Behavioral approach to market and default risks modeling

Taguedong, Sylvain Chamberlain

27 December 2009

BEHAVIORAL APPROACH TO MARKET AND DEFAULT RISKS MODELING

By ***Sylvain Chamberlain Taguedong***

A thesis presented to the Department of Economics of the University of Lugano in partial fulfillment of the requirements for the degree of ***Master of Sciences in Finance***.

Advisor:
Prof. ***Giovanni Barone-Adesi*** Director of Swiss Finance Institute, University of Lugano.

Co-Advisor:
Prof. ***Enrico De Giorgi***
Universities of St. Gallen and Lugano.

Lugano, Switzerland

Abstract

In this paper we discuss popular market and default risks modeling. We highlight some shortcomings. Then, we present the prospect and cumulative prospect theories. We discuss again the previous models under behavioral finance framework and get different results. Based on these results, we propose a new Value at Risk measure and make suggestions on other measures.

Table of Contents

Dedication

To my father, His Majesty *Jean-Pierre Taguedong*;
and to my mother, *Regine Kenfack*

Acknowledgements

I thank my brothers and sisters for their support;
my Dear friend, the Chartered Accountant *Elvira Siani*;
Antonietta Castelnuovo, Swatch Group S.A.;
Fra' Martini Dotta, Director of SOS Ticino.

Preface

The initial project was a study of the most popular risk measures with assets' dynamics linked to main fundamental exogenous risks. Which turned out too complex after an explorative analysis. But thanks to behavioral finance course, the desire has been almost compelling to integrate classical models into a framework of decision under uncertainty and irrationality.

I want to express my sincere gratitude to my supervisors, especially to Prof. ***Giovanni Barone-Adesi.***

I'm also grateful to the United States Bureau of Labor and Statistics (U.S. BLS) and Moody's Analytics for their feedbacks during the explorative phase of this project, they gave a twist to the scope of the project.

Chapter 1.

Introduction

In past decades, market and credit risks have become a major concern for regulators and financial companies. Their impact, frequency and spread have prompted banks and regulators to manage and control them in order to prevent bankruptcy and preserve the financial system stability. After the Great Depression, banks are regulated by the Glass Steagall act in 1933 and Bretton Woods agreements in 1944. From early 1990, they mostly follow the Basel guidelines. Unlike regulations in the mid twentieth century, which were mostly restrictions on financial transactions, Basel Accords are recommendations regarding the Regulatory Capital. In this framework, internationally active banks are allowed to operate either with the standard approach or with the internal rating-based approach.

Regarding the internal rating-based regime, many models for assessing risk are proposed in academe and in the industry. In practice, the Value at Risk (VaR) has been adopted as the standard for market risk and most of the biggest banks have their own model for credit risk. At portfolio level, the main models used to assess credit risk are: JP Morgan's CreditMetrics, Moody's KMV, Credit Suisse's CreditRisk$^+$ and McKenzy & Company's Credit Portfolio View. Despite precautionary measures taken in recent years, banks and financial system aren't yet sound. The demise of some wall street giants as Bear Stearns and Lehman Brothers and the record fall in macroeconomics indicators in many countries are illustrations of some deficiencies present in the current regulatory framework. The assessment of risk poses many challenges in, mainly within the internal rating based approach, both the modeling process and the estimation technique.

Our intention in this paper isn't to give exhaustively the causes of models' failure. Rather, we try to highlight some flaws in the modeling process and in classical finance theory. Then, under behavioral finance theory, present a new methodology which circumvents partly those imperfections.

First of all, we define what do market and credit risk mean and highlight features of some models widely used in the industry. Secondly, we analyze traditional estimation techniques for both parametric and non parametric models. Thirdly, we introduce behavioral finance and present the cumulative prospect theory. In chapter fourth, we talk about modeling under behavioral finance assumptions. In the fifth chapter, we analyze the implications of irrational behavior to Value at Risk and probability of default. Then, we present a simple Value at Risk model estimator based on judgment and some usually non accounted facts in estimation. In the end, we draw some concluding remarks.

Chapter 2.

Market and credit risks models

1. Market risk

Market risk is the possibility of losses in on-balance-sheet and off-balance-sheet positions driven by market prices fluctuations. Markets risks can be distinguished in:

- Risks related to interest rate instruments in trading book
- Stock prices risks in trading book
- Foreign exchange and commodities risks

There are many models for assessing market risk. But, due to their importance, we only consider the Value at Risk (VaR) and the Expected Shortfall.

1.1. Value at Risk

Value at risk is traced back to early 1990s when JP Morgan's chairman demanded a report, "Report 4.15", summarizing market movements and providing the bank's risk for the next day.

In a financial point of view, it is the margin needed to cover both expected loss and unexpected loss.

From a mathematical standpoint VaR is, given a confidence level and a time horizon, the maximum loss that a portfolio incurs based on the distribution of price changes over a given historical period. This is formulated as follows:

Discrete case

Let X be a discrete random variable with observations $x_i \in \mathbb{R}, i = 1, 2, 3, \dots, n$;

$x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$ is the ordered sequence of x_i; $\mathbb{P}[X = x_i]$ is the probability mass function and α is the confidence level;

then

$$VaR = -x_{k^*};$$

where

$$k^* = min\{k \in \mathbb{N} | \sum_{i=1}^{k-1} \mathbb{P}[X = x_i] < \alpha \leq \sum_{i=1}^{k} \mathbb{P}[X = x_i]\}.$$

Continuous case

Let X be a continuous random variable; $\mathbb{P}[X \leq x]$ the probability density function and α the confidence level

then

$$VaR = -inf\{x \in \mathbb{R} | \mathbb{P}[X \leq x] \geq \alpha \}.$$ [1]

Value at Risk is recommended by Bank for International Settlements and the Group of Thirty as a good method for market risk mitigation. Conceptually, VaR is simple. Although recently some doubt has been cast on its validity, it remains one of the best practice for risk management. From a managerial viewpoint, it's the required margin to cover partially potential losses. Value at Risk is particularly important from the regulatory standpoint in the sense that it's a tool to protect depositors and preserve financial stability. VaR and other risk measures are becoming more and more demanded (especially when it comes to control proprietary trading on a rational basis) as the financial environment is growing in complexity.

Given that VaR is not always sub-additive outside the Gaussian world, which means that the benefit of diversification is mitigated (Since it's determined on a consolidated basis), coherent models such as expected shortfall have been proposed.

1.2. Expected shortfall

Provided a VaR with confidence level α the expected shortfall is defined as follows:

$$ES = -\frac{(\mathbb{E}[X|X \leq -VaR_\alpha] - VaR_\alpha (\alpha - \mathbb{P}[-X|X \leq -VaR_\alpha]))}{\alpha}.$$

This measure has two main properties. First, it's more conservative than VaR. Second, it captures the effect of diversification. Which means that a diversified portfolio would imply less capital cushion rather than more capital cushion.

[1] Note that the continuous and discrete cases can be consolidate to $VaR = -sup\{x \in \mathbb{R} | \mathbb{P}[X \leq x] \leq \alpha\}$.

2. Credit risk

Credit risk is "the possibility of loss incurred as result of a borrower or a counterparty failing to meet its financial obligations" (Credit Suisse group, 2002).

As it appears from this definition, credit risk permeates all financial contracts and is one of the main concern. The literature on credit risk measurement is vast. In recent times, it has shown a particular vigor due to a more complex environment and the need of pricing more and more complicated financial instruments (for instance, Credit Default Swap). The credit risk methodologies, although different from an estimation viewpoint, all address one or more of the following risk aspects:

- Probability of default
- Loss given default
- Exposure at default

Borrowing the classification in CreditMetrics™—Technical Document (2007), we can distinguish the techniques for estimating the probability of default in:

- Accounting analytics methods
- Statistical methods
- Option theoretic methods

Accounting analytics rely on financial statements measures such as coverage ratios, earnings, cash flows etc... There are then adjusted by an expert judgment to obtain ratings. The prominent are S&P and Moody's ratings.

Statistical methods can be distinguished in:

- Discriminant analysis, a combination of variables that determines default state is found. For instance Edward Altman's zeta-scores
- Regression analysis, a combination of variables that estimates the default likelihood is found. For instance McKenzy's Credit Portfolio View
- Neural network

Option theoretic methods consider that the company's equity is a call option on the asset value and that default occurs when asset value hits a given threshold (value of the obligation). The likelihood of default is derived from that assumption. The most notable model is Moody's KMV.

In the following, we illustrate some popular portfolio credit risk models for estimating the likelihood of default.

2.1. Merton model

Model assumptions:

1. Perfect market:
    - No transactions costs, taxes and indefinite divisibility of assets.
    - Sufficient number of investors who can trade, at the market price, as much as they desire.
    - Existence of a market of borrowing and lending at the same rate.
    - No short-sale constraints
2. Continuous trading
3. Modigliani-Miller theorem holds
4. Flat term structure
5. Firm value follows a diffusion process

$$dV_t = (\mu V_t - c)dt + \sigma V_t dz;$$

where V_t is the firm value at time t, c is the dividend pay-out (in dollar), μ is the annualized expected rate of return, σ is the volatility.

Furthermore, for technical tractability of the model, it's assumed that the capital structure is simple (one class of homogenous debt, say D at time t, and the residual claim).

By Itô lemma and for $c = 0$ we have

$$\ln(V_t) - \ln(V_0) = \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma z_t;$$

thus

$$\mathbb{P}[V_t \leq D] = \Phi\left(-\frac{\ln(V_0) - \ln(D) + \left(\mu - \frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}}\right).$$

2.2. Moody's KMV EDF

It's a tool to obtain the probability of default (EDF) where technical assumptions in Merton model are made more realistic. The model considers:

- Complex capital structure
- Default can happen at any time before maturity
- The default point is firm specific, depends on firm's liability structure and it is derived empirically
- The distance to default (DD) is matched to the probability of default via an empirical distribution

By definition

$$DD = \frac{\ln\left(\frac{V_t}{D_t + at}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}};$$

where a are cash outflows per unit time (interest payments, coupons and dividends).
Then,Vasicek and Kealhofer found that

$$EDF \simeq 2\Phi(-DD).$$

Some claimed features of the model are:

- EDF is a forward-looking measure
- EDF is an actual probability of default
- Non statistical fitted model, based on cause and effect

2.3. CreditMetrics

Credit metrics is mainly a methodology of computing credit risk from already computed default probabilities. Transition matrices are given by other entities like S&P and Moody's. The value payoff is derived by discounting at the corresponding forward zero curves (which is derived from zero coupon prices).

2.4. Jarrow Turnbull model

Firms are grouped into classes, default state is absorbing and the default frequency is assumed to follow a Poisson process (or a Cox process, i.e. Poisson process with stochastic intensity).

Let N_t be the default frequency of a class during the time interval $(t, t+\Delta t]$; and $\lambda(t)$, the annualized mean default frequency. $\lambda(t)$ is a function of default free interest rates and an equity market index.

Then N_t has probability mass function

$$\frac{(\lambda(t)\Delta t)^n}{n!} e^{-\lambda(t)\Delta t} \qquad\qquad t \in [0, \infty).$$

Let I_t be the default indicator at time t; τ the waiting time until default occurs during the time interval $(t, t+\Delta t]$

$$I_t = \begin{cases} 1, & if\ default \\ 0, & otherwise \end{cases}$$

τ is a stopping time such that

$$\mathbb{P}[\tau > t] = \mathbb{P}[no\ default\ before\ t] = \prod_{i=1}^{k} \frac{(\lambda(t_i)\Delta t)^0}{0!} e^{-\lambda(t_i)\Delta t};$$

where

$$k \in \mathbb{N}\ and\ k\Delta t = t$$

i.e.

$$\mathbb{P}[\tau > t] = e^{-\sum_{i=1}^{k} \lambda(t_i)\Delta t}$$

which is equivalent to

$$\mathbb{P}[\tau \le t] = 1 - e^{-\sum_{1}^{k} \lambda(t_i)\Delta t}.$$

The above function is the exponential distribution. Since default state is absorbing, the likelihood of default before time t, $\mathbb{P}[default\ before\ t]$, is

$$1 - \ e^{-\ \sum_1^k \lambda(t_i)\Delta t}.$$

By independence of default events we have

$$\mathbb{P}[no\ default\ before\ t + \Delta t]$$
$$= \mathbb{P}[no\ default\ before\ t]\mathbb{P}\big[no\ default\ during\ (t, t + \Delta t]\big]$$

Which implies that

$$\mathbb{P}\big[no\ default\ during\ (t, t + \Delta t]\big] = e^{\lambda(t)\Delta t}$$

thus

$\mathbb{P}\big[default\ during\ (t, t + \Delta t]\big] = 1 - e^{\lambda(t)\Delta t}$, an Exponential Distribution

We have

$$\ln\big(\mathbb{P}\big[no\ default\ during\ (t, t + \Delta t]\big]\big) = \lambda(t)\Delta t$$

For $\lambda(t)\Delta t$ small (i.e. $\Delta t \to 0$), we have

$$\mathbb{P}\big[no\ default\ during\ (t, t + \Delta t]\big] \simeq 1 - \lambda(t)\Delta t$$

we deduce that

$$\mathbb{P}\big[default\ during\ (t, t + \Delta t]\big] \simeq \lambda(t)\Delta t.$$

So, default is completely characterized by the hazard function, $\lambda(t)$.

2.5. Credit risk+

Assets in a portfolio are divided by sectors. Where a sector is a group of obligors affected by the same main factor (like domiciliation in a country). The main assumption in this model is that asset default rate is assumed to be binomially distributed. The probability generating function is used as a device to compute aggregate default.

Granted a portfolio of N obligors, $\mathbb{P}_i$ the annual probability of obligor i defaulting; the probability generating function of obligor i is

$$\mathcal{P}_i(t) = (1 - \mathbb{P}_i)t^0 + \mathbb{P}_i t^1 = 1 + \mathbb{P}_i(t-1);\ |t| < 1.$$

The model offers two measures according to whether the default state is stable or changing.

For constant default rate

By independence of default between different obligors, the probability generating function of the portfolio is

$$\mathcal{P}(\mathrm{t}) = \prod_{\mathrm{i}=1}^{\mathrm{N}} \mathcal{P}_i(t) = \prod_{\mathrm{i}=1}^{\mathrm{N}} (1 + \mathbb{P}_i(t-1))$$

$$\ln\big(\mathcal{P}(\mathrm{t})\big) = \sum_{i=1}^{N} \ln(1 + \mathbb{P}_i(t-1))$$

For $\mathbb{P}_i$ small

$$\ln(1 + \mathbb{P}_i(t-1)) = \mathbb{P}_i(t-1)$$

thus

$$\mathcal{P}(\mathrm{t}) = \mathrm{e}^{\sum_{i=1}^{N} \mathbb{P}_i(t-1)} = \mathrm{e}^{(t-1)\sum_{i=1}^{N} \mathbb{P}_i} = \mathrm{e}^{(t-1)\mu} \mathrm{e}^{t\mu} \mathrm{e}^{-\mu} = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu} \mu^n}{n!}\, t^n,$$

where $\sum_{i=1}^{N} \mathbb{P}_i = \mu$ is the expected number of defaults of the sector (representative obligor).

We deduce that the portfolio probability of n defaults is

$$\mathbb{P}[n\ defaults] = \frac{\mathrm{e}^{-\mu}\mu^n}{n!}, \text{ a Poisson Distribution.}$$

Empirically, the average default probability μ is random. This is due to background factors such as the state of the economy. We mean that an obligor will more likely default in an economy downturn than in an economy upturn. Thus, the variability of default rate must be taken into account in assessing the portfolio probability of default. This leads to the determination of probability with changing default rate.

For variable default rate

We have $\mathcal{P}_{\mathrm{X}}(\mathrm{t}) = \sum_{n=0}^{\infty} t^n\, \mathbb{P}[n\ defaults|X] = \mathrm{e}^{(t-1)X}$, where X is the variable default rate and $\mathcal{P}_{\mathrm{X}}(\mathrm{t})$ is the conditional probability generating function.

Let f(x) be the density function of X, then the unconditional probability generating function is

$$\mathcal{P}(\mathrm{z}) = \int_0^{\infty} \mathrm{e}^{(t-1)x}\, \mathrm{f(x)}dx.$$

Assuming that $X \sim \Gamma(\alpha, \beta)$, i.e. $\mathrm{f(x)} = \frac{\mathrm{e}^{\frac{-x}{\beta}} \mathrm{x}^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}$, we have

$$\mathcal{P}(\mathrm{t}) = \int_0^{\infty} \mathrm{e}^{(t-1)x}\, \mathrm{f(x)}dx$$

$$= (1-\mathrm{q})^{\alpha} \sum_{n=1}^{\infty} \binom{n+\alpha-1}{n} q^n t^n, \text{ by the generalized binomial theorem}$$

where

$$q = \frac{\beta}{1+\beta}.$$

We deduce that

$$\mathbb{P}[n\ defaults] = (1-\mathrm{q})^{\alpha}\binom{n+\alpha-1}{n}q^{n}, \text{ a Negative Binomial Distribution.}$$

2.6. Portfolio-view (by Financial Analytics)

The idea is that the likelihood to default of a company, in a specific industry, depends mostly on a combination of macroeconomic factors. Macro-variables can be growth rate of the economy, interest rates, foreign exchange rate, unemployment, etc...(Wilson, 1997). The default rate in the sector is given by a logistic distribution. Then, to obtain the company default probability, the default probability in the sector is accommodated to company specific factors.

Let $X_t^i \ |i = 1,2,\dots,n$ be the main n macro variables affecting the sector at a given time t; a the vector of parameters. We have

$$\mathbb{P} = \frac{1}{1+e^{(Y_t)}};$$

where

$$Y_t = a'\begin{pmatrix}1\\ X_t\end{pmatrix} + \varepsilon_t$$

and

$$X_t = (X_t^1,\ X_t^2, \dots,\ X_t^n)'\ ; a = (a_0, a_1, \dots, a_n)'.$$

ε_t are assumed to be multivariate normal distributions with mean 0 and variance-covariance matrix Σ.

Then, the probability of default of an individual is given by adapting $\mathbb{P}$ to the individual's characteristics such as size and capital structure.

Summarizing, the main feature of structural models is that the probability of default is drawn from a normal distribution, under some conditions; and they allow us to mark to market the default risk.

Whereas the reduced form models depend on default history of a pool of companies. The probability of default is obtained from an Exponential Distribution in Jarrow Turnbull case, Poisson Distribution and Negative Binomial Distribution in Credit Suisse model.

Financial analytics derives the probability of default from a Logistic Distribution.

Chapter 3.

Estimation under classical finance

The main challenge resides in statistical estimation. Any measurement process is separable in three phases:

1. Investigation of data statistical properties
2. Estimation of distribution function
3. Computation

In practice, all methods rely on some regularities found in financial data. Some stylized facts commonly met in stocks' prices are:

- Low autocorrelation of returns
- Significant autocorrelation of squared returns
- Negative skewness
- Fat tails

Since the remarkable work by Bachelier, "Theory della speculation", there is an extensive literature describing statistical features of financial prices. Bachelier describes prices fluctuations as independent and Gaussian distributed.

Over the past decade, ARCH type models have been the prevailing way of representation of financial prices. The main feature of price behavior captured by arch models, Engle (1982), is volatility clustering (that is, large price changes are followed by large price changes and small price changes are followed by small prices changes) with undefined direction. GARCH model renders arch more flexible by including long lag length. Other variants like FIGARCH are designed to accommodate the long memory (that is the longest cycle is proportional to the sample) and structural breaks observed in some data.

1. Non parametric models

According to this methodology there is not an a priori distribution. The goal is to determine an empirical distribution. Non parametric methods rely heavily on the strong law of large numbers, I mean, on the i.i.d and stationary assumptions. This is mostly met when determining losses distribution and transition matrices in credit risk models such as Creditmetrics. Unfortunately, the assumption of independency is called into question by the cyclicities that often appear in financial phenomena. The dependency and even the long dependency are well supported by Mandelbrot's works with Van Ness on Fractional Brownian Motion (FBM); and with Fischer and Calvet on Multifractality of Asset Returns (MMAR).

The assumption of stationarity is inconsistent with structural changes in the economy, in sectors of the economy and in businesses. For instance, this is illustrated by the fluctuations of S&P 500 on the graph below in which periods of high volatility contrast with periods of low volatility. Particularly, the behaviors before 1980, during 1980-1995 and after 1995 are remarkable.

Another drawback of this approach is the lack of data. In theory, this approach works, all else being equal, when we have an infinite number of observations. The minimum number of observations that can reveal the properties of the phenomenon is generally unknown. Notwithstanding Basel II recommends at least one year period of historical observation for market risk. Concerning reduced-form models, the estimation is even worse provided that we have only one year horizon (that is yearly observations).

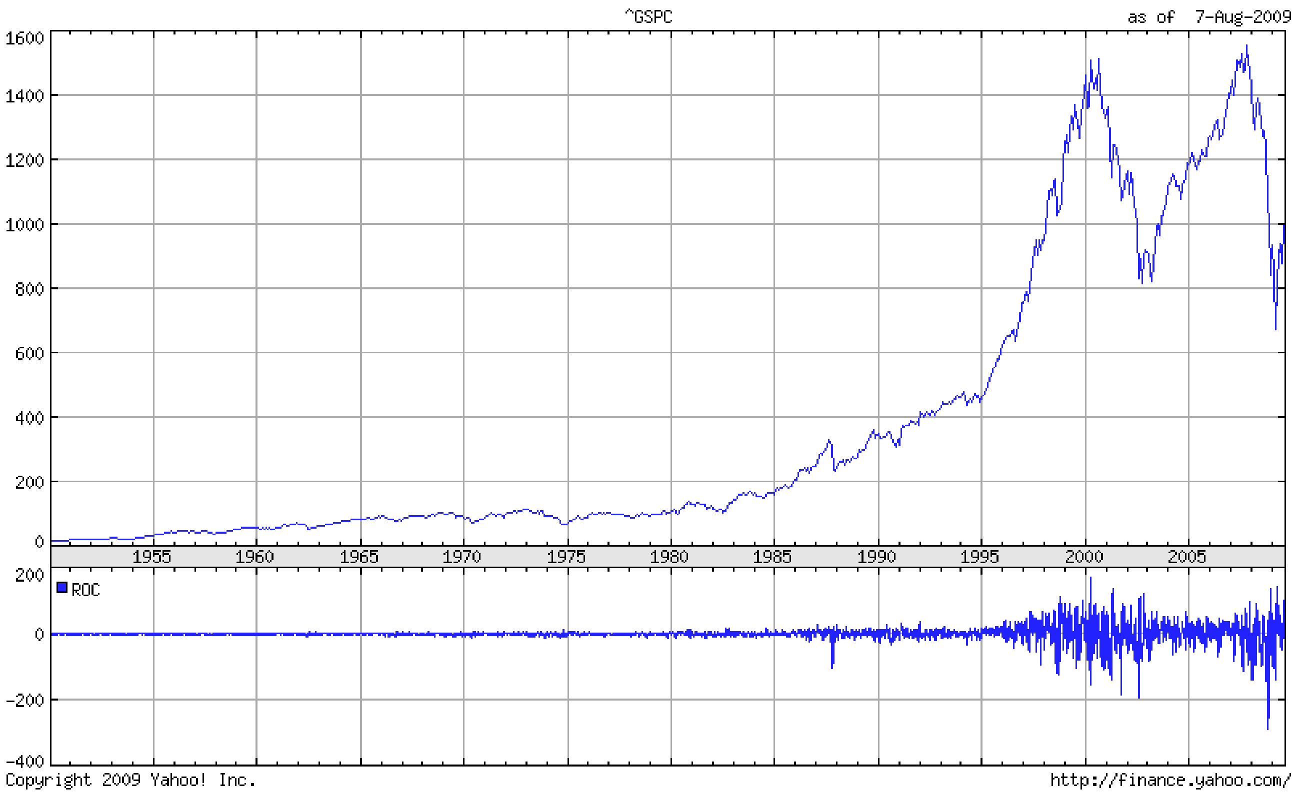

Figure1: S&P 500 index. Prices and returns from 08/08/1951 to 04/09/2009.

Source: yahoo finance

2. Parametric models

The idea is to infer model parameters for an assumed distribution of risk factors. Commonly the assumed distribution is tested on observations (Kolmogorov Smirnov test). In fact, temporal aggregation of the logarithm of absolute returns, for sufficiently large horizon and under other regularity conditions, are used to justify the normality of the logarithm of return during that period.

In principle, central limit theorem works for infinite number of observations. Which is barely found in practice. Also, the large number of observations that can give a good approximation, specially for tail events, is unknown. Even provided a good number of observations, the central part of the distribution of the aggregate fit the Gaussian better than the tails. Bouchaud, J.P. and Potter, M (2000). Hopelessly, this makes the Gaussian approximation at the tails (the most important parts from a risk management viewpoint) less meaningful.

3. Thought experiment

We've seen that is difficult to estimate the distribution, especially at the tails. The implication is that any prior distribution or posterior distribution may be wrong. In contrast to distributions usually assumed in practice (Gaussian, Student-t), this leaves room to a variety of distributions. So, any inferred distribution is to some extent subjective (L. J. Savage, 1962). Thus, the risk measure inferred from the analyst model might be very different from the true one (which is unknown due to the almost impossibility to determine the real distribution). In the following we consider the cases in which a risk manager, provided its distribution inferred, may end up.

First, let us point that the VaR doesn't give any protection for event with amplitude larger than VaR. We mean that, no matter the coincidence of the true and inferred VaR and how small are the likelihoods of events above VaR, still we aren't immune at any time from catastrophic events. This is simply due to the randomness of the position and even worst when it shows up like an ergodic Markov process[2] (meaning that a past catastrophe doesn't reveal enough information which would help to its predictability).

[2] Aperiodic recurrent markov process.

As we have seen before, the true distribution of a portfolio is basically unknown. Since there is no reason to confine the set of possible distributions to 'bell curve' and 'Student-t' distributions and, in general, to well-behaved distributions, we consider that a portfolio can in theory have any kind of distribution. To see the implication of this fact on risk control procedure, we consider two classes of distributions:

- Type1 distribution, with finite mean and variance
- Type2 distribution, with infinite or undefined mean and variance

Also let assume the standard deviation as our risk measure.

In situation where the real distribution is of type2, our portfolio is exposed to huge loss at any time. There is not safety margin that can reduce the risk we are bearing, because we expect catastrophic event at any time. In this case, the only strategies to mitigate the risk is to hedge it out, avoid short positions [3](especially in derivative markets) or simply exit from the market.

When the distribution is of type1, traders would keep lower margin than required and banks would keep less capital than required if the inferred VaR is smaller than the true one. Which also means that they would be taking more risk instead of reducing it.

Besides some deficiencies given precedently , the assumption of efficient market hypothesis (which is key in VaR and structural models) has been seriously drawn into question. This leads us to the behavioral finance approach to decision under uncertainty.

[3] Let recall protection sellers in credit markets.

Chapter4.

Behavioral finance

Laboratory and real-life experiments have shown that not always decision-makers behave in a rational manner. Owing to limitations in processing information, agents often make choices different from what should be made given information available. Psychological biases arising from this process could be:

- Cognitive bias, which is the judgmental distortion due to cognitive factors (cognitive dissonance, anchoring, framing, etc.)
- Emotional bias, distortion due to emotional factors (greed, fear, etc.)

In practice, there have been found many anomalies which are inconsistent with traditional finance settings. Some notable of them in financial markets are:

- The value premium puzzle (price-earnings ratio effect): historical outperformance of value stocks over growth stocks
- Size premium puzzle (small firm effect): historically, small capitalization stocks have outperformed large capitalization stocks
- Equity premium puzzle: historically, equity has outperformed virtually default-free debt (government bonds). Mehra and Prescott (1985)
- Under-reaction or over-reaction of stock prices to earnings announcement: stocks with higher past returns are overvalued and those with lower past return are undervalued. De Bondt and Thaler (1985)
- Asset allocation puzzle: investors don’t hold the same composition of risky assets. Canner, Mankiw and Weil (1997)
- Price fluctuations with no information: Harris and Gurel (1986) found that there is an average abnormal return of 3.5% for stocks added to S&P 500. Another emblematic example is the stock market crash on Monday, 19, 1987

Attempts have been made to explain anomalies met in decision making under uncertainty. We can recall the Bernoulli method in solving the St Petersburg paradox by a decreasing marginal utility of wealth. We also recall the rank-dependent model which, in order to explain Allais paradox, generalizes the expected utility by applying a cumulative probability weighting function. The most satisfactory explanation of decision making is given by the cumulative prospect theory.

1. Prospect theory and cumulative prospect theory

Following the concept of irrational behavior, Kahneman and Tversky posited the Prospect Theory. Unlike the expected utility theory of Von Neumann and Morgenstern, which is a normative paradigm, the prospect theory is a descriptive pattern of decision making. According to this theory, the decision process follows two steps:

- Editing phase
- Evaluation phase

The editing phase is the preliminary investigation. Prospects are reformulated. Outcomes and probabilities are transformed. The main operations during the editing stage are:

- Coding, outcomes are perceived as gains and losses with respect to some reference point
- Combination, simplification by combining probability with identical outcomes
- Segregation, separation of the riskless component from the risky one
- Cancellation, discarding common components to the prospects
- Simplification, rounding outcomes and probabilities. Discarding highly uncertain outcomes
- Detection of dominance, dominated prospects are discarded

After the editing phase, which mostly alters preferences, prospects are evaluated. During the evaluation stage, weights are assigned to probabilities and values are assigned to outcomes.

Values denote changes in a given portfolio, rather than final states of the portfolio. The value function (Figure2) has the following properties:

- Concave in the domain of gains and convex in the domain of losses, although individual preferences can cause the shape of value function be the other way round
- Loss aversion. That is, losses are bigger threats than gains. This makes the function steeper in the domain of losses

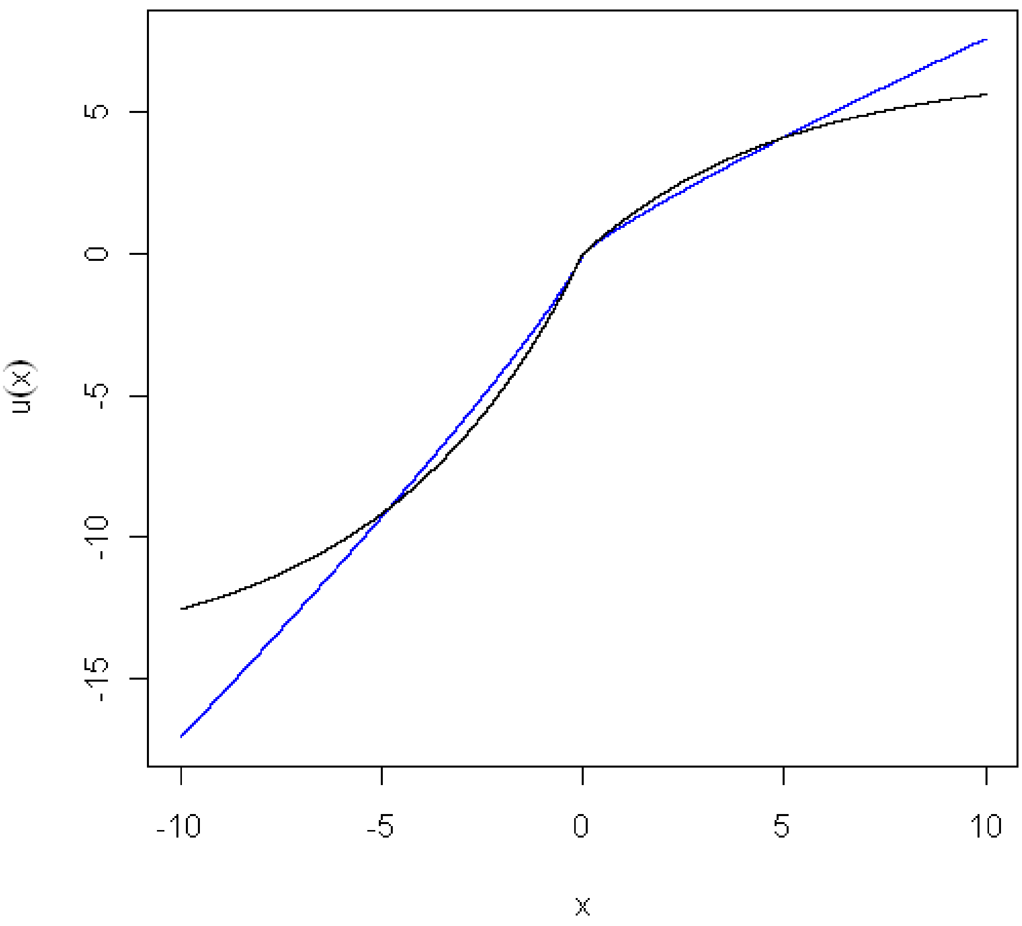

Figure2: Utility function

Tversky Kahneman (1992) utility function (blue line): $u(x) = \begin{cases} -\lambda(-x)^{\beta}, & x < 0 \\ x^{\alpha}, & x \geq 0 \end{cases}$

$\alpha = \beta = 0.88$ **and** $\lambda = 2.25$

Haim Levy, Enrico Degiorgi and Thorsten Hens (2003) utility function (black line): $u(x) = \begin{cases} \lambda_2 e^{-\alpha x} - \lambda_2, & x < 0 \\ -\lambda_1 e^{-\alpha x} + \lambda_1, & x \geq 0 \end{cases}$

$\alpha = 0.2, \lambda_1 = 6.5$ **and** $\lambda_2 = 14.5$

The weights are inferred from choices. They are not probabilities, since they don't obey probability laws. The properties of the weighting function (Figure3), say π, are:

- π is an increasing function of stated probabilities
- $\pi(0) = 0$ and $\pi(1) = 1$
- Overweighting for small likelihood events and underweighting for the other events
- Subcertainty

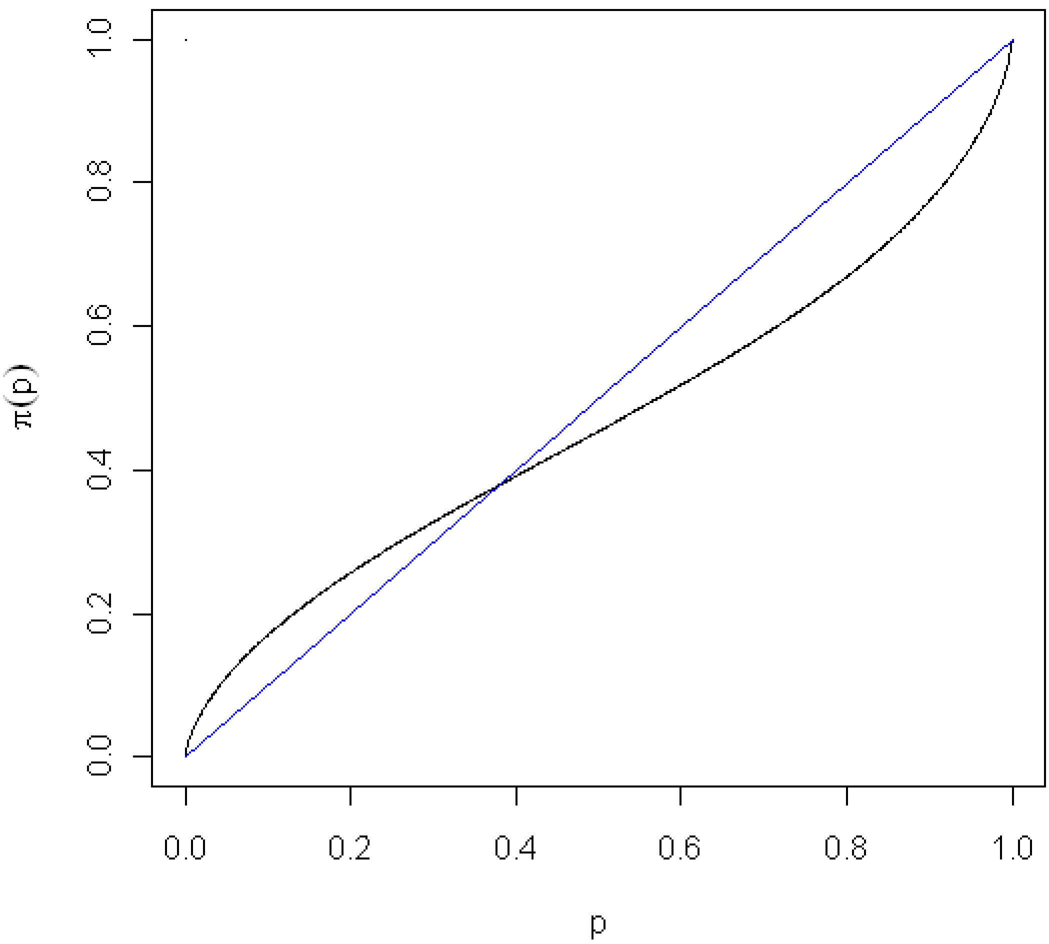

Figure3: Tversky Kahneman (1992) probability weighting function.

$\pi(p) = \frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{1/\gamma}}$. **$\gamma = 0.69$ for both gains and losses**

The prospect theory also accounts for many other situations like the shift of reference point, observed risk attitudes and cases in which probabilities are not provided (in these cases weights are linked to specific events).

Later, the same authors proposed an extension of the prospect theory called the Cumulative Prospect Theory. The main features of the new version are:

- Weights are linked to cumulative probabilities. So, it is applicable to uncertain as well as to risky prospects with continuous distribution. It is coherent with stochastic dominance.
- Different decision weights for gains and losses

We have just seen that not always agents behave rationally and that it is well explained by Cumulative Prospect Theory. Next, we'll see the implications of rational and irrational behaviors for price formation. For that we borrow the model in "Noise trader risk in financial market" by J. Bradford Delong et al..

2. Price formation with noise trader risk

Let consider a market with two assets and two groups of agents, arbitrageurs and noise traders, with the following characteristics:

- Every agent has a portfolio containing exclusively a safe and an unsafe assets
- The unsafe asset supply is perfectly elastic, that is supply is variable when price is always fixed, with price one
- Traders maximize their expected utility one period ahead
- Arbitrageurs are expected utility maximizers with correct beliefs on prices
- Noise traders are utility maximizers misperceiving the price of the risky asset
- Consistency between preference and trading. This is an important assumption for price formation. Because, as pointed out by Kahneman and Tversky (1979), there is a possible inconsistency between choice and bid. That is, we can prefer something but act differently.

Let call

p_t the price of the risky asset at time t, [4]

ϵ_{t+1} noise trader misperception of time t+1 risky asset price at time t

n the proportion of noise traders

α^{nt} the share of the unsafe asset hold by the noise traders

ω^{nt} noise traders perceived wealth at time $t+1$

d_{t+1} the real dividend per share at time $t+1$

r the real risk free rate

By definition, noise traders are investors mistaking noise for information Kyle (1985) and Black (1986); and arbitrageurs are investors with correct beliefs.

[4] On occasion, just for rhetoric, we will replace

- time t by today
- time t+1 by tomorrow or next period
- and time t-1 by yesterday or last period

Traders' tomorrow perceived wealth is

$$\omega = \alpha(p_{t+1} + \epsilon_{t+1}) + \alpha d_{t+1} + (1 - \alpha p_t)(1 + r)\ ;$$

note that $\epsilon_t = 0$ for arbitrageurs.

Let assume that all traders have the same constant absolute risk aversion with utility function defined as

$$\mathbb{U}(\omega) = -e^{-2\gamma\,\omega};$$

where 2γ is absolute risk aversion coefficient.

We stress that the latter assumption is not made only for seek of simplicity. As it has been demonstrated by Haim Levy, Enrico Degiorgi and Thorsten Hens (2003), this functional form does guarantee experimental findings of Kahneman and Tversky (1992) and the existence of equilibra.

We have

$$\mathbb{E}[\mathbb{U}(\omega)]$$

$$= \mathbb{E}[-e^{-2\gamma\,\omega}]$$

$$= \mathbb{E}\left[-e^{-2\gamma\,\alpha p_{t+1} - 2\gamma\,(\alpha\epsilon_{t+1} + \alpha d_{t+1} + (1-\alpha p_t)(1+r))}\right]$$

In our model, maximizing the expected utility one period ahead means that today traders evaluate tomorrow utility when they are already long the unsafe asset, they have information on tomorrow dividend and misperception (meaning that they are not uncertain at that time).

Therefore

$$\mathbb{E}_{\mathrm{t}}[\mathbb{U}(\omega)]$$

$$= \mathbb{E}_{\mathrm{t}}\left[-e^{-2\gamma\,\alpha p_{t+1} - 2\gamma\,(\alpha\epsilon_{t+1} + \alpha d_{t+1} + (1-\alpha p_t)(1+r))}\right]$$

$$= e^{-2\gamma\,(\alpha\epsilon_{t+1}+\alpha d_{t+1}+(1-\alpha p_t)(1+r))}\mathbb{E}_{\mathrm{t}}[-e^{-2\gamma\,\alpha p_{t+1}}]$$

$$= -e^{-2\gamma\,(\alpha\epsilon_{t+1}+\alpha d_{t+1}+(1-\alpha p_t)(1+r))}e^{-2\gamma\,\alpha\mu_t+\frac{\sigma_t^2(2\gamma\,\alpha)^2}{2}}$$

where $p_{t+1}|t$, have mean μ_t and variance σ_t^2; d_{t+1} and ϵ_{t+1} are known at time t.
Agents maximize their expected utility, $\max_\alpha \mathbb{E}_{\mathrm{t}}[\mathbb{U}(\omega)]$, at the following number of shares

$$\alpha = \frac{\epsilon_{t+1} + \mu_t + d_{t+1} - p_t(1+r)}{2\gamma\,\sigma_t^2}$$ [5]

Once determined the number of shares bought by both group of traders, now we see at which price they traded. For one unit of risky asset in the market, at equilibrium, we have:

$$n\alpha^{nt} + (1-n)\alpha^{a} = 1$$

Inserting arbitrageurs and noise traders' shares into the preceding equation gives

$$n\,\frac{\epsilon_{t+1}+\mu_t+d_{t+1}-p_t(1+r)}{2\gamma\,\sigma_t^2} + (1-n)\,\frac{\mu_t+d_{t+1}-p_t(1+r)}{2\gamma\,\sigma_t^2} = 1.$$

Then, it results that the equilibrium price is

$$p_t = \frac{\mu_t+d_{t+1}}{1+r} + n\,\frac{\epsilon_{t+1}}{1+r} - \frac{2\gamma\,\sigma_t^2}{1+r}.$$ [6]

The above equation simply says that agents trade at a price which is the sum of:

- Tomorrow dividend present value, $\frac{d_{t+1}}{1+r}$
- Tomorrow price present value adjusted by its risk premium, $\frac{\mu_t}{1+r} - \frac{2\gamma\,\sigma_t^2}{1+r}$
- The present value of today noise trader misperception, $n\,\frac{\epsilon_{t+1}}{1+r}$

[5] Let note that σ_t^2 must be non zero. Otherwise there is not equilibrium. Proof in Appendices.
[6] Derivation in Appendices.

The conditional expectation and volatility can be written with unconditional ones as follows

$$\mu_t = \mu + z_{t+1} \text{ and } \sigma_t^2 = \sigma^2 + w_{t+1}.$$

Which means that

$$p_t = \frac{\mu + z_{t+1} + d_{t+1}}{1+r} + n\frac{\epsilon_{t+1}}{1+r} - \frac{2\gamma(\sigma^2 + w_{t+1})}{1+r};$$

where $\mathbb{E}[p_t] = \mu$ and $Var(p_t) = \sigma^2$.

z_{t+1} and w_{t+1} denote respectively the uncertainty of expected price and the uncertainty of volatility.

Moreover, we note that z_{t+1} and w_{t+1} are two random variables with respective means,

$$\mathbb{E}[p_{t+1}] - \mathbb{E}[p_t] \text{ and } Var(p_{t+1}) - Var(p_t).$$

Let see how this affects the price by computing the first and second moments.

For old generation of traders (to whom tomorrow dividend and misperception are uncertain), the dividend expected value is $\overline{d_{t+1}}$ with volatility $\sigma^2_{d_{t+1}}$ and noise traders' misperception expected value is $\overline{\epsilon_{t+1}}$ with volatility $\sigma^2_{\epsilon_{t+1}}$.

Let call

$\overline{z_{t+1}}$ and $\overline{w_{t+1}}$ respectively expected values of z_{t+1} and w_{t+1}

$\sigma^2_{z_{t+1}}$ and $\sigma^2_{w_{t+1}}$ respectively volatilities of z_{t+1} and w_{t+1}

Considering z_{t+1}, w_{t+1}, d_{t+1}, ϵ_{t+1} independent we have

$$\mathbb{E}[p_t] = \frac{\mathbb{E}[\mu_t + d_{t+1}]}{1+r} + n\frac{\mathbb{E}[\epsilon_{t+1}]}{1+r} - \frac{2\gamma\,\mathbb{E}[\sigma_t^2]}{1+r};$$

which is equivalent to

$$\mu = \frac{\overline{d_{t+1}}}{r} + \frac{\overline{z_{t+1}}}{r} + n\frac{\overline{\epsilon_{t+1}}}{r} - \frac{2\gamma\overline{w_{t+1}}}{r} - \frac{2\gamma\,\sigma^2}{r}.$$

Concerning the volatility, we have

$$Var(p_t) = \frac{1}{(1+r)^2}\left(Var(\mu_t + d_{t+1}) + n^2 Var(\epsilon_{t+1}) + (2\gamma\,)^2 Var(\sigma_t^2)\right);$$

hence

$$\sigma^2 = \frac{\sigma^2_{d_{t+1}} + \sigma^2_{z_{t+1}} + n^2\sigma^2_{\epsilon_{t+1}} + (2\gamma\,)^2\sigma^2_{w_{t+1}}}{(1+r)^2}.$$

The latter equation departs from the noise traders approach by De Long et al (1987) in which the assumption of prices process stationarity implies that changes in noise traders risk is compensated by change in fundamentals.

Finally we get the price function

$$p_t = \frac{d_{t+1}}{1+r} + \frac{\overline{d_{t+1}}}{r(1+r)} - \frac{2\gamma\sigma^2_{d_{t+1}}}{r(1+r)^2} + \frac{n\epsilon_{t+1}}{1+r} + \frac{n\overline{\epsilon_{t+1}}}{r(1+r)} - \frac{2\gamma n^2\sigma^2_{\epsilon_{t+1}}}{r(1+r)^2} + \frac{z_{t+1}}{1+r} + \frac{\overline{z_{t+1}}}{r(1+r)} - \frac{2\gamma\sigma^2_{z_{t+1}}}{r(1+r)^2} - \frac{2\gamma w_{t+1}}{1+r} - \frac{2\gamma\overline{w_{t+1}}}{r(1+r)} - \frac{2\gamma^3\sigma^2_{w_{t+1}}}{r(1+r)^2}.$$ [7]

This price functional form is separable in four main components: a dividend price, a misperception price, the price due to expected price uncertainty, and the price due to uncertainty of price volatility.

- The dividend price, $\frac{d_{t+1}}{1+r} + \frac{\overline{d_{t+1}}}{r(1+r)} - \frac{2\gamma\sigma^2_{d_{t+1}}}{r(1+r)^2}$. Which is the sum of discounted next period dividend and its discounted continuation value

$$\frac{d_{t+1}}{1+r} + \frac{\overline{d_{t+1}}}{r(1+r)}$$

and adjusted by the risk premium

[7] Note that discounting at risk free rate is coherent since it is applied to the certainty equivalent. For instance, $\overline{d_{t+1}} - 2\gamma\sigma^2_{d_{t+1}}$ for uncertain dividend. So the valuation is risk-neutral (therefore, no-arbitrage).

$$-\frac{2\gamma\sigma_{d_{t+1}}^2}{r(1+r)^2}.$$

This price is paid by investors if there were not noise traders in the market and the dividend payout is stable through time.

- The misperception price, $\frac{n\epsilon_{t+1}}{1+r}+\frac{n\overline{\epsilon_{t+1}}}{r(1+r)}-\frac{2\gamma n^2\sigma_{\epsilon_{t+1}}^2}{r(1+r)^2}$.

  Which is the sum of discounted next period misperception and its discounted noise continuation value

$$\frac{n\epsilon_{t+1}}{1+r}+\frac{n\overline{\epsilon_{t+1}}}{r(1+r)}$$

and the misperception risk adjustment (misperception risk premium)

$$-\frac{2\gamma n^2\sigma_{\epsilon_{t+1}}^2}{r(1+r)^2}.$$

It says that the fundamental price can be distorted by investors sentiments. It's the additional price asked by investors in case there are noise traders misperceiving the price through time. It may also explain the empirical findings of Mehra Prescot (1986) that market prices diverge from fundamental values.[8] This component is central in our model. Therefore, let analyze more in details its implications.

The fundamental price is affected when there are enough noise traders, i.e., n must not be too small. This may be an explanation of the fact that prices fluctuate even when there are no information, Harris and Gurel (1986).

The term $\frac{n\overline{\epsilon_{t+1}}}{r(1+r)}$ represents the price pressure, the value due to average noise misperception. It can also be interpreted as the continuation value of noise.

Because of the fact that noise traders can change their mind, investors ask a premium for the noise risk they are bearing regardless the uncertainty of fundamentals. This is also called "create space effect", $\frac{2\gamma n^2\sigma_{\epsilon_{t+1}}^2}{r(1+r)^2}$.

[8] See also Roll (1984), Campbell and Kyle (1987).

- The price due to the uncertainty (volatility) of expected price, $\frac{z_{t+1}}{1+r} + \frac{\overline{z_{t+1}}}{r(1+r)} - \frac{2\gamma\sigma^2_{z_{t+1}}}{r(1+r)^2}$.

  Although the price is affected by noise and fundamentals, the resulting price can be stationary if the effects from these two sources of risk compensate themselves. Therefore this component is also very important as it captures the fact that the combined action (from fundamentals and from noisy information) is not expected to continue affecting the price like they are doing.

  The term z_{t+1} captures the future average behavior of fundamentals and next generations of traders. Thus, it's essentially too difficult to estimate based on past information. In practice, the success in timing prices in the spot market depends essentially on the ability in guessing this component.
- The price due to the uncertainty (volatility) of volatility of price, $-\frac{2\gamma w_{t+1}}{1+r} - \frac{2\gamma\overline{w_{t+1}}}{r(1+r)} - \frac{2\gamma^3\sigma^2_{w_{t+1}}}{r(1+r)^2}$

  It represents the premium due to uncertainty in prices' volatility. By its nature, it is a key driver of performance in derivative markets.[9]

Relation to asset price modeling

We saw in the first part of this paper that, in classical finance, returns are usually modeled on one risk factor (recall the diffusion process from which the famous Black-Scholes formula is derived). Moreover, it is often assumed that the risk factor has some stability properties through time hence allowing the application of statistical inference. Unfortunately, it is also well known that most of models fail in practice. Very often many estimation techniques assume that the future will be the same like the past, even parametric methods (like Montecarlo Simulation) still rely to some extent on the past to derive their parameters. Although other attempts have been made to improve the modeling of the behavior of asset prices (let mention mixed process like jump diffusion), we argue that they remain simply descriptive approaches. Their predictive power is still arguable mostly due to uncertainty of misperception.

A behavioral approach like the noise trader risk model we analyzed above is very interesting because it segregates prices into two background factors. The advantage is huge

[9] We know from the Black-Scholes-Merton's fornula that derivatives instruments depend essentially on volatility.

as it helps to explain many classical theory inconsistencies found in practice and can help to better control the behavior of asset prices (at least from a risk management viewpoint).

3. Measurement of sentiment

Estimating investor sentiment is as difficult as estimating the fundamental value. Nonetheless, several approaches are proposed in literature and some of which are used by options traders for contrarian investment. Existing methods can be discriminated in market-based approaches and direct data surveys. Without going through the validity of the methods, in the following we sketch some of them.

3.1. Market-based proxies

i. Closed-end funds discount

This was first observed by Charles Lee, Andrei Shleifer, and Richard Thaler (1991) in “Investor Sentiment and the Closed-End Fund Puzzle". Realizing that closed-end funds[10] trade at a premium when they are first introduced and at a discount latter, the claim that this can only be explained by investors misperceptions. Because, according to traditional theories the net asset value (NAV) must equal the market price of the fund.

ii. Index based on new equity issue inconsistency with the law of one price, Welch, Ivo and Lily Qiu (2004); and IPO first-day returns or IPO volume.

iii. Baker and Wurgler Sentiment index (2007)

It is an average of six measures: trading volume based on NYSE turnover, the dividend premium, the closed-end fund discount, first-day returns on IPOs, and the equity share in new issues (equity issues over equity plus long term debt issues). They argue that since single measures are not reliable, they can be averaged out to get a better proxy.

3.2. Survey-based proxies

i. University of Michigan consumer confidence index. The purpose is to estimate consumers’ current confidence in the economy.

[10] Closed-end funds are investment companies that issue a fixed number of shares which are then publicly traded.

ii. Yale school of management's stock market confidence index. This method, designed by a group led by Robert Shiller, attempts to gauge investors sentiment. Questionnaires regard stock market outlook.

Chapter 5.

Behavioral finance; market and credit risk measures

1. Behavioral finance and Value at Risk

So far, we have studied price and return behaviors in a market with noise traders. Now let check what does that involve from a risk management viewpoint. To that end, we first see how Value at Risk and the probability of default (in option-theoretic models) are affected.

Let assume that all risk factors determining the return are bounded, we mean, with finite means and variances. Without this assumption, as we've seen in the thought experiment, the traditional risk measurement process is useless. Moreover we assume that the return has standard distribution ϕ, α is the confidence level, $\hat{\mu}_{t+1}$ is the tomorrow estimated mean and $\hat{\sigma}_{t+1}$ is the tomorrow estimated standard deviation from a sample with length m.

The parametric Value at Risk is

$$VaR_{t+1} = -\mu_{t+1} + \sigma_{t+1}\phi^{-1}(1-\alpha).$$

Now, we compute the mean and standard deviation of the return. Yet for simplicity, we assume that we already know the initial price, p_t, when computing the VaR. Thus, the only risk factors are those of the final price p_{t+1}. Therefore, the relative return between two periods is

$$r_t = \frac{p_{t-p_{t-1}}}{p_{t-1}}.$$

Thus

$$r_t =$$

$$\frac{r(1+r)\Delta_{d_{t+1}}+(1+r)\Delta_{\overline{d_{t+1}}}-2\Delta_{\gamma_{t+1}\sigma^2_{d_{t+1}}}+r(1+r)\Delta_{n_{t+1}\epsilon_{t+1}}+(1+r)\Delta_{n_{t+1}\overline{\epsilon_{t+1}}}-2\Delta_{\gamma_{t+1}{n_{t+1}}^2\sigma^2_{\epsilon_{t+1}}}}{D_{t-1}}+$$

$$\frac{r(1+r)\Delta_{z_{t+1}}+(1+r)\Delta_{\overline{z_{t+1}}}-2\Delta_{\gamma_{t+1}\sigma^2_{z_{t+1}}}-2r(1+r)\Delta_{\gamma_{t+1}w_{t+1}}-2(1+r)\Delta_{\gamma_{t+1}\overline{w_{t+1}}}-8\Delta_{\gamma_{t+1}{}^3\sigma^2_{w_{t+1}}}}{D_{t-1}}$$

[11]

So

$$\mu_{t+1} = \mathbb{E}[r_{t+1}]$$

$$= \frac{(1+r)\Delta_{\overline{d_{t+2}}} - 2\Delta_{\gamma_{t+2}\sigma^2_{d_{t+2}}} + (1+r)\Delta_{n_{t+2}\overline{\epsilon_{t+2}}} - 2\Delta_{\gamma_{t+2}n_{t+2}{}^2\sigma^2_{\epsilon_{t+2}}}}{D_t}$$
$$+ \frac{(1+r)\Delta_{\overline{z_{t+2}}} - 2\Delta_{\gamma_{t+2}\sigma^2_{z_{t+2}}} - 2(1+r)\Delta_{\gamma_{t+2}\overline{w_{t+2}}} - 8\Delta_{\gamma_{t+2}{}^3\sigma^2_{w_{t+2}}}}{D_t}$$
$$+ \frac{r(1+r)\left(\overline{d_{t+2}} + n_{t+2}\overline{\epsilon_{t+2}} + \overline{z_{t+2}} - 2\gamma_{t+2}\overline{w_{t+2}}\right)}{D_t}$$
$$- \frac{r(1+r)(d_{t+1} + n_{t+1}\epsilon_{t+1} + z_{t+1} - 2\gamma_{t+1}w_{t+1})}{D_t}$$

and

$$\sigma^2{}_{t+1} = \frac{\left(r(1+r)\right)^2\left(\sigma^2_{d_{t+2}} + n_{t+2}{}^2\sigma^2_{\epsilon_{t+2}} + \sigma^2_{z_{t+2}} + 4\gamma_{t+2}{}^2\sigma^2_{w_{t+2}}\right)}{D_t{}^2}.$$

The expected return just computed can be divided up in four main components:

- A fundamental component

$$\frac{(1+r)\Delta_{\overline{d_{t+2}}}+r(1+r)\overline{d_{t+2}}-r(1+r)d_{t+1}-2\Delta_{\gamma_{t+2}\sigma^2_{d_{t+2}}}}{D_t},$$

  which is simply the fundamentals' contribution to the return.

- A component due to noise traders misperception

$$\frac{(1+r)\Delta_{n_{t+2}\overline{\epsilon_{t+2}}}+r(1+r)n_{t+2}\overline{\epsilon_{t+2}}-r(1+r)n_{t+1}\epsilon_{t+1}-2\Delta_{\gamma_{t+2}n_{t+2}{}^2\sigma^2_{\epsilon_{t+2}}}}{D_t},$$

[11] Derivation in Appendices.

Which is the misperception's contribution to the return

- A component due to uncertainty of expected price

$$\frac{(1+r)\Delta_{\overline{z_{t+2}}}+r(1+r)\overline{z_{t+2}}-r(1+r)z_{t+1}-2\Delta_{\gamma_{t+2}\sigma^2_{z_{t+2}}}}{D_t},$$

That is the expected price uncertainty's contribution to the return

- A component due to uncertainty of volatility

$$\frac{-2(1+r)\Delta_{\gamma_{t+2}\overline{w_{t+2}}}-8\Delta_{\gamma_{t+2}{}^3\sigma^2_{w_{t+2}}}-r(1+r)2\gamma_{t+2}\overline{w_{t+2}}+r(1+r)2\gamma_{t+1}w_{t+1}}{D_t},$$

that is the volatility uncertainty's contribution to the return

Contrary to the efficient market hypothesis, it is obvious that the risk measure would depend both on fundamentals and investors' misperception. This hints us to analyze more in detail the goodness of VaR under a market with noise traders.[12]

Under market with no misperception by some investors, as it appears from the expected return and the volatility, Value at Risk would be a good measure since it increases with the riskiness in fundamentals and decreases with good expectations in fundamentals (all else being equal, we mean that the combined effect of uncertainty of expected price and volatility is negligible). However the estimated Value at Risk, as usually computed based on available information, may be still biased due the difficulty in estimating the uncertainty of expected price and price volatility.

Under markets in which prices satisfy our pricing function, Value at Risk decreases proportionally to the combined effect of price pressure and create space θ_{t+1} (all else being equal); where

$$\theta_{t+1}(R_{t+1},I)=\frac{(1+r)\Delta_{n_{t+2}\overline{\epsilon_{t+2}}}+r(1+r)n_{t+2}\overline{\epsilon_{t+2}}-r(1+r)n_{t+1}\epsilon_{t+1}-2\Delta_{\gamma_{t+2}n_{t+2}{}^2\sigma^2_{\epsilon_{t+2}}}}{D_t}.$$ [13]

[12] This section can be considered as follow-up to our thought experiment.

[13] $\hat{\theta}_{t+1}(R_{t+1},I)$, means that the systematic misperception depends on the exposure (R_{t+1}), current and future information (I). Normally it includes also volatility uncertainty's contribution to the return, since this component behaves in the risk measure like the expected return from misperception . But we assumed it is negligible just to highlight the effect of systematic noise trading.

We quickly realize that this combined effect is nothing but the misperception's contribution to the return.

The result is striking because it means that:

- First, risk perception would be reduced when a bubble is booming. Since θ_{t+1} would appear like an opportunity (positive cash flow) while being a hidden threat
- And second, the risk measure misinterpreting θ_{t+1} would be an incentive to bubble

We believe that this component might be significative for some assets (like glamour stocks). Its main properties are:

- Homogeneousness. It is proportional to the exposure.
- Translation invariant. It is not affected by a risk free asset
- It is sub-additive. Because any position can be hedged out in complete market

These features are crucial when it comes to defining and constructing coherent risk measures under behavioral framework.

Also, we can say from the mean and variance equations that the risk measure increases with risk aversion and proportion of noise traders. This suggests to redefine our risk measure in order to include the badly captured effect of noise trading. Hence, we propose the following adjusted market risk measure

$$VaR_{t+1} = -\mu_{t+1} + \sigma_{t+1}\phi^{-1}(1-\alpha) + \theta_{t+1}(R_{t+1}, I). \qquad (0)$$

The latter measure can be easily translated to non parametric methods. In this case we' ld have

$$VaR_t = -r_{t^*} + \theta_{t^*}$$

where r_{t^*} satisfies requirements in non parametric Value at Risk, and

$$\theta_{t^*} = \frac{r(1+r)n_{t^*}\epsilon_{t^*} + (1+r)\Delta_{n_{t^*}\overline{\epsilon_{t^*}}} - 2\Delta_{\gamma_{t^*}n_{t^*}{}^2\sigma^2_{\epsilon_{t^*}}}}{D_{t^*}}$$

In practice, the measure used in parametric methods is the estimated Value at Risk. Thus, provided our pricing function, let see its implication to our traditional risk measurement process. We assume an estimation based on method of moments.

Method of moment estimation

We know that the first moment estimator is

$$\hat{\mu}_{t+1} = \frac{\sum_{i=1}^{m} r_i}{m},$$

the correct second moment estimator is

$$\hat{\sigma}_{t+1} = \sqrt{\frac{\sum_{i=1}^{m}(r_i - \hat{\mu}_{t+1})^2}{m-1}}$$

and Value at Risk estimator is

$$\widehat{VaR}_{t+1} = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\phi^{-1}(1-\alpha)$$

where m is the number of observations during the historical period.

It is clear that, ahead of high volatility period these two moment estimators would tend to be undervalued because they incorporate average past information. So, provided that the standard deviation estimator dominates the mean estimator, VaR estimator would tend to be undervalued ahead of high volatility period.

In the first section of this paper, we say that it is usually claimed that VaR is coherent in a Gaussian world (more generally in world where distribution is stable square-integrable). Clearly, since the risk value is given by the estimator, the estimator coherence matters more than the measure coherence. Yet, we may not have diversification benefit even in a Gaussian world if the estimator is not appropriate (inferred or chosen such to be coherent). In the sequel we illustrate some evidences. To that end we use the method of moment estimator and the GARCH(1,1).

Let $X\ and\ Y$ be our standardized exposures returns; λ a real number. $X\ and\ Y$ are Gaussians and not strictly positive, so we are left we sub-additivity, positive homogeneity and translation equivariant.

VaR with method of moments

We have

$$\widehat{VaR}_{t+1}(X) = -\frac{\sum_{i=1}^{m} x_i}{m} + \sqrt{\frac{\sum_{i=1}^{m}(x_i - \hat{\mu}_{t+1})^2}{m-1}}\,\phi^{-1}(1-\alpha)$$

i. Translation equivariant

We have

$$\widehat{VaR}_{t+1}(X+\lambda)$$

$$= -\frac{\sum_{i=1}^{m}(x_i+\lambda)}{m} + \sqrt{\frac{\sum_{i=1}^{m}\left((x_i+\lambda) - \frac{\sum_{i=1}^{m}(x_i+\lambda)}{m}\right)^2}{m-1}}\,\phi^{-1}(1-\alpha)$$

$$= \widehat{VaR}_{t+1}(X) - \lambda$$

ii. Positive homogeneity

When $\lambda > 0$, we have

$$\widehat{VaR}_{t+1}(\lambda X)$$

$$= -\frac{\sum_{i=1}^{m} \lambda x_i}{m} + \sqrt{\frac{\sum_{i=1}^{m}(\lambda x_i)^2}{m-1}}\,\phi^{-1}(1-\alpha)$$

$$= \lambda\widehat{VaR}_{t+1}(X)$$

iii. Sub-additivity

$$\widehat{VaR}_{t+1}(X+Y)$$

$$= -\frac{\sum_{i=1}^{m}(x_i+y_i)}{m} + \sqrt{\frac{\sum_{i=1}^{m}\left((x_i+y_i) - \hat{\mu}_{t+1}(X) - \hat{\mu}_{t+1}(Y)\right)^2}{m-1}}\,\phi^{-1}(1-\alpha)$$

$$= -\frac{\sum_{i=1}^{m}(x_i + y_i)}{m} + \sqrt{\frac{\sum_{i=1}^{m}\left(\left(x_i - \hat{\mu}_{t+1}(X)\right) + \left(y_i - \hat{\mu}_{t+1}(Y)\right)\right)^2}{m-1}}\phi^{-1}(1-\alpha)$$

$$= -\frac{\sum_{i=1}^{m}(x_i + y_i)}{m}$$
$$+ \sqrt{\hat{\sigma}^2{}_{t+1}(X) + \hat{\sigma}^2{}_{t+1}(Y) + \frac{2\sum_{i=1}^{m}\left(x_i - \hat{\mu}_{t+1}(X)\right)\left(y_i - \hat{\mu}_{t+1}(Y)\right)}{m-1}}\phi^{-1}(1$$
$$- \alpha)$$

Thanks to Cauchy-Schwarz inequality we have

$$\left(\sum_{i=1}^{m}\left(x_i - \hat{\mu}_{t+1}(X)\right)\left(y_i - \hat{\mu}_{t+1}(Y)\right)\right)^2$$
$$\leq \sum_{i=1}^{m}\left(x_i - \hat{\mu}_{t+1}(X)\right)^2 \sum_{i=1}^{m}\left(y_i - \hat{\mu}_{t+1}(Y)\right)^2$$

we deduce that

$$\widehat{VaR}_{t+1}(X+Y) \leq \frac{\sum_{i=1}^{m}(x_i + y_i)}{m} + \left(\hat{\sigma}_{t+1}(X) + \hat{\sigma}_{t+1}(Y)\right)\phi^{-1}(1-\alpha)$$

which is simply equivalent to

$$\widehat{VaR}_{t+1}(X+Y) \leq \widehat{VaR}_{t+1}(X) + \widehat{VaR}_{t+1}(Y)$$

Hence, the VaR estimator with method of moments is coherent when the distribution is stable.

VaR with GARCH(1,1)

To show that VaR is not coherent, it is sufficient to show that the variance estimator is not homogeneous or translation equivariant. Again, let consider our standardized exposure X and the positive real λ. Provided our variance estimator

$$\hat{\sigma}^2{}_{t+1}(X) = \omega + \alpha x_t{}^2 + \beta \hat{\sigma}^2{}_t(X),$$

where ω, α and β are defined in the usual manner
We have

$$\hat{\sigma}^2{}_{t+1}(\lambda X) = \omega_\lambda + \alpha_\lambda (\lambda x_t)^2 + \beta_\lambda \hat{\sigma}^2{}_t(\lambda X).$$

Like the optimization, the proof is essentially done numerically.

Numerical evidence

We consider 400 S&P500 returns sample from 04/09/2009 backward. We confront $\widehat{VaR}_{05/09/2009}(X)$ with $\widehat{VaR}_{05/09/2009}(10 * X)$. We get

$$\widehat{VaR}_{05/09/2009}(X) \; = 0.02981085$$

$$\widehat{VaR}_{05/09/2009}(10 * X) \; = 0.2515701.$$

We note that

$$\widehat{VaR}_{05/09/2009}(10 * X) \; < 10 * \widehat{VaR}_{05/09/2009}(X).$$

Hence the risk measure is not positive homogeneous
This inequality says that risk increases less proportionally to the exposure size. Therefore alluding to some concentration benefit. Which is counter-intuitive.
Also we have

$$\widehat{VaR}_{05/09/2009}(X + 0.01) = 0.03021208$$

We note that

$$\widehat{VaR}_{05/09/2009}(X + 0.01) \; \neq \widehat{VaR}_{05/09/2009}(X) - 0.01$$

We just found that, in parametric estimation, the VaR estimator coherence is more important than the VaR coherence. We backed it with a numerical evidence.[14] This can be easily extended to any parametric risk measure.

2. Behavioral finance and probability of default in option-theoretic models

The probability of default is the most important and difficult element to estimate in Merton-type models. In the following, we assume that assumptions in Merton-type are met (continuous time trading, diffusion process, etc…) and under a market with noise traders risk, we sound the effects on probability of default.
In this framework a company defaults at time t when

$$V_t \leq D \quad . \tag{1}$$

We also know that

$$V_t = D + E_t \tag{2}$$

where E_t is the company equity at time t.
From (1) and (2) we have

$$E_t \leq 0.$$

Since

$$E_t = \max\,(V_t - D, 0),$$

default is triggered when

$$E_t = 0. \tag{3}$$

[14] Using Boeing and Apple adjusted returns, we found numerically that the measure is sub-additive. Caveat: these findings are not to be generalized.

Equation (3) says that the company defaults wherever

$$\frac{d_{t+1}}{1+r}+\frac{\overline{d_{t+1}}}{r(1+r)}-\frac{2\gamma\sigma_{d_{t+1}}^2}{r(1+r)^2}+\frac{n\epsilon_{t+1}}{1+r}+\frac{n\overline{\epsilon_{t+1}}}{r(1+r)}-\frac{2\gamma n^2\sigma_{\epsilon_{t+1}}^2}{r(1+r)^2}+\frac{z_{t+1}}{1+r}+\frac{\overline{z_{t+1}}}{r(1+r)}-\frac{2\gamma\sigma_{z_{t+1}}^2}{r(1+r)^2}-\frac{2\gamma w_{t+1}}{1+r}-\frac{2\gamma\overline{w_{t+1}}}{r(1+r)}-\frac{2\gamma^3\sigma_{w_{t+1}}^2}{r(1+r)^2}=0 \tag{4}$$

By definition the price is always positive. From the pricing function and equation (4) we see that even with strong fundamentals, with negligible impact of the third and fourth components of the pricing function, a negative price pressure (when noise traders are over-pessimistic) may cause an unnecessary downgrade of the company. And a positive price pressure (when noise traders are over-optimistic) may unduly upgrade the company. The result is that traditional mark to market models (in which it is believed that the more liquid an asset is the better is), and in particular option-theoretic approach, are not reliable in presence of noise traders.

Some empirical tests on structural models have been conducted. For instance, it results from the test conducted by Young Ho Eom, Jean Elwege and Jing-Zhi Huang (2004) that probabilities are too low in Merton model. It would be premature to relate this fact to asset price behavior. Because, as we’ve seen earlier in VaR estimation, the risk element would depend also on the manner model’s parameters are estimated.

Chapter 6.

Application

1. Estimation technique

We have seen that risk measurement process is a very ambitious task from a modeling and estimation viewpoint. Recently, new approaches are requested and some are emerging from the literature. Wilson Sy (2008) suggests that credit risks must be based on a causal framework. But, as also noted by John Maynard Keynes (1937), the complexity of financial and economic environment makes this approach almost impossible. The distribution and its estimate are in some extent based on weak assumptions. This makes the measurement process fundamentally a judgmental process. (L. J. Savage, 1962). As supported by John C. Hull (2007), since prices are affected by behavioral biases, risk models must be tempered with judgment and stress testing.

According to equation (0), VaR parametric estimation is reduced to the estimation of the mean μ, the volatility σ and the additional parameter θ, obviously with the distribution given a priori or already estimated. The new parameter θ in our proposed market risk measure which is driven by the systematic misbehavior of noise traders could be essentially estimated on judgmental basis. The mean is usually approximated as average of observations in a given historical period or simply imposed by the analyst[15]. Contrary to the mean, there is a colorful literature on volatility estimation. It can be classified in historical volatility, GARCH-type volatility, implied volatility, stochastic volatility and realized volatility models (with related OHLC (open-high-low-close) volatility). As found by Torben G. Andersen et al.(1999) and Stephen Figlewski (2004), there is not a standard volatility estimator (neither an universal loss function). The estimator goodness criterion can be statistic (for instance, the root mean square or the mean absolute error) or economic (proportion of overestimation)[16]. The volatility estimator is ad hoc, we mean that it is chosen by the analyst according to the problem at hand.

[15] See Black, Fischer (1976).
[16] For the predictability of volatility forecast, see Torben G. Andersen et al.(1999) and Stephen Figlewski (2004). See Engle et al. (1993) and West et al. (1993) for economic criterion.

Under some regularity condition (bounded prices' process) we try to estimate the market risk of the S&P 500 index.[17] In the following, we perform two tasks. First, to see the effect of sample size, we compare different daily-VaRs estimations. GARCH(1,1) VaRs with (sample sizes 10000, 400 and 30), Yang and Zhang VaR (with Yang and Zhang volatility estimator) and HL VaR (with a naïve estimator called HL estimator).[18] Secondly, provided that we found that misperception which is an important risk factor is normally ill impounded in the traditional Value at Risk, we see how our VaR with 400 observations from 04/09/2009 backward is affected by the factor θ. Moreover, we use shorter period (one month) to estimate the moments of the distribution in the naïve estimator. This is because longer series tends to average out short term trends, affecting the first moment therefore. We mean that the effect of recent important information (like signals of high volatility periods) is mitigated with longer series.[19]

HL VaR estimator:

Regarding the distribution, the cumulative prospect theory says that it depends mostly on the weighs assigned by investors in a given market condition. Thus, it would normally be skewed towards positive returns in bull markets and negative returns in bear markets. Nevertheless, for simplicity and limited information about the true distribution, we consider that the distribution is Gaussian. Also, provided that we have limited information about returns' process we assumed only extreme scenarios given by a historical period (one month in our case) to approximate the first and second moments. We get,

$$\widehat{VaR}_{t+1} = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\phi^{-1}(1-\alpha) + \hat{\theta}_{t+1}(R_{t+1}, I)$$

with

$$\hat{\mu}_{t+1} = \frac{max\big((r_i)_{i=1,..,30}\big) + min\big((r_i)_{i=1,..,30}\big)}{2}$$

$$\hat{\sigma}_{t+1} = \sqrt{\frac{max\big((r_i)_{i=1,..,30}\big)^2 + min\big((r_i)_{i=1,..,30}\big)^2}{2} - \hat{\mu}_{t+1}{}^2}.$$

17 This choice is intentional. Because S&P 500 index reflects the most the risky asset in our model.

18 See A. Saichev, D. Sornette and V. Filimonov (2009) for a comprehensive study on homogeneous OHLC estimators.

19 Similar realized volatility approaches have been done by Merton (1980), Poterba and Summers (1986).

Although this estimator may be inefficient (It would depend mostly on the sampling frequency. We mean whether with high frequency or low frequency data), we are compensated with some desirable economic features such as:

- High sensitivity to extreme events and their signals (especially for high volatility periods).[20] This is very important since these events are the essence of Value at Risk.
- sensitivity to mild events
- Coherence. Obviously this measure is coherent. The proof is similar to the aforementioned method of moments
- Flexibility. As we'll be seen below, the conservativeness can be controled through an appropriate adjustment of the volatility

2. **Backtest: GARCH(1,1) daily-VaRs (with sample sizes 10000, 400 and 30) and daily-VaR with Yang and Zhang estimator.**

Backtests are from *05/01/1962 to 04/09/2009*. Except for the one with sample size 10000 which ranges from *16/10/1989* to *04/09/2009* due to sample length.

Graphically, there is not a significative difference in GARCH(1,1) when changing sample size. Except that the smaller sample size seems to cover slightly better the left tail (represented by peaks in Figure4 in Appendices). Also, Figure4 and Figure5 do not indicate significative difference between models. Except the naïve estimator appears to be more conservative at the peaks.

Now, we perform a statistical comparison between models based on a metric. Since an implicit economic requirement underpinning the VaR is to be more conservative in higher volatility periods a loss function we can appeal to is the Mean Conditional Absolute Distances[21]. Where it is defined as

$$\mathrm{MCAD} = \mathrm{average}[|(\mathrm{VaR_t} + \mathrm{r_t})_\mathrm{n}|;\ |\mathrm{r_t}| > 1\%]$$

[20] In accordance with Basel2, page200.
[21] The concept of Mean Absolute Distances, although copied from the Mean Absolute Deviation, is different in the sense that it is more economically orientated and it doesn't represent an error.

In words, it is the average of the sequence given by the differences between VaRs and negative returns with amplitudes above a threshold value (1% in this case is arbitrary). The results are illustrated in the next table.

Caveat: the volatility in Yang and Zhang is based on 2-days horizon, assuming a constant volatility during 2-days; and the mean is zero.

Mean Conditional absolute distances					
Scaling parameter[22]	**VaR (10000)**	**VaR(400)**	**VaR(30)**	**VaR (Yang&Zhang)**	**VaR(HL)**
1	*0.01115671*	*0.01471326*	*0.01238488*	*0.008055*	*0.05914015*
0.8	-	-	-	-	*0.04415261*
0.6	-	-	-	-	*0.02939166*
0.5	-	-	-	-	*0.02230908*
0.4	-	-	-	-	*0.01588664*

Table1: Mean Conditional Absolute Distances.

As it appears in the table above, VaR(HL) is the most moderate and VaR(Yang&Zhang) is the less moderate. Scaling the volatility in VaR(HL) estimator with parameter ranging from *1* to about *0.4* reduces the conservativeness without losing other features (like covering the peaks with probability one). These results are consistent with features of the estimators and what have been seen graphically.

3. Sensitivity analysis

To check the fashion in which the risk measure can be affected by the misperception, let estimate the Adjusted Value at Risk of *05/09/2009* with Filter Historical Simulation. The

[22] Scale parameter to render the HL-VaR estimator less conservative.

400 returns from *04/09/2009* backward are filtered through a GACH(1,1) with variance targeting. We obtain a non-adjusted VaR of ***0.0357659***. Then, adding arbitrary values of $\hat{\theta}_{05/09/2009}$ we get the table below.

FHS Adjusted Value at Risk (to systematic noise misperception)					
$\hat{\theta}_{05/09/2009}$	*-0.05*	*-0.02*	*0*	*0.02*	*0.05*
$\widehat{VaR^{\theta}}_{05/09/2009}$	*-0.0257659*	*0.0157659*	*0.0357659*	*0.0557659*	*0.0857659*

Table2: FHS Adjusted VaR, $VaR^{\theta}{}_{05/09/2009}$. $\hat{\theta}_{05/09/2009}$ is arbitrary.

The $\hat{\theta}$ arbitrary values in the table are not far from possible real ones. In chapter4, we saw that fluctuations may be sometimes driven mostly by the misperception factor. A plausible evidence is documented by Harris and Gurel (1986). They found that there is an average abnormal return of *3.5%* for stocks added to S&P 500. A value which is almost the double of risk value usually obtained through traditional models, even when performing an a posteriori measurement after a burst bubble.

The negative risk in the second column is not surprising because it is the typical case of risk free positions, allowing risk reduction therefore. But in our case above it may be the value of an undervalued asset.

Chapter 7.

Conclusion

In this paper we went through the modeling process of market and default risk. We realized that it is not an easy task to get the clear representation of a phenomenon of interest. Even when we have an idea on its behavior, the used estimator matters the most in the final result. We construct a simple estimator, based on judgment, that may be as useful as many classical estimators for market risk control.

Moreover we found that, under behavioral assumptions, risk measures must be revised to account for hidden risk due to systematic misperception. Although notions developed in this paper are ad hoc, they can be extended to other assets. We believe that a profound study of these concepts might improve the understanding of assets' behavior and risk control.

Appendices

Credit risk moment generation function with variable default rate
For variable default rate

$$\mathcal{P}(\mathrm{t}) = \sum_{n=0}^{\infty} \mathbb{P}[n\ defaults] z^n$$

$$= \sum_{n=0}^{\infty} t^n \int_0^{\infty} \mathbb{P}[n\ defaults|X]\, \mathrm{f(x)} dx$$

$$= \int_0^{\infty} \sum_{n=0}^{\infty} t^n\, \mathbb{P}[n\ defaults|X]\, \mathrm{f(x)} dx$$

$$= \int_0^{\infty} \mathrm{e}^{(t-1)x}\, \mathrm{f(x)} dx.$$

Assuming that $X \sim \Gamma(\alpha, \beta)$, i.e. $\mathrm{f(x)} = \frac{\mathrm{e}^{\frac{-x}{\beta}} \mathrm{x}^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}$, we have

$$\mathcal{P}(\mathrm{t}) = \int_0^{\infty} \mathrm{e}^{(t-1)x}\, \mathrm{f(x)} dx$$

$$= \int_0^{\infty} \mathrm{e}^{(t-1)x} \frac{\mathrm{e}^{\frac{-x}{\beta}} \mathrm{x}^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} dx$$

$= \frac{\Gamma(\alpha)}{\beta^{\alpha}\Gamma(\alpha)(1+\beta^{-1}-\mathrm{z})^{\alpha}}$; by generalized binomial theorem

$$= \frac{1}{\Gamma(\alpha)(1+\beta^{-1}-\mathrm{t})^{\alpha}}$$

$$= \left(\frac{1-\mathrm{q}}{1-\mathrm{tq}}\right)^{\alpha}$$

$$= (1-\mathrm{q})^{\alpha} \sum_{n=1}^{\infty} \binom{n+\alpha-1}{n} q^n t^n$$

where

$$q = \frac{\beta}{1+\beta}.$$

Finding share that maximizes expected utility function

We have

$$\mathbb{E}_{\mathrm{t}}\left[-e^{-2\gamma\, \alpha p_{t+1} - 2\gamma\, (\alpha \epsilon_t + \alpha d_t + (1-\alpha p_t)(1+r))}\right]$$

$$= -e^{-2\gamma\, (\alpha \epsilon_t + \alpha d_t + (1-\alpha p_t)(1+r))} e^{-2\gamma\, \alpha \mu_t + \frac{\sigma_t^2 (2\gamma\, \alpha)^2}{2}}$$

Computing the derivative at α

$$\frac{\partial \mathbb{E}_{\mathrm{t}}[\mathbb{U}(\omega)]}{\partial \alpha}$$

$$= 2\gamma\, \big(\epsilon_t + \mu_t + d_t - p_t(1+r)\big) - 4\sigma_t^2 \gamma\,^2 \alpha$$

thus

$$\frac{\partial \mathbb{E}_{\mathrm{t}}[\mathbb{U}(\omega)]}{\partial \alpha} = 0$$

$$2\gamma\, \big(\epsilon_t + \mu_t + d_t - p_t(1+r)\big) - 4\sigma_t^2 \gamma\,^2 \alpha = 0$$

Which yields

$$\alpha = \frac{\epsilon_t + \mu_t + d_t - p_t(1+r)}{2\gamma\ \sigma_t^2}$$

Finding the equilibrium price

$$n\alpha^{nt} + (1-n)\alpha^{a} = 1$$

Is equivalent to

$$n\frac{\epsilon_t + \mu_t + d_t - p_t(1+r)}{2\gamma\ \sigma_t^2} + (1-n)\frac{\mu_t + d_t - p_t(1+r)}{2\gamma\ \sigma_t^2} = 1$$

the same

$$n\frac{(\epsilon_t+\mu_t+d_t)}{2\gamma\ \sigma_t^2} + (1-n)\frac{(\mu_t+d_t)}{2\gamma\ \sigma_t^2} - p_t(1+r)\left(\frac{n}{2\gamma\ \sigma_t^2} + \frac{1-n}{2\gamma\ \sigma_t^2}\right)=1$$

thus

$$n\frac{\epsilon_t}{2\gamma\ \sigma_t^2} + \frac{\mu_t + d_t}{2\gamma\ \sigma_t^2} - p_t\left(\frac{1+r}{2\gamma\ \sigma_t^2}\right) = 1$$

Hence

$$p_t = \frac{\mu_t + d_t}{1+r} + n\frac{\epsilon_t}{1+r} - \frac{2\gamma\ \sigma_t^2}{1+r}$$

Breaking down the equilibrium price in fundamental value and noise risk

$$p_t = \frac{\mu_t + d_{t+1}}{1+r} + n\frac{\epsilon_{t+1}}{1+r} - \frac{2\gamma\sigma_t^2}{1+r}$$

$$p_t = \frac{\mu_t + d_{t+1} - 2\gamma\sigma_t^2}{1+r} + \frac{n\epsilon_{t+1}}{1+r}$$

$$p_t = \frac{\frac{\overline{d_{t+1}}}{r} + n\frac{\overline{\epsilon_{t+1}}}{r} - \frac{2\gamma\sigma_t^2}{r} - 2\gamma\sigma_t^2}{1+r} + \frac{n\epsilon_{t+1} + d_{t+1}}{1+r}$$

$$p_t = \frac{\frac{\overline{d_{t+1}}}{r} + n\frac{\overline{\epsilon_{t+1}}}{r} - \frac{2\gamma\sigma_t^2\,(1+r)}{r}}{1+r} + \frac{n\epsilon_{t+1} + d_{t+1}}{1+r}$$

$$p_t = \frac{\overline{d_{t+1}}}{r(1+r)} + \frac{n\overline{\epsilon_{t+1}}}{r(1+r)} - \frac{2\gamma\sigma_t^2}{r} + \frac{n\epsilon_{t+1} + d_{t+1}}{1+r}$$

$$p_t = \frac{d_{t+1}}{r(1+r)} + \frac{n\overline{\epsilon_{t+1}}}{r(1+r\,)} - \frac{2\gamma\left(n^2\sigma_{\epsilon_{t+1}}^2 + \sigma_{d_{t+1}}^2\right)}{r(1+r)^2} + \frac{n\epsilon_{t+1} + d_{t+1}}{1+r}$$

$$p_t = \frac{d_{t+1}}{1+r} + \frac{\overline{d_{t+1}}}{r(1+r)} + \frac{n(\epsilon_{t+1} - \overline{\epsilon_{t+1}}\,)}{1+r} + \frac{n\overline{\epsilon_{t+1}}}{r(1+r)} + \frac{n\overline{\epsilon_{t+1}}}{1+r} - \frac{2\gamma\left(n^2\sigma_{\epsilon_{t+1}}^2 + \sigma_{d_{t+1}}^2\right)}{r(1+r)^2}$$

$$p_t = \frac{d_{t+1}}{1+r} + \frac{\overline{d_{t+1}}}{r(1+r)} - \frac{2\gamma\sigma_{d_{t+1}}^2}{r(1+r)^2} + \frac{n(\epsilon_{t+1} - \overline{\epsilon_{t+1}}\,)}{1+r} + \frac{n\overline{\epsilon_{t+1}}}{r} - \frac{2\gamma n^2\sigma_{\epsilon_{t+1}}^2}{r(1+r)^2}$$

Return formula

Let say

$\Delta_{d_t} = d_t - d_{t-1}$ $\qquad \Delta_{\overline{d_t}} = \overline{d_t} - \overline{d_{t-1}}$

$\Delta_{\gamma_t\sigma_{d_t}^2} = \gamma_t\sigma_{d_t}^2 - \gamma_{t-1}\sigma_{d_{t-1}}^2$ $\qquad \Delta_{n_t\epsilon_t} = n_t\epsilon_t - n_{t-1}\epsilon_{t-1}$

$\Delta_{n_t\overline{\epsilon_t}} = n_t\overline{\epsilon_t} - n_{t-1}\overline{\epsilon_{t-1}}$ $\qquad \Delta_{\gamma_t {n_t}^2\sigma_{\epsilon_t}^2} = \gamma_t {n_t}^2\sigma_{\epsilon_t}^2 - \gamma_{t-1}{n_{t-1}}^2\sigma_{\epsilon_{t-1}}^2.$

$\Delta_{z_t} = z_t - z_{t-1}$ $\qquad \Delta_{\overline{z_t}} = \overline{z_t} - \overline{z_{t-1}}$

$\Delta_{\gamma_t\sigma_{d_t}^2} = \gamma_t\sigma_{d_t}^2 - \gamma_{t-1}\sigma_{d_{t-1}}^2$ $\qquad \Delta_{\gamma_t w_t} = \gamma_t w_t - \gamma_{t-1}w_{t-1}$

$\Delta_{\gamma_t\overline{w_t}} = \gamma_t\overline{w_t} - \gamma_{t-1}\overline{w_{t-1}}$ $\qquad \Delta_{{\gamma_t}^3\sigma_{d_t}^2} = {\gamma_t}^3\sigma_{d_t}^2 - {\gamma_{t-1}}^3\sigma_{d_{t-1}}^2$

We have

$$D_t = r(1+r)d_{t+1} + (1+r)\overline{d_{t+1}} - 2\gamma_{t+1}\sigma^2_{d_{t+1}} + r(1+r)n_{t+1}\epsilon_{t+1} + (1+r)n_{t+1}\overline{\epsilon_{t+1}} - 2\gamma_{t+1}{n_{t+1}}^2\sigma^2_{\epsilon_{t+1}} + r(1+r)z_{t+1} + (1+r)\overline{z_{t+1}} - 2\gamma_{t+1}\sigma^2_{z_{t+1}} - 2r(1+r)\gamma_{t+1}w_{t+1} - 2(1+r)\gamma_{t+1}\overline{w_{t+1}} - 8{\gamma_{t+1}}^3\sigma^2_{w_{t+1}}.$$

therefore

$$r_t = \frac{r(1+r)\Delta_{d_{t+1}} + (1+r)\Delta_{\overline{d_{t+1}}} - 2\Delta_{\gamma_{t+1}\sigma^2_{d_{t+1}}} + r(1+r)\Delta_{n_{t+1}\epsilon_{t+1}} + (1+r)\Delta_{n_{t+1}\overline{\epsilon_{t+1}}} - 2\Delta_{\gamma_{t+1}{n_{t+1}}^2\sigma^2_{\epsilon_{t+1}}}}{D_{t-1}} + \frac{r(1+r)\Delta_{z_{t+1}} + (1+r)\Delta_{\overline{z_{t+1}}} - 2\Delta_{\gamma_{t+1}\sigma^2_{z_{t+1}}} - 2r(1+r)\Delta_{\gamma_{t+1}w_{t+1}} - 2(1+r)\Delta_{\gamma_{t+1}\overline{w_{t+1}}} - 8\Delta_{{\gamma_{t+1}}^3\sigma^2_{w_{t+1}}}}{D_{t-1}}$$

Figures

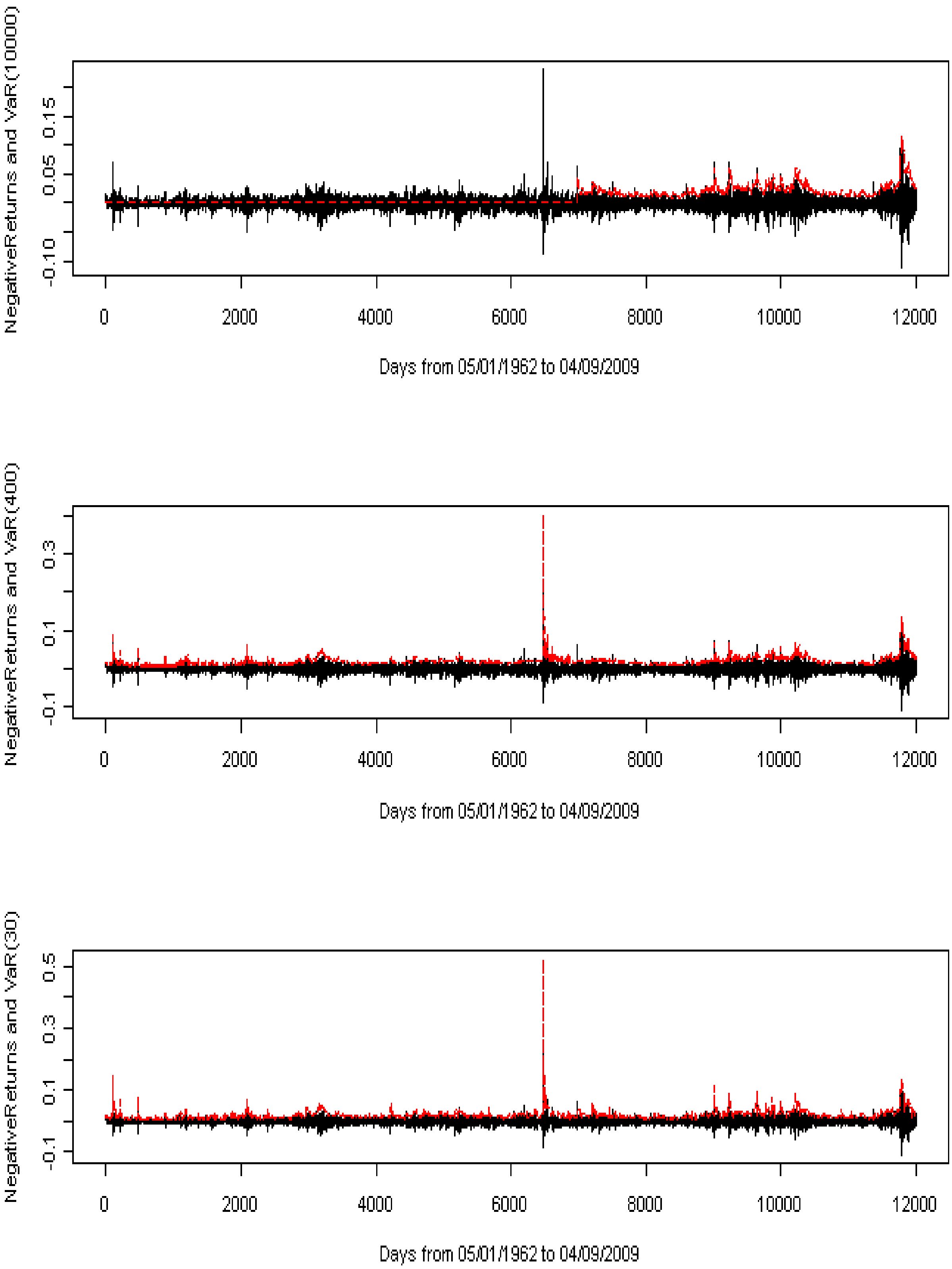

Figure4: Backtest of GARCH(1,1) VaRs (with sample sizes 10000, 400 and 30).

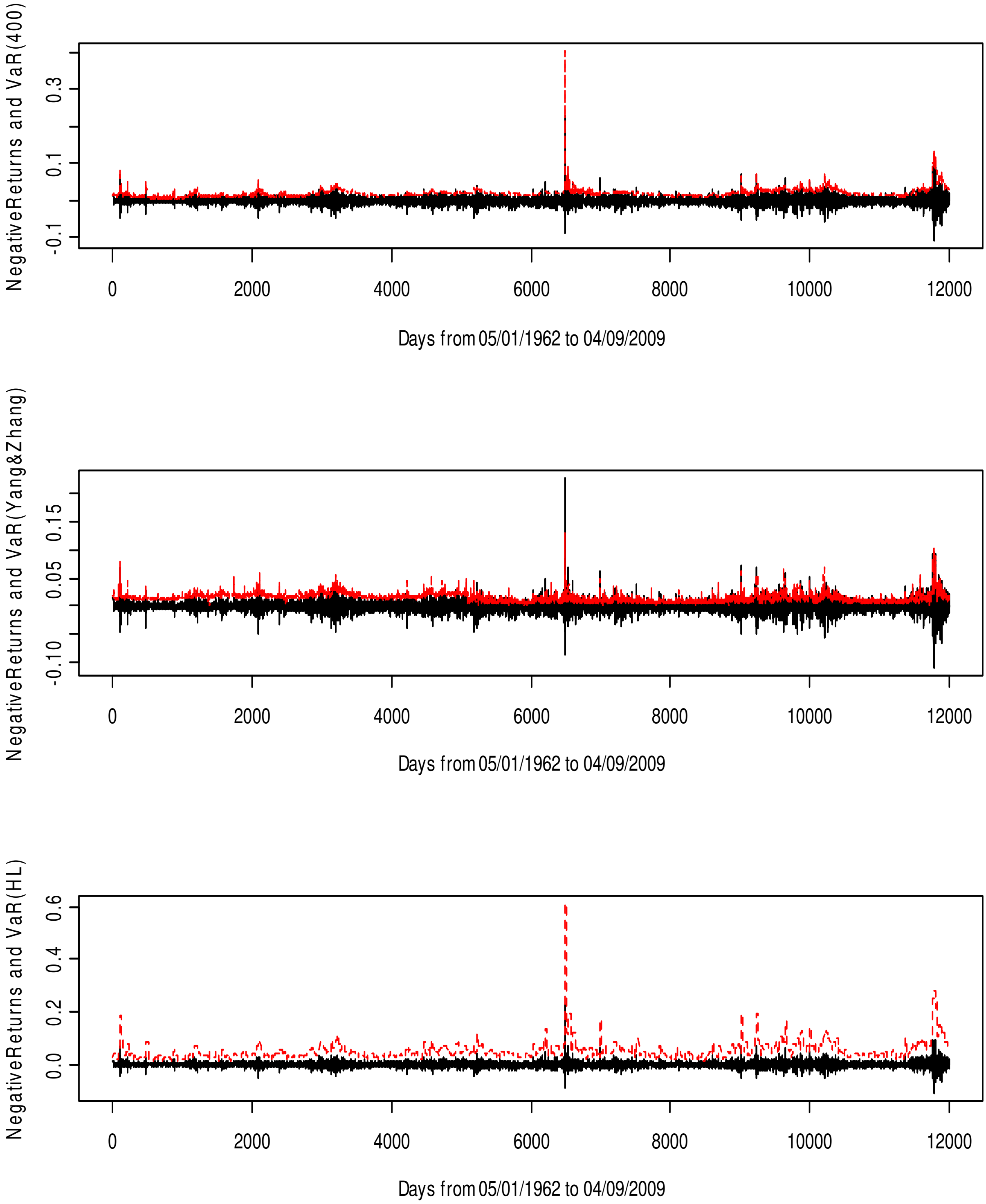

Figure5: Backtest of GARCH(1,1) (with sample size 400), Yang & Zhang and HL VaRs.

Bibliography

A. Saichev, D. Sornette and V. Filimonov (2009) *Most efficient homogenous volatility estimators.*

Acerbi, C., Tasche, D. (2001) *On the Coherence of Expected Shortfall.*

Baker Malcolm and Jeffrey Wurgler (2007) *Investor Sentiment in the Stock Market.* Journal of Economic Perspectives volume 21 (2), pg. 129-151.

Basel Committee on Banking Supervision (2006) *International Convergence of Capital Measurement and Capital Standards.*

Benoit B. Mandelbrot; John W. Van Ness (1968) *Fractional Brownian motions, fractional noises and applications.*

Benoit Mandelbrot, Adlai Fischer and Laurent Calvet (1997) *A Multifractal Model of Asset Returns.*

Bernoulli (1738) *Commentaries of the Imperial Academy of Science of Saint Petersburg*.

Carlo Acerbi Dirk Tasche (2001) *Expected Shortfall: a natural coherent alternative to Value at Risk.*

Charles Lee, Andrei Shleifer, and Richard Thaler (1991) *Investor Sentiment and the Closed-End Fund Puzzle.*

Credit Suisse First Boston International (1997) *Creditrisk^{+TM} – Technical Document.*

Garman, M. B., and Klass, M. J. (1980) *On the estimation of security price volatilities from historical data.* Journal of business.

Giovanni Barone-Adesi, Kostas Giannopoulos and Les Vosper (2000) *Filtering Historical Simulation*. *Backtest Analysis.*

Haim Levy, Enrico Degiorgi and Thorsten Hens (2003). *Prospect theory and the CAPM: A contradiction or coexistence?*

J. Bradford De long; Andrei Shleifer; Lawrence H. Summers and Robert J. Waldmann (1987) *Noise trader risk in financial markets.*

Jean-Philippe Bouchaud and Marc Potters (2000) *Theory of financial risks: From Statistical Physics to Risk Management.*

John C. Hull (2007) *The Credit Crunch of 2007: What Went Wrong? Why? What Lessons Can Be Learned?*

Kahneman, D. and A. Tversky(1979) *Prospect theory: an analysis of decisions under Risk*. Econometrica 47, 263–291.

Keynes, John Maynard (1937) *The General Theory*. In Quarterly Journal of Economics.

L. J. Savage (1962) *Foundation of Statistical Inference.*

Moody's KMV (2007) *EDF™ 8.0 Model enhancements.*

Parkinson, M. (1980) *The extreme value method for estimating the variance of the rate of return*. Journal of Business 53:61–65.

Peterson, M. (2004) *Transformative Decision Rules, Permutability, and Non-Sequential Framing*, Synthese 139, 387–403.

Quiggin, J. (1993) *Generalized Expected Utility Theory. The Rank-Dependent Mode.*

RiskMetrics Group, Inc (2007) *CreditMetrics™—Technical Document.*

Robert A. Jarrow and Stuart Turnbull (1995) *Pricing Derivatives on Financial Securities Subject to credit risk.* Journal of Finance, vol. 50.

Robert Geske, Walter Torous (1987) *Volatility and Mispricing: Robust Variance Estimation and Black-Scholes Call Option Pricing.*

Rogers, L. C. G., and Satchell, S. E. (1991) *Estimating variance from high, low and closing prices*. Annals of Applied Probability 1:504–12.

Shiller, Robert J. (2000) *Measuring Bubble Expectations and Investor Confidence.*

Stephen Figlewski (1994) *Forecasting Volatility Using Historical Data.* Journal of Business.

Stephen Figlewski (2004) *Forecasting volatility*

Torben G. Andersen , Tim Bollerslev , Steve Lange. (1999) *Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon.*

Tversky, A and Kahneman, D (1992) *Advances in Prospect Theory: Cumulative Representation of Uncertainty.*

Von Neumann J, Morgenstern O (1944) *Theory of games and economic behavior* Princeton University Press, Princeton.

Welch, Ivo and Lily Qiu (2004) *Investor Sentiment Measures*

Wilson Sy (2008) *Credit risk models: why they failed in the credit crisis?*

Yang, Dennis, and Qiang Zhang, (2000) *Drift-independent volatility estimation based on high, low, open, and close prices*. Journal of Business 73, 477-491.

Young Ho Eom, Jean Elwege and Jing-Zhi Huang (2004) *Structural models of corporate bonds pricing: An empirical analysis.*

http://icf.som.yale.edu/confidence.index

https://customers.reuters.com/community/university/default.aspx

R-functions and computations are available on request