Long-Term Skewness and Systemic Risk

ROBERT F. ENGLE Stern School of Business at New York University

ABSTRACT

Financial risk management has generally focused on short-term risks rather than long-term risks, and arguably this was an important component of the recent financial crisis. Econometric approaches to measuring long-term risk are developed in order to estimate the term structure of value at risk and expected shortfall. Long-term negative skewness increases the downside risk and is a consequence of asymmetric volatility models. A test is developed for long-term skewness. In a Merton style structural default model, bankruptcies are accompanied by substantial drops in equity prices. Thus, skewness in a market factor implies high defaults and default correlations even far in the future corroborating the systemic importance of long-term skewness. Investors concerned about long-term risks may hedge exposure as in the Intertemporal Capital Asset Pricing Model (ICAPM). As a consequence, the aggregate wealth portfolio should have asymmetric volatility and hedge portfolios should have reversed asymmetric volatility. Using estimates from VLAB, reversed asymmetric volatility is found for many possible hedge portfolios such as volatility products, long- and shortterm treasuries, some exchange rates, and gold. (*JEL*: G01)

KEYWORDS: ARCH, GARCH, Hedge portfolios, long-term risk, ICAPM, skewness, systemic risk

The financial crisis that engulfed the global economy in 2008 and 2009 has been described and diagnosed by many. A thoughtful discussion of the causes and remedies can be found in Acharya and Richardson (2009), Acharya et al. (2010), or Brunnermeier et al. (2009) in the Geneva Report. While there are many complex underlying causes, there are two key features of all analyses. First is the failure of many risk management systems to accurately assess the risks of financial positions,

Address to correspondence to Robert F. Engle, Stern Schoolf of Business at New York University, 44 West Fourth Street, New York, NY 0012, or e-mail: rengle@stern.nyu.edu.

doi: 10.1093/jjfinec/nbr002 © The Author 2011. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org. and the second is the failure of economic agents to respond appropriately to these risks. The explanations for the failure to act are based on distorted incentives due to institutional structures, compensation practices, government guarantees, and regulatory requirements. In short, many agents were paid well to ignore these risks. These two features are inextricably linked, and it is unlikely that we will be able to quantify the relative importance of miss-measurement of risk from the incentives to ignore risk. The wide-ranging legislative approach to reregulating the financial sector is designed to reduce the mismatch of incentives. This includes systemic risks, "too big to fail" firms, counterparty risks in Over the Counter (OTC) derivatives markets, government bank deposit guarantees as a distorting factor in proprietary trading, and many other institutional features.

In this paper, I would like to focus primarily on the question of whether risks were then or are now being measured appropriately. Was this financial crisis forecast by risk managers? Was this crisis in a 99% confidence set of possible outcomes? Was there economic or econometric information available that was not being used in risk assessment and how could risk management systems be augmented to take such factors into account?

In approaching this question, I will examine the performance of simple risk measurement systems that are widely used but will not attempt to document what was in place at what institution. The most widely used measure and one that has been frequently criticized is the value at risk (VaR) of a firm or portfolio. This is the 1% quantile of the distribution of future values and is typically defined over the next trading day. A preferable number for theoretical reasons is the expected shortfall or the loss that is expected to occur if the VaR is exceeded. Both of these measures are based on a volatility forecast, and it is natural to examine the accuracy of volatility forecasts in this economic environment.

Each of these measures is defined over a 1-day horizon. In some cases, longer horizons are used such as a 10-day horizon, but the risk measures for these longer horizons are simply scaled up on the assumption that the risk is constant. For example, the 10-day VaR is typically computed as the square root of 10 times the 1-day VaR. Both measures require statistical assumptions in order to estimate and forecast risk. The precision and stability of the measures depends upon the methodology being used.

A recent study, Brownlees, Engle, and Kelly (2009), has examined the real-time forecasting performance of volatility models for a variety of different methods and assets. These are daily models based on a range of asymmetric Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models reestimated periodically. The surprising finding is that these models showed almost no deterioration during the financial crisis. Using a statistical loss function, volatility forecasts were just as accurate during the crisis as before it, and risk management based on these forecasts was also robust to the crisis. This finding appears in conflict with the quoted observation that risk managers underestimated risk. The important issue is that both volatility forecasting and calculations of VaR and ES are computed as 1 day ahead forecasts. *The financial crisis was predictable one day ahead*.

It is immediately clear that this is not enough notice for managers who invest in illiquid assets such as the mortgage backed securities at the center of the crisis. Risk management must give more warning if it is be useful.

The short-run nature of VaR, and ES is a very important feature. Positions held for more than one day will have additional risk due to the fact that risk itself can change. Engle (2009b) describes this as the "risk that the risk will change." Most investors and financial firms hold positions much longer than one day and consequently, changes in risk will be a very important determinant of returns. A market timing approach to the holding period of assets might attempt to close positions before the risk increases, but this is bound to be unsuccessful for the average investor.

Consider the situation prior to this crisis. From 2003 until mid 2007, volatilities were very low, and from 2002 until mid 2004, short-term interest rates were very low. Thus, highly leveraged positions in a wide range of securities had low short-term risk. Many financial institutions everywhere in the globe took on these positions, but when volatilities and interest rates rose, the value of the positions fell dramatically precipitating the financial crisis. Many economists and analysts viewed the rise in interest rates and volatility as likely and the options market and fixed income market certainly forecast rises, yet our simple risk measures did not have a way to incorporate these predictions. Thus, an important challenge to risk assessment is to develop measures that can capture both short- and long-term risk.

In the next section, a term structure of risk will be estimated for various volatility models. Section 4 will develop a test for long-term skewness to see if standard volatility models are capable of modeling this characteristic of the data and the risk it generates. In Section 5, the economic underpinning of the asymmetric volatility models is developed along the lines first suggested by French, Schwert, and Stambaugh (1987). Section 6 extends this logic to hedge portfolios as in the Merton (1973) ICAPM. These hedge portfolios provide a solution to asset allocation with changing state variables. It is argued that this solution would reduce the systemic risk of the financial system if it were widely adopted.

1 TERM STRUCTURE OF RISK

To reflect the variety of investment horizons available to financial firms, this paper will follow Guidolin and Timmermann (2006) and define a term structure of risk which is a natural analogy with the term structure of interest rates and the term structure of volatility. Risk measures are computed for horizons from one day to many years which illustrate both short- and long-term risks. The measures are defined to consider the losses that would result from extreme moves in underlying asset prices over these different horizons. Major risks that will take time to unfold will lead to high long-term risks relative to short-term risks.

There are some conceptual issues in defining such risks, but the serious challenge is how to measure them. The ultimate goal is to choose investment strategies that are optimal in the face of such term structures of risk. If the value of a portfolio at time t is S_t and its log is s_t , then the loss is the difference between the date T value and the current value. Hence, the VaR at level alpha and horizon T per dollar invested is given by

$$P_t(S_{t+T}/S_t - 1 < -\operatorname{VaR}_{\alpha,T}) = \alpha.$$
(1)

Equivalently,

$$P_t(s_{t+T} - s_t < \log(1 - \operatorname{VaR}_{\alpha,T})) = \alpha.$$
⁽²⁾

The evaluation of this expression requires a model of the distribution of this asset at time t+T. Because this expression reflects the distribution of returns, it incorporates information on both expected returns and deviations from expected returns. For short horizons, it does not matter whether expected returns are considered or not but for longer horizons, this distinction is very important. After all, if expected returns are positive then long horizon returns will be proportional to T and since volatility is proportional to the square root of T, ultimately for sufficiently large T, VaR will be negative.

A variation on traditional VaR will be used here to separate forecasts of risk and return. VaR⁰ will be defined as the α quantile of losses assuming expected continuously compounded returns are zero. Expected losses from holding the asset therefore can be estimated without a model of expected returns. This definition becomes

$$P_t(s_{t+T} - E_t s_{t+T} < \log(1 - \operatorname{VaR}^0_{\alpha,T})) = \alpha.$$
(3)

Consequently, the relation between the two measures of VaR is approximately

$$\operatorname{VaR}_{\alpha,T} \cong \operatorname{VaR}^{0}_{\alpha,T} - \mu T, \tag{4}$$

where μ is the expected return. Clearly, the larger the estimate of μ , the less risky the asset appears thereby confounding the separate estimation of risk and return. Since riskier assets will generally have higher expected returns, the measure of risk and return should be constructed separately and then combined in portfolio choice.

Measures based on VaR are widely criticized from at least two additional points of view. The first is theoretical in that this measure is not coherent in the sense that it does not satisfy natural axioms of diversification. The second criticism is that by ignoring the size of the risk in the alpha tail, VaR encourages traders to take positions that have small returns almost always and massive losses a tiny fraction of the time. A measure designed to correct these flaws is ES which is defined as the expected loss should VaR be exceeded. Hence,

$$ES_{\alpha,T} = -E_t(S_{t+T}/S_t - 1|S_{t+T}/S_t < 1 - VaR_{\alpha,T})$$
(5)

and

$$\mathrm{ES}^{0}_{a,T} = -E^{0}_{t}(S_{t+T}/S_{t} - 1|S_{t+T}/S_{t} < 1 - \mathrm{VaR}^{0}_{a,T}). \tag{6}$$

Again the superscript 0 refers to the assumption that the mean logarithmic return is zero. The advantage of ES is also its disadvantage—it is difficult to estimate the losses in the extreme tail of the distribution.

An approach to estimating these losses is to use extreme value theory. In this approach, the tails of the distribution have the shape of a Pareto or Frechet distribution regardless of the underlying distribution within a wide and reasonable class of distributions. Only one parameter is required in order to have an estimate of the probability of extremes even beyond any that have been experienced. The Pareto distribution is defined by

$$F(x) = 1 - (x/x_0)^{-\lambda}, \ x \ge x_0 > 0.$$
(7)

The larger the parameter λ , the thinner the tails of the distribution. As well known, x has finite moments only for powers strictly smaller than λ , see, for example, Mc-Neil, Embrechts, and Frey (2005). The tail parameter of the student-t distribution is its degrees of freedom. The Pareto density function is

$$f(x) = \lambda (x/x_0)^{-\lambda - 1} / x_0, \quad x > x_0 > 0.$$
(8)

And the mean of x, given that it exceeds the threshold, is given by

$$E(x|x > x_0) = \frac{\lambda x_0}{\lambda - 1}, \ \lambda > 1.$$
(9)

The estimation of tail parameters is done on losses not on log returns. Thus x is interpreted as

$$x_t = 1 - S_{t+T} / S_t \tag{10}$$

Assuming that the tail beyond VaR is of the Pareto form and recognizing that it can directly be applied to the lower tail by focusing on losses as in Equation (10), an expression for the expected shortfall is simply

$$\mathrm{ES}^{0}_{_{\alpha,T}} = \mathrm{VaR}^{0}_{_{\alpha,T}} \left(\frac{\lambda}{\lambda - 1}\right). \tag{11}$$

A common approach to estimating λ is the Hill estimator. See McNeil, Embrechts, and Frey (2005). This strategy defines a truncation point x_0 such that all data exceeding this value are considered to be from a Pareto distribution. These observations are used to form a maximum likelihood estimator (MLE) of the unknown tail

parameter. Assuming all losses exceeding VaR follow such a Pareto distribution, the Hill estimator is

$$\hat{\lambda} = \frac{\sum\limits_{t \in \text{Exceed}} 1}{\sum\limits_{t \in \text{Exceed}} \log(z_t)}, \ z_t = \frac{1 - S_{T+t}/S_t}{VaR_{\alpha,T}^0}.$$
(12)

It has been shown that under some regularity conditions, see Resnick and Starica (1995), and references therein, Hill's estimator is still consistent for the tail index of the marginal distribution in a time series context if this marginal distribution is assumed to be time invariant.

2 ESTIMATING THE TERM STRUCTURE OF RISK

In order to estimate these measures for a real data set at a point in time, a simulation methodology is natural. This of course requires a model of how risk evolves over time. A familiar class of models is the set of volatility models. These models are empirically quite reliable and find new applications daily. However, they do not have a rigorous economic underpinning although some steps along this path will be discussed later in this paper. Each volatility model could be the datagenerating process for the long-term risk calculation. For a recent analysis of the forecasting performance of these volatility models during the financial crisis, see Brownlees, Engle, and Kelly (2009).

The simulation methodology can be illustrated for the simple GARCH(1,1) with the following two equations:

$$r_t = \sqrt{h_t}\varepsilon_t,$$

$$h_{t+1} = \omega + (\alpha \varepsilon_t^2 + \beta)h_t.$$
(13)

These equations generate returns, r, as the product of the forecast standard deviation and an innovation that is independent over time. The next day conditional variance also depends on this innovation and forecast standard deviation. Thus, the random shock affects not only returns but also future volatilities. From a time series of innovations, this model generates a time series of returns that incorporates the feature that volatility and risk can change over time. The distribution of returns on date T incorporates the risk that risk will change. Two simulations will be undertaken for each model, one assumes that the innovations are standard normal random variables, and the other assumes that they are independent draws from the historical distribution of returns divided by conditional standard deviations and normalized to have mean zero and variance one. This is called the bootstrap simulation. It is also called the Filtered Historical Simulation (FHS) by Barone-Adesi, Engle, and Mancini (2008).

A collection of asymmetric volatility processes will be examined in this paper but the list is much longer than this and possibly the analysis given below will suggest new and improved models. See, for example, Bollerslev (2008) for a comprehensive list of models and acronyms.

TARCH:
$$h_{t+1} = \omega + \alpha r_t^2 + \gamma r_t^2 I_{r_{t<0}} + \beta h_t$$
,
EGARCH: $\log(h_{t+1}) = \omega + \alpha |\varepsilon_t| + \gamma \varepsilon_t + \beta \log(h_t)$, $\varepsilon_t = r_t / \sqrt{h_t}$,
APARCH: $h_{t+1}^{\delta/2} = \omega + \alpha (|r_t| - \gamma r_t)^{\delta} + \beta h_t^{\delta/2}$, (14)
NGARCH: $h_{t+1} = \omega + \alpha (r_t - \gamma \sqrt{h_t})^2 + \beta h_t$,
ASQGARCH: $h_{t+1} = \omega + \beta h_t + (\alpha (r_t^2 - h_t) + \gamma (r_t^2 I_{r_t<0} - h_t/2)) h_t^{-1/2}$.

Each of these models can be used in place of the second equation of (13).

In the calculations below, parameters are estimated using data on S&P500 returns from 1990 to a specified date. These parameter estimates and the associated distribution of standardized returns are used to simulate 10,000 independent sample paths 4 years or 1000 days into the future. For each of these horizons, the risk measures described above are computed.

In Figure 1, the 1% quantile is plotted for the normal and bootstrap version of Threshold Autoregressive Conditional Heteroskedasticity (TARCH) from August 1, 2007 which is often thought to be the onset of the financial crisis. The quantile is superimposed on the S&P stock price over the next 2 years.

It is clear from this picture that the 1% quantile was very pessimistic for a year or more. However, the market minimum in March 2009 was substantially below



Figure 1 1% Quantiles of S&P500 starting August 1, 2007.

the simulation based on normal errors and approximately equal to the bootstrap simulation of the TARCH. The probabilistic statement here should be taken carefully. These bands have the property that at any horizon, the probability of exceeding the band is 1%. This does not mean that the probability of crossing the band somewhere is 1%, presumably it is substantially greater. Furthermore, by choosing the very beginning of the crisis as the starting point, I am again being nonrandom in the choice of which curves to present.

In Figure 2, the same quantiles are plotted from three other starting points, January 2008, September 2008, and June 2009, the end of the sample.

Clearly, the quantile starting in September 2008 just before the Lehman bankruptcy was quickly crossed. Now in 2010, it appears that the quantile starting in June will not be crossed at all but perhaps we should wait to see.

In comparison, several other quantile estimates are computed. Particularly interesting are the quantiles when returns are independent over time. Also computed are quantiles from EGARCH and APARCH models. These are shown in Figure 3 clustered between the TARCH and the i.i.d. quantiles.

The quantiles for i.i.d. shocks show far less risk than all the other models and they do not depend upon whether the shocks are normal or bootstrapped. The quantiles for the EGARCH and APARCH also depend little on whether the shocks are normal or bootstrapped. In fact, these models perform rather similarly.

For each of these simulations and for each horizon, the ES can be estimated empirically and the tail parameter can be estimated by the Hill estimator thus giving an alternative estimate of the ES using Equation (11).



Figure 2 1% Bootstrap Quantiles from various starting points.



Figure 3 1% Quantiles for TARCH, APARCH, EGARCH, and i.i.d. sampling.

For each method, the tail parameter is estimated for each horizon and is presented in Figure 4. The smaller the number, the fatter the tail. As can be seen, the tails for all methods become thinner as the horizon increases and the distribution gradually approaches the normal. For the bootstrapped TARCH, the tails are most extreme. When this parameter drops below four, the fourth moment does not exist. The two i.i.d. methods have distributions that converge with the number of replications to normal distributions rather than Frechet, hence, the tail parameter is not well defined and these estimates are presumably not consistent. The other methods are broadly similar which suggests that the tail parameter can perhaps be inferred from the horizon.

Based on these tail parameters and the quantiles computed above, the ES can be easily estimated. As the tail parameters should be approximately invariant to the choice of tail quantile, this provides a measure of ES that can be computed for any probability. These are shown in Figure 5.

Clearly, the highest losses are for the bootstrapped TARCH model and the lowest are for the i.i.d. models. The others are broadly similar.

3 LONG-TERM SKEWNESS

It is now widely recognized that asymmetric volatility models generate multiperiod returns with negative skewness even if the innovations are symmetric. This



Figure 4 Tail parameters at different horizons for different models.



Figure 5 ES by horizon and method computed from tail parameter.

argument is perhaps first emphasized by Engle (2004). The explanation is simple. Since negative returns predict higher volatilities than comparable positive returns, the high volatility after negative returns means that the possible market declines are more extreme than the possible market increases. In Berd, Engle, and Voronov (2007), an expression for unconditional skewness at different horizons is given. Here, a simulation will be used to estimate how skewness varies over horizon.

Skewness is defined in terms of long horizon continuously compounded or log returns as

$$sk_t(T) = \frac{E_t(s_{t+T} - s_t - \mu)^3}{[E_t(s_{t+T} - s_t - \mu)^2]^{3/2}}, \ \mu = E(s_{t+T} - s_t).$$
(15)

In a simulation context, this is easily estimated at all horizons simply by calculating the skewness of log returns across all simulations. This is naturally standardized for the mean and standard deviation of the simulation. By using log return, this measure focuses on asymmetry of the distribution. A distribution is symmetric if an x% decline is just as likely as an x% increase for any x% change. Hence, skewness is a systematic deviation from symmetry and negative skewness means that large declines are more likely than similar size increases.

For the horizons from 1 to 100 days, Figure 6 presents the results. All the skewness measures are negative and are increasingly negative for longer horizons. The simulations based on the normal distribution start at zero while the measures based on the bootstrap start at -0.4. However, after 50 days, the TARCH has fallen substantially below the other methods. The unconditional skewness of the data on SP500 returns from 1990 to 2009 is also plotted in this graph. This is also strongly negatively skewed becoming more negative as the term of the return is increased. The striking feature, is that the data are more negatively skewed than any of the models, at least between 5 and 45 days.

In Figure 7, the same plot is shown for longer horizons. Now, it is clear that the bootstrap TARCH is the most negative and the data are the least negative. It may also be apparent in the simulation that the bootstrap TARCH is very unstable. Presumably, it does not have finite sixth moments so the skewness is not consistently estimated. The other five estimators are rather similar. The most negative skewness is for approximately 1-year horizon and for still longer horizons, the skewness gradually approaches zero. However, this is extremely slow so that the skewness at a 4-year horizon is still -1.

4 TESTING LONG-TERM SKEWNESS

It is not clear that these models give accurate estimates of skewness of time aggregated returns. The data appears to be more negatively skewed for short horizons and less negatively skewed for long horizons than the models. The models are estimated by maximum likelihood but this focuses on the one step ahead



Figure 6 Skewness of different methods.



Figure 7 Skewness of different methods and horizons.

volatility forecast and may not reveal misspecification at long horizons. In this section, an econometric test is developed to determine whether the long-term skewness implied by a set of parameter estimates is significantly different from that in the data. It is a conditional moment test and has a well-defined asymptotic distribution. However, because it uses overlapping data and in some cases, highly overlapping data, it is unlikely that the asymptotic distribution is a good guide to the finite-sample performance. Thus, Monte Carlo critical values are also computed. In brief, it is found that the models cannot be rejected for misspecification of long-term skewness.

The test is based on the difference between the third moment of data simulated from the model with estimated parameters and the third moment of the data. For a horizon k and estimated parameters $\hat{\theta}$, the process is simulated N times with bootstrapped innovations to obtain a set of log asset prices $\{s_k^i(\hat{\theta})\}$. The moment to be tested is therefore

$$m_{t}^{k}(\hat{\theta}) = \frac{\left(s_{t} - s_{t-k} - \frac{1}{T}\sum_{j=k}^{T}(s_{j} - s_{j-k})\right)^{3}}{\left[\frac{1}{T}\sum_{t=k}^{T}\left(s_{t} - s_{t-k} - \left(\frac{1}{T}\sum_{j=k}^{T}(s_{j} - s_{j-k})\right)\right)^{2}\right]^{3/2}} - \frac{\frac{1}{N}\sum_{i=1}^{N}\left(s_{k}^{i}(\hat{\theta}) - \overline{s_{k}^{i}(\hat{\theta})}\right)^{3}}{\left[\frac{1}{N}\sum_{i=1}^{N}\left(s_{k}^{i}(\hat{\theta}) - \overline{s_{k}^{i}(\hat{\theta})}\right)^{2}\right]^{3/2}}$$
(16)

The expected value of this moment should be zero if the model is correctly specified and the sample is sufficiently large.

To test whether the average of this moment is zero, a simple outer product of the gradient test can be used with simulated moments. This test approach was initially used by Godfrey and Wickens (1981) and formed the core of the Berndt et al. (1974) maximization algorithm , now called BHHH. It was nicely exposited in Davidson and MacKinnon (1993, Chapter 13.7 and 16.8) and Engle (1984). It is widely used to construct conditional moment tests under the assumption that the moment condition converges uniformly to a normal mean-zero random variable with a variance that is estimated by its long-run variance.

The test is then created by regressing a vector of 1s on the first derivative matrix of the log likelihood function and the added sample moments.

$$L(\theta) = \sum_{t=1}^{T} L_t(\theta), \ G_t(\theta) \equiv \partial L_t / \partial \theta', \ 1 = G(\hat{\theta})c + M(\hat{\theta})b + \text{resid},$$
(17)

where M is the matrix of moment conditions from Equation (16). The test statistic is simply the joint test that all the coefficients b = 0.

Because of the dependence between successive moment conditions and potential heteroskedasticity, this distribution theory only applies when Heteroskedasticity and Autocorrelation Consistent Covariance (HAC) standard errors are used.



Figure 8 Critical values for long-term skewness test at 2.5% in each tail.



Figure 9 T-statistics for long-term skewness test.

Even with such technology, there is a great deal of experience that suggests that the finite-sample distribution is not well approximated by the asymptotic distribution which is a standard normal. Thus, I will compute Monte Carlo critical values as well as asymptotic critical values.

The test is done for one horizon k at a time. Parameters are estimated by MLE using the full sample and then the skewness at horizon k appropriate for these parameters is computed by simulating 10,000 sample paths of length k. The moment condition (16) can then be constructed for each data point to obtain a vector of moments M. Equation (17) is estimated by ordinary least squares (OLS) but the standard errors are constructed to correct for heteroskedasticity and autocorrelation. This is done with prewhitening and then applying Newey West with a variable bandwidth. The standard error associated with b is used to construct a *t*-statistic and this is tested using either the standard normal distribution or the Monte Carlo distribution.

To compute the Monte Carlo critical values, 1000 replications are computed with a sample size of 2000 and skewness computed in Equation (16) from 10,000 bootstrapped simulations. An upper and a lower 2.5% quantile are extracted and used as alternative critical values for the *t*-statistic. The parameters used for this Monte Carlo are close approximations to the data-based parameters.

The Monte Carlo critical values for the TARCH and EGARCH corresponding to the upper and lower tails using bootstrapped simulations are rather far from the standard normal. The critical values are given in the following table:

Thus, the critical values for TARCH models using a horizon of 10 days would be (-2.7,2.8), while for the 500-day horizon, it would be (-0.2, 10). Clearly, the asymptotic (-2,2) is not a very good approximation, particularly, for long horizons where the overlap is very large. The EGARCH is more symmetrical and somewhat closer to the asymptotic distribution at least up to 50 days.

When this test is applied for long time series on the SP500 for the five models used in this paper for the five horizons, the *t*-statistics are given in Figure 9. Notice that the largest *t*-statistic is for the 500 day overlap with TARCH. Its value of 5.5 exceeds any reasonable approximation to the asymptotic assumption of standard normality, however, from Figure 8, it is clear that the finite-sample distribution is dramatically skewed and that this value does not exceed the critical value. No other t-statistics exceed two or exceed the Monte Carlo critical values. Thus, the conclusion is that all these models appear to generate long-term skewness that is consistent with the SP data.

5 LONG-TERM SKEWNESS AND DEFAULT CORRELATIONS

In a Merton style model of default, the probability of default is the probability that the equity price will fall below a default threshold. The location of the default threshold can be estimated from the probability of default which is priced in the market for credit default swaps. In simple models designed to measure the frequency of defaults over a time period, it is often assumed that all firms are identical. Consequently, all firms must have the same market beta, idiosyncratic volatility, probability of default, and default threshold. The data-generating process is simply the one factor model

$$r_{i,t} = r_{m,t} + \sigma_i \varepsilon_{i,t}. \tag{18}$$

The distribution of defaults depends only on the properties of market returns and idiosyncratic returns. When the market return is positive, there are only defaults for companies with large negative idiosyncratic returns. When the market return is very negative, there are defaults for all companies except the ones with very large positive idiosyncrasies. The probability of default is the same for all companies, but it is much more likely for some simulations than for others. This model is developed in more detail in Berd, Engle, and Voronov (2007).

For normal and constant volatility idiosyncracies and defaults that occur whenever the stock price is below the default threshold at the end of the period, the probability of default is given by

$$\pi_{i} = \pi = P(r_{i,t} < k_{t}) = E[P(r_{i,t} < k_{t} | r_{m,t})]$$

= $E[P(\varepsilon_{i} < (k_{t} - r_{m,t}) / \sigma_{i} | r_{m,t})]$
= $E[\Phi((k_{t} - r_{m,t}) / \sigma_{i})].$ (19)

The value of k each period is chosen to satisfy Equation (19) for a set of simulations of market returns. The default correlation is defined as the correlation between indicators of default for any pair of companies. It is the covariance between these indicators divided by the variance. This can be estimated using the independence between idiosyncrasies using the following expression:

$$\rho_t^d = (P(r_{i,t} < k_t \& r_{j,t} < k_t) - \pi^2) / (\pi - \pi^2)
= (E[\Phi((k_t - r_{m,t}) / \sigma)^2 | r_{m,t}] - \pi^2) / (\pi - \pi^2)
= (V[\Phi((k_t - r_{m,t}) / \sigma)]) / (\pi - \pi^2).$$
(20)

Assuming the annual probability of default is 1%, the default frontier can be evaluated for different models by solving Equation (19) for k. Then the number of defaults calculated for any simulation. Idiosyncratic volatility is assumed to be 40% so the average correlation of two names is 18%. In Figure 10, the number of defaults that occur on the 1% worst case simulation are tabulated. This simulation corresponds to big market declines and many defaults.

In this worst case, the proportion of companies defaulting after 5 years is 40%. This bad outcome is certainly systemic. After 1 year where the probability of default is 1%, there is a nonnegligible probability that there will be 15% defaults. If the bad outcomes continue, this will rise to dramatic levels.

The default correlations in these simulations can also be computed and are tabulated in Figure 11. These correlations are highest at the first year and then gradually decline. The correlation between returns in the simulation is 18% so only the TARCH BOOT model has default correlations greater than equity correlations.

This analysis implies that a 99% confidence interval for the number of defaults 5 years out would include something like 1/3 of the existing firms when the mean default rate is only approximately 5%. For policy makers to reduce this systemic risk, either the stochastic process of market volatility must be changed or the unconditional probability of default which is assumed to be 1% per year must be reduced. Proposed capital requirements for financial firms could have that effect. In reality, the default probability is endogenously chosen by firms as they select their capital structure. A firm with less debt and more equity will have a lower probability of default. It is therefore natural for firms with lower volatility of their earnings and correspondingly lower equity volatility to choose more leverage to lower the cost of capital structure is made based on short-term volatility rather



Figure 10 1% Worst case proportion of defaults by various models.



Figure 11 Default correlations for various models.

than long-term volatility as argued in this paper, then in low volatility periods, firms will systematically choose excessive leverage and higher actual bankruptcy costs than would be optimal. Thus, an important outcome of the development of effective long-term risk measures could be a reduction in the actual default probabilities.

In practice, the identical firm model with independent idiosyncrasies may not be sufficiently rich to reveal the correlation structure. Certainly both equity correlations and default correlations are higher within industry than across industries. To encompass these observations, a richer model is needed such as the Factor Dynamic Conditional Correlation (DCC) model described in *Anticipating Correlations* by Engle (2009a) and Engle (2009c).

6 AN ECONOMIC MODEL OF ASYMMETRIC VOLATILITY

6.1 Asset Pricing

French, Schwert, and Stambaugh (1987) were the first to recognize the impact of changes in risk premia on asset prices. If volatilities are changing and risk premia depend upon expectations of future volatility, then a rise in expected future volatility should coincide with a drop in asset prices since the asset is less desirable. An important stylized fact is that volatilities and returns are very negatively correlated. This feature was tested by FSS using simple realized volatility measures and found to be important. Subsequently, there have been many papers on this topic including Campbell and Hentschel (1992), Smith and Whitelaw (2009), and Bekaert and Wu (2000). While alternative theories for asymmetric volatility based on firm leverage have received much attention including Christie (1982), Black (1976), recently Choi and Richardson (2008), I will focus on the risk premium story as it is particularly relevant for broad market indices where risk is naturally associated with volatility and where systemic risk corresponds to big market declines.

An economic model of risk premia can be based on a pricing kernel. Letting m_t represent the pricing kernel, the price of any asset today, S_t , can be expressed in terms of its cash value tomorrow x_{t+1} .

$$S_t = E_t(m_{t+1}x_{t+1}). (21)$$

The expected return can be written as

$$E_t r_{t+1} = r_t^f - (1 + r_t^f) \operatorname{Cov}_t(r_{t+1}, m_{t+1}).$$
(22)

In a one-factor model such as the Capital Asset Pricing Model (CAPM) where r^m is the single factor, the pricing kernel can be written as

$$m_{t+1} = \frac{1}{1+r_t^f} - b_t (r_{t+1}^m - E_t r_{t+1}^m).$$
(23)

Expected returns can be expressed as

$$E_t r_{t+1}^m = r_t^f + (1 + r_t^f) b_t \operatorname{Var}_t(r_{t+1}^m) = r_t^f + \delta_t h_{t+1},$$
(24)

which is the familiar Autoregressive Conditional Heteroskedasticity in Mean (*ARCH-M*) model implemented in Engle, Lilien, and Robins (1987) when δ is constant and Chou, Engle, and Kane (1992) when it is not.

To describe unexpected returns, a forward-looking model is required, and we will employ the widely used Campbell and Shiller (1988) log linearization. In this approximate identity, the difference between returns and what they were expected to be, is decomposed into a discounted sum of surprises in expected returns and in cash flows or dividends. Notice that even if dividends are perfectly predictable, there will be surprises in returns through changes in risk premia. If the risk free rate is also changing, this will be another component of innovations, however, it can be implicitly incorporated in the cash flow innovation. For this derivation, the risk free rate and coefficient of risk aversion b will be assumed constant.

$$r_{t+1} = E_t(r_{t+1}) - \eta_{r,t+1} + \eta_{d,t+1},$$

$$\eta_{r,t+1} = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+j+1},$$

$$\eta_{d,t+1} = (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j+1}.$$
(25)

Substituting Equation (24) into Equation (25) yields an expression for the changes in risk premium.

$$\eta_{r,t+1} = (E_{t+1} - E_t)\delta \sum_{j=1}^{\infty} \rho^j h_{t+j+1}.$$
(26)

It can be seen immediately from Equations (25) and (26) that returns will be negatively correlated with the innovation in the present discounted value of future conditional variance. To simplify this expression, I will assume that the variance follows a general linear process.

7 GENERAL LINEAR PROCESS VARIANCE MODELS

This framework suggests a different way to write volatility models. In some cases, this is just a restatement of the model, but it also suggests new members of the family. The idea is to write the model in terms of the variance innovations. Consider a model given by the following equations:

$$h_t = \sum_{j=1}^{\infty} \varphi_j \nu_{t-j}, \quad E_t(\nu_{t+1}) = 0, \ V_t(r_{t+1}) \equiv h_{t+1}, \ \varphi_1 \equiv 1.$$
(27)

For this class of models, the innovations given by v_t are Martingale difference sequences and are measurable with respect to r_t . That is, the innovations are nonlinear functions of r and other information from the past.

Clearly,

$$(E_{t+1} - E_t)h_{t+2} = \nu_{t+1}, (E_{t+1} - E_t)h_{t+3} = \varphi_2\nu_{t+1}, (E_{t+1} - E_t)h_{t+k} = \varphi_{k-1}\nu_{t+1}.$$

For general linear variance processes, substitution in Equation (26) yields

$$\eta_{r,t+1} = \nu_{t+1} \left[\delta \sum_{j=1}^{\infty} \rho^j \phi_j \right], \tag{28}$$

$$r_{t+1} = E_t(r_{t+1}) - \nu_{t+1} \left[\delta \sum_{j=1}^{\infty} \rho^j \phi_j \right] + \eta_{t+1}^d.$$
⁽²⁹⁾

Here, we see that the innovation in variance is negatively correlated with returns. If this innovation has substantial persistence so that volatility can be expected to be high for a long time, then the ϕ parameters will be large and the correlation will be very negative. If δ is large, indicating a high level of risk aversion, then again, the correlation will be very negative. If cash flow innovations are small, then the correlation will approach negative one.

In the common case where the variance process can be written as a first-order process,

$$h_{t+1} = \omega + \theta h_t + \nu_t, \tag{30}$$

then

$$\varphi_k = \theta^{k-1}$$
, and $\left[\delta \sum_{j=1}^{\infty} \rho^j \varphi_j\right] = \delta \left[\sum_{j=1}^{\infty} \rho^j \theta^{j-1}\right] = \frac{\delta \rho}{1 - \theta \rho}.$ (31)

This model can be estimated indirectly under several additional assumptions. Assuming a volatility process for r, the parameters of this process and the variance innovations can be estimated. From these parameters and an assumption about the discount rate, the constant in square brackets can be computed. Then the series of cash flow innovations can be identified from

$$\eta_{d,t} = r_t - E_{t-1}(r_t) + \nu_t \left[\delta \sum_{j=1}^{\infty} \rho^j \varphi_j \right].$$
(32)

Defining

$$\varepsilon_t = (r_t - E_{t-1}(r_t)) / \sqrt{h_t}, \tag{33}$$

and
$$A = \left[\delta \sum_{j=1}^{\infty} \rho^j \varphi_j\right]$$
, (34)

then Equation (32) can be rewritten as

$$\frac{\eta_{d,t}}{\sqrt{h_t}} = \varepsilon_t + \frac{\nu_t}{\sqrt{h_t}}A.$$
(35)

If ε_t is i.i.d. and $v_t / \sqrt{h_t}$ is a function only of ε_t , then the cash flow volatility process must be proportional to *h*. In this case, the covariance matrix of the innovations can be expressed as

$$E\begin{bmatrix}\varepsilon\\\nu/\sqrt{h}\\\eta_d/\sqrt{h}\end{bmatrix}\begin{bmatrix}\varepsilon\\\nu/\sqrt{h}\\\eta_d/\sqrt{h}\end{bmatrix}^T = \begin{bmatrix}1&\chi&A\chi+1\\\chi&\psi&A\psi+\chi\\A\chi+1&A\psi+\chi&A^2\psi+2A\chi+1\end{bmatrix}.$$
 (36)
$$\chi = E\left(\varepsilon\nu/\sqrt{h}\right), \quad \psi = V\left(\nu/\sqrt{h}\right)$$

Of the models considered in this paper, only the Asymmetric Square Root GARCH (ASQGARCH) satisfies the assumption that $v_t/\sqrt{h_t}$ is a function only of ε_t . This will be shown below.

In the case where this assumption is not satisfied, then the volatility process of cash flow innovations will generally be more complicated. In the GARCH family of models, the conditional variance of η_d will be the sum of a term linear in *h* and a term quadratic in *h*. The quadratic part is from the variance of *v*. If the quadratic term is small compared with the linear term, then the process will probably look like a GARCH process in any case. Thus, it is important to determine the relative importance of the risk premium and the cash flow terms.

Many asymmetric GARCH processes can be written as first-order autoregressive variance processes, see Medahi and Renault (2004). It is the case for three variance processes used in this paper: TARCH, NGARCH, and ASQGARCH. For the TARCH model,

$$h_{t+2} = \omega + \alpha r_{t+1}^2 + \gamma r_{t+1}^2 I_{r_{t+1}<0} + \beta h_{t+1}$$

= $\omega + (\alpha + \gamma/2 + \beta)h_{t+1} + [\alpha (r_{t+1}^2 - h_{t+1}) + \gamma (r_{t+1}^2 I_{r_{t+1}<0} - h_{t+1}/2)]$ (37)
= $\omega' + \theta h_{t+1} + [\nu_{t+1}].$

The innovation is given by the expression in square brackets. It is easy to see that if h_{t+1} is factored out of the innovation term, then the remainder will be i.i.d. Consequently, the variance of the variance innovation is proportional to the square of the conditional variance.

A similar argument applies to the NGARCH model:

$$h_{t+2} = \omega + \alpha (r_{t+1} - \gamma h_{t+1}^{1/2})^2 + \beta h_{t+1}$$

= $\omega + (\alpha (1 + \gamma^2) + \beta) h_{t+1} + [((\varepsilon_{t+1} - \gamma)^2 - (1 + \gamma^2)) \alpha h_{t+1}]$ (38)
= $\omega' + \theta h_{t+1} + [\nu_{t+1}], \quad \varepsilon_t = r_t / \sqrt{h_t}.$

Again, the variance of the innovation is proportional to the square of the variance.

The asymmetric SQGARCH or ASQGARCH model introduced in Engle and Ishida (2002) and Engle (2002) is designed to have the variance of the variance proportional to the variance as in other Affine models. This model can be expressed in the same form

$$h_{t+2} = \omega + (\alpha + \gamma/2 + \beta)h_{t+1} + [\alpha(r_{t+1}^2 - h_{t+1}) + \gamma(r_{t+1}^2 I_{r_{t+1}<0} - h_{t+1}/2)]h_{t+1}^{-1/2}$$

= $\omega' + \theta h_{t+1} + [\nu_{t+1}].$ (39)

This model is the same as the TARCH except that the innovation term is scaled by $h_t^{-1/2}$. The innovation can be written as

$$\nu_{t+1}/\sqrt{h_{t+1}} = [\alpha(\varepsilon_{t+1}^2 - 1) + \gamma(\varepsilon_{t+1}^2 I_{\varepsilon_{t+1}<0} - 1/2)].$$
(40)

Clearly, $v_t / \sqrt{h_t}$ is a function only of ε_t and will have a time invariant variance. If gamma is much bigger than alpha, then the asymmetry will be very important. There is no guarantee that this model will always produce a positive variance since when the variance gets small, the innovation becomes bigger than in the TARCH and can lead to a negative variance. Various computational tricks can eliminate this problem although these all introduce nonlinearities that make expectations impossible to evaluate analytically.

It is apparent that the same approach can be applied to the other models to create affine volatility models such as the Square Root Non Linear GARCH (SQNGARCH).

$$h_{t+2} = \omega + (\alpha(1+\gamma^2)+\beta)h_{t+1} + [((\varepsilon_{t+1}-\gamma)^2 - (1+\gamma^2))\alpha h_{t+1}]h_{t+1}^{-1/2}$$

= $\omega' + \theta h_{t+1} + [\nu_{t+1}], \quad \varepsilon_t = r_t / \sqrt{h_t}.$ (41)

7.1 Empirical Estimates

The ASQGARCH model estimated for the S&P500 from 1990 through November 2010 gives the following parameter estimates and *t*-statistics. Returns are in percent. An intercept in the mean was insignificant.

$$r_{t+1} = \underbrace{0.0358h_{t+1} + \sqrt{h_{t+1}}\varepsilon_{t+1}}_{(3.46)},$$

$$h_{t+1} = \underbrace{0.0224}_{(12.53)} + \underbrace{0.976h_t + \left[\underbrace{0.0049(r_t^2 - h_t) + \underbrace{0.102}_{(15.9)}(r_t^2 I_{r_t < 0} - h_t / 2)}_{(15.9)} \right] h_t^{-1/2}.$$
(42)

The variance innovations in square brackets have a mean approximately zero with a standard deviation of 0.20 which is considerably smaller than the daily standard

deviation of returns measured in percent which over this period is 1.17. However, the correlation between returns and the innovation in variance is -0.65 which is very negative. This is a consequence of the strong asymmetry of the variance innovation which follows from gamma being much bigger than alpha.

The size of the coefficient A can be constructed from these estimates and a discount rate. The persistence parameter is about 0.976 which is quite high although considerably lower than often found in GARCH estimates. The coefficient of risk aversion is estimated to be 0.0358 which is quite similar to that estimated in Bali and Engle (2010) Table II which must be divided by 100 to account for the use of percentage returns. Using an estimate of the discount rate as 0.97 annually or 0.9998 daily gives A = 0.913. With this parameter, estimates of the cash flow innovation can be computed. It has a standard deviation of 1.06 and a correlation with the variance innovation of -0.50 when weighted by conditional variance. The variance of the cash flow innovation. Thus, there are many daily news events that move returns beyond the information on risk. The estimate of the covariance matrix from Equation (36) is

$$\frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} \varepsilon_t \\ \nu_t / \sqrt{h_t} \\ \eta_{d,t} / \sqrt{h_t} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \nu_t / \sqrt{h_t} \\ \eta_{d,t} / \sqrt{h_t} \end{bmatrix}^T = \begin{bmatrix} 1 & -0.119 & .888 \\ -0.119 & .036 & -.086 \\ .888 & -.086 & .809 \end{bmatrix}.$$
(43)

The implication of these estimates is that the volatility innovations are a relatively small component of return innovations but are highly negatively correlated with returns and with cash flow dividends. Economic events that give positive news about cash flows are likely to give negative news about volatility. Once again, high volatility occurs in the same state as low earnings.

Other volatility processes will have variance innovations that are also negatively correlated with returns. Calculating these correlations over different sample periods and comparing them with log changes in VIX gives Figure 12. The



Figure 12 Correlations between returns and volatility innovations.

correlations are quite stable over time. For the VIX itself the correlation is -0.7 while for some of the volatility models it is as high as -0.8. The weighted correlations are similar.

8 HEDGING LONG-TERM RISKS IN THE ICAPM

The analysis presented thus far shows that the one factor model naturally implies asymmetric volatility in the factor. This asymmetric volatility process means that long horizon returns will have important negative skewness. This negative skewness in the factor means that the probability of a large sustained decline in asset prices is substantial. Such a tail event may represent systemic risk.

Naturally, investors will seek to hedge against this and other types of longterm risk. Financial market instability, business cycle downturns, rising volatility, short-term rates, and inflation are all events that investors fear. Even risks in the distant future such as global overheating and insolvency of social security or medicare can influence asset prices. This fear induces additional risk premia into the financial landscape. Portfolios that provide hedges against these longterm risks become desirable investments, not because they are expected to outperform, but because they are expected to outperform in the event of concern. These might be called defensive investments. I will develop a pricing and volatility relationship for these hedge assets and show how it differs from conventional assets.

Merton (1973) has suggested that changes in the investment opportunity set can be modeled with additional state variables and consequently additional factors. These factors are hedge portfolios that are designed to perform well in a particular state. Some examples help to clarify the issues. Investors often invest in gold to hedge both inflation and depression. Investments in strong sovereign debt and currency are considered to be hedges against financial crisis and global recession. Investments in volatility as an asset class are natural hedges against any recession that leads to rising volatilities. Similarly, Credit Default Swap (CDS) provide simple hedges against deteriorating credit quality and systemic risks. Any of these hedge portfolios could be shorted. In this case, the shorted portfolios will have higher risk premiums as they increase the risk in bad times.

Consider the ICAPM of Merton where there are a set of state variables and a set of k factors that can be used to hedge these long-term risks. The pricing kernel has the form

$$m_t = a_t + b_{1,t} f_{1,t} + \dots + b_{k,t} f_{k,t}.$$
(44)

The factor reflecting aggregate wealth portfolio will have a coefficient that is negative. I show next that hedge portfolios will have coefficients b that are positive, at least in the leading cases. The pricing relation is

$$1 = E_t(m_{t+1}(1+r_{t+1})),$$

$$E_t(r_{t+1}) = r_t^f - (1+r_t^f)(\operatorname{cov}_t(m_{t+1}, r_{t+1}))$$

$$= r_t^f - (1+r_t^f) \sum_{i=1}^k b_{i,t} \beta_{i,t} \operatorname{var}_t(f_{i,t+1}),$$

$$\beta_{i,t} = \operatorname{cov}_t(f_{i,t+1}, r_{t+1}) / \operatorname{var}_t(f_{i,t+1}).$$
(45)

If the bs and betas are constant over time, then the risk premia of all assets will depend only on the conditional variances of the factors.

Consider now the pricing of the second factor in a two-factor model. The expected return is given by

$$E_t(f_{2,t+1}) = r_t^f - (1 + r_t^f)(b_{1,t} \operatorname{cov}_t(f_{1,t+1}, f_{2,t+1}) + b_{2,t} \operatorname{var}_t(f_{2,t+1})), \quad (46)$$

where $b_1 < 0$. In a one-factor world, this would be a conventional asset priced by setting $b_2 = 0$. The risk premium will be proportional to the covariance with the first factor. If this asset provides a desirable hedge, then it should have a lower expected return and a higher price than if $b_2 = 0$. This can only happen if $b_2 > 0$. Thus, the pricing kernel for a two-factor model will have a positive coefficient for the second asset. The risk premium associated with the variance of this asset is negative. Of course, a short position in this asset will have a conventional positive risk premium for taking the risk associated with the second factor as well as the covariance with the first factor.

With multiple risks and factors, the same analysis can be applied to the last factor. If it hedges some event that is of concern to investors that is not already hedged, then it should have a risk premium that is based in part on its covariance with other factors, and then is reduced by its own variance.

Equation (46) provides an econometric setting to examine the process of an asset that can be used as a hedge. An innovation to the conditional variance of the hedge asset will reduce its risk premium and increase its price, holding everything else equal. That is, when the risk or variance of the hedge portfolio is forecast to increase, the hedge becomes more valuable and its price increases. Consequently, increases in variance will be correlated with positive returns rather than negative returns. This volatility is asymmetric but of the opposite sign and might be called reverse asymmetric volatility. Now, positive returns have a bigger effect on future volatility than negative returns.

This can be derived more formally using the Campbell Shiller log linearization as in Equation (25) where discounted variance innovations provide a model of unexpected returns. The important implication of this analysis is that the asymmetric volatility process of a hedge portfolio should have the opposite sign as that of a conventional asset. The long-term skewness of hedge portfolios will therefore be positive rather than negative.

This argument is only strictly correct when the covariance between factors is held constant. Normally, when a variance for one asset is forecast to increase, then the covariance between this asset, and a second asset should increase unless the correlation falls or the variance of the second asset falls. A positive shock to covariances will increase the risk premium and reduce the price movement thereby offsetting some of the direct effect. Thus, the empirical ability to see this asymmetry is limited by the natural offsetting covariance effects. Furthermore, the volatility innovations in the first asset will influence the covariance in a way that is simply noise in a univariate model.

To examine this empirical measure, I look at estimates for almost 150 assets which are followed every day in VLAB. This can easily be seen at www.vlab.stern.nyu.edu. The assets include equity indices, international equity indices, individual names, sector indices, corporate bond prices, treasuries, commodities, and volatilities, as well as currencies. The names of the series are given in Table 1. These are sorted by the estimated value of Gamma, the asymmetric volatility coefficient in the TARCH model. When this coefficient is negative, then positive returns predict higher volatility and when it is positive, negative returns predict higher volatility.

From the argument given above, hedge portfolios should have negative gamma. From the list, it is clear that many of the assets naturally thought to be hedge assets do indeed have negative gamma. This is true of eleven out of thirteen volatility series. These series are constructed like the Market Volatility Index (VIX) but on a variety of assets including international equities and commodities. In each case, these assets will be expected to increase dramatically in value in another financial crisis. The next highest asset on the list is GOLD. It has a negative gamma.

US Treasuries appear to be a hedge, whether at the short or long end of the maturity. The swap rate variables are approximately converted to returns by using the negative of the log change in yield. In this way, these look very similar to the bond indices published by Barclays at the short, intermediate or long end of the term structure.

Many exchange rates have negative gamma. As these are measured as dollars per foreign currency, these should increase in a crisis if the foreign currency is expected to weaken in a future crisis. We see that relative to the Kenyan, Japanese, Thai, Russian, and Swiss currencies, the volatility is asymmetric. Relative to many others, the exchange rate has the conventional positive sign. For most of these other exchange rates, the coefficient is close to zero.

The large positive coefficients are typically broad-based equity indices, both US and global indices. This is consistent with the economic observation that indices that approximate the global wealth portfolio must have a positive gamma so that returns and volatilities are negatively correlated.

The evidence that these assets are priced as Merton hedge portfolios suggests that important parts of the investment community do indeed respond to long-run risks by taking positions in hedge portfolios. The high cost of volatility derivatives is further evidence of the value of these contracts in protecting against various forms of tail risk.

Name	Gamma
AEX volatility index return series	-0.14403
CBOE volatility index (OLD) return series	-0.14172
SMI volatility index return series	-0.1354
CBOE NASDAQ-100 volatility index return series	-0.13041
CBOE volatility index return series	-0.126
CBOE DJIA volatility index return series	-0.12559
FTSE 100 volatility index return series	-0.1003
iShares gold return series	-0.04837
US Dollar to Kenyan Shilling return series	-0.04168
Kospi volatility index return series	-0.0404
DAX volatility index return series	-0.03757
5Y interest rate swap return series	-0.03625
iShares Barclays 1–3 year return series	-0.03567
10Y interest rate swap return series	-0.03548
2Y interest rate swap return series	-0.03479
US Dollar to Japanese Yen return series	-0.03309
CBOE crude oil volatility index return series	-0.02868
US Dollar to Thai Baht return series	-0.02313
iShares Barclays 10–20 year return series	-0.01594
US Dollar to Swiss Franc return series	-0.01416
Barclays US Agg. Government return series	-0.00958
US Dollar to Russian Rouble return series	-0.00932
BEL 20 volatility index return series	-0.00741
iShares Barclays short return series	-0.00553
NYSE Euronext return series	-0.00422
Barclays US Agg. Credit AA return series	0.002092
CAC40 volatility index return series	0.003436
US Dollar index return series	0.003664
US Dollar to British Pound return series	0.00737
iShares Silver return series	0.008139
US Dollar to EURO return series	0.008159
US Dollar to Canadian Dollar return series	0.008498
iShares S&P GSCI commodity-indexed trust return series	0.014338
Bovespa return series	0.014887
Intel return series	0.017776
	(continued)

 Table 1
 Asymmetric volatility models from VLAB November 29, 2010.

(continued)

Table 1 (continued)

Name Ga	amma
US Dollar to Australian Dollar return series 0.	.020536
Hewlett-Packard return series 0.	.021733
Barclays US Agg. credit BAA return series 0.	.02187
US Dollar to Singapore dollar return series 0.	.02245
MSCI Canada return series 0.	.023255
Alcoa return series 0.	.028223
Wal Mart stores return series 0.	.028549
Caterpillar return series 0.	.031276
AT&T return series 0.	.033125
Du Pont return series 0.	.033618
Charles Schwab return series 0.	.034797
McDonalds return series 0.	.03503
US Dollar to South Korean Won return series 0.	.036069
Blackrock return series 0.	.037161
CBOE Gold volatility index return series 0.	.039308
Bank Of America return series 0.	.040501
ML HY Cash Pay C All return series 0.	.042471
General Motors return series 0.	.042776
Boeing return series 0.	.043467
CME Group return series 0.	.047192
Franklin Resources return series 0.	.04838
Walt Disney return series 0.	.048408
General Electric return series 0.	.048818
MBIA return series 0.	.048849
Citigroup return series 0.	.049284
Procter & Gamble return series 0.	.049897
iShares Cohen & Steers Realty Majors return series 0.	.050613
US Bancorp return series 0.	.052038
Exxon Mobil return series 0.	.052686
iShares Barclays MBS return series 0.	.052709
MSCI FTSE/Xinhua China 25 return series 0.	.054102
Coca Cola return series 0.	.055971
Goldman Sachs return series 0.	.060912
MSCI Taiwan return series 0.	.062269
American Express return series 0.	.063519
Materials SPDR return series 0.	.064714
Korea composite stock index return series 0.	.064814
Wells Fargo return series 0.	.064971
Merck & Co. return series 0.	.065279
iShares Dow Jones U.S. Real Estate return series 0.	.065537
MSCI Japan return series 0.	.065731
SPDR S&P Metals & Mining return series 0.0	.066303
Budapest Stock Exchange Index return series 0.	.066486

(continued)

Table 1 (continued)

Name	Gamma
MSCI Italy return series	0.067246
MasterCard return series	0.067339
Energy SPDR return series	0.06862
JP Morgan Chase return series	0.068975
Johnson & Johnson return series	0.070134
Home Depot return series	0.070361
MSCI UK return series	0.073132
iShares iBoxx \$ High Yield Corporate return series	0.07341
Fannie Mae return series	0.073971
E-Trade return series	0.074877
Technology SPDR return series	0.076638
Utilities SPDR return series	0.077278
MSCI Sweden return series	0.078147
IBM return series	0.078657
NASDAQ Composite return series	0.079284
iShares S&P Global Energy return series	0.080881
American Internation Group return series	0.081977
Hang Seng return series	0.082658
MSCI Spain return series	0.083193
Consumer Staples SPDR return series	0.083752
Russell 2000 return series	0.085299
MSCI Germany return series	0.08642
CIT Group return series	0.086607
Consumer Discretionary SPDR return series	0.089509
MSCI Australia return series	0.090098
MSCI Switzerland return series	0.090125
US Dollar to Indian Rupee return series	0.090441
US Dollar to Indian Rupee return series	0.090441
MSCI Netherlands return series	0.091247
Morgan Stanley return series	0.091704
MSCI France return series	0.091707
MSCI Emerging Markets return series	0.094782
FTSE 100 return series	0.097009
MSCI Belgium return series	0.098846
MSCI Hong Kong return series	0.099802
CAC 40 return series	0.103081
MSCI Brazil return series	0.103465
MSCI World return series	0.103797
Industrials SPDR return series	0.104001
Prudential Financial return series	0.104334
Financials SPDR return series	0.104509
Dow Jones Industrials return series	0.105162
SPDR S&P Oil & Gas Exp. & Prod. return series	0.105526

(continued)

Table 1 (continued)

Name	Gamma
Wilshire Small Cap 250 return series	0.10599
DAX 30 return series	0.106108
iShares iBoxx \$ Investment Grade Corporate return series	0.106886
MSCI Singapore return series	0.107279
MSCI EAFE return series	0.108903
S&P 500 Composite return series	0.110646
iShares S&P New York AMT-Free Municipal return series	0.111022
MetLife return series	0.113888
MSCI BRIC return series	0.113948
Healthcare SPDR return series	0.115262
United Technologies return series	0.11589
MSCI FTSE China (HK Listed) return series	0.12127
MSCI Mexico return series	0.140785
ML HY Cash Pay B All return series	0.147277
MIB 30 return series	0.191248

9 Conclusions

This paper introduces the idea that short- and long-run risk can be separately measured and used to achieve investment goals. Failure to pay attention to long-term risk is a potential explanation for the financial crisis of 2007–2009 where low short term risk allowed financial institutions to load themselves with leverage and illiquid assets such as subprime Collateralized Debt Obligation (CDOs). Recognizing that time-varying volatility models are able to measure how fast risk is likely to evolve, the term structure of risk can be estimated by simulation. VaR, ES, and tail parameters can all be estimated. Because long-term skewness is more negative than short-term skewness for typical volatility models, this becomes a key feature of long-term risk since negative skewness makes a major decline in asset values more likely. In a Merton style model of defaults, it is shown that the negative long-term skewness can give rise to massive defaults and high correlations of defaults reinforcing the possibility of another episode of systemic risk.

Although long-term negative skewness is a property of asymmetric volatility models, there is no test for the strength of this effect. In this paper, such a test is developed as a simulated method of moments test. From the asymptotic distribution and a Monte Carlo estimate of the distribution, it is found that all the asymmetric models are consistent with the long-term skewness in the data. It does not seem that this will distinguish between models.

The economic underpinning of the asymmetric volatility model is developed following French Schwert and Stambaugh and simple asset pricing theory. The model has implications for the proportion of return volatility that is due to changes in risk and the proportion due to changes in other factors such as cash flow. Affine style models in discrete time provide a useful parameterization although this is not necessarily the only way to express this relation.

Finally, economic agents that seek to hedge their long-term risk will naturally look for Merton hedges that will outperform in bad states of nature. From standard asset pricing theory, it is possible to develop testable implications about the volatility of such hedge portfolios. Remarkably, hedge portfolios should have asymmetric volatility of the opposite sign from ordinary assets. From examining close to 150 volatility models in VLAB, it is discovered that the only ones with reverse asymmetric volatility are volatility derivatives themselves, treasury securities both long and short maturities, some exchange rates particularly relative to weak countries and gold. These are all assets likely to be priced as hedge portfolios. There are caveats in this procedure that argue for more research to investigate the power of this observation.

REFERENCES

- Acharya, V., T. Cooley, M. Richardson, and I. Walter. 2010. *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance*. New York: Wiley Publishing.
- Acharya, V., and M. Richardson. 2009. *Restoring Financial Stability: How to Repair a Failed System*. New York: Wiley Publishing.
- Bali, T., and R. Engle. 2010. The Intertemporal Capital Asset Pricing Model with Dynamic Conditional Correlations. *Journal of Monetary Economics* 57: 377–390.
- Barone-Adesi, G., R. Engle, and L. Mancini. 2008. A GARCH Option Pricing Model with Filtered Historical Simulation. *Review of Financial Studies* 21: 1223–1258.
- Bekaert, G., and G. Wu. 2000. Asymmetric Volatility and Risk in Equity Markets. *Review of Financial Studies* 13: 1–42.
- Berd, A., R. Engle, and A. Voronov. 2007. The Underlying Dynamics of Credit Correlations. *Journal of Credit Risk* 3: 27–62.
- Berndt, E., B. Hall, R. Hall, and J. Hausman. 1974. Estimation and Inference in Nonlinear Structural Models. *Annals of Social Measurement* 3: 653–665.
- Black, F. 1976. "Studies in Stock Price Volatility Changes." Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economics Statistics Section. pp. 177–181.
- Bollerslev, T. 2008. "Glossary to ARCH (GARCH)." CREATES Research Papers. 2008–49. School of Economics and Management. University of Aarhus.
- Brownlees, C., R. Engle, and B. Kelly. 2009. Evaluating Volatility Forecasts over Multiple Horizons. Working Paper, New York University.
- Brunnermeier, M., A. Crockett, C. Goodhart, A. Persaud, and H. Shin. 2009. "The Fundamental Principles of Financial Regulation." Geneva Reports on World Economy 11. International Center for Monetary and Banking Studies.
- Campbell, J., and L. Hentschel. 1992. No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics* 31: 281–318.

- Campbell, J., and R. Shiller. 1988. The Dividend Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies* 1: 195–228.
- Choi, J., and M. Richardson. 2008. "The Volatility of a Firm's Assets." AFA 2010 Atlanta Meetings Paper.
- Chou, R, R. Engle, and A. Kane. 1992. Measuring Risk Aversion from Excess Returns on a Stock Index. *Journal of Econometrics* 52: 201–224.
- Christie, A. 1982. The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. *Journal of Financial Economics* 10: 407–432.
- Davidson, R., and J. G. MacKinnon. 1993. Estimation and Inference in Econometrics. New York: Oxford University Press. 874 pages.
- Engle, R. 1984. Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics. *Handbook of Econometrics* 2: 776–826.
- Engle, R. 2002. New Frontiers in ARCH Models. *Journal of Applied Econometrics* 17: 425–446.
- Engle, R. 2004. Risk and Volatility: Econometric Models and Financial Practice. Nobel Lecture. *American Economic Review* 94: 405–420.
- Engle, R. 2009a. "Anticipating Correlations: A New Paradigm for Risk Management." Princeton, New Jersey: Princeton University Press.
- Engle, R. 2009b. The Risk that Risk Will Change. *Journal of Investment Management* 7: 24–28.
- Engle, R. 2009c. *High Dimension Dynamic Correlations. The Methodology and Practice of Econometrics: A Festschrift in David F. Hendry.* Oxford: Oxford University Press.
- Engle, R., and I. Ishida. 2002. "Forecasting Variance of Variance: The Square-root, the Affine and the CEV GARCH Models." New York University Working Paper.
- Engle, R., D. Lilien, and R. Robins. 1987. Estimation of Time Varying Risk Premia in the Term Structure: the ARCH-M Model. *Econometrica* 55: 391–407.
- French, K., G. Schwert, and R. Stambaugh. 1987. Expected Stock Returns and Volatility. *Journal of Financial Economics* 19: 3–30.
- Godfrey, L. G., and M. R. Wickens. 1981. Testing Linear and Log-Linear Regressions for Functional Form. *Review of Economic Studies* 48: 487–496.
- Guidolin, M. and A. Timmermann. 2006. Term Structure of Risk under Alternative Econometric Specifications. *Journal of Econometrics* 131: 285–308.
- Kelly, B. 2009. "Risk Premia and the Conditional Tails of Stock Returns." New York University Working Paper.
- McNeil, A., P. Embrechts, and R. Frey. 2005. *Quantitative Risk Management*. Princeton, New Jersey: Princeton University Press, Princeton Series in Finance.
- Medahi, N., and E. Renault. 2004. Temporal Aggregation of Volatility Models. *Journal of Econometrics* 119: 355–379.
- Merton, R. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41: 867–887.
- Resnick S. I., and C. Starica. 1995. Consistency of Hill's Estimator for Dependent Data. *Journal of Applied Probability* 32: 139–167.
- Smith, D., and R. Whitelaw. 2009. "Time-Varying Risk Aversion and the Risk-Return Relation." New York University Working Paper.