

# Cross-sectional Skewness

Sangmin Oh\*      Jessica A. Wachter<sup>†</sup>

June 18, 2019

## Abstract

This paper evaluates skewness in the cross-section of stock returns in light of predictions from a well-known class of models. Cross-sectional skewness in monthly returns far exceeds what the standard lognormal model of returns would predict. In spite of the fact that cross-sectional skewness is positive, aggregate market skewness is negative. We present a model that accounts for both of these facts. This model also exhibits long-horizon skewness through the mechanism of nonstationary firm shares.

---

\*Booth School of Business, The University of Chicago. Email: sangmino@chicagobooth.edu

<sup>†</sup>The Wharton School, University of Pennsylvania. Email: jwachter@wharton.upenn.edu. We thank Hendrik Bessembinder, John Campbell, Marco Grotteria, Nishad Kapadia, Yapai Zhang, and seminar participants at the Wharton School for helpful comments.

# 1 Introduction

Underlying the cross-section of stock returns is a universe of heterogeneous entities commonly referred to as firms. What is the most useful approach to modeling these firms? For the aggregate market, there is a wide consensus concerning the form a model needs to take to be a plausible account of the data. While there are important differences, quantitatively successful models tend to feature a stochastic discount factor with stationary growth rates and permanent shocks, combined with aggregate cash flows that, too, have stationary growth rates and permanent shocks.<sup>1</sup> No such consensus exists for the cross-section.

We start with a simple model for stock returns to illustrate the puzzle. The model is not meant to be the final word on the cross-section, but rather to show that the most straightforward way to extend the consensus for the aggregate to the cross-section runs quickly into difficulties both with regard to data and to theory. Because the literature has focused on the lognormal distribution, we assume asset returns have a common lognormal shock and an idiosyncratic shock. This model is consistent with an equilibrium where agents have constant relative risk aversion and dividends also feature a common and idiosyncratic shock.

We calibrate this model to the CRSP universe on stock returns. Despite the fact that the lognormal model implies return skewness, we find that the degree of skewness implied by the model is far less than monthly cross-sectional skewness in the data. This turns out to be the case even when we consider that the skewness in the data might be a time series rather than a cross-sectional phenomenon. That is, in the time series, it is possible that skewness results from “superstar months,” rather than “superstar stocks.” We show that even when adjusting for the former, there are far too many “superstar stocks” to be explained by the lognormal model.

Interestingly, the high degree of cross-sectional skewness co-exists with negative skewness in monthly aggregate market returns.<sup>2</sup> We extend the basic lognormal model to

---

<sup>1</sup>See, for example, Bansal and Yaron (2004), Campbell and Cochrane (1999), Wachter (2013).

<sup>2</sup>Albuquerque (2012) also notes the coexistence of positive firm skewness with negative aggregate

account for both negative aggregate market skewness and idiosyncratic skewness in the cross-section of returns. We show that large, rare idiosyncratic jumps appear to be required to explain cross-sectional skewness.

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.

Our work also relates to the research on power laws and firm values (Axtell, 2001; Gabaix, 2009). We show that both the lognormal model and its generalizations produce skewed distributions in long-run returns. This implies that a potential role for idiosyncratic volatility in aggregate outcomes, as in Gabaix (2011).

Since Fama (1965) established that stock returns did not approximate a normal distribution, the literature has examined the empirical linkages between this skewness and expected returns. The initial focus was on co-skewness (Harvey and Siddique, 2000; Dittmar, 2002), while more recent papers examine idiosyncratic skewness as well (Bali et al., 2011; Boyer et al., 2010; Kapadia, 2006). Others work measures ex ante skewness through options (Chang et al., 2013; Conrad et al., 2013). The focus of these papers is on the measurement of conditional skewness for a particular stock at a given point in time. This is a difficult measurement problem. In this paper, by contrast, we focus on the degree of unconditional skewness relative to various benchmark hypotheses on the return data generating process.<sup>3</sup> We find that it is very large.

Indeed, while most of the literature has focused on the cross-section of expected re-  
skewness. He presents a model that builds on the CARA-normal framework to capture this phenomenon.

---

<sup>3</sup>An underlying assumption in this literature, based on early work of Kraus and Litzenberger (1976), is that non-increasing absolute risk aversion implies that positive co-skewness is negatively priced and that positive idiosyncratic skewness has a price of zero. However, a lognormal distribution features both types of skewness and admits a CAPM-type result with constant relative risk aversion. Thus the choice of benchmark is important.

turns with the goal of establishing a correct pricing model, we focus on the measurement of cross-sectional skewness for its own sake. Cross-sectional skewness has received little attention, perhaps due to the view that it is not relevant in diversified portfolios because of the central limit theorem; that is, a portfolio of sufficiently many assets is close to normally distributed, and idiosyncratic risk does not matter. Our results indicate, however, that this reasonable intuition is model-dependent. There is no *ex ante* reason to dismiss skewness as irrelevant in a diversified portfolio. Correctly characterizing the distribution of returns, therefore, is important for portfolio decisions; it is also important for the reliability of statistics such as the mean and standard deviation, particularly the conditional mean and standard deviation. Finally, if one wants to simultaneously understand the cross-section of stock returns as well as the aggregate market, it is necessary to model both simultaneously. In this regard, one must think about how the cross-section aggregates, and here, the distributional assumptions on the cross-section are of first-order importance.

## 2 Some facts about skewness

We assume a time series of stock market return data  $t = 1, \dots, T$ . There are a total of  $N$  stock returns available, but at any point in time, only a subset are available. That is  $\mathcal{J}_t \subset \{1, \dots, N\}$  is the set of stocks available at time  $t$ .

Furthermore, we 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:

$$\bar{R}_j \equiv \frac{1}{T_j - t_{0j} + 1} \sum_{s=t_{0j}}^{T_j} R_{js} \quad (1)$$

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

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

where  $|\mathcal{J}_t|$  denotes the number of elements in  $\mathcal{J}_t$ . While (1–3) do involve abuse of notation, the intent should be clear. 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 return at time  $t$ .

- Time-series skewness for stock  $j$ :

$$\gamma_j^{TS} \equiv \frac{\frac{1}{T_j - t_{0j} + 1} \sum_{s=t_{0j}}^{T_j} (R_{js} - \bar{R}_j)^3}{\left[ \frac{1}{T_j - t_{0j}} \sum_{s=t_{0j}}^{T_j} (R_{js} - \bar{R}_j)^2 \right]^{3/2}} \quad (4)$$

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

- Cross-sectional skewness (measured using pooled returns)

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

- Cross-sectional skewness at time  $t$ :

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

Harvey and Siddique (2000) and Dittmar (2002) focus on the first of these, whereas Kapadia (2006) focuses on the third. Following Bessembinder (2018), we 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 is a fixed quantity belonging to firm  $j$ , then (4) will be a consistent estimator of it. However, it may not be an efficient 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). One might object, however, that (5) does not only capture skewness in returns, 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.

In what follows, we do not take a stand on which of these is the best, or even if they are consistent estimators of some fixed quantity as  $N$  or  $T$  becomes large. We simply calculate the values in the data, and compute the sampling distribution under various null hypotheses.

As a first look at the data, Table 1 shows cross-sectional skewness  $\gamma^{CS}$  across various subsets of the CRSP universe.<sup>4</sup> Monthly skewness equals about 6. We also confirm the result of Bessembinder (2018) that most stock-month combinations deliver lower returns than the average 1-month Treasury bill.

Table 2 takes a deeper look at these results by examining statistics on  $\gamma^{TS}$  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  $\gamma^{CS}$ , 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

In what follows, 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.<sup>5</sup>

---

<sup>4</sup>See section 5 for a description of these subsets when not obvious.

<sup>5</sup>Alternative tractable ways of capturing skewness include centered Gamma shocks (Bekaert et al., 2019) and skew-normal shocks (Colacito et al., 2016).

### 3.1 The Compound Poisson process

Let  $Q_t$  be a compound Poisson process parameterized by  $\lambda_t$  and jump size  $\zeta$ . Namely,  $\lambda_t$  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:

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

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  $u \in \mathbb{R}$ ,

$$\mathbb{E}_t [e^{uQ_{t+1}}] = e^{\lambda_t (\mathbb{E}[e^{u\zeta}] - 1)} \quad (7)$$

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

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

so that

$$\mathbb{E} [Q_{t+1} | \mathfrak{F}_t] = \lambda_t \mathbb{E} [\zeta]$$

We compute the conditional variance of  $Q_{t+1}$  using the law of total variance:

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

### 3.2 Dividend growth and the stochastic discount factor

Consider a cross-section of  $N$  assets. Define  $D_{jt}$  as the dividend on asset  $j$  at time  $t$ .

Assume log dividend growth on asset  $j$ ,  $j = 1, \dots, N$  is as follows:

$$\Delta d_{j,t+1} = \left( \mu_j - \frac{1}{2} \sigma_{i,j}^2 - \frac{1}{2} \beta_j^2 \sigma_c^2 \right) + \beta_j \sigma_c \epsilon_{t+1}^c + \sigma_{ij} \epsilon_{j,t+1}^i - \beta_j^Q Q_{t+1} + Q_{j,t+1}^i \quad (8)$$

where  $\epsilon_{ct}$  and  $\{\epsilon_{jt}^i | j = 1, \dots, N\}$  are iid standard normal variables,

$$Q_{t+1} \sim \text{Compound Poisson}(\lambda_t, \zeta)$$

and

$$Q_{j,t+1}^i \sim \text{Compound Poisson}(\lambda^i, \zeta_j^i)$$

such that, conditional on time- $t$  information  $Q_{t+1}$  and  $\{Q_{j,t+1}^i | j = 1, \dots, N\}$  are independent of the normal shocks and of each other.<sup>6</sup>

First note that, absence compound Poisson realizations, the mean of dividend growth (as opposed to log dividend growth) equals  $\mu_j$ . Normal systematic shocks  $\epsilon_{c,t+1}$  affect firm- $j$  dividend growth, as do iid idiosyncratic shocks  $\epsilon_{j,t+1}^i$ . The loading on the systematic shock differs across firms and is given by  $\beta_j$ . We calibrate  $Q_{t+1}$  and  $Q_{j,t+1}^i$  so that they are positive, and assume  $\beta_j^Q < 0$ . The compound Poisson random variable  $Q_{t+1}$  is systematic, whereas  $Q_{j,t+1}^i$  is idiosyncratic.

The above specification allows  $\zeta$  and  $\zeta_j^i$  to be random variables. We will assume that these are drawn from time-invariant distributions. Note that the intensity of  $Q_{t+1}$  is time-varying. We assume:

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

where

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

In recent work, Salgado et al. (2019) document considerable skewness in firm-level employment and sales. This skewness could be either positive or negative depending on the state of the business cycle. These skewed shocks could be the origin of the skewed shocks in asset-level cash flows in (8).

---

<sup>6</sup>This is a “multiple-trees” model similar to that considered by Cochrane et al. (2008) and Martin (2013).



Define the stochastic discount factor (SDF) as follows

$$M_{t+1} = \exp \left\{ -r_f - \frac{1}{2} x_t^2 - \lambda_t (\mathbb{E}_t [e^\zeta] - 1) \left( x_t \epsilon_{t+1}^c + Q_{t+1} \right) \right\} \left( \right) \quad (9)$$

where  $r_f$  is the one-period riskfree rate, which we assume to be constant, and where

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

for a standard normal  $\epsilon_{t+1}^x$  is independent of all other shocks. Applying (7), along with properties of the lognormal distribution, we have  $E[M_{t+1}] = e^{-r_f}$  as required.

### 3.3 Solution for prices

Let  $P_{n,j,t}$  be the time- $t$  price of the dividend on asset  $j$  that will be paid  $n$  periods from now. Absence of arbitrage implies

$$\mathbb{E}_t \left[ M_{t+1} \left( \frac{P_{n-1,j,t+1}}{P_{n,j,t}} \right) \right] = 1 \quad (10)$$

Equation 10 implies the following recursion for ex-dividend prices.

$$\frac{P_{n,j,t}}{D_{j,t}} = \mathbb{E}_t \left[ M_{t+1} \left( \frac{D_{j,t+1}}{D_{j,t}} \right) \left( \frac{P_{n-1,j,t+1}}{D_{j,t+1}} \right) \right] \left( \right) \quad (11)$$

Note also that

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

where  $M_{t,t+n} = \prod_{i=1}^n M_{t+i}$ . The Markov assumptions on the state variables imply that we can define functions

$$F_n^j(x_t, \lambda_t) = \frac{P_{n,j,t}}{D_{j,t}}$$

It follows from (11) and  $P_{0,j,t} = D_{j,t}$  that

$$F_n^j(x_t, \lambda_t) = \exp \left\{ a^j(n) + b_{xj}(n) x_t + b_{\lambda j}(n) \lambda_t \right\} \quad (12)$$

where

$$\begin{aligned} b_{x_j}(n) &= -\beta_j \sigma_c \frac{1 - \varphi^n}{1 - \varphi} \\ b_{\lambda_j}(n) &= \mathbb{E} \left[ e^\zeta - e^{-(\beta_j^Q - 1)\zeta} \right] \frac{1 - \varphi_\lambda^n}{1 - \varphi_\lambda} \end{aligned}$$

Note that  $b_x^j(n), b_\lambda^j(n) < 0$  and decreasing.

Define the return

$$R_{j,n,t+1} = \frac{P_{j,n-1,t+1}}{P_{jnt}} = \frac{F_{n-1}^j(x_{t+1}, \lambda_{t+1}) D_{j,t+1}}{F_n^j(x_t, \lambda_t) D_{jt}}$$

It follows from (12) that

$$\begin{aligned} \log R_{j,n,t+1} &= \left[ r_f + \beta_j \sigma_c x_t + \left( \mathbb{E}_t \left[ e^{-(\beta_j^Q - 1)\zeta} \right] \left( -\mathbb{E}_t [e^{-\zeta}] \right) \lambda_t \right) \right. \\ &\quad + b_{x_j}(n-1) \sigma_x \epsilon_{t+1}^x + b_{\lambda_j}(n-1) Q_{t+1}^\lambda + \beta_j \sigma_c \epsilon_{t+1}^c + \sigma_{i,j} \epsilon_{j,t+1}^i - \beta_j^Q Q_{t+1} + Q_{j,t+1}^i \\ &\quad - \frac{1}{2} \sigma_{i,j}^2 - \frac{1}{2} \beta_j^2 \sigma_c^2 - \frac{1}{2} (b_{x_j}(n-1) \sigma_x)^2 \\ &\quad \left. - \nu \left( \mathbb{E}_t \left[ e^{b_{\lambda_j}(n-1)\zeta^\lambda} \right] - 1 \right) \left( -\lambda_j^i \left( \mathbb{E}_t [e^{\zeta^i}] - 1 \right) \right) \right] \quad (13) \end{aligned}$$

Then the risk premium on the dividend strip for asset  $j$  is as follows:

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

Aggregate parameters  $\sigma_x, \nu$  do not appear, and the idiosyncratic parameters  $\sigma_{ij}$  and  $\lambda^i$  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  $\lambda_t$  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.

### 3.4 Pricing long-lived assets

In order to calibrate stock returns, we need to move from dividend strips to long-lived assets. The price-dividend ratio on asset  $j$  is equal to the sum of the price-dividend ratios on the strips for asset  $j$ :

$$\frac{P_{jt}}{D_t} = \sum_{n=1}^{\infty} F_n^j(x_t, \lambda_t) \quad (14)$$

Define the return

$$R_{j,t+1} = \frac{\frac{P_{j,t+1}}{D_{j,t+1}} + 1}{\frac{P_{jt}}{D_{jt}}} \frac{D_{j,t+1}}{D_{jt}}.$$

Because asset  $j$  is a portfolio of strips, the return is a weighted average of the strip returns:

$$R_{j,t+1} = \sum_{n=1}^{\infty} \left( w(x_t, \lambda_t, n) R_{j,n,t+1}, \quad (15) \right.$$

with weights:

$$w(x_t, \lambda_t, n) = \frac{F_n^j(x_t, \lambda_t)}{\sum_{n=1}^{\infty} F_n^j(x_t, \lambda_t)}. \quad (16)$$

Strictly speaking, (15) and (16) fully define the return. However, these are difficult to implement. In order to calibrate the model, we approximate the return.

Note that all terms in the summation in (13) have dividend growth in common. Write the log return as:

$$\log R_{j,t+1} = \log \sum_{n=1}^{\infty} \left( w(x_t, \lambda_t, n) \frac{F_{n-1}^j(x_{t+1}, \lambda_{t+1})}{F_n^j(x_t, \lambda_t)} \right) \left( + \Delta d_{j,t+1} \right) \quad (17)$$

For the first term in (17), we use the approximation:

$$\log \sum_{n=1}^{\infty} \left( w(x_t, \lambda_t, n) \frac{F_{n-1}^j(x_{t+1}, \lambda_{t+1})}{F_n^j(x_t, \lambda_t)} \right) \left( \approx \sum_{n=1}^{\infty} \left( w(x_t, \lambda_t, n) \log \frac{F_{n-1}^j(x_{t+1}, \lambda_{t+1})}{F_n^j(x_t, \lambda_t)} \right) \right)$$

which is accurate for small shocks. Variation in the weights is second-order for the effects of interest. Define

$$b_x^* \approx \sum_{n=1}^{\infty} \left( w(x_t, \lambda_t, n) b_{x_j}(n-1) \right)$$

$$b_\lambda^* \approx \sum_{n=1}^{\infty} \left( w(x_t, \lambda_t, n) b_{\lambda_j}(n-1) \right)$$

as weighted averages of the coefficients.

Using (13), we find that

$$\log R_{j,t+1} \approx \text{constant as of time } t + b_x^* \sigma_x \epsilon_{t+1}^x + b_\lambda^* Q_{t+1}^\lambda + \beta_j \sigma_c \epsilon_{t+1}^c + \sigma_{ij} \epsilon_{j,t+1}^i - \beta_j^Q Q_{t+1} + Q_{j,t+1}^i \quad (18)$$

To ensure that the return is consistent with no-arbitrage, we then write:

$$\log R_{j,t+1} = r_f + \beta_j \sigma_c x_t + \mathbb{E}_t \left[ e^\zeta - e^{-(\beta_j^Q - 1)\zeta} \right] \left( \begin{aligned} & - \frac{1}{2} \sigma_{i,j}^2 - \frac{1}{2} \beta_j^2 \sigma_c^2 - \frac{1}{2} (b_{x_j}^* \sigma_x)^2 \\ & - \nu \mathbb{E}_t \left[ e^{b_{\lambda_j}^* \zeta^\lambda} - 1 \right] - \lambda_j^i \mathbb{E}_t \left[ e^{\zeta_j^i} - 1 \right] \left( \right. \end{aligned} \right)$$

$$\left. + b_{x_j}^* \sigma_x \epsilon_{t+1}^x + b_{\lambda_j}^* Q_{t+1}^\lambda + \beta_j \sigma_c \epsilon_{t+1}^c + \sigma_{i,j} \epsilon_{j,t+1}^i - \beta_j^Q Q_{t+1} + Q_{j,t+1}^i \quad (19) \right)$$

The time- $t$  constants in (19) imply

$$\mathbb{E}_t [M_{t+1} R_{j,t+1}] = 1,$$

given the expression in (18).

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:

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

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:

$$\begin{aligned} \log R_{t+1} = r_f + \sigma_c x_t + \mathbb{E}_t [e^\zeta - 1] \lambda_t & \\ - \frac{1}{2} \sigma_c^2 - \frac{1}{2} (b_x^* \sigma_x)^2 - \nu \mathbb{E}_t [e^{b_\lambda^* \zeta^\lambda} - 1] & \left( \right. \\ & \left. + b_x^* \sigma_x \epsilon_{t+1}^x + b_\lambda^* Q_{t+1}^\lambda + \sigma_c \epsilon_{t+1}^c - Q_{t+1} \right) \quad (20) \end{aligned}$$

### 3.5 Summary of the models

- *Consumption CAPM*. This model has dividend process (8) with the compound Poisson processes  $Q_t$  and  $Q_{jt}^i$  identically zero.<sup>7</sup> The model has a limiting version of SDF (9) 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 (8) with the compound Poisson processes  $Q_t$  and  $Q_{jt}^i$  equal to zero. Moreover, in the SDF (9)  $Q_t^\lambda = 0$ . We call this the *Lognormal model*.
- LW model with rare dividend disasters. This model takes (8) and sets the idiosyncratic compound Poisson process  $Q_{jt}^i$  to zero. Moreover, in the SDF (9),  $Q_t^\lambda = 0$ .
- LW model with normal-times negative systematic skewness. This model takes (8)

---

<sup>7</sup>That is, we set  $\lambda_t = 0$  and  $\lambda^i = 0$ .

and sets  $Q_{jt}^i$  to zero. The SDF is given by (9). We call this the *Lognormal-N model*.

- LW model with negative market skewness and positive idiosyncratic skewness. This model uses (8) and (9) 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.<sup>8</sup>

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

---

<sup>8</sup>One micro-foundation for the lognormal-NP model is learning and recursive utility. See Wachter and Zhu (2019).

## 5 Simulation

We generate fictitious samples from the models described in Section 3. We consider 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. We then estimate firm-level parameters as described in the next section. Clearly, there are many more firms in the estimation than there are stocks in the simulation (14,786 versus 5447). We thus use the following bootstrap procedure. At the start of each fictitious sample, we draw 5447 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  $\tau_j = T_j - t_{0j} + 1$  is the number of months that firm  $j$  is listed, we draw from the estimated parameters of firm  $j$  with probability  $\frac{\tau_j}{\sum_{j=1}^{14,786} \tau_j}$ . Thus, across fictitious samples, we should roughly capture the true distribution of firms in the cross section. The first of these strategies 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 surprisingly 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.

Given a set of firm-level parameters, we simulate fictitious samples assuming returns are distributed as in Section 3. For each sample, we draw the aggregate shocks  $\epsilon_c$ ,  $\epsilon_x$ ,  $Q$ , and  $Q_\lambda$ . We then draw the appropriate number of sequences of firm-specific shocks, and use (19) 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.

## 5.1 Calibration

We calibrate the aggregate parameters of each model to match the mean, standard deviation, and autocorrelation of the price-dividend ratio, similarly to Lettau and Wachter (2007), but with adjustments given that our calibration is monthly.<sup>9</sup> The unconditional mean of the price of risk,  $\bar{x}$ , is chosen such that when  $x_t$  is at its long-run mean, the maximal Sharpe ratio is 0.20. Thus, setting  $\sqrt{e^{\bar{x}} - 1} = 0.20$ , implying  $\bar{x} = 0.198$ . The parameter  $\sigma_c$  is determined by annual dividend volatility. Given an annual persistence of 0.87,  $\varphi = \varphi_\lambda = 0.988 = 0.87^{1/12}$ .

For simplicity, we calibrate the lognormal-N and lognormal-NP models so that there is about one Poisson occurrence per year, and that the effect size is 4.5%. In order to identify volatility parameters, for the purposes of the calibration, we set the loading  $b_x^* = 1$  in (20) and  $b_{xj}^* = \beta_j$ .<sup>10</sup> These assumptions imply a one-factor model. We thus estimate  $\beta_j$  and  $\beta_j^Q$  from a CAPM regression of log market returns on log asset returns. Given these aggregate loadings, we set idiosyncratic volatility parameters to match total volatility for the lognormal and lognormal-N model. That is,  $\sigma[R]^2 = \sigma_{i,j}^2 + \beta_j^2 \sigma[R_m]^2$ , where  $\sigma[R]^2$  is the variance of  $R$ . For the lognormal-NP model, we assume idiosyncratic jumps happen with one tenth of the probability of the aggregate jumps, and draw parameters from a power law distribution that is sufficient to match the skewness in the data. Given these values, we scale down  $\sigma_{i,j}$  accordingly, so that we match individual stock volatility. Table 3 gives our calibrated values.

Table 4 reports results for the aggregate market in the data, and compares these 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.

---

<sup>9</sup>We thus set  $\sigma_c$  and  $\sigma_x$  to be lower in the lognormal-N and lognormal-NP models as compared with the lognormal model.

<sup>10</sup>In effect, we assume that assets' exposure to  $x_t$  is the same as to  $\epsilon^c$ . For assets of fixed duration, this assumption is correct. However, assets with longer durations will have more exposure to  $x_t$ . Given that our primary purpose is skewness in individual stock returns rather than explaining average returns, accounting for this effect seems unlikely to make a difference.



## 5.2 Cross-sectional skewness in pooled returns

Tables 5 and 6 shows our main results. We discuss only Table 6, as the results for Table 5 are similar.

The data column of Table 6 shows the mand 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 most stock returns on most days perform worse than Treasury bills (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

$$\Pr(R_j > R_f) = \Pr(\log R_M > r_f), \quad (21)$$

where the inequality holds because the log in a monotonic transformation. Consider (19), specialized to the case of no Poisson shocks:

$$\begin{aligned} \log R_{j,t+1} - r_f &= \beta_j \sigma_c x_t - \frac{1}{2} \sigma_{i,j}^2 - \frac{1}{2} \beta_j^2 \sigma_c^2 - \frac{1}{2} (b_{xj}^* \sigma_x)^2 \\ &\quad + b_{xj}^* \sigma_x \epsilon_{t+1}^x + \beta_j \sigma_c \epsilon_{t+1}^c + \sigma_{i,j} \epsilon_{j,t+1}^i. \end{aligned}$$

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  $-\frac{1}{2}\sigma_{i,j}^2 - \frac{1}{2}\beta_j^2\sigma_c^2 - \frac{1}{2}(b_{xj}^*\sigma_x)^2$ . It is the log return that matters when comparing the median return to the riskfree rate, given (21). 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. While neither can match it exactly, the numbers are in the same order of magnitude. It is in the pooled cross-sectional skewness that the models are so far off. The maximum value generated across simulations of these models is about 1, whereas pooled cross-sectional skewness is close to six in the data. Only the lognormal-NP model, with the positive idiosyncratic Poisson events, can match the cross-sectional skewness in the data.

### 5.3 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.

To confront this concern, we compute skewness in the cross-section at each point in time. That is, we consider the measure (6). Figure 3 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 2, which shows analogous

results for the 404 firms, tells a similar story. Finally, Table 7 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.<sup>11</sup>

## 5.4 Skewness in long-horizon returns

Motivated by findings of Bessembinder (2018), studies of the skewed size distribution in market capitalizations (Axtell, 2001; Gabaix, 2009), and granularity in returns (Gabaix, 2011), we examine the implications of the models for long-horizon returns, namely the cumulative returns for asset  $j$  between the beginning and end of the sample:

$$CR_{j,t,t+T} = \prod_{i=1}^T R_{j,t+i}.$$

We use this measure to construct the percent value created by each asset over total value:

$$\frac{CR_{j,t,t+T}}{\sum_{k=1}^N (CR_{k,t,t+T})}. \quad (22)$$

We compare the sum of these statistics for the top 10 firms in the model and in the data.

All three models generate skewness in cumulative returns that exceed the values in the data. This comes about because all three models imply non-stationary firm shares. Because log cumulative returns are (close to) a random walk, firm shares in (22) tend to either zero or 1 over time.

In the data, it is difficult to tell if firm shares are stationary or not.<sup>12</sup> Moreover, the data equivalent of (22) uses market capitalization, not returns, leaving open the question of whether payouts, either in the form of dividends or special distributions, can reconcile

---

<sup>11</sup>In this analysis, we simulate from a lognormal model with constant idiosyncratic volatilities. One might argue then that we are not perhaps 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.

<sup>12</sup>The question of firm size stationarity is also connected to the question of stationarity in top wealth shares. See Gomez (2019).

the model with the data. Ultimately, though, the models as stated imply that, over time, one firm will take over the economy with probability 1. To create stationary firm shares, one would most likely require a model where firms exit and other firms enter at, say, an average size. Such a model could potentially have asset pricing implications, which we leave for future work.

## 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 8 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. We leave these topics to future research.

## References

- 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.
- Bekaert, G., Engstrom, E., and Xu, N. (2019). The Time Variation in Risk Appetite and Uncertainty. NBER Working Paper No. 25673.
- 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. 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., 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.
- 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.
- 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.
- Fama, E. F. (1965). Portfolio analysis in a stable paretian market. *Management science*, 11(3):404–419.
- 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.
- Gomez, M. (2019). Displacement and the Rise in Top Wealth Inequality . Working paper, Columbia University.

- 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.
- Kapadia, N. (2006). The next microsoft? skewness, idiosyncratic volatility, and expected returns. Working paper, Tulane University.
- Kogan, L., Papanikolaou, D., and Stoffman, N. (2019). Left behind: Creative destruction, inequality, and the stock market. *Journal of Political Economy*.
- 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.
- Martin, I. (2013). The lucas orchard. *Econometrica*, 81(1):55–111.
- Salgado, S., Guvenen, F., and Bloom, N. (2019). Skewed Business Cycles. Working paper, University of Minnesota, University of Pennsylvania, and Stanford University.
- Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance*, 68(3):987–1035.
- Wachter, J. A. and Zhu, Y. (2019). Learning with rare disasters. Working paper, University of Pennsylvania.

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

Skewness Type	Statistic	404 Firms (1973.01 - 2016.12)		14,786 Firms (1973.01 - 2016.12)	
		skew[ $R$ ]	skew[log $R$ ]	skew[ $R$ ]	skew[log $R$ ]
Time-series	Min.	-1.01	-4.02	-3.66	-17.2
	5th	-0.188	-1.13	-0.179	-1.46
	50th	0.362	-0.204	0.891	0.021
	95th	2.12	0.376	3.63	1.18
	Max.	7.58	1.64	16.4	6.33
	S&P500	-0.44	-0.71	-0.44	-0.71
Cross-section	Min.	-2.56	-10.1	-0.288	-4.15
	5th	-0.69	-1.61	0.393	-1.71
	50th	0.69	0.11	2.40	-0.085
	95th	3.35	1.67	8.49	1.52
	Max.	12.9	5.79	46.0	2.66
Pooled		1.32	-0.42	5.99	-0.24

*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 3:** Calibrated Parameters of the Model

Variable	Value
$\bar{x}$	0.198
$\bar{\lambda}$	0.083
$\beta_j, \beta_j^Q$	Estimated from CAPM Regression
$\varphi, \varphi^\lambda$	0.988
$\sigma_c$	0.0419
$\sigma_x$	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
Additional Parameters for Lognormal-N	
$\zeta$	0.045
$\zeta^\lambda$	0.045
$\nu$	Set to $\bar{\lambda}$
Additional Parameters for Lognormal-NP	
$\zeta_j^i$ (Simulation with 404 Firms)	Drawn from a power-law distribution with mean $7\zeta$
$\zeta_j^i$ (Simulation with 14,786 Firms)	Drawn from a shifted power-law distribution with mean $15\zeta$ and minimum $6.5\zeta$
$\lambda_j^i$	Set to $0.1 \bar{\lambda}$

*Source:* CRSP, Ken French's Website

g

*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.  $\bar{\lambda}$  is set to 1 per year or 0.083 monthly.  $\beta_j$  and  $\beta_j^Q$  are estimated from a CAPM regression of log returns, and the persistence variables  $(\phi, \phi^\lambda)$  are set to match the autocorrelation of the price-dividend ratio.  $\sigma_c$  and  $\sigma_x$  are chosen to match the data explicitly following Lettau and Wachter (2007), and they are scaled by 0.7 in simulations for Lognormal-N and Lognormal-NP. Aggregate jump intensities  $(\zeta, \zeta^\lambda)$  are set to 0.045, while the idiosyncratic jump intensities are drawn independently across stocks from a power-law distribution.

**Table 4:** 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	-1.79	-0.168	1.27	2.79	3.92	1.18
$\sigma[R_m]$	4.55	4.04	4.23	4.52	4.89	5.47	4.57
$\gamma^{TS}[R_m]$	-0.518	-0.217	-0.056	0.130	0.293	0.488	0.120
Panel B. Lognormal-N Model							
$E[R_m] - 1$	0.922	0.224	0.548	1.03	1.46	1.75	1.01
$\sigma[R_m]$	4.55	3.52	3.68	3.94	4.23	4.39	3.98
$\gamma^{TS}[R_m]$	-0.518	-0.716	-0.549	-0.351	-0.147	0.013	-0.370
Panel C. Lognormal-NP Model							
$E[R_m] - 1$	0.92	0.257	0.632	1.07	1.50	1.96	1.09
$\sigma[R_m]$	4.55	3.46	3.68	3.94	4.22	4.36	3.92
$\gamma^{TS}[R_m]$	-0.518	-0.815	-0.579	-0.370	-0.156	-0.028	-0.366

*Source:* CRSP and simulations

*Notes:* We conduct 400 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 100 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 5:** Inference on Pooled Monthly Returns (Simulation with 404 Firms)

	Empirical	Simulated Values					Simulation Population
	Value	Min	5th	50th	95th	Max	
Panel A. Lognormal Model							
$E[R] - 1$	1.33	-1.52	-0.21	1.30	2.72	3.83	1.28
$\sigma[R]$	10.3	9.79	9.95	10.2	10.4	10.8	10.2
$\gamma^{CS}[R]$	1.32	0.322	0.402	0.48	0.56	0.634	0.487
$\gamma^{CS}[\log R]$	-0.42	-0.172	-0.102	-0.021	0.052	0.114	-0.020
$\% \log R > \log R_f$	54.5	40.1	45.6	51.8	57.4	61.7	49.4
$\tilde{\gamma}^{TS}[R_j]$	0.36	0.234	0.252	0.275	0.301	0.321	.
Panel B. Lognormal-N Model							
$E[R] - 1$	1.33	0.35	0.633	1.06	1.46	1.73	1.05
$\sigma[R]$	10.3	9.69	9.76	9.89	10.0	10.1	9.90
$\gamma^{CS}[R]$	1.32	0.35	0.382	0.419	0.455	0.503	0.420
$\gamma^{CS}[\log R]$	-0.42	-0.14	-0.119	-0.079	-0.045	-0.018	-0.080
$\% \log R > \log R_f$	54.51	48.1	49.3	51.2	53.0	54.2	49.1
$\tilde{\gamma}^{TS}[R_j]$	0.362	0.189	0.202	0.225	0.245	0.260	.
Panel C. Lognormal-NP Model							
$E[R] - 1$	1.33	0.265	0.662	1.04	1.47	1.81	1.06
$\sigma[R]$	10.3	11.5	11.7	12.0	12.2	12.6	11.9
$\gamma^{CS}[R]$	1.32	2.93	3.26	3.80	4.67	10.5	3.91
$\gamma^{CS}[\log R]$	-0.42	0.752	0.944	1.08	1.22	1.33	1.08
$\% \log R > \log R_f$	54.5	46.2	47.8	49.5	51.3	52.7	47.4
$\tilde{\gamma}^{TS}[R_j]$	0.362	0.503	0.568	0.706	0.893	1.01	.

*Source:* CRSP and simulations

*Notes:* We conduct 400 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 100 simulations.  $E[R] - 1$  and  $\sigma[R]$  are reported in percentages.  $\gamma^{CS}$  denotes cross-sectional skewness across pooled returns, while  $\gamma^{TS}[R_j]$  denotes the time-series skewness of returns for firm  $j$ . We report the  $\tilde{\gamma}^{TS}[R_j]$ , the median value of  $\gamma^{TS}[R_j]$  across the 404 firms.

**Table 6:** Inference on Pooled Monthly Returns (Simulation with 14,786 Firms)

	Empirical Value	Simulated Values					Simulation Population
		Min	5th	50th	95th	Max	
Panel A. Lognormal Model							
$E[R] - 1$	1.33	-1.97	-0.214	1.37	3.04	4.30	1.263
$\sigma[R]$	17.7	16.3	16.5	16.9	17.3	17.8	16.8
$\gamma^{CS}[R]$	5.99	0.806	0.879	0.953	1.033	1.10	0.949
$\gamma^{CS}[\log R]$	-0.235	-0.255	-0.202	-0.144	-0.079	-0.035	-0.148
% $\log R > \log R_f$	48.0	40.5	45.3	49.7	54.3	57.7	51.7
$\tilde{\gamma}^{TS}[R_j]$	0.875	0.402	0.411	0.424	0.438	0.467	.
Panel B. Lognormal-N Model							
$E[R] - 1$	1.33	0.225	0.570	1.10	1.59	1.88	1.07
$\sigma[R]$	17.7	16.2	16.3	16.5	16.7	16.9	16.5
$\gamma^{CS}[R]$	5.99	0.849	0.879	0.912	0.950	0.979	0.910
$\gamma^{CS}[\log R]$	-0.235	-0.218	-0.200	-0.175	-0.153	-0.133	-0.178
% $\log R > \log R_f$	48.0	46.6	47.7	49.2	50.5	51.29	51.2
$\tilde{\gamma}^{TS}[R_j]$	0.875	0.374	0.386	0.396	0.406	0.418	.
Panel C. Lognormal-NP Model							
$E[R] - 1$	1.33	0.232	0.653	1.14	1.62	2.13	1.16
$\sigma[R]$	17.7	17.5	17.7	17.9	18.1	18.3	17.9
$\gamma^{CS}[R]$	5.99	5.13	5.34	5.77	6.48	15.2	5.91
$\gamma^{CS}[\log R]$	-0.235	1.041	1.10	1.15	1.21	1.25	1.16
% $\log R > \log R_f$	48.0	44.3	45.7	47.4	48.9	50.6	49.5
$\tilde{\gamma}^{TS}[R_j]$	0.875	2.04	2.11	2.20	2.32	2.43	.

*Source:* CRSP and simulations

*Notes:* We conduct 400 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 100 simulations.  $E[R] - 1$  and  $\sigma[R]$  are reported in percentages.  $\gamma^{CS}$  denotes cross-sectional skewness across pooled returns, while  $\gamma^{TS}[R_j]$  denotes the time-series skewness of returns for firm  $j$ . We report the  $\tilde{\gamma}^{TS}[R_j]$ , the median value of  $\gamma^{TS}[R_j]$  across the 14,786 firms.

**Table 7:** Inference on Monthly Cross-sectional Skew

	Empirical $\tilde{\gamma}_{cs}$	$\tilde{\gamma}_{cs}$ from Simulated Values					% of Months with Empirical $\gamma_{cs}$ $\geq$ Max
		Min	5th	50th	95th	Max	
Panel A. Simulation with 404 Firms (1973.01 - 2016.12)							
Lognormal		0.234	0.252	0.275	0.301	0.321	64.8
Lognormal-N	0.694	0.189	0.202	0.225	0.245	0.26	66.7
Lognormal-NP		0.503	0.568	0.706	0.893	1.01	37.5
Panel B. Simulation with 14,786 Firms (1973.01 - 2016.12)							
Lognormal		0.402	0.411	0.424	0.438	0.467	93.9
Lognormal-N	2.400	0.374	0.386	0.396	0.406	0.418	94.5
Lognormal-NP		2.04	2.11	2.20	2.32	2.43	49.8

*Source:* CRSP and simulations

*Notes:* We conduct 400 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  $\gamma_{cs}$  is greater than the maximum  $\tilde{\gamma}_{cs}$  obtained from the simulations.

**Table 8:** Simulated Results from Each Model

Stylized Facts	Lognormal	Lognormal-N	Lognormal-NP
<i>A-1. Monthly Returns (Stocks)</i>			
Risk Premium	○	○	○
Volatility	○	○	○
Positive TS Skewness	○	○	○
Positive Pooled Skewness	×	×	○
% > Risk-free Rate	○	○	○
<i>A-2. Monthly Returns (Market)</i>			
Risk Premium	○	○	○
Volatility	○	○	○
Negative TS Skewness	×	○	○
<i>B. Monthly Cross-section</i>			
Positive CS Skewness	×	×	○
<i>C. Long-run Returns</i>			
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-1 is relevant to moments of pooled monthly returns across stocks, and Panel A-2 pertains to analogous moments for the market returns. Panel B refers to the median monthly cross-sectional skewness, and Panel C pertains to distributions of long-run returns and firm sizes across all stocks.

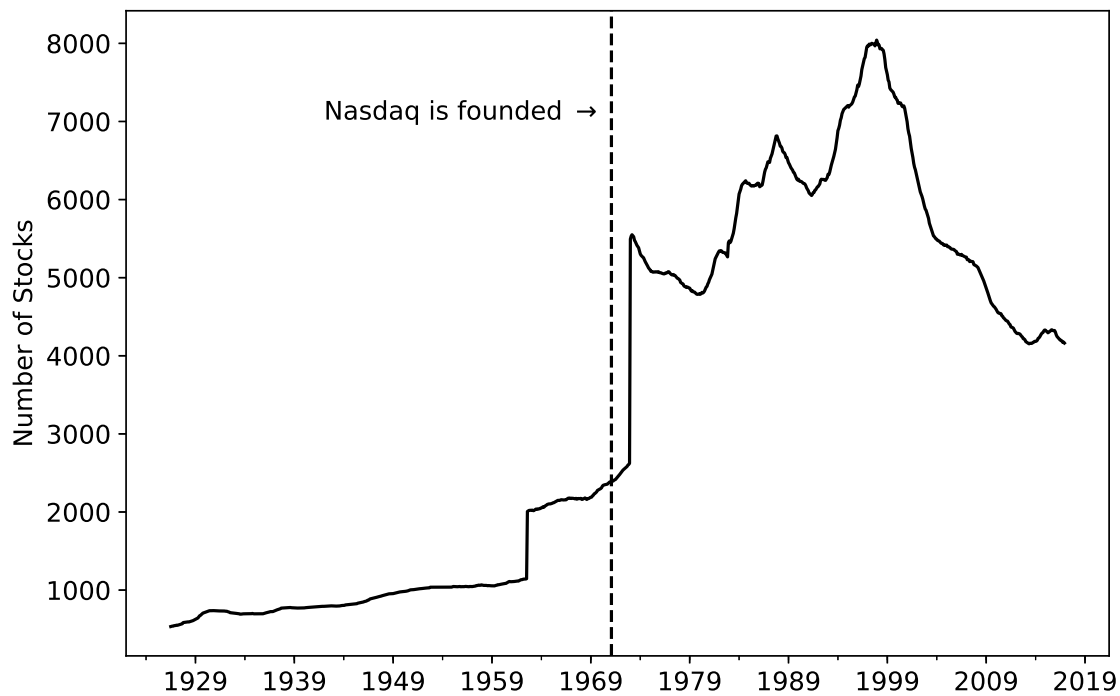
**Table 9:** Percent of total increase in value accounted for by the top 10 firms

	Empirical Value	Simulated Values				
		Min	5th	50th	95th	Max
Panel A. Simulation with 404 Firms (1973.01 - 2016.12)						
Lognormal		28.4	35.0	61.5	99.5	100.0
Lognormal-N	31.06	25.4	34.4	51.2	83.5	97.0
Lognormal-NP		30.2	40.4	60.9	88.7	98.4
Panel B. Simulation with 14,786 Firms (1973.01 - 2016.12)						
Lognormal		16.4	24.5	65.0	100.0	100.0
Lognormal-N	16.24	17.7	27.3	50.8	90.4	100.0
Lognormal-NP		18.5	30.3	56.7	93.1	100.0

*Source:* CRSP and simulations

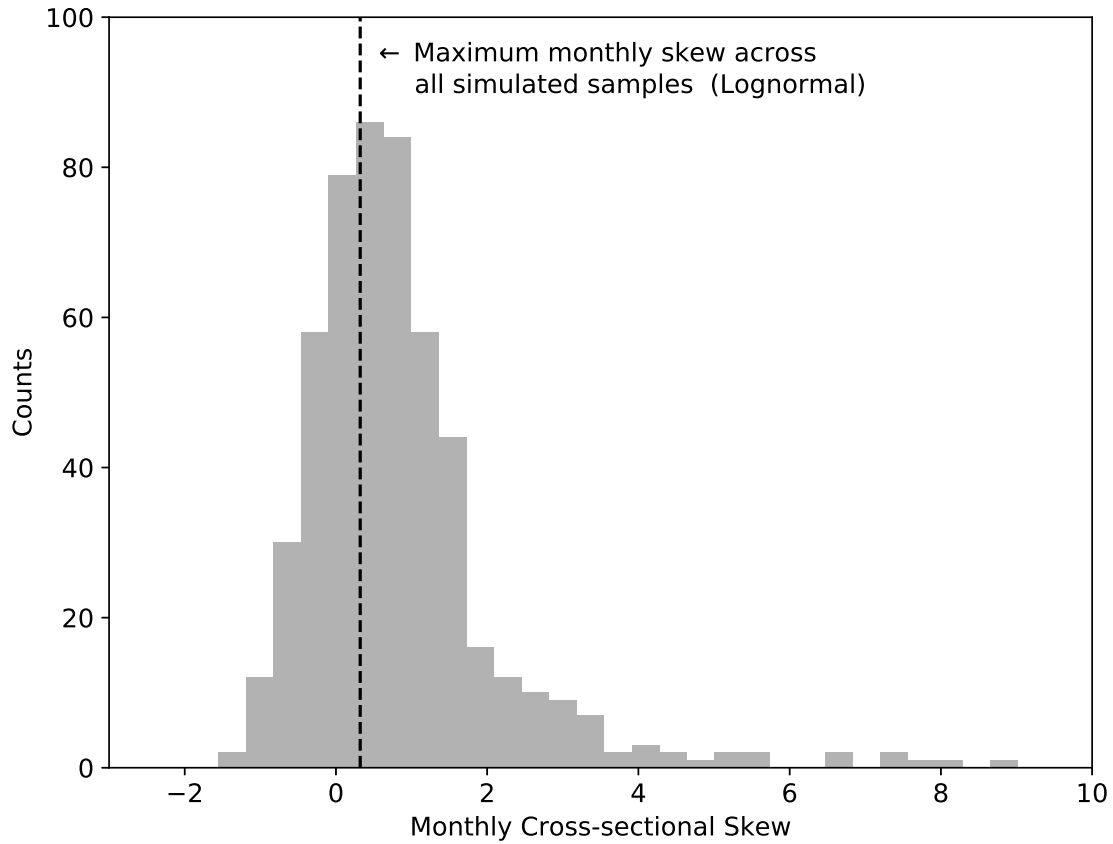
*Notes:* We conduct 400 monthly simulations of the stock market for the three models. For each simulation, we examine the percentage of the total increase in value contributed by the top ten firms, where we measure the increase in cumulative returns. Data values assume we begin with an equal-weighted portfolio, and the sampling distributions of the percentages are obtained from the simulations. A greater percentage contributed by the top ten firms implies a greater asymmetry in the distribution of long-term returns.





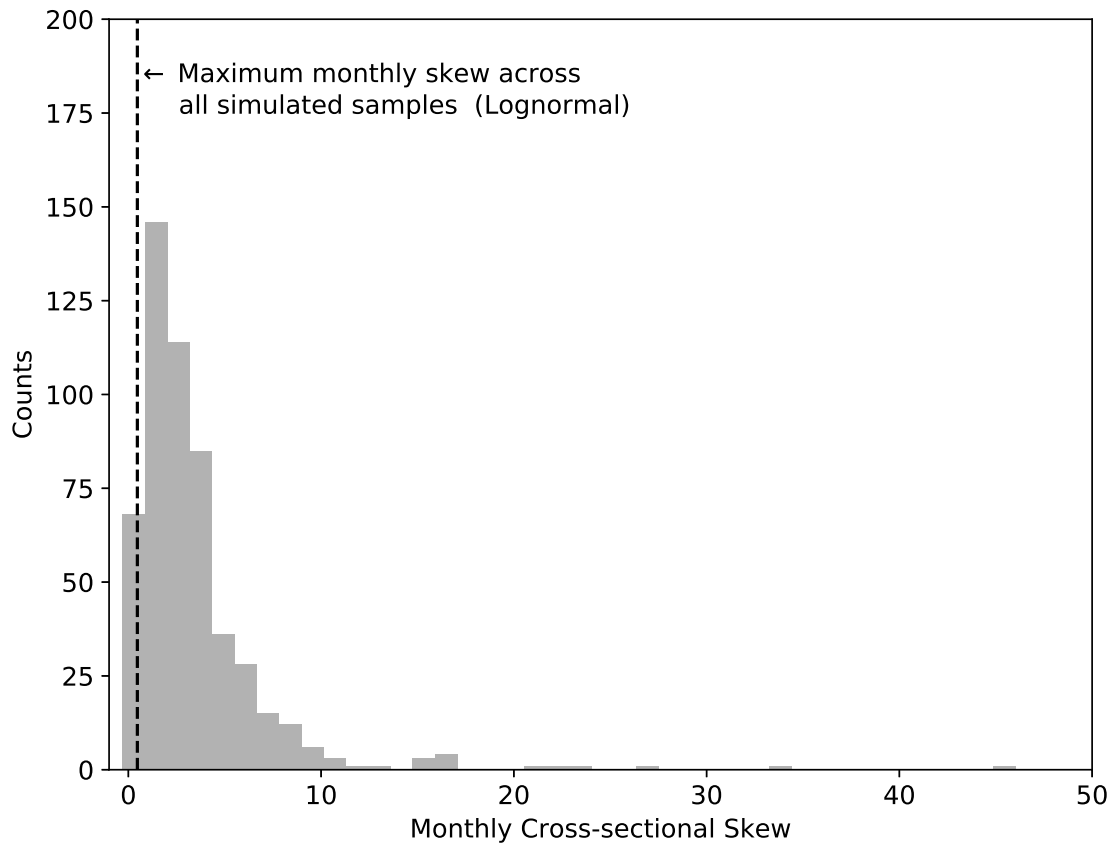
**Figure 1: Historical Number of CRSP Common Stocks**

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 (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 400 simulations.



**Figure 3: 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 400 simulations.