NBER WORKING PAPER SERIES

CROSS-SECTIONAL SKEWNESS

Simon Oh
Jessica A. Wachter

Working Paper 25113
http://www.nber.org/papers/w25113

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2018

We thank Hendrik Bessembinder, John Campbell, Marco Grotteria, Nishad Kapadia, Yapai Zhang, and seminar participants at the Wharton School for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Cross-sectional Skewness
Simon Oh and Jessica A. Wachter
NBER Working Paper No. 25113
September 2018
JEL No. G12

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. However, skewness in long-run returns substantially understates what the lognormal model would predict. Nonstationary share dynamics imply a breakdown in the distinction between market and idiosyncratic risk in the lognormal model. We present an alternative model that matches the skewness in the data and implies stationary wealth shares. In this model, idiosyncratic risk is the primary driver of growth in the economy.

Simon Oh
University of Chicago
sangmino@chicagobooth.edu

Jessica A. Wachter
Department of Finance
2300 SH-DH
The Wharton School
University of Pennsylvania
3620 Locust Walk
Philadelphia, PA 19104
and NBER
jwachter@wharton.upenn.edu

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.[1] No such consensus exists for the cross-section.

In this paper we argue that two types of cross-sectional skewness are important for addressing the question of how to model the cross-section. These are the skewness in short-term returns and the skewness in long-term returns. We argue that it is useful to consider both types of skewness together, since explanations for one type of skewness potentially lead to problems for the other type.

We start with a simple model for stock returns to illustrate the puzzle. 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. 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. This causes a problem for definitions of market risk and idiosyncratic risks. We show that this result is robust to idiosyncratic risk that

[1] See, for example, Bansal and Yaron (2004), Campbell and Cochrane (1999), Wachter (2013).

decreases in market size, and firm payout that policies that change based on the life cycle, both features that are present in the data.

We then present an alternative model based on the binomial distribution, and on shocks that are additive for individual firms but growing with the economy. In contrast to the lognormal model, the binomial model implies stationary dividend shares and substantial cross-sectional skewness. It also implies decreasing standard deviation as a function of size, and increasing payout through the life cycle. We show that this model also implies a role for idiosyncratic risk in the aggregate economy.

Our paper relates to recent work by Bessembinder (2017), who shows that most stocks underperform Treasury bills most of the time, thereby highlighting the importance of skewness in the data. He also notes the high degree of skewness in market capitalizations. 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.

Our work also relates to the research on power laws and firm values (Axtell, 2001; Gabaix, 2009) in that the same mechanism that generates the power law (independent, permanent shocks to growth rates) lies behind our results on long-run returns. Like Gabaix (2011), who focuses on GDP, we show how idiosyncratic volatility can play a role in aggregate outcomes through a failure of diversification.

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.[2] We find that it is very large.

Indeed, while most of the literature has focused on the cross-section of expected returns 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.

The paper proceeds as follows. Section 2 describes the benchmark lognormal for asset returns. Section 3 takes this model to the data on return skewness. Section 3.6 discusses the mechanism in the model that leads to failure for long-horizon returns, and tests various potential fixes. Section 5 builds an alternative model that matches cross-sectional skewness, and nonetheless possesses stationary firm dynamics. We discuss a connection between aggregate growth and idiosyncratic risk. Section 6 concludes.

[2] 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.

2 The standard lognormal model

Recent research in finance provides numerous models for stock returns. Most of these models, however, take a similar form. Stock returns are subject to a common shock and an idiosyncratic shock. In many models, stock returns are an equilibrium outcome of an economy populated by investors who receive an endowment or who make an investment decision subject to a technology.

Assume there are N firms indexed by j.[3] Specifically, let $R_{j,t+1}$ denote the gross return on firm j between time t and $t+1$. We assume a one-factor statistical model for log returns:

$$\log R_{jt} = \alpha_j + \beta_j \log R_{Mt} + \epsilon_{jt}, \tag{1}$$

where ϵ_{jt} are mean zero and iid across time, where ϵ_{jt} and ϵ_{kt} are independent for $j \neq k$, and where ϵ_{jt} is independent of R_{Mt}. We also assume a one-period riskfree asset with constant (gross) return R_f. We use the notation R_{Mt} to denote the common factor without, for now, taking a stance on whether this denotes the true market return. While our main results do not depend on specific distributional assumptions, we assume that ϵ_{jt} are normally distributed for all j and t, and that $\log R_{Mt}$ is also normally distributed. That is, for each $j = 1, \ldots, N$, and for all t, we assume $\epsilon_{jt} \sim N\big(0, \sigma_{\epsilon j}^2\big)$. We assume, for all t, $\log R_{Mt} \sim N\big(\mu_M - \frac{1}{2}\sigma_M^2, \sigma_M^2\big)$, so that $\mu_M = \log E[R_{Mt}]$. Under these assumptions (1) is the discrete-time equivalent of a geometric Brownian motion for the stock price. We assume a one-factor model for convenience, but our results do not depend on this assumption.

To illustrate how this structure can arise, we consider an iid lognormal model. Assume a complete-markets endowment economy with a representative agent with utility

$$\sum_{t=0}^{\infty} \delta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma}, \tag{2}$$

[3] We use the terminology firm and stock interchangeably to mean the entity with returns (1). In the data, we identify firms with CRSP PERMNOs.

where the growth of the aggregate endowment is lognormally distributed

$$\log \frac{C_{t+1}}{C_t} \overset{iid}{\sim} N\big(\mu_c - \frac{1}{2}\sigma_c^2, \sigma_c^2\big). \tag{3}$$

The parameter $\delta > 0$ represents the time discount factor and $\gamma > 0$ the coefficient of relative risk aversion. We consider assets whose dividends are characterized by:

$$\log \frac{D_{j,t+1}}{D_{j,t}} = \left(\mu_j^d - \frac{1}{2}(\sigma_j^d)^2\right) + \beta_j^d \left(\log \frac{C_{t+1}}{C_t} - \mu_c - \frac{1}{2}\sigma_c^2\right) + \epsilon_{j,t+1}^d, \tag{4}$$

with $\epsilon_{j,t+1}^d \sim N(0, \sigma_{\epsilon^d j}^2)$ are i.i.d across time, independent of $\log C_{t+1}/C_t$, and where ϵ_{jt}^d and ϵ_{kt}^d are independent for $j \neq k$.[4] We use the superscript d to denote dividend-related parameters and shocks to distinguish them from return-related parameters and shocks.[5]

Appendix A shows that (1) holds with R_{Mt} being the return on the aggregate consumption claim. Because the aggregate consumption claim, is, by definition, total wealth, this justifies our use of the M subscript. Moreover, as we show in Appendix A, a capital-asset pricing model holds. We summarize these results in Theorem 1.

Theorem 1. *Assuming (2) characterizes utility, (3) characterizes consumption, and (4) characterizes dividends, returns satisfy (1) with*

$$\log E[R_{jt}/R_f] = \beta_j \log E[R_{Mt}/R_f] \tag{5}$$

and

$$\log E[R_{Mt}/R_f] = \gamma \sigma_M^2, \tag{6}$$

where R_f is the equilibrium return on the riskfree asset, and where $\beta_j = \beta_j^d$. Moreover, $\epsilon_{jt} = \epsilon_{jt}^d$.

[4]This is a "multiple-trees" model similar to that considered by Cochrane et al. (2008) and Martin (2013).

[5]Recent work focuses on the separation between risk aversion and the inverse of the elasticity of intertemporal substitution (Epstein and Zin, 1989). However, in this iid model, it is well-known that allowing for this separation nonetheless is observationally equivalent to the form (2), with γ as risk aversion.

The model approximates the CAPM of Sharpe (1964). It is also a special case of the Consumption CAPM of Breeden (1979) and the ICAPM of Merton (1973) when investment opportunities are constant.[6] While this model has been generalized in various ways, these generalizations mainly pertain to the form of the utility function (2) or the distribution of the aggregate endowment.[7] These generalizations are unlikely to make a difference in cross-sectional skewness of firms, which is our focus.

The results above sharpen predictions of previous work on skewness. Since the pioneering work of Kraus and Litzenberger (1976), studies have linked a preference for skewness, and thus deviations from the CAPM, to the condition of non-increasing absolute risk aversion. Here, we assume utility with constant *relative* risk aversion, implying strictly decreasing absolute risk aversion. Asset returns feature both co-skewness and idiosyncratic skewness through the lognormal distribution. However, (5) implies the CAPM holds. How can this be? It turns out that a preference for skewness, and lognormal returns interact precisely to as deliver the CAPM (5), which is of course not exactly the same as the mean-variance CAPM of Sharpe (1964). In this study, we evaluate skewness relative to the lognormal benchmark as opposed to the normal benchmark. Lognormality is arguably a more plausible null hypothesis than normality.

While simple, this model captures some important features of asset price data. First, gross returns always are positive, as they must be because of limited liability (this would not be true if we assumed, for example, a normal distribution). Second, stock returns are unpredictable. This assumption has been extensively evaluated in the data, and there does appear to be statistically significant predictability. However, for our purposes what is important is that most return variance arises from unpredictable changes in prices, as one would expect in an environment with utility maximizing agents.[8] Third, stocks returns have heterogeneous volatilities and are subject to idiosyncratic risks. Fourth, the

[6]Campbell (2008) notes that (5) holds provided that returns are lognormal. The above result justifies the assumption of lognormality.

[7]For example, it is will known that this model cannot explain the equity premium puzzle (Mehra and Prescott, 1985) or the volatility puzzle (Shiller, 1981; Campbell and Shiller, 1988).

[8]See recent work by Cochrane (2008) and Welch and Goyal (2008) and the survey by Campbell (2008). Wachter and Warusawitharana (2015) discuss economically motivated priors on predictability.

distribution of stock returns is stationary. Consistent with this assumption, risk premia and interest rates appear stationary in the data in spite of exponential growth in wealth over the last century and longer (Shiller, 1989; Golez and Koudijs, 2017) and are also similar internationally (Campbell, 2018, Chapter 6). Finally, the model is micro-founded in a model with utility maximizing agents where consumption and dividends are also consistent with important data features (i.e. dividends and consumption are positive, growing over time, with rates of change that are largely unpredictable).[9]

3 Empirical results

3.1 Data

The data consists of monthly returns on ordinary common shares of stocks traded on all major exchanges, available on CRSP, from July 1926 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 also consider two subsets when calibrating our simulations.

The first subset consists of all firms with at least 60 months of returns. We choose this cutoff because it allows for plausible estimation of the parameters. This yields a universe of 16,087 firms. We also consider a much smaller subset of the firms continuously in existence between January 1973 and December 2016. For the second subset, we restrict our universe to stocks without missing data for monthly returns. This gives us 404 firms. This set of stocks is clearly subject to survivor bias. However, it has the advantage that we can directly compare data to our simulated firms without the need to consider entry and exit.

Table 1 reports return moments for all CRSP stocks, for the subset of stocks for which

[9]Like the predictability of stock returns, the predictability of consumption growth, and, more to our point, dividend growth, has been extensively evaluated (van Binsbergen and Koijen, 2010). In Section 3.6 we discuss how introducing predictability in dividend growth affects our results.

we have more than 5 years of data, and for the stocks with continuous returns between 1973 and 2016. The mean of monthly net returns is 1.1% across all stocks. It is slightly higher–1.3%–across the stocks with more than 5 years of data. For both sets of stocks, the median is indistinguishable from zero. The standard deviation of returns for both sets is 17% per month. Both are quite positively skewed, with a skewness coefficient of about 6.

Overall, the stocks with more than 5 years of data have similar cross-sectional moments compared with the full universe. The stocks in continuous existence, interestingly, have a similar mean return. However, their median return is higher: it is 1% whereas the median return is indistinguishable from zero for the wider population. The standard deviation is 10% versus 17% for the wider population, and the skewness is 1.3 versus 5.8.

In recent work, Bessembinder (2017) notes that on average, individual stocks do not outperform Treasury bills. Because this finding has a connection to skewness, as Bessembinder discusses and as we elaborate on below, we follow his study and report the percent of stock returns exceeding the one-month Treasury bill return. Confirming his results, we find that only 48% of stock returns exceed the Treasury bill return of that month.

3.2 Simulation strategy

We now ask how likely these and other features of the data are to occur in the economy described in Section 2. To do this, we create fictitious samples using Monte Carlo simulation.

An immediate question arises: how many stocks do we consider, and how do we handle the fact that stocks move in and out of the sample? These questions make clear the benefit of our small sample of 404 companies in continuous existence. In the case of this sample, we can calibrate our firms to these companies. However, this approach has the problem of survivorship bias. The distribution of means and standard deviations, for example, may not represent the true ex-ante distribution facing investors.

Figure 1 illustrates one challenge in modeling the cross-section of individual stocks.

This figure shows how the number of CRSP stocks has fluctuated over time. The number of stocks, below 1000 prior to 1950, reached a peak of about 8000 in the late 1990s, and has since declined to around 4000. There is a jump in the number in January 1973 corresponding to the establishment of the NASDAQ.

To roughly capture the number of firms trading in the stock market, we set the number of firms in our simulation equal to the median value in Figure 1, which is 2,440. 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 (16,087 versus 2,440). We thus use the following bootstrap procedure. At the start of each fictitious sample, we draw 2,440 stocks from the universe of 16,087 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 N_j is the number of months that firm j is listed, we draw from the estimated parameters of firm j with probability $\frac{N_j}{\sum_{i=1}^{16,087} N_j}$. Thus, across fictitious samples, we should roughly capture the true distribution of firms in the cross section.

Given a set of firm-level parameters, we simulate fictitious samples assuming returns are distributed as in Section 2. We consider samples of two different lengths to represent the sample with 16,087 firms from January 1973 to December 2016 and the sample with 404 firms from July 1926 to December 2016. For each type of simulation, we consider 400 such fictitious samples.

For each sample, we draw random normal shocks ϵ_{Mt}. We then draw N_{firm} sequences of iid idiosyncratic shocks ϵ_{jt}. Stock j's time-t return is then

$$\log R_{jt} = \mu_j - \frac{1}{2}\sigma_j^2 + \beta_j \sigma_M \epsilon_{Mt} + \sigma_{j,\epsilon}\epsilon_{jt}. \tag{7}$$

Below, we describe the parameter estimation. We set the riskfree rate to a constant in the simulations and equal to the average rate on the 1-month Treasury bill.

3.3 Estimation

We estimate the firm-level parameters using CRSP data. We take the sample mean return on the value weighted portfolio to estimate $e^{\mu_M} = E[R_{Mt}]$ and the variance of the log return to estimate $\sigma_M^2 = \mathrm{Var}(\log R_{Mt})$.

For each stock j, we estimate β_j from the OLS regression (1). That is, we estimate β_j using log returns. While most studies use level returns, running the regression in log returns is what the theory suggests. Given β_j and μ_M, we compute the expected log return on stock j as

$$\mu_j = (1 - \beta_j) \log R_f + \beta_j \mu_M. \tag{8}$$

We estimate $\sigma_{j,\epsilon}^2$ as the variance of the residuals in (1). Then if σ_j^2 is the variance of the total return of $\log R_{jt}$, it must be that

$$\sigma_j^2 = \sigma_{j,\epsilon}^2 + \beta_j^2 \sigma_M^2,$$

thus the estimates of β_j, σ_M^2, and $\sigma_{\epsilon,i}^2$ provide a sample estimate of σ_j^2. Standard OLS regression results tell us that this is the same estimate we would obtain if we estimated the sample standard deviation of log returns directly. This completes the set of parameter values we need to simulate from the model.

As a robustness check, we directly estimate μ_j using the sample average. Namely,

$$\mu_j = E\left[\log R_j\right] + \frac{1}{2}\sigma_j^2, \tag{9}$$

where we compute $E\left[\log R_j\right]$ using the sample average of log returns and σ_j^2 as the sample variance of log returns.

3.4 Pooled monthly returns

In this section, we compare characteristics of monthly returns in the model and in the data. In each fictitious sample path, we compute the mean, the variance, and other statistics.

Thus for each statistic we have a sampling distribution consisting of 400 "observations." We report the median value, the 5th and 95th percentile values, and the minimum and maximum across all 400 simulations. We also report population statistics computed from pooling the entire set of simulations.

Table 2 reports the results from the simulations. Panel A and C rely on the CAPM for estimation, while Panels B and D use direct estimation of mean returns. Direct estimation successfully replicates the mean return, while relying on the CAPM appears to somewhat understate the mean return. Both methods, and both types of simulations, replicate the standard deviation of returns almost exactly. Thus the model can successfully capture the first two moments of returns, perhaps not surprisingly.

The model cannot, however, capture the skewness of returns. For the full set of firms, the skewness is 5.846, while the model predicts a skewness of mere 0.938. Moreover, the model implies that skewness is well-estimated, with the minimum, across all 400 draws being 0.849 while the maximum is 1.287. Skewness in the data is therefore far above what could be possibly implied by the model. This discrepancy occurs whether or not the mean is estimated using the CAPM or directly.

One may think that the skewness is a feature of the potentially short-lived firms that populate the larger sample. However, skewness that is excessive relative to the model is also present for the 404 long-lived firms. Here, the skewness is much smaller: 1.321. It is still far above what the model generates, however (the maximum in this calibration is 0.561).

Note that the model fails to match skewness in the data despite the fact that returns are positively skewed in the model due to the lognormal distribution. To highlight this point, we also examine the skewness of log returns. Skewness of log returns in the model is very slightly negative.[10] Interestingly, skewness of log returns is also very slightly negative in the data. Along this dimension, the model successfully matches the data, but

[10]Skewness is not exactly zero for log returns because we are sampling from a distribution with heterogeneous variances. The skewness of the return distribution depends on properties of the variance distribution.

it is misleading: the low log skewness in the data apparently occurs not because there is a heavy right tail, but because the left tail outweighs the right in the skewness calculation. Looking purely at log returns, one would miss the important result that the model fails to match the skewness of level returns.

Motivated by results of Bessembinder (2017), we also calculate the percent of observations that exceed the riskfree rate. The model succeeds in matching the data for both sets of simulations. For the full set of firms, the fraction of returns exceeding the riskfree rate is 48% in the data. It is also 48% in population in the model.

The model's success in matching the proportion of returns exceeding the riskfree rate is important for two reasons. First, it shows that it is possible, in an equilibrium model with risk averse agents, to have returns on equities fall below the riskfree rate most of the time. Second, the skewness in the data appears unrelated to this result; the model can capture the percent of observations above the riskfree rate, but it fails to capture the skewness.

How is it that, in the fully rational equilibrium model, in which there is an equity premium, stocks underperform Treasury bills most of the time? There are two reasons why this occurs: the first is due to lognormality, and the second, which enhances the first, is due to undiversified risk.

Consider first the case of the market portfolio. In this case, there is no undiversified risk.[11] Define an $N(0,1)$ random variable ϵ_M and let $r_f = \log R_f$. Because the log is a monotonic transformation,

$$\Pr(R_M > R_f) = \Pr(\log R_M > r_f) = \Pr\big(\sigma_M \epsilon_M \geq -(\mu_M - r_f) + \frac{1}{2}\sigma_M^2\big).$$

Because $\epsilon_M \sim N(0,1)$, the probability that $\sigma_M \epsilon_M \geq -(\mu_M - r_f) + \frac{1}{2}\sigma_M^2$ exceeds 50% if and only if $-(\mu_M - r_f) + \frac{1}{2}\sigma_M^2 < 0$, or equivalently if $\mu_M - \frac{1}{2}\sigma_M^2 - r_f > 0$. If $\mu_M - \frac{1}{2}\sigma_M^2 - r_f < 0$, then we would expect the market portfolio to underperform Treasury bills more than half

[11]Bessembinder (2017) also analyzes this result, but assumes that the excess return is lognormally distributed. This is a less appealing assumption because the excess return could, theoretically, be negative, while the return itself cannot.

of the time.

Is it possible, under the model above, to have $\mu_M - \frac{1}{2}\sigma_M^2 - r_f < 0$? It is, because the model only requires that $\mu_M - r_f > 0$. The mean return is greater than the riskfree rate, but the median return is not.

However, while this analysis shows that it is theoretically possible for more than half of the observations on the market portfolio to fall below the riskfree rate, it is not likely given the data. The equity premium, $\mu_M - r_f$ is about 0.07 per annum, while $\frac{1}{2}\sigma_M^2$ is about 0.02 per annum. The magnitude of the equity premium is simply too large.

However, for an individual stock, the effect above is enhanced because of idiosyncratic volatility. Recall that, for the market return, the condition that returns exceed the riskfree rate more than half the time is $\mu_M - \frac{1}{2}\sigma_M^2 - r_f > 0$. For an individual stock j, this condition naturally becomes $\mu_j - \frac{1}{2}\sigma_j^2 - r_f > 0$. However, while μ_M is in some sense an average of μ_j across j,[12] σ_M^2 is far below the average of σ_j^2 on account of diversification.

Finally, note that none of this analysis depends on the horizon. If returns are log-normally distributed, means and standard deviations both scale linearly in the horizon, as does the riskfree rate return. The probability of long-horizon returns exceeding the return on the Treasury bill position is the same as the probability of short-horizon returns exceeding the Treasury bill position.

3.5 Monthly returns 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.

[12] To be precise: $\mu_M = \log(E[R_M]) = \log(\sum w_j E[R^j]) = \log(\sum w_j e^{\mu_j})$.

To test this hypothesis, we compute skewness in the cross-section at each point in time. We calculate

$$\text{skew}_t^{cs} = \frac{\frac{1}{n}\sum_{i=1}^{n}(R_{j,t} - \bar{R}_t)^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(R_{j,t} - \bar{R}_t)^2\right]^{3/2}}, \tag{10}$$

where $\bar{R}_t = \frac{1}{n}\sum_{i=1}^{n} R_{i,t}$, that is, the cross-sectional mean of the return. In a sample of length T, we obtain T observations of cross-sectional skewness. We first examine results for the 16,087 firms; 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 vast majority of data observations exceed the maximum value implied in the 400 model simulations. Figure 3, which shows analogous results for the 404 firms, tells a similar story. Finally, Table 3 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]

3.6 Long horizon returns

The previous section shows that the lognormal model vastly under-represents positive return skewness in monthly returns in the data.

We now consider long-horizon returns. One advantage of the iid lognormal model is that it makes very clear predictions about long horizon returns. Both the mean and the variance of log returns scale perfectly in T. For a given stock j, skewness increases in T because

$$\text{skew}(\sigma_j^2) = \left(e^{\sigma_j^2} + 2\right)\sqrt{e^{\sigma_j^2 - 1}}.$$

[13]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. Moreover, the computation of idiosyncratic volatility also depends on the underlying return generating process, a point to which we will return in Section 5.

However, the percent of returns of stock j exceeding the riskfree rate depends only on the properties of the log return, and thus is horizon-invariant.

One striking feature of the data, is the skewness in the size of firms. The recent striking growth in the stock-market capitalization of so-called FANG[14] companies, despite their already large size have called wide attention to this phenomenon. However, it is not a new feature of the data. Axtell (2001) discusses the extreme skewness in firm sizes, which he characterizes with a power law distribution.[15] Bessembinder (2017) discusses the right tail in a metric of wealth creation.

We now ask whether our model can capture this skewness. We examine the cross-section of cumulative returns and growth in market capitalization implied by the model of Section 2. We consider, for each statistic, the proportion of the total value of that statistic captured by the top 10 firms. We first compute this quantity for cumulative returns. Table 4 confirms the skewness reported in previous studies holds for cumulative return measures: the ten firms with the highest returns constitute 29% of the total when measured by growth in market capitalization and 72% of the total when measured by cumulative return. It would seem unlikely that the model, which is unable to capture the cross-sectional skewness, could capture how a small fraction of firms could so dominate when measured according to these statistics.

To compare the model with the data, we cumulate returns in each simulation, with returns given by the exponential of (1). This equation does not take a stand on whether the return comes from price appreciation or from the dividend yield, and therefore we report the same growth in market capitalization and cumulative return in the model. We discuss how introducing dividend payouts would affect the model's results later.

Table 4's results for methods 1 and 2 illustrate that the model overstates long-run skewness, rather than understating it. [16] In the data, the top 10 firms account for 72%

[14]Facebook, Amazon, Netflix, Google

[15]In our paper, firm sizes do not satisfy a power law distribution because there does not exist a stationary distribution of firm sizes, as discussed below. From a technical point of view, it is helpful to have a stationary distribution of firm sizes. It is less clear, however, that the true distribution of firm sizes is in fact stationary.

[16]Independently and concurrently, Bessembinder (2017) shows that a model with normally distributed

of the cumulative return and 29% of growth in market capitalization. For most samples in the simulations, the top 10 firms account for a full 99% of the value.

4 Discussion

4.1 Skewness in long-horizon returns

Where does the skewness in long-horizon returns come from? It arises from the cumulative effects of randomness. The log difference between any two returns is a random walk, and thus, by a well-known mathematical result, must wander arbitrarily far away from zero. It follows that "most" of the time, an economy that has run sufficiently long will feature a highly skewed, and in fact nonconvergent distribution of returns.

To see this more precisely, define the cumulative return for asset j between t and T:

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

From (1), for any two firms j, k

$$\log \frac{CR_{k,t,t+T}}{CR_{j,t,t+T}} = T(\alpha_k - \alpha_j) + (\beta_k - \beta_j) \log CMR_{t,t+T} + \sum_{i=1}^{T} (\epsilon_{k,t+i} - \epsilon_{j,t+i}) \tag{11}$$

where $CMR = \prod_{i=1}^{T} R_{m,t+i}$ is the cumulative market return. We then have the following result

Lemma 1. *For any two firms j, k with $j \neq k$, the difference between log cumulative returns is a random walk with drift.*

Proof. The result follows because both $\log CMR_{t,t+T}$ and $\sum_{i=1}^{T} (\epsilon_{k,t+i} - \epsilon_{j,t+i})$ are random walks. □

The random walk is known to be nonstationary. The variance of $\sum_{i=1}^{T} (\epsilon_{k,t+i} - \epsilon_{j,t+i})$

idiosyncratic shocks but otherwise identical firms can also overstate the skewness in the data.

is $T(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2)$, namely it increases linearly in the horizon. This means that the difference between log returns can wander anywhere on the real line, and so in practical terms it will spend most of its time at very large positive or large negative numbers. To be precise:

Corollary 1. *There is a subsequence $s(T)$ such that for every j, k, $j \neq k$, $\log CR_{k,t,t+s(T)} - \log CR_{j,t,t+s(T)}$ approaches negative or positive infinity as $s(T)$ approaches infinity.*

Proof. The result follows from Lemma 1 and results in Feller (1968). □

Now consider ratios of the form

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

which give the percent contribution of the cumulative return on firm j to the total across the market. Corollary 1 implies that these ratios to be either very close to 1, or very close to zero most of the time, provided that the economy has run long enough. The large percentages of the total gain accounted for by merely ten assets shown in Table 4 demonstrate this effect in the simulations.

Skewness in cumulative returns is important for two reasons. First, it is at odds with the data, as Table 4 shows. Even if we did not reject the lognormal model because it delivers insufficient short-horizon skewness, we could reject it because of excessive long-horizon skewness. Note, however, that while the data rejects the lognormal model, it is not possible to tell from the data whether the underlying process for cumulative returns, or for that matter, market capitalization growth, is stationary.

This leads to the second reason why skewness in cumulative returns is important: because of what it says about market values. Divergent market values cause a theoretical problem that we describe in the next section. One could, in a model, have divergent cumulative returns, but a stationary distribution of relative market values. Unfortunately, however, under the assumptions in Section 2, relative market values do diverge.

4.2 Skewness in market capitalization

The model for returns in (1) does not distinguish between returns including and excluding dividends. To address the distinction between market capitalization and cumulative return, we need a model for the dividend yield. The iid endowment economy model described in Section 2 is one such model.

The model in Section 2 implies that ratios of prices to dividends are constant. Namely,

$$\frac{P_{jt}}{D_{jt}} = \frac{1}{\Phi_j - 1} \tag{12}$$

for an asset-specific constant Φ_j such that $\Phi_j > 1$ (see Appendix A). The finding of a constant price-dividend ratio is intuitive because the growth rate of dividends and the discount rate applied to future cash flows are both constant. Equation (12) implies that growth rates in dividends and in market capitalizations are the same. Moreover, (4) implies that dividend growth follows a law of motion that is analogous to that of returns. Therefore, the results for cumulative returns also hold for market capitalizations:

Lemma 2. *For any two firms j, k with $j \neq k$, the log ratios of market capitalizations $\log P_{jt}/P_{kt}$ are a random walk with drift.*

Proof. The result follows from (4) and the fact that the ratio of prices to dividends is a constant (12). $\square$

Corollary 2. *There is a subsequence $s(T)$ such that for every j, k, $j \neq k$, $\log P_{j,s(T)}/P_{k,s(T)}$ approaches negative or positive infinity as $s(T)$ approaches infinity.*

This result once again follows from Feller (1968) and the fact that log price differences are random walks. Thus a small number of firms, and eventually a single firm, will come to dominate the economy.

Finally note that cumulative returns and market capitalizations are closely related:

$$
\begin{aligned}
CR_{j,t,t+T} &= \prod_{i=1}^{T} \left(\frac{P_{j,t+i} + D_{j,t+i}}{P_{j,t+i-1}} \right) \\
&= \prod_{i=1}^{T} \left(\frac{P_{j,t+i}/D_{j,t+i} + 1}{P_{j,t+i-1}/D_{j,t+i-1}} \frac{D_{j,t+i}}{D_{j,t+i-1}} \right) \\
&= \Phi_j^T \frac{D_{j,T}}{D_{j,t}}
\end{aligned}
$$

and thus

$$
CR_{j,0,T} = \Phi_j^T D_{j,T} = \Phi_j^T (\Phi_j - 1) P_{j,T}
$$

where we have normalized the initial price in the economy to equal 1.

Intuitively, since returns and dividends are driven, in equilibrium, by the same set of shocks. Thus the cross-section of firm values is nonstationary and, as with cumulative returns, with probability one a single firm will come to take over the economy.

4.3 Idiosyncratic and market risk

A consequence of divergent ratios of market capitalization is two incompatible notions of a market portfolio. Consider two possible definitions:

(a) The claim to aggregate consumption

(b) The value-weighted portfolio of dividend claims.

Note that (a) is identified with R_{Mt} in (1). It is the common factor of returns and dividends. It represents overall wealth in the economy. Relative to this notion of the market portfolio, the terms ϵ_{jt} are idiosyncratic and unpriced.

That said, (b) is directly analogous to the CRSP value-weighted return typically used in studies of the stock market. However, (b) is necessarily dominated by a very few firms. For these firms ϵ_{jt} represents market risk, not idiosyncratic risk.

How is it possible that (a) and (b) are different? Recall that in this model, the CAPM holds, at least approximately. It then seems that investors should indeed hold the value-weighted portfolio. The answer is that they do hold the portfolio with return R_{Mt}, which pays the consumption claim as the dividend. The model in Section 2 does not enforce that the the dividends add up to consumption. There must be another traded asset, a hidden one, with potentially strange properties, that makes up the difference. The hypothesis that the market portfolio does not correspond to the observed stock market is an old one (Breeden, 1979; Jagannathan and Wang, 1996). This discussion shows not only that such nontraded assets must exist, but that their share in the economy either goes to zero (in which case consumption is negligible) or 1 (in which case the stock market is negligible) as time goes by.

Even though the definitions of the market portfolio are incompatible, it may still seem surprising that the idiosyncratic shocks ϵ_{jt} do not diversify away in the market portfolio. After all, (1), because it is based on log returns, holds equally for long and short-horizon returns. The Central Limit Theorem, and the Arbitrage Pricing Theory of Ross (1976) should imply that ϵ_{it} is diversified away. The problem is that the variance of the idiosyncratic risk grows linearly in the horizon. It is necessary to add assets at a faster rate than the the length of the horizon for the diversification to take place. Figure 1 suggests that this has not occurred. [17]

How then can one model the market and individual firms at the same time? There are other modeling options, but these come at arguably greater cost. One partial solution is to assume away permanent shocks, a standard approach in the macroeconomics literature (Zhang (2005) is an example of such an approach in the finance literature). In such a model the cross-section of cumulative returns still has a nonstationary distribution. However, market capitalization does have a stationary distribution. This itself is at potentially at odds with the data, which suggests a heavy right tail in cumulative growth rates, just not as heavy as in the model. Moreover, in assuming away permanent economic shocks,

[17] Gabaix (2011) also argues that firm-level fluctuations could be important for aggregate shocks. However, his argument is different from ours. He starts with the assumption of a stationary power law distribution in firm sizes. Because some firms are very large, their contributions to GDP growth are significant, are not diversified away. In our framework, there is a nonstationary distribution of firm sizes, which creates a theoretical inconsistency in the definition of market risk.

these models assume away a feature of the economy that is both undeniably present and of great consequence to asset pricing.[18]

A second approach is to assume away idiosyncratic shocks. Many papers take this approach to modeling the cross section. Different implementations of this approach include Hansen et al. (2008), Lettau and Wachter (2007), Santos and Veronesi (2010), Tsai and Wachter (2016), Wachter and Zhu (2017).[19] When one is modeling features of portfolios, not individual firms, this is arguably reasonable, though it does take the distribution of firm sizes in the data as given. However, in assuming away firm-specific shocks, this approach also assumes away an important feature of the economy: the empirical results in Section 3.4 make clear that there do exist large, firm-specific shocks. By assuming them away, we ignore the important question of the relation between these shocks and long-run uncertain growth, which is ultimately what generates risk premia in the first place.

4.4 Mechanisms to reduce long-horizon skewness

In Section 3.6, we have seen that the lognormal model overstates long-run skewness: while in the data the top ten firms account for 72% of the cumulative return and 29% of the growth in market capitalization, in the simulations they account for a full 99% in most samples. We now examine if such discrepancy persists even after incorporating mechanisms that are designed to reduce the long-run skewness in our simulations. Specifically, we consider the effect of dividends and time-varying idiosyncratic volatility.

In the first set of modified simulations, we consider dividends. While this would not affect skewness in long-run returns, it has the potential to decrease skewness in the growth rate of market capitalizations. There is empirical support for the argument that

[18] Under minor deviations from time-additive utility, temporary shocks receive a much reduced price of risk relative to permanent shocks. Zhang (2005) and other papers avoid this problem by assuming an exogenous stochastic discount factor.

[19] Hansen et al. (2008) model consumption-dividend ratios as stationary and do not address the question of the market portfolio. Lettau and Wachter (2007) and Santos and Veronesi (2010) model portfolio shares as stationary. Tsai and Wachter (2016) model two claims that by definition add up to the market. While the sizes are non-stationary, claims are re-issued every time a shock takes place place, and so the resulting distribution of portfolio market capitalization is in fact stationary. Wachter and Zhu (2017) assume portfolios differ only in their exposure to a rare event which is not realized in the sample.

firms initiate dividends after reaching maturity in their life cycles.[20] Our simulation procedure allows us to account for this. We use the following simple definition of maturity: we assume firm has matured when its market capitalization exceeds that of the median market capitalization in the previous month. We then subtract the dividend yield from the return when computing growth in market capitalization. Once a firm initiates dividends, it continues to pay, even if its value drops below the median in subsequent months.

For the purposes of calibration, we compute the monthly dividend-price ratio for the 16,087 firms from our baseline simulation.[21] We set the dividend-price ratio used in our simulation to be equal to the average value computed from the CRSP data, which is 0.1786%.

We also consider the possibility that idiosyncratic volatility might shrink as firms become larger. Herskovic et al. (2016) document an inversely monotonic relation between firm size quintile and return volatility. To account for this effect, we assign firms lower values of idiosyncratic volatility as they become large relative to other firms. Specifically, we sort firms into size deciles in the simulations. We assign firms an idiosyncratic volatility based on their size decile. Those in the smallest market cap decile are assigned the highest idiosyncratic volatility, and those in the largest market cap decile are assigned the smallest idiosyncratic volatility.[22] Firms idiosyncratic volatility can change, month-by-month in the simulation.

We obtain the idiosyncratic volatilities through the following calibration procedure. We calculate the idiosyncratic volatility for firms in our sample, and sort these into idiosyncratic-volatility deciles. The equal-weighted average idiosyncratic volatility then

[20] Fama and French (2001) and Bulan et al. (2007) find that larger and mature firms are more likely to initiate dividends.

[21] We define the monthly dividend-price ratio as follows:

$$\frac{1}{12}\frac{\sum_{i=0}^{11} D_{t-i}}{P_t},$$

where P_t and D_t are price and dividend in period t, respectively, and where we follow the standard practice of using the previous twelve months of dividends to eliminate seasonality.

[22] In the first month of the simulation, all firms are identical and thus assigned the average of the idiosyncratic volatility.

becomes the assigned value for that decile. We report these values in Table 5.[23]

Note that our procedure biases us against finding skewness in the simulations in two ways. First, we no longer capture the true dispersion in volatilities reflected in CRSP data because we assign the same volatility to firms in the same decile. This dispersion is crucial for capturing skewness. Second, we assign firms that are small to the highest idiosyncratic volatility decile and those that are large to the lowest. This overstates the relation between size and volatility, thereby shrinking the volatility of larger stocks by more than they may shrink in the data. Our procedure allows us to conduct numerically feasible simulations while capturing the size-volatility relation in a way that works in favor of the lognormal model.

Results for both modifications are shown in Table 4. As might be expected, allowing for monthly dividends implies that it is less likely for a small number of firms to take over the economy. This is reflected in a lower percentage for the minimum and 5th percentile. However, in most simulations, the top ten firms still entirely take over the economy when the model is calibrated to the larger number of firms over the longer data sample. The reason is that the effect of random growth described in the previous section still holds among the very largest firms, even when firms pay dividends.[24] Allowing for time-varying volatility based on size has a more noticeable effect. In fact, market cap growth and cumulative return in the data are now above the minimum that can be achieved in the simulation. However, the median percentage, when the economy is calibrated to the larger group of firms over the longer time period, is still 99%. The persistence of the discrepancy between model and data seems to imply that understatement of skewness at short horizons and its overstatement at long horizons require a more fundamental reconciliation.

[23]For our baseline simulation, we compute these parameters *after* we draw the 2,440 stocks that constitute each sample; for the simulation with 404 firms, the parameters are computed beforehand to circumvent any redundant calculations.

[24]We could have potentially found larger effects by forcing a strictly increasing relation between dividend yield and size. However, not such relation exists in the data.

5 A binomial model

In this section we present a simple model that qualitatively captures many features of that the data that are out of reach for the lognormal model. Because our intent is to capture cross-sectional skewness, we abstract away from features like aggregate volatility, and therefore risk premia. As will be clear, incorporating these features into the model is straightforward.

Consider a cross-section of firms indexed by $j = 1, \dots n$, and let $D_{j,t}$ denote the dividend paid by firm j at time t. For each j and t, define a Bernoulli random variable $X_{j,t}$ with parameter λ (that is, $X_{j,t} = 1$ with probability λ and zero otherwise). The Bernoulli random variables are independently distributed across times t and firms j. Let R denote the (constant) discount rate in the economy.

We recursively define the processes $D_{j,t}$ and D_t as follows. Let $D_0 = 1$ and $D_{j0} = 1/n$. Given D_t, define

$$D_{j,t+1} = D_{j,t} + \frac{1}{n} D_t e^z X_{jt} \tag{13}$$

Given $D_{j,t+1}$, define

$$D_{t+1} = \sum_{j=1}^{n} D_{j,t+1}. \tag{14}$$

By construction, individual firm dividends add up to the market dividend, which can then be set equal to consumption in a representative agent model. Thus this model avoids the pitfalls of the lognormal model described above.

The following results characterize the aggregate dividends.

Proposition 1. *The aggregate dividend follows the process:*

$$D_{t+1} = D_t + D_t e^z \frac{1}{n} \sum_{j=1}^{n} X_{j,t+1}.$$

Standard binomial results can then be used to characterize $\frac{1}{n}\sum_{j=1}^{n} X_{j,t+1}$. As n approaches infinity, this quantity converges quickly to a normal distribution with mean λ and standard deviation which falls with $\sqrt{n}$. However, given that this quantity is easy to compute numerically, characterizing its distribution is not necessary for our purposes.

In this model, firms are subject to unusual large shocks represented by $X_{j,t}$. These shocks generate a high degree of cross-sectional skewness, so much so that it is possible to calibrate the skewness in the model to the data. Though shocks to the model are identically distributed over time, firm-level volatility, as measured in the usual way by squared differences in returns over short intervals, appears to be time-varying (Figure 4). Time-varying volatility, however, simply reflects the rare realizations of the skewed shocks.

From the point of view of an individual firm, the realizations of $X_{j,t}$ are rare and unpredictable (in our calculation, they occur in 2.5% of simulation months for a given firm). A calculation of growth rate in firm's dividends, most of the time, would lead to a conclusion of zero growth. Yet, from the point of view of the economy as a whole, growth occurs with near-certainty, at a rate of λe^z. In this model, growth in the economy is due entirely to the rare idiosyncratic firm-level shocks.

In expectation, firms do participate in economy-wide growth, and this is reflected in their price:

Proposition 2. *Given discount factor R, the ex-dividend price of firm j at time t equals*

$$P_{j,t} = \frac{D_{j,t}}{R-1} + \frac{1}{n} D_t \left(\frac{1+\lambda e^z}{R-(1+\lambda e^z)} - \frac{1}{R-1} \right), \tag{15}$$

whereas the value of the aggregate market equals

$$P_t = D_t \frac{1+\lambda e^z}{R-(1+\lambda e^z)}. \tag{16}$$

Equation 15 shows that the price of firm j is given by the sum of its "cash-cow" value, namely, the value of its current dividend discounted into the future, plus a term that accounts for future growth. In this simple model, the term that accounts for future growth is equal across firms.

Define the share of firm j's dividend in the aggregate economy as $\alpha_{jt} = D_{jt}/D_t$. Then

α_{jt} evolves according to

$$\alpha_{j,t+1} = (\alpha_{jt} + \frac{1}{n} e^Z X_{j,t+1}) \left(\frac{D_{t+1}}{D_t} \right)^{-1} \tag{17}$$

Define $\kappa = E\left[(D_{t+1}/D_t)^{-1}\right]$. Assuming $\kappa < 1$, that is, assuming that on average the economy grows over time, α_{jt} has a stationary mean given by

$$\bar{\alpha} = \frac{1}{n} e^Z \lambda \frac{1}{1-\kappa}.$$

We can expect that this model will generate a well-defined stationary distribution.

By definition, the return on firm j equals

$$R_{j,t+1} = \frac{P_{j,t+1} - P_{jt} + D_{j,t+1}}{P_{jt}} \tag{18}$$

It is straightforward to show that this model captures the dependence between size and idiosyncratic volatility.

It follows from (13) that the the conditional variance of $D_{j,t+1}$ is the same for all firms. Thus the conditional variance of $P_{j,t+1}$ is the same for all firms. It follows that the conditional volatility of returns decreases in P_{jt}.

Intuitively, investment opportunities in this economy are of a fixed size. In contrast, in the lognormal economy, the size of investment opportunities scales with the size of the firm. The decreasing relation between firm size and idiosyncratic volatility suggests a mechanism such as the one presented here might be at work.

Table 6 summarizes the properties of this model relative to the lognormal model. Unlike the lognormal model, this model is capable of accounting for cross-sectional skewness. It also accounts for the decreasing relation between firm size and volatility, which the lognormal model in Section 2 does not do (in the simulations, we hard-wired this relation to hold). It produces a stationary distribution of firm shares, which is convenient for modeling purposes. However, it does not produce the long-horizon skewness that characterizes the data.

The model has other potential implications. For instance, variation in the parameter λ or z would lead to time-varying idiosyncratic volatility. If these shocks were priced (as they would be under Epstein and Zin (1989) utility, or if they were correlated with realized consumption), then the model would produce both common variation in idiosyncratic volatility, and pricing effects for this factor, as shown in Herskovic et al. (2016).

Nonetheless, the lack of long-horizon skewness suggests that this model is not a sufficient description of the data. It may be that the true process lies between this model and the lognormal model.

6 Conclusion

We have investigated skewness in the cross section of asset returns through the lens of a lognormal model of returns with common and idiosyncratic shocks. We show a seemingly paradoxical result: the model both dramatically understates the skewness in the cross-section and overstates the skewness in cumulative returns. The understatement is at short horizons, while the overstatement is at long horizons.

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. In models, this creates a nonstationary distribution of relative firm values. Assuming a separate market portfolio that is not the value-weighted average of firms solves the problem in so far as one can proceed with the calculation of equilibrium expected returns. However, such a market portfolio diverges from the value weighted portfolio of firms, so that either the stock market, or consumption becomes negligible. Besides being unrealistic, this inevitably creates difficulties in a true general equilibrium setting when consumption is constrained to equal firms' production. It also ignores the connection between innovation (which takes place at the firm level) and aggregate growth. We present a model that confronts this problem by implying stationary firm sizes. However, this does not capture the long-horizon skewness in the data. The long-run distribution of firm sizes, should it exist, and the mechanisms keeping that distribution stationary (if indeed it is stationary) are thus of fundamental importance in asset pricing.

A Asset prices in the lognormal model

Consider an infinite-horizon complete-markets endowment economy. Assume a representative agent with utility (2) and endowment with iid log growth rates $\log C_{t+1}/C_t \overset{iid}{\sim} N\left(\mu_c - \frac{1}{2}\sigma_c^2, \sigma_c^2\right)$.

Let R_{Mt} denote the gross return on the asset that pays consumption as its dividend (namely, aggregate wealth). The Euler equation implies

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{M,t+1}\right] = 1$$

Let P_{Ct} be the ex-dividend price of the consumption claim. Conjecture that P_{Ct}/C_t is a constant, and define

$$\Phi_C \equiv \frac{P_{Ct}/C_t + 1}{P_{Ct}/C_t}$$

Then

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{1-\gamma} \Phi_C\right] = 1.$$

Thus the price-dividend ratio is (indirectly) characterized by

$$\Phi_C^{-1} = E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{1-\gamma}\right] \tag{A.1}$$

which has a solution as long as parameters are such that the right hand side is less than 1. This confirms the conjecture. Because

$$R_{M,t+1} = \Phi_C \frac{C_{t+1}}{C_t}, \tag{A.2}$$

R_{Mt} is also lognormally distributed. Equation A.2 also implies that $R_{M,t+1}$ is perfectly correlated with $\frac{C_{t+1}}{C_t}$, and that $\log R_{M,t+1}$ and $\log \frac{C_{t+1}}{C_t}$ have equal variance.

The Euler equation for the riskfree asset equals

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1}\right] = 1,$$

where $R_{f,t+1}$ is the return on the riskfree asset between time t and $t+1$. Note that $R_{f,t+1}$ is known at time t. Therefore $R_{f,t+1}$ equals a constant R_f with

$$\log R_f = -\log\delta + \gamma(\mu_c - \frac{1}{2}\sigma_c^2) - \frac{\gamma^2}{2}\sigma_c^2. \tag{A.3}$$

The risk premium equation, (6), follows from

$$\log E_t\left[R_{M,t+1}/R_f\right] = \log E_t\left[\Phi_C \frac{C_{t+1}}{C_t}\right] - \log R_f,$$

(A.3), and (A.1). The fact that the risk premium is constant justifies replacing the conditional expectation with the unconditional expectation.

Now consider an asset paying (4) as its dividend. Let P_{jt} be the ex-dividend price. Conjecture that P_{jt}/D_{jt} is a constant and define

$$\Phi_j \equiv \frac{P_{jt}/D_{jt} + 1}{P_{jt}/D_{jt}}$$

Arguments similar to above show that the price-dividend ratio is characterized by

$$\Phi_j^{-1} = E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \frac{D_{j,t+1}}{D_{jt}}\right] \tag{A.4}$$

provided again that the right hand side is less than 1. It follows that

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

Thus $\log R_{j,t+1}$ is normally distributed, and is perfectly correlated with $\log\frac{D_{j,t+1}}{D_{jt}}$. Moreover, they have equal variance. Because the same is true for $\log R_{M,t+1}$ and $\log C_{t+1}/C_t$, it follows that $\beta_j = \beta_j^d$ and $\epsilon_{jt} = \epsilon_{jt}^d$.

To show the equilibrium expected return of R_{jt}, note that

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{j,t+1}\right] = 1,$$

and therefore

$$-\log\delta - \gamma(\mu_C - \frac{1}{2}\sigma_C^2) + \gamma^2\frac{1}{2}\sigma_C^2 + \mu_j + \gamma\beta_j\sigma_M^2 = 1. \tag{A.5}$$

We have used the fact that consumption growth and market returns are perfectly correlated. Substituting in for (A.3) shows

$$\mu_j - \log R_f = \gamma\beta_j\sigma_M^2.$$

Note that this is a restatement of the Consumption CAPM of Breeden (1979). In the general case, σ_M is replaced by the standard deviation of consumption growth, and the β is with respect to consumption growth. The CAPM (5) follows from substituting in for γ using (6).

B Proofs for the binomial model

Proof of Proposition 1 It follows from (14) that

$$\begin{aligned}
D_{t+1} &= D_t + (D_{t+1} - D_t) \\
&= D_t + \left(\sum_{j=1}^{n} D_{j,t+1} - \sum_{j=1}^{n} D_{j,t}\right) \\
&= D_t + \sum_{j=1}^{n}(D_{j,t+1} - D_{j,t}) \\
&= D_t + \frac{1}{n}D_t e^z \sum_{j=1}^{n} X_{j,t+1}
\end{aligned}$$

Proposition 3. $\forall s \geq t, E_t[D_s] = D_t(1+\lambda e^z)^{s-t}$

Proof. Clearly the result holds for $s = t$. Assuming the result holds for $s - 1$,

$$E_t[D_{t+s}] = E_t[E_{t+s-1}[D_{t+s}]] = E_t[D_{t+s-1}](1 + \lambda e^z) = D_t(1 + \lambda e^z)^{t+s}$$

$\square$

Proof of Proposition 2

$$\begin{aligned}
P_{j,t} &= \sum_{s=t+1}^{\infty} R^{-(s-t)} D_{j,t} + \frac{1}{n}\lambda e^z \sum_{s=t+1}^{\infty} R^{-(s-t)} D_t + \frac{1}{n}\lambda e^z \sum_{s=t+2}^{\infty} R^{-(s-t)} E_t\left[D_{t+1}\right] \cdots \\
&= D_{j,t} \sum_{s=t+1}^{\infty} R^{-(s-t)} + \frac{1}{n} D_t \left(\sum_{s=t+1}^{\infty} R^{-(s-t)} (1 + \lambda e^z)^{s-t} - \sum_{s=t+1}^{\infty} R^{-(s-t)} \right) \\
&= D_{j,t} \frac{1}{R-1} + \frac{1}{n} D_t \left(\frac{1 + \lambda e^z}{R - (1 + \lambda e^z)} - \frac{1}{R-1} \right)
\end{aligned}$$

References

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.

Bessembinder, H. (2017). Do stocks outperform treasury bills? Working paper, Arizona State University.

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.

Bulan, L., Subramanian, N., and Tanlu, L. (2007). On the timing of dividend initiations. *Financial Management*, pages 31–65.

Campbell, J. Y. (2008). Viewpoint: Estimating the equity premium. *Canadian Journal of Economics*, 41(1):1–21.

Campbell, J. Y. (2018). *Financial Decisions and Markets: A Course in Asset Pricing*. Princeton University Press, Princeton, NJ.

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.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1:195–228.

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. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, 21(4):1533–1575.

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

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

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.

Epstein, L. and Zin, S. (1989). Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57:937–969.

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

Fama, E. F. and French, K. R. (2001). Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial economics*, 60(1):3–43.

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

Golez, B. and Koudijs, P. (2017). Four centuries of return predictability. forthcoming, *Journal of Financial Economics*.

Hansen, L. P., Heaton, J. C., and Li, N. (2008). Consumption strikes back? Measuring long run risk. *Journal of Political Economy*, 116(2):260–302.

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.

Jagannathan, R. and Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, 51:3–53.

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

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.

Mehra, R. and Prescott, E. (1985). The equity premium puzzle. *Journal of Monetary Economics*, 15:145–161.

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

Ross, S. A. (1976). The arbitrage theory of capital market asset pricing. *Journal of Economic Theory*, 13:341–360.

Santos, T. and Veronesi, P. (2010). Habit formation, the cross section of stock returns and the cash-flow risk puzzle. *Journal of Financial Economics*, 97(3):385–413.

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

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436.

Shiller, R. J. (1989). *Market Volatility*. MIT Press, Cambridge, MA.

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

van Binsbergen, J. H. and Koijen, R. S. J. (2010). Predictive regressions: A present-value approach. *Journal of Finance*, 65(4):1439–1471.

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 Warusawitharana, M. (2015). What is the chance that the equity premium varies over time? Evidence from regressions on the dividend-price ratio. *Journal of Econometrics*, 186(1):74–93.

Wachter, J. A. and Zhu, Y. (2017). The macroeconomic announcement premium. Rodney L. White Center Working Paper.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21:1455–1508.

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

Table 1: Selected Statistics on Pooled Monthly Level Returns

	All CRSP (1926 - 2016)	All CRSP (1945 - 2016)	16,087 Select (1926 - 2016)	404 Select (1973 - 2016)
Mean (in %)	1.118	1.297	1.315	1.331
Median (in %)	0.000	0.000	0.000	0.936
Std. Dev (in %)	17.83	16.82	16.95	10.27
Skewness	6.335	5.987	5.846	1.321
% Positive	48.40	48.91	48.94	54.51
% $\geq$ 1-Month T-Bill	47.75	48.19	48.26	53.07
% $\geq$ VW Mkt Return	46.34	46.71	46.71	50.11
% $\geq$ EQ Mkt Return	45.83	46.19	46.13	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 July 1926 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: Inference on Pooled Monthly Returns

	Empirical Value	Simulated Values Min	5th	50th	95th	Max	Simulation Population
Panel A. Simulation with 16,087 Firms - Method 1 (CAPM)							
$E[R]-1$	1.315	0.51	0.71	1.056	1.36	1.59	1.00
$\sigma[R]$	16.95	16.00	16.18	16.46	16.75	16.97	16.42
$skew[R]$	5.846	0.849	0.882	0.938	1.043	1.287	0.942
$skew[\log R]$	-0.196	-0.214	-0.192	-0.166	-0.142	-0.122	-0.164
$\% \log R > \log R_f$	48.94	48.51	49.18	50.07	50.94	51.51	50.01
Panel B. Simulation with 16,087 Firms - Method 2 (Direct Estimation)							
$E[R]-1$	1.315	0.810	1.002	1.311	1.645	1.859	1.290
$\sigma[R]$	16.95	16.05	16.24	16.51	16.80	17.02	16.48
$skew[R]$	5.846	0.826	0.857	0.917	1.009	1.180	0.916
$skew[\log R]$	-0.196	-0.245	-0.225	-0.191	-0.166	-0.146	-0.192
$\% \log R > \log R_f$	48.94	49.59	50.30	51.20	52.11	52.67	51.13
Panel C: Simulation with 404 Firms - Method 1 (CAPM)							
$E[R]-1$	1.331	0.376	0.567	0.889	1.204	1.491	0.867
$\sigma[R]$	10.27	9.98	10.06	10.19	10.32	10.46	10.19
$skew[R]$	1.321	0.396	0.429	0.465	0.502	0.561	0.464
$skew[\log R]$	-0.423	-0.100	-0.073	-0.042	-0.010	0.036	-0.043
$\% \log R > \log R_f$	54.51	49.69	50.58	51.90	53.21	54.51	51.82
Panel D: Simulation with 404 Firms - Method 2 (Direct Estimation)							
$E[R]-1$	1.331	0.814	1.006	1.329	1.645	1.934	1.307
$\sigma[R]$	10.27	10.03	10.11	10.23	10.37	10.50	10.24
$skew[R]$	1.321	0.395	0.428	0.463	0.501	0.559	0.462
$skew[\log R]$	-0.423	-0.102	-0.076	-0.044	-0.012	0.034	-0.045
$\% \log R > \log R_f$	54.51	51.65	52.55	53.87	55.20	56.47	53.80

Source: CRSP and simulations

Notes: We conduct 400 monthly simulations of the stock market for the two methods of estimating the drift parameter (μ_i). Sampling distribution of each statistic is obtained from the simulations. Panels A and B are relevant to simulations using 16,087 firms with at least 60 monthly returns from July 1926 to December 2016; Panels C and D pertain to the simulations using 404 firms without missing data for monthly returns from January 1973 to December 2016. The first column shows the statistic for the corresponding firms. 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.

Table 3: Inference on Monthly Cross-sectional Skew

	Empirical $\bar{\gamma}_{cs}$	$\bar{\gamma}_{cs}$ from Simulated Values					% of Months with Empirical γ_{cs} $\geq$ Max
		Min	5th	50th	95th	Max	
Panel A. Simulation with 16,087 Firms (1926.07 - 2016.12)							
$skew[R]$ (Method 1)	2.381	0.804	0.844	0.894	0.965	1.044	69.15
$skew[R]$ (Method 2)	2.381	0.766	0.816	0.871	0.934	1.015	70.26
Panel B. Simulation with 404 Firms (1973.01 - 2016.12)							
$skew[R]$ (Method 1)	0.961	0.348	0.374	0.400	0.426	0.453	86.19
$skew[R]$ (Method 2)	0.961	0.349	0.375	0.401	0.427	0.454	86.19

Source: CRSP and simulations

Notes: We conduct 400 monthly simulations of the stock market for the two methods of estimating the drift parameter (μ_i). The first column shows the average monthly cross-sectional skewness ($\bar{\gamma}_{cs}$) for the corresponding universe of stocks and sample period. The next five columns illustrate the distribution of $\bar{\gamma}_{cs}$ obtained from simulations. The final column reports the percentage of months in the sample period in which empirical γ_{cs} is greater than the maximum $\bar{\gamma}_{cs}$ obtained from the simulations.

Table 4: 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 16,087 Firms (1926.07 - 2016.12)						
Method 1 (Baseline)						
Market Cap Growth	29.40	56.46	89.14	99.62	99.96	99.99
Cumulative Return	71.67	56.46	89.14	99.62	99.96	99.99
Method 2 (Direct Estimation)						
Market Cap Growth	29.40	99.11	99.94	99.99	99.99	99.99
Cumulative Return	71.67	99.11	99.94	99.99	99.99	99.99
Method 3 (Baseline with Monthly Dividends)						
Market Cap Growth	29.40	56.55	89.10	99.61	99.99	99.99
Cumulative Return	71.67	56.46	89.14	99.62	99.96	99.99
Method 4 (Baseline with Time-varying Volatility)						
Market Cap Growth	29.40	45.95	76.39	98.98	99.99	99.99
Cumulative Return	71.67	45.95	76.39	98.98	99.99	99.99
Panel B: Simulation with 404 Firms (1973.01 - 2016.12)						
Method 1 (Baseline)						
Market Cap Growth	41.13	25.97	31.66	47.16	81.53	96.36
Cumulative Return	29.97	25.97	31.66	47.16	81.53	96.36
Method 2 (Direct Estimation)						
Market Cap Growth	41.33	40.63	48.71	69.87	93.91	99.20
Cumulative Return	29.97	40.63	48.71	69.87	93.91	99.20
Method 3 (Baseline with Monthly Dividends)						
Market Cap Growth	41.33	25.90	31.57	47.19	81.51	96.40
Cumulative Return	29.97	25.97	31.66	47.16	81.53	96.36
Method 4 (Baseline with Time-varying Volatility)						
Market Cap Growth	41.33	15.92	21.62	23.86	32.35	38.60
Cumulative Return	29.97	15.92	21.62	23.86	32.35	38.60

Source: CRSP and simulations

Notes: We conduct 400 monthly simulations of the stock market for both methods of estimating the drift parameter (μ_i). We also consider simulations with monthly dividends and time-varying volatility. 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 value alternatively as market capitalization growth and cumulative return. 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.

Table 5: Deciles for Stocks Formed on Idiosyncratic Volatility

Decile	Average Idiosyncratic Volatility (Monthly)	
	16,087 Select (1926 - 2016)	404 Select (1973 - 2016)
1	0.285	0.143
2	0.219	0.112
3	0.189	0.097
4	0.166	0.088
5	0.146	0.082
6	0.128	0.075
7	0.111	0.069
8	0.095	0.064
9	0.079	0.058
10	0.060	0.051

Source: CRSP

Notes: The table reports the deciles for stocks formed on monthly idiosyncratic volatility ($\sigma_{j,\epsilon}$) and its equal-weighted average within each decile. The second column pertains to the set of 16,087 firms with at least 60 monthly returns from July 1926 to December 2016, the values of which are calculated across the entire universe. In the simulations, the volatility parameters are estimated only from the 2,440 firms that constitute each sample. The third column pertains to the set of 404 firms without missing data for monthly returns from January 1973 to December 2016.

Table 6: Comparison of Selected Features in Data and Models

	Data	Model (Lognormal)	Model (Binomial)
Standard Deviation of Monthly Returns ($\sigma[R]$)	16.95	16.46	2.29
Skewness of Monthly Returns (skew$[R]$)	5.85	0.95	5.75
% of Market Cap. Growth By Top Ten Firms	29.40	99.63	2.62
Stationary Distribution of Market Cap.?	Unknown	No	Yes
Inverse Relationship in Firm Volatility and Size?	Yes	No	Yes

Source: CRSP and simulations

Notes: The table compares selected features in the data and the model. The first column examines monthly returns of all CRSP common stocks from July 1926 to December 2016. The second column pertains to simulations using 16,087 firms from the lognormal model and the CAPM. The third column pertains to simulations from the binomial model, whose parameters were calibrated to match the skewness in the data. The calibrated parameters are reported in Table 7. For both models, the reported statistics are median values obtained from 400 simulation samples.

Table 7: Parameters for Binomial Model

Parameter	Value
R	1.30
λ	0.3
λe^z	0.25
N	500
T	1200

Notes: The table shows the parameters used in the binomial model (see Table 6), which were calibrated to match the skewness in the data. Reported values of $R, \lambda,$ and λe^z have been annualized.

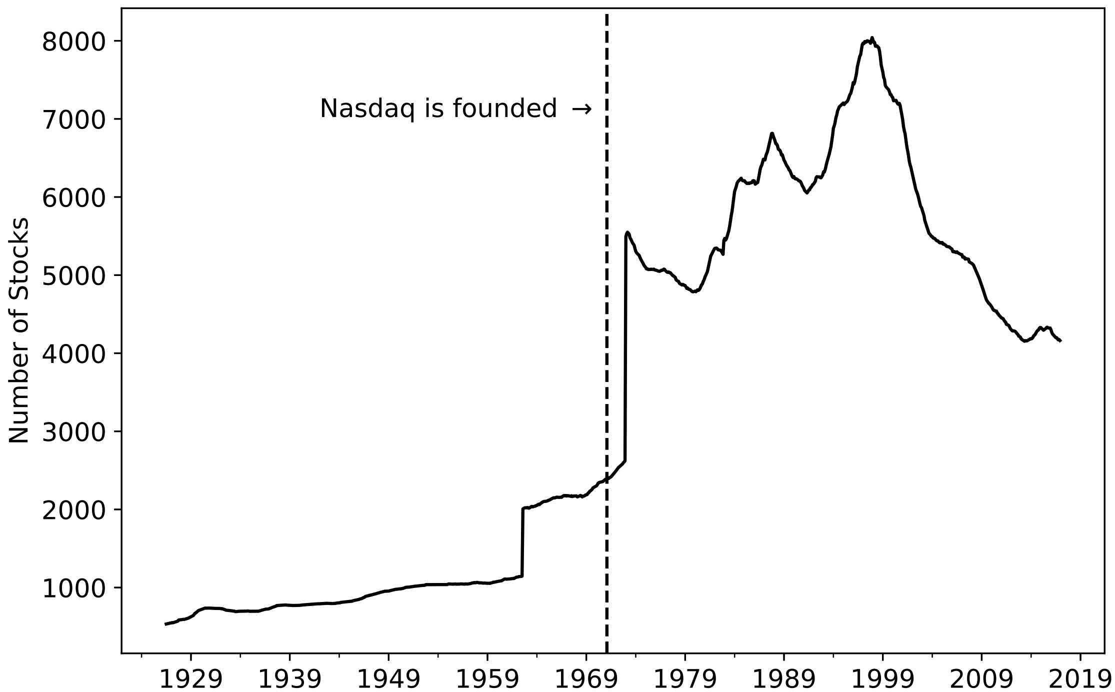

Figure 1: **Historical Number of CRSP Common Stocks**
On the first day of each month from July 1926 to December 2016, we count the number of unique common stocks in the cross-section, as available in CRSP. The jump on January 1973, from 2,623 to 5,494, roughly corresponds to the establishment of Nasdaq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

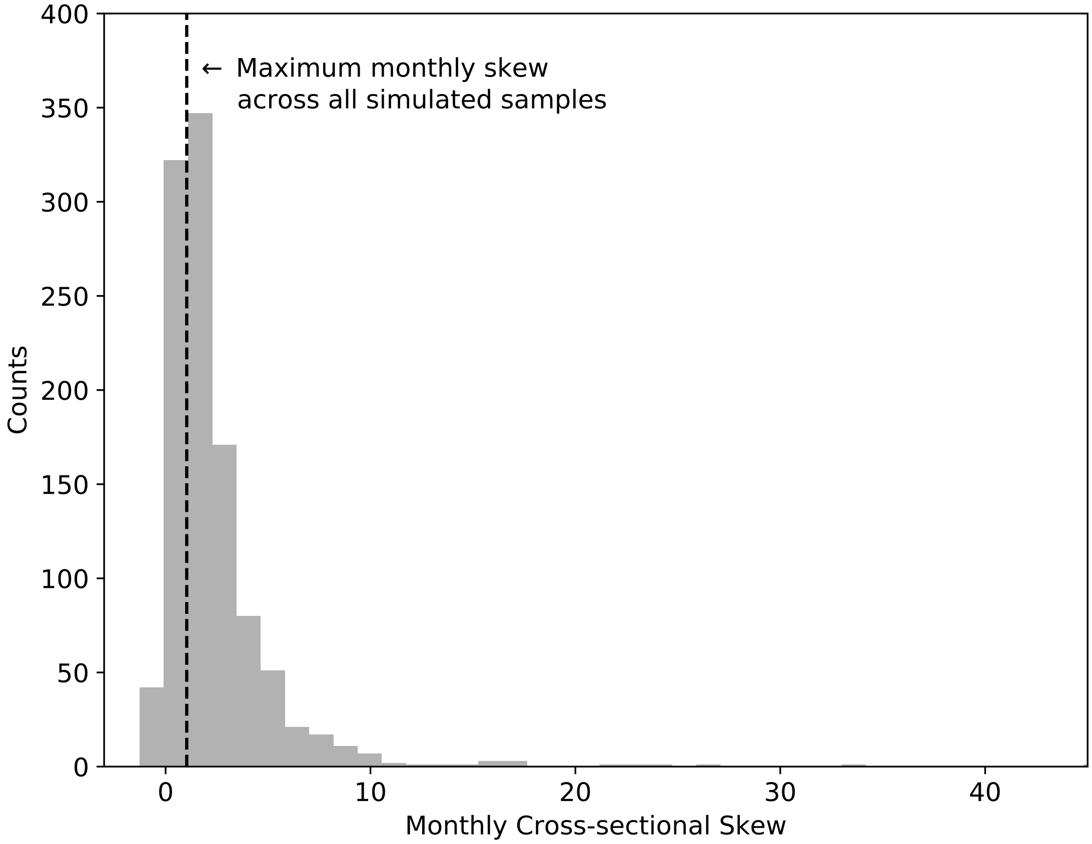

Figure 2: **Distribution of Monthly Cross-sectional Skewness (16,087 Firms)**
The figure illustrates the distribution of monthly cross-sectional skewness, defined as the skewness of monthly level returns for the cross-section of firms in each given month. The graph pertains to set of 16,087 firms with at least 60 monthly returns from July 1926 to December 2016. The vertical line on the graph represents the maximum of the average monthly cross-sectional skewness obtained from the 400 simulations.

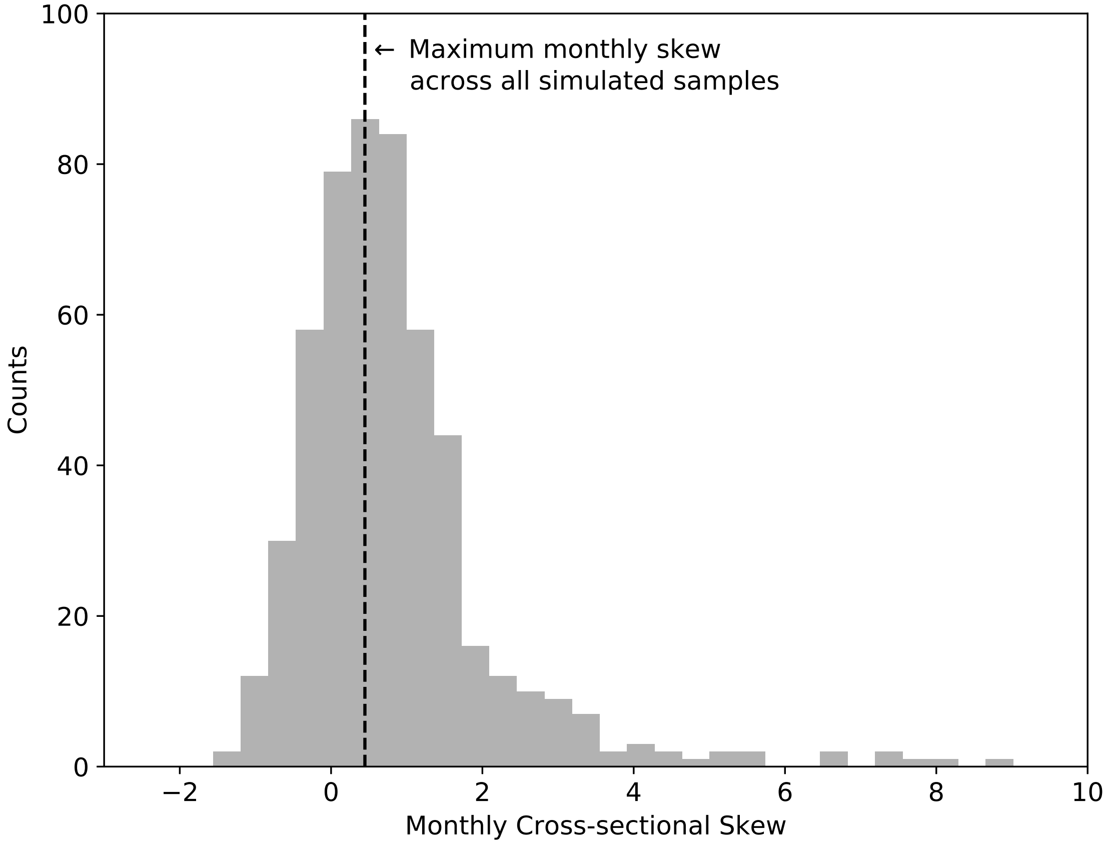

Figure 3: **Distribution of Monthly Cross-sectional Skewness (404 Firms)**
The figure illustrates the distribution of monthly cross-sectional skewness, defined as the skewness of monthly level returns for the cross-section of firms in each given month. The graph pertains to set of 404 firms without missing 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.

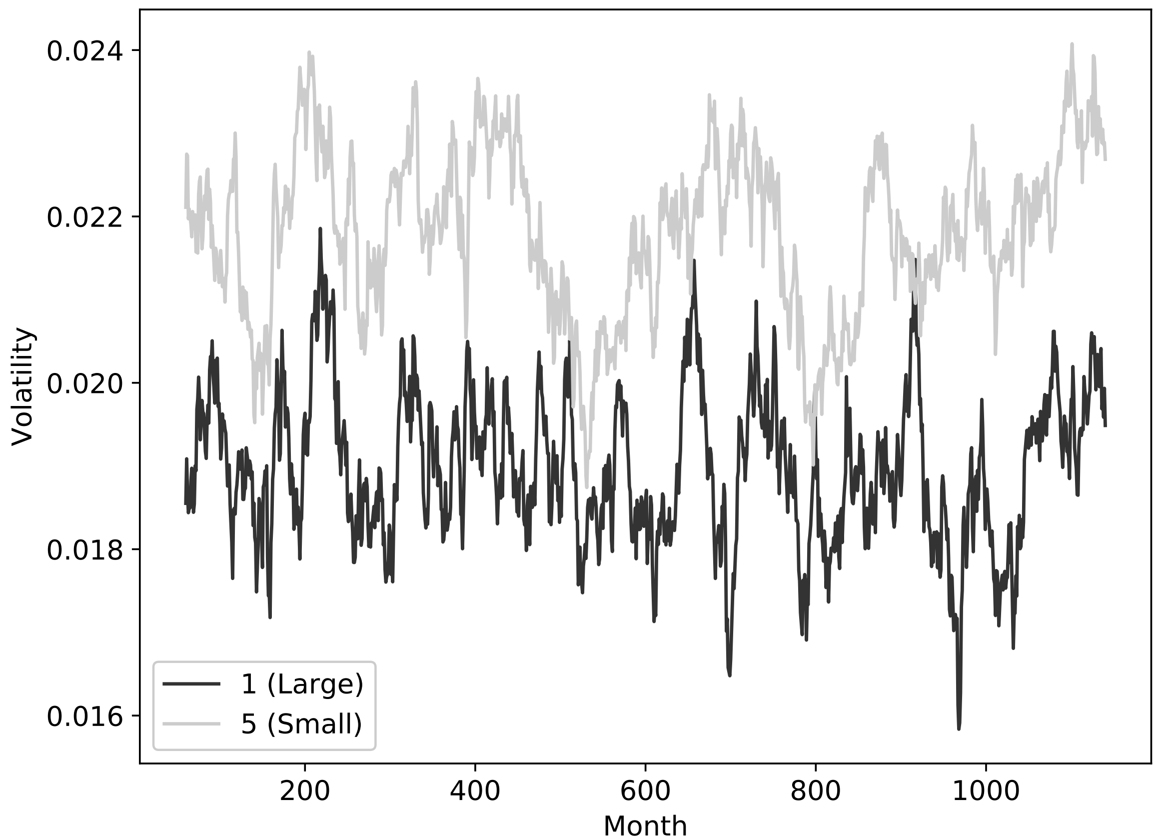

Figure 4: **Total Volatility by Size Group (Binomial Model)**
The figure plots monthly firm-level total volatility averaged within the top and bottom market capitalization quintiles. The returns are simulated from the binomial model with 500 fictitious firms and for 1,200 months, using the calibration in Table 7. At each month, volatilities are estimated as the standard deviation of monthly returns over the next 60 months.