PRELIMINARY
PLEASE DO NOT DISTRIBUTE

Estimating the Moments of Long Horizon Returns

Anthony Neuberger, Richard Payne

Cass Business School,
City, University of London

March 2017

Abstract

Moments of long-horizon returns are important for asset pricing but are hard to measure. Proxies for these moments are often used but none is wholly satisfactory. We show analytically that high-frequency returns can be used to make more much precise estimates for long-horizon moments, under rather general conditions. Skewness comprises two components: the skewness of short-horizon returns, and a leverage effect, the covariance between contemporaneous variance and lagged returns. The kurtosis of long-horizon returns comprises three components: kurtosis in short-horizon returns, the covariance between cubed short-horizon returns and lagged returns, and the covariance between squared short-horizon returns and lagged squared returns (which we call the GARCH effect). Applying the framework to US stock index returns, we show that variation in monthly moments is essentially driven by the leverage effect and the GARCH effect while daily return skewness and kurtosis are largely irrelevant. We also show that monthly skewness has power to forecast US index returns at monthly and annual horizons.

1. Introduction

There is good reason to believe that higher moments of returns – not just second moments – are important for asset pricing. A large theoretical literature, starting with Kraus and Litzenberger (1976), and continuing with the macroeconomic disaster research of Rietz (1988), Longstaff and Piazzesi (2004), and Barro (2006), hypothesises that heavy-tailed shocks and left-tail events in particular have an important role in explaining asset price behaviour. Barberis and Huang (2007) and Mitton and Vorkink (2007) argue that investors look for idiosyncratic skewness, seeking assets with lotto-type pay-offs. There is much empirical evidence suggesting that market skewness is time varying, and that it predicts future returns in both the time series (Kelly and Jiang, 2014) and in the cross-section (Harvey and Siddique, 2000, and Ang, Hodrick, Xing and Zhang, 2006). Boyer, Mitton and Vorkink (2010) and Conrad, Dittmar and Ghysels (2013) show that high idiosyncratic skewness in individual stocks too is correlated with positive returns. Ghysels, Plazzi and Valkanov (2016) show similar results for emerging market indices.

But there are two serious problems in measuring these moments at the long horizons (months, years) of interest to asset pricing. First, the higher the moment, the more sensitive the estimate is to outliers. Second, the longer the horizon, the smaller the number of independent observations. We show how these problems can be mitigated by using the information in high frequency returns to estimate the skewness and kurtosis of long horizon returns with greater precision.

As is well known, in the case of the second moment the use of high frequency information can help one to overcome the problems associated with a limited data span. Under the assumption that the price process is martingale, the annualized variance of returns is independent of the sampling frequency, so one can use the high frequency returns to estimate the variance of long horizon returns. But this simple result does not hold for higher moments - there is no necessary relationship between the higher moments of long and short horizon returns. If daily returns are volatile, then annual returns are also volatile. But if daily returns are highly skewed and i.i.d., then annual returns will show little skew. Conversely, daily return distributions can be symmetric, while annual returns are skewed (e.g. in a Heston-type model where volatility is stochastic and shocks to volatility are correlated with shocks to prices). Similar examples could be given for kurtosis.

The purpose of this paper is to demonstrate how to exploit the information in high frequency returns to estimate skewness and kurtosis of long horizon returns. The only assumption we make about the price process is that it is martingale, and that the relevant moments exist. We prove that the skewness of long horizon returns can come from one of only two sources: the skewness of short horizon returns; and what has been called the leverage affect, i.e. the covariance between lagged returns and squared returns. Similarly, the kurtosis of long horizon returns has just three possible sources: the kurtosis of short horizon returns; the covariance between cubed returns and lagged returns; and the covariance between squared returns and lagged squared returns (which we refer to as the GARCH effect). When we take these theoretical results to the data, we show that the skewness of the US stock market at long horizons is due almost entirely to the leverage effect, and its kurtosis to the GARCH effect. Thus, the left-skew and excess kurtosis in annual stock market returns owe virtually nothing to the skew and kurtosis of daily returns.

To date, the literature has used a variety of approaches to measure the higher moments of long horizon returns. The most straightforward is to apply the standard estimators to historic returns. Kim and White (2004) show that these estimators are subject to large estimation errors[1] and advocate the use of robust estimators such as those developed by Bowley (1920), which are based on quantiles of the observed distribution. The attraction is that quantiles can be estimated with much greater precision than moments. This solution is used in Conrad, Dittmar and Ghysels (2013) and the methodology is further developed in Ghysels, Plazzi and Valkanov (2016). The weakness of the approach is that it assumes that the body of the distribution, which is captured by the quantiles, is highly informative about the behaviour of the tails, which determine the higher moments.

Kelly and Jiang (2014) follow an alternative approach. They focus on the tails. They get power not by taking a very long time series, but rather by exploiting the information in the cross-section. They assume that tail risk for individual stocks is a combination of stable stock specific tail risk and time-varying market-wide tail risk. They can therefore exploit the existence of a large number of stocks to get a much more precise estimate of market-wide tail risk. The validity of the inference depends not only on the assumed decomposition of the tail component, but also on assumptions about the dependence of returns across stocks.The options market is an attractive source of information about moments since, unlike the underlying market which shows just one realization of the price process, the options market reveals the entire implied density of returns at any point in time. The technology for extracting implied skewness and kurtosis from options prices is well-established (Bakshi, Kapadia and Madan, 2003). The method can only be used on assets – such as the major market indices - that support a liquid options market, and cannot be used for managed portfolios. But there is a more fundamental issue: implied measures reflect risk premia as well as objective probabilities. As demonstrated by Broadie, Chernov and Johannes (2007), the wedge between the objective price process and the process as implied by option prices (the so called risk neutral process) can be very wide.

We show by simulation that our measures of skewness and kurtosis are indeed substantially more powerful than standard estimators, reducing standard errors on skewness by around 60% and on kurtosis by around 30%. We apply our technology to the US stock market over the last 90 years, and show that monthly skewness has averaged -0.7, and monthly excess kurtosis has average +1.3, with substantial time variation. We also show that monthly skew and kurtosis owe virtually nothing to their daily counterparts; rather monthly skew is due almost entirely to

[1] To estimate the skewness (kurtosis) of a normally distributed random variable with a standard error of 0.1 requires a sample size of 600 (2400). Even for monthly returns, this would require 50 (200) years of returns data. If returns are non-normal, the standard errors are generally substantially higher. Monthly returns on the US market over the last 50 years have a skew coefficient of -0.98; the bootstrapped standard error is 0.3.

the Leverage effect, and monthly kurtosis to the GARCH effect. Finally, we show that monthly skew, when used in conjunction with the tail risk measure of Kelly and Jiang (2014) predicts returns, with more negative skew forecasting more positive returns.

The rest of the paper proceeds as follows. In Section 2 we develop the theoretical relationship between low frequency skewness and kurtosis and their high-frequency counterparts. In Section 3 we demonstrate the power of the technique through simulation. Section 4 provides an empirical application to the US stock market. Section 5 concludes.

2. THE THEORY

2.1 Moments of price changes

We work in a discrete time setting, $t \in \mathrm{Z}$. The asset has discounted price P_t ("the price").We are concerned with the distribution of returns from time t to $t+T$. For brevity, we refer to the time increment as a day, and the long horizon as a month, but obviously nothing hangs on this. The term kurtosis is used specifically for excess kurtosis.

The problem we are interested in is:

> [**P**]: Let $P := \{P_t | t = ...,0,1,...\}$ be a strictly positive martingale process, whose associated returns process r, where $r_t := P_t/P_{t-1}$, is strongly stationary. The long horizon returns process R is defined by $R_t := P_t/P_{t-T}$. How can one estimate higher moments of long horizon returns R efficiently, assuming that these moments exist?

Problem **P** is difficult because it deals with returns (ratios) rather than with price changes (differences). We therefore first address a simpler problem, **P***, and use the solution as a guide to solving **P**.

The simpler problem is:

> [**P***]: Let $P := \{P_t | t = ...,0,1,...\}$ be a real-valued (not necessarily positive) martingale process whose associated *difference* process d, where $d_t := P_t - P_{t-1}$, is strongly stationary. The long horizon difference process D is defined by $D_t := P_t - P_{t-T}$. How can one estimate higher moments of D efficiently, assuming that these moments exist?

The solution to **P*** is given by

Proposition 1

The volatility, skewness and kurtosis of monthly price changes is related to the distribution of daily price changes in the following way

$$
\begin{aligned}
&\operatorname{vol}\left[D_t\right]=\operatorname{vol}\left[d_t\right]; \\
&\operatorname{skew}\left[D_t\right]=\left(\operatorname{skew}\left[d_t\right]+3\frac{\operatorname{cov}\left[y_{t-1}^*,d_t^2\right]}{\operatorname{var}\left[d_t\right]^{3/2}}\right)T^{-1/2}; \\
&\operatorname{kurt}\left[D_t\right]=\left(\operatorname{kurt}\left[d_t\right]+4\frac{\operatorname{cov}\left[y_{t-1}^*,d_t^3\right]}{\operatorname{var}\left[d_t\right]^2}+6\frac{\operatorname{cov}\left[z_{t-1}^*,d_t^2\right]}{\operatorname{var}\left[d_t\right]^2}\right)T^{-1};
\end{aligned}
\tag{1}
$$

where

$$
\begin{aligned}
y_{t-1}^* &:= \sum_{u=1}^{T}\left(P_{t-1}-P_{t-u}\right)/T; \quad \text{and} \\
z_{t-1}^* &:= \sum_{u=1}^{T}\left(P_{t-1}-P_{t-u}\right)^2/T.
\end{aligned}
\tag{2}
$$

Proof: the full proof is in the Appendix.

Proposition 1 gives expressions for the volatility (the square root of the variance rate), the skewness and the excess kurtosis of monthly price changes. The first result is familiar: the volatility of price changes is the same whether computed from monthly or daily data. The second result says that skew at the monthly horizon has just two sources: daily skew and a term we call leverage. Daily skew attenuates with horizon with the square root of time. The leverage term is proportional to the covariance between squared price changes and the quantity y^*, which is equal to the difference between the opening price on the day and the average price over the last month.

The final result says that the kurtosis of monthly returns has just three sources: daily kurtosis attenuating with time, the covariance between cubed price changes and y^*, and the covariance between squared price changes and z^*. z^* is a measure of the average squared price change over the last month.

In order to demonstrate the logic underlying Proposition 1 (and indeed the main result in this paper, Proposition 2) and the role of the assumptions we make (martingale, strict stationarity), it is useful to review the proof of one of the elements of the proposition, that concerning skewness.

Start with an algebraic decomposition of the third power of the monthly price change

$$
D_t^3=\sum_{u=0}^{T-1}d_{t-u}^3+3\sum_{u=0}^{T-1}\left(P_{t-u-1}-P_{t-T}\right)d_{t-u}^2+3\sum_{u=0}^{T-1}\left(P_{t-u-1}-P_{t-T}\right)^2 d_{t-u}. \tag{3}
$$

Taking conditional expectations of both sides, the third term drops out because of the martingale assumption[2], so

$$\mathrm{E}_{t-T}\left[D_t^3\right]=\sum_{u=0}^{T-1}\mathrm{E}_{t-T}\left[d_{t-u}^3\right]+3\sum_{u=0}^{T-1}\mathrm{E}_{t-T}\left[\left(P_{t-u-1}-P_{t-T}\right)d_{t-u}^2\right]. \tag{4}$$

Define

$$y_{t-1}^* := \sum_{u=1}^{T}\left(P_{t-1}-P_{t-u}\right)/T. \tag{5}$$

y_{t-1}^* is the difference between today's opening price and the T-day moving average[3]. Using strict stationarity, the conditional expectations can be replaced by unconditional expectations. Substituting y_t^* in to (4) gives the following expression for the unconditional third moment

$$\mathrm{E}\left[D_t^3\right]=T\mathrm{E}\left[d_t^3\right]+3T\mathrm{E}\left[y_{t-1}^* d_t^2\right]. \tag{6}$$

y_{t-1}^* is mean zero, so the expectation can be replaced by the covariance, giving

$$\mathrm{E}\left[D_t^3\right]=T\left(\mathrm{E}\left[d_t^3\right]+3\,\mathrm{cov}\left(y_{t-1}^*,d_t^2\right)\right). \tag{7}$$

A similar argument shows that

$$\mathrm{E}\left[D_t^2\right]=T\mathrm{E}\left[d_t^2\right]. \tag{8}$$

The result in proposition 1 then follows immediately from the definition of the skewness coefficient.

2.2 Unconditional Moments of Returns

The objective is to produce a result akin to Proposition 1, one that applies to moments of returns rather than to price changes. We now work with daily returns, $r_t = P_t / P_{t-1}$, and monthly returns, $R_t = P_t / P_{t-T}$. The problem is intractable if we stay with the standard definitions of moments. It is necessary to modify the definition of moments.

Define

[2] If the price process were not martingale, there would be an additional term in the skew, the covariance between price changes and past volatility. But there is reason to believe that any such term would be small, at least in the case of the equity market. As Bollerselv *et al* (2013, p210) say: "The most striking empirical regularities to emerge from this burgeon literature are that …returns are at best weakly positively related, and sometimes even negatively related, to past volatilities."

[3] The asterisk is used to link this variable with the corresponding variable in the problem **P**.

$$\text{var}^L[r] := \text{E}\left[x^{(2,L)}(r)\right] \text{ where } x^{(2,L)}(r) := 2(r - 1 - \ln r);$$
$$\text{var}^E[r] := \text{E}\left[x^{(2,E)}(r)\right] \text{ where } x^{(2,E)}(r) := 2(r\ln r + 1 - r);$$
$$\text{skew}[r] := \frac{\text{E}\left[x^{(3)}(r)\right]}{\text{var}^L[r]^{3/2}} \text{ where } x^{(3)}(r) := 6\left((r+1)\ln r - 2(r-1)\right); \tag{9}$$
$$\text{kurt}[r] := \frac{\text{E}\left[x^{(4)}(r)\right]}{\text{var}^L[r]^2} - 3 \text{ where } x^{(4)}(r) := 12\left((\ln r)^2 + 2(r+2)\ln r - 6(r-1)\right);.$$

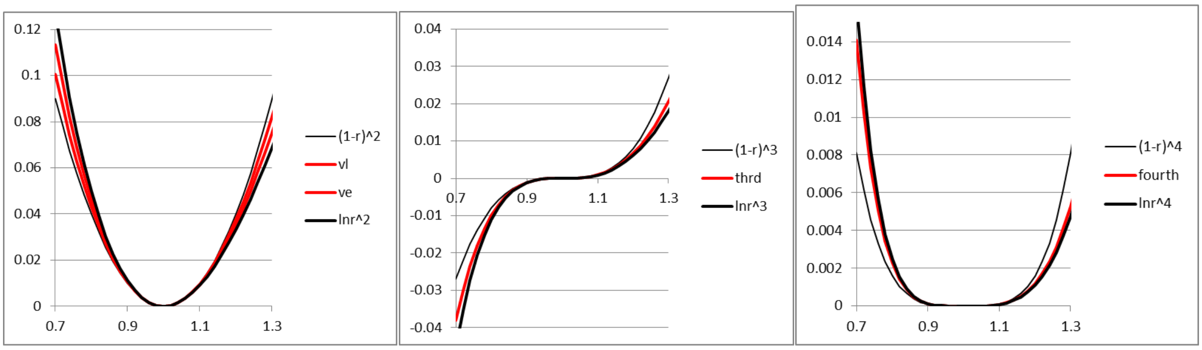
$x^{(2,L)}$ approximates the second power of log returns, as does $x^{(2,E)}$. Similarly, $x^{(3)}$ and $x^{(4)}$ approximate the third and fourth powers, as can be seen by doing a Taylor expansion, and is shown graphically in Figure 1. Modifying the definitions of moments in this way is not unprecedented. The Model Free Implied Variance that is widely used by both academics and practitioners. It is defined as

$$MFIV[R] \coloneqq \text{E}^{\text{Q}}\left[x^{(2,L)}(R)\right], \tag{10}$$

where the Q superscript denotes that the expectation is under the pricing measure. It also follows the definition of realized variance in Bondarenko (2014). A definition of skewness similar to the above is seen in Neuberger (2012).

We also define volatility of the return (vol[r]) as the square root of the variance rate.

Figure 1



With these definitions, we can now state the main theoretical result of this paper

Proposition 2

If P is a strongly stationary martingale process, the volatility, skewness and kurtosis of monthly returns (as defined in (9)) is related to the distribution of daily returns as follows

$$
\begin{aligned}
&\operatorname{var}^L\left[R_t\right]=T\operatorname{var}^L\left[r_t\right];\\
&\operatorname{skew}\left[R_t\right]=\left(\operatorname{skew}\left[r_t\right]+3\frac{\operatorname{cov}\left[y_{t-1},x^{(2,E)}\left(r_t\right)\right]}{\operatorname{var}^L\left[r_t\right]^{3/2}}\right)T^{-1/2};\\
&\operatorname{kurt}\left[R_t\right]=\left(\operatorname{kurt}\left[r_t\right]+4\frac{\operatorname{cov}\left[y_{t-1},x^{(3)}\left(r_t\right)\right]}{\operatorname{var}^L\left[r_t\right]^{2}}+6\frac{\operatorname{cov}\left[z_{t-1},x^{(2,L)}\left(r_t\right)\right]}{\operatorname{var}^L\left[r_t\right]^{2}}\right)T^{-1};
\end{aligned}
\tag{11}
$$

where

$$
\begin{aligned}
&y_{t-1}:=\sum_{u=1}^{T}\left(\frac{P_{t-1}}{P_{t-u}}-1\right)\Big/T \quad \text{and}\\
&z_{t-1}:=\sum_{u=1}^{T}2\left(\frac{P_{t-1}}{P_{t-u}}-1-\ln\left(\frac{P_{t-1}}{P_{t-u}}\right)\right)\Big/T.
\end{aligned}
\tag{12}
$$

Proof: the proof is similar to that of Proposition 1; details in the Appendix.

This is virtually identical to Proposition 1. It can be seen that

- Volatility of monthly returns is identical to the volatility of daily returns.
- The skew in daily returns generates a much smaller $\left(1/\sqrt{T}\right)$ skew in monthly returns.
- If monthly returns have significant skew, it must be through the leverage effect, the correlation between volatility and past returns. Past returns are measured by y, which is the net return relative to the one month moving average[4].
- Kurtosis in daily returns generates a much smaller ($1/T$) kurtosis in monthly returns.
- If annual returns are significantly leptokurtic, it is for one of two reasons:
  - because daily skew is correlated with past returns (as measured again by y);
  - or because of a GARCH effect whereby current variance is correlated with past variance. Past variance is measured by z, which is a function of the average realized variance over horizons of up to one month, again with more recent experience having more weight.

The results are quite general; there is no presumption about any functional form for the stochastic process driving the price. In a Merton (1976) jump-diffusion model, the asymmetric jump creates skewness and kurtosis. The absence of any covariation between volatility and lagged returns and lagged squared returns (volatility is constant) ensures that there is no leverage or GARCH effect, so skewness and kurtosis attenuate rapidly with the horizon. A Heston (1993) model has no skewness or conditional kurtosis in short horizon returns, but does generate a leverage effect and hence skewness in longer horizon returns because correlation between innovations in returns and innovations in volatility, coupled with the persistence of volatility, creates a correlation between volatility and lagged returns. It also generates kurtosis

[4] The moving average in this case is the rolling harmonic mean.

because the volatility of volatility and the persistence of volatility shocks creates a correlation between current volatility and lagged volatility. GARCH processes generate heteroscedasticity in long horizon returns through the persistence of volatility shocks. To generate skewness the various GARCH variants (NGARCH, QGARCH etc) have volatility reacting asymmetrically to positive and negative return shocks; as with Heston, this creates a correlation between volatility and lagged returns.

2.3 Conditional moments

Proposition 2 applies to unconditional moments. There are strong grounds for believing that moments are time varying. In this section we show how to use high frequency returns to make improved estimates of conditional moments. The results are somewhat less neat than before, but there is an offsetting advantage. We no longer need to assume that the price process is strictly stationarity – indeed, the unconditional moments need not necessarily exist.

The concept of realized moments is of central importance. In the case of the second moment, the natural definition of realized (daily) variance over some period (t, w) is

$$rx_{t,w}^{(2,L)} := \sum_{s=t+1}^{w} x^{(2,L)}\left(r_s\right). \tag{13}$$

This corresponds to the standard definition, except that we use $x^{(2L)}(r)$ in place of $(\ln r)^2$. We have already shown, in the proof of Proposition 2, that the realized variance $rx_{t,t+T}^{(2L)}$ over the month (t, $t+T$) is an unbiased estimator of the conditional variance of monthly returns at the beginning of the month, $\mathrm{var}_t^L\left[R_{t+T}\right]$. It follows that

$$rx_{t,t+T}^{(2,L)} = \mathrm{var}_t^L\left(R_{t+T}\right) + \varepsilon_{t+T}. \tag{14}$$

ε_{t+T} is an estimation error with zero mean.

With moments varying over time, the statistic of interest is the average moment over some period (t, w)

$$\mathrm{var}_{(t,w)}^L\left(R\right) := \frac{\sum_{u=t}^{w-1} \mathrm{var}_u^L\left(R_{u+T}\right)}{w-t}. \tag{15}$$

Using (14), we can construct an unbiased estimator of the average monthly conditional variance over the period as

$$\frac{\sum_{u=t}^{w-1} rx_{u,u+T}^{(2,L)}}{w-t} = T\frac{rx_{t,w}^{(2,L)}}{w-t} + \phi \text{ where } \phi = \frac{\sum_{u=1}^{T-1} u\left(x_{w+T-u}^{(2,L)} - x_{t+T-u}^{(2,L)}\right)}{w-t}. \tag{16}$$

The estimated monthly variance over a period is composed of two terms. The first is the average realized daily variance expressed as a monthly rate. The second term is a boundary or edge effect, adding in realized daily variances in the month at the end of the period and subtracting comparable terms from the first month. With a strictly stationary process this boundary effect would be mean zero. We do not assume stationarity but will assume that its impact is second order. When we say "ignoring boundary effects", we mean dropping such terms.

Define the realized third and fourth moments

$$\begin{aligned} rx_{t,w}^{(3)} &:= \sum_{s=t+1}^{w} x^{(3)}\left(r_s\right) + 3y_{s-1}x^{(2,E)}\left(r_s\right); \\ rx_{t,w}^{(4)} &:= \sum_{s=t+1}^{w} x^{(4)}\left(r_s\right) + 4y_{s-1}x^{(3)}\left(r_s\right) + 6z_{s-1}x^{(2,L)}\left(r_s\right). \end{aligned} \tag{17}$$

Proposition 3

If P is a martingale process, the average conditional second, third and fourth moments of monthly returns over a period (t, w) can be estimated from the corresponding realized moments of daily returns (as defined by (14) and (17)), multiplying them by $T/(w-t)$. Ignoring boundary effects, the estimators are unbiased.

Proof: substantially the same as for Proposition 2.

Although T, the length of the "month", only appears explicitly in proposition 3 as a scaling factor, it is implicit in the definitions of y and z – the monthly third and fourth moments depend on the covariances of daily variance and skewness with returns and variances over the previous month.

In the empirical work, we will generally normalize the third and fourth conditional moments by the conditional variance, and refer to the results as realized skewness and kurtosis. However, these are not unbiased estimators of skewness and kurtosis unless the skewness and kurtosis coefficients are uncorrelated with variance.

3. Simulation Results

3.1 Results for variance, skewness and kurtosis

We now evaluate the performance of our estimators of higher moments through a series of simulation experiments. We compare the estimators both with standard estimators and with quantile-based estimators.

Returns are simulated from three different models; a geometric Brownian motion (GBM), a Heston model and an EGARCH specification. For each model we simulate 10,000 paths for daily returns, each of length 5000 (i.e. roughly 20 years of return data in each path).

The parameters for each model are derived from fitting them to recent spans of daily US stock market returns. For the GBM and the Heston model, the parameters are taken from Eraker (2004). Those estimations use daily S&P-500 returns from January 2nd 1980 to December 31st 1999. The EGARCH parameters are obtained from our own fit of such a model to daily value-weighted CRSP US stock returns covering the period from January 2nd 1980 to the end of December 2015.

Given the parameters for a particular model, the objects that we wish to measure are the standard deviation, skewness and kurtosis of 25-day (i.e. roughly monthly) returns, where these are as defined in equation (9). We employ three estimation techniques for each moment. First, we construct the sample moments of non-overlapping 25-day returns (and we refer to these subsequently as 'Monthly' estimates). Second, we measure the sample moments using overlapping 25-day returns (referred to later as 'Overlapping' estimates).[5] Finally, we implement the estimators from Proposition 2 (which we label 'NP').

Results from these simulations are given in Tables 1 to 3 and Figures 2 to 4. Table 1 and Figure 2 show the simulation results when returns are generated by a GBM. Table 2 and Figure 3 give simulation results for the Heston model. Finally, Table 3 and Figure 4 show the EGARCH results. Each table gives statistics on the distribution of estimates from all three estimation techniques and for each of the three moments from across the 10,000 sample paths. The figures show the histograms of estimates from the 10,000 paths. In the discussion below, we focus on skewness and kurtosis estimates.

Table 1: simulation results for Geometric Brownian Motion

Standard deviation			
	NP	Monthly	Overlap
Mean	0.0469	0.0469	0.0469
STDEV	0.0005	0.0021	0.0019
5th prctile	0.0461	0.0434	0.0438
95th prctile	0.0477	0.0504	0.0501
Coefficient of Skewness			
	NP	Monthly	Overlap
Mean	-0.0053	-0.0043	-0.0047
STDEV	0.0351	0.2480	0.2420
5th prctile	-0.0625	-0.4084	-0.4015
95th prctile	0.0533	0.4120	0.3957
Excess Kurtosis			
	NP	Monthly	Overlap
Mean	-0.0023	-0.0248	-0.0211
STDEV	0.0714	0.3103	0.2164
5th prctile	-0.1121	-0.4600	-0.3248
95th prctile	0.1216	0.5323	0.3629

[5] So for each simulated return path of 5,000 data points, the 'Monthly' estimator uses 200 non-overlapping 25-day returns and the 'Overlapping' estimator uses 4,976 overlapping 25-day returns.

Under the assumption that daily returns follow a GBM, 25-day skewness and kurtosis should both be zero. Table 1 confirms that, on average, this is true for all three estimation techniques. More importantly, the dispersion of the estimates for the NP method are greatly reduced relative to those from monthly and overlapping estimations. The standard deviations of estimates from our method are between 70% and 90% smaller than those from other methods (and the distances between the 5th and 95th percentiles of the distributions of estimates are similarly reduced). The improvement in estimation accuracy for the NP method is most striking for skewness, but only slightly less impressive for kurtosis. Overall, for both skewness and kurtosis and regardless of which dispersion measure you favour, the use of high-frequency data to improve low-frequency moment estimates provides a substantial improvement in accuracy. Figure 2 shows the distributions of estimates under the GBM assumption and the improved accuracy of the NP estimation method over the others is clear.

Table 2: simulation results for Heston model

Standard deviation			
	NP	Monthly	Overlap
Mean	0.0469	0.0468	0.0468
STDEV	0.0023	0.0032	0.0032
5th prctile	0.0431	0.0417	0.0418
95th prctile	0.0508	0.0524	0.0522
Coefficient of Skewness			
	NP	Monthly	Overlap
Mean	-0.2728	-0.2616	-0.2584
STDEV	0.1025	0.3543	0.3478
5th prctile	-0.4454	-0.8341	-0.8204
95th prctile	-0.1147	0.3178	0.3097
Excess Kurtosis			
	NP	Monthly	Overlap
Mean	1.0769	1.0014	1.0194
STDEV	0.4156	0.8426	0.6910
5th prctile	0.5485	0.0444	0.2442
95th prctile	1.8369	2.4540	2.2193

For the Heston model, we expect excess kurtosis (as the variance of daily returns is changing through time) and negative skew (as innovations to the variance equation and the return equation are negatively correlated). Again all three estimation techniques pick these features up, but again use of the NP method results in a significant reduction in the spread of estimation errors. For skewness, the standard deviation of estimates for the new method is around 70% smaller than those of the monthly or overlapping methods, while for kurtosis improvements are between 40% and 50%. Again, the improved precision generated by the new estimator is clearly visible in Figure 3.

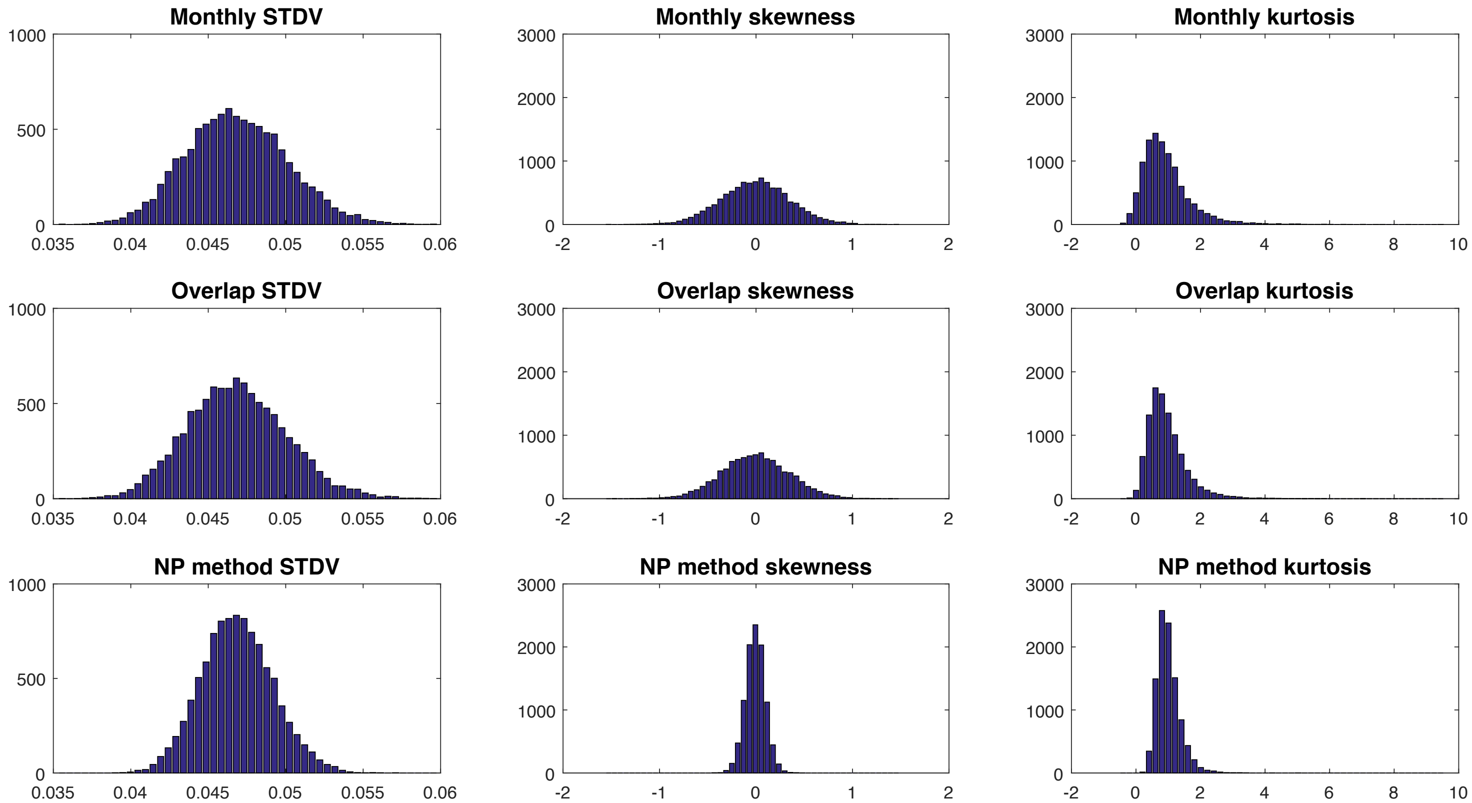
Finally, the estimated EGARCH model also generates negative skew and excess kurtosis and these appear in all estimation methods. Table 3 shows that the NP estimation technique dominates in terms of accuracy under this model also but the improvements in accuracy are less pronounced. The standard deviation of skewness estimates is around 60% smaller for the new method but the standard deviation of the kurtosis estimates drops by only 20 to 30%. Figure 4 plots these results.

Table 3: simulation results for EGARCH model

	NP	Monthly	Overlap
Standard deviation			
Mean	0.0605	0.0603	0.0603
STDEV	0.0031	0.0048	0.0048
5th prctile	0.0555	0.0529	0.0531
95th prctile	0.0657	0.0686	0.0687
Coefficient of Skewness			
	NP	Monthly	Overlap
Mean	-0.6656	-0.6348	-0.6280
STDEV	0.1786	0.4314	0.4246
5th prctile	-0.9875	-1.3672	-1.3408
95th prctile	-0.4311	0.0082	0.0171
Excess Kurtosis			
	NP	Monthly	Overlap
Mean	2.1650	1.9278	1.9424
STDEV	1.4409	2.0072	1.7886
5th prctile	0.8391	0.2024	0.4155
95th prctile	4.5723	5.2941	4.9854

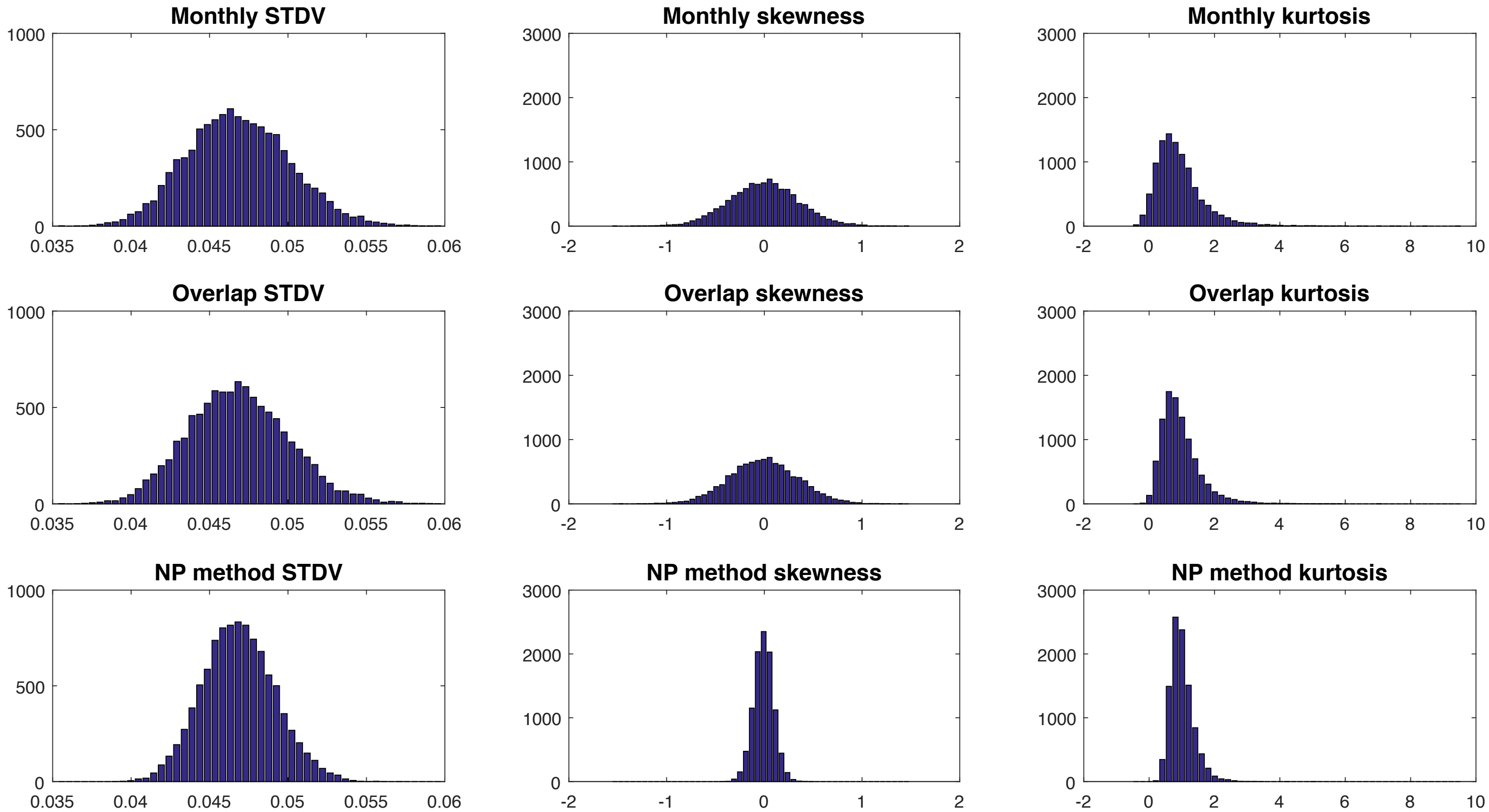
Overall, regardless of which model we choose or which moment one focusses on, use of the estimators described in Proposition 2 leads to much more precise estimates of monthly return moments. Improvements are greater for skewness estimates than they are for kurtosis and are larger for the GBM and Heston models than they are for the EGARCH specification. But in almost all cases, use the NP moment estimators leads to the dispersion of estimated coefficients being reduced by 50% or more.

Figure 2



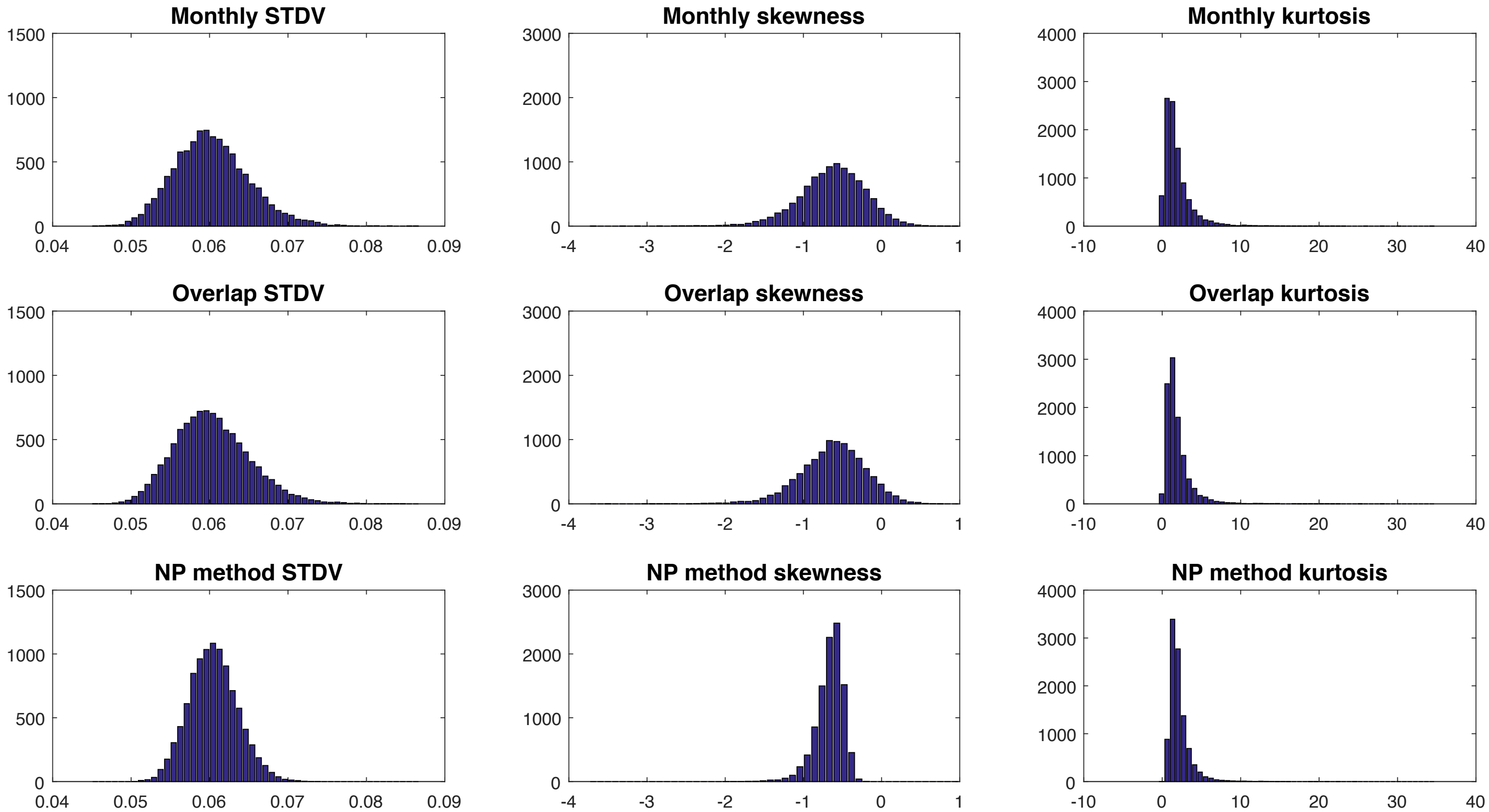
Monthly STDV
Monthly skewness
Monthly kurtosis
Overlap STDV
Overlap skewness
Overlap kurtosis
NP method STDV
NP method skewness
NP method kurtosis

Figure 3



Monthly STDV
Monthly skewness
Monthly kurtosis
Overlap STDV
Overlap skewness
Overlap kurtosis
NP method STDV
NP method skewness
NP method kurtosis
1000
500
0
0.035
0.04
0.045
0.05
0.055
0.06
3000
2000
1000
0
-2
-1
0
1
2
-2
0
2
4
6
8
10

Figure 4


Monthly STDV
Monthly skewness
Monthly kurtosis
Overlap STDV
Overlap skewness
Overlap kurtosis
NP method STDV
NP method skewness
NP method kurtosis

3.2 Focus on skewness and comparison with quantile-based estimation

As the application to the US stock market that we provide later in the paper will use the NP method applied to measurement of skewness, in this section we focus on estimation of the third moment. First we investigate how the performance of the estimator changes as the level of skew in returns changes and second we compare performance with a quantile-based estimate of skew recently proposed by Ghysels, Plazzi and Valkanov (2016).

3.2.1 Performance of the NP estimator across skew levels

In order to investigate how the performance of our estimator changes with the level of skew in returns, we take a Heston model and vary the correlation between return and variance innovations between -0.9 and +0.9 (with the former giving large negative skewness and the latter generating large positive skewness). All other parameters are set at the values from Eraker (2004). For each parameter set, our simulation contains 1,000 replications of 1,000 daily returns and from these we estimate 25-day skew.

The results are summarised in Figure 5. The x-axis of this figure shows the correlation parameter from the Heston model. Against each correlation parameter, we plot the average estimated skewness from our 1,000 runs, as well as the inter-quartile range and the 5^{th} and 95^{th} percentiles of the distribution of skew estimates. Also plotted on Figure 5 is the theoretical value of the coefficient that one should obtain from the Heston model at each parameter value.

The Figure demonstrates that the NP estimator does an excellent job of tracking skewness, on average, across the range of parameters. There is a slight tendency for the estimator to be biased towards zero when the theoretical skew is large, though, with the largest bias around 0.1 when theoretical skewness is at a value of 0.9. The inter-quartile range is fairly stable at a value of around 0.25. The bias in the estimation of the coefficient of skewness arises due to the fact that it is a ratio of the estimated third moment to the cube of the estimated standard deviation. Estimates of both of these moment measures using the NP method are unbiased, but estimation errors in second and third moments are correlated and it is this that causes the bias in the estimated skewness.

Obviously, as one changes the quantity of high-frequency data points one uses to construct low frequency skew, the estimation becomes more accurate. Figure 6 is an identical plot of results but from simulations of time-series of length 10,000, rather than 1,000, and comparison of Figures 5 and 6 clearly shows the improvement in estimation precision, with the bias dropping to close to zero and the inter-quartile range to around 0.1.

Figure 5

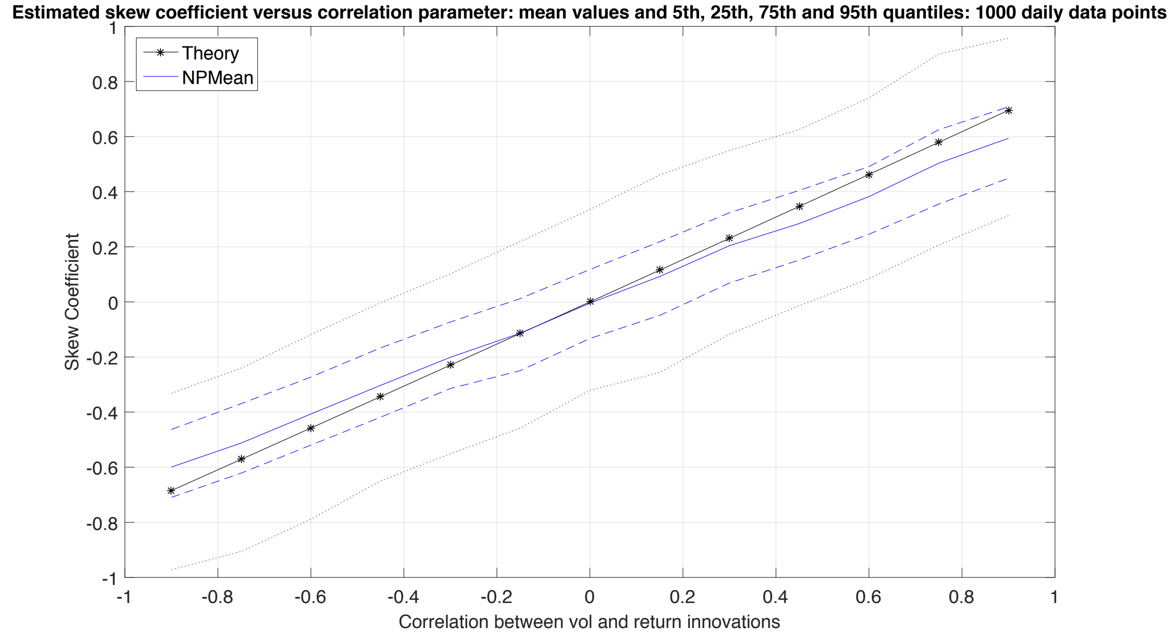


Figure 6

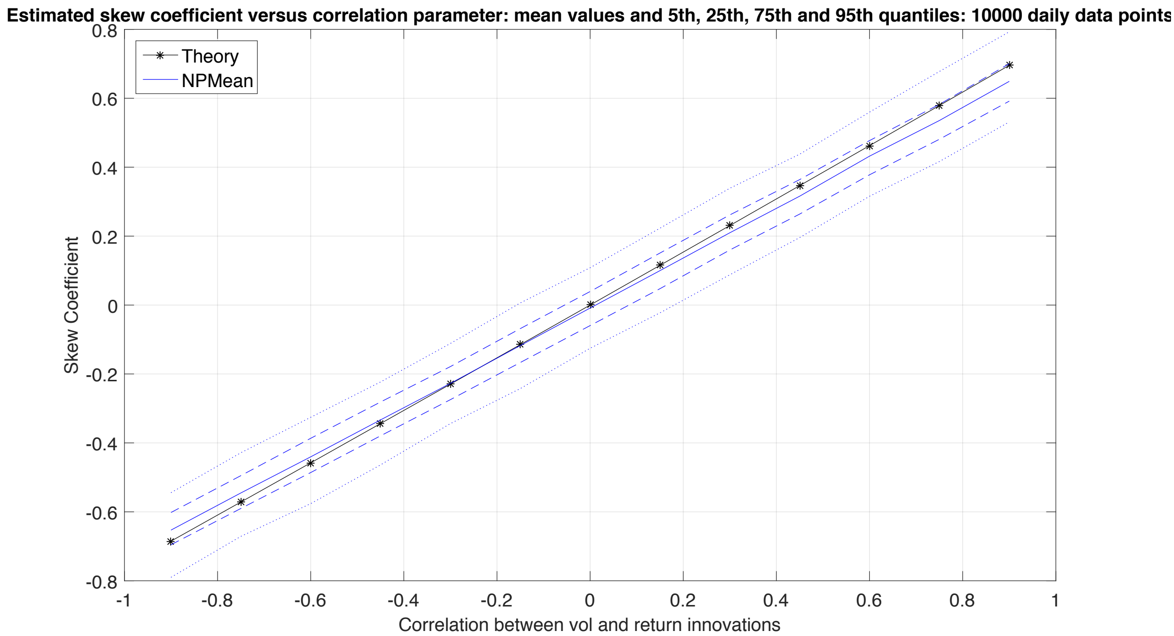


Figure 7

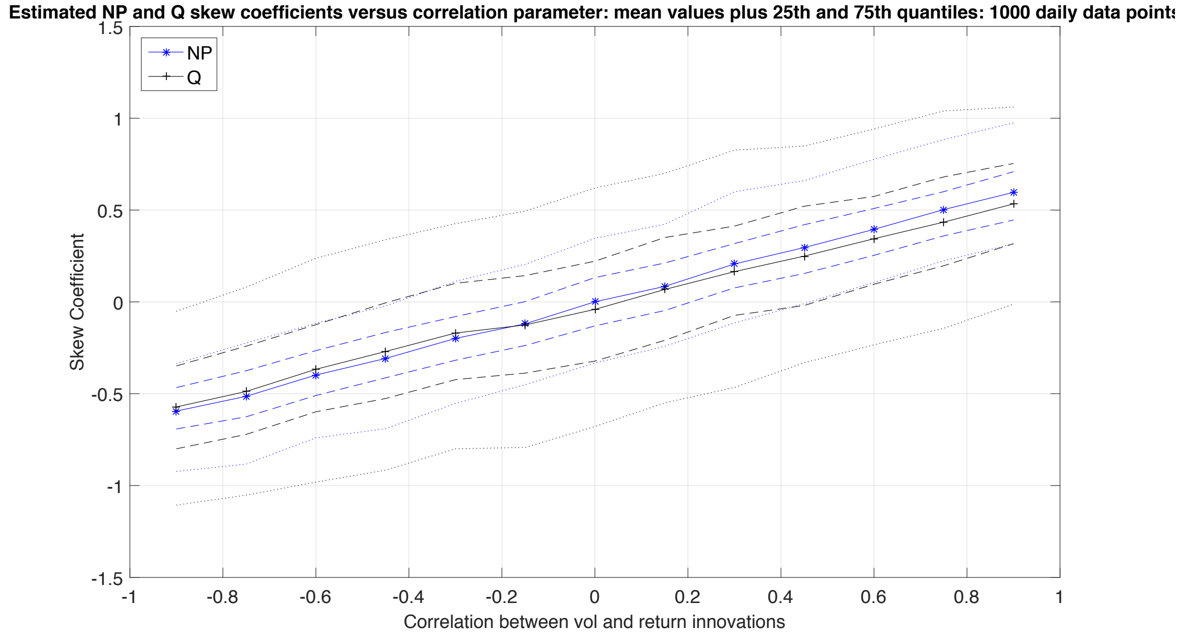
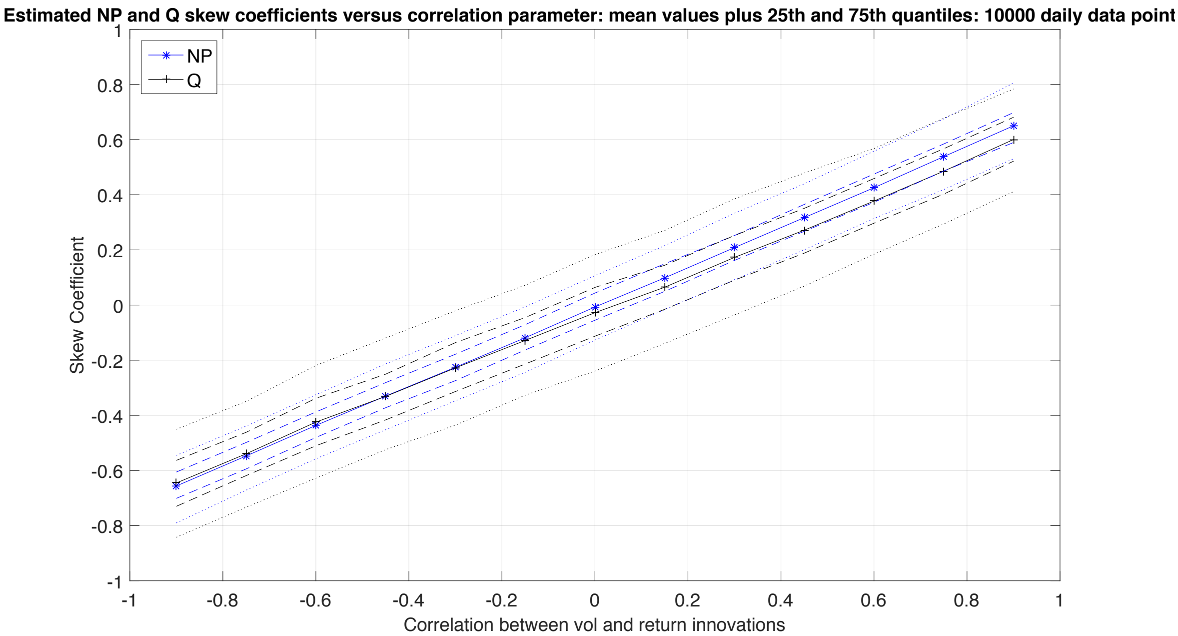


Figure 8



3.2.2 Comparison of NP and quantile-based skew measures

Ghysels, Plazzi and Valkanov (2016) (hereafter GPV) propose a skewness estimator based on the quantiles of the return distribution in their recent work on international asset allocation. We now compare the performance of our estimator and their preferred estimator based on data simulated from a Heston model for returns. We use exactly the same simulation setup as in Section 3.2.1, except now we estimate NP skewness and GPV's quantile skewness for each simulated set of data.

The GPV skew estimator is as follows;

$$6 \times \frac{\int_0^{0.5} \{[q_\alpha(r_t) - q_{0.5}(r_t)] - [q_{0.5}(r_t) - q_{1-\alpha}(r_t)]\} d\alpha}{\int_0^{0.5} [q_\alpha(r_t) - q_{1-\alpha}(r_t)] d\alpha} \times \frac{\int_0^{0.5} q_\alpha(z)\, d\alpha}{\int_0^{0.5} q_\alpha^2(z)\, d\alpha}$$

where r_t are returns measured at the frequency of interest (e.g. monthly), $q_\alpha(x)$ is the αth quantile of the distribution of x and the $q_\alpha(z)$ are the quantiles of the standard Normal distribution. In their implementation, GPV approximate the integrals in the first ratio by aggregating across the following set of quantiles: [0.99, 0.975, 0.95, 0.90, 0.85, 0.80, 0.75].

As this formula makes clear, the GPV estimator estimates skew by looking directly at the symmetry (or lack of it) of α and 1-α quantiles with respect to the median. This is captured by the numerator of the first ratio in the equation while the other terms are just scaling factors.

For each simulation run, we apply the GPV estimator to overlapping 25-day returns. It is worth re-iterating that the GPV estimator and the estimator proposed here are designed to target slightly different measures of skewness. GPV propose an estimator of the traditional skewness coefficient whereas our estimator is of the modified skew coefficient as defined in equation (9). However, differences in these targets are minor.

The results from our comparison are displayed in Figure 7. As before, the x-axis values of Figure 7 are the Heston correlation parameters and skewness is on the y-axis.

The results from the Figure are encouraging. The NP and the GPV mean estimates lie very close together in Figure 7, but the precision of the NP estimator is much greater. The inter-quartile range of the GPV estimates average around 0.5 i.e. around twice as large as that for the NP estimator. In Figure 8, where simulations are based on 10,000 rather than 1,000 high-frequency returns, the results on precision are similar, with the NP estimator having an inter-quartile range around half that of the GPV estimator.

4. Application to the US Equity Market

In this section, we apply our technology to the US stock market. In particular, we document how the moments of long horizon (monthly) returns have evolved over the last ninety years, and the extent to which the moments of monthly returns derive from the moments of short horizon (daily) returns. We then focus on skew, and explore the relationship between time variation in skew and future market returns.

The returns cover the period from 1926 to the end of 2015 and they were retrieved from Ken French's data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

4.1 Moments of 25-day returns for non-overlapping sample years

To document the behaviour of the moments of long horizon returns, we take non-overlapping years of data and use the daily data within each year to estimate 25-day return standard deviations, skewness and excess kurtosis. We choose 25-day returns to approximate a monthly horizon. Time-series plots of the three moments estimated using the NP method, and skewness estimated using the quantile method, are presented in Figure 9.[6]

The top left panel of Figure 9 indicates that, over the 90 or so years that our data cover, US stock market volatility follows a U-shaped pattern, punctuated by infrequent upward spikes. Stock markets were most volatile in the Great Depression and the recent Financial Crisis, but there was also substantial market volatility around the oil price shocks of the early 1970s and the stock market crash of 1987.

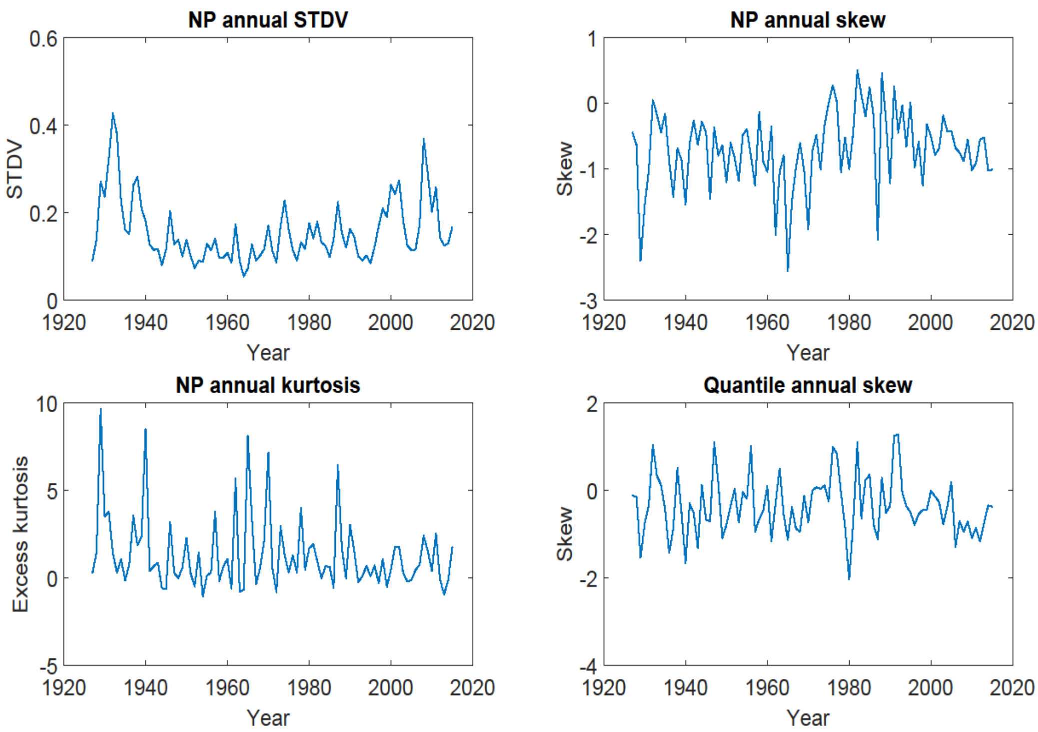
The NP skew measures indicate that monthly stock market skew is almost always negative (with a mean of -0.7) over our 90 years of data. The only exception to this is a period covering the late 1970s to mid-1990s where several years of slight positive skew are apparent (although one would expect this positive skew to be statistically insignificant). Times of particularly severe negative skewness include the Great Depression, the mid-1960s and the stock market crash of 1987. Interestingly, the recent Financial Crisis does not appear to be associated with large negative skew. If anything, our estimates suggest that since the mid-1990s, US stock market skew has been relatively subdued. The quantile-based skew measure, in the bottom right panel, is also negative on average (averaging -0.35), but it is less easy to see a pattern in the monthly skews here than it is in the NP estimates. The quantile skew and NP skew measures are positively correlated, with a correlation coefficient of 0.50.

Finally, the bottom left panel of Figure 9 show estimates of excess monthly return kurtosis estimated from daily data. As expected, excess kurtosis is positive on average, with a mean of 1.34, and the Great Depression and the 1987 crash show up prominently. Again, though, the recent Financial Crisis does not appear to display any unusual excess kurtosis.

[6] The quantile based skew measure is estimated from the set of overlapping 25 day returns that can be constructed from the year of daily returns.

Finally, comparing the top right and bottom left panels, there is clear evidence that years of large negative skew are also those with large excess kurtosis. In fact the correlation between the annual time-series for monthly skew and monthly excess kurtosis is -0.73. Thus at times when the weight in the tails of the distribution of returns is rising, the left tail is becoming heavier than the right.

Figure 9



4.2 The components of 25-day skew and kurtosis

As Proposition 2 makes clear, skew in long horizon returns is driven by skew in high-frequency returns and by the leverage effect. Long-horizon kurtosis has three possible sources: kurtosis in high-frequency returns, covariation between lagged returns and current cubed returns (which we refer to as the 'Cube' component) and covariation between current and lagged squared returns (which we will call the GARCH effect).

Figure 10 shows the time-series variation in the two skew components for the years 1927 to 2015. It is clear that all of the action in generating negative skewness comes from the leverage effect. The influence of skew in daily returns is negligible. Thus, both the average level of 25-day skew in our sample and its significant variation through time are attributable to correlations between lagged returns and current squared returns.

Figure 11 shows the decomposition of 25-day kurtosis into its three components (i.e. daily kurtosis, the cube term and the GARCH term). As with skew, the contribution of the daily moment is close to zero and its time-series variation is small. The Cube term is also close to zero on average and so almost all of the mean excess kurtosis apparent in the data, as well as the time-variation in that excess kurtosis, comes from the GARCH component.

Figure 10

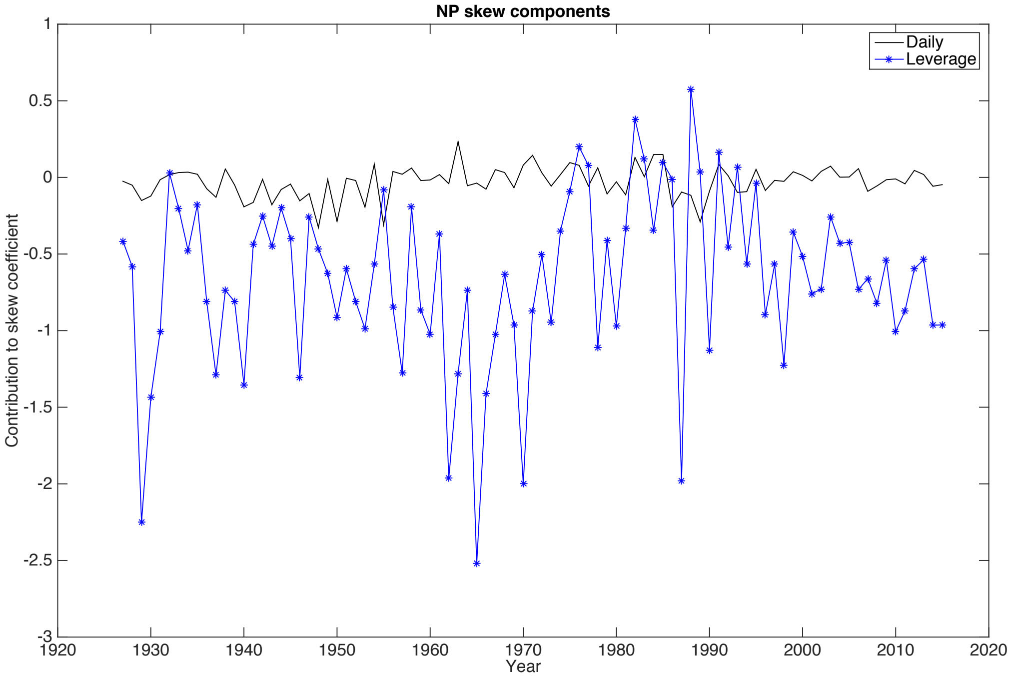
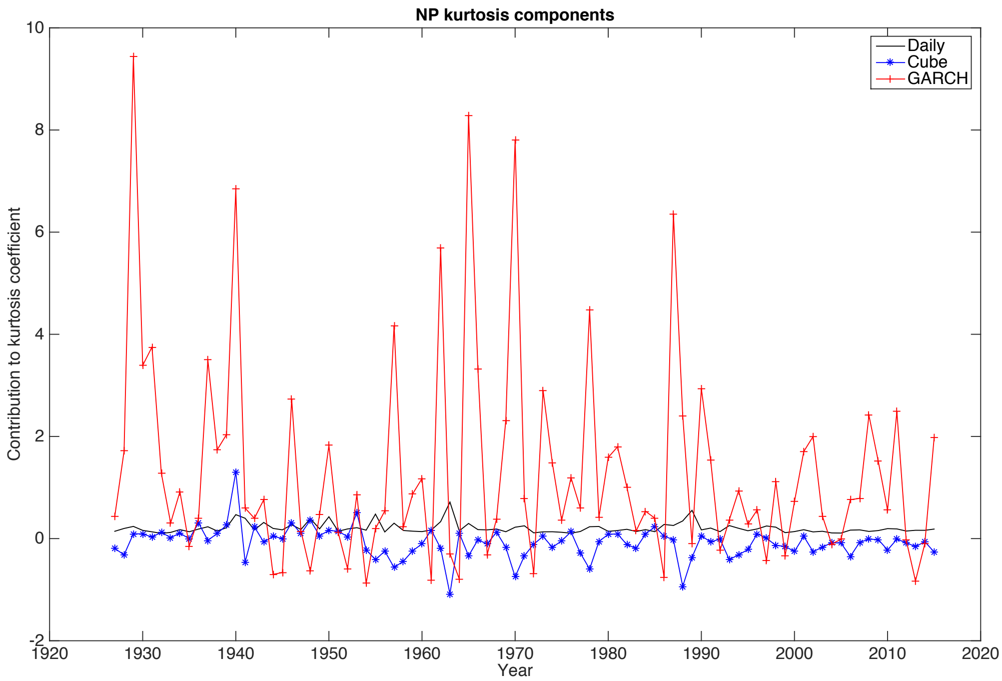


Figure 11



We have shown that the higher moments of long-frequency returns and daily returns are essentially unconnected. Low-frequency skewness and kurtosis are driven by leverage and GARCH effects respectively.

4.3Forecasting stock market returns using skew

As argued by Kraus and Litzenberger (1976) and much of the subsequent literature, there are good theoretical grounds for believing that if market skewness is time-varying, it should forecast future market returns. The skewness that is relevant for asset pricing returns is the skewness of long horizon returns. Our new technology enhances our ability to explore the relationship between skewness and returns.

We take as a starting point the work of Kelly and Jiang (2014). These authors propose a measure of monthly tail risk for the US stock market that is based on information in the left-tails of index constituents. Specifically, they assume that all constituent stocks have common tail shapes so that the tail shape of the index can be accurately estimated by aggregating tail events across stocks. With this assumption, they build their ‘tail risk’ measure, which they describe as measuring the risk of ‘extreme market movements’ built from the ‘crash events’ of individual companies. Importantly, using data beginning in 1963, they show that tail risk is a powerful positive predictor of US stock market returns so that the greater the measured risk at *t*, the higher the mean returns from *t* to t+k (where *k* ranges from 1 month to 5 years).

We build upon the Kelly and Jiang (2014) analysis. As our brief discussion of their paper suggests, building a risk measure from the left tail of stock returns is akin to measuring crash risk, and a more direct measure of index crash risk can be obtained from the skewness of the index return. The more negative the skew of the index return, the greater the crash risk. Thus we augment the Kelly and Jiang (2014) forecasting framework with a skew variable. We run monthly regressions such as that below;

$$r_{t,t+k} = \alpha + \beta_1 Tail_t + \beta_2 Skew_t + u_t$$

where $r_{t,t+k}$ is the return on the US stock market from *t* to *t* plus *k* months and the two right-hand side variables are tail risk measured over the month ending at *t* and 25-day index skewness measured using our proposed estimator. Skewness is measured using data from the past 500 days and *k* varies from 1 month to 60 months.

Table 4 contains the results from running our forecasting regressions. First we run regressions with tail risk only, thus mimicking the work of Kelly and Jiang (2014) and then we add our skew measure to the specification.

A first observation is that we entirely corroborate the results in Kelly and Jiang (2014). Their tail risk measure is positively and significantly related to returns across all forecasting horizons, with the explanatory power of the regressions rising from 1% at the 1 month horizon to 28% at the 5 year horizon. Thus greater tail risk robustly predicts greater returns.

When we add 25-day index skew into the picture, however, some interesting additional results emerge. First of all, the coefficient on skew is significant and negative for forecasting horizons of 1 year or less and is insignificant otherwise. Tail risk retains its sign and significance in all specifications. Thus, skewness has incremental forecasting power for monthly and annual returns, such that greater negative skew today leads to greater positive returns in the future. At those horizons, the addition of skewness to the regression improves the explanatory power (as measured by R^2) by about 50% .

Two insights emerge from this analysis. First, tail risk is not simply a measure of asymmetry in index returns. Tail risk and 25-day skewness simultaneously contribute to forecasting returns so tail risk is not encoding the same information as 25-day skew. Indeed, if one simply measures the correlation between tail risk and 25-day skew, it is positive i.e. greater tail risk leads to greater positive skew. Given these facts, it is clear that while tail risk is telling us something important about the distribution of index returns, its power does not come from measuring asymmetry in returns.

The second insight is that investors appear to demand compensation for the risk associated with greater negative monthly skew in index returns. Thus, higher moments do matter for asset prices.

5. Conclusions

Measures of the higher moments of low-frequency (i.e. monthly or quarterly) returns on stock indices or currencies or managed portfolios are important in a variety of contexts, including risk management, portfolio selection and asset pricing. But these moments are hard to measure due to the limited data available on low-frequency returns and due to the fact that the higher moments of low-frequency returns need bear little relationship to the higher moments of high frequency returns.

In this paper we show how, under fairly general conditions, high-frequency (e.g. daily) returns can be used to estimate low frequency skewness and kurtosis with impressive precision. This precision is demonstrated via a set of simulation experiments in which returns are generated from a few popular data generating processes (e.g. a Heston model and a GARCH model).

The analysis demonstrates that the skewness of low frequency returns has two components, the skewness in high-frequency returns and the covariance between lagged returns and current squared high-frequency returns (i.e. the leverage effect). Empirically, the latter is shown to be much more important than the former when measuring the skewness of annual or monthly US stock index returns using daily data. Low frequency kurtosis has three components. These are the kurtosis of high-frequency returns, the correlation between lagged returns and current cubed high-frequency returns and the correlation between lagged and current squared high-frequency returns. The last of these, which we call the GARCH effect, is the dominant contributor to the time-series variation in low frequency kurtosis in US stock index returns. Thus, we show, both analytically and empirically, that information on high-frequency skewness and kurtosis is close to irrelevant when it comes to measuring low frequency skew and kurtosis. Low frequency moments are instead driven by the dynamic relationships between (powers of) past high-frequency returns and (powers of) current high-frequency returns.

Table X: forecasting market returns with tail risk and skew

Estimates from regressions where tail risk and monthly skewness are used to forecast US stock index returns at various horizons. All regressions are estimated using OLS with standard errors estimated using the Hansen-Hodrick procedure (to account for the overlap in the dependent variable when using annual, 3 year or 5 year returns).

	1 month		1 year		3 years		5 years	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	6.033	6.031	5.827	5.853	5.556	5.546	5.304	5.298
t-stat	2.692	2.699	3.218	3.287	4.387	4.430	5.548	5.657
Tail	5.030	6.722	4.251	5.643	3.832	3.694	3.122	3.065
t-stat	2.393	3.100	2.489	3.351	3.836	3.663	5.681	6.309
Skew		-4.230		-3.418		0.323		0.132
t-stat		-2.317		-2.843		0.558		0.275
R^2	0.009	0.014	0.066	0.103	0.208	0.209	0.281	0.281

Finally, we demonstrate the utility of our low-frequency moment measures by investigating the ability of monthly skewness to forecast stock index returns. We show that, in a model including both skewness and the tail risk measure proposed by Kelly and Jiang (2014), monthly skewness is a significant negative forecaster of monthly and annual index returns. Thus the more negative is monthly skewness, the larger are expected monthly and annual returns. As Kelly and Jiang (2014) show, finding a risk measure that retains significance when used alongside their tail risk measure is unusual and in our case the improvements in explanatory power contributed by skewness are impressive. Thus, when measured correctly, low frequency moments do correlate with future mean returns, implying that investors are sensitive to the asymmetry in low-frequency return distributions.

REFERENCES

Ang, A., Chen, J. and Xing, Y., 2006. Downside risk. *Review of Financial Studies*, 19(4), pp.1191-1239.

Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X., 2006. The cross-section of volatility and expected returns. *The Journal of Finance*, *61*(1), pp.259-299.

Bakshi, G., Kapadia, N. and Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, *16*(1), pp.101-143.

Barberis, N. and Huang, M., 2007. Stocks as lotteries: The implications of probability weighting for security prices (No. w12936). *National Bureau of Economic Research*.

Barro, R.J., 2006. Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, pp.823-866.

Bollerslev, T., Osterrieder, D., Sizova, N. and Tauchen, G., 2013. Risk and return: Long-run relations, fractional cointegration, and return predictability. *Journal of Financial Economics*, 108(2), pp.409-424.

Bondarenko, O., 2014. Variance trading and market price of variance risk. *Journal of Econometrics*, *180*(1), pp.81-97.

Bowley, A.L., 1920. *Elements of statistics* (Vol. 2). PS King.

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

Broadie, M., Chernov, M. and Johannes, M., 2007. Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, *62*(3), pp.1453-1490.

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

Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance*, *59*(3), pp.1367-1403.

Ghysels, E., Plazzi, A. and Valkanov, R.I., 2016. Why invest in emerging markets? The role of conditional return asymmetry. *The Journal of Finance, 71(5),* 2145-2192.

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

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

Kelly, B. and Jiang, H., 2014. Tail risk and asset prices. *Review of Financial Studies*, *27*(10), pp.2841-2871.

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

Kim, T.H. and White, H., 2004. On more robust estimation of skewness and kurtosis. *Finance Research Letters*, *1*(1), pp.56-73.

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

Neuberger, A., 2012. Realized skewness. *Review of Financial Studies*, *25*(11), pp.3423-3455.

Longstaff, F.A. and Piazzesi, M., 2004. Corporate earnings and the equity premium. *Journal of financial Economics*, *74*(3), pp.401-421.

Rietz, T.A., 1988. The equity risk premium a solution. *Journal of Monetary Economics*, *22*(1), pp.117-131.

APPENDIX

Proof of Proposition 1

The monthly price change is the sum of daily price changes so

$$
\begin{aligned}
D_t^2 &= \sum_{u=0}^{T-1} d_{t-u}^2 + 2\sum_{u=0}^{T-1}\left(P_{t-u-1} - P_{t-T}\right) d_{t-u};\\
D_t^3 &= \sum_{u=0}^{T-1} d_{t-u}^3 + 3\sum_{u=0}^{T-1}\left(P_{t-u-1} - P_{t-T}\right) d_{t-u}^2 + 3\sum_{u=0}^{T-1}\left(P_{t-u-1} - P_{t-T}\right)^2 d_{t-u};\\
D_t^4 &= \sum_{u=0}^{T-1} d_{t-u}^4 + 4\sum_{u=0}^{T-1}\left(P_{t-u-1} - P_{t-T}\right) d_{t-u}^3 + 6\sum_{u=0}^{T-1}\left(P_{t-u-1} - P_{t-T}\right)^2 d_{t-u}^2\\
&\quad +4\sum_{u=0}^{T-1}\left(P_{t-u-1} - P_{t-T}\right)^3 d_{t-u}.
\end{aligned}
\tag{18}
$$

Taking conditional expectations at time *t-T*, the last term drops out so

$$
\begin{aligned}
\mathrm{E}_{t-T}\left[D_t^2\right] &= \sum_{u=0}^{T-1} \mathrm{E}_{t-T}\left[d_{t-u}^2\right];\\
\mathrm{E}_{t-T}\left[D_t^3\right] &= \sum_{u=0}^{T-1} \mathrm{E}_{t-T}\left[d_{t-u}^3\right] + 3\sum_{u=0}^{T-1} \mathrm{E}_{t-T}\left[\left(P_{t-u-1} - P_{t-T}\right) d_{t-u}^2\right];\\
\mathrm{E}_{t-T}\left[D_t^4\right] &= \sum_{u=0}^{T-1} \mathrm{E}_{t-T}\left[d_{t-u}^4\right] + 4\sum_{u=0}^{T-1} \mathrm{E}_{t-T}\left[\left(P_{t-u-1} - P_{t-T}\right) d_{t-u}^3\right]\\
&\quad +6\sum_{u=0}^{T-1} \mathrm{E}_{t-T}\left[\left(P_{t-u-1} - P_{t-T}\right)^2 d_{t-u}^2\right].
\end{aligned}
\tag{19}
$$

Taking unconditional expectations and rearranging terms

$$
\begin{aligned}
&\mathrm{E}\left[D^2\right] = T\mathrm{E}\left[d^2\right];\\
&\mathrm{E}\left[D^3\right] = T\mathrm{E}\left[d^3\right] + 3T\mathrm{E}\left[y^* d^2\right];\\
&\mathrm{E}\left[D^4\right] = T\mathrm{E}\left[d^4\right] + 4T\mathrm{E}\left[y^* d^3\right] + 6T\mathrm{E}\left[z^* d^2\right];\\
&\text{where } y_t^* := \sum_{u=1}^{T}\left(P_{t-1} - P_{t-u}\right)\big/T; \quad \text{and } z_t^* := \sum_{u=1}^{T}\left(P_{t-1} - P_{t-u}\right)^2\big/T.
\end{aligned}
\tag{20}
$$

Now

$$
\begin{aligned}
&\mathrm{E}\left[y^*\right] = 0 \quad \text{and}\\
&\mathrm{E}\left[z^*\right] = \sum_{u=1}^{T}(u-1)\mathrm{E}\left[d^2\right]\big/T = \tfrac{1}{2}(T-1)\mathrm{E}\left[d^2\right].
\end{aligned}
\tag{21}
$$

Replacing expectations of products in (20) by covariances and products of means we get

$$\begin{aligned}
&\mathrm{E}\left[D^2\right]=T\mathrm{E}\left[d^2\right];\\
&\mathrm{E}\left[D^3\right]=T\left\{\mathrm{E}\left[d^3\right]+3\operatorname{cov}\left(y^*,d^2\right)\right\};\\
&\mathrm{E}\left[D^4\right]-3\mathrm{E}\left[D^2\right]^2=T\left\{\mathrm{E}\left[d^4\right]-3\mathrm{E}\left[d^2\right]^2+4\operatorname{cov}\left(y^*,d^3\right)+6\operatorname{cov}\left(z^*d^2\right)\right\}.
\end{aligned}\tag{22}$$

Using the standard definitions of skewness and excess kurtosis, the result follows.

Proof of Proposition 2

Applying similar arguments to those used in the previous proof (algebraic decomposition, take conditional expectations, drop terms using the martingale property, replace conditional by unconditional expectations) we get

$$\begin{aligned}
&\mathrm{E}\left[\ln R\right]=T\mathrm{E}\left[\ln r\right];\\
&\mathrm{E}\left[R\ln R\right]=T\mathrm{E}\left[r\ln r\right]+T\mathrm{E}\left[yr\ln r\right];\\
&\mathrm{E}\left[\left(\ln R\right)^2\right]=T\mathrm{E}\left[\left(\ln r\right)^2\right]+2T\mathrm{E}\left[w\ln r\right];\\
&\text{where } y_t:=\sum_{u=1}^{T}\left(\frac{P_{t-1}}{P_{t-u}}-1\right)\Big/T \text{ and } w_t:=\sum_{u=1}^{T}\ln\left(P_{t-1}/P_{t-u}\right)\Big/T.
\end{aligned}\tag{23}$$

Substituting these into the definitions of v^L, v^E, s and k gives

$$\begin{aligned}
&\mathrm{E}\left[x^{(2L)}(R)\right]=T\mathrm{E}\left[x^{(2L)}(r)\right];\\
&\mathrm{E}\left[x^{(2E)}(R)\right]=T\left\{\mathrm{E}\left[x^{(2E)}(r)\right]+2\mathrm{E}\left[yr\ln r\right]\right\};\\
&\mathrm{E}\left[x^{(3)}(R)\right]=T\left\{\mathrm{E}\left[x^{(3)}(r)\right]+6\mathrm{E}\left[yr\ln r\right]\right\};\\
&\mathrm{E}\left[x^{(4)}(R)\right]=T\left\{\mathrm{E}\left[x^{(4)}(r)\right]+24\mathrm{E}\left[w\ln r\right]+24\mathrm{E}\left[yr\ln r\right]\right\}.
\end{aligned}\tag{24}$$

With the definition of var^L the first line gives the first part of Proposition 2.

Note that r and y are independent, and both are mean zero so

$$6\mathrm{E}\left[yr\ln r\right]=3\operatorname{cov}\left(y,x^{(2E)}(r)\right).\tag{25}$$

This together with the definition of skewness and the third line of equation (24) gives the second part of the Proposition.

Finally, the fourth line of (24) gives

$$\begin{aligned}
&\mathrm{E}\left[x^{(4)}(R)\right]=T\left\{\mathrm{E}\left[x^{(4)}(r)\right]+6\mathrm{E}\left[zx^{(2L)}(r)\right]+4\mathrm{E}\left[ys(r)\right]\right\}\\
&\text{where } z:=2\left(y-w\right).
\end{aligned}\tag{26}$$

Now

$$
\begin{aligned}
&\mathrm{E}\left[y\right]=0 \quad \text{and} \\
&\mathrm{E}\left[z\right]=\sum_{u=1}^{T}\left(u-1\right)\mathrm{var}^{L}\left[r\right]/T=\tfrac{1}{2}\left(T-1\right)\mathrm{var}^{L}\left[r\right].
\end{aligned} \tag{27}
$$

Replace expectations of products by covariances, and products of expectations, substitute into the definition of kurtosis, and the final part of the proposition follows.