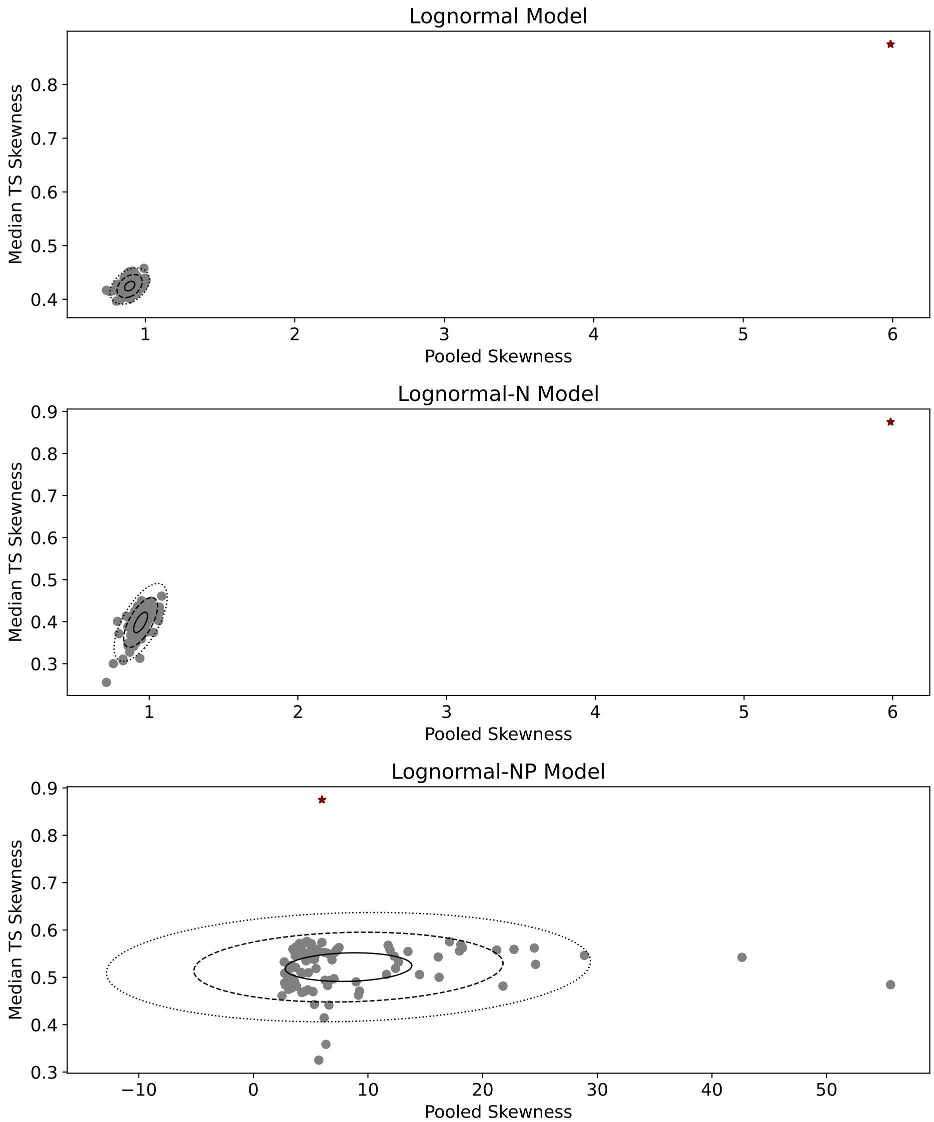
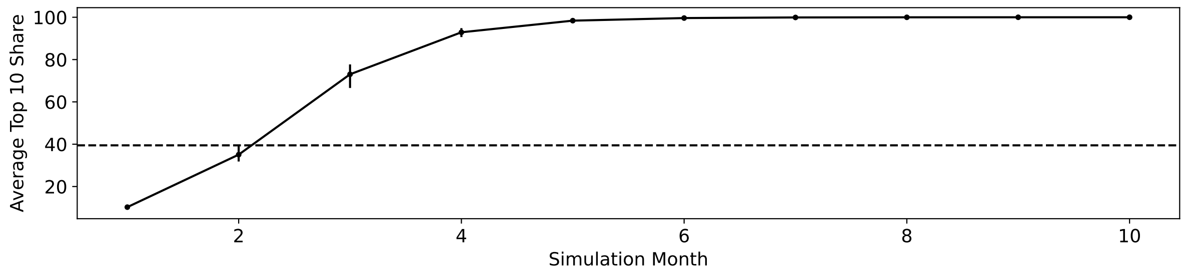
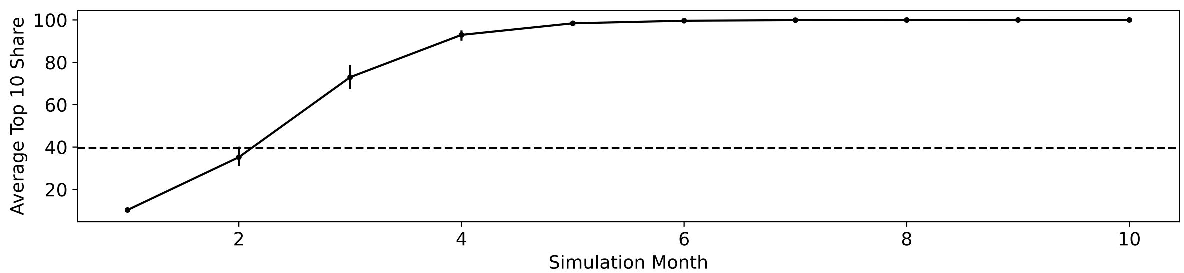


Figure 9: Joint Tests of Pooled and TS Skewness (14,786 Firms)
This figure plots the pooled cross-sectional skewness against the median time-series skewness from 100 simulations. Each gray circle represents the moment from one simulated sample, while the red star represents the data value. The ellipse with the solid border represents the 50% confidence level based on simulation samples, followed by 95% and 99% confidence ellipses represented by the dotted borders.

Panel A. Lognormal Model



Panel B. Lognormal-N Model



Panel C. Lognormal-NP Model

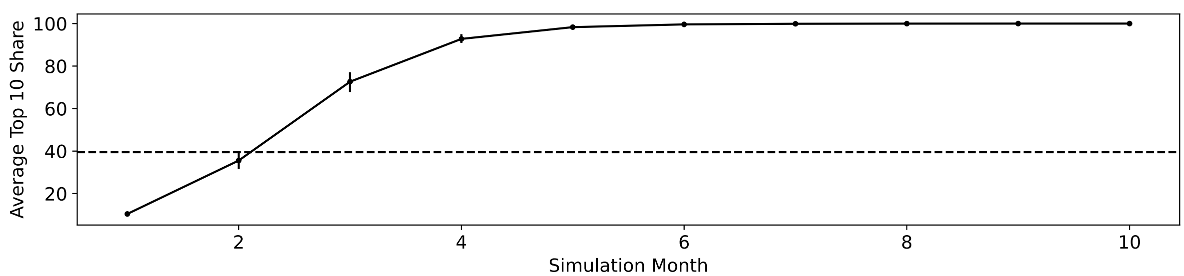
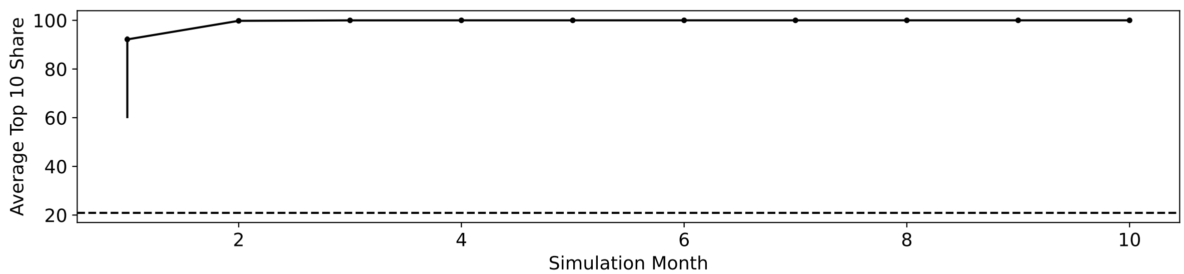
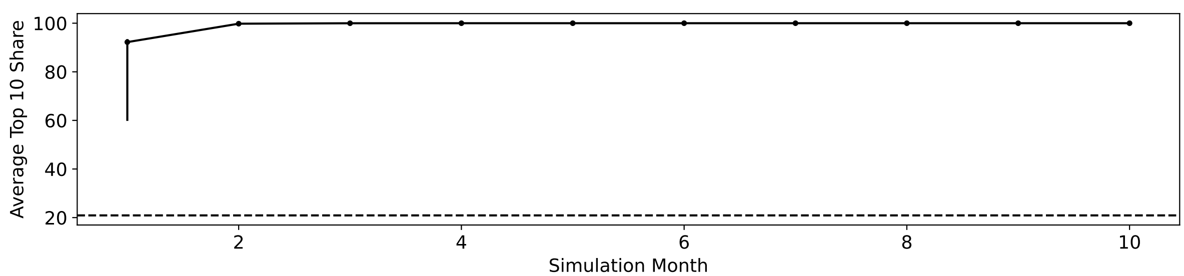


Figure 10: Average Top 10 Share in Market Capitalization (404 Firms)
The figure illustrates the median top 10 share in market capitalization computed at each simulation month for the first 10 months. The error bars represent the 5th and 95th percentile values across simulations. The graph pertains to set of 404 firms without missing data for monthly returns from January 1973 to December 2016, and the horizontal line represents the corresponding data value.

Panel A. Lognormal Model



Panel B. Lognormal-N Model



Panel C. Lognormal-NP Model

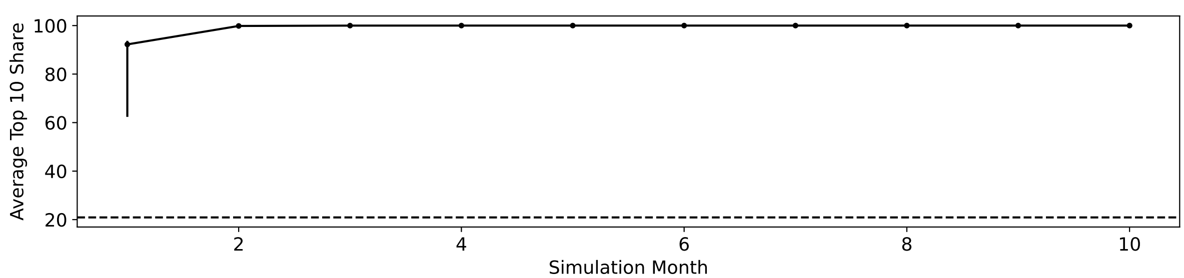


Figure 11: Average Top 10 Share in Market Capitalization (14,786 Firms)
The figure illustrates the median top 10 share in market capitalization computed at each simulation month for the first 10 months. The error bars represent the 5th and 95th percentile values across simulations. The graph pertains to set of 14,786 firms with at least 60 monthly returns from January 1973 to December 2016, and the horizontal line represents the corresponding data value.

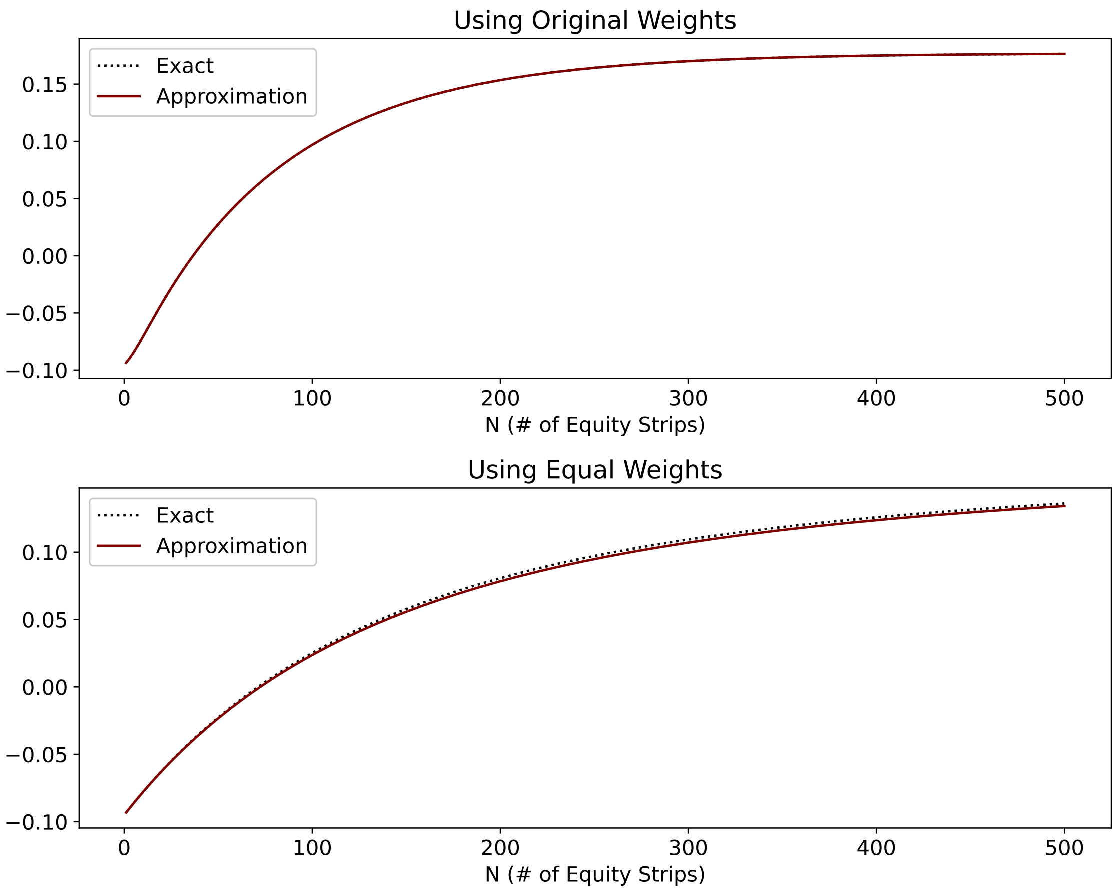


Figure 12: Evaluation of Log Approximation Assumption
This figure plots the exact log equity return computed using (24) with strips of $N = 1$ to 500 maturities and the approximate log equity return computed using (27) for a sample month. Using the calibrated and estimated parameters for the lognormal-NP model, we simulate returns to equity strips of varying maturities for a single stock j. For simplicity, we set $\beta_j = \beta_j^Q = 0.8$ and set $\sigma_j = 0.078$, the median value in the calibration from the set of 404 firms. λ_j and ζ_j are also set to $\bar{\lambda}$ and ζ respectively.

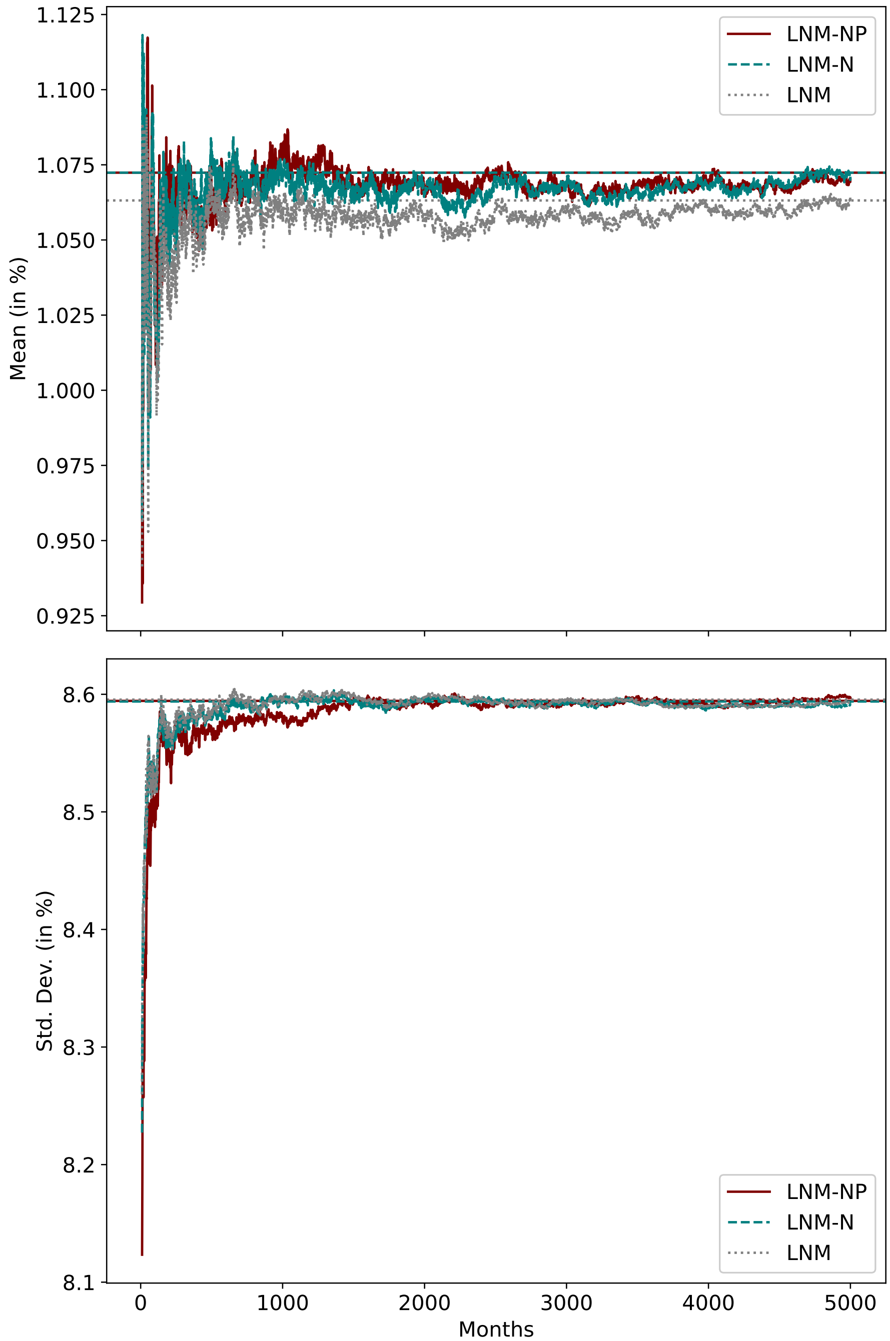


Figure 13: Estimates of Mean and Standard Deviation in Stylized Simulations

This figure plots the median estimate of the time-series mean and standard deviation across 1,000 stylized simulations as a function of the number of months (T). The horizontal lines denote the true value for the three models.

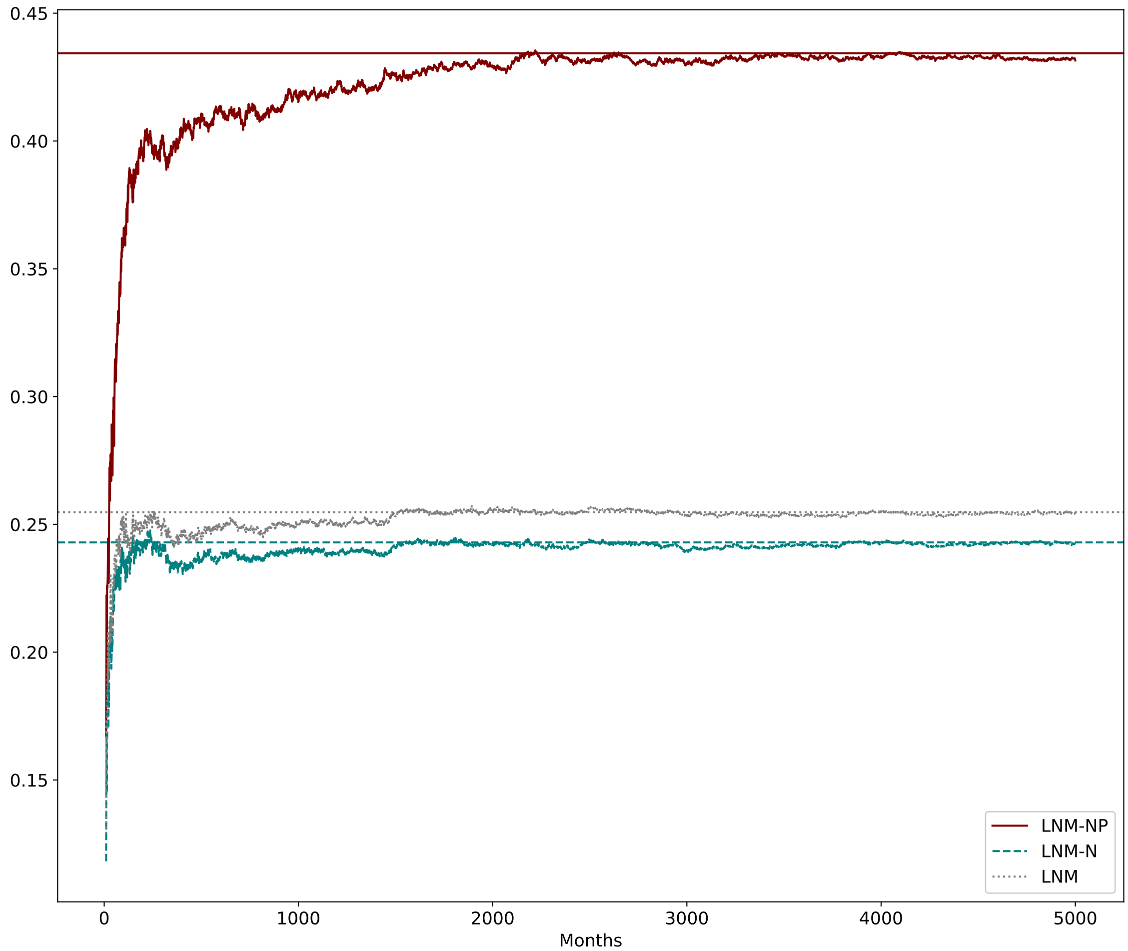


Figure 14: Estimates of Time-series Skewness in Stylized Simulations
This figure plots the median estimate of the time-series skewness across 1,000 stylized simulations as a function of the number of months (T). The horizontal lines denote the true value for the three models.