Stability Analytics of Vulnerabilities in Financial Time Series

Diethelm Würtz, Tobias Setz, Yohan Chalabi

No. 2011-01
Stability Analytics of Vulnerabilities in Financial Time Series

Diethelm Würtz, Tobias Setz, Yohan Chalabi
Institute for Theoretical Physics
Econophysics Group
Swiss Federal Institute of Technology Zurich

First Version September 2011, Final Version August 2012

Abstract

In this article we give a brief overview to the recently introduced stability analytics to analyze and quantify vulnerabilities in financial time series. We present the underlying statistical concepts for analyzing change points and structural breaks, for detecting extreme values and outliers and for exploring non-stationarities and multi-fractal behavior. The analytics are applied to the long term U.S. Large Cap Stocks Index and the Commodity Research Bureau Index ranging back to 1925.

Introduction

To identify and quantify risk in a financial time series several statistical measures were introduced in literature. These include rolling measurements of sample variances, volatilities, value at risk, expected shortfall, spectral risk measures and distortion risk measures amongst others. In this paper we like to follow a different approach which allows to quantify "risk" by analyzing (i) structural changes and break points, (ii) extreme values and outliers, and (iii) non-stationarities caused by multifractal behavior in financial time series. These three measures can be considered as indicators for the stability of the generating process underlying a financial time series.

In the following we investigate the stability of prices and indices of financial market instruments. Note, this should not be confused with financial stability analysis where stress testing is used to explore the stability conditions of financial systems. The major uncomfortable feelings of any investor are those of large losses, deep drawdowns, and long recovery times. The measures presented in the following are concerned with risk as well but they also describe unstable changing conditions of the environment which influences the time series process. It is the goal of this paper to define and explore measures which define risk from a stability point of view.

In section 1 we briefly discuss methods to detect turning points in a financial time series process. We use a retroactive spline smoothing approach to determine the up and downs. In section 2 we apply the Bayesian product partition model of Barry and Hartigan [1992] in the Monte Carlo Markov Chain approximation to detect change and break points in a financial return series. In section 3 we embed the univariate time series in phase space and apply the robust principal component analysis to detect multivariate patterns of extreme values and outliers. The formulation we use is that of Filzmoser, Maronna, and Werner [2008]. In section 4 we compute the Morlet wavelet spectrum to make non-stationarities caused by multifractal behavior visible. We use the Morlet wavelet spectrum as formulated by Torrence and Compo [1998].

A preliminary description of the Stability Analytics was already presented by Würtz et al. [2010ab, 2011a] and applied in the chapter about "Advanced Analytics" in the handbook Long Term Statistical Analysis of U.S. Asset Classes by Diethelm Würtz et al. [2011]. From this book we have also overtaken the data.
sets for the US large cap stocks [2011] and the CRB commodity research bureau index [2011] used in this paper. All calculations were done with the R statistical software environment [2012] and the financial computing environment Rmetrics [2012].

1 Turning Points and Cycles

In the business cycle analysis of economic indicators, turning points play an important role. They identify the datings and durations of changing economic conditions like expansions and contractions. For financial data we can use similar ideas and concepts to identify increasing (bull) and decreasing (bear) market periods. The method we use is a retrospective approach which smooths the index (or price) series to a given degree. This helps to identify the peaks and/or the pits of the series. From the forecasting point of view, a turning points analysis can help to predict changing market conditions.

Vanhaelen, Dresse and De Mulder [2000] formulate the following general principles for the detection of turning points: (i) The choice of turning points should not be affected by aberrant observations, extremes or outliers, (ii) irregular movements in the series should be excluded and (iii) the application of the method should not delay the identification of the turning points for too long. The authors also suggest additional requirements which may be imposed in accordance with stylized facts of the financial time series. This may concern the succession of peaks and troughs, a minimum length for the phases and cycles and possibly a minimum amplitude for the movements.

In the business cycle literature we find that the most used approach for detecting turning points is the method of Bry and Boschan [1971]. This approach (i) first involves the detection of extremes and outliers and their replacement by a moving average. (ii) Then an initial set of turning points is identified in a smoothed series, by applying a moving average filter. (iii) The turning points detected at the different stages are checked for the alternation of peaks and troughs, for a minimum span of 15 months between two successive peaks or two successive troughs, and of 5 months between two successive turning points. Finally, (iv) turning points in the first or the last 6 months of the series are rejected, Vanhaelen, Dresse and De Mulder [2000].

Other approaches use filters to smooth the financial time series. Two examples are the (i) HP filter of Hodrick and Prescott [1997], and the (ii) CF filter of Christiano and Fitzgerald [1999], which yield smooth cycles, and as a result the Bry-Boschan routine simplifies. Usually the filtered cycles do not require several iteration steps since phase and cycle length-constraints are getting better satisfied by smoothing, Gomyai and Guidetti [2008].

Here we use an alternative approach. We found out, that one can achieve with spline smoothing, see Hastie and Tibshirani [1990], Chambers and Hastie [1992], and Green and Silverman [1994], the same effects as obtained by filtering. Therefore we applied first spline smoothing to the index values, according to Ripley and Mächler [2011], and then we searched for the peak and trough points, Ibanez [1982]. The result of the retroactive turning points analysis using the formalism of Ibanez, Grosjean, and Etienne [2011], is shown in figure 1 for two asset classes, i.e. for the (i) U.S. large cap stocks (LCAP) and the (ii) commodity research bureau index (CRB). The turning points (red bullets) as shown in figure 1 were extracted from the spline smoothed monthly index values (red curve). Note, that the index (black curve) is plotted on a logarithmic scale. The shaded areas (blue) reflect bear markets. Also, the axis for the returns (orange curve) is suppressed.

The number of bear markets depends on the level of smoothing. For a medium smoothing parameter, we observe in the case of the U.S. large cap stocks 6 down periods. The major down-periods are related to the U.S. great depression, to the U.S. sub prime crisis, and to the recent European debt crisis. For the commodities we observe 11 down-periods. The most significant ones are related to the great depression, to the oil crisis, and the recent food speculation. To identify these periods have a look at figure 1.
2 Bayesian Structural Change and Break Points Analytics

The aim of change point analysis is to partition the series into different segments that are connected and ordered with increasing time. This is done by clustering the elements of the series keeping the time order. The result is a set partition where within each segment the elements have a common parameter or a set of parameters. In the detection of change points this parameter is usually the mean. The cuts between the segments are known as change or break points. There are several models for segmentation and partitioning by piecewise mean, variance, correlation or distributional parameters.

Frequentist change point detection methods rely on hypothesis testing. The major frequentist tests are summarized by Zeileis et al. [2007]. The procedures they discuss are concerned with testing or assessing deviations from stability in the classical linear regression model. In these methods the breakpoints are located by minimizing the residual sum of squares of a piecewise linear regression model. The ideas behind the implementation can be found in Zeileis [2003]. Another approach is Circular Binary Segmentation, see Sen [1975] and Olshen [2004]. Circular binary segmentation methods use likelihood ratio statistics to test the null hypothesis of no change point. If the null hypothesis is rejected the series is split and the test is repeated on the sub-segments till no additional changes are detected, see Olshen [2007].

Here we use a Bayesian Change Point Detection method to identify change points in the mean and variance of the series. The Bayesian Product Partition Model introduced by Barry and Hartigan [1992] allows us to determine the probability of there being a change point at each location in the series. This becomes very advantageous in the stability analysis context as information about which time periods have frequently high change point probabilities and are therefore less stable than when the probability of a change is low. The Bayesian product partition model may also provide estimates of the parameters at each moment in time as well as the posterior distributions of the partitions and the number of change points. This was later extended by Loschi [2005] to give the posterior distribution of the probability that each instant in time is a change point. While frequentist procedures for change point analysis estimate specific locations of change points, the Bayesian procedure offers the probability of a change point at each location in the series. In a retroactive stability study this probability information may be more important than the location of historical changes which are often already known and are easily located. For further discussions and applications we refer to the work of Erdman and Emerson [2008].

For the two asset classes, U.S. large cap stocks (LCAP) and the commodity research bureau index (CRB) we have computed the posterior mean, the posterior variance, and the probability for observing a change point or break points in the dynamical price process of the two instruments, LCAP and CRB, respectively. The results are shown in figure 2 and 3. From the ratio of the posterior mean and the square root of the posterior variance we have also calculated a modified performance measure in the sense of the Sharpe Ratio. Further performance measures as described in the book of Bacon [2006] can be derived in the same spirit. For both indices the stability analytics give a clear indication for the periods of instabilities.

The posterior probabilities for the LCAP and CRB logarithmic returns have been smoothed by splines on several levels and are expressed by the rainbow colored curves in figure 4. The black curve about the 0.5 level expresses the difference between the strongest smoothed curve and the weakest smoothed curve which serves as a measure for stability. Less stable periods are below the 0.5 levels and above this level stability increases. The troughs of this curve indicate the lows in the stability of the LCAP and CRB indices.

The first unstable periods in the LCAP index can be found for the Great Depression. The great depression started in about 1929 and lasted until the late 1930s or early 1940s. The dot-1 in figures 2 and 4 indicates October 24, 1929, known as the Black Thursday when the stock market in New York abruptly felt. The next severe instabilities appeared 1973 as a consequence of the Oil Price Shock along with the stock bear market between January 1973 and December 1974. This crisis can be regarded as the first event since the Great Depression which had a persistent effect on the world economy. The dot-2 marks end of August 1974. The dot-3 event was caused by the Black Monday October 19, 1987, when stock markets around the world crashed while the dot-4 marks the collapse of Lehman Brothers during the U.S. Sub Prime crisis.
In figure 3 and 4 where we have plotted the posterior probabilities for the CRB Commodity Research Bureau Index, the dots mark the following events: as of December 1932 the Great Depression, as of mid of 1947 the Post World War II period, as of December 1973 the Oil Crisis, and as of March 2008 the world Food Price Crisis. Food price speculation increased dramatically in 2007 and during the 1st and 2nd quarter of 2008 creating a global crisis and causing political and economical instabilities and social unrest in both poor and developed nations. For further information and discussions on the economic and financial instabilities we refer to Wikipedia, and the references given there.

3 PCA Extreme Values and Outlier Analytics

Outlier detection methods are a way of searching for extreme and outlying patterns in the data. Identifying possible outliers and periods with many outlying data points helps to identify periods of instability. The principal component outlier analytics can be regarded in this respect. It helps to identify and locate extreme and outlying values. It is a robust method which gives an effective indication when extremes and outliers start to dominate the financial or economic time series process and create periods of high risk.

There are many different methods of outlier detection. In this study we focus on a principal components (PC) outlier detection method where we use the Mahalanobis distance as the measure of “outlyingness”. The method is able to detect outlying points in a way that is both computationally efficient and robust, Filzmoser, Maronna and Werner [2008]. A principal components analysis returns the set of principal components. The PC’s are a set of orthogonal variables which each maximize the variance in their respective directions. They are selected such that the first component explains the most variance in the data, the second component the second most and so on. Outlying data points increase the variance of the data. As the PC’s are the directions which maximize the amount of variance along each component, it is plausible that the extremes will appear more visible in PC space than in the original data space.

For the LCAP and CRB indices we performed a principal component analysis on overlapping weekly patterns binded row by row to a multivariate time series. We identified the outlying patterns using the algorithm of Filzmoser, Maronna and Werner [2008]. Based on the robustly sphered data, semi-robust principal components are computed within their approach which are needed for determining distances for each observation. Separate weights for location (mean) and scatter (variance) outliers are then computed based on these distances. The combined weights are finally used for outlier identification. For each observation the algorithm returns a numeric vector with final weights where small values indicate potential (multivariate) outliers. Also returned are two vectors with weights for the locations and scatter from which we can derive weighted means and variances.

In figure 5 we show the results from the PCA extreme value and outlier analytics for the monthly log-arithmetic returns calculated from the LCAP and CRB indices. We have also added to this graph, as in the case of the BCP Analytics, the rainbow colored smoothed curves of the numeric weights for each observation. The maximum dispersion of these curves gives us again an indication of the stability of the underlying time series process.

4 Non-Stationarities and Multifractal Wavelet Analytics

Stationarity is the fundamental requirement for any real world forecastable time series model. In this case we assume that the distribution of previous observations can be used to predict future ones. In other terms, the joint probability distribution of the observations is assumed to be stationary over time, i.e. its moments including the mean and volatility are constant. However, we often observe events in financial returns that cannot be described by the distribution of past observations. When such points are observed, the time series is considered to be non-stationary.

Our choice to analyze non-stationary behavior in asset returns is the multi-resolution approach in time / frequency space using wavelet analysis. Wavelet analysis is a widespread technique that offers the user the possibility to decompose a time series into time/frequency space, see Gençay, Selcuk, and Whitcher [2002].
Mallat [2009], and Nason [2008]. This allows one to determine the dominant modes of variability in the series and how they vary over time. The signal can be examined in time and frequency simultaneously. The wavelet transform has been found to be particularly useful for series that are aperiodic, noisy, irregular, intermittent and so on, see Addison [2002].

Wavelet functions are localized waveforms that can be altered in two ways, either moved in location along the series or altered in frequency - a stretching or squeezing of the wavelet form. At each desired location and frequency (scale) the resulting wavelet is compared to the original series to determine how well it matches. A good match results in a large transform value and vice versa. The transform values for each particular location and scale are plotted on a contour plot. For the Morlet wavelet form, the transform value is a complex number and different possibilities to plot those values are thinkable. We use the square of the absolute value (power spectrum) which has an expectation value that is approximately equal to the variance of the series, Torrence, and Compo [1998]. By examining the patterns in the wavelet contour plot it is possible to see changes in the frequency and variance of the signal as well as jumps and discontinuities very clearly.

For the monthly logarithmic returns of the LCAP and CRB indices we have computed the wavelet power spectrums. Figure 6 shows the transform values of the time series for each location and scale considered. The location or x-axis marks end-of-month dates, the scale or y-axis counts periods given in months. The contours of the image plot were taken on equidistant quantile levels ranging from yellow to red. The thick line belongs to the 95% significance level. The cross-hatched regions on either end of the graph indicate the cone of influence, where edge effects become important.

The big depression is highly visible for both indices, the LCAP and CRB. For the CRB the instabilities last much longer until the 50s. After World War II we observe a prosperous increase in the indices going together with a high stability. The first oil crisis in the 1973/74 becomes quite evident in the CRB wavelet spectrum with minor influence in the large cap stock market. The period of instability caused by the oil crisis as marked by the 95% level contours extends to more than a decade. The most recent instabilities in the commodities can be related to the 2010 food speculation. The short but nevertheless significant crashes in the stock market can be identified in the LCAP wavelet spectrum: Black Monday in October 1987, Dot Com Bubble in April 2000, 9/11 Attack in September 2001, and Sub Prime Crisis in February 2007.

5 Summary and Outlook

Bayesian change point analytics, principal component extreme value and outlier analytics, and the wavelet analytics are three flavors of analysing and quantifying the vulnerabilities to external forces of a time series process. In all three cases we can define a number to measure the strength of instabilities appearing over time.

Evenmore, the time dependent process of stability measured on a rolling window can be used to define (like for the volatility) a value for the average (mean) and variability (variance) of the changing stability over time. We can derive a cumulated stability index, and from the distribution of its changes we can derive and calculate measures like value-at-stability or an expected shortfall stability as we know them from the financial returns for the value-at-risk and the expected shortfall risk. Also many performance measures as described by Bacon [2006] can be designed from stabilities, like we know it for example from the Sharpe ratio of financial returns.

The stability measures can also be used for rating and ranking equities or other financial assets. In this case we select those instruments with the highest stability and with lowest fluctuations. On the other hand we can formulate a stability parity approach in portfolio optimization or for the creation of new indices based on stability parity indexation concepts. Like in risk budgeting, see for example Demey, Mailllard, and Roncalli [2011], we can introduce the concept of stability budgeting which can help to design better diversified portfolios and indices. Finally we like to remark that stability analytics can also be used to compute the sensitivity with respect to market stability, and to construct scenarios for the identification of vulnerabilities.
Figure 1: Monthly turning points analytics of the U.S. Large Cap Stocks Index (upper graph) the CRB Commodity Research Bureau Index (lower graph) starting in December 1925 and ending in December 2010. For a detailed description we refer to the text.
Figure 2: Monthly Bayesian change points analytics of the logarithmic returns of the Large Cap Stocks Index starting in December 1925 and ending in December 2010. From up to down the curves show the posterior mean, the posterior variance, the posterior probability, and the probability weighted ratio of the posterior mean and standard deviation (root of variance). The four dots mark (1) the Black Thursday 1929, (2) the Oil and Stock Crisis 1973/74, (3) the Black Monday 1987, and (4) the collapse of Lehman Brothers.
Figure 3: Monthly Bayesian change points analytics of the logarithmic returns of the Commodity Research Bureau Index starting in December 1925 and ending in December 2010. From up to down the curves show the posterior mean, the posterior variance, the posterior probability, and the probability weighted ratio of the posterior mean and standard deviation (root of variance). The four dots mark (1) the Great Depression as end of 1932, (2) the Post World War II period as mid of 1947 (3) the Oil Price Crisis as end of 1973, and (4) the food price speculation as end of the first quarter of 2008.
Figure 4: Monthly Bayesian change points analytics of the U.S. Large Cap Stocks Index (upper graph) and the CRB Commodity Research Bureau Index (lower graph) starting in December 1925 and ending in December 2010. For a detailed description we refer to the text. The dots number the same events as in figures 2 (LCAP) and 3 (CRB).
Figure 5: PC Outlier Analytics - U.S. Large Cap Stocks and CRB Commodity Research Bureau Index: The figures show the posterior probability for each date of having an outlier within the time series. The rainbow colored time series are smoothed curves of the original probability series. The black curve is the difference between the strongest smoothed curve and the weakest smoothed curve which serves as a measure for stability. The dots number the same events as in figures 2 (LCAP) and 3 (CRB).
Figure 6: Wavelet Analytics - U.S. Large Cap Stocks (upper) and CRB Commodity Research Bureau Index (lower image). The dots number the same events as in figures 2 (LCAP) and 3 (CRB).
References


Bry G., and Boschan C., [1971], Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, Technical Paper 20, NBER.


Erdman Chandra, and Emerson John W. [2008], A fast Bayesian change point analysis for the segmentation of microarray data, Bioinformatics 24, 2143–2148.


R, Statistical computing environment, www.cran-project.org


US Large Cap Stocks [2011], Sources: Ibbotson SBBI Yearbook 2010, Standard and Poors, and iShares ETF Database.


About the Authors

Diethelm Würtz is Private Lecturer at the “Institute for Theoretical Physics” at the Swiss Federal Institute of Technology (ETH) in Zurich. His research interests are in the field of risk management and stability analysis of financial markets. He teaches computational science and financial engineering. He is senior partner of the ETH spin-off company “Finance Online” and president of the “Rmetrics Association in Zurich”.

Tobias Setz has a Bachelor and Master degree in Computational Science from ETH Zurich. He started his PhD Thesis in the Econophysics Group of Professor Diethelm Würtz. His major research interests are in Bayesian statistics and Wavelet analytics. He is an experienced R programmer and co-maintainer of the Open Source Rmetrics software environment.

Yohan Chalabi has a master in Physics from the Swiss Federal Institute of Technology in Lausanne. He is now a PhD student in the Econophysics group at ETH Zurich at the Institute for Theoretical Physics. Yohan is a co-maintainer of the Rmetrics packages.

Acknowledgement

The work presented in this article was partly supported by grants given by ETH Zurich and Rmetrics Association Zurich.
Disclaimer:
This document is copyrighted and its content is confidential and may not be reproduced or provided to others without the express written permission of the authors. This material has been prepared solely for informational purposes only and it is not intended to be and should not be considered as an offer, or a solicitation of an offer, or an invitation or a personal recommendation to buy or sell any stocks and bonds, or any other fund, security, or financial instrument, or to participate in any investment strategy, directly or indirectly. It is intended for use in research only by those recipients to whom it was made directly available by the authors of the document.