Last Revised: May 26, 2011

Financial Modeling and the Crisis I: "Black Swans"

Terry Marsh and Paul Pfleiderer[1]

1. Introduction

A visit to a bookstore these days, be it in Tokyo, London, Sydney or New York, finds the New Arrivals shelves stacked with fallout from the financial crisis. Books with attention-grabbing titles typically contain colorful narratives of the near-destruction of Wall Street and Main Street at the hands of bankers and "quants" armed with their weapons of mass destruction. In many of these accounts it is asked "how did economists and others get it so wrong." One alleged Achilles heel of financial models is what has come to be known as the problem of "unknown unknowns"[2] or "black swan events," i.e., extreme events that are not well enough understood for their probabilities to be accurately modeled. As Goldman Sachs' Chief Financial Officer famously lamented in August 2007: "we were seeing things that were 25 standard deviation moves, several days in a row." Of course, according to most of the risk forecasting models in general use at the time (many based on Gaussian assumptions), "25 standard deviation events occurring several days in a row" would have been deemed well nigh impossible, and yet these events occurred. The obvious implication is model failure.

In this paper, we argue that the situation is not as bleak as the above portrayals suggest. In particular, we argue that extant models that properly encompass predictable *shifts* in investment risk ameliorate some of the alleged "black swan" problems. We provide a simple illustration using U.S. S&P Index returns and the VIX Index. We also discuss empirical evidence that suggests that securities markets do in fact seem to be attuned to the risk of rare economic disasters and "price in" their risk. In a companion article, we extend the discussion here to a multivariate setting in which extreme events such as those occurring in the recent financial crisis increase the correlation amongst asset returns.

2. Black Swan Events and "Unknown Unknowns"

The non-Gaussian behavior of short-term asset returns is well-known empirically and is associated with "extreme events," "peso problems," "fat tails," and "leptokurtic distributions." It is also well-known that estimating the probability of extreme events is quite difficult when using most historical data: small- or modest-length samples contain few rare and extreme observations almost by definition. In the limiting case of "never-seen-before" black swan events, no sample observations exist. In principle, when there is little precision in estimating higher order moments of return probabilities from available data, confidence in point estimates for risk and pricing

---

[1] Terry Marsh is an Emeritus Professor of Finance at U.C. Berkeley and C.E.O. of Quantal International Inc. Paul Pfleiderer is the C.O.G. Miller Distinguished Professor of Finance at the Graduate School of Business, Stanford University.

[2] A term used by former U.S. defense secretary Donald Rumsfeld in a February 2002 news conference that helped earn him a "Foot in Mouth" award. A charitable interpretation is that he was referring to the Knightian concept of uncertainty.

models that are sensitive to those estimates should also be low; indeed these estimates may be well-nigh useless for many applications. In short, the implications of "black swans" are hardly a novelty to anyone understanding the limitations of small historical samples. Of course, there are those who claim that bad luck is more to blame than bad judgment – the historical data did not reveal the extreme event so how could one have anticipated it? But such excuses fly in the face of a healthy body of statistics dealing with model misspecification and its diagnosis, in which both technique and judgment play a role. As might be expected, in the presence of significant model specification risk and estimation risk,[3] decisions made by risk-averse decision-makers generally tend to be more conservative. Given the reported spotty use of readily available background knowledge for incorporating model risk and the decidedly non-conservative risks taken on by some financial institutions, one can't help but wonder whether financial model failure on account of "black swan" events has had more to do with transparency and disincentives for deploying competent models than with lack of technical know-how *per se*, a possibility we discuss elsewhere (Marsh and Pfleiderer (2009)).

"Black swan"-related issues cannot be avoided in the absence of perfect information. This is the message of most commentary, obvious though the point may be. Given the potential importance of black swan risk, it is to be expected that many investors will want to protect themselves against it. Since most of the events that are (*ex post*) labeled black swan events are systematic in nature – the last financial crisis being a good example – these generally are not risks that can be diversified away. A particular investor can obtain insurance against these extreme events – but only if someone else takes the other side of the transaction and provides the insurance. (These are not insurable risks in the sense that they can be diversified away by all investors holding diversified portfolios.) There are many ways insurance against extreme events might be delivered, from explicit insurance contracts like swaps and options to targeted "tail risk" hedge funds. It is, of course, not obvious what the "right" prices for these insurance positions are, or even if the market for many of them will ultimately be viable, since the quantitative risk characteristics of what is being insured against are difficult to estimate and perhaps "unknowable" – the potential uninsurability may be the point of the black swan label, but the label *per se* helps not at all in resolving what is, and what is not, insurable. In addition, purchasers of such insurance must worry about the performance of counterparties when the risk event occurs, a worry that exists in any setting involving insurance but which may be especially pronounced when the insured risks are extreme "black swan" events.

A good case can be made that the asset markets already "price in" extreme events in the form of rare economic disasters at the rate at which they have been observed in the twentieth century; for example, Barro (2008) calibrates a version of a reasonably standard representative investor model to the data for these crises (crediting earlier work by Rietz (1988)) and argues that the average stock market premium and real interest rates are consistent with the data for past disasters. In Barro's model, the disasters arrive identically and independently over time. Gabaix (2007) generalizes the Barro model to incorporate a time-varying intensity of disasters, which "…in turn generates time-varying risk premia and thus volatile asset prices and return predictability."

[3] Whilst "…an agent who knows the state space but not the appropriate probabilities and behaves 'conservatively' because of this lack of knowledge…[differs from one] not knowing the state space and behaving conservatively," it is difficult to distinguish the two; see Deckel, Lipman, and Rustichini (1998, p. 540) for discussion.

Black swan adherents may argue that it is one thing to put forward models in which historical securities prices are consistent with investors acting "as if" they factor disasters into their decision-making. But it is another to point to evidence that risk modelers have the wherewithal to account for these disasters, rather than just delude ourselves that investors can do the impossible. In John Kay's (2011) strong language: "[have]…risk modelers …created an industry whose intense technical debates with each other lead gullible outsiders to believe that this is a profession with genuine expertise"? This question is rhetorical, but does pose the practical question of what is genuinely do-able without deluding ourselves.

In the spirit of answering the question, we believe that there are in fact extant models that account much better for the extreme events in security returns observed over time than the black swan discussion might have one believe – including events having the order of magnitude seen in the recent financial crisis. The probabilities for returns in these models are given by various mixtures of probability distributions that account for varying "degrees of intensity" in the flow of economic events at different times; the subordinated model discussed by Clark (1973) is an early example. More recent are the hidden Markov models (HMM) that follow Hamilton (1989) (1990) in which the distribution of daily returns is governed by Markovian switching among conditional Gaussian distributions. Bulla (2008) analyzes an extension of the basic HMM approach where daily returns *conditional* on being in a crisis scenario themselves follow a *t*-distribution. That is, extreme returns arise from both the switch into a high volatility crisis regime and additionally because in that high volatility regime extreme returns are more likely than predicted by a Gaussian distribution, even given the regime's already-high volatility. Quantal uses a multivariate mixture model like this in its global multi-factor risk system. The *t*-distribution for the crisis scenario arises from a "crisis factor" that dominates across stocks and asset classes. In the crisis scenario, correlations increase substantially across stocks, so in effect the dimensionality of the factors affecting cross-sectional asset returns in large part "collapses" to a single dominant factor: "everything moves together" in a crisis.

We now provide a simple, readily-verifiable illustration of how the subordinated model might be applied to explain observed returns on a market-wide stock index. The index setting provides a simple context to illustrate how subordinated and other richer models can be used to fit the data well. The Quantal model for individual stocks has a more complicated structure, but is based on the same principles illustrated here for the index. For the illustration, we use the CBOE's daily VIX Index,[4] extending from January 2, 1990, through May 2, 2011, along with the log of daily returns for the corresponding S&P Index over the same period. As we'll now show, the evidence is that a Gaussian distribution model for these daily *conditional* S&P index returns is reasonable.

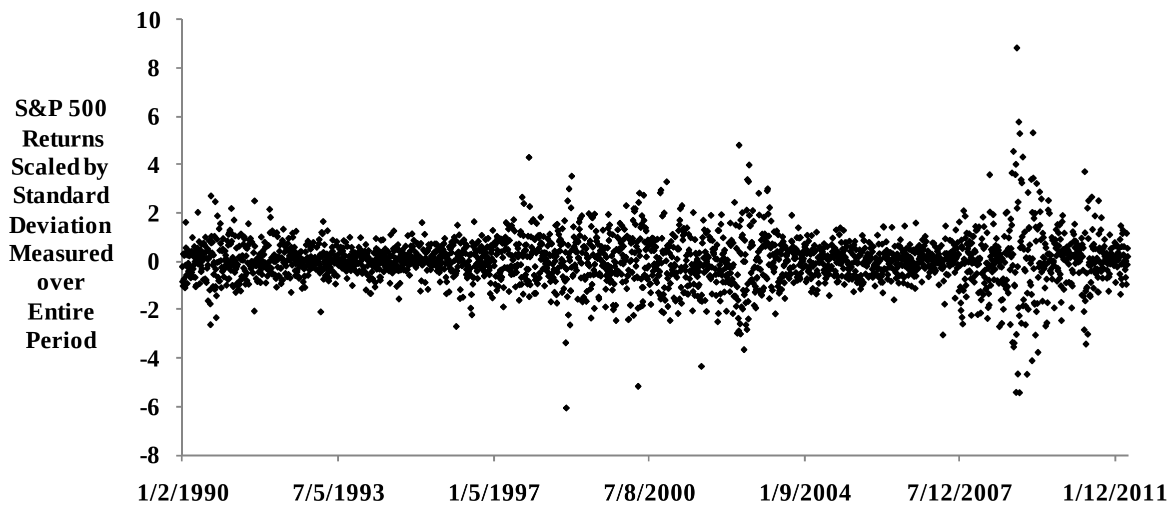
Figure 1 contains a plot of the S&P Index over the 21-year period, and Figure 2 shows the time series plot of returns scaled by a single 21-year estimate of standard deviation of the returns.

[4]It is constructed with the new VIX methodology.

Figure 1

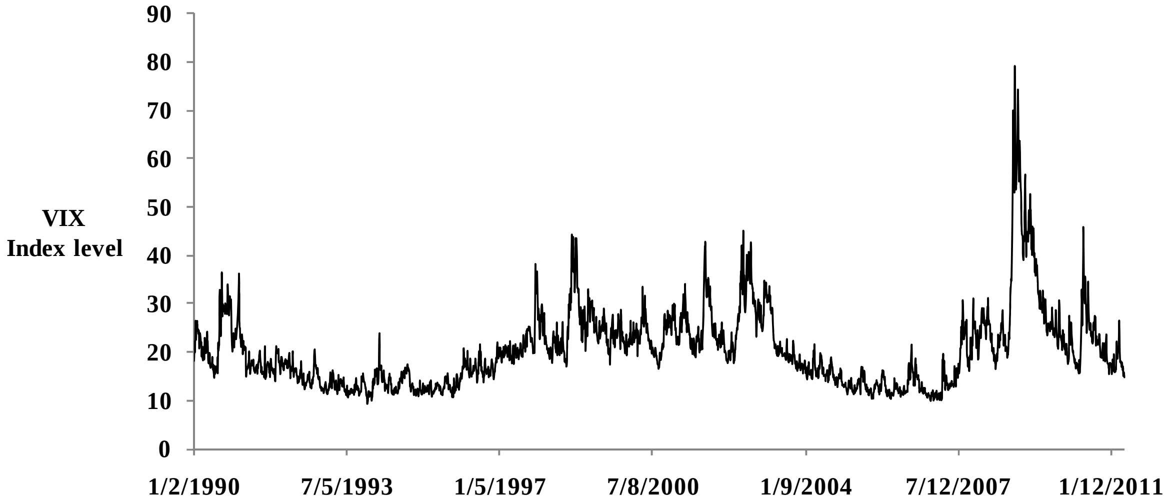
Figure 2 shows that an unconditional model (which by necessity has constant volatility) will give rise to frequent "four sigma" events along with occasional "eight sigma" events as seen during the financial crisis.

Figure 2



The graph in Figure 2 would look roughly the same if the standard deviation used in the normalization was based on a slow moving-average estimate, e.g. a standard deviation based on a five–year window. Figure 3 shows the time series for the VIX Index over the twenty year period.

Figure 3



Clearly the level of the VIX Index shifts around considerably over the period, with a "high VIX regime" prevailing during the financial crisis. Given that the VIX is based on the market's perception of the probability distribution for one-month ahead S&P Index price changes as revealed by the traded prices for one-month index options, it is reasonable to take the time series of the VIX as a proxy for the market's shifting forecast of S&P Index volatility. Figure 4 contains an analog to the plot in Figure 2 where, in Figure 4, S&P returns have been scaled by the *prior-day* VIX as a proxy for a conditional forecast of return volatility. This prior-day VIX as a forecast of volatility potentially includes the "EGARCH effect" of S&P Index change on the forecast volatility.[5] Just eyeballing[6] Figure 4, it is obvious that the returns that appeared extreme relative to an *i.i.d.* Gaussian distribution over the twenty-year period are dramatically less extreme when standardized by their *one-day lagged*-VIX proxy for forecast volatility – in more formal language, when the daily returns are subordinated to the stochastic volatility process.

[5] That is, changes in VIX are persistently negatively correlated with changes in the level of the S&P, e.g. if the S&P on day *t* goes down, the VIX on day *t* tends to go up, all else equal, and thus our forecast of volatility on day *t+1*.
[6] The late University of Chicago statistician Harry Roberts coined the term "interocular analysis" as a tongue-in-cheek label one might use to appear sophisticated when eyeballing the data.

Figure 4

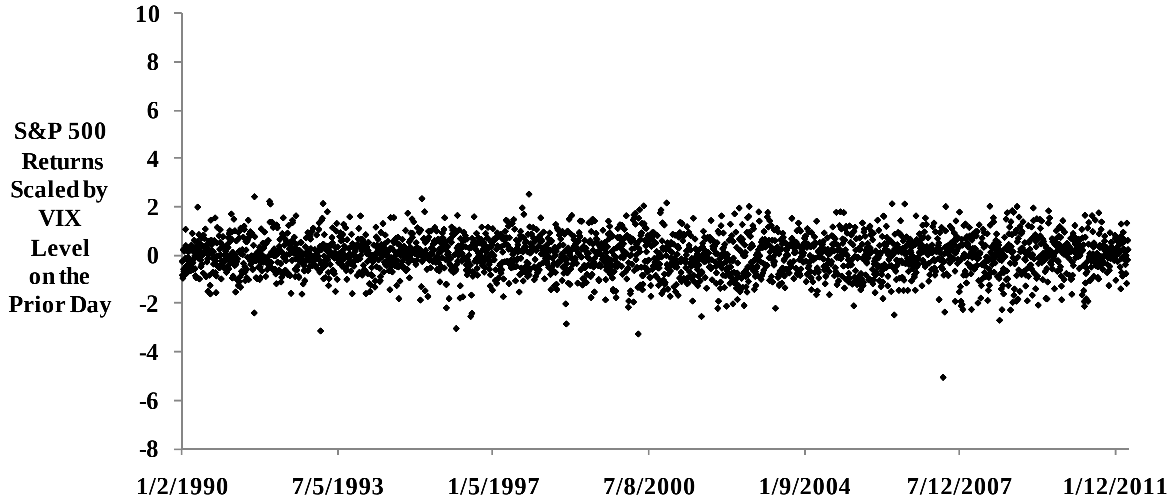


Table 1 contains summary statistics for the standardized returns; the kurtosis for the daily returns normalized by prior-day VIX is 4.47, formalizing one's visual impression from Figure 4 that the standardized returns look relatively Gaussian. In the language above, the prior-day VIX incorporates "knowable unknowns" about S&P returns, and the distribution of these known unknowns conditional on prior-day information is rather well-described as Gaussian uncertainty. As mentioned above, Bulla (2008) shows how, at the S&P Index level, daily returns may be slightly better fit with a t-distribution alternative to the Gaussian conditional on being in a crisis.

Table 1
Summary Statistics for S&P returns over the period January 2, 1990 to May 2, 2011. Statistics for S&P returns scaled by a single constant in-sample standard deviation are in the left-hand column; for S&P returns scaled by prior-day VIX as a conditional forecast of volatility are in the right-hand column

	S&P 500 Returns Scaled by Standard Deviation Measured over Entire Period	S&P 500 Returns Scaled by VIX on Preceding Day
Mean	0.021240	0.019613
Median	0.046141	0.048874
Standard Deviation	1.000000	0.772626
Kurtosis	11.912724	4.474777
Skewness	-0.203650	-0.365445
Minimum	-8.121683	-5.031819
Maximum	9.397617	3.307484
Chance of Seeing Minimum or Less if Normal	1 Chance in 838,000,000,000	1 Chance in 825
Chance of Seeing Maximum or More if Normal	Essentially Zero Chance	1 Chance in 1.086

The subordinated model also has intuitive appeal, in the sense that in practice we observe episodes of "intense action" in the market (many events compounding each other) interspersed with "slow days." It is apparent from Figure 2 that there is clustering of extremes in realized price changes associated with episodes of high volatility and increased flow of information.[7] The same simple illustration just given for S&P returns has also been used in studying exchange rates, where GARCH forecasts of volatility are used to scale daily exchange rates, with roughly similar results. Quantal PRO uses a model for the conditional variance-covariance matrix that incorporates a conditional multivariate Gaussian factor structure, and incorporates a non-linear "crisis factor" in crisis regimes (which has the same practical import as positing a student-t distribution in the crisis scenario).

It is well known that the measured kurtosis of the distribution of stock market returns decreases when the returns are measured over longer intervals. The convergence of longer-interval returns to a Gaussian distribution, i.e., the amelioration of extreme events, would be foreordained if the distribution of daily returns satisfies the conditions of one of the versions of

[7] The subordinated model is arguably more transparent than a complicated copula or extreme value distribution, and we argue in Marsh and Pfleiderer (2010) that risk model transparency is particularly desirable when financial institutions have incentives to understate risk.

the central limit theorem. Practitioners often use a maximum drawdown measure to account for the extreme value risk associated with a given strategy. Figure 5 contains a plot of the maximum drawdown in *annual* returns on the S&P 500 Index, extended by the Global Financial database to go back to 1801. The reference bar on the figure is the drawdown of 50.95% that occurred in 2008. Notice that a drawdown event of the 2008 magnitude has occurred 3 times in 211 years, i.e., the chance of a drawdown has historically been about once every 70 years or 14% each decade.

Figure 5
Maximum S&P 500 Drawdown over 24 Months from 1801 to Present

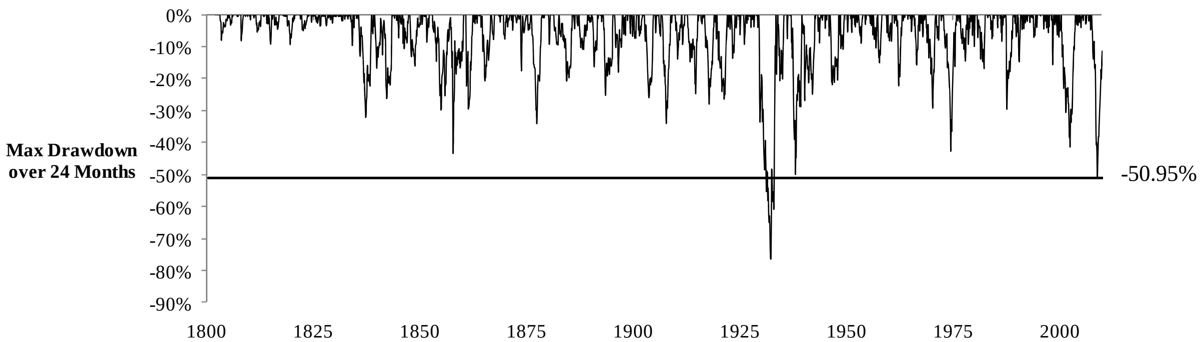
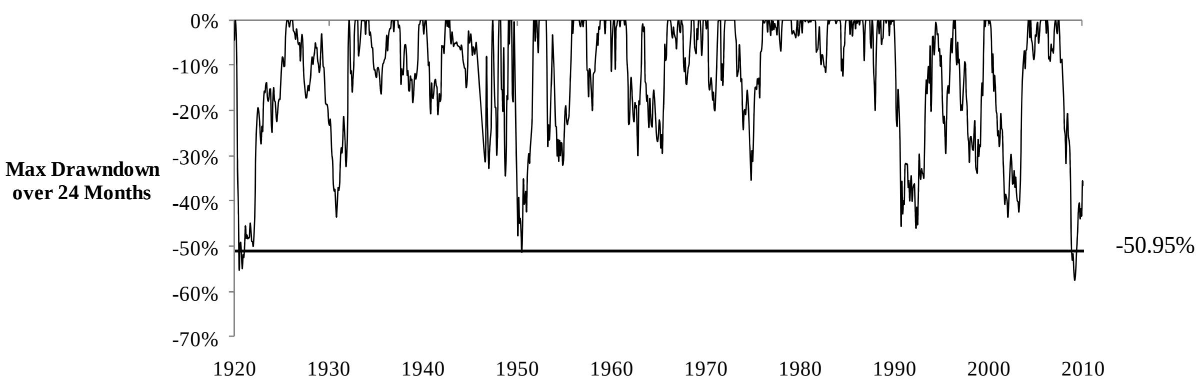


Figure 6 provides a graph of maximum drawdown similar to that in Figure 5, but instead for Japan's Topix Index from 1920. There, a maximum drawdown comparable to the 50.95% in the U.S. in 2008 has occurred 3 times in the 91 years of twentieth century data, the same rate as for the S&P 500 Index measured over the same century. At least two of the three points of maximum drawdown in the U.S. and Japan are co-incident, evidence of the importance of global catastrophe factors in the twentieth century.

Figure 6
Maximum Topix Drawdown over 24 Months from 1920 to Present



The preceding discussion has centered on equity index returns. As is well known, equity returns are already amplified somewhat by the leverage due to corporate debt outstanding. Fund management strategies or equity derivatives can, in turn, multiply that in-built equity leverage many times over, thereby accentuating the sensitivity to extreme event model misspecification. The beginnings of the financial crisis are usually traced to the over-the-counter (OTC) market for structured debt instruments. The freeze-ups in that market and the feedback to the risk of highlylevered and poorly-hedged financial institutions had a knock-on systemic effect. And while the financial crisis may have "started out" in the OTC market, it appears to be "finishing up" in the sovereign debt market. Some people profess to have clearly foreseen the crisis, and this may lead one to question whether the precipitating causes of the crisis were indeed unforeseeable black swan events. We have argued elsewhere, along with many others, that an important factor in understanding the crisis was the distorted incentives to control and monitor risk, "fat-tailed" or otherwise. At the same time, one "knowable unknown" that was arguably missing in many existing debt models was illiquidity risk, the "…risk…that…a…bond's value will drop when aggregate liquidity deteriorates" (Lin, Wang, and Wu (2011, p. 629)). But this debt exposure to drops in aggregate liquidity in a crisis seems likely not to be an unforeseeable black swan; indeed Lin, Wang, and Wu's and Bao, Pan, and Wang's (2010) evidence is that it is also already priced in the yield premiums on corporate bonds.

5. Summary and Discussion

We have shown that the frequency of "black swan" surprises in financial returns is reduced when conditional forecasts of risk are updated in a timely manner. The logic for updating the forecasts is that there are non-transitory shifts in risk over time, which in turn is consistent with there being variations over time in intensity of the "bursts" of economic events and trading that affect asset price changes. The nature of these "bursts" is almost surely changing over time due to rapid improvements in technology for collecting, disseminating, and analyzing structured and so-called "unstructured" data. These improvements can be expected to further reduce the frequency of some remaining types of "black swan" extreme events – previously inconceivable events can become conceivable with better statistics and better research. Of course neither the best information technology, nor the best risk model based on the information available, will be a panacea against investment losses. Nor does progress on the data side or the model side mean that we won't sometimes see abrupt changes in the levels of investment risk.[8] Indeed, in some models, enhanced information flows potentially speed up contagion effects and "systemic" black swan events. The bottom line, however, is that "black swans" are hardly immutable, nor does that moniker *per se* provide any insight as to whether they are becoming more, or less, manageable.

[8]Samuelson (1976) warned that it is not optimal to fixate on forecasting turning points in economic series, the analogy here being forecasts of abrupt changes in risk; intuitively, doing so generates too many false predictions given a mean-square-error loss function.

REFERENCES

1) Bao, Jack, Jun Pan, and Jiang Wang, 2010, "The Illiquidity of Corporate Bonds," Unpublished paper, Ohio State University, July 15; forthcoming: *Journal of Finance*.

2) Barro, Robert J., 2006, "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 121(3), pp. 823-866.

3) Bateman, Hazel, Jordan Louviere, Stephen Satchell, Towhidul Islam, and Susan Thorp, 2010, "Retirement investor risk tolerance in tranquil and crisis periods: experimental survey evidence," School of Finance and Economics, University of Technology, Sydney, March.

4) Bank for International Settlements, 2009, "Revisions to the Basel II Market Risk Framework," July: http://www.bis.org/publ/bcbs158.pdf?noframes=1

5) Bulla, Jan, 2011, "Hidden Markov models with t components: Increased persistence and other aspects," *Quantitative Finance*, 11 (3), 459-475, July 1.

6) Chako, George, 2008, "Liquidity Risk in Financial Markets," Chapter 7 in: Innovations in Investment Management, edited by H. Gifford Fong, Bloomberg Press.

7) Clark, Peter K., 1973, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), January, 135-155.

8) Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, 1998, "Recent developments in modeling unforeseen contingencies, *European Economic Review*, 42, 523-542.

9) Golub, Bennett W., and Conan C. Crum, 2010, "Risk Management Lessons Worth Remembering From the Credit Crisis of 2007 – 2009," Working Paper, Blackrock, February. http://ssrn.com/abstract=1508674

10) Kay, John, 2011, "Don't blame luck when your models misfire," *Financial Times*, March 1, p. 9.

11) Lin, Hai, Junbo Wang, and Chunchi Wu, 2011, "Liquidity risk and expected corporate bond returns," *Journal of Financial Economics*, 99, 628-650.

12) Marsh, Terry, and Paul Pfleiderer, 2009, "The 2009-2009 Financial Crisis: Risk Model Transparency and Incentives," Preface in: *The Recent Trend of Hedge Fund Strategies*, Ed: Yasuaki Watanabe, Nova Publishers.

13) Marsh, Terry, and Paul Pfleiderer, 2011, "Financial Modeling and the Crisis II: Asset Allocation ," Manuscript, Stanford University.

14) Reitz, Thomas A., 1988, "The Equity Risk Premium: A solution," *Journal of Monetary Economics*, 22, 117-131.

15) Samuelson, Paul A., 1976, "Optimality of Sluggish Predictors under Ergodic Probabilities," *International Economic Review*, 17(1), February, 1-7.