Natural disasters, insurance stocks and the numeraire portfolio

Jason West
Department of Accounting, Finance and Economics
Griffith Business School
Griffith University
Nathan QLD Australia 4111
Ph: +61 7 37354272
j.west@griffith.edu.au

Eckhard Platen
School of Finance and Economics
University of Technology, Sydney
PO Box 123
Broadway NSW Australia 2007
Ph: +61 2 9514 2271
Eckhard.Platen@uts.edu.au

Abstract

This study employs a generalized numeraire portfolio to benchmark insurance stocks to detect abnormal returns in the presence of natural disasters and insurable losses. We show that under the benchmark approach the efficient markets hypothesis holds in the presence of extreme insurable loss whereas other common methods such as the market model and Fama-French three factor model often fail due to the accumulation of estimation errors. We construct a portfolio of US insurance firms and observe the market reaction to a set of major insured natural disasters. Numeraire denominated or benchmarked returns are shown to be are natural measures of abnormal returns. Using the benchmark approach we observe no significant trend in the cumulative abnormal returns of insurance securities following a natural disaster. Using both the traditional market model and the Fama-French three factor model however, we observe significantly positive cumulative abnormal returns following an insured event. The errors inherent in the market model and three-factor model for event studies are shown to be eliminated using the benchmark approach.

JEL Classification: G10, G13
Key words: Event Study, Benchmark Model, Efficient Markets Hypothesis, Market Model, Factor Models, Growth Optimal Portfolio.

Natural disasters, insurance stocks and the numeraire portfolio

Abstract

This study employs a generalized numeraire portfolio to benchmark insurance stocks to detect abnormal returns in the presence of natural disasters and insurable losses. We show that under the benchmark approach the efficient markets hypothesis holds in the presence of extreme insurable loss whereas other common methods such as the market model and Fama-French three factor model often fail due to the accumulation of estimation errors. We construct a portfolio of US insurance firms and observe the market reaction to a set of major insured natural disasters. Numeraire denominated or benchmarked returns are shown to be are natural measures of abnormal returns. Using the benchmark approach we observe no significant trend in the cumulative abnormal returns of insurance securities following a natural disaster. Using both the traditional market model and the Fama-French three factor model however, we observe significantly positive cumulative abnormal returns following an insured event. The errors inherent in the market model and three-factor model for event studies are shown to be eliminated using the benchmark approach.

JEL Classification: G10, G13
Key words: Event Study, Benchmark Model, Efficient Markets Hypothesis, Market Model, Factor Models, Growth Optimal Portfolio.

1. Introduction

In an efficient capital market the prices of securities observed at any time are based on the correct evaluation of all information available at that time. Fama (1976) showed that prices should therefore fully reflect all available information in an efficient market. Given a set of assets, a profit opportunity arises if, by trading assets from a given set, with positive probability one can purchase a claim to a non-negative future payoff for a non-positive current price, or if one can sell a claim to a certain zero future payoff or a positive current price. Only the absence of such profit opportunities implies that the share market is efficient under Fama's assertion. Asset pricing models that have emerged from this notion of market efficiency express the idea that there is an equality of risk adjusted expected rates of return among a given set of assets.

To capitalize on this idea in an effort to simplify asset pricing models Long (1990) showed that a given set of assets offers no profit opportunities, if and only if a numeraire portfolio can be formed from the set. A numeraire portfolio is defined to be a self-financing portfolio such that, if current and future asset prices and dividends are denominated in units of the numeraire, the expected rate of return of every asset within the given set of assets is always zero. By simply dividing each asset by the contemporaneous value of the so-called numeraire portfolio, the best forecast of each asset's numeraire denominated rate of return is zero. This construction is quite general with respect to discrete or continuous time asset pricing model assumptions and return distributions.

Event studies are an important tool for measuring how security prices change in response to information following an event (Binder, 1998). Some commonly used approaches in event studies include the one-factor market model, the Fama-French three-factor model and the Carhart four-factor model. The one-factor model, developed by Sharpe (1963) and applied in Brown and Warner (1985) describes the process generating returns during a prescribed event window. Variations of the market model have been used in a number of studies to detect departures from the market efficiency assumption, and to detect other peculiar behaviours in security returns following a specific event (Chandra and Balachandran, 1990; Boehmer et al. 1991; Ang and Zhang, 2004; Nam et al. 2005; Bartholdy et al. 2007). The Fama and French (1996) three-factor model uses a market index, a size index and a book-to-market index to explain stock returns. There have been several other models that extend these approaches, the most prominent being the characteristic-based benchmark estimate of Daniel et al. (1997)and the four-factor model of Carhart (1997) that appends the Fama-French three-factor model with a short-run momentum index. In this study we confine our analysis to the performance of the market model and the Fama-French three factor model against the benchmark approach theorized in Long (1990).

Ahern (2009) showed that there was little additional explanatory power in multifactor models that extend the Fama-French three-factor approach. While multi-factor regression models may alleviate the omitted variable bias of a simple market model, they may also introduce additional estimation errors (Fama and French, 1997). The primary advantage of the benchmark approach is that no regressors need to be estimated. This helps eliminate estimation error and there is no requirement to choose a normal estimation period, either pre-or post-event, as there is in traditional approaches.

The basis for inference in event studies is the test statistic. This is generally the ratio of the mean excess return to its estimated standard deviation. The aim of the test in event studies is to detect the presence of statistically significant departures from the assumption of market efficiency. Tests of the efficient market hypothesis (EMH) are always joint tests of market efficiency, the underlying equilibrium model and a related market or factor model. Studies that reject the EMH may incorrectly do so if the errors from the market model accumulate above the rejection limit.

The benchmark approach offers a robust and reliable alternative to the traditional models used for examining investor behaviour following an event under the EMH. The approach uses the growth optimal portfolio (GOP) as the numeraire portfolio, first suggested by Kelly (1956). This approach avoids the problems associated with bias and efficiency when using more traditional market models and other problems such as the presence of momentum when using the Fama-French three-factor model as the basis for the test. We use a variety of testing methods in this analysis to detect adjustments of security prices to specific kinds of new information. We employ this unique methodology to natural disasters causing insurable losses in the US insurance sector.

The advantages of using a numeraire proxy to isolate market adjusted returns are the same as the advantages of using various market portfolio proxies to obtain abnormal returns. Just as market model proxies are used to estimate abnormal returns for individual assets, a numeraire proxy can be used for obtaining the market adjusted returns from the benchmarked returns of individual assets. Benchmarked returns have the same qualitative interpretation as conventional abnormal return measures. Ignoring measurement and estimation errors, both approaches measure asset performance relative to the average contemporaneous performance of other assets in the market. Both statistics are constructed so that, in equilibrium, the best short-term forecast for the statistic is zero.

The main contribution of this paper is that it applies the benchmark approach to event study methodology in a practical setting. We identify a growth optimal portfolio to act as the numeraire and then show that the benchmark approach strongly dominates market model and factor model approaches. The so-called benchmark approach of Bühlmann and Platen (2003) provides a general framework for financial market modelling which extends beyond standard risk-neutral pricing theory. It permits a unified treatment of portfolio optimisation, derivative pricing, integrated risk management and insurance risk modelling. The existence of an equivalent risk-neutral pricing measure is not required. Instead, it leads to pricing formulae with respect to the real-world probability measure. This yields important modelling freedom which is necessary for the derivation of realistic and parsimonious market models. In addition to the fact that market model regressions of each asset are not necessary, an important advantage of the benchmark approach is that analysis may still be conducted using firms who cannot obtain reliable market model regressions.

In order to demonstrate the applicability of this approach, we use a portfolio of 7 large insurable disasters as events to observe the influence on stock returns in the insurance sector. This study utilises a short window positioned around relatively cleanly-dated events to demonstrate the tractability of the benchmark approach. This methodology provides a robust system for detecting abnormal returns to test market efficiency and is simpler that traditional approaches. Natural disasters as insurable events are used to measure market efficiency for two reasons. First, the magnitude of the insurance losses for each of the selected events are sizeable, therefore market inefficiencies in the form of mispricing should be apparent. Second, the availability of each insurer's unearned premium reserve account will have a direct but unknown impact on the firm's expected earnings which will immediately affect insurer stock returns after a significant event. We avoid problems associated with a contaminated estimation period discussed in Aktas et al. (2007) by using an appropriate test statistic.

This paper is organised as follows. Section 2 describes the benchmark approach in broad terms. Section 3 outlines the structure of the insurance market used in this study and the nature of natural disaster insurance events. Section 4 describes the testing methodology of both the benchmark approach (numeraire-denominated returns) and the market model approach. Section 5 constructs the test statistics used for this analysis and Section 6 presents the results of both approaches. Section 7 discusses and tests the differences between both approaches. Section 8 tests for a relationship between the size of natural disaster events and abnormal returns. Section 9 offers some concluding remarks.

2. Numeraire portfolios and efficient markets

Previous event studies have almost exclusively used the market model to measure abnormal returns, however more recent analysis have employed the Fama and French (1996) three-factor model and the Carhart (1997) four-factor model. The Carhart (1997) model will not be employed in this analysis as it has been shown to offer only a modest increase in explanatory power relative to the Fama-French three-factor model (Schneider and Gaunt, 2011). For the market model the intercept and slope from the regression of a stock's return on the market return, estimated outside the event period, are used to estimate the stock's expected returns conditional on market returns during the event period. We refer to this as the market model approach (MMA) in this analysis. For the Fama-French three-factor model the estimation of abnormal returns is defined as the difference between the ex post return of the underlying asset minus the normal return defined under the three-factor model using the size (SMB),book-to-market ratio (HML) and the risk premium on the market portfolio as determinants of asset returns.

As with the CAPM, market model prediction errors are biased estimates of the true abnormal returns when security returns are generated by a multifactor model. They may also be noisier than multifactor model prediction errors. The bias will, however, generally average to zero in a large sample. Our sample here is relatively small so a more natural measure of detecting abnormal returns is required.

If a satisfactory numeraire proxy of the type suggested by Long (1990) can be identified then benchmarked returns offer a viable alternative to the above conventional measures to estimate abnormal returns. Our approach is conceptually similar to the benchmarking of assets in Bühlmann and Platen (2003) to empirically measure market efficiency. The most poignant feature of the numeraire portfolio is the property that zero is always the best conditional forecast of the future numeraire denominated rate of return on an asset (Long, 1990). This condition is consistent with cross-sectional heterogeneity and material variation in expected dollar rates of return, so long as these are not indicative of riskless profit opportunities.

Long (1990) showed that a set of assets offers no profit opportunities if and only if a numeraire portfolio can be formed from the set. A numeraire portfolio is defined to be a self-financing portfolio such that, if current and future asset prices and dividends are denominated in units of the numeraire (that is, divided by the contemporaneous value of the numeraire portfolio), the expected rate of return of every asset on the list is always equal to zero. Numeraire-denominated returns are natural measures of abnormal returns. If there are no profit opportunities, the best forecast of future numeraire-denominated returns is zero. Numeraire-denominated or benchmarked returns are nominal returns adjusted to reflect the contemporaneous return on the market (as measured by the nominal return on the numeraire portfolio), and the numeraire-denominated rate of return on itself is zero by construction. In this sense, numeraire-denominated returns measure asset-specific returns in the same sense as market-model residuals. The numeraire portfolio approach has been successfully applied to a number of event studies in Gerard et al. (2000), Hentschel and Long (2004) and Christensen (2005).

Benchmarked returns have two advantages over existing approaches to estimate abnormal returns. Because the multivariate process of benchmarked returns depends only on relative gross returns, it must be stationary under a broader range of circumstances than the multivariate process of nominal returns. For instance, over longer time horizons, the impact of inflation on the parameter estimates for the market model may be significant, even though relative gross returns, and therefore benchmarked returns, are not affected. Benchmarked returns are also independent of expectations and realisations of pure price level inflation. Long (1990) showed that benchmarked rates of return can be a stationary stochastic process even if the nominal rates of return are not. Also, numeraire denominated returns for individual assets are computed through simple division which eliminates the requirement for market model parameters to be estimated for each security across each event in the analysis. The feature of computational and interpretive simplicity is, in fact, one of its most attractive features.

More broadly the numeraire portfolio permits simple derivations of the main results of financial theory. The prices of self-financing portfolios, when the optimal growth portfolio is the numeraire, are martingales in the 'true' probability and given the dynamics of the traded securities, the composition of the numeraire portfolio as well as its value are easily computable. Among its numerous properties, the numeraire portfolio is instantaneously mean variance efficient which allows a simple derivation of standard continuous time CAPM, C-CAPM, APT and contingent claim pricing. The proof underlying the use of the numeraire portfolio as a substitute for the market model is derived in Long (1990) and is not reproduced here.

The obvious problem in estimated market adjusted returns obtained using the benchmark approach is finding an appropriate GOP proxy. However, as shown in Breymann et al. (2004) an excellent proxy for the GOP is available which will be exploited.

3. Large insurable losses

Several natural disasters and catastrophic events have recorded losses greater than US$5 billion and subsequently caused insurance firm bankruptcies in the US. Very little research has been dedicated to examining the market's reaction to such events. Only three studies have examined the impact of large losses and other significant events on insurers' stock values using the market model. Sprecher and Pertl (1983) and Davidson et al. (1987) find that large losses due to natural disasters and airline incidents are rapidly incorporated into stock prices with significant negative returns however Shelor et al. (1992) find that property-liability insurer stock values tended to increase after an isolated catastrophic loss event. Over the long-term however, no specific evidence has been produced to examine the market's reaction to insurance stocks which face large insurance losses. Previous studies measuring the reaction to security returns have focused on only a single event, however, to obtain an unbiased measure of efficiency to avoid factors particular to that event, a number of independent events is used in this analysis. This study will show that using the benchmark approach of Long (1990), impending large insurance losses are efficiently incorporated into stock prices following catastrophic events.

3.1 Insurance coverage for natural disaster events

Natural disasters are used in the analysis for two reasons. Firstly such events may in fact be of some benefit to property casualty insurers. The hypothesis that an increased demand for insurance following a significant event driven by consumer awareness is supported in Kunreuther et al. (1978) who found that consumer demand for flood insurance increased following a flood event. Secondly when a catastrophe is significantly large the entire insurance sector may experience a net loss, however the smaller firms and firms who have a geographically concentrated customer base are limited by statutory accounting requirements in the amount they can deploy from surplus accounts to pay claims. Smaller and geographically concentrated insurers are usually more negatively affected than the larger insurers.

Furthermore statutory accounting requirements, which are designed to ensure sufficient reserves to cover policies, mandate the maintenance of an unearned premium reserve account to limit the amount of surplus assets an insurer can deploy to settle claims. The amount is then reduced through a credit to earnings as the policy period progresses to expiry, however to balance the unearned premium reserves account an insurer must transfer funds from other accounts, usually from surplus reserves. Importantly these reserves cannot be used to settle claims and so when an insurer expects a large number of claims from a catastrophic event, the insurer suffers potential liquidity shortfalls which may result in the insurer entering the reinsurance market to obtain capital. Reinsuring increases liquidity by freeing up surplus reserves from the unearned premium reserve account and the insurer receives a ceding commission from the reinsurer that rebates the insurer for costs to issue the policy. The ceding commission increases surplus reserves immediately, but a financially distressed insurer may be forced to accept a lower ceding commission from reinsurers which may not sufficiently recover surplus accounts, which in turn affects solvency. Catastrophic events therefore tend to affect the traded stock of insurers in a relatively more complex way than for non-insurers. Large losses may result in a fall in an insurer's net worth, however the incident itself may present an insurer with profit opportunities, particularly the larger and more liquid insurers who may exploit a temporary liquidity crisis to achieve greater than expected returns.

3.2 A portfolio of catastrophic losses

The need to observe multiple events over a period of time is highlighted by the shortcomings evident in single event studies. The returns around a single event may be susceptible to contamination by major economic events independent of, but occurring simultaneously with, the natural disaster event date. Since many of the assets held by insurance companies are typically highly concentrated in interest rate sensitive instruments relative to the market, an unanticipated change in market conditions may considerably affect insurance stock returns. To combat this, we examine insurance security returns for the US insurance sector using a portfolio of events over a 20-year period.

A catastrophic loss in insurance terms has a number of competing definitions (Chen et al. 2011). In order to encompass all existing meanings of the notion of a catastrophe to insurers and the subsequent reported losses to market participants, we chose the seven largest insurance loss events in the US that occurred over a twenty-year period, 1989-2008. These events represent significant losses, the smallest being Hurricane Hugo at US$7.3 billion and the largest being Hurricane Katrina at US$45 billion.

3.3 Information leakage concerns

The events used in the analysis must avoid information leakage effects. This means events cannot be anticipated by the market, thus allowing the reporting of event details to send new and relevant information to the market. Aktas et al. (2007) extended the Markov switching regression framework developed by Hamilton (1989) to address contamination in the estimation window by proposing a two-state version of the classical market model as a return-generating process. Their results highlight the importance of explicitly controlling for unrelated events occurring during the estimation window, especially in the presence of event-induced volatility. To address the potential existence of bias from other events we employ the two-state market model (TSMM) test in addition to two relatively standard but powerful test statistics.

To obtain a cleanly dated event for each catastrophe, the event date varies depending on the nature of the actual event. For the earthquake event used in this analysis the event date $t = 0$ is the actual date of the event as first reported in news announcements in Reuters, Bloomberg and the wider financial news media. Typically, no prior warning is possible and therefore the problem of information leakage to the market is avoided. The event date is therefore the actual date in which the event occurred and was reported in the appropriate news channel as having occurred. For hurricanes however, anticipated losses which may represent information leakage, can occur up to several days prior to the actual event. By the time a hurricane reaches the coastal region the information surrounding the impending event may be old news to the market. For both of these types of events, we carefully set the event date in such a way that the information concerning the event has not been widely disseminated. For consistency we set the event date two days prior to the instant that the majority of damage occurs. This may be somewhat of an arbitrary correction but 48 hours represents the upper limit of forecast accuracy for hurricanes (Considine et al. 2006), which helps capture the true market reaction where information leakage is possible.

Avoiding the problem of information leakage is important when the reaction of the market to an event is observed only on the event day, constituting a one-day event window. Assuming it takes the market longer to assimilate all information surrounding an event, particularly a large loss insurance event with extensive damage to a widespread area, a longer event window is required. We will utilise a 31 day trading window, 10 days of which precedes each event $t \in \{-10, \dots, -1\}$ one day to represent the event date $t = 0$, and the remaining 20 days superseding each event $t \in \{1, \dots, 20\}$. Using this trading window will also help avoid the problem of estimating the exact event date for events, such as hurricanes, which do not have the characteristics of inflicting damage, and therefore conveying information to market participants, at any particular instant.

Using the above criteria for the US insurance industry from 1989-2008 we observe 7 large loss events that can be classified as a catastrophe. The market proxy used to represent the benchmark in this analysis is the Morgan Stanley Capital Growth World Index (MSCI). The MSCI World Index includes stocks from 22 countries and is based on around 1200 stocks over 1989-2008. The MSCI World Index has been empirically shown in Le and Platen (2006) and Breymann et al. (2006) to be a good approximation of the GOP, on both a daily and intraday basis. As discussed in Platen (2002) however, any broadly diversified portfolio can be shown to approximate the GOP. In principle, the GOP is the portfolio that cannot be beaten in any reasonable systematic way.

4. Testing methodology

If tests based on different market models result in different conclusions about market efficiency, then a correctly specified model is of vital importance (Brenner, 1979). However if the conclusions about market efficiency are insensitive to the model used then a robust theory of efficient market exists. Since the correct market model is unknown, we cannot separate truly inefficient market behaviour from an observed indication of inefficiency due to biases. One may assume the market to be efficient and choose the model that coincides with it to be the correct one. Alternatively, one may assume that a particular market model is true and test the efficiency of the market using this model. We lean towards the former approach in the initial part of this analysis, due to the increasing weight of evidence supporting the EMH (Fama, 1998).

The Central Limit Theorem guarantees that if the excess returns in the cross section of securities are independent and identically distributed drawings from finite variance distributions, the distribution of the sample mean excess return converges to normality as the number of securities included in the drawings increases (Beard et al. 1984). There is evidence of the distribution of cross sectional US insurance stock excess returns converging to the normal, and as such the use of a number of security returns across a number of events is methodologically prudent. Indeed, of concern here is that the assumptions underlying the Central Limit Theorem are not violated while testing market efficiency.

We use the return of the MSCI World Index to represent the value-weighted market index return $R_m(t)$ at time *t*, since it is our proxy for the market. If $\mathrm{Cov}\left(\tilde{R}_m(t), \tilde{\varepsilon}_i(t)\right) = 0$ during the period under consideration, then $\tilde{\beta}_i$ will be an unbiased estimate of β_i, regardless of the true underlying model. We let *t* represent the time passage before and after the event and let $i, i = 1, \dots, N$ be the number of securities represented for each event. We do not distinguish between types of events. The value for *N* is therefore the number of securities *i* multiplied by the number of events. Not all securities were publicly listed for every event. In total, 420 sets of security prices are used in this analysis.

Our data consists of the security returns of several dozen US insurance firms over 7 individual catastrophic events. When there is positive cross-sectional dependence in security prices, failure to make an adjustment results in a systematic underestimation of the variance of the mean excess return, implying too many rejections of the null hypothesis, both when it is true and when abnormal performance is present, see Brown and Warner (1985). There are time dependent variations in the correlation between each security and the market proxy across each event. This is likely to induce errors in the market model parameter estimates. The market model goes some way in correcting for this however, by re-estimating the parameters for each security and for each event. In addition the hurricane season in 2004 and 2005 presents multiple events that overlap the event window. For these years we examined excess returns for the first event of the season, the last event of the season and each event and found little difference in the significance of the results. Hence all overlapping events can be used to construct a portfolio of large insurable loss events.

4.1 Excess returns from the numeraire portfolio

The benchmark approach can be formulated to retrieve evidence of abnormal returns in an efficient market. A numeraire portfolio *N* is defined as a self-financing portfolio with always positive value such that, for each asset *j* and each time t, $0 \leq t \leq T$,

$$\frac{P_{j,t}}{V_{N,t}} = E_t\left\{\frac{P_{j,t+1} + D_{j,t+1}}{V_{N,t+1}}\right\}, \qquad (1)$$

with probability one, where $P_{j,t}$ is the ex-dividend price of asset *j*, $D_{j,t}$ is the dividend per unit of asset *j*, $V_{N,t}$ represents the value of a self-financing market portfolio and $E_t\{.\}$ denotes the expected value conditional on all information available at time *t*. If a numeraire portfolio exists, and if

$$\hat{P}_{j,t} = \frac{P_{j,t}}{V_{j,t}}, \hat{D}_{j,t} = \frac{D_{j,t}}{V_{j,t}},$$

and

$$1 + \hat{R}_{j,t+1} = \frac{\hat{P}_{j,t+1} + \hat{D}_{j,t+1}}{\hat{P}_{j,t}} = \frac{1+R_{j,t+1}}{1+R_{N,t+1}},$$

then the numeraire's definition and the law of iterated expectations imply

$$\hat{P}_{j,t} = E_t\{\hat{P}_{j,t+1} + \hat{D}_{j,t+1}\} = E_t\{\sum_{\tau \approx t+1}^{T} \hat{D}_{j,\tau} + \hat{P}_{j,T}\},$$

and

$$E_t\{\hat{R}_{j,t+1}\} = 0$$

with probability one for all j and t. In other words, when prices and dividends are denominated in units of a numeraire portfolio, asset prices equal the undiscounted sum of expected future payoffs.

The numeraire denominated rate of return $\hat{R}_i(t)$ on security i for the period $[t-1, t]$ is

$$\hat{R}_i(t) = \frac{1+R_i(t)}{1+R_{GOP}(t)} - 1, \quad (2)$$

where $\hat{R}_i(t)$ is the nominal rate of return on security i for the period $t \in \{-10, \ldots, 20\}$, and $R_{GOP}(t)$ is the contemporaneous nominal rate of return on the numeraire portfolio or MSCI, which is also the GOP. The values for $R_{GOP}(t)$ are essentially equivalent to the values for $R_m(t)$ used in the market model approach. However, we distinguish between the two for notational convenience since the same index is used in different contexts.

Equation (2) can be expressed as

$$\ln\left(1 + \hat{R}_i(t)\right) = \ln\left(1 + R_i(t)\right) - \alpha - \beta \ln\left(1 + R_{GOP}(t)\right), \quad (3)$$

and by setting $\alpha = 0$ and $\beta = 1$ we obtain

$$\ln\left(1 + \hat{R}_i(t)\right) = \ln\left(1 + R_i(t)\right) - \ln\left(1 + R_{GOP}(t)\right), \quad (4)$$

which is known as the 'zero-one market model' based on log returns, see Brenner (1979). We use this specification as the alternative to the traditional market model, however we will refer to this method as the benchmark approach for consistency.

As shown in Long (1990), the expected value of the one step ahead numeraire denominated rate of return of a fair price process satisfies the equation

$$E_t\left[1 + \hat{R}_i(t+1)\right] = E_t\left[\frac{1+R_i(t+1)}{1+R_{GOP}(t+1)}\right] = 1, \quad (5)$$

for $t \in \{0,1,2,\ldots\}$. An asset's numeraire denominated gross return, defined as one plus its rate of return, for a given period is calculated as its nominal gross return divided by the numeraire's nominal gross return. Thus, benchmarked returns are nominal returns adjusted to reflect the contemporaneous return on the market, as measured by the nominal return on the numeraire portfolio. The benchmark return on itself is zero by construction. In this sense, benchmark returns measure asset specific returns in the same context as market model residuals. Benchmarked or numeraire denominated returns are therefore a natural measure of abnormal returns. If there are no profit opportunities or more broadly, no arbitrage, the best forecast of future benchmarked returns is zero.

The securities observed over all catastrophic events are combined into an equally weighted portfolio $\hat{\hat{R}}_p(t)$, which has a return

$$\hat{R}_p(t) = \frac{1}{N}\sum_{i=1}^{N} \hat{R}_i\,(t), \tag{6}$$

where N is the number of securities multiplied by the number of events. This creates a portfolio of 420 securities, representing a time series of all publicly listed stocks over the 30 day event period for 7 catastrophic events. We use an equally weighted portfolio because on and immediately after the event date there is unreliable information concerning the actual level of insurable loss exposure to each insurer, therefore an equally weighted portfolio serves as an average loss for the whole sector. The market share of each insurer is roughly equal and we therefore assume the insurable loss for each insurer, averaged over a portfolio of 7 events, will also be roughly equal. We shall also form an inverse variance portfolio $\hat{R}_p^{\sigma^{-1}}(t)$ by

$$\hat{R}_p^{\sigma^{-1}}(t) = \sum_{i=1}^{N} w_i\,\hat{R}_i(t)\,, \tag{7}$$

where w_i represents the weight applied to each GOP denominated security return, the calculation of which will be discussed in section 5.1.

The notion of an abnormal return, from the market model approach, is misleading in the context of numeraire denominated returns, since the short-term expected numeraire denominated return is zero by construction, see (5). However, any nonzero return observed under the benchmark approach will be viewed as an abnormal return for the purposes of consistency in this study.

The *cumulative abnormal return* (CAR) for an equally weighted portfolio $\widehat{CAR}(t)$ is calculated as

$$\widehat{CAR}(t) = \sum_{t=-10}^{20} \hat{R}_p\,(t)\,, \tag{8}$$

and the CAR for the inverse variance weighted index $C\hat{A}R^{\sigma^{-1}}(t)$ is

$$\widehat{CAR}^{\sigma^{-1}}(t) = \sum_{t=-10}^{20} \hat{R}_p^{\sigma^{-1}}\,(t)\,, \tag{9}$$

for $t \in \{-10, \dots, 20\}$. These measures are useful for observing trends in abnormal returns during the event period. The use of CARs instead of other measures such as the buy and hold abnormal returns (BHARs) is because the BHAR measure can give false impressions of the speed of price adjustment to an event, see Fama (1998). The relative computational and interpretive simplicity of the benchmark approach is a feature of this method.

4.2 Excess returns and factor models

Assuming capital market efficiency in the context of Fama (1976), security prices will adjust rapidly to new information in an unbiased manner. Capital market efficiency can be represented as

$$E_t\{[P_i(t+\Delta) - E_t(P_i(t+\Delta)|\mathcal{A}_t)]|\mathcal{A}_t\} = 0\,, \tag{10}$$

where $P_i(t+\Delta)$ is the price of security i at time $t+\Delta$, $\mathcal{A}_t$ is the information set at time t and E_t is the conditional expectations operator at time t. The difference between the expected price and the actual price at $t+\Delta$ based on the information set $\mathcal{A}_t$ is expected to be zero, assuming market equilibrium.

In order to gain a more relevant view of market efficiency for this empirical study we use rates of return rather than prices in the above model. Equation (10) is restated as

$$E_t\{[R_i(t+\Delta) - E_t(R_i(t+\Delta)|\mathcal{A}_t)]|\mathcal{A}_t\} = 0\,,$$

where $R_i(t+\Delta)$ is the return of security i at $t+\Delta$.

The advantage of this method is that it is able to abstract the effects of the unique event from that of general market conditions. The disadvantage is that errors in the expected returns during the event period may be significant if the covariance between the security of interest and the market proxy is diminutive and insignificant. The market model first cited by Sharpe (1963), introduced a relationship between $R_i(t)$ and $R_m(t)$ implied by bivariate normality, and is represented by

$$R_i(t) = \alpha_i + \beta_i R_m(t) + \varepsilon_i(t)$$

where $R_i(t)$ is the return on security i at time t, $R_m(t)$ is the return on the value-weighted market index at time t, α_i is the intercept of security i, β_i is the beta of security i equivalent to Cov(R_i, R_m)/Var (R_i, R_m), and $\varepsilon_i(t)$ is the disturbance term. The disturbance term $\varepsilon_i(t)$ has mean zero and is independent of $R_m(t)$ so that

$$E(\tilde{\varepsilon}_i(t)|R_m(t)) = E\big(\tilde{\varepsilon}_i(t)\big) = 0$$

and

$$\sigma^2\left(\tilde{R}_i(t)|R_m(t)\right) = \sigma^2(\tilde{\varepsilon}_i(t)|R_m(t)) = \sigma^2\big(\tilde{\varepsilon}_i(t)\big) = \sigma^2(\tilde{\varepsilon}_i)$$

and Cov$\big(\tilde{\varepsilon}_i(t), R_m(t)\big) = 0$ for $t \in \{-150, \dots, 20\}$. Time zero $t = 0$ is the date of the event. The abnormal returns are examined from 150 days prior to the event $t = -150$, to 20 days after the event $t = 20$. The first pass regression estimates the parameters of the market model using $t = -150$ days to $t = -11$ days, and we call this period the estimation period. This regression provides estimates for α_i and β_i, which are denoted as a_i and b_i, respectively, for each security i over each of the 7 events. These parameter estimates are applied to the actual market return $R_m(t)$ for days $t = -10, \dots, 0, \dots, 20$ to obtain the normal returns $\tilde{R}_i(t)$ for security i over each of the 7 events. The estimated normal returns are compared to the actual returns for each of the i securities for days $t = -10, \dots, 0, \dots, 20$. This subset of time is referred to as the event period.

An equally-weighted index has historically been employed in actual event studies to represent the market index, and due to the evidence of bias using a value-weighted index in Brown and Warner (1980) we employ the US S&P 500 Equal Weighted Index (EWI) for the regressions. The S&P500 EWI is the equal-weight version of the widely regarded S&P500 and has the same constituents as the capitalization weighted S&P500 with each company in the index allocated a fixed weight.

The difference between the normal returns and the actual returns for security i at time t is called an *abnormal return*, $AR_i(t)$, and is determined by

$$AR_i(t) = R_i(t) - \big(a_i + b_i R_m(t)\big), \tag{11}$$

where $R_i(t)$ represents the actual return on security i at time t, and $a_i + b_i R_m(t)$ is the expected return for each of the 7 events.

The average abnormal return, or more accurately the portfolio excess return, is computed by summing the $AR_i(t)$ across all i, $i = 1, \dots, N$ firms for each day both before and after the catastrophic event, and dividing by N.

$$AR_p(t) = \frac{1}{N}\sum_{i=1}^{N} AR_i(t). \tag{12}$$

The cumulative abnormal return (CAR) is defined as

$$CAR(t) = \sum_{t=-10}^{20} AR_p(t). \tag{13}$$

The Fama and French (1993) three-factor model is constructed similarly. The *ex post* Fama French model is given as $R_i(t) = \hat{\alpha}_i + \hat{\beta}_i R_m(t) + \hat{s}_i SMB(t) + \hat{h}_i HML(t) + \varepsilon_i(t)$ for $i = 1, \dots, N$ where $R_m(t)$ is the value-weighted index, SMB (Small Minus Big) is a mimicking portfolio to capture risk related to size and HML (High Minus Low) is a mimicking portfolio to capture risk associated with book-to-market characteristics.

The coefficient estimates $\hat{\alpha}_i$, $\hat{\beta}_i$, $\hat{s}_i$ and $\hat{h}_i$ are regression coefficients and ε_i is the error term. The coefficient estimates are obtained using OLS regression on estimation period returns. We used a daily time series of the Fama-French factors (HML, SMB and market excess return) for the US from the Fama and French factor database. Abnormal returns and cumulative abnormal returns are estimated for each security over the 7 event windows as per equations (12) and (13).

4.3 Event-induced variance

In general, event study tests are reasonably powerful, see Brown and Warner (1985), but there are potential testing problems created by event induced increases in the variances of returns, particularly when using the market model approach (Lee and Varela, 1997). If the variance is underestimated, the test statistic may lead to an incorrect rejection of the null hypothesis of zero abnormal returns (a Type I error). To remedy this problem, estimation period residual variance is ignored and the cross sectional variance over the event period itself forms the test statistic. We will also employ an extension of this method that weights the abnormal returns in inverse proportion to their variance for the market model approach, the Fama-French three factor model and the benchmark approach.

The cross-sectional standard deviation in the event period can increase by up to four times the standard deviation observed during the estimation period (Boehmer et al. 1991). The increase in standard deviation from event-induced variance observed in US insurance stock returns rises by up to 55 percent for some stocks, although many experience insignificant changes. This amount is, however, enough to invoke an incorrect rejection of the null hypothesis. The potential for event induced variance to contaminate the true returns will be rectified using an alternative weighted least squares testing method.

5 Test statistics

The basis for inference in traditional event studies is the use of a test statistic. Many existing tests are simple and easily constructed but they lack the ability to deal with event induced variance and abnormal return dependencies among security returns. The weighted least squares test employed here is relatively powerful because it can better account for event induced variance and more importantly, cross sectional dependencies.

5.1 Weighted least squares

Two different approaches to account for cross-sectional dependencies are the generalised and non-generalised least squares tests. The generalised least squares test uses the covariance in weighting the mean excess returns while in contrast, the non-generalised test disregards the correlations among abnormal returns when assigning portfolio weights. This approach is preferred to the generalised least squares approach if there are strong dependencies among contemporaneous returns. There is some dependence among insurance security returns due to the high degree of industry concentration. There are, however, problems associated with incorporating covariance effects into a multifactor market model, the main one being that inferences about market efficiency can be sensitive to the assumed model for expected returns, see Fama and French (1996).

A measure of cumulative abnormal returns used by Shelor et al. (1992) is the modified weighted least squares method, initially suggested by Chandra and Balachandran (1990). This approach weights the abnormal returns in inverse proportion to their variance. Shelor et al. (1992) assert that if no systematic relationship is assumed between the mean and variance of abnormal returns, then this test appears to be the most accurate. Chandra and Balachandran (1990) claim that generalised least squares tests are inappropriate for event studies because these tests are highly sensitive to errors in specifying the abnormal return model, and also because we do not know enough about how securities should react to information to specify the correct model for an event.

If the adjusted abnormal returns are assumed to be independent, then the weighted least squares portfolio is simply the minimum variance portfolio. In the presence of the correlation between abnormal returns, Chandra and Balachandran (1990) find that this method is still proficient in observing the true excess abnormal returns.

We construct a portfolio of abnormal returns with weights

$$w_i = \frac{(\sigma_{i,t})^{-2}}{\sum_{i=1}^{N}(\sigma_{i,t})^{-2}}, \tag{14}$$

where

$$\sigma_{i,t} = \sigma_i \left[1 + \frac{1}{T_i} + \frac{(R_M^e(t) - \bar{R}_M)^2}{\sum_{t=1}^{T_i}(R_M(t) - \bar{R}_M)^2}\right], \tag{15}$$

and where T_i is the number of days in security is estimation period, $R_M^e(t)$ is the market return on day t during the event period, $\bar{R}_M$ is the average market return during the estimation period and $R_M(t)$ is the market return on day t. From this, a portfolio of abnormal returns for all securities across all events is constructed as

$$AR_p(t) = \sum_{N_t=1}^{420} w_i AR_i(t), \tag{16}$$

The weighted least squares regression model produces heteroscedastic and slightly skewed excess returns that approximate the normal. The weighted least squares cumulative abnormal returns are therefore computed as

$$CAR_p(t) = \sum_{t=0}^{T} AR_p(t). \tag{17}$$

Assuming that the abnormal returns are independent and identically distributed, as well as Gaussian, the standard error of $AR_p(t)$ is

$$\sigma_p = \frac{1}{\sum_{i=1}^{N}(\sigma_{i,t})^{-2}}, \tag{18}$$

and the z-statistic for the weighted least squares cumulative abnormal returns is

$$z_p(t) = CAR_p \sqrt{\frac{\sum_{i=1}^{N}(\sigma_{i,t})^{-2}}{T_i+1}}. \tag{19}$$

The standardised cumulative abnormal returns for firm i during event e is

$$SCAR_i = \left(\frac{1}{1+T}\right)\sum_{t=0}^{T}\frac{AR_i(t)}{\sigma_p}. \quad (20)$$

The standardised cumulative abnormal return is assumed to be normally distributed with mean 0 and variance 1. A quasi z-statistic can be obtained if we sum the $SCAR_i$ across all firms and divide by the square root of the number of firms, since the portfolio is also assumed to be normally distributed with mean 0 and variance 1.

5.2 Two-state market model

To deal with possible bias due to contamination of returns during the event window we compare the weighted least squares test with the two state market model (TSMM) test of Aktas et al. (2007) which relies on the Markov switching regression framework of Hamilton (1989). This test assumes that the return generating process can be adequately modeled by a two-state process in which one regime has normal variance and the other high variance. The market model parameters are assumed to be the same in the two regimes such that

$$R_i(t) = \alpha_i + \beta_i R_m(t) + \gamma_i D_i(t) + \epsilon_{i,S}(t), \quad (21)$$

where $\epsilon_{i,S}(t) \xrightarrow{d} N\left(0, \sigma_{i,S}^2\right)$ and S is a state variable assuming a value of 1 for the low variance state and 2 for the high variance state, as per Aktas et al. (2007). The γ_i coefficient is the estimated event-day abnormal return and the standard error of γ_i is used to standardise the abnormal return. The standardised abnormal return is $\hat{S}(AR_i) = \hat{\gamma}_i / SE(\gamma_i)$ where $SE(\gamma_i)$ is the standard error of the coefficient γ_i. The test statistic from Aktas et al. (2007) based on a maximum likelihood approach is estimated for each event window. The TSMM test statistic has been shown to dominate other standard tests during contaminated event windows. The results are discussed in section 6.2.

6 Results

6.1 The benchmarking approach

The expected returns of daily benchmarked (numeraire-denominated) prices over a short window are theoretically zero. We should therefore observe zero or statistically insignificant returns over the event period, using the MSCI World Index as the GOP proxy or numeraire. Figure 1 shows the numeraire denominated returns $\hat{R}_p(t)$ for an equally-weighted portfolio. Tests for significance are similar to those used for the market model approach. After the event day $t = 0$, we observe what appears to be a marked increase in the variability of returns that persists for about 7 days, however, all ARs were not significant. For the equally-weighted portfolio in figure 2, we recorded no significant CARs throughout the period and we can observe no significant trend. So, while the variability of returns increases, no significant trend in either direction is actually detected.

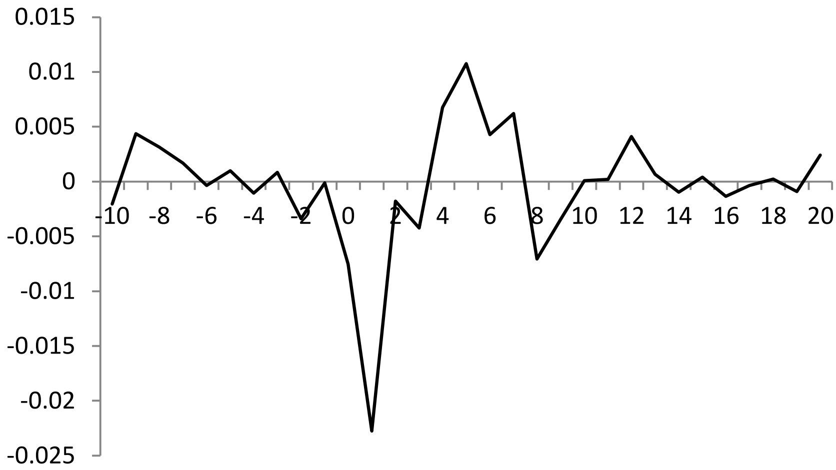

Figure 1: Numeraire denominated returns $\hat{R}_p(t)$ for an equally weighted portfolio.

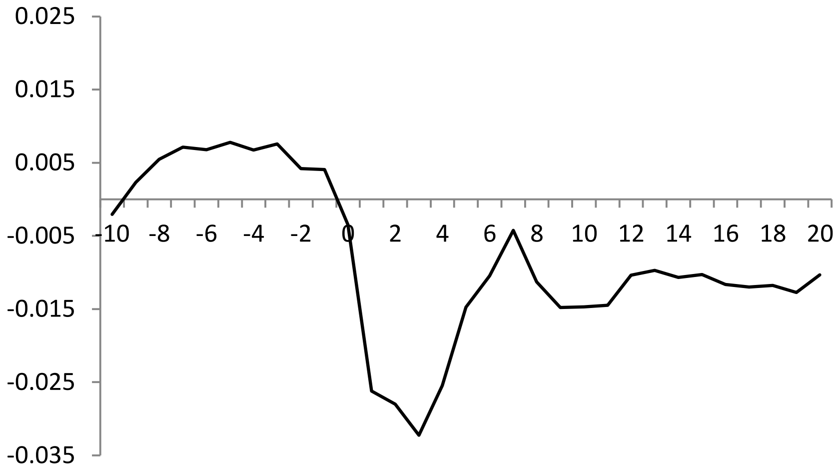

Figure 2: Numeraire denominated cumulative abnormal returns $\widehat{CAR}(t)$ for an equally weighted portfolio.

The same weighting technique applied to the market model ARs in section 5.1, when used for the benchmarked returns, eliminates the apparent variability observed in figure1. The new portfolio, defined here as an inverse variance weighted portfolio of numeraire denominated returns, was constructed with the returns shown in figure 3. This portfolio eliminates event induced variance from the observed returns, see section 5. The graphical variability after the event day $t = 0$ is absent and insignificant ARs and CARs are observed over the event period. No particular trend is observed either, as shown by the CAR in figure 4. In fact, the variability of benchmarked returns has been reduced by a factor of 4 using the inverse variance weighted portfolio. The tables have not been provided for brevity; no individual abnormal return was statistically significant for either portfolio.

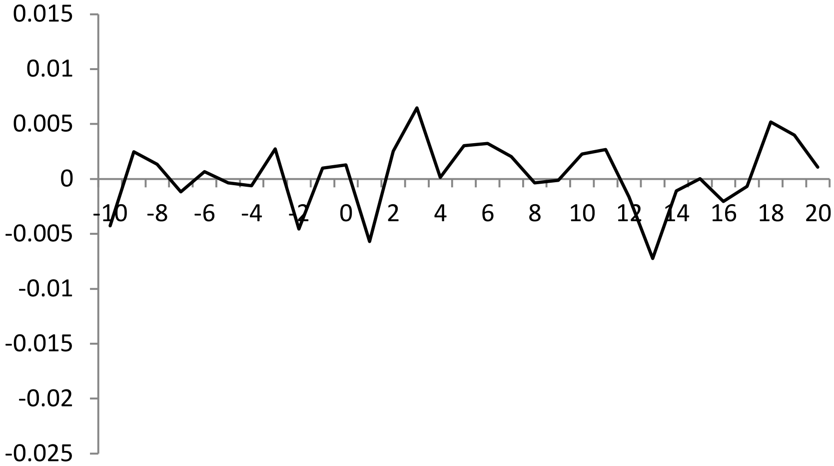

Figure 3: Numeraire denominated abnormal returns $\hat{R}_p^{\sigma^{-1}}(t)$ for an inverse variance weighted portfolio.

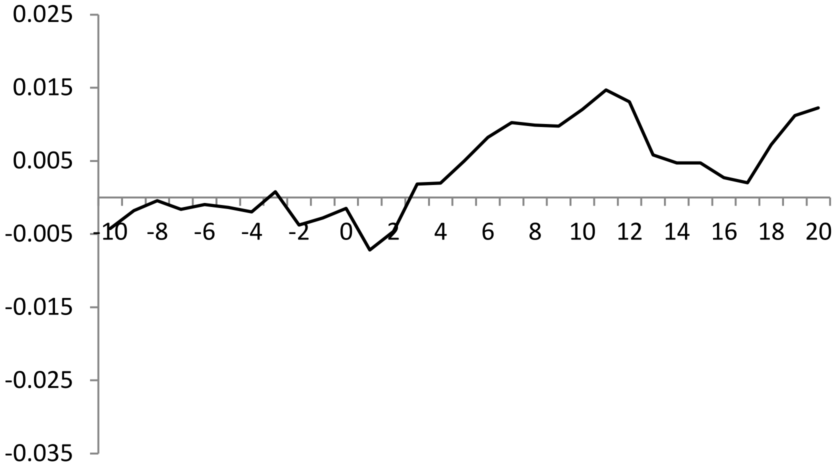

Figure 4: Numeraire denominated cumulative abnormal returns $\widehat{CAR}^{\sigma^{-1}}(t)$ inverse variance weighted portfolio.

6.2 Market and Fama-French models

Under the market model approach, catastrophic insurance events appear to have a positive impact on insurance firm value. Statistically significant positive abnormal returns were detected at various days after the event. The cumulative abnormal returns were also statistically significant and positive after some delay, using the two testing methods that eliminate event-induced variance and cross-sectional dependencies, outlined in section 5. Though statistically significant, no attempt is made in this analysis to determine if these results are economically significant. The absence of transaction costs and other market frictions in observing the positive reaction is therefore assumed.

Table 1 illustrates the results from the least squares method, which compensates for strong dependencies across securities. There are significantly positive abnormal returns on the event day $t = 0$, and also at $t = 1, 2, 3, 4, 5, 8, 11, 12, 19$ and 20 after the event day. In addition, there are significantly positive cumulative abnormal returns from $t = 1$ to $t = 20$. For all tables in this section, * and † denote significance at the 0.05 and 0.10 levels respectively.

Day (t)	$AR_n(t)$	t-stat	$CAR_n(t)$	t-stat
-10	-0.00425	-1.24	-0.00425	-3.65*
-9	0.00346	0.96	-0.00079	-0.46
-8	0.00150	0.36	0.00071	0.03
-7	-0.00140	-0.44	-0.00069	-0.31
-6	0.00360	1.04	0.00291	0.34
-5	-0.00194	-0.57	0.00096	0.00
-4	-0.00099	-0.33	-0.00002	-0.15
-3	0.00830	2.32*	0.00828	0.71
-2	0.00074	0.18	0.00902	0.68
-1	0.00202	0.59	0.01104	0.78
0	0.00939	2.71*	0.02043	1.41
1	0.02471	7.11*	0.04514	3.01*
2	0.00861	2.41*	0.05375	3.31*
3	0.01764	5.07*	0.07140	4.12*
4	-0.00788	-2.28*	0.06352	3.40*
5	-0.00586	-1.71†	0.05765	2.88*
6	-0.00457	-1.37	0.05308	2.47*
7	0.00010	0.05	0.05319	2.33*
8	0.01594	4.58*	0.06913	2.90*
9	0.00460	1.33	0.07373	2.94*
10	-0.00025	-0.08	0.07348	2.79*
11	0.00865	2.46*	0.08213	2.98*
12	-0.00589	-1.78†	0.07624	2.63*
13	-0.00481	-1.43	0.07143	2.34*
14	-0.00052	-0.15	0.07091	2.23*
15	0.00157	0.45	0.07249	2.19*
16	-0.00122	-0.39	0.07127	2.07*
17	0.00319	0.84	0.07446	2.08*
18	0.00368	1.03	0.07814	2.11*
19	0.00701	2.03*	0.08515	2.23*
20	-0.00950	-2.75*	0.07565	1.90†

Table 1: Market model AR and CAR using the weighted least squares method.

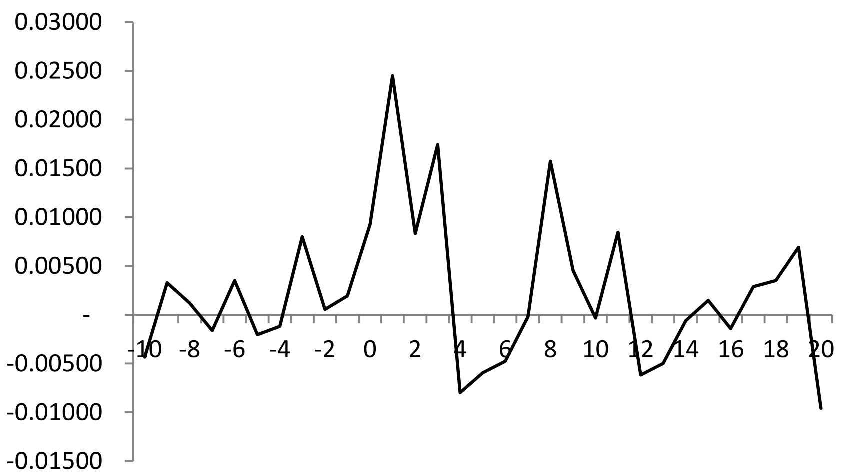

Figure 5: Weighted least squares abnormal returns for all events.

Noting the event day at $t = 0$, figure 5 illustrates the abnormal returns using the least squares method for the market model approach.

After a delay of 12 days, the CAR for the insurance sector relative to the market appears to show significant positive returns. Similar AR and CAR profiles are observed when using the Fama-French three-factor model (FF3F) approach, although the errors are much reduced. These results are broadly aligned with those obtained in Shelor et al. (1992). They attribute the positive increase in their study to anticipated increases in the demand for insurance dominating the expected rapid depletion of surplus accounts and perceived losses to insurance firms. The statistically significant results using the traditional market model approach and the FF3F model are, however, caused by biases in the regression and are not always reliable, as illustrated in section 4.

Table 2 illustrates the AR and CAR results over the event window using the FF3F model. Similar to the market model results in table 1 positive ARs are observed around the event date and for several days after the event. Positive CARs are observed from $t = 1$ to $t = 20$.

Day (t)	$AR_p(t)$	t-stat	$CAR_p(t)$	t-stat
-10	0.0350	0.59	0.0350	1.35
-9	0.0950	1.04	0.1300	1.54
-8	0.0261	1.09	0.1561	1.64
-7	0.0330	0.98	0.1891	1.61
-6	0.0516	1.57	0.2407	1.59
-5	0.0512	2.36*	0.2919	2.73*
-4	0.0326	0.71	0.3245	2.86*
-3	0.0451	1.60	0.3696	2.04*
-2	0.0363	1.41	0.4059	3.06*
-1	0.0241	0.39	0.4300	2.97*
0	0.0221	2.29*	0.4521	1.94†
1	-0.0110	-0.36	0.4411	2.06*
2	0.0966	3.40*	0.5377	1.92†
3	-0.0376	-1.24	0.5001	1.73†
4	0.0898	3.05*	0.5899	2.89*
5	0.0735	2.36*	0.6634	2.94*
6	0.0303	1.06	0.6937	3.09*
7	-0.0072	-0.14	0.6865	3.16*
8	0.0427	1.48	0.7292	3.14*
9	0.0587	1.88†	0.7879	3.15*
10	0.0451	1.52	0.8330	3.23*
11	0.0765	2.56*	0.9095	3.13*
12	-0.0306	-0.99	0.8789	3.28*
13	-0.0011	-0.03	0.8778	3.13*
14	-0.0015	-0.19	0.8763	2.91*
15	0.0487	1.55	0.9250	3.00*
16	-0.0073	-0.12	0.9177	3.10*
17	-0.0273	-0.87	0.8904	3.19*
18	0.0515	1.71†	0.9419	3.18*
19	0.0680	2.67*	1.0099	3.50*
20	0.0493	1.43	1.0592	3.23*

Table 2: Fama-French three factor model (FF3F) AR and CAR.

In contrast to the benchmark approach results of no significant returns, both the MMA and FF3F models show positive ARs and CARs around the event date and for several days after the event.

To control for potential contamination of the portfolio of event windows the TSMM statistic of Aktas et al. (2007) is used. The TSMM test statistic results weakly dominates the least squares test. However, in our analysis the degree of dominance was significantly less than the levels identified in Aktas et al. (2007).

The number and timing of statistically significant AR and CAR results during each event window did not change, which suggests that on a portfolio basis the event windows were largely free from contamination. TSMM test results are extensive and in the interests of brevity can be obtained from the author upon request.

7 Testing differences between models

These results demonstrate that statistically significant abnormal returns are observed for catastrophic insurance events using the market model and Fama-French approach. However, when using the benchmark approach the errors induced from parameter estimation under the market model disappear and the true abnormal returns are observed, with reference to a numeraire. The numeraire used here is the growth optimal portfolio (GOP).We now consider if the results obtained under the benchmark approach are robust for use in testing market efficiency.

It is clear that the abnormal returns obtained under both the MMA and the FF3F model violate the results produced in the benchmark theory of Long (1990) and Bühlmann and Platen (2003). This suggests that the market model and Fama-French three factor model provide an inefficient representation of expected returns. The abnormal returns or, more formally, the *supermartingale* property of benchmarked returns is, however, maintained using the benchmark approach. Benchmarked returns therefore provide a more comprehensive view of expected returns for testing market efficiency. We now examine each approach with reference to the EMH.

Fama (1998) identified that the presence of long-term anomalies in ARs in a data set is sensitive to the methodology used. We can extend this assertion here since we have shown that in the short-term, detecting anomalies in observed ARs is also sensitive to the methodology. ARs tend to become marginal or, as shown here, even disappear when exposed to different models for expected returns or when different statistical approaches are used to measure them. The incorrect specification of the market model for a number of securities, whose covariance with the market proxy is limited, contaminates the test for efficiency, and is thus a 'bad' model problem (Fama, 1998). The 'bad' model problem grows as the return horizon increases, such that spurious ARs eventually become statistically reliable in CARs. This is due to the fact that while the mean of the CAR increases by the number of days included in the event period at the rate t, the standard error of the CAR increases at the rate of $\sqrt{t}$.

The anomalies observed under the market model and Fama-French model have clearly disappeared under the benchmark approach. Given the evidence in favour of the EMH it remains for us to strengthen our argument in testing the dependence of the conclusions with reference to the EMH on the particular model used.

Like all asset pricing models, the market model and other factor models provide an incomplete description of expected returns. While the shortcomings of this approach are well known, it is useful to compare the results obtained under this approach with the results obtained under the benchmark approach.

If we can show that the results underlying each model are significantly different in a statistical sense, we can say with some confidence that one model must offer a better representation of the true abnormal returns than the other. In this section, we shall use some parametric tests to investigate the differences between the two sample distributions. Specifically, we shall test the differences in the basic statistics that characterise the distribution of residuals obtained under both approaches.

The construction of each approach outlined in section 4 implicitly assumes that

$$E(\tilde{\eta}_z) = 0, \tag{22}$$

where z is the residual from model $z, z \in \{a, b\}$. Here a refers to the market model approach and b to the benchmark approach. If both approaches yield similar results, then the pairwise comparison test of the form

$$E(\tilde{\eta}_a - \tilde{\eta}_b) = 0,\ a \neq b \tag{23}$$

will hold, where $\tilde{\eta}_a$ is the residual obtained from the market model approach and $\tilde{\eta}_b$ is the residual obtained from the benchmark approach.

While the specification and power of the above tests are relatively straightforward, the economic interpretation of each test is less so because the event study is a joint test of whether abnormal returns are zero and whether the assumed model of normal returns (e.g., the MMA and the NDRA) are correct. If the alternative hypothesis is rejected this means either that abnormal returns were really non-zero and the MMA and NDRA are wrong or that the event really does produce significant returns. We conduct a test for model robustness for both the MMA and the NDRA using non-events.

We know from the empirical tests in section 6 that both models yield different results. Testing for the 'correct' model is however, dependent on the underlying assumptions for the EMH. If catastrophic events convey new information to the market, then we are directly testing different conclusions about market efficiency. If catastrophic events do not convey new information to the market, then we are only testing differences between models. If one strongly believes that markets are efficient, then the conclusion would be that the true model is the one that fails to reject the EMH at a given confidence level. If we therefore assume the null hypothesis of the EMH then the model which supports the null of equation (22) is the 'correct' model. If, however, we find that the null hypothesis assumes that all models fail to reject the EMH, therefore, if validated, we say the evidence for the EMH is not conditional on the underlying model and the null hypothesis of equation (23) should hold. Given the recent evidence in the capital markets which fail to reject the EMH, with reference to the null hypothesis implicit in equation (22), it is clear that, from these assumptions, the benchmark approach is the more appropriate or 'correct' model. The market model approach has been shown to be subject to significant calibration errors, and it is also limited by the stability of the correlation between individual securities and the market proxy. The Fama-French three factor model suffers from the above limitations but has been shown to be a significantly more efficient model than the market model. The use of the benchmark approach however offers a natural measure of abnormal returns with respect to the market itself. When compared with the factor models, this is a powerful argument in favour of the benchmark approach. We test the differences between models to show that the market model approach is inaccurate and is, therefore the 'incorrect' model using US insurance data.

The benchmarked returns for both the equally-weighted portfolio and the inverse variance weighted portfolio are Gaussian. The Jarque-Bera statistics are 2.6133 and 0.2257 for the equally weighted and inverse variance weighted portfolio residuals respectively which, as expected under the EMH, suggests the residuals generated under the benchmark approach are normally distributed. In fact, the inverse variance weighted portfolio of residuals exhibit statistics for its distribution that are almost indistinguishable from the normal. For the market model, both the standardised cross-sectional portfolio and the weighted least squares portfolio of returns are also Gaussian with Jarque-Bera statistics of 3.9535 and 1.5372, respectively.

Assuming therefore that the residuals for both models are drawn from a normal distribution, we can conduct tests on the differences between the MMA and benchmark approach models. Firstly, we obtain a t-statistic for the mean difference between the two residuals, which is actually a test on the differences between means. This is computed as

$$t(D) = \frac{(\tilde{\eta}_{at} - \tilde{\eta}_{bt})\sqrt{n}}{\sqrt{\tilde{\sigma}^2(\eta_{at}) + \tilde{\sigma}^2(\eta_{bt}) - 2\widetilde{cov}(\eta_{at}, \eta_{bt})}}, \tag{24}$$

for $t \in \{-10, \ldots, 20\}$, where $\tilde{\eta}_a$ is the residual obtained from the market model approach, $\tilde{\eta}_b$ is the residual obtained from the benchmark approach and n is the number of residuals in the set, see Brenner (1979).

Next, we estimate the correlation between the two sets of residuals via

$$\tilde{\rho}(\eta_a, \eta_b) = \frac{\widetilde{cov}(\eta_{at}, \eta_{bt})}{\tilde{\sigma}(\eta_{at})\tilde{\sigma}(\eta_{bt})}. \tag{25}$$

For $t \in \{-10, \dots, 20\}$.

Finally, we calculate the Pitman statistic, which tests whether two dependent variances belong to the same population, as

$$P(\eta_a, \eta_b) = \frac{\tilde{\sigma}^2(\eta_{at}) - \tilde{\sigma}^2(\eta_{bt})}{\sqrt{[\tilde{\sigma}^2(\eta_{at}) + \tilde{\sigma}^2(\eta_{bt})]^2 - 4\tilde{\rho}(\eta_{at}, \eta_{bt})\tilde{\sigma}^2(\eta_{at})\tilde{\sigma}^2(\eta_{bt})}}, \tag{26}$$

for $t \in \{-10, \dots, 20\}$, see Snedecor and Cochran (1967). Table 3 presents the absolute averages of the above three statistics over the event window. The value for $|\bar{t}(D)|$ is significant at the 5 percent level, which indicates a significant difference between the mean residuals of each model. The Pitman statistic $|\bar{P}(\eta_a, \eta_b)|$ is close to 1 since there is minimal correlation between the residuals of each model. From these tests, we can conclude that both approaches yield statistically significant different results and therefore the different approaches cannot simultaneously support the EMH.

	$\lvert\bar{t}(D)\rvert$	$\lvert\bar{\rho}(\eta_a, \eta_b)\rvert$	$\lvert\bar{P}(\eta_a, \eta_b)\rvert$
MMA a - b	0.6961	0.0385	0.9977
FF3F a - b	0.7729	0.0467	0.9683

Table 3: Test statistics for the differences between both the market model (MMA) and the Fama French model (FF3F) with the benchmarked returns approach.

The, so called, robustness hypothesis for testing the EMH under different approaches in this example therefore fails. This is clearly a result of the inaccuracies in estimating step ahead returns using the market model and Fama-French approach.

Assuming the EMH holds, as research evidence suggests, it is clear that the benchmark approach provides a better description of expected returns over the short term. The various ‘bad’ model problems that arise under the market model and Fama-French approach generate inaccurate expected returns in relation to the movement of the market proxy. Similar results to the market model approach were reached using the above analysis for the Fama-French model. ‘Bad’ model problems are unavoidable however their effect can be vastly reduced through the use of a numeraire portfolio and the benchmark technique.

8 Catastrophe size and market reaction

Before concluding, one obvious question that needs to be addressed concerns the relation between the size of the natural disaster and the subsequent reaction implicit in ex post security returns by the market. It is generally assumed by most investors and other market participants that larger catastrophes will have a greater impact on security prices than smaller events (Shelor et al. 1992). If the market model is in fact the ‘correct’ model, then one would expect a relationship between the size of the loss and the reaction by the market to exist. Therefore, a simple regression was conducted where the insured loss, total loss and the insured to total loss ratio were regressed against the abnormal returns and cumulative abnormal returns, obtained using the market model and Fama-French approach, at three points in time after each event -the event day $t = 0$, at $t = 5$ and at $t = 15$. Day $t = 5$ was chosen because this would represent sufficient time to conduct a crude analysis on the extent of insurable losses while day $t = 15$ would represent sufficient time to conduct a detailed analysis of insurable losses.

The null hypothesis H_0 we seek to reject is that the abnormal returns and cumulative abnormal returns on days $t \in \{0, 5, 15\}$ are not dependent on the size of the insured loss, the size of the total loss, and the size of the ratio of insured loss to total loss. Thus we test the regressions

$$AR_{j,t} = \gamma_0 + \gamma_1 IL_{j,t} + \gamma_2 TL_{j,t} + \gamma_3 (IL/TL)_{j,t} + \varepsilon^1_{j,t}, \tag{27}$$

for $t \in \{0, 5, 15\}$, and

$$CAR_{j,t} = \delta_0 + \delta_1 IL_{j,t} + \delta_2 TL_{j,t} + \delta_3 (IL/TL)_{j,t} + \varepsilon^2_{j,t}. \tag{28}$$

For $t \in \{5,15\}$, where $IL_{j,t}$ represents insured loss for event j, $TL_{j,t}$ represents total loss for event j and $(IL/TL)_{j,t}$ represents the ratio of insured loss to total loss for event j. The first null hypothesis we are attempting to reject is that the level of insured losses $IL_{j,t}$ have no impact on the $AR_{j,t}$ or $CAR_{j,t}$ at $t \in \{0, 5, 15\}$ for event j, such that

$$H_0^1 : \gamma_1 = 0;\ \delta_1 = 0. \tag{29}$$

The second null hypothesis we are attempting to reject is that the level of total losses $TL_{j,t}$ have no impact on the $AR_{j,t}$ or $CAR_{j,t}$ at $t \in \{0, 5, 15\}$, such that

$$H_0^2 : \gamma_2 = 0; \delta_2 = 0. \tag{30}$$

Thirdly, we will attempt to reject the null hypothesis of the disparity between insured losses against total losses $(IL/TL)_{j,t}$ having no impact on the $AR_{j,t}$ or $CAR_{j,t}$ at $t \in \{0, 5, 15\}$, such that

$$H_0^3 : \ \gamma_3 = 0; \delta_3 = 0. \tag{31}$$

From table 4 we fail to reject the null hypotheses H_0^1, H_0^2, H_0^3 for days $t \in \{0, 5, 15\}$ for all events j since these results are not significantly different from zero. The lack of significance in the statistics implies that there is no real relationship between abnormal returns observed in insurance securities using the market model and the Fama-French model, and the readily observable factors that surround identifiable catastrophic insurance events. The regression statistics are presented in tables 4 and 5. The t-statistics are in parentheses.

Variable	$\gamma_0;\ \delta_0$	$\gamma_1;\ \delta_1$	$\gamma_2;\ \delta_2$	$\gamma_3;\ \delta_3$	R^2
AR_0	0.0018	-2.20E-06	1.20E-06	-0.0229	0.0373
	(0.3814)	(-0.2299)	(0.3928)	(-0.7224)	
AR_5	-0.0019	-1.40E-06	1.40E-06	0.0091	0.0202
	(-0.3475)	(-0.0719)	(0.3176)	(0.6194)	
CAR_5	0.0225	5.60E-06	2.10E-06	-0.0246	0.0135
	(0.8587)	(0.1500)	(-0.1692)	(-0.5097)	
AR_{15}	0.0103	-6.13E-08	1.60E-06	0.0008	0.0066
	(0.2781)	(-0.0209)	(0.1425)	(0.4475)	
CAR_{15}	0.0276	-1.90E-06	3.30E-06	0.0192	0.0064
	(0.3983)	(-0.0964)	(0.1489)	(0.2889)	

Table 4: Regression of losses against AR_t and CAR_t for $t = 0, 5, 15$ for the market model (MMA).

Variable	$\gamma_0;\ \delta_0$	$\gamma_1;\ \delta_1$	$\gamma_2;\ \delta_2$	$\gamma_3;\ \delta_3$	R^2
AR_0	0.0192	-1.90E-06	1.90E-06	-0.0999	0.0143
	(0.6724)	(-0.4289)	(0.7888)	(-1.0344)	
AR_5	-0.0205	-0.94E-06	1.74E-06	0.0992	0.0082
	(-0.7525)	(-0.5937)	(0.2006)	(0.5499)	
CAR_5	0.0525	2.76E-06	-1.70E-06	-0.0116	0.0295
	(0.9007)	(0.1850)	(-0.1002)	(-0.5187)	
AR_{15}	0.0103	-3.30E-06	-1.20E-06	0.0198	0.0486
	(0.2611)	(-0.0277)	(-0.1491)	(0.2715)	
CAR_{15}	0.0166	-1.90E-06	-2.40E-06	0.0992	0.0054
	(0.3121)	(-0.0174)	(-0.1991)	(0.3389)	

Table 5: Regression of losses against AR_t and CAR_t for $t\ = 0, 5, 15$ for the Fama-French three factor model (FF3F).

A regression was also conducted using market adjusted returns under the benchmark approach corresponding to the same hypotheses, which obtained similar results, as per table 6.

Variable	$\gamma_0;\ \delta_0$	$\gamma_1;\ \delta_1$	$\gamma_2;\ \delta_2$	$\gamma_3;\ \delta_3$	R^2
AR_0	0.2002	-1.50E-06	1.79E-06	-0.1377	0.0113
	(0.9794)	(-1.2981)	(0.9678)	(-1.2097)	
AR_5	-0.0290	-1.14E-07	2.89E-06	0.2332	0.0232
	(-0.6915)	(-1.4827)	(0.3899)	(0.5295)	
CAR_5	0.1262	1.06E-05	-1.20E-06	-0.0016	0.0295
	(0.6417)	(0.7291)	(-1.1012)	(-1.0187)	
AR_{15}	0.0271	-3.20E-05	-2.20E-07	0.0299	0.0116
	(0.9082)	(-0.1219)	(-0.3820)	(1.0023)	
CAR_{15}	0.1282	-2.80E-06	-1.50E-07	0.0712	0.0160
	(0.3661)	(-0.4176)	(0.1489)	(0.3918)	

Table 6: Regression of losses against AR_t and CAR_t for $t\ = 0, 5, 15$ for the numeraire denominated returns approach (benchmark approach).

This strengthens the argument in favour of failing to reject the EMH, and it also confirms that the benchmark approach for event studies, where securities are somewhat dependent and are subject to event induced variance, is a more efficient alternative than the traditional market model or factor model approach.

9 Conclusion

We have shown that using the benchmark approach offers a natural measure of abnormal returns. This method is more powerful than the traditional market model and the Fama-French three factor model. We have shown how this can be applied to test the EMH in the US insurance sector following natural disaster insurable events. The benchmark approach for testing the EMH, given the proposition that expected proxy-denominated stock returns are zero, can be applied to test a variety of anomalies in both the short-and long-term. This yields results consistent with the findings in this paper, thus strengthening the EMH assumption in capital markets. Furthermore, the use of the benchmark approach avoids the step-ahead estimation errors implicit in using the market model approach and estimation errors due to, among other things, the presence of momentum when using the Fama-French three factor model. The benchmark approach is flexible, simple and robust and can be applied to many other areas of finance to test for inefficiencies, so long as a representative numeraire or GOP is available.

References

Ahern, K. (2009). Sample selection and event study estimation. *Journal of Empirical Finance* 16, 466-482.

Aktas, N., de Bodt, E. and Cousin, J. (2007). Event studies with a contaminated estimation period. *Journal of Corporate Finance* 13(1), 129-145.

Ang, J.S. and Zhang, S. (2004). An evaluation of testing procedures for long horizon event studies. *Review of Quantitative Finance and Accounting* 23(3), 251-274.

Bartholdy, J., Olson, D. and Peare, P. (2007). Conducting event studies on a small stock exchange. *European Journal of Finance* 13(3), 227-252.

Beard, R., Pentikainen, T. and Pesonen, E. (1984). *Risk Theory*. third edn. Chapman and Hall, New York.

Binder, J. (1998). The event study methodology since 1969. *Review of Quantitative Finance and Accounting* 11(2), 111-137.

Boehmer, E., Musumeci, J. and Poulsen, A. (1991). Event study methodology under conditions of event-induced variance. *Journal of Financial Economics* 30, 253-272.

Breymann, W., Kelly, L. and Platen, E. (2006). Intraday empirical analysis and modelling of diversified world stock indices. *Asia Pacific Financial Markets* 12, 1-28.

Brenner, M. (1979). The sensitivity of the efficient market hypothesis to alternative specifications of the market model. *Journal of Finance* 34(4), 915-929.

Brown, S. and Warner, J. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics* 14, 3-32.

Buhlmann, H. and Platen, E. (2003). A discrete time benchmark approach for insurance and finance. *ASTIN Bulletin* 33(2), 153-172.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance* 52, 57-82.

Chandra, R. and Balachandran, B. (1990). A synthesis of alternative testing procedures for event studies. *Contemporary Accounting Research* 6, 611-640.

Chen, H.Y., Lee, C.F., Tai, T. and Wang, K. (2011). Fiscal and monetary policies in reaction to the financial tsunami by the Taiwanese Government. *Review of Pacific Basin Financial Markets & Policies* 14(1), 153-169.

Christensen, M.M. (2005). On the history of the Growth Optimal Portfolio. Working Paper, University of Southern Denmark.

Considine, T.J., Jablonowski, C., Posner, B. and Bishop, C.H. (2006). The efficiency gains from probabilistic weather forecasts: A case study of oil and gas producers in the gulf of Mexico. National Science Foundation Report ATM-9908963, Pennsylvania State University.

Daniel, K., Grinblatt, M., Titman, S. and Wermers, R.(1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52, 1035-1058.

Davidson, W., Chandy, P. and Cross, M. (1987). Large losses, risk management and stock returns in the airline industry. *Journal of Risk and Insurance* 54, 163-172.

Fama, E. (1976). *Foundations of Finance*. Basic Books, New York.

Fama, E. (1998). Market efficiency, long term returns and behavioral finance. *Journal of Financial Economics* 49, 283-306.

Fama, E. and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E. and French, K. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51, 55-84.

Fama, E. and French, K. (1997). Industry costs of equity. *Journal of Financial Economics* 43, 153-193.

Gerard, B., Santis, G.D. and Ortu, F. (2000). Generalized numeraire portfolios. UCLA Working Papers in Finance.

Hamilton, J. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica* 57, 357-384.

Hentschel, L. and Long, J. (2004). Numeraire portfolio measures of size and source of gains from international diversification. Working Paper, University of Rochester.

Kelly, J. (1956). A new interpretation of information rate. *Bell Systems Technology Journal* 35(1), 917-926.

Kunreuther, H., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkan, B. and Katz, N. (1978). *Disaster Insurance Protection: Public Policy Lessons*. Wiley, New York.

Le, T. and Platen, E. (2006). Approximating the growth optimal portfolio with a diversified world stock index. *Journal of Risk Finance* 7(5), 559-574.

Lee, S. and Varela, O. (1997). An investigation of event study methodologies with clustered events and event day uncertainty. *Review of Quantitative Finance and Accounting* 8(3), 211-228.

Long, J.(1990). The numeraire portfolio. *Journal of Financial Economics* 26, 29-69.

Nam, C., Yang, D.H., Park, M.C., Oh, G.H. and Park, J.H. (2005). Stock market reaction to mergers and acquisitions in anticipation of a subsequent related significant event: evidence from the Korean telecommunications industry. *Review of Pacific Basin Financial Markets & Policies* 8(2), 185-200.

O'Brien, M.A., Brailsford, T. and Gaunt, C. (2009). Interaction of size, book-to-market and momentum effects in Australia. *Accounting and Finance* 50(1), 197-219.

Platen, E. (2002). Arbitrage in continuous complete markets. *Advances in Applied Probability* 34, 540-558.

Schneider, P. and Gaunt, C. (2012). Price and earnings momentum in Australian stock returns. *Accounting and Finance* 52(2), 495-517.

Sharpe, W. (1963). A simplified model of portfolio analysis. *Management Science* 9(2), 277-294.

Shelor, R., Anderson, D. and Cross, M. (1992). Gaining from loss: property-liability insurer stock values in the aftermath of the 1989 California earthquake. *Journal of Risk and Insurance* 59(3), 476-488.

Snedecor, G. and Cochran, W. (1967). *Statistical Methods*. Ames Iowa State University Press, Iowa.

Sprecher, C. and Pertl, M. (1983). Large losses, risk management and stock prices. *Journal of Risk and Insurance* 50, 107-117.

Appendix: Large Loss Events -US 1989-2008

Event	Type	Date	Insured loss ($bn)
Hugo	Hurricane	22-September-1989	7.3
Andrew	Hurricane	24-August-1992	23.8
Northridge	Earthquake	17-January-1994	18.2
Charley	Hurricane	13-August-2004	8.5
Katrina	Hurricane	29-August-2005	45.3
Wilma	Hurricane	24-October-2005	11.3
Ike	Hurricane	12-September-2008	12.5

Table 7: Hurricane events USA 1989-2008 (in 2009 US$).

Table7 summarizes the catastrophic events that qualify for analysis in this study. The conditions set were estimated insurance losses greater than US$4 billion and a minimum degree of information leakage. The dates given are the actual dates where it became apparent that substantial loss was likely (this can be several days prior to the actual incident). These events constitute 7 of the top 10 most costly disasters in US history (Insured Losses, 2009, US$ billions). *Source:* PCS; Insurance Information Institute.