REALIZED SKEWNESS

Anthony Neuberger
Warwick Business School
Warwick University
Coventry
CV4 7AL
United Kingdom

anthony.neuberger@wbs.ac.uk

January 2012

The third moment of returns is important for asset pricing. But skewness, particularly of long horizon returns, is hard to measure precisely. In the case of the second moment, the measurement problem is addressed through the use of realized volatility. This paper proposes an analogous definition of the realized third moment that is computed from high frequency returns, and provides an unbiased estimate of the third moment of long horizon returns. The realized skewness of index returns is shown to be predictable and time varying. The skewness of index returns increases with the length of the horizon (at least up to one year).

The author is grateful to comments from Carol Alexander, Mark Britten-Jones, Peter Carr, David Hobson, Ingmar Nolte, Roman Kozhan, Eberhard Mayerhofer, Alessandro Palandri, Paul Schneider, Neil Shephard, Hao Zhou and seminar participants at the UBC Summer Conference, the Frankfurt MathsFinance Conference, the Oxford-Man Institute, the Universities of Brunel, Reading, Piraeus and Warwick, and Morgan Stanley.

.

The third moment of returns is important for asset pricing. But skewness, particularly of long horizon returns, is hard to measure precisely. In the case of the second moment, the measurement problem is addressed through the use of realized volatility. This paper proposes an analogous definition of the realized third moment that is computed from high frequency[1] returns, and provides an unbiased estimate of the third moment of long horizon returns.

While standard approaches to asset pricing concentrate largely on the first and second moment of returns, there is mounting evidence that the higher moments are also important. There is a substantial empirical literature going back to Kraus and Litzenberger (1976), and including more recently Harvey and Siddique (2000), Ang, Hodrick, Xing and Zhang (2006), Ang, Cheng and Xing (2006) and Xing, Zhang and Zhao (2010), that suggests that the asymmetry of the returns distribution both for individual stocks and for the equity market as a whole is important for asset pricing and investment management. Skewness is central to the debate on the role of large rare disasters in explaining the equity risk premium (Rietz (1988), Longstaff and Piazzesi (2004), Barro (2009), and Backus, Chernov and Martin (2011)). There is also a growing literature starting with Carr and Wu (2007) on skew in foreign exchange markets.

Our tools for measuring the third moment of returns are much less sophisticated than they are for the second moment. We know how to use high frequency data to compute realized volatility and so make good estimates of the actual variance of returns. We can compute the model free implied variance (MFIV) from option prices, and use the difference between implied and realized variance to test for the existence of variance risk premia. By contrast, there is no established methodology for estimating the third moment of period returns from high frequency returns, and no recognized concept of the realized third moment that can be traded against the option implied third moment. The aim of this paper is to fill the gap. In particular I propose definitions of the realized and implied third moment that are analogous to realized and implied variance.

[1] I use the term "high frequency" to mean high frequency relative to the horizon of interest, rather than restricting it to intra-day returns.

The parallel between the second and third moments is not straightforward. Whereas the annualized volatility of returns is substantially independent of horizon, the same is not true of skew. It is possible for daily returns to be highly skewed but for monthly returns to be virtually symmetric. If daily returns are identically and independently distributed (iid) then the skew in N day returns is $1/\sqrt{N}$ of the daily skew. Conversely, even if instantaneous returns are symmetric, long horizon returns can exhibit a pronounced skew, as in a Heston (1993) stochastic volatility model.

The skew in the Heston model is generated by the correlation between returns and the volatility of returns. The strong negative correlation between returns and volatility in the equity index market (the so-called leverage effect) has been documented and discussed since it was identified by Black (1976) and Christie (1982). As I will show, it is this leverage effect, and not the skew in short horizon returns, that generates most of the negative skew in long horizon returns. If realized skew computed from high frequency returns is to provide a useful estimate of the skew at lower frequencies, it must capture the leverage effect.

Schoutens (2005) and Schoutens, Simons, and Tistaert (2005) describe swap contracts that pay the sum of cubed daily returns. They capture the third moment of high frequency returns, but in failing to take account of the leverage effect the pay-offs bear no relation to the third moment of returns over the life of the swap. Bakshi, Kapadia and Madan ("BKM", 2003) show how the implied moments and skewness coefficient of the risk-neutral probability density can be recovered from option prices. However, due to risk premia in option prices, the implied density in general differs substantially from the density under the physical measure (see Carr and Wu, 2009). There is no obvious way of relating the implied BKM skew at any horizon to the actual behavior of high frequency returns up to that horizon.

Realized variance of high frequency returns over a period is a good measure of the variance of the period return because it has the Aggregation Property – the realized variance of high frequency returns is equal in expectation to the realized variance of the period return, at least for price processes that are martingales. Requiring the realized third moment to have the Aggregation Property turns out to be sufficient to define it uniquely.

Just as there is a model-free strategy to replicate a variance swap, a swap that pays the difference between option implied variance and realized variance, so the Aggregation Property ensures that there is a model-free strategy to replicate a third moment or skewness swap, one that pays the difference between the option implied third moment and the realized third moment. From the perspective of the practitioner, this allows the development of a useful risk management tool. From the academic perspective, just as Carr and Wu (2009) use variance swaps to explore variance risk premia, so Kozhan, Neuberger and Schneider (2011) use skewness swaps to explore the existence and behavior of risk premia associated with skewness.

Using high frequency data to compute the realized second and third moments makes it possible to measure the skewness of long horizon returns with far greater precision than one could using long horizon returns alone. I show, using daily data, that the skew in the equity index market, far from declining with horizon, is higher at one year than it is at one month. Furthermore there is evidence that skewness is time varying.

These results are relevant to the debate on the role of large disasters in explaining the equity premium. In reviewing the evidence, Backus, Chernov and Martin (2011, p1970, “BCM”) note that “in virtually all of this research [starting with Rietz (1988), followed by Longstaff and Piazzesi (2004), Barro (2009) and others], the distribution [of log returns] is modeled by combining a normal component with a jump component. The jump component, in this context, is simply a mathematical device that produces nonnormal distributions”. There are good modeling reasons to induce extreme events through a jump process. The returns process remains iid. With a static investment opportunity set, the optimal strategy of the representative agent is myopic. This makes the problem tractable, and the investor’s investment horizon irrelevant.

But if the evidence presented here is correct it means that the model is seriously mis-specified. The parameters chosen by BCM for the process for the equity index (under the physical measure) imply that the skewness of daily returns is -0.6. This is consistent with US data. But, coupled with the iid assumption, and the consequent absence of any leverage effect, it implies a skewness of annual returns of -0.04 (see BCM, Table 3), in contrast with a figure of around -1 suggested by the analysis presented here. If the horizon of the representative agent is of the order of years rather than days, the calibration of the model to daily returns may understate the role that rare disasters play in asset pricing.

The paper is structured as follows: after presenting some empirical evidence on the skewness of equity index returns to motivate the rest of the analysis, section 2 presents the Aggregation Principle and applies it to price changes. The mathematics of price changes is simpler than the mathematics of returns. Section 3 extends the theory to apply to returns and defines realized skewness. In section 4, I use simulations to analyze the performance of realized skewness, and in section 5 the methodology is taken direct to the data to investigate time variation in the term structure of skewness of index returns. The sixth section concludes.

1. Skewness of Equity Index Returns

Table 1 shows estimated skew coefficients (using standard definitions) for excess log returns on the S&P500 index for various different periods and frequencies. The data come from the CRSP data base (June 2011), and excess log returns are defined as $\log(R^M/R^F)$ where R^M is the total daily return on the S&P500 and R^F is the total daily return on T-bills, computed from the monthly rate. The p-values (against the null that the skew is positive) and the confidence intervals are obtained by bootstrapping.

It is striking that the evidence for the negative skewness of daily index returns rests heavily on one observation – 19 October 1987. In the period as a whole, the distribution of daily returns appears to be significantly skewed, but in each of the twelve year sub-periods that excludes the 1987 Crash, the estimated skew coefficient is not significantly different from zero, and indeed in one of the periods is positive (though not significantly so). The weakness of the evidence for negative skewness has been noted by other authors including Kim and White (2004).

Monthly[2] returns are significantly negatively skewed in the sample as a whole and in two of the four sub-periods; the central estimate is negative in all the sub-periods. But the standard errors are high; for the sample as a whole, comprising 48 years of data, the 90% confidence interval has a range of -0.29 to -1.14. Given the difficulty of measuring skewness, it is hard to tell whether skewness is constant or varies over time.

One other feature of the data that is striking is the high level of skewness in monthly relative to daily returns. If stock returns were iid, the moments of *n*-day returns would be proportional to *n* and the skewness coefficient would be proportional to $1/\sqrt{n}$. There is no evidence of such a decline in the data. To explore the relation between skewness coefficients of daily and monthly returns more fully, I bootstrap the daily returns to generate a synthetic time series, and compute the population skewness of both daily and monthly returns. The results are plotted in top panel of Figure 1. The lower panel does the same thing using monthly and annual rather than daily and monthly returns.

The hypothesis that returns are iid and the skew at the monthly horizon derives from the skew at the daily horizon is firmly rejected by the data (*p*-statistic of 0.02%), and the hypothesis that annual skew is generated by monthly skew is also rejected (*p*-statistic of 4.76%). This is clear evidence that the distribution of low frequency returns is heavily influenced by serial dependence in high frequency returns. If high frequency returns are to be used to improve the estimate of the skewness of low frequency returns, it must be done in a way that reflects the serial dependencies that are manifest in the data.

2. Arithmetic Contracts

This section introduces the Aggregation Property – the property that ensures that the quantity measured using high frequency returns is an unbiased estimate of its low frequency counterpart. The goal is to get a good measure of the third moment of returns,

---

[2] A month meaning a period of 22 trading days.

and that will be achieved in the following section. In this section we look at moments of price changes rather than returns because the mathematics are simpler.

2.1 Notation and Terminology

S_t $\left(t \in [0,T]\right)$ is a positive adapted variable defined on a standard filtered probability space $\left(\Omega, \mathfrak{F}, \left(\mathfrak{F}_t\right)_{t\in[0,T]}, \mathbb{P}\right)$. $\mathbb{E}_t[.]$ denotes $\mathbb{E}\left[.\middle|\mathfrak{F}_t\right]$. I assume that the distribution of S_T is such that expectations of S_T and functions of S_T such as S_T^2, $\ln(S_T)$ and $S_T\ln(S_T)$ exist.

$\mathbf{T} = \{t_0 = 0 < t_1 < ... < t_N = T\}$ is a partition of $[0,T]$. $\|\mathbf{T}\| = \max_i \{t_i - t_{i-1}\}$ is the mesh of $\mathbf{T}$. If x is a process, x_i is shorthand for x_{t_i} in order to avoid excessive subscripting. δx_i denotes $x_i - x_{i-1}$. Given a function $g(.)$, $\sum^{\mathbf{T}} g(\delta x)$ denotes $\sum_{i=1}^{N} g(\delta x_i)$

I will sometimes refer to the period $[0,T]$ as a month, and the length of the sub-period as a day, but obviously nothing hangs on this. In the present section, references to variance and skewness relate to the distribution of price changes and not of returns. To avoid irrelevant complications with interest rates and dividends, I work throughout with forward prices, so all trades, whenever entered into, are for settlement at time T.

2.2 The Aggregation Property

For any martingale S

$$\mathbb{E}_0\left[\left(\sum^{\mathbf{T}} \delta S\right)^2\right] = \mathbb{E}_0\left[\sum^{\mathbf{T}} (\delta S)^2\right]. \tag{1}$$

There are several interpretations of this equation. The left hand side is $\mathbb{E}_0\left[\left(S_T - S_0\right)^2\right]$. With S being martingale, this is equal to $\mathbb{E}_0\left[\left(S_T - S_0 - \mathbb{E}_0\left[S_T - S_0\right]\right)^2\right]$, the *true variance* of the monthly price change S_T-S_0 under the physical measure $\mathbb{P}$ conditional on information $\mathfrak{F}_0$ available at time 0. Equation (1) then says that if prices are martingales, the sum of squared daily price changes (the *realized variance*) is an unbiased estimator of the true variance. Alternatively, interpreting the measure $\mathbb{P}$ as a pricing measure, it says that the fair price of a one-month variance swap computed daily (a swap that pays the realized daily variance over a month) is the same as the price of a contingent claim that pays $\left(S_T - S_0\right)^2$. Indeed, since the relationship holds under any pricing measure, it also implies that a variance swap can be perfectly replicated if the contingent claim exists (or can be synthesized from other contingent claims), and the underlying asset is traded. It is reasonable to call the time 0 price of the claim the *implied variance*.

The relation between true variance of the monthly price change, the realized variance of daily changes over the month and the implied variance of one-month options at the beginning of the month holds exactly, whatever the price process and whatever the length and number of sub-periods, provided only that *S* is a martingale. It depends on the interchangeability of the summation and the square function under the expectations operator.

In order to generalize the notion of true, realized and implied characteristics, some more definitions are needed. If *g* is a real-valued function and *X* is an adapted (scalar or vector) process, then (*g*; *X*) has the *Aggregation Property* if for any times $0 \le s \le t \le u \le T$

$$\mathbb{E}_s\left[g\left(X_u - X_s\right) - g\left(X_u - X_t\right) - g\left(X_t - X_s\right)\right] = 0. \tag{2}$$

Applying the Law of Iterated Expectations, if (*g*; *X*) has the Aggregation Property then $\mathbb{E}_0\left[g\left(X_T - X_0\right)\right] = \mathbb{E}_0\left[\sum^{\mathbf{T}} g\left(\delta X\right)\right]$ for any partition **T**. $\sum^{\mathbf{T}} g\left(\delta X\right)$ is the *realized* characteristic; and $\mathbb{E}_0\left[g\left(X_T - X_0\right)\right]$ is the *implied* characteristic (if the measure is a pricing measure) or the *true* characteristic if the measure is the physical measure.

We have seen that (g; S) has the Aggregation Property for any martingale S when $g(x) = x^2$. The obvious question is whether there are any other interesting functions g for which this holds. The answer (as demonstrated formally below in Proposition 1) is no.

However, we do not need to restrict ourselves to the case where $X = S$. Suppose that $X(S)$ is a vector-valued process $\{S_t, V_t(S_T) : t \in [0,T]\}$ where $V_t(S_T) = Var_t[S_T]$, the variance of S_T conditional on information at t. Define **G** to be the set of analytic functions g such that $(g; X(S))$ has the Aggregation Property for all martingales S.

Proposition 1: **G** *consists of the functions*[3]

$$g(\delta S, \delta V) = h_1 \delta S + h_2 \delta V + h_3 \delta S^2 + h_4 \left(\delta S^3 + 3\delta S \delta V\right). \tag{3}$$

where the $\{h_i\}$ *are arbitrary constants.*

Proof: see Appendix.

Proposition 1 shows that **G** is spanned by four functions: $g = \delta S$ and $g = \delta V$, which are uninteresting, $g = \delta S^2$, which is the variance and is familiar, and $g = \delta S^3 + 3\delta S \delta V$, whose properties we now explore.

The true characteristic of this function is

$$\begin{aligned} \mathbb{E}_0\left[g\left(S_T - S_0, V_T - V_0\right)\right] &= \mathbb{E}_0\left[g\left(S_T - S_0, -V_0\right)\right] \text{ (since } V_T = 0\text{)} \\ &= \mathbb{E}_0\left[\left(S_T - S_0\right)^3 - 3V_0\left(S_T - S_0\right)\right] \text{ (definition of } g\text{)} \\ &= \mathbb{E}_0\left[\left(S_T - S_0\right)^3\right] \text{ (since } S \text{ is a martingale).} \end{aligned} \tag{4}$$

Hence, the characteristic captured by g is the third moment. The left hand side of equation (4) is the true third moment of the price change over the month. The realized

---

[3] The convention is that the δ operator has priority over all others, so δS^3 is read as $(\delta S)^3$ and not as $\delta(S^3)$.

third moment is $\sum^{\mathbf{T}} g\left(\delta S, \delta V\right) = \sum^{\mathbf{T}}\left(\delta S^3 + 3\delta S \delta V\right)$. The Aggregation Property means that the realized third moment equals the true third moment in expectation

$$\mathbb{E}_0\left[\sum^{\mathbf{T}}\left(\delta S^3 + 3\delta S\delta V\right)\right] = \mathbb{E}_0\left[\left(S_T - S_0\right)^3\right]. \tag{5}$$

Equation (5) is significant in several respects. It shows that skewness in low frequency returns derives only in part from the skewness in high frequency returns. The second source (and indeed the only source when *S* is a continuously sampled continuous martingale) is the covariation between shocks to the price level and shocks to future variance.

It also shows how high frequency data can be used to get more efficient estimates of the skewness in price changes over a period than can be obtained from just the price change over the period. This improvement rests on two assumptions: the discounted price process is a martingale, and the variance of the terminal price is in the observer's information set. Third, if a pricing measure is used, the right hand side is the implied third moment (which can be inferred from the prices of options that mature at time *T*). The difference between the implied and realized third moments can be used to detect and analyze risk premia associated with skewness.

Fourth, again interpreting the equation under a pricing measure, it suggests how a third moment swap can be designed and replicated. A third moment swap pays the difference between the implied and the realized third moments. To replicate the swap, both *S* and *V* have to be tradable as well as observable. To make *V* tradable, assume the existence of a traded square contract, a security that pays S_T^2. With $\mathbb{P}$ being a pricing measure, the price of the contract on day *i* is

$$\begin{aligned} P_i &= \mathbb{E}_i\left[S_T^2\right] \\ &= S_i^2 + \mathbb{E}_i\left[S_T^2 - S_i^2\right] \\ &= S_i^2 + V_i \text{ since } S_i = \mathbb{E}_i\left[S_T\right]. \end{aligned} \tag{6}$$

The gain from holding one square contract for one period is

$$\delta P_{i+1} = \delta S_{i+1}{}^2 + 2S_i \delta S_{i+1} + \delta V_{i+1}. \tag{7}$$

Suppose an agent enters into a one month third moment swap at day 0, paying floating and receiving fixed, with the realized third moment being computed from daily prices. She uses the fixed payment to buy a contract that pays the cube of the price change over the month. She hedges by holding -3(S_i - S_0) square contracts and $3\left(\left(S_i - S_0\right)^2 - V_i\right)$ forward contracts over day *i*. If her initial wealth W_0 = 0, her terminal wealth is

$$\begin{aligned} W_T &= \left\{\left(S_T - S_0\right)^3\right\} - \left\{\sum^{\mathbf{T}} \left(\delta S\right)^3 + 3\delta S \delta V\right\} \\ &-3\left\{\sum_{i=0}^{N-1}\left(S_i - S_0\right)\left(\left(\delta S_{i+1}\right)^2 + 2S_i\delta S_{i+1} + \delta V_{i+1}\right)\right\} + 3\left\{\sum_{i=0}^{N-1}\left(\left(S_i - S_0\right)^2 - V_i\right)\delta S_{i+1}\right\} = 0. \end{aligned} \tag{8}$$

(8) shows that the agent can hedge the swap exactly for any finite partition **T**. So in particular it works perfectly with discrete monitoring and with jumps in the price of the underlying or in its variance. The only requirements are that the market is frictionless, and that the asset and the square contract on it can be traded at any time *t* that is used for computing the realized third moment.

3. Geometric Contracts

3.1 Generalized Variance

Financial economists are interested in the behavior of returns, not price changes. It is tempting to apply the theory in the previous section directly to the log price, $s_t \equiv \ln S_t$. But the log price is not a martingale either under a pricing measure or, in general, under the physical measure. In looking for a definition of implied and realized variance of returns, I start from the premise that it is important to keep the Aggregation Property. The price for this is a relaxation of the definition of variance.

Let f be an analytic function on the real line with the property that $\lim_{x\to 0} f(x)/x^2 = 1$. Given a process s, define the process $v_t^f(s_T) = E_t\left[f(s_T - s_t)\right]$. I will call $v_t^f(s_T)$ a *generalized variance* process for s.[4]

The variance measures that are widely used by academics and practitioners (squared net returns and squared log returns) conform to the definition of generalized variance. Two other generalized variance measures, v^L and v^E, turn out to be important

$$\begin{aligned} v_t^L &= \mathbb{E}_t\left[L(s_T - s_t)\right] \text{ where } L(x) = 2\left(e^x - 1 - x\right), \\ v_t^E &= \mathbb{E}_t\left[E(s_T - s_t)\right] \text{ where } E(x) = 2\left(xe^x - e^x + 1\right). \end{aligned} \tag{9}$$

To explain the use of the letters L and E, rewrite (9) as

$$\begin{aligned} v_t^L &= 2\mathbb{E}_t\left[\frac{S_T}{S_t} - 1 - \ln\left(\frac{S_T}{S_t}\right)\right] \text{ so } \mathbb{E}_t\left[\ln S_T\right] = \ln S_t - v_t^L/2; \\ v_t^E &= 2\mathbb{E}_t\left[\frac{S_T}{S_t}\ln\left(\frac{S_T}{S_t}\right) - \frac{S_T}{S_t} + 1\right] \text{ so } \mathbb{E}_t\left[S_T \ln S_T\right] = S_t \ln S_t + S_t v_t^E/2. \end{aligned} \tag{10}$$

In a Black-Scholes world, where the underlying asset has constant volatility σ, the price of a log contract, one that pays $\ln S_T$, is $\ln S_t - \sigma^2(T-t)/2$, so v_t^L is the implied Black-Scholes variance of the log contract. It is the same as the model-free implied variance of Britten-Jones and Neuberger (2000). Similarly, v_t^E is the implied Black-Scholes variance of the contract that pays $S_T \ln S_T$. I call it entropy because of the functional similarity to entropy as used in thermodynamics and information theory.

[4] To simplify the algebra, the variance is not annualized; it tends to increase in magnitude with time to maturity.

The following properties of the log and entropy variance will be useful. From the first line of (10)

$$\begin{aligned} v_t^L &= 2s_t - 2\mathbb{E}_t\left[s_T\right] \text{ so} \\ \mathbb{E}_t\left[\delta v_{t+1}^L - 2\delta s_{t+1}\right] &= 0. \end{aligned} \tag{11}$$

Similarly, from the second line

$$\begin{aligned} & e^{s_t}\left(v_t^E + 2s_t\right) = 2\mathbb{E}_t\left[s_T e^{s_T}\right] \\ & \text{so } e^{s_t}\mathbb{E}_t\left[e^{\delta s_{t+1}}\left(v_t^E + \delta v_{t+1}^E + 2s_t + 2\delta s_{t+1}\right) - v_t^E - 2s_t\right] = 0 \\ & \text{which implies that } \mathbb{E}_t\left[e^{\delta s_{t+1}}\left(\delta v_{t+1}^E + 2\delta s_{t+1}\right)\right] = 0. \end{aligned} \tag{12}$$

3.2 Aggregation with Returns

Let **G*** be the set of analytic functions g where $\left(g; X\left(S\right)\right)$ has the Aggregation Property for any positive martingale S, where $X(S)$ is the vector process (s, v) with $s = \ln S$ and v is a generalized variance process for s.

Proposition 2: *the set* **G*** *comprises the following functions*

$$\begin{aligned} & g\left(\delta s, \delta v\right) = h_1\delta s + h_2\left(e^{\delta s} - 1\right) + h_3\delta v + h_4\left(\delta v - 2\delta s\right)^2 + h_5\left(\delta v + 2\delta s\right)e^{\delta s} \\ & \text{where } \left\{h_i\right\} \text{ are arbitrary constants, with the following constraints:} \\ & \quad - \text{ if } h_4 \neq 0, v = v^L \text{ and } h_5 = 0; \\ & \quad - \text{ if } h_5 \neq 0, v = v^E \text{ and } h_4 = 0; \\ & \quad - \text{ if } h_4 = h_5 = 0, v \text{ is any generalized variance measure.} \end{aligned} \tag{13}$$

Proof: see Appendix.

I now examine the properties of three particular members of **G***.

Proposition 3: *the function* $g^M(\delta s) \equiv e^{\delta s} - 1$ *is a measure of expected return and has the Aggregation Property.*

Proof: The Aggregation Property follows immediately from Proposition 2 with $h_2 = 1$, and $h_1 = h_3 = h_4 = h_5 = 0$. The implied characteristic is $\mathbb{E}_0\left[S_T/S_0 - 1\right]$ and is the mean return. ■

With g^M the *implied return* is zero. The *realized return* is the sum of daily net returns over the month. The swap contract corresponding to g^M takes the form of a standard equity for floating swap – the receiver receives the total return on the underlying (net of the riskless interest rate) each "day" on a fixed nominal amount. The payer can hedge perfectly by going long $1/S_i$ forward contracts each day.

3.3 Variance of Returns

Proposition 2 can also be used to develop a measure of variance that aggregates perfectly, even in the presence of jumps.

Proposition 4: *the function* $g^V(\delta s) \equiv 2\left(e^{\delta s} - 1 - \delta s\right)$ *is a measure of variance and has the Aggregation Property.*

Proof: The Aggregation Property follows from Proposition 2 with $h_1 = -2$, $h_2 = +2$, $h_3 = h_4 = h_5 = 0$. $g^V(x) = x^2 + O\left(x^3\right)$, so g^V is a generalized variance measure. ■

With this unconventional definition of variance, the *implied variance* at time t, IV_t, is the price of a contract that pays $g^V(s_T - s_0)$, so $IV_t = v_t^L$. The *realized variance* is $RV_t = \sum^{\mathbf{T}} g^V(\delta s) = \sum^{\mathbf{T}} 2\left(e^{\delta s} - 1 - \delta s\right)$. The corresponding definition of *true variance* is $TV_t = \mathbb{E}_t\left[2\left(S_T/S_0 - 1 - \ln\left(S_T/S_0\right)\right)\right]$ where the expectation is under the physical or objective measure, rather than under some risk-adjusted measure. The true variance is unobservable; the realized variance is an unbiased estimate of the true variance under the assumption that the price process is martingale; the implied variance is an unbiased estimate of the true variance in the absence of a variance risk premium.

The definition of implied variance is the same as the standard MFIV but differs from the definition of Bakshi, Kapadia and Madan (2003) which is the variance (conventionally defined) of the risk neutral distribution of log returns. The realized variance is the same as in Bondarenko (2007), but differs from the more conventional definition, $\sum^{T} \delta s^2$. The conventional definition has the merit that it corresponds to the definition used in practice in the variance swap market. But, as noted by Jiang and Tian (2005; see particularly footnote 7), the replication of a standard variance swap is imperfect. The replication is only perfect in the limit case when the mesh of the partition goes to zero, and with the added assumption that the price process is continuous. By contrast, the fact that the new measure of realized variance has the Aggregation Property means that replication is perfect for every price path and every partition, and in particular is robust to jumps.

In practice, with reasonably frequent rebalancing, the two measures of realized variance are very similar. This is not surprising since they are both generalized measures of variance. Suppose one computes the monthly realized volatility of the S&P500 using daily returns. Using either measure, the mean volatility over the last ten years (2001-2010) is just over 18% (annualized). The root mean square difference between the two is 0.06%. Since the 1950s the biggest difference between the two measures was in the month of October 1987 when the conventional measure of realized volatility was 101.2%, while the alternative measure registered 98.8%.

There do not appear to be any strong theoretical arguments for preferring the conventional definition of realized variance (apart from the fact that it is well established both in the academic and practitioner communities). The main justification for the conventional measure of realized variance is that it converges to the quadratic variation as the mesh size becomes small. The quadratic variation is important because it is an unbiased estimate of the conditional variance of the log price process under certain conditions (see Andersen, Bollerslev, Diebold and Labys, 2003, Theorem 1 and Corollary 1).[5] But, as the following Proposition states, *RV* also has this property when the process is a diffusion.

Proposition 5: *Let f be an analytic function on the real line such that* $\lim_{x \to 0} f(x)/x^2 = 1$. *For any continuous semimartingale s, the associated realized variance* $\sum^{\mathbf{T}} f(\delta s)$ *converges in probability to the quadratic variation of s over the period [0, T] as the mesh of the partition* **T** *goes to zero.*

Proof: see Appendix.[6]

The realized variance *RV*, with $f = g^V$, like the standard definition, converges to the quadratic variation as the mesh becomes fine; it has the added advantage that, with finite mesh, it is an unbiased estimate of the quadratic variation under the assumption that the price is a martingale. In the empirical sections of this paper, I will use the term variance in the sense of Proposition 4.

3.4 The Third Moment of Returns

Proposition 2 also shows the way to construct a definition of the realized third moment of returns that closely resembles the definition already established for price changes.

Proposition 6: $g^Q(\delta s, \delta v^E) \equiv 3\delta v^E(e^{\delta s} - 1) + K(\delta s)$ *, where* $K(\delta s) \equiv 6(\delta s e^{\delta s} - 2e^{\delta s} + \delta s + 2)$*, approximates the third moment of log returns and has the Aggregation Property.*

[5] The key condition is that the predictable component of returns is predetermined. If volatility is stochastic then this condition is certainly violated under any pricing measure, and cannot be expected to hold under the physical measure, and the quadratic variation is not an unbiased estimator of the conditional variance of log returns. An implication of Propositions 4 and 5 together is that, if the price process is continuous, the quadratic variation of returns is an unbiased estimate of the logarithmic variance $\mathbb{E}_0[L(s_T - s_0)]$, under any pricing measure.

[6] I am much indebted to Eberhard Mayerhofer for the proof of this Proposition.

Proof: g^Q has the Aggregation Property by Proposition 2 with $h_1 = 6$, $h_2 = -12$, $h_3 = -3$, $h_4 = 0$, and $h_5 = 3$. $g^Q\left(s_T - s_t, v_t^E - v_T^E\right) = -3v_t^E\left(S_T/S_t - 1\right) + K\left(s_T - s_t\right)$. With the price following a martingale, the first term is zero in expectation. $K\left(x\right) = x^3 + O\left(x^4\right)$ and converges to the third moment of returns when x is small. ■

The *implied third moment*, ITM_t, is the price of a claim that pays $g^Q\left(s_T - s_t, -v_t^E\right)$ which can be replicated exactly from forward contracts, entropy contracts and log contracts. The price of the claim is

$$ITM_t = 3\left(v_t^E - v_t^L\right). \tag{14}$$

In a Black-Scholes world, all options trade on the same implied volatility, the log and entropy variances are equal and $ITM_t = 0$. As Bakshi and Madan (2000) show, any general claim can be replicated by a portfolio of vanilla options; the replicating portfolio for the third moment claim is

$$6\left\{\int_{S_t}^{\infty} \frac{k - S_0}{S_t k^2}\mathbf{C}(k)dk - \int_0^{S_t} \frac{S_t - k}{S_t k^2}\mathbf{P}(k)dk\right\} - 3v_t^E\mathbf{F}, \tag{15}$$

where $\mathbf{C}(k)$ and $\mathbf{P}(k)$ denote a call and a put with maturity T and strike k, and $\mathbf{F}$ is an at-the-money forward contract with the same maturity.

The implied third moment is thus the price of a portfolio that is long out of the money calls and short out of the money puts. By dividing the implied third moment by the implied variance to the power of 3/2, the implied skewness coefficient (ISC) can be computed

$$ISC_t = \frac{ITM_t}{\left(IV_t\right)^{3/2}} = \frac{3\left(v_t^E - v_t^L\right)}{\left(v_t^L\right)^{3/2}}. \tag{16}$$

The realized third moment is $RTM_t = \sum^{\mathbf{T}} g^Q\left(\delta s, \delta v^E\right)$; it is natural then to define the realized skewness coefficient

$$RSC_t = \frac{RTM_t}{\left(RV_t\right)^{3/2}} = \frac{\sum^{\mathbf{T}}\left\{3\delta v^E\left(e^{\delta s}-1\right)+K\left(\delta s\right)\right\}}{\left(\sum^{\mathbf{T}} 2\left(e^{\delta s}-1-\delta s\right)\right)^{3/2}}. \tag{17}$$

Finally, we can define the true third moment and true skew coefficient as $TTM_t = \mathbb{E}_t\left[g^Q\left(s_T - s_t, -v_t^E\right)\right]$ and $TSC_t = TTM_t\left(TV_t\right)^{-3/2}$.

The realized third moment is an unbiased estimator of the true third moment if the underlying asset and the entropy contract defined on it follow a martingale. The realized skew coefficient is not necessarily an unbiased estimate of the true skew coefficient since the ratio of two unbiased estimators is not in general an unbiased estimate of the ratio.

A third moment swap, where the floating leg is the realized third moment and the fixed leg is the implied third moment, can be replicated perfectly by dynamic hedging, trading the entropy contract and the forward contract. To replicate the swap, it is necessary that an entropy contract with maturity *T* is traded (or equivalently, that calls or puts with maturity *T* and all possible strikes are traded, so that the entropy contract can be replicated). An agent who writes the swap, receiving fixed and paying floating, receives net

$$3\left(v_0^E - v_0^L\right) - \sum^{\mathbf{T}}\left\{3\delta v^E\left(e^{\delta s}-1\right)+K\left(\delta s\right)\right\}. \tag{18}$$

To hedge her position she needs to buy six log contracts, with terminal pay-off $6\ln S_T$. This costs $6\ln S_0 - 3v_0^L$ (by (10)). She also needs to hedge dynamically by holding a long position of $6/S_i$ entropy contracts on day *i* (contracts that have a pay-off of $S_T \ln S_T$) and also a short position in $3\left(2s_i + v_i^E + 4\right)\big/ S_i$ forward contracts. The replication of the swap is exact.

3.4 True and Realized Third Moments when Prices are not Martingales

We are interested in using the realized third moment to estimate the true third moment of returns. Proposition 6 shows that the realized third moment is an unbiased estimator of the true third moment when prices are martingale. But when prices are not martingale, it is a biased estimator. The following proposition characterizes the bias under the much weaker assumption that the process is ergodic.

Proposition 7 *The difference between the true and realized third moment depends on the cross-correlation between returns on the underlying asset and the returns on a hedged option position. Specifically*

$$TTM = \mathbb{E}\left[RTM\right] + 6\mathbb{E}\left[\sum_{u=1}^{T-1}\sum_{t=u}^{T-1}\left\{\frac{\delta E_u - \Delta_0 \delta S_u}{S_0}\frac{\delta S_{t+1}}{S_t} + \frac{\delta S_u}{S_0}\frac{\delta E_{t+1} - \Delta_t \delta S_{t+1}}{S_t}\right\}\right] \tag{19}$$

where E_t is the price of the entropy contract at time *t*, and $\Delta_t \equiv \frac{\partial E_t}{\partial S_t} = 1 + \ln S_t + \frac{1}{2} v_t^E$.

If changes in the term structure of volatility are parallel, and the correlation between returns on the underlying asset on day t and returns on a fixed maturity delta-hedged entropy contract on day t+n is $\rho_{ry}(n)$then

$$\frac{TTM}{\mathbb{E}\left[RTM\right]} \approx \frac{\sum_{n=-T}^{T} w_{TM}\left(n,T\right)\rho_{ry}\left(n\right)}{\rho_{ry}\left(0\right)} \text{ where } w_{TM}\left(n,T\right) \equiv \frac{\left(T-|n|\right)\left(T-n-1\right)}{T\left(T-1\right)}. \tag{20}$$

The difference between the true variance of a price series and its realized variance depends on the auto-correlation function of returns, $\rho_{rr}\left(u\right)$.

$$\frac{TV}{\mathbb{E}\left[RV\right]} \approx \sum_{n=-T}^{T} w_V\left(n,T\right)\rho_{rr}\left(n\right) \text{ where } w_v\left(n,T\right) \equiv \frac{\left(T-|n|\right)}{T}, \tag{21}$$

Proof: *in Appendix.*

The first part of the proposition says that if the expected return on an asset is positively (negatively) correlated with its variance risk premium in prior or in subsequent periods, then the realized third moment is a downward (upward) biased estimate of the true third moment. Under the hypothesis that the underlying asset and the entropy contract are martingales, the expected return on the asset and its variance risk premium are zero, and hence there is no bias.

The second part expresses the bias in terms of the cross-correlation between the expected return and the variance premium. The assumption that shifts in term structure are parallel allows us to approximate the return on an option contract with a maturity that declines in time by a duration matched constant maturity option.

The third part, concerning the relation between realized and true variance is well-known in the literature (see Campbell, Lo and MacKinlay (1996) page 49). The relationship between realized variance estimates computed using different horizons has long been used as a test for serial correlation in returns (Lo and MacKinlay, 1988).

4. Simulations

As we saw in section 1, direct estimates of the skewness of monthly returns are very noisy even with many years of data. Realized skewness, which makes use of higher frequency data, offers the possibility of measuring skewness with much greater precision. But Proposition 7 shows that the estimates will be biased in the presence of equity and variance risk premia. In this section we use simulation to get a sense of the bias and precision that may occur in practice.

The process that is simulated is the SVCJ stochastic volatility model of Duffie, Pan and Singleton (2000) with contemporaneous jumps in the underlying and volatility

$$\frac{dS_t}{S_t} = \gamma dt + \sqrt{V_t} dW_t^S + \left(e^{Z_t^S} - 1\right) dN_t$$
$$dV_t = \kappa\left(\theta - V_t\right) dt + \sigma_V \sqrt{V_t} dW_t^V + Z_t^V dN_t \tag{22}$$

where S is the price of the underlying, V is its spot variance, W^S and W^V are Brownian processes with constant correlation ρ, N is a Poisson process with intensity λ, Z^S is normally distributed with mean μ_S and standard deviation σ_S, while Z^V is distributed exponentially with mean μ_V.

The parameters used for the simulation, shown in Table 2, are taken from Broadie, Chernov and Johannes (2007; Table I "EJP", and Table IV). They are estimated from S&P futures and options data from 1987 to 2003. The drift γ is chosen so that S is martingale under the physical measure as well as the risk-adjusted measure. As a result, the realized variance is an unbiased estimate of the true variance. The difference in the parameters of the volatility process under the two measures implies the existence of variance risk premia that create a bias in the realized third moment.

The risk-adjusted parameters are used to compute option prices and the implied log and entropy variances. Paths for the asset price and implied variances are then simulated[7] under the physical measure, using daily increments over 200 months, each of 22 days. For each run, the distribution of the 200 monthly returns is then used to compute the sample second and third moments as well as the skewness coefficient. In addition, the monthly realised and implied moments are calculated and averaged over the 200 months.

The top panel of Table 3 shows that the sample variance of monthly returns computed over 200 months provides a fairly precise estimate of the actual variance. The unconditional variance of monthly returns implied by the parameters in Table 2 is 0.214 x 10^{-2}; the sample estimate has a standard deviation of 0.038 x 10^{-2}. By contrast, the standard deviations both of the third moment and of the coefficient of skewness are of similar magnitude to their true values, implying that 200 observations would be insufficient to reject the hypothesis that the distribution of returns is positively skewed.

[7] The QE algorithm recommended in Andersen (2007) is used for the discretization; the results are substantially the same if Euler discretization is used.

The implied moments have far smaller standard deviations than their sample counterparts; this is to be expected since they are beginning-of-month expectations rather than end-of-month out-turns. But they are strongly biased because of the risk premia in option prices as manifest in the large difference between the physical and risk-adjusted parameters in Table 2.

The realized second moment is an unbiased estimate of the unconditional second moment. This is expected since the simulated price process is martingale. It is also less noisy than the sample second moment. The realized third moment is also much less noisy than its sample counterpart, showing its value as a tool for estimating skewness. But it does exhibit substantial bias. With the implied variance being larger than the expected variance one might expect the realized third moment (which is the covariation between implied variance and returns) to be biased upwards in absolute terms. This is not the case. The reason for the downward is that the speed of mean reversion (κ) is twice as large under the risk adjusted measure as under the physical measure. This reduces the sensitivity of implied variance to a shock in instantaneous volatility, and so reduces the magnitude of the covariation between implied variance and returns.

As noted by Broadie, Chernov and Johannes (2007, p 1476) the diffusive risk premium is hard to estimate. So while it is clear that this premium may give rise to significant bias, the sign of the bias that arises in practice cannot be identified with confidence. The variance jump risk premium in this model, though substantial, does not give rise to bias in the realized skewness since it affects only the level of the implied variance, but does not affect its covariation with returns.

The implications for the estimates of the skewness coefficient follow immediately from the estimates of the second and third moments. The sample estimate is very noisy, with a standard deviation almost equal to its mean; the implied estimate is much more precise, but more than three times too large, while the realized estimate is somewhat biased (-27%) but with a standard deviation of only 9% of its mean.

The second panel is similar to the first except that it is concerned with annual returns estimated using 20 years of data. The true unconditional skewness of annual returns is somewhat lower than of monthly returns, but with only 20 rather than 200 observations, the sample estimate is now much noisier. The implied skewness is still biased, but the bias is somewhat smaller than in the monthly data. The magnitude of the bias in the realized skewness is much greater. At longer horizons, the increased speed of mean reversion in volatility greatly attenuates the covariation between returns and implied variance.

The main conclusions that can be drawn from the simulations is that realized skewness provides a far more precise estimate of skewness than does the sample skewness, but it is subject to bias if the risk neutral speed of mean reversion in volatility differs from its value under the physical measure. However, other variance risk premia, that do not affect the covariation between returns and implied volatility (such as the variance jump risk premia in the SVCJ model) have no effect on realized skewness though they do give rise to bias in the implied skewness.

5. Skewness of Equity Index Returns

5.1 The Term Structure of Skewness

The empirical analysis in this section is based on European options written on the S&P500 index traded on the CBOE obtained from OptionMetrics. The one month options mature every month, while the 3, 6 and 12 month options mature every three months. In each of the series, the first period starts in December 1997 and the last starts in September 2009. The data set includes closing bid and ask quotes for each option contract along with the corresponding strike prices, Black-Scholes implied volatilities, the zero-yield curve, and closing spot prices of the underlying. Entries with non-standard settlements are deleted.

Put and call option prices for every strike at each maturity are computed by interpolating implied volatilities between quoted strike prices using a cubic spline. Outside the quoted range, the implied volatilities at the lowest and the highest strike price are used. The Log and Entropy Contracts are synthesized from the continuum of conventional options and v^E and v^L are calculated. This is the procedure used by Carr and Wu (2009).

Table 4 shows summary statistics at different maturities. Realized variance is on average lower than implied variance, which is consistent with a positive variance risk premium, and both increase linearly with maturity as one might expect. The realized and implied third moments are negative at all maturities[8]. They increase with maturity faster than linearly. The implied third moment is on average larger (in absolute size) than the realized third moment at short maturities, but the difference appears to vanish or reverse at longer maturities. The skewness, both implied and realized, exceeds -1 on average at all maturities. Implied skewness appears to decline with maturity, whereas realized skewness on average actually increases with maturity.

Both the second and third moments, whether realized or implied, are highly skewed and variable. They are very highly (negatively) correlated with each other, with correlations in excess of -0.9. The skewness coefficients tend to be much less variable and to be distributed symmetrically. They are much less highly correlated with variance; the sign of the correlation coefficient is positive implying that the higher the variance the less skewed the distribution, but the correlations are all below 0.5.

To analyze the term structure of skewness, define $Y_{t,n}$ as the realized third moment accumulated over the quarter ending in month t using options that mature at the end of month $t+n$. The Y are all negative, and are highly correlated in the cross-section, with the magnitude increasing with maturity. Table 5 reports the results of the regression

$$\frac{Y_{t,n}}{Y_{t,0}} = \left(\beta_3 D_3 + \beta_6 D_6 + \beta_9 D_9\right) e^{\tilde{\varepsilon}_{t,n}} \quad \left(n = 3,6,9\right). \tag{23}$$

where D_i is a dummy that takes the value of 1 if $n = i$, and zero otherwise, and $\tilde{\varepsilon}_{t,n}$ is an error term. The estimated value of β_3 is 2.18 implying that the realized third moment over six months is 3.18 its value over 3 months. Realized variance is linear in horizon, so the realized skewness coefficient over six months is $3.18/2^{1.5} = 1.12$ times its value over three months. The increase is statistically significant; $\log\beta_3$ would need to be below 0.60 for the skewness to be lower at six months than at three, and $t(\log\beta_3\text{-}0.60) = 3.21$.

[8] In the entire sample there are eight out of 154 months in which the realized third moment is positive, and no cases at longer maturities.

The point estimate of relative skewness increases from 1 at three months to 1.12 at six months, and to 1.15 at nine and twelve months, though the increases beyond six months are not statistically significant. The panel regression of Table 5 therefore confirms the impression given by the population averages of Table 4 that realized skewness rises with horizon up to six months, and shows no evidence of significant decline in horizons up to one year.

Before concluding that true skewness too increases with horizon, it is necessary to address the issue of bias. Proposition 7 shows how cross correlations between index returns and returns on a hedged constant maturity entropy contract can be used to estimate the size of any bias. To compute the option returns, I price a notional entropy contract whose maturity is set equal to two-thirds of the horizon of interest[9] by interpolating linearly between the entropy contracts with neighboring maturities. To put some error bounds round the point estimates, I also compute the same statistics by bootstrapping the two series, destroying their auto-correlation structure. The results are set out in Table 6.

This shows that the realized second and third moments both over-estimate the true moments at all maturities, with the effect being more important (and more significant statistically) for the variance than for the third moment, and for shorter horizons than for longer horizons. It is not possible to draw firm conclusions about bias in the skewness coefficient since it is the ratio of two biased estimates, but at the very least one can say that the data do not suggest that the realized skewness over-estimates the true skewness, and they also suggest that any bias in the moment estimates attenuates with maturity.

The analysis of realized skewness of index returns at horizons of up to one year shows that returns are strongly negatively skewed at all horizons (with a skew coefficient in excess of -1), that there is no evidence at all that skewness declines with horizon and some evidence that it actually increases for horizons up to six months.

---

[9] The figure of two thirds is chosen to ensure that the constant maturity contract has the same weighted average maturity as the entropy contracts used for estimating the third moment. The conclusions are not sensitive to the choice of maturity.

5.2 Time Variation in Skewness

Realized skewness is quite volatile. Table 4 shows that the realized skewness of the S&P500 on a quarterly horizon has averaged -1.39 with a standard deviation of 0.5. It is important to know whether this variation is predictable or whether it is just noise. To investigate this we test the power of a simple predictive model of realized skewness using implied skewness and lagged realized skewness as predictors.

Table 7 reports the regression

$$RSC_{t,t+n} = \alpha + \beta_1 ISC_{t,t+n} + \beta_2 RSC_{t-m,t} + \varepsilon_{t+n} \quad \text{for } n = 1, 3, 6 \text{ and } 12, \text{ and } m = \min\{3, n\}, \tag{24}$$

where $RSC_{t,t+n}$, the realized skew coefficient from time t to $t+n$ is forecast using $ISC_{t,t+n}$ the skew implied by the time t prices of options that expire at $t+n$, and the lagged realized skew. The table shows that there is predictable time variation in the realized skew coefficient, with between 10% and 25% of the variation in the realized skew being predictable using simple explanatory variables. The lagged realized skew enters in with positive sign for shorter maturities suggesting the possibility the skew risk premium is predictable, but the magnitude is small. The coefficient is not statistically significant at longer horizons. The regression results are similar if other measures of lag realized skew are used.

Figure 2 shows the relationship between realized and predicted skew graphically at the quarterly horizon, and demonstrates that the time variation is significant economically as well as statistically, with predicted skew varying between -1.0 and -1.8 over the period.

6. Conclusions

Skewness of returns at long horizons is important for both asset pricing and risk management, yet measuring skewness is very difficult. With almost fifty years of data the evidence that daily index returns are negatively skewed just meets standard significance levels, but even then the results are heavily influenced by the Crash of 1987.

For longer horizon returns, there is also evidence that returns are negatively skewed. But the confidence intervals are wide, and the results are insignificant in some 12 year sub-periods. There is however strong evidence that the distribution of long horizon returns can only be explained by serial dependencies in the data, and is inconsistent with an iid model of the returns process. There is a need for a methodology to measure the skewness of long horizon returns that is far more precise than simply using the sample skewness, and which respects the serial dependencies.

In the case of the second moment of returns, the problem has been addressed by using high frequency returns to compute the realized variance over long horizons. The key property of realized variance is that it is an unbiased estimate of the true conditional variance. In the paper this property has been formalized as the Aggregation Property, and it has been shown that there is a unique definition of the realized (and true) third moment under which the realized third moment is an unbiased estimate of the true third moment. We have also seen that the standard definition of realized variance needs to be slightly modified so that it too has the Aggregation Property.

The realized second and third moments of index returns are highly correlated, but there is significant variation in the realized skewness coefficient and it is predictable; implied skewness from option prices (as measured by the slope of implied volatility with strike) predicts realized skewness (which roughly corresponds to the correlation between returns and volatility shocks). The realized skewness of the market index actually increases with the horizon, at least from one to twelve months. This is inconsistent with an iid model in which the twelve month skewness would be less than 30% of the one month skewness.

There are many further research questions that could usefully be explored with the use of realized skewness. The examination of the relationship between implied variance and realized variance, and the existence of variance risk premia by Carr and Wu (2009) could readily be extended to skewness risk premia both in the equity index market (as in Kozhan, Neuberger and Schneider (2011)) and in other financial markets. The evidence on the pricing of skewness and coskewness presented by Ang, Hodrick, Xing and Zhang

(2006) could be refined using a definition of skewness with a more rigourous theoretical basis.

It would also be nice to be able to extend the analysis to higher order moments (or indeed to the cumulants of the distribution as in Backus, Chernov and Martin (2011)). This would not be straightforward; as Proposition 2 shows, the set of functions that possess the Aggregation Property is quite limited; the way forward here may be to include other traded claims in addition to those on the variance of the distribution.

APPENDIX

Proof of Proposition 1

It is straightforward to prove that all members of **G** have the Aggregation Property; all that is needed is to substitute (3) into (2). Proving the converse, that all analytic functions that have the Aggregation Property are in **G**, is more complicated.

Let $\tilde{\eta}$ be a random variable with $\mathbb{E}\left[\tilde{\eta}\right]=0$ and $\mathbb{E}\left[\tilde{\eta}^2\right]=\alpha$, and g an analytic function that has the Aggregation Property. We consider two processes with $t\in\{0,1,2\}$. The first is given by S_0 = S_1 = 0, S_2 = $\tilde{\eta}$. S is clearly martingale. The process (S, V) is (0, α), (0, α), ($\tilde{\eta}$, 0). For the Aggregation Property to hold

$$\mathbb{E}\left[g\left(\tilde{\eta},-\alpha\right)\right]=\mathbb{E}\left[g\left(0,0\right)+g\left(\tilde{\eta},-\alpha\right)\right]. \tag{A-1}$$

It follows that $g(0, 0) = 0$. The second process for S is

$$0\rightarrow\begin{cases}u & \rightarrow \quad u+\tilde{\eta} \qquad \Pr=\pi \\ d & \rightarrow \quad d \qquad \Pr=1-\pi\end{cases}, \text{ with } ud<0 \text{ and } \pi u+\left(1-\pi\right)d=0. \tag{A-2}$$

S is martingale. The process (S, V) is

$$\left(0,V_0\right)\rightarrow\begin{cases}\left(u,\alpha\right) & \rightarrow \quad \left(u+\tilde{\eta},0\right) \qquad \Pr=\pi \\ \left(d,0\right) & \rightarrow \quad \left(d,0\right) \qquad \Pr=1-\pi\end{cases}$$
$$\text{where } V_0=\pi\left(u^2+\alpha^2\right)+\left(1-\pi\right)d^2. \tag{A-3}$$

For the Aggregation Property to hold

$$\mathbb{E}\left[\pi g\left(u+\tilde{\eta},-V_0\right)+\left(1-\pi\right)g\left(d,-V_0\right)\right]=$$
$$\mathbb{E}\left[\pi g\left(u,\alpha-V_0\right)+\pi g\left(\tilde{\eta},-\alpha\right)+\left(1-\pi\right)g\left(d,-V_0\right)+\left(1-\pi\right)g\left(0,0\right)\right]. \tag{A-4}$$

Simplifying, and using the fact that $g(0, 0) = 0$, gives

$$\mathbb{E}\left[g\left(u+\tilde{\eta},-V_0\right)\right]=\mathbb{E}\left[g\left(\tilde{\eta},-\alpha\right)\right]+g\left(u,\alpha-V_0\right), \tag{A-5}$$

for arbitrary u and V_0[10]. Take the limit of (A-5) as $u \to 0$

$$\mathbb{E}\left[g\left(\tilde{\eta},-V_0\right)\right]=\mathbb{E}\left[g\left(\tilde{\eta},-\alpha\right)\right]+g\left(0,\alpha-V_0\right). \tag{A-6}$$

Take the derivative of (A-6) with respect to V_0

$$\mathbb{E}\left[g_2\left(\tilde{\eta},-V_0\right)\right]=g_2\left(0,\alpha-V_0\right), \tag{A-7}$$

where the subscript denotes the partial derivative. Now take limits as $V_0 \to \alpha$

$$\mathbb{E}\left[g_2\left(\tilde{\eta},-\alpha\right)\right]=g_2\left(0,0\right), \tag{A-8}$$

(A-8) holds for any random variable $\tilde{\eta}$ with $\mathbb{E}\left[\tilde{\eta}\right]=0$ and $\mathbb{E}\left[\tilde{\eta}^2\right]=\alpha$. So for any positive function p

$$\begin{aligned}&\int_{-\infty}^{+\infty} p\left(S\right)g_2\left(S,V\right)dS \text{ is constant provided that } \int_{-\infty}^{+\infty} p\left(S\right)dS=1,\\&\int_{-\infty}^{+\infty} Sp\left(S\right)dS=0 \text{ and } \int_{-\infty}^{+\infty}\left(S^2+V\right)p\left(S\right)dS=0.\end{aligned} \tag{A-9}$$

The Lagrangian of system (A-9), $g_2\left(S,V\right)-\lambda_1-\lambda_2 S-\lambda_3\left(S^2+V\right)$, is zero, so $g_2(S, V)$ must take the form

$$g_2\left(S,V\right)=a+B\left(V\right)S+C\left(V\right)\left(S^2+V\right), \tag{A-10}$$

[10] We require that u and V_0 can take any values within some neighborhood; we do not require that they can take any value. We derive g's properties within that neighbourhood and then use the assumption that g is analytic to extend the function to the real plane.

for some constant a and functions B and C. But substituting (A-10) back into (A-7) shows that $C(v)$ is a constant, denoted by $2c$. Integrating (A-10) gives

$$g(S,V) = aV + S\int_0^V B(W)dW + c\left(2S^2 + V\right)V + D(S), \qquad \text{(A-11)}$$

where D again is an arbitrary function. It is easy to verify that (A-11) does indeed satisfy (A-6) provided that $D(0) = 0$. Substituting it into the more general (A-5) shows that $c = 0$, and that the following must be satisfied if g is to have the Aggregation Property

$$u\int_{\alpha-V_0}^{-V_0} B(W)dW + \mathbb{E}\left[D(u+\tilde{\eta}) - D(\tilde{\eta}) - D(u)\right] = 0, \qquad \text{(A-12)}$$

for arbitrary u, V_0 and random variable $\tilde{\eta}$ with zero mean. For this to hold, differentiating (A-12) with respect to V_0 gives $B(\alpha - V_0) = B(-V_0)$, so $B(V)$ must be a constant, denoted by $3b$. Let

$$\tilde{\eta} = \tilde{\eta}^*(\kappa) \equiv \begin{cases} 1+\sqrt{\kappa} & \Pr = 1/2 \\ 1-\sqrt{\kappa} & \Pr = 1/2 \end{cases} \qquad \text{(A-13)}$$

for some $\kappa \in (0,1)$. Substitute into (A-12), divide by $\kappa/2$ and take limits as $\kappa \to 0$

$$\begin{aligned} &D''(u) - D''(0) = 6bu, \text{ so} \\ &D(u) = d_1 u + d_2 u^2 + bu^3, \end{aligned} \qquad \text{(A-14)}$$

for any d_1, and d_2. So g must take the form

$$\begin{aligned} &g(S,V) = h_1 S + h_2 V + h_3 S^2 + h_4\left(S^3 + 3SV\right) \\ &\text{where } h_1 = d_1, h_2 = a, h_3 = d_2 \text{ and } h_4 = b. \end{aligned} \qquad \text{(A-15)}$$

. ■

Proof of Proposition 2

The proof is similar to the proof of Proposition 1, but with the added problem that the form of the variance function is not known. The proof that all members of **G*** have the Aggregation Property is straightforward. We focus on the converse.

Let $\tilde{\eta}$ be a random variable with $\mathbb{E}\left[e^{\tilde{\eta}}\right]=1$ and $\mathbb{E}\left[f\left(\tilde{\eta}\right)\right]=\alpha$, and g a function that has the Aggregation Property. We consider two processes with $t\in\{0,1,2\}$. The first is given by s_0 = s_1 = 0, s_2 = η. S = e^s is clearly martingale. The process (s, v) is (0, α), (0, α), (η, 0). For the Aggregation Property to hold

$$\mathbb{E}\left[g\left(\tilde{\eta},-\alpha\right)\right]=\mathbb{E}\left[g\left(0,0\right)+g\left(\tilde{\eta},-\alpha\right)\right]. \tag{A-16}$$

It follows that $g(0, 0) = 0$. The second process for s is

$$0\rightarrow\begin{cases}u & \rightarrow & u+\tilde{\eta} & \Pr=\pi\\ d & \rightarrow & d & \Pr=1-\pi\end{cases}\text{, with } ud<0 \text{ and } \pi e^{u}+\left(1-\pi\right)e^{d}=1. \tag{A-17}$$

S is martingale. The process (s, v) is

$$\left(0,v_0\right)\rightarrow\begin{cases}\left(u,\alpha\right) & \rightarrow & \left(u+\tilde{\eta},0\right) & \Pr=\pi\\ \left(d,0\right) & \rightarrow & \left(d,0\right) & \Pr=1-\pi\end{cases}$$
$$\text{where } v_0=\pi\mathbb{E}\left[f\left(u+\tilde{\eta}\right)\right]+\left(1-\pi\right)f\left(d\right). \tag{A-18}$$

For the Aggregation Property to hold

$$\begin{aligned}&\mathbb{E}\left[\pi g\left(u+\tilde{\eta},-v_0\right)+\left(1-\pi\right)g\left(d,-v_0\right)\right]=\\ &\mathbb{E}\left[\pi g\left(u,\alpha-v_0\right)+\pi g\left(\tilde{\eta},-\alpha\right)+\left(1-\pi\right)g\left(d,-v_0\right)+\left(1-\pi\right)g\left(0,0\right)\right].\end{aligned} \tag{A-19}$$

Simplify, and use the fact that $g(0, 0) = 0$, to give

$$\mathbb{E}\left[g\left(u+\tilde{\eta},-v_0\right)\right]=\mathbb{E}\left[g\left(\tilde{\eta},-\alpha\right)\right]+g\left(u,\alpha-v_0\right), \tag{A-20}$$

for arbitrary u and $v_0 > 0$. Take the limit of (A-20) as $u \to 0$

$$\mathbb{E}\left[g\left(\tilde{\eta},-v_0\right)\right]=\mathbb{E}\left[g\left(\tilde{\eta},-\alpha\right)\right]+g\left(0,\alpha-v_0\right), \tag{A-21}$$

Take the derivative of (A-21) with respect to v_0

$$\mathbb{E}\left[g_2\left(\tilde{\eta},-v_0\right)\right]=g_2\left(0,\alpha-v_0\right), \tag{A-22}$$

where the subscript denotes the partial derivative. Now take limits as $v_0 \to \alpha$

$$\mathbb{E}\left[g_2\left(\tilde{\eta},-\alpha\right)\right]=g_2\left(0,0\right), \tag{A-23}$$

Since (A-23) holds for any random variable $\tilde{\eta}$ with $\mathbb{E}\left[e^{\tilde{\eta}}\right]=1$ and $\mathbb{E}\left[f\left(\tilde{\eta}\right)\right]=\alpha$, using the same Lagrangian argument as in (A-9), $g_2(s, v)$ must take the form

$$g_2\left(s,v\right)=a+B\left(v\right)\left(e^s-1\right)+C\left(v\right)\left(f\left(s\right)+v\right), \tag{A-24}$$

for some constant a and functions B and C. Substituting (A-24) back into (A-22) shows that $C(v)$ is a constant, which we denote by $2c$. Integrating (A-24) gives

$$g\left(s,v\right)=av+\left(e^s-1\right)\int_0^v B\left(w\right)dw+cv\left(2f\left(s\right)+v\right)+D\left(s\right), \tag{A-25}$$

where D again is an arbitrary function. It is easy to verify that (A-25) does indeed satisfy (A-21). Substituting it into the more general (A-20) shows that the following equation must be satisfied if g is to have the Aggregation Property

$$\begin{aligned}&\left(e^u-1\right)\int_{\alpha-v_0}^{-v_0} B\left(w\right)dw-2c\left\{v_0\mathbb{E}\left[f\left(u+\tilde{\eta}\right)-f\left(u\right)\right]+\left(f\left(u\right)-v_0\right)\mathbb{E}\left[f\left(\tilde{\eta}\right)\right]\right\}\\&+\mathbb{E}\left[D\left(u+\tilde{\eta}\right)-D\left(\tilde{\eta}\right)-D\left(u\right)\right]=0\end{aligned} \tag{A-26}$$

For a random variable $\tilde{\eta}$ and $p \in [0,1]$ define

$$\tilde{\eta}_p \equiv \begin{cases} \tilde{\eta} & \Pr = p \\ 0 & \Pr = 1-p \end{cases}. \tag{A-27}$$

If $\mathbb{E}\left[e^{\tilde{\eta}}\right]=1$ and $\mathbb{E}\left[f(\tilde{\eta})\right]=\alpha$ then $\mathbb{E}\left[e^{\tilde{\eta}_p}\right]=1$ and $\mathbb{E}\left[f(\tilde{\eta}_p)\right]=\alpha p$. Putting $\tilde{\eta}_p$ into (A-26) gives

$$\begin{aligned}\left(e^u-1\right)\int_{p\alpha-v_0}^{-v_0} B(w)\,dw-2cp\left\{v_0\mathbb{E}\left[f(u+\tilde{\eta})-f(u)\right]+\left(f(u)-v_0\right)\mathbb{E}\left[f(\tilde{\eta})\right]\right\} \\ +p\mathbb{E}\left[D(u+\tilde{\eta})-D(u)-D(\tilde{\eta})\right]-(1-p)D(0)=0.\end{aligned} \tag{A-28}$$

By setting p = 0, we can see that $D(0)$ = 0. Since the other terms in (A-28) are linear in the arbitrary scalar p , the first term must be so too, which implies that B is constant, and can be denoted by -b. So (A-26) can be simplified to

$$\begin{aligned}b\left(e^u-1\right)\mathbb{E}\left[f(\tilde{\eta})\right]-2c\left\{v_0\mathbb{E}\left[f(u+\tilde{\eta})-f(u)\right]+\left(f(u)-v_0\right)\mathbb{E}\left[f(\tilde{\eta})\right]\right\} \\ +\mathbb{E}\left[D(u+\tilde{\eta})-D(\tilde{\eta})-D(u)\right]=0\end{aligned} \tag{A-29}$$

Let

$$\tilde{\eta}=\tilde{\eta}^*(\kappa)\equiv\begin{cases}\ln\left(1+\sqrt{\kappa}\right) & \Pr=1/2 \\ \ln\left(1-\sqrt{\kappa}\right) & \Pr=1/2\end{cases} \tag{A-30}$$

for some $\kappa \in (0,1)$. Substitute into (A-29), divide by $\kappa/2$ and take limits as $\kappa \to 0$

$$\begin{aligned}2b\left(e^u-1\right)-2c\left\{v_0\left(f''(u)-f'(u)-2\right)+2f(u)\right\} \\ +\left(D''(u)-D'(u)-D''(0)-D'(0)\right)=0\end{aligned} \tag{A-31}$$

Since (A-31) holds for arbitrary v_0

$$c = 0, \text{ or } f''(u) - f'(u) - 2 = 0. \tag{A-32}$$

If $c \neq 0$, we can solve for f using its limit properties at 0 to give

$$f(u) = 2\left(e^u - 1 - u\right) = L(u). \tag{A-33}$$

The general solution for D from (A-31) is

$$D(u) = d_1 u + d_2\left(e^u - 1\right) + \left(8c - 2b\right) u e^u + 4cu^2 \tag{A-34}$$

where d_1 and d_2 are arbitrary scalars. Putting (A-34) into (A-25) gives

$$\begin{aligned} &g(s,v) = h_1 s + h_2\left(e^s - 1\right) + h_3 v + h_4\left(v - 2s\right)^2 + h_5 e^s\left(v + 2s\right), \text{ where} \\ &h_1 = d_1; \quad h_2 = d_2; \quad h_3 = a + b - 4c; \quad h_4 = c; \quad h_5 = 4c - b. \end{aligned} \tag{A-35}$$

Finally, substituting for g into (A-20) gives

$$\begin{aligned} &\mathbb{E}\left[g\left(u + \tilde{\eta}, -v_0\right)\right] - \mathbb{E}\left[g\left(\tilde{\eta}, -\alpha\right)\right] - g\left(u, \alpha - v_0\right) = \\ &h_4\left(4u + 2v_0 - 2\alpha\right)\left(\mathbb{E}\left[2\tilde{\eta}\right] + \alpha\right) + h_5\left(e^u - 1\right)\left(\mathbb{E}\left[2\tilde{\eta}e^{\tilde{\eta}}\right] - \alpha\right). \end{aligned} \tag{A-36}$$

For this to be zero, as required for Aggregation, one of three conditions is necessary

$$\begin{aligned} &1) \quad h_4 = h_5 = 0; \\ &2) \quad h_4 = 0, \text{ and } \mathbb{E}\left[f\left(\tilde{\eta}\right)\right] = \mathbb{E}\left[2\tilde{\eta}e^{\tilde{\eta}}\right], \text{ so } f = E; \\ &3) \quad h_5 = 0, \text{ and } \mathbb{E}\left[f\left(\tilde{\eta}\right)\right] = \mathbb{E}\left[-2\tilde{\eta}\right], \text{ so } f = L. \end{aligned} \tag{A-37}$$

■

Proof of Proposition 5

By assumption, f is an analytic function which can be written as

$$f(x) = x^2 + h(x), \quad \text{where } h(x) = \sum_{k \geq 3} a_k x^k. \tag{A-38}$$

and the convergence radius of f is infinity, so

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = 0. \tag{A-39}$$

Let X be a continuous semimartingale. Assume, without loss of generality, that $X(0) = 0$. For a function g, define the *g-variation* of X as

$$\begin{aligned} V(X,g)_t &\equiv \lim_{n\to\infty} V^{(n)}(g)_t \text{ (if it exists), where} \\ V^{(n)}(X,g)_t &\equiv \sum_{i=1}^{[nt]} g\left(X_{i/n} - X_{(i-1)/n}\right), \end{aligned} \tag{A-40}$$

where $[x]$ denotes the integer part of x. Let T_N be the following increasing sequence of stopping times

$$T_N \equiv \inf\left\{t > 0 \,\middle|\, |X_t| \geq N\right\}, \tag{A-41}$$

and define the stopped process $X_t^N \equiv X_{t\wedge T_N}$. For $k \geq 3$,

$$\begin{aligned} \sum_{i=1}^{[nt]} \left|X_{i/n}^N - X_{(i-1)/n}^N\right|^k &= \sum_{i=1}^{[nt]} (2N)^k \left|\frac{X_{i/n}^N - X_{(i-1)/n}^N}{2N}\right|^k \\ &\leq (2N)^k \sum_{i=1}^{[nt]} \left|\frac{X_{i/n}^N - X_{(i-1)/n}^N}{2N}\right|^3 \\ &= (2N)^{k-3} \sum_{i=1}^{[nt]} \left|X_{i/n}^N - X_{(i-1)/n}^N\right|^3. \end{aligned} \tag{A-42}$$

Hence

$$\left|\sum_{k=1}^{M}\left(a_k\sum_{i=1}^{[tn]}\left(X_{i/n}^{N}-X_{(i-1)/n}^{N}\right)^k\right)\right|\leq\left\{\sum_{k=1}^{M}|a_k|\left(2N\right)^{k-3}\right\}\left\{\sum_{i=1}^{[tn]}\left|X_{i/n}^{N}-X_{(i-1)/n}^{N}\right|^3\right\}. \qquad \text{(A-43)}$$

(A-39) implies that the first term on the right hand side tends to a finite limit as M tends to infinity. Denoting the limit by c_N,

$$\left|V^{(n)}\left(X^N,h\right)_t\right|\leq c_N\left\{\sum_{i=1}^{[nt]}\left|X_{i/n}^{N}-X_{(i-1)/n}^{N}\right|^3\right\}\text{ for all }N. \qquad \text{(A-44)}$$

So for each N and for all $t < T_N$ the h-variation of X is bounded. Since X is continuous by assumption, its cubic variation converges to zero in probability as n tends to infinity (see Lepingle, 1976, and Jacod, 2008). Hence, by (A-44), the h-variation of X equals 0 on $[0, T_N]$, and the f-variation of X equals the quadratic variation on the same interval. Since T_N is an increasing sequence of stopping times tending to infinity, $T_N \wedge t$ increases to t a.s.. Furthermore, for $N' > N$, we have $V(X^N, f)_t = V(X^{N'}, f)_t$ on $\left[0, T_N \wedge t\right]$, hence the f-variation of X is well defined by the sequence $V(X^N, f)_t$ and equals the quadratic variation of X. The result holds true for any function f that satisfies (A-38) and (A-39). In particular, it holds for $f = L$ and $f = E$. ■

Proof of Proposition 7

$$\begin{aligned}&g^Q\left(s_T-s_0,v_T^E-v_0^E\right)=\sum_{t=0}^{T-1}g^Q\left(\delta s_{t+1},\delta v_{t+1}^E\right)+Y\\&\text{where }Y\equiv 6\sum_{u=1}^{T-1}\sum_{t=u}^{T-1}\left\{\frac{\delta E_u-\Delta_0\delta S_u}{S_0}\frac{\delta S_{t+1}}{S_t}+\frac{\delta S_u}{S_0}\frac{\delta E_{t+1}-\Delta_t\delta S_{t+1}}{S_t}\right\}.\end{aligned} \qquad \text{(A-45)}$$

is an algebraic identity. The first part of the proposition immediately follows by taking expectations at time 0, and applying the definitions of realized and true third moment. Let $x_{u+1}\equiv\dfrac{\delta E_{u+1}-\Delta_u\delta S_{u+1}}{S_u}$ and $r_{t+1}\equiv\delta S_{t+1}/S_t$. We can approximate Y by:

$$Y \approx 6\sum_{u=1}^{T-1}\sum_{t=u}^{T-1}\left\{\operatorname{cov}\left(x_u, r_{t+1}\right)+\operatorname{cov}\left(r_u, x_{t+1}\right)\right\} = 6\sum_{u=0}^{T-1}\sum_{t=0,\neq u}^{T-1}\operatorname{cov}\left(r_{t+1}, x_{u+1}\right). \tag{A-46}$$

Now x_{u+1} is the return on a delta hedged option position with duration T-u-1

$$\begin{aligned} x_{u+1} &= \left(1+r_{u+1}\right)\ln\left(1+r_{u+1}\right)-r_{u+1}+\tfrac{1}{2}\left(1+r_{u+1}\right)\delta v_{u+1}^E \\ &\approx \tfrac{1}{2}\left\{r_{u+1}^2-\sigma_u^2+\left(T-u-1\right)\delta\sigma_{u+1}^2\right\} \end{aligned} \tag{A-47}$$

where σ_u is the daily implied Black-Scholes volatility of the Entropy Contract at time u. If shifts in the volatility term structure are parallel, then to remove the dependence of x_u on T we need to define $y_t \equiv x_t/\left(T-t\right)$ and write

$$\begin{aligned} Y &\approx 3\sum_{u=0}^{T-1}\sum_{t=0,\neq u}^{T-1}\left(T-u-1\right)\operatorname{cov}\left(r_{t+1}, \delta y_{u+1}\right) \\ &= \tfrac{3}{2}\operatorname{var}\left(y\right)\operatorname{var}\left(r\right)\sum_{n=-T,\neq 0}^{T}\left(T-n-1\right)\left(T-\left|n\right|\right)\rho_{ry}\left(n\right) \text{ where } \rho_{r\sigma}\left(n\right)\equiv\rho\left(r_t, y_{t+n}\right) \end{aligned} \tag{A-48}$$

The realized third moment is simply the diagonal term where n=0, giving

$$\frac{TTM}{\mathbb{E}\left[RTM\right]} \approx \frac{\sum_{n=-T}^{T}\left(T-n-1\right)\left(T-\left|n\right|\right)\rho_{ry}\left(n\right)}{\left(T-1\right)T\rho_{ry}\left(0\right)} \tag{A-49}$$

The third part of the proposition comes from the identity

$$\begin{aligned} &g^V\left(s_T-s_0\right)=\sum_{t=0}^{T-1}g^V\left(\delta s_{t+1}\right)+Y \\ &\text{where } Y \equiv 2\sum_{u=1}^{T-1}\sum_{t=u}^{T-1}\frac{\delta S_u}{S_0}\frac{\delta S_{t+1}}{S_t}, \end{aligned} \tag{A-50}$$

and then reasoning is as before. ■

References

Andersen, L. B. G., 2007, “Efficient Simulation of the Heston Stochastic Volatility Model”, available at *SSRN*: http://ssrn.com/abstract=946405.

Andersen, T., T. Bollerslev, F. Diebold and P. Labys, 2003, “Modeling and Forecasting Realized Volatility”, *Econometrica*, 71(2), 579-625.

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

Ang, A., R. Hodrick, Y. Xing and X. Zhang, 2006, “The Cross-Section of Volatility and Expected Returns”, *Journal of Finance*, 61(1), 259–299.

Backus, D., M. Chernov and I. Martin, 2011, “Disasters Implied by Equity Index Options”, *Journal of Finance*, 66(6), 1969-2012.

Bakshi, G., N. Kapadia, and D. Madan, 2003, “Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options”, *Review of Financial Studies*, 16(1), 101-143.

Bakshi, G., and D. Madan, 2000, “Spanning and Derivative Security Valuation”, *Journal of Financial Economics*, 55, 205-238.

Barro, R., 2009, “Rare Disasters, Asset Prices, and Welfare Costs”, *American Economic Review*, 99(1), 243-264.

Black, B., 1976, “Studies of Stock Price Volatility Changes”, *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics*, 177–181.

Bondarenko, O., 2007, “Variance Trading and the Market Price of Variance Risk”, working paper, http://tigger.uic.edu/~olegb/research/vartrad.pdf.

Britten-Jones, M., and A. Neuberger, 2000, “Option Prices, Implied Price Processes, and Stochastic Volatility”, *Journal of Finance*, 55(2), 839-866.

Broadie, M., M. Chernov and M. Johannes, 2007, “Model Specification and Risk Premia: Evidence from Futures Options”, *Journal of Finance*, 62(3), 1453-1490.

Campbell., J.Y., A.W. Lo and A.C. MacKinlay, 1996, “The Econometrics of Financial Markets”, *Princeton University Press*.

Carr, P., and L. Wu, 2007, “Stochastic Skew in Currency Options”, *Journal of Financial Economics*, 86(1), 213-247.

Carr, P., and L. Wu, 2009, “Variance Risk Premiums”, *Review of Financial Studies*, 22, 1311-1341.

Christie, A. C., 1982, “The Stochastic Behavior of Common Stock Variances—Value, Leverage and Interest Rate Effects”, *Journal of Financial Economics*, 3, 145–166.

Duffie, D., J. Pan and K. Singleton, 2000, “Transform Analysis and Option Pricing for Affine Jump-Diffusions”, *Econometrica*, 68, 1343-1376.

Harvey, C.R., and A. Siddique, 2000, “Conditional Skewness in Asset Pricing Tests”, *Journal of Finance*, 55(3), 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), 327-343.

Jacod, J., 2008, “Asymptotic Properties of Power Variations and Associated Functionals of Semimartingales”, *Stochastic Processes and their Applications*, 118, 517-559.

Jiang, G. J., and Y.S. Tian, 2005, “The Model-Free Implied Volatility and Its Information Content”, *Review of Financial Studies*, 18(4), 1305-1342.

Kim, T.. and H. White, 2004, “On More Robust Estimation of Skewness and Kurtosis”, *Finance Research Letters*, 1(1), 56-73.

Kozhan, R., A. Neuberger and P. Schneider, 2011, "The Skew Risk Premium in Index Option Prices", *AFA 2011 Denver Meetings Paper*, available at http://ssrn.com/abstract=1571700.

Kraus, A., and R. H. Litzenberger, 1976, "Skewness Preference and the Valuation of Risk Assets", *Journal of Finance*, 31(4), 1085-1100.

Lepingle, D., 1976, " La variation d'ordre p des semi-martingales", *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 36, 295-316.

Lo, A., and A. C. MacKinlay, 1988, "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test", *Review of Financial Studies*, 1(1), 41-66.

Longstaff, F., and M. Piazzesi, 2004, "Corporate Earnings and the Equity Premium", *Journal of Financial Economics*, 74(3), 401-421.

Rietz, T., 1988, "The Equity Risk Premium a Solution", *Journal of Monetary Economics*, 22(1), 117-131.

Schoutens, W., 2005, "Moment Swaps", *Quantitative Finance*, 5(6), 525-530.

Schoutens, W., E. Simons, and J. Tistaert, 2005, "Model Risk for Exotic and Moment Derivatives", in Kyprianou, A.E. et al. (eds.) "*Exotic Option Pricing and Advanced Levy Models*", Wiley, 67-97.

Xing, Y., X. Zhang, and R. Zhao, 2010, "What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns?", *Journal of Financial and Quantitative Analysis*, 45(3), 641-662.

Table 1: Skewness of Daily and Monthly S&P500 returns 1963-2011

	Period	Skewness Coefficient	p-value	[5%,95%] confidence intervals
Daily	7/1963 - 6/2011	-0.832	0.043	[-1.81 ,-0.02]
	7/1963 -6/1975	+0.145	0.818	[-0.10, +0.40]
	7/1975 -6/1987	-0.085	0.224	[-0.28, +0.10]
	7/1987 -6/1999	-3.312	0.026	[-6.17, -0.18]
	7/1999 -6/2011	-0.202	0.241	[-0.66, +0.29]
Monthly	7/1963 - 6/2011	-0.736	0.001	[-1.14, -0.29]
	7/1963 -6/1975	-0.669	0.002	[-1.03, -0.27]
	7/1975 -6/1987	-0.212	0.293	[-0.86, +0.47]
	7/1987 -6/1999	-0.868	0.240	[-1.89, +0.49]
	7/1999 -6/2011	-1.008	0.007	[-1.64, -0.20]

Returns are log excess returns, computed from CRSP data. The skewness coefficient is the standard estimate of the population skewness. The p-value is against a null of positive skewness. p-values and confidence intervals are obtained by bootstrapping 10,000 times.

Table 2: Parameter estimates used in simulations

	γ	κ	θ	σ_V	μ_V	μ_S (%)	σ_S (%)	ρ	λ (%)
$\mathbb{P}$ measure	0.015	0.026	0.54	0.08	1.48	-2.63	2.89	-0.48	0.6
$\mathbb{Q}$ measure	0.031	0.057	0.246	0.08	8.78	-5.39	5.78	-0.48	0.6

The parameters for the SVCJ model $dS_t/S_t = \gamma dt + \sqrt{V_t}dW_t^S + \left(e^{Z_t^S} - 1\right)dN_t$ where V follows $dV_t = \kappa\left(\theta - V_t\right)dt + \sigma_V\sqrt{V_t}dW_t^V + Z_t^V dN_t$. Parameters are under the physical and risk-adjusted ($\mathbb{P}$ and $\mathbb{Q}$) measures. The numbers correspond to daily percentage returns and are taken from Tables I (row "EJP") and IV of Broadie, Chernov and Johannes (2007). The drift parameter γ is chosen to make S a martingale under both measures.

Table 3: Simulation Results

	Sample	Implied	Realized
	Monthly returns		
Second Moment	0.214	0.301	0.213
(x100)	(0.038)	(0.013)	(0.026)
Third Moment	-0.044	-0.238	-0.037
(x1000)	(0.045)	(0.001)	(0.010)
Skewness	-0.427	-1.594	-0.307
	(0.367)	(0.063)	(0.030)
	Annual Returns		
Second Moment	2.46	3.81	2.44
(x100)	(0.88)	(0.03)	(0.29)
Third Moment	-1.49	-4.81	-0.81
(x1000)	(3.15)	(0.01)	(0.18)
Skewness	-0.359	-0.648	-0.194
	(0.833)	(0.006)	(0.020)

The top panel reports the results of 10,000 simulations of a stochastic volatility process with jumps as described in equation (22). Each simulation comprises 200 months of 22 trading days. For each simulation, the second and third moments and skewness coefficient of the 200 monthly returns are computed. The average of these statistics across the 10,000 simulations are then recorded in the first column, together with their standard deviation in parentheses. The second and third columns do the same for the implied statistics (using one month options) and for the realized characteristics (using daily returns).

The lower panel is similar except that it shows statistics of annual returns computed over 20 year simulated histories. A year is defined as a period of 252 trading days.

Table 4: Realized and Implied Variance and Skewness of the S&P500 1997-2009

	Sample	Implied	Realized	Correl with 2nd moment Implied	Correl with 2nd moment Realized
			Monthly returns		
Second Moment (x100)	0.23	0.47 (0.53)	0.39 (0.63)		
Third Moment (x1000)	-0.16	-0.64 (1.36)	-0.38 (1.21)	-0.954	-0.934
Skewness	-1.98	-1.90 (0.88)	-1.10 (0.92)	0.297	0.074
			Three monthly returns		
Second Moment (x100)	0.82	1.44 (1.00)	1.16 (1.84)		
Third Moment (x1000)	-0.43	-2.99 (3.22)	-2.39 (7.53)	-0.952	-0.976
Skewness	-0.58	-1.69 (0.48)	-1.39 (0.50)	0.451	0.166
			Annual returns		
Second Moment (x100)	4.11	5.50 (3.17)	4.75 (4.76)		
Third Moment (x1000)	-6.21	-15.00 (15.13)	-18.14 (29.22)	-0.916	-0.954
Skewness	-0.74	-1.10 (0.35)	-1.60 (0.51)	0.235	0.403

The implied second and third moments are calculated from quoted option prices each month on the trading day following the previous option expiry date in December 1997 to September 2009 each month for monthly expiries, and three monthly for the longer maturities. The realized statistics are calculated using daily returns over the remaining life of the option. The sample statistics are computed using returns to expiry. Each cell shows the mean with the standard deviation in parentheses. Moments are computed using the model free definitions from this paper. The correlations are with the corresponding second moment – implied with implied, and realized with realized.

Table 5: Estimation of the Term Structure of the Realized Third Moment

n	*Log* β_n		β_n	*Relative Skew*
	Estimate	*Std error*		
0	0		1.00	1
3	0.777	0.054	2.17	1.122
6	1.030	0.063	2.80	1.150
9	1.163	0.074	3.20	1.147

The table reports the results of the panel regression $\log\left(Y_{t,n}/Y_{t,0}\right) = \sum_{i=3,6,9} \log\left(\beta_i\right) D_i + \tilde{\varepsilon}_{t,n}$ where D_i is a dummy that takes the value of 1 if $n = i$, and zero otherwise, $\tilde{\varepsilon}_{t,n}$ is an error term, and $Y_{t,n}$ is the realized third moment over the quarter to time t, computed using the implied variance of options that expire at the end of month $t+n$. The standard errors are White cross-section standard errors. The relative skew is $\sum_{j \le n} \beta_j \Big/ \left(\tfrac{n}{3} + 1 \right)^{3/2}$.

Table 6: Estimated Ratio of True Moment to Realized Moment

Horizon	Variance	Third Moment
1 month	0.71 [0.87,1.13]	0.77 [0.83,1.16]
3 months	0.67 [0.76,1.23]	0.88 [0.72,1.30]
1 year	0.88 [0.51,1.42]	1.01 [0.44,1.54]

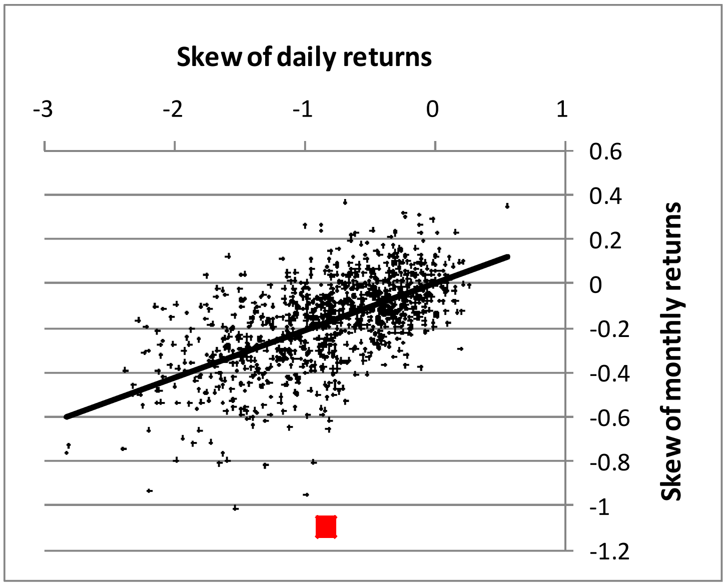
The estimates of the ratio of true variance and the third moment of returns to their realized counterparts make use of Proposition 7, and are computed from the auto- and cross-correlation structure of returns on the index and on hedged entropy contracts on the S&P500 over the period January 1996 to October 2010. The entropy contract return is based on a notional maturity equal to two thirds of the horizon, with prices obtained by linearly interpolating between the prices of the two nearest maturing entropy contracts. The numbers in square parentheses are the 5% and 95% intervals obtained by bootstrapping the returns data 1000 times. They show the confidence intervals under the null hypothesis of zero auto-correlation, and hence unbiased estimates.

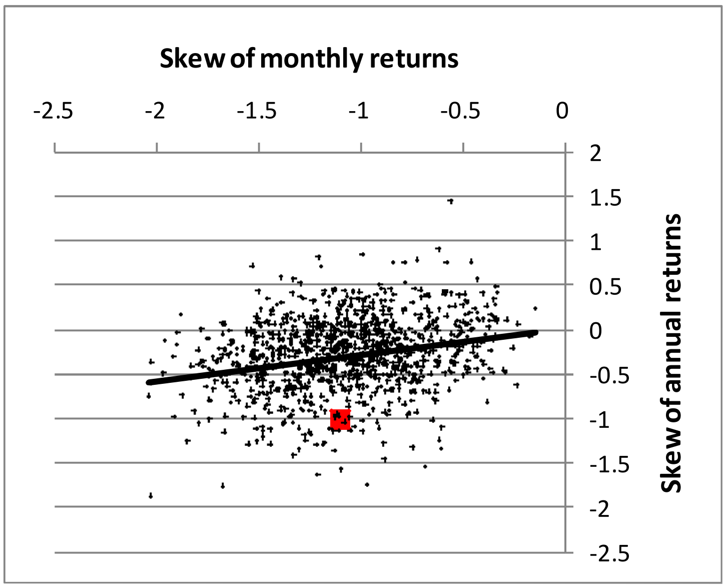
Table 7: Regression of Realized Skew Coefficient on Implied

	α	β_1	β_2	adjR^2
1 month	-0.480	0.328		9.4%
	(-2.74)	(3.70)		
	-0.240	0.297	0.264	15.9%
	(-1.49)	(3.33)	(2.46)	
3 month	-0.595	0.470		19.1%
	(-2.61)	(3.16)		
	-0.541	0.359	0.161	19.0%
	(-2.70)	(2.55)	(1.48)	
6 month	-0.364	0.838		25.5%
	(-1.30)	(4.35)		
	-0.382	0.904	-0.084	24.2%
	(-1.34)	(3.65)	(-0.72)	
12 month	-0.857	0.672		19.4%
	(-2.87)	(3.01)		
	-0.658	0.657	0.152	20.2%
	(-2.41)	(3.21)	(0.98)	

The table reports OLS results for the regression $RSC_{t,t+n} = \alpha + \beta_1 ISC_{t,t+n} + \beta_2 RSC_{t-m,t} + \varepsilon_{t+n}$ for n = 1, 3, 6 and 12 months, and m = min(n, 3) months. t–statistics in parentheses using Newey-West standard errors. Regressions are monthly for n =1, and quarterly otherwise, over the period 1998-2009.

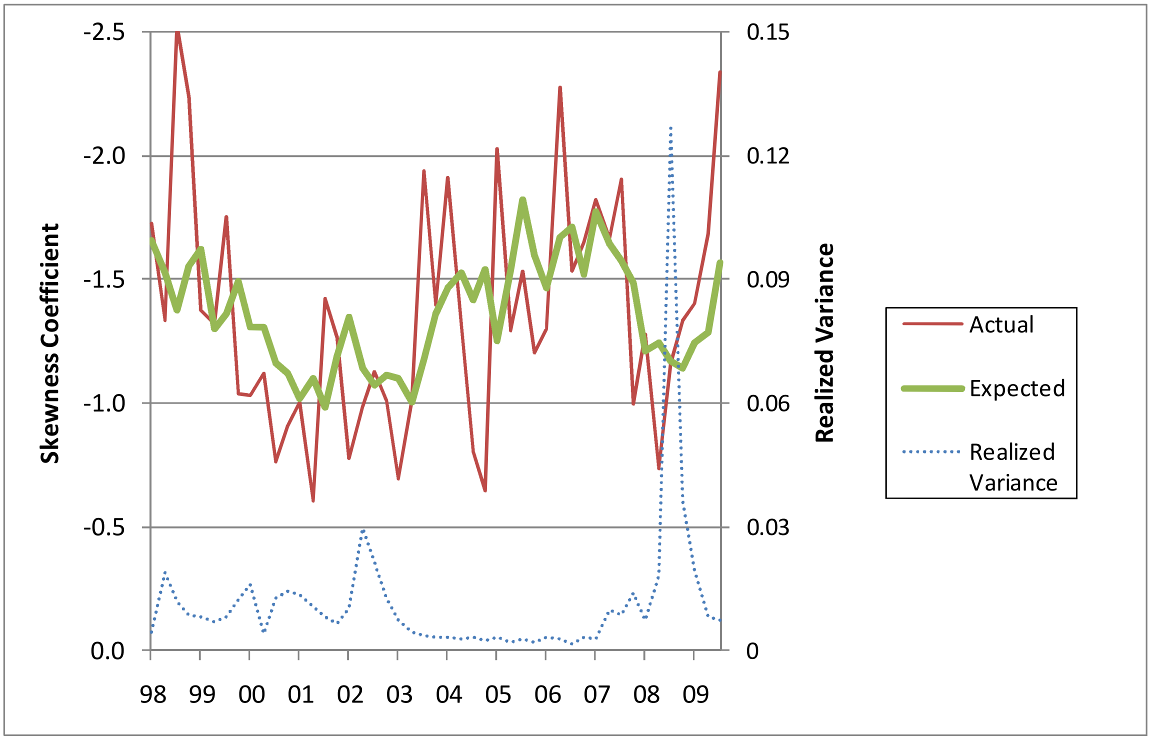
Figure 1:Skewness of index returns at different intervals





The upper plot uses bootstrapped daily returns to simulate the joint distribution of the skewness coefficient of daily and monthly log excess returns on the S&P500 index 1963-2011. The large square shows the actual skewness of daily and monthly returns over the period. The bold line plots the line $y = \sqrt{22}x$.The lower plot is similar except it shows annual returns against monthly returns and the line is $y = \sqrt{12}x$.

Figure 2: Realized Skewness of 3 monthly returns on the S&P 500



The graph shows the realized skewness coefficient of 3 monthly returns on the S&P500 computed using daily returns and implied variance changes over successive quarters from March 1998 to December 2009, together with the expected skewness coefficient at the beginning of the quarter. The expected skew is computed using the model in Table 6, where expected realized skew is a linear function of lagged realized skew and implied skew. The realized variance over the quarter is plotted on the right hand axis for comparison purposes.