Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 09-09

Comovements of Different Asset Classes During Market Stress

Jan Piplack
Stefan Straetmans

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Janskerkhof 12
3512 BL Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the first author.

Jan Piplack
Utrecht University
Utrecht School of Economics
Janskerkhof 12
3512 BL Utrecht
The Netherlands.
E-mail: J.Piplack@uu.nl
RABO bank
Croeselaan 18
3521 CB Utrecht
The Netherlands.

Stefan Straetmans
University of Maastricht
Department of Finance
P.O. Box 616
6200 MD Maastricht
The Netherlands
s.straetmans@finance.unimaas.nl

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 09-09

Comovements of Different Asset Classes During Market Stress

Jan Piplack[a]
Stefan Straetmans[b]

[a]Utrecht School of Economics
Utrecht University
RABO bank,
Utrecht the Netherlands

[b]Department of Finance
Maastricht University

May 2009

Abstract

This paper assesses the linkages between the most important U.S. financial asset classes (stocks, bonds, T-bills and gold) during periods of financial turmoil. Our results have potentially important implications for strategic asset allocation and pension fund management.
We use multivariate extreme value theory to estimate the exposure of one asset class to extreme movements in the other asset classes. By applying structural break tests to those measures we study to what extent linkages in extreme asset returns and volatilities are changing over time. Univariate results andch bivariate comovement results exhibit significant breaks in the 1970s and 1980s corresponding to the turbulent times of e.g. the oil shocks, Volcker's presidency of the Fed or the stock market crash of 1987.

Keywords: Flight to quality, financial market distress, extreme value theory.

JEL classification: G01, G15.

1 Introduction

Joint crashes in different asset markets can have severely destabilizing effects on countries and the international financial architecture. Strong financial market linkages during crisis periods can severely increase the risk of bank failures through a joint deterioration of their assets possibly leading to domino effects in countries' financial systems even when their banks' assets are well diversified. In general, extreme comovements of financial markets on a national and international level crucially determine the systemic risk of these markets. The amount and size of jointly affected markets together with potential difficulties and bottlenecks in the financial system and payment process determines the severeness of any real effects that may follow. There have been periods with financial and political instability like the oil crises in 70s and 80s, the Asian Flu and Russian Cold (1997 and 1998, respectively) or more recently the subprime mortgage crises in 2007, where such effects have been witnessed. Thus, the study of extreme (co)movements in asset markets is not only important to investors but also to policy makers and financial regulators that care about overall economic and financial stability.

Possibly the first systematic study of cross-country financial crisis spillovers is Morgenstern (1956, Chapter X). He explicitly refers to "statistical extremes" of the 23 stock markets and their effects on foreign stock markets. More recently, the econometric literature utilizes correlation analysis based, for example, on ARCH and GARCH-type models. Such contributions usually examine if stock market comovements differ between crisis and non-crisis episodes and typically also try to determine the direction of possible spill-over effects. Contributions like King and Wadhwani (1990), Hamao et al. (1990), Mallaris and Urrutia (1992), Lin et al. (1994) and Engle and Susmel (1993) belong to this strand of the literature. Papers focussing on foreign exchange markets and currency crises include Eichengreen et al. (1996), Sachs et al. (1996), Kaminski and Reinhart (2000). However, little work has been done on linkages across asset classes. Hartmann et al. (2004) constitutes a notable exception.

This paper extends the literature by increasing the amount of asset classes considered. This allows us to study and compare phenomena like "flight to quality", and "flight to liquidity". We define flight to quality as the simultaneous event of a stock market crash and a boom in either government bond or gold markets; whereas flight to liquidity stands for a stock market crash coinciding with a boom in the market for T-bills. Compared to the scant existing literature on cross-asset linkages, we use more assets and longer time series. This allows us to implement extreme value techniques and to apply tests for structural change on our linkage measures.

The used methodology combines extreme value theory (EVT) with a structural stability test developed by Quintos et al. (2001). Contributions using similar approaches include Hartmann et al. (2004; 2005), and Straetmans et al. (2006). Bivariate extreme value theory captures the dependence structure in the tails of multivariate distributions by means of the so-called tail dependence parameter. This parameter is able to capture both linear and nonlinear dependence in the tails whereas traditional correlation analysis only measures linear dependence and is predisposed toward the multivariate normal distribution. Another advantage constitutes the nonparametric character of the used methodology, i.e., we leave the joint asset return process unspecified and thereby limit the scope for miss-specification (model risk).

Anticipating on our results, we find relatively small tail indexes for gold and T-bills as compared to stocks and bonds. Bivariate results indicate that the likelihood of co-crashes dominates flight to quality and flight to liquidity phenomena. As concerns structural change, both univariate and bivariate tails are found to be nonstable over time for certain asset pairs. The breaks suggest a mean reverting pattern in the amount of tail thickness and tail dependence: initially the probability mass has risen (oil shocks) to decline later on towards the end of the 80s. Tail asymmetries as well as cross sectional differences in tail estimates are found to be statistically insignificant from zero.

The paper is organized as follows. Section 2 introduces the theoretical basis for the extreme value analysis. Section 3 explains the tail-dependence measure for extreme financial market comovements in more detail. In Section

4 we introduce the stability test that we perform in order to check for structural breaks in the univariate and multivariate series. Section 5 presents the results obtained by applying those techniques to the data. Section 6 concludes.

2 Asset linkages: Theory

We measure the dependencies between returns of different asset classes for extreme price movements, i.e. in the bivariate tail of the asset pairs. They are constructed either as conditional tail probabilities or conditionally expected extreme co-events. We will argue that the two indicators are perfectly correlated and are two alternative ways to presenting the same empirical outcomes. The techniques used are not new and have partly been used in, for example, Poon et al. (2004) and Hartmann et al. (2004; 2005).

In this section and Section 3 we assume constancy/stationarity of the tail behavior of assets over time. We also focus on the unconditional marginal return distributions and do not condition any statistic on time. Therefore, we refrain from using time subscripts even though the reader should bear in mind that the assumed asset return series evolve over time. In Section 4 we introduce time subscripts t because we relax the assumption of constancy of the tail behavior and allow (test) for structural breaks.

2.1 Conditional tail probabilities

Consider a pair of different asset types, i.e., stocks and bonds. Denote the return of stocks and bonds by the random variables X_i $(i = 1, 2)$, respectively. Each series X_i is assumed to have n observations. For sake of convenience and when necessary, we take the negative of returns, so that we can define all used formulae in terms of upper tail returns. Crisis levels or extreme percentiles Q_i $(i = 1, 2)$ are chosen such that the tail probabilities are equal across assets, i.e., $P\{X_1 > Q_1\} = P\{X_2 > Q_2\} = p$.

With common marginal exceedance probabilities, crisis levels Q_i (Value-at-Risk/VaR) will generally not be the same across assets, because the marginal distribution functions $P\{X_i > Q_i\} = 1 - F_i(Q_i)$ are nonidentical. Crisis levels can be interpreted as 'barriers' that will on average only be broken once in $1/p$ time periods, i.e., p^{-1} days in case of daily data frequency. Suppose now that we want to measure the dependence between two assets beyond the crisis levels (Q_1, Q_2). A natural measure is the conditional tail probability

$$\begin{aligned} \beta_\tau &: = P\{X_1 > Q_1(p) \,|X_2 > Q_2(p)\} \\ &= \frac{P\{X_1 > Q_1(p), X_2 > Q_2(p)\}}{P\{X_2 > Q_2(p)\}} \\ &= \frac{P\{X_1 > Q_1(p), X_2 > Q_2(p)\}}{p}, \end{aligned} \tag{1}$$

which measures the likelihood that an asset's value (in this case X_1) falls sharply, if there is an extreme negative shock to a second asset. In case of independence the conditional tail probability reduces to $p^2/p = p$, which constitutes a lower bound that helps to judge the strength of assets' tail dependence.

2.2 Conditionally expected extreme events

Alternatively, suppose we would like to find the expected number of assets' extremes (booms or busts) given that one observes a boom or bust in at least one asset class. Using the same notation as before, we represent random asset returns by X_1 and X_2. Q_1 and Q_2 are the corresponding percentiles (or 'thresholds') above which we speak of a market boom or crash (in case of a loss) and that will only be exceeded with probability p. Let κ stand for the number of assets with extreme returns, i.e. κ equals one or two. Our extreme linkage indicator is the conditional expectation $E[\kappa|\kappa \geq 1]$. From elementary probability theory (starting from the standard definition of conditional probability) we can state that

$$\begin{aligned}E[\kappa|\kappa \quad &\geq \quad 1] := \frac{E[\kappa]}{P\{\kappa \geq 1\}} \\ &= \quad \frac{P\{X_1 > Q_1, X_2 \leq Q_2\} + P\{X_1 \leq Q_1, X_2 > Q_2\}}{P\{X_1 \geq Q_1 \text{or } X_2 \geq Q_2\}} \\ &\quad + \frac{2P\{X_1 > Q_1, X_2 > Q_2\}}{P\{X_1 \geq Q_1 \text{or } X_2 \geq Q_2\}} \\ &= \quad \frac{2p}{P\{X_1 \geq Q_1 \text{or } X_2 \geq Q_2\}} \end{aligned} \tag{2}$$

with $P\{X_1 \geq Q_1 \text{or } X_2 \geq Q_2\} = 1 - P\{X_1 \leq Q_1, X_2 \leq Q_2\}$. Notice that the conditional expectation reduces to $2/(2-p)$ under the benchmark of independence. It is also easily observed that $E[\kappa|\kappa \geq 1] = P\{\kappa = 2|\kappa \geq 1\} + 1$, so that an alternative interpretation of our extreme linkage indicator is in terms of (1 plus) the conditional probability that both assets simultaneously boom or bust given that at least one asset exhibits extreme behavior. For higher dimensions than two $E[\kappa|\kappa \geq 1]$ is still equal to the ratio of the sum of the marginal excess probabilities divided by the joint failure probability. The relation between both extreme linkage measures (1) and (2) easily follows from the following chain of equalities:

$$\begin{aligned}E[\kappa|\kappa \quad &\geq \quad 1] = \frac{2p}{P\{X_1 \geq Q_1 \text{or } X_2 \geq Q_2\}} \\ &= \quad \frac{2p}{2p - P\{X_1 \geq Q_1, X_2 \geq Q_2\}} \\ &= \quad \frac{2}{2-\beta_\tau}.\end{aligned}$$

Clearly, $1 \leq E \leq 2$ corresponds with $0 \leq \beta_\tau \leq 1$.

3 Estimation of the linkage indicators

The estimation of (1) and (2) reduces to the estimation of the joint probability $P\{X_1 \geq Q_1, X_2 \geq Q_2\}$. Within the framework of a parametric probability law, the calculation of the proposed multivariate probability measures is straightforward, because one can estimate the distributional parameters by, e.g., maximum likelihood techniques. However, if one makes the wrong distributional assumptions, the linkage estimates may be severely biased due to misspecification. As there is no clear evidence that all asset returns follow the same distribution − even less so for the crisis situations we are interested in here − we want to avoid very specific assumptions for assets' returns. Therefore, we implement the semi-parametric EVT approach proposed by Ledford and Tawn (1996); see also Draisma et al. (2001), and Poon et al. (2004) for recent applications). Loosely speaking, their approach consists of generalizing some 'best practice' in univariate extreme value analysis.

Before proceeding with the modeling of the extreme dependence structure, however, it is worthwhile to eliminate any possible influence of marginal aspects on the joint tail probabilities by transforming the original variables to a common marginal distribution. After such a transformation, differences in joint tail probabilities can be solely attributed to differences in the tail dependence structure of the extremes. Thus our dependence measures, unlike e.g. correlation, are no longer influenced by the differences in marginal distribution shapes. To this aim we transform asset returns (X_1, X_2) to unit Pareto marginals:

$$\widetilde{X}_i = \frac{1}{1 - F_i\left(X_i\right)},\ i = 1, 2, \tag{3}$$

with $F_i\left(\cdot\right)$ representing the marginal cumulative distribution function (cdf) for X_i.[1] This variable transform leaves the joint tail probability in the numerator of (1) invariant because

$$P\left\{X_1 > Q_1\left(p\right), X_2 > Q_2\left(p\right)\right\} = P\left\{\widetilde{X}_1 > s, \widetilde{X}_2 > s\right\},$$

with $s = 1/p$.[2] The estimation problem can now be simplified toward esti-

[1] Since $F_{1,2}$ are unknown, we replace them with their empirical counterparts. For each X_i this leads (with a small modification to prevent division by 0) to:

$$\widetilde{X}_i = \frac{1}{1 - R_{X_i}/(n+1)},\ i = 1,\, 2,$$

where $R_{X_i} = rank(X_{ij}, j = 1, \cdots, n)$.

[2] The joint probability stays invariant under any monotonically increasing transforma-

mating a univariate exceedance probability for the cross-sectional minimum of the two return series, i.e., it is always true that:

$$P\left\{\widetilde{X}_1 > s, \widetilde{X}_2 > s\right\} = P\left\{\min\left(\widetilde{X}_1, \widetilde{X}_2\right) > s\right\} = P\left\{Z_{\min} > s\right\} \tag{4}$$

The marginal tail probability at the right-hand side can now be easily calculated by making an additional assumption on the univariate tail behavior of $Z_{\min}$. Ledford and Tawn (1996) argue that the bivariate dependence structure is a regularly varying function under fairly general conditions. Draisma et al. (2001) give sufficient conditions and further motivation. Therefore, we assume that the auxiliary variable $Z_{\min}$ has a regularly varying tail. An intuitive justification of the regular variation assumption for the bivariate tail lies in the generally observed regular variation (heavy tails or non-normality) of the original return series X_1 and X_2. Thus, it is reasonable to assume that the transformed series $\widetilde{X}_i$ and hence the series of the cross-sectional minimum in (4) should also inherit this property. Upon assuming that $Z_{\min}$ exhibits a fat tail, the regular variation assumption means that the marginal excess probability for the tail of the auxiliary variable in (4) has a Pareto tail decline:

$$P\left\{Z_{\min} > s\right\} \approx L\left(s\right) s^{-\alpha}, \quad \alpha \geq 1 \tag{5}$$

with s large (p small) and where $L\left(s\right)$ is a slowly varying function.[3] Distributions with a Pareto-type tail decline have bounded moments only up to α, where α is the 'tail index' of $Z_{\min}$. In contrast, distributions with exponentially decaying tails (e.g. the normal df) or with finite endpoints have all moments bounded. So, the larger α the thinner is the tail of a distribution.[4] We can now distinguish two cases in which the $\widetilde{X}_i$ $(i = 1, 2)$ are either *tail dependent or independent.* In the former case, $\alpha = 1$ and

$$\lim_{s\to\infty} P\left\{\widetilde{X}_1 > s \,\middle|\, \widetilde{X}_2 > s\right\} > 0.$$

tion of the marginals.

[3]i.e., $lim_{s\to\infty} L\left(ts\right)/L\left(s\right) = 1$ for all fixed $t > 0$.

[4]Such an interpretation holds for the univariate and for the multivariate case.

Stated otherwise, the tail probability defined on the pair of random variables (X_1, X_2) does not vanish in the bivariate tail. Examples of asymptotically dependent random variables include the multivariate student-t distribution and the multivariate logistic distribution, see e.g. Longin and Solnik (2001), Poon et al. (2004). For *asymptotic independence* of the random variables $\alpha > 1$, and we have that

$$\lim_{s\to\infty} P\left\{\widetilde{X}_1 > s \,\middle|\, \widetilde{X}_1 > s\right\} = 0.$$

Examples of this class of distributions include the bivariate standard normal distribution or the bivariate Morgenstern distribution. For the bivariate normal with nonzero correlation coefficient ρ, the auxiliary variable's tail descent in (4) will be governed by $\alpha = 2/(1+\rho)$ whereas the bivariate Morgenstern corresponds with $\alpha = 2$. Notice that we only reach $\alpha = 2$ for the bivariate standard normal when $\rho = 0$. In general, whenever the $\widetilde{X}_i$ $(i = 1, 2)$ are fully independent, $\alpha = 2$ and $P\{Z_{\min} > s\} = p^2$. But the reverse is not true, i.e., there are joint distributions with nonzero pairwise correlation that nevertheless have $\alpha = 2$. The above-mentioned Morgenstern model provides an example. When the normal random variables are independent $(\rho = 0)$, the joint excess probability is also governed by $\alpha = 2$.

The steps (3), (4) and (5) show that the estimation of joint probabilities like in (4) can be reduced to a univariate estimation problem. Univariate excess probabilities can be estimated by using the semi-parametric probability estimator from De Haan et al. (1994):

$$\widehat{p}_s = \frac{m}{n} \left(Z_{n-m,n}\right)^{\alpha} s^{-\alpha}, \tag{6}$$

where the 'tail cut-off point' $Z_{n-m,n}$ is the $(n-m)$-th ascending order statistic (or loosely speaking the m-th smallest return with m being the amount of returns belonging to the tail of the distribution) of the auxiliary variable Z_{min}.[5] Below we explain how we chose m.

The probability estimator (6) still needs a tail index estimate α as an

[5]Such a procedure can also be used for more than two return series as is done in, for example, Hartmann et al. (2005).

input. We estimate the tail index of the $Z_{\min}$ series by means of the popular Hill (1975) statistic:

$$\widehat{\alpha} = \left(\frac{1}{m} \sum_{j=0}^{m-1} \ln \left(\frac{Z_{n-j,n}}{Z_{n-m,n}} \right) \right)^{-1}, \tag{7}$$

where m has the same value and interpretation as in (6). Further details on the Hill estimator and related procedures to estimate the tail index are provided in Jansen and de Vries (1991) or the monograph by Embrechts et al. (1997).[6]

The above discussion demonstrates that the pair of estimators in (6)-(7) both characterizes univariate and multivariate tail behavior. This is because the estimation of a joint exceedance probability can be reduced to estimating a univariate exceedance probability. In the latter case, the tail index α not only signals the tail thickness of the auxiliary variable $Z_{\min}$ but it also reflects the strength of the dependence in the tails of the original return pair (X_1, X_2) in the tail area $[Q_1, \infty\rangle \times [Q_2, \infty\rangle$. The smaller the value of α the higher the probability mass in the tail of $Z_{\min}$ and thus also the higher the value of the joint probability in (1). One therefore often calls the inverse parameter $\eta = 1/\alpha$ the tail dependence coefficient. An estimator of the bivariate tail probability measure in (1) now easily follows by combining (6) and (7):

$$\begin{aligned} \widehat{\beta}_\tau &= \frac{\widehat{p}_s}{p} \\ &= \frac{m}{n} \left(Z_{n-m,n} \right)^{1/\widehat{\eta}} s^{1-1/\widehat{\eta}} \end{aligned} \tag{8}$$

for large but finite $s = 1/p$. When the original pair of returns exhibit asymptotic independence ($\eta < 1$), the tail probability is a declining function of the threshold s and converges to zero if $s \to \infty$. On the other hand, in the polar case of asymptotic or tail dependence ($\eta = 1$), the tail probability will always be above zero (regardless of the value of the conditioning percentile).

[6]Hill (1975) derived asymptotic consistency and normality of the Hill estimator under an i.i.d. assumption. Hsing (1991) and Resnick and Stărică (1998) derive similar results for the case of dependent data. For technical details we refer to the respective papers.

However, in this paper we will not focus on the asymptotic dependence vs. independence debate and will leave the tail dependence coefficient unrestricted. Moreover, Poon et al. (2004) already noticed that wrongly imposing asymptotic dependence ($\eta = 1$), if the returns are actually independent in the limit, might lead to severe overestimation of extreme linkage measures like (1). Thus, our approach is more flexible and we avoid the risk of overestimation.

Notice that the Hill statistic (7) still requires the choice of a nuisance parameter m, i.e., where do we let the tail start? Goldie and Smith (1987) suggest to select m such as to minimize the asymptotic mean-squared error (AMSE) of the Hill statistic in (7). Such a minimum should exist because of the bias-variance trade-off that is characteristic for the Hill estimator. This idea of balancing the bias and variance has become the cornerstone for most empirical techniques to determine m. We opted for the Beirlant et al. (1999) algorithm who proposed to use an exponential regression model (ERM) on the basis of scaled log-spacings between subsequent extreme order statistics from a Pareto-type distribution. Running Least Squares regressions on this exponential regression model allows one to estimate the AMSE for different m-values and to choose the optimal m that minimizes the AMSE. For more details on the algorithm we refer to the cited reference.

4 Hypothesis testing

In this section we introduce tests that can be used to assess various hypotheses regarding the temporal stability and cross sectional equality of the considered asset linkage indicators. The first one allows to test for the structural stability of the two indicators whereas the second test compares linkage indicators both across asset pairs and across time.

4.1 Time variation

The theory up to now assumed stationarity of tail behavior over time. From e.g. a strategic asset allocation perspective, however, it is important to know whether these interdependencies stay constant over time. As the discussion of the Ledford and Tawn (1996) approach toward estimating (1) has shown, the structural (in)stability of the indicators will critically depend on whether the tail dependence parameter η is constant or not. We therefore study possible temporal shifts in η with a recently developed structural stability test for the Hill statistic (7).

Quintos et al. (2001) present a number of tests for identifying a single unknown break in the estimated tail index $\widehat{\alpha}$. As our estimation approach allows to map the multivariate dependence problem into a univariate estimation problem, we can choose from them the best test procedures for our tail dependence parameter η. Balancing the prevention of type I and type II errors we opt for their recursive test.

Let t denote the endpoint of a sub-sample of size $w_t < n$. The recursive estimator for the tail dependence parameter η is calculated from (7) for subsamples $[1; t] \subset [1; n]$[7]:

$$\widehat{\eta}_t = \frac{1}{m_t} \sum_{j=0}^{m_t - 1} \ln \left(\frac{Z_{t-j,t}}{Z_{t-m_t,t}} \right), \qquad (9)$$

with $m_t = \kappa t^{2/3}$.[8]

The value of the recursive test statistic equals the supremum of the following series:

$$Y_n^2(t) = \left(\frac{t m_t}{n} \right) \left(\frac{\widehat{\eta}_n}{\widehat{\eta}_t} - 1 \right)^2 . \qquad (10)$$

Expression (10) compares the recursive value of the estimated tail parameter (7) with its full sample counterpart $\widehat{\eta}_n$. The null hypothesis of interest is that the tail dependence parameter does not exhibit any temporal changes.

[7]Subscripts t now indicate that we relaxed the assumption of stationary tail behavior. All variables with t as subscript now refer to a subsample of the full sample $1, ..., n$.

[8]Full sample values of m are determined by means of the Beirlant et al. (1999) exponential regression algorithm. In accordance with the minimization criterion of Goldie and Smith (1987), the theoretical value of m should be related to the sample size in a nonlinear way, i.e., $m = \kappa n^{\gamma}$. Setting $\gamma = 2/3$ and having obtained an estimate of m from the Beirlant algorithm we can solve for the scaling factor $\kappa = m/n^{2/3}$. Finally, subsample values for the recursive test can be determined using the scaling variable κ, i.e., $m_t = \kappa t^{2/3}$ with t the recursive subsample size.

More specifically, let η_t be the dependence in the left tail of Z.[9] The null hypothesis of constancy then takes the form

$$H_0 : \eta_{[nr]} = \eta, \quad \forall r \in R_\varepsilon \ = [\varepsilon; 1-\varepsilon] \subset [0;1] \ , \tag{11}$$

where $[\cdot]$ is the integer part operator. Without prior knowledge about the direction of a break, one is interested in testing the null against the two-sided alternative hypothesis $H_A : \eta_{[nr]} \neq \eta$. For practical reasons the above test is calculated over compact subsets of $[0;1]$, i.e., t equals the integer part of nr for $r \in R_\varepsilon = [\varepsilon; 1-\varepsilon]$ and for small $\varepsilon > 0$. Sets like R_ε are often used in the construction of parameter constancy tests (see, e.g., Andrews (1993)).[10] In line with Quandt's (1960) pioneering work on endogenous breakpoint determination in linear time series models, the candidate break date r can be selected as the maximum value of the test statistic (10), because at this point in time the constancy hypothesis is most likely to be violated.

Quintos et al. (2001) derived asymptotic critical values for the sup-value of (10) but these are not applicable in our framework. First, Quintos et al. (2001) assume that m is selected in such a way that the Hill estimator, stability test and resulting critical values are not marred by asymptotic bias. In practice, however, nearly all algorithms (including the Beirlant et al. (1999) algorithm that we implement) based on Asymptotic Mean Squared Error (AMSE) minimization induce an asymptotic bias term in the critical values. Also, the critical values can be further biased by nonlinear dependencies like, e.g., ARCH effects (volatility clustering).

We decided to determine the critical values by means of a parametric bootstrap of the recursive test while m and it subsample counterpart m_t are chosen by means of the Beirlant algorithm. In order to take account of the

[9] In case one uses this for the univariate return series one just has to replace Z by X.

[10] The restricted choice of r implies that $\varepsilon n \leq t \leq (1-\varepsilon)\, n$. When the lower bound would be violated the recursive estimates might become too unstable and inefficient because of too small sub-sample sizes. On the other hand, the test will never find a break for t equal or very close to n, because the test value (10) is close to zero in that latter case. Thus, for computational efficiency one might stop calculating the tests beyond the upper bound of $(1-\varepsilon)\, n < n$. We search for breaks in the $[0.15n; 0.85n]$ subset of the total sample, as Andrews does.

temporal dependence in the data and the possibility of volatility spillovers from one series to another, we use bivariate GARCH models as the basis for our parametric bootstrap. In order to keep the amount of parameters to be estimated as low as possible we chose for a diagonal BEKK(1,1,1) model first described by Engle and Kroner (1995). In general the BEKK(1,1,K) model can be defined as:

$$H_t = C'C + \sum_{k=1}^{K} A_k' \epsilon_{t-1} \epsilon_{t-1}' A_k + \sum_{k=1}^{K} G_k' H_{t-1} G_k, \tag{12}$$

where C, A_k, and G_k are $N \times N$ matrices, C is upper triangular and H_t is the conditional covariance matrix at time t. Thereby, the full model is being characterized by the following equation:

$$Y_t = \Gamma + \sigma_t, \tag{13}$$

where $\sigma_t \sim N(0, H_t)$ and Y_t represents the a 2×1 vector of the assets' returns at time t and Γ gives the average daily return. After estimating this model for all possible asset combinations we use the estimated coefficients and saved residuals for the parametric bootstrap of (10).

Quintos et al. (2001) report a Monte Carlo study that indicates good small sample power, size and bias properties of the recursive break test. Only in the case of a decrease of extreme tail dependence under the alternative hypothesis ($\eta_1 > \eta_2$) they detect less acceptable power properties. We solve this problem by executing the recursive test both in a "forward" version and a "backward" version. The forward version calculates the sub-sample ηs in calendar time, and the backward version in reverse calendar time. If a downward break in η occurs and the forward test does not pick it up, then the backward test corrects for this.

4.2 Cross-sectional variation

We would also like to know whether cross-sectional differences in linkage indicators for various asset pairs are statistically and economically significant.

The asymptotic normality of $\widehat{\eta}$ enables some straightforward hypothesis testing.[11] However, equality tests based on the full sample values of the tail dependence parameter η are expected to be distorted if η values exhibit structural breaks. A test for the cross sectional equality of tail dependence parameters (null hypothesis) over time seems therefore more appropriate and can be based on the following statistic:

$$Q_t = \frac{\widehat{\eta}_{1,t} - \widehat{\eta}_{2,t}}{s.e.\left(\widehat{\eta}_{1,t} - \widehat{\eta}_{2,t}\right)}, \tag{14}$$

with $\widehat{\eta}_{1,t}$ and $\widehat{\eta}_{2,t}$ standing for recursive estimates of the tail dependence of asset pairs to be compared. The test statistic should be close to normality provided t is sufficiently large.[12] Accordingly, the asymptotic critical values are 1.65, 1.96 and 2.58 for the 10%, 5% and 1% significance levels, respectively. In the empirical applications below the asymptotic standard error in the test's denominator (14) is estimated using a nonparametric asymptotic variance estimator proposed by Drees (2003) that is robust for general nonlinear temporal dependence in the data.

5 Extreme asset linkage results: Stocks, bonds, T-bills, gold

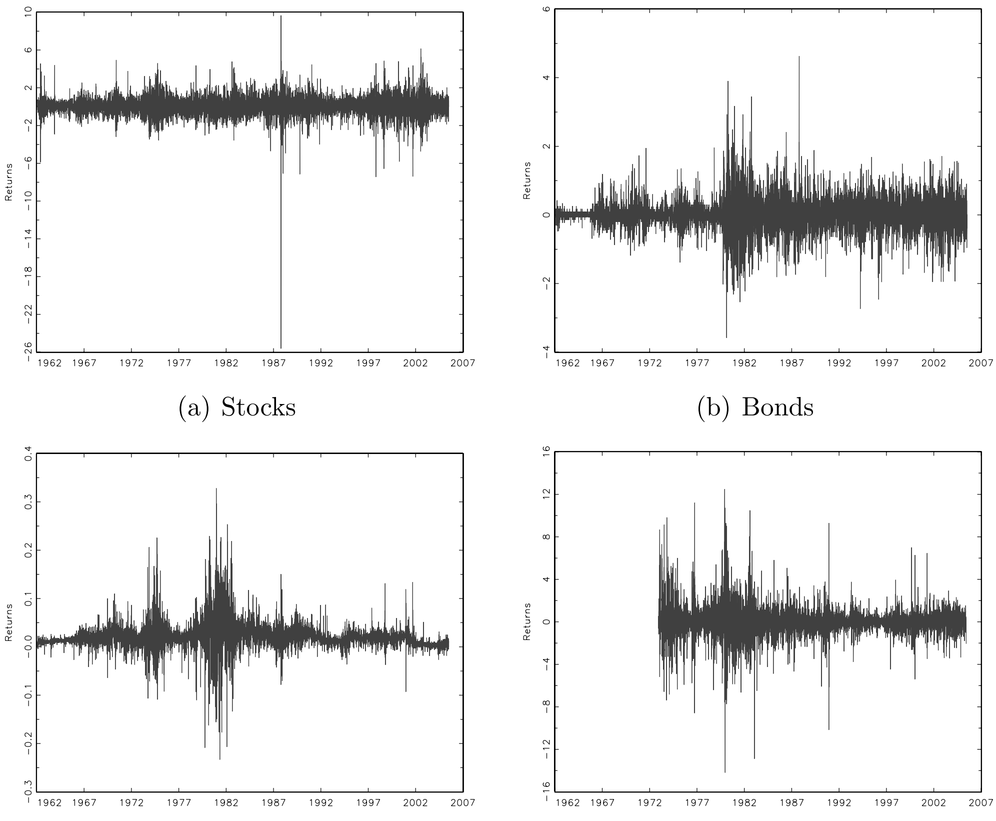
In this section we assess the likelihood of extreme return exceedances and co-exceedances for different asset classes in the U.S. financial markets. The data consist of 11,327 daily observations for stocks, bonds, and T-bills and 8,480 daily observations for gold. Time series for stocks, bonds and T-bills roughly span the period 1962-2005. We take the Dow Jones Industrials Index, ten year constant maturity government bonds, and three month constant maturity US government T-bills, respectively. Gold price series are significantly shorter and only start after the demise of the Bretton Woods system and the

[11] Asymptotic normality of the estimator has been established in, for example, Hsing (1991), Quintos et al. (2001) and Drees (2003).

[12] One can safely assume that Q comes sufficiently close to normality for empirical sample sizes as the one used in this paper (see, e.g., Hall (1982), or Embrechts et al. (1997).

related abolishment of gold-US$ convertibility in the beginning of the 1970s. A more detailed description of the data is given in the appendix. The series are plotted in Figure 1.

Figure 1: Returns of stocks, bonds, T-bills, and gold



(a) Stocks (b) Bonds

(c) T-bills (d) Gold

Note: Returns have been calculated as explained in the appendix.

5.1 Univariate results

In this section we analyze the tail behavior of univariate distributions of a sample of asset returns as a preliminary step for detecting possible extreme co-exceedances across asset classes. We also analyze the squared return tails and interpret it as a proxy of "extreme" volatility.

Descriptive statistics for all daily asset returns and squared returns are reported in Table 1. Only stock returns clearly exhibit negative skewness.

Stocks and gold show the largest spread in their return distribution, which might be concluded from the maximum, minimum, and standard deviation measures. All series have excess kurtosis which indicates deviations from normality (fat tails).

Table 1: Descriptive statistics for daily US asset returns

Panel A: Returns

Asset	Mean	Median	Max	Min	Std.Dev.	Skewness	Kurtosis	Obs.
Stocks	0.0237	0.0000	9.67	-25.63	0.9578	-1.73	53.33	11327
Bonds	0.0261	0.0227	4.63	-3.59	0.4285	0.18	10.03	11327
T-bills	0.0213	0.0186	0.32	-0.23	0.0244	1.49	21.81	11327
Gold	0.0221	0.0000	12.50	-14.20	1.3016	0.30	16.02	8480

Panel B: Squared Returns

Asset	Mean	Median	Max	Min	Std.Dev.	Skewness	Kurtosis	Obs.
Stocks	0.00918	0.002260	6.570	0.0	0.06630	86.44	8497.66	11327
Bonds	0.00184	0.000290	0.210	0.0	0.00550	13.06	325.76	11327
T-bills	0.00001	0.000004	0.001	0.0	0.00003	12.17	240.79	11327
Gold	0.01690	0.002460	2.010	0.0	0.06570	13.40	268.82	8480

Note: The data are daily from the beginning of 1962 until the end of 2005. The observations for gold start in 1973.

5.1.1 Univariate extreme value analysis

Table 2 summarizes the magnitude and timing of the two most extreme in-sample events together with the tail index and the percentile estimates based on equations (6) and (7), respectively. Panel A contains the results for the returns whereas the squared return results are reported in Panel B. Within Panel A we further distinguish between the left and right tail of the unconditional return distributions in order to account for possible asymmetries. Panel A shows that extreme losses and gains for stocks and gold are generally much higher than for bonds and T-bills. Even excluding the most extreme stock returns in October 1987 would not change this result. Moreover, for stocks and gold the historical extremes point toward tail asymmetries. The extreme negative returns are much larger in absolute value that the respective positive returns. For bonds and T-bills this is not so clear cut and tends to be the other way around.

Table 2: Minima, maxima, tail index, and univariate tail estimates for daily US asset returns

Panel A: Returns

	Left Tail						Right Tail					
					percentile						percentile	
Asset	Min 1 (%)	Min 2 (%)	opt. m	$\hat{\alpha}$	$\frac{1}{10.000}$	$\frac{1}{1.000}$	Max 1 (%)	Max 2 (%)	opt. m	$\hat{\alpha}$	$\frac{1}{10.000}$	$\frac{1}{1.000}$
Stocks	-25.63	-8.38	287	3.42	-9.24	-4.71	9.67	6.15	189	3.69	8.76	4.69
	10/19/87	10/26/87					10/21/87	07/24/02				
Bonds	-3.59	-2.74	158	3.84	-4.01	-2.20	4.63	3.90	145	3.68	4.37	2.34
	02/19/80	04/04/94					10/20/87	04/16/80				
T-bills	-0.23	-0.20	51	2.64	-0.27	-0.11	0.32	0.29	781	2.64	0.61	0.25
	05/04/81	10/09/79					12/19/80	01/05/81				
Gold	-14.20	-12.89	103	3.12	-16.22	-7.76	12.50	11.21	191	2.73	20.92	9.00
	01/22/80	02/28/83					01/03/80	11/03/76				

Panel B: Squared Returns

					percentile	
Asset	Max 1 (%)	Max 2 (%)	opt. m	$\hat{\alpha}$	$\frac{1}{10.000}$	$\frac{1}{1.000}$
Stocks	6.57	0.93	579	1.74	1.26	0.34
	10/19/87	10/21/87				
Bonds	0.2143	0.1525	92	2.42	0.1621	0.063
	10/20/87	04/16/80				
T-bills	0.0010	0.0008	56	2.67	0.0009	0.0004
	12/19/80	01/05/81				
Gold	2.0156	1.66	60	2.02	2.72	0.87
	01/22/80	02/28/83				

Panel C: Test for tail index equality

Test st.	P-value in %
-0.27	39.36
0.10	45.84
-0.00	50.07
0.31	37.81

Note: $\hat{\alpha}$ is the reciprocal of the Hill estimator in equation (7). The columns "percentiles" are the percentiles with marginal probabilities of $p = 1/10.000$ and $p = 1/1.000$, respectively. Max and Min 1 and 2 are the two most extreme positive and negative return observations in the sample, respectively. Panel A shows estimation results for the left and right tail. Panel B those for the squared returns as a proxy for volatility. Panel C gives test statistics and p-values for the test for equality of the left and right tail indexes as shown in (14).

A somewhat different picture emerges when we consider the estimated tail indices $\hat{\alpha}$. Stock returns seem to be asymmetric but the left tail index $\hat{\alpha} = 3.42$ only slightly falls below its right tail counterpart ($\hat{\alpha} = 3.69$). Bond and gold tail behavior suggest a fatter right tail. T-bills seem to exhibit symmetric tails. The left-tail index estimates are highest in the case of bond returns. Otherwise stated, long-term government-bond investments exhibit more limited downside risk than stocks. These results seem to confirm earlier research by e.g. Longin and Solnik (2001).

In Panel B we show the results for squared returns. Squared returns can be interpreted as a measure for assets' volatility. Engle (1982) pointed out that asset return volatility is likely to change over time but in a persistent manner. He developed a test for the so-called ARCH effect by choosing squared returns as a volatility proxy and regressing squared returns on lagged squared returns. It has been theoretically shown that there is a relation between volatility clustering and fat tails, see Koedijk and Schafgans (1973). Moreover, it can be shown that the squared returns should also be heavy tailed and that the probability mass in the tails of the return squares is even higher. Panel B reveals that the estimated tail indices for the squared returns are below the Panel A tail indices indeed.

The table also provide some casual evidence for cross asset linkages during crisis periods. The calendar dates of the extreme events, as recorded below the minima and maxima, suggest the presence of a 'flight-to-quality'effect from stocks to bonds after Black-Monday. Stocks crashed on 10/19/1987 and bonds boomed on 10/20/1987. Notice also that the US stock market showed a strong technical upward correction on 10/21/1987 partly offsetting the exaggerated slump from two days before. Another interesting observation is that from the twelve most extreme events in the case of bonds, T-bills, and gold eight fall in the years between 1979 and 1981 which probably reflects that extreme volatility was at its highest around the second oil crisis. Similar results hold for the squared returns (volatilities). The most volatile period for stocks and bonds was in 1987. As for bonds, T-bills, and gold four out of the six most volatile days were in the period between 1979 and 1981.

The economic issue of interest, both for the general assessment of financial market stability and for financial investors' and institutions' risk management, is the likelihood and size of extreme returns as reflected by the tail probabilities and corresponding percentiles. The percentiles reflect possible extreme events or scenarios whose expected waiting time to occur equals the inverse of the corresponding marginal excess probabilities p. For example, a daily meltdown in the Dow Jones Industrial Average of -4.70% or more is expected to happen only once every 1,000 days or 3.9 years. So, the reported values can be interpreted as value-at-risk (VaR) estimates for given marginal significance levels p.

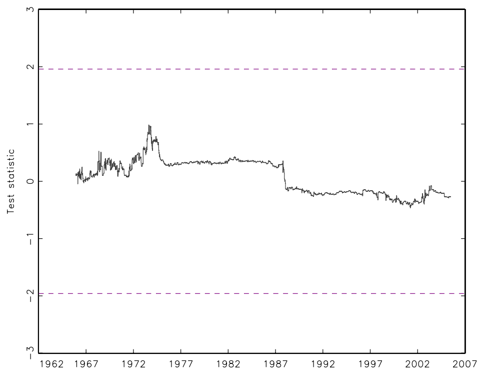
The question remains whether the observed differences in tail index point estimates are statistically significant across tails. In order to test the null hypothesis of equal tail indices, we report corresponding test statistics for the tail asymmetry test and p-values (Panel C of the same Table 2). Additionally we show the recursive test version over the sample in Figure 2 for all four assets. Both the (full sample) test statistics in the table and the recursive statistics in the figure show that none of the assets exhibits significant tail asymmetry.

5.1.2 Stability of EVT estimates

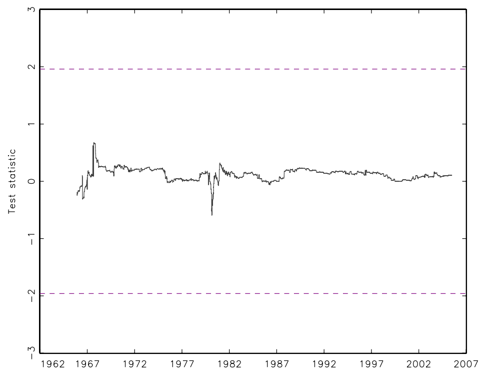
In order to check for structural breaks in the tail behavior of the unconditional return distribution we utilize a test developed by Quintos et al. (2001) as described in Section 4. Results of the test are summarized in Table 3. The table is again split into left and right tail results (except for the squared returns). In the most right part of the table we report the results for our volatility measure. Panel A states the results for the recursive test which checks for an increase in the thickness of the respective tail so a decrease in the tail index α. In Panel B one can see the results for the reverse recursive test testing for the opposite. For every asset and tail[13] we report the test statistic, the bootstrap simulated critical value and the date of the break if the test is significant at least at the 5% level. Asterisks indicate the significance level of the test statistic. The only cases where the null hypothesis of tail index

[13] In the case of squared returns obviously only the right tail is being considered.

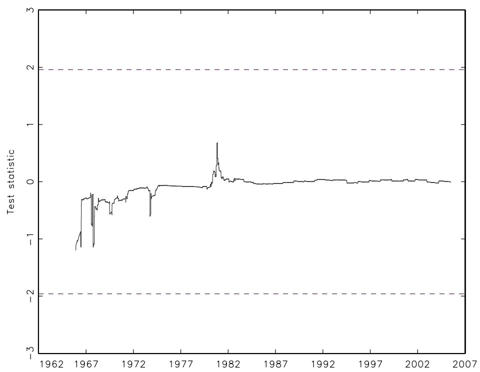
Figure 2: Recursive cross-section test: Univariate



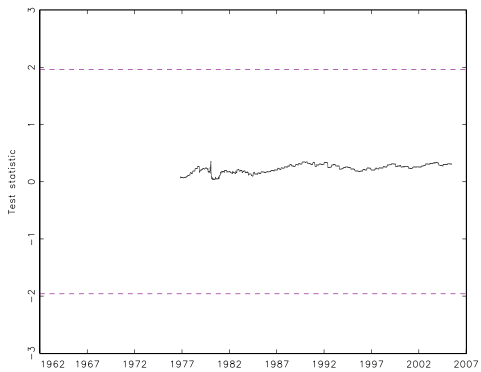
(a) Stocks



(b) Bonds



(c) T-bills



(d) Gold

Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels.

stability can be rejected at the 1% level involve the right tail of bonds and T-bills and the squared returns of bonds. Other identified breaks are significant at the 5% but not at the 1% level. Therefore, evidence for a structural break in the cases of the left tail for stocks and T-bills, the right tail for gold and the squared returns for T-bills and gold is much less convincing. In the case of bonds and T-bills one can clearly see that there appears to be a break in both series in the beginning of the 80s in the recursive test for the right tail. Later, in the mid and end 80s both right tails of the return distributions are detected to show again a break but now in the reverse recursive test. An obvious interpretation could be that Paul Volcker's structural change in the Fed's monetary policy and the second oil crisis had a major impact on the return behavior of bonds and T-bills but much less on stocks and gold. Volcker's shift from targeting the interest rates to rather limiting the growth rate of money supply had a strong influence on obligations' returns but also on their volatility. Thus, the turbulent times from the beginning of the 80s until black Monday in October 1987 are more strongly reflected in the return behavior of T-bills and bonds than in stocks and gold. Already a visual inspection of the return series in Figure 1 supports this result. As one can see, in the case of bonds the period starting in the early 80s until the end of the 80s shows stronger variability than the rest of the sample. For T-bills this unusual period is shorter, which is also reflected in the statistical test results. Somewhat surprisingly the test results only identify significant breaks for the right but not the left tail of the bonds and T-bills return distributions. One might attribute this to a relatively low power of the stability test.

Table 3: Univariate results for stability test for daily US asset returns and squared returns

Panel A: Recursive Test

Asset	Left Tail			Right Tail			Squared Returns		
	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date
Stocks	6.12*	4.59	04/14/1986	1.39	3.58	-	1.74	5.15	-
Bonds	2.44	2.93	-	4.79**	2.11	02/18/81	4.08**	2.14	06/05/80
T-bills	1.20	2.24	-	24.83**	19.36	09/22/82	2.93*	2.87	05/19/80
Gold	0.53	3.68	-	0.27	2.12	-	0.21	2.77	-

Panel B: Reverse Recursive Test

Asset	Left Tail			Right Tail			Squared Returns		
	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date	Test stat.	BT cr.val.	Break date
Stocks	1.42	3.90	-	1.49	3.90	-	1.89	5.13	-
Bonds	1.63	3.02	-	3.43**	1.93	06/05/89	1.17	2.12	-
T-bills	3.43*	2.62	09/29/80	102.40**	19.36	07/08/85	4.18*	2.79	01/27/82
Gold	1.92	4.83	-	2.87*	2.35	06/14/00	3.24*	2.33	12/27/79

Note: Test statistics are based on equation (10). The bootstrap (BT) critical values were simulated as described in section 4 in the text. * and ** indicate rejection of the null hypothesis of tail index constancy at the 5% and 1% significance levels, respectively.

5.2 Bivariate results

In this section we examine the propensity for co-exceedances across US asset classes. We start with standard correlation analysis followed by an identification of asset linkages during crisis periods using bivariate extreme value analysis. The extreme linkages allow us to assess the potential for cross asset substitution effects during market stress (the likelihood of flight to quality, flight to liquidity etc.).

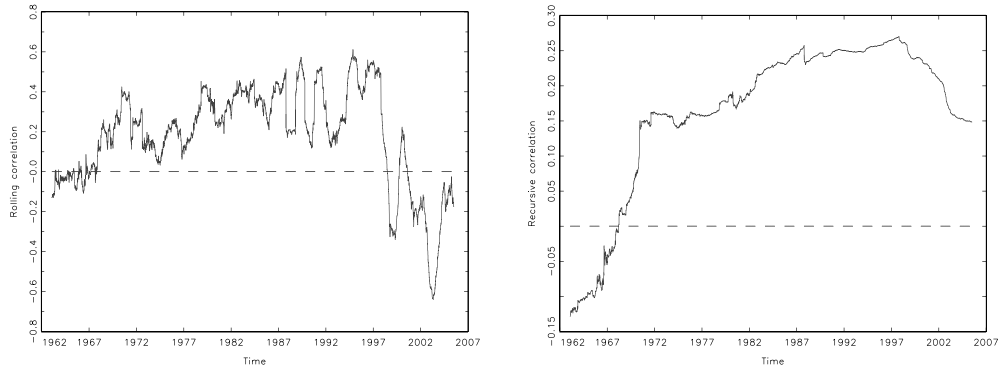
5.2.1 Correlation analysis

The results of the correlation analysis are summarized in Figures 3 to Figure 8. For each of the six possible bivariate asset combinations we calculate rolling and recursive correlations. The rolling correlation is a yearly correlation in the sense that it is calculated using a time window of 260 trading days, which corresponds to one trading year. The recursive measure gives the correlation between the returns from the beginning of the sample period until point t. The correlation plots can be used to get some preliminary evidence for possible substitution effects between asset classes during periods of market stress.

Correlations between stocks and bonds (Figure 3) or stocks and T-bills (Figure 4) are slightly more often positive than negative which does not seem to provide much evidence for substitution effects like flight to quality or flight to liquidity. However, the rolling correlations become negative around some crisis periods. For example, stock-bond (rolling) correlations turn negative after the Asian crisis and the negativity aggravates after the dotcom bubble burst. Stock-T-bill (rolling) correlations have similar signs around the same periods and also turn negative in the aftermath of the 1987 stock market crash. Thus, the rolling correlations provide some evidence of substitution from stocks into bonds or T-bills during crisis periods.

Correlations between stocks, bonds and gold more strongly point towards gold as a safe haven during times of market stress. First, the stock-gold (Figure 5) and bond-gold (Figure 7) rolling correlations tend to be more often negative than positive (both over crisis and noncrisis periods). Second,

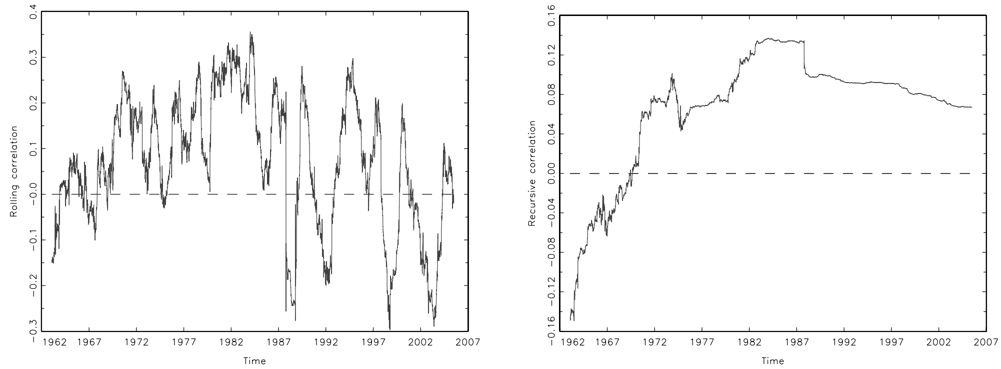
Figure 3: Stocks and bonds: Rolling and recursive correlation



(a) Rolling (b) Recursive

Note: Rolling correlations have been calculated with a window length of 260 trading days.

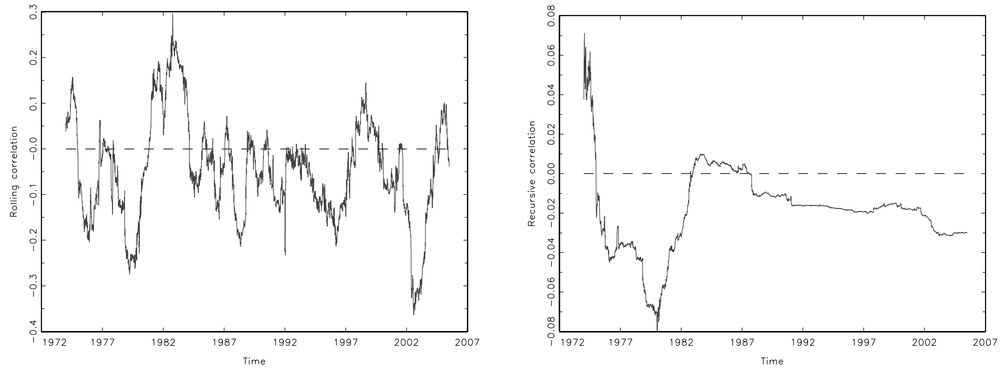
Figure 4: Stocks and T-bills: Rolling and recursive correlation



(a) Rolling (b) Recursive

Note: Rolling correlations have been calculated with a window length of 260 trading days.

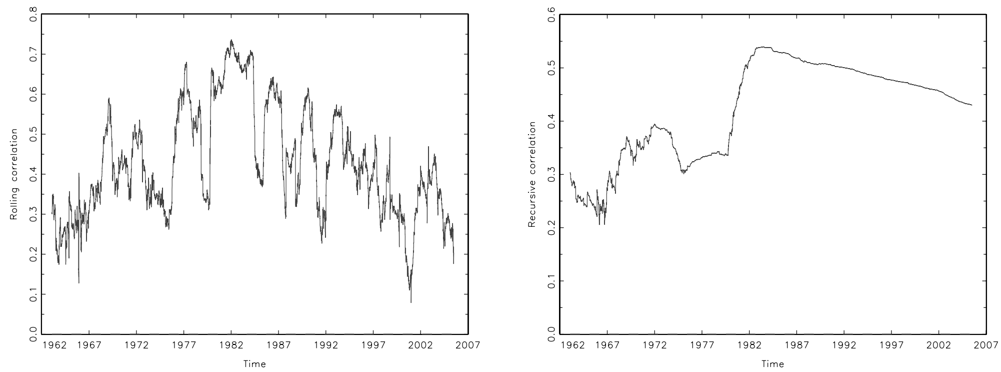
Figure 5: Stocks and gold: Rolling and recursive correlation



(a) Rolling (b) Recursive

Note: Rolling correlations have been calculated with a window length of 260 trading days.

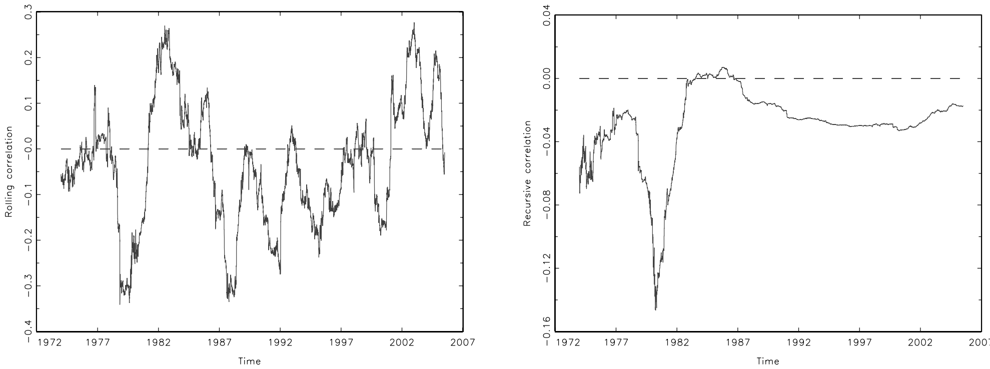
Figure 6: Bonds and T-bills: Rolling and recursive correlation



(a) Rolling (b) Recursive

Note: Rolling correlations have been calculated with a window length of 260 trading days.

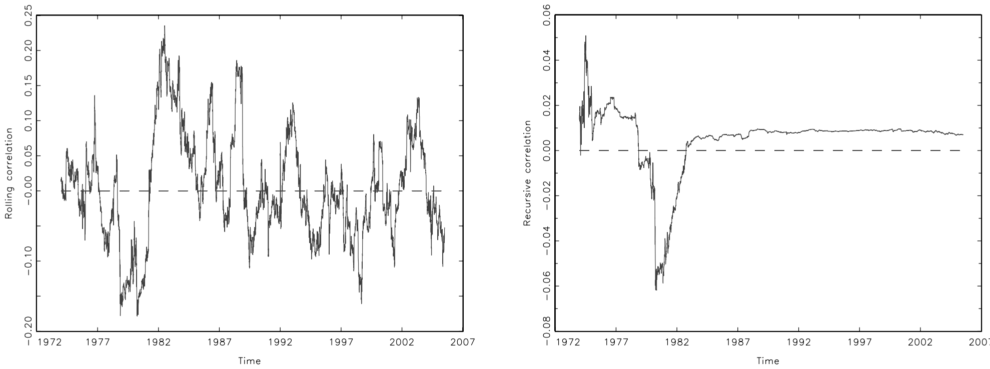
Figure 7: Bonds and gold: Rolling and recursive correlation



(a) Rolling (b) Recursive

Note: Rolling correlations have been calculated with a window length of 260 trading days.

Figure 8: T-bills and gold: Rolling and recursive correlation



(a) Rolling (b) Recursive

Note: Rolling correlations have been calculated with a window length of 260 trading days.

the negative correlations seem particularly present during the oil shocks in the 70s, the stock market crash of 1987, the 1997 Asian crisis and the 2000-2001 dotcom bubble burst.

Unsurprisingly, the correlations between bond and T-bill returns (Figure 6) are positive due to their linkage via the term structure. What the plots show is that this linkage can become weaker or stronger over time but it never becomes negative.

Finally, the correlations between T-bills and gold are not indicative of any substitution effects. This should not surprise given that both assets are considered as interesting investment objects in times of distress.

There are numerous problems, however, with the use of correlations as dependence measures. First, they typically measure linear dependencies whereas it is often suggested that linkages during stress periods might be nonlinear phenomena. Otherwise stated, correlation figures not necessarily give a good indication for co-dependence of the extremes of the marginal return distributions. We therefore decided to apply a alternative framework that only exploits information for the bivariate tail.

5.2.2 Bivariate extreme value analysis

Table 4 reports estimates for the conditional probability measure (1) and the conditional expectation measure (2) for all possible asset pairs in our sample assuming stationarity. The extreme measures are conditioned on different marginal excess probabilities p allowing us to evaluate the extreme dependence measures for different crisis levels.[14] Bivariate measures also allow us to compare the propensity towards co-crashes across assets with that towards substituting for a potentially safer asset (flight to quality or flight to liquidity effects).

The table reports the values for the tail index calculated in Equation (7), optimal amount of extremes m, conditional probabilities and expected values for occurrence of the mentioned co-exceedances corresponding to equation (1) and (2), respectively.[15]

[14] The lower the value of p the further we look into the bivariate tail.

[15] The optimal m is determined by means of the Beirlant et al. (1999) algorithm.

Interpretation of the conditional probabilities is straightforward. For example the entry 2.33 for stock-bond co-crashes means that there is a 2.33% chance of a sharp joint drop in stock and bond values. "Sharp" in this context means that the crash levels correspond to 0.1% VaR for stock and bond tails. In Table 2 we saw that the univariate percentiles, on which we condition for calculating the stock-bond cocrash probability, correspond to -4.71% for stocks and -2.20% for bonds. The reader might be tempted to interpret the potential for stock and bond co-crashes as small. However, if the extreme events were independent, we would expect a conditional probability of around 0.1%. So, conditioning on a crash in one market, increases the probability that the other one also collapses by a factor 23. The co-exceedance probabilities in Table 4 all exceed the benchmark level of 0.1% implying that there is significant tail dependence. Stated otherwise, the probability of having an extreme gain or loss in one asset category suddenly becomes much higher once another "domino stone" has fallen.

By further inspecting the table one can see that on average gold is less linked to the other assets during extreme events. The conditional probabilities are on average lower than in the other asset co-event cases. As such, gold looks as a reasonable hedge against all other asset classes considered.

When we have a look at the estimates for stocks and bonds, a co-crash is more likely than flight to quality from stocks into bonds. For stocks and T-bills we have lower estimates for the conditional probabilities and expected values, with flight to liquidity being close to the independence case. As for stock-T-bill co-crashes the probabilities are higher but still lower than for stock-bond co-crashes. So, capital leaving the stock market does not seem to cause a run on T-bills, i.e., the flight to liquidity hypothesis. Bonds and T-bills strongly co-move in the lower tails.

The bivariate results show that multivariate return distributions are not symmetric.

Notice that extreme event linkages can strongly differ from traditional dependence measures like correlation which is based on the full distributional support. Otherwise stated, conclusions drawn from a simple full sample correlation or even a rolling correlation analysis can be misleading for the

Table 4: Bivariate results: Full sample

Panel A1: Stocks and Bonds, Stocks and T-bills, Stocks and Gold

	Stocks and Bonds		Stocks and T-bills		Stocks and Gold	
	Co-crash	FtQ S/B	Co-crash	FtL S/TB	Co-crash	FtQ S/G
Tail Index	1.45	1.622	1.52	1.84	1.82	1.79
Optimal m	290	404	397	424	371	437
Con.Pro. in %						
$p = 0.1$	2.33	0.68	1.54	0.23	0.29	0.37
$p = 0.05$	1.68	0.44	1.07	0.13	0.16	0.22
$p = 0.01$	0.79	0.16	0.47	0.03	0.04	0.06
E-Values						
$p = 0.1$	1.0118	1.0034	1.0082	1.0011	1.0014	1.0019
$p = 0.05$	1.0085	1.0022	1.0051	1.0009	1.0008	1.0011
$p = 0.01$	1.004	1.0008	1.0022	1.0004	1.0002	1.0003

Panel A2: Bonds and T-bills, Bonds and Gold, Gold and T-bills

	Bonds and T-bills		Bonds and Gold		Gold and T-bills	
	Co-crash	FtL B/TB	Co-crash	FtQ B/G	Co-crash	FtL G/TB
Tail Index	1.18	1.97	1.58	1.33	1.59	1.48
Optimal m	397	477	296	121	420	499
Con.Pro. in %						
$p = 0.1$	14.61	0.07	0.89	2.89	0.92	2.08
$p = 0.05$	12.83	0.03	0.61	2.31	0.61	1.51
$p = 0.01$	9.60	0.00	0.24	1.35	0.23	0.72
E-Values						
$p = 0.1$	1.0792	1.0003	1.0045	1.0146	1.0046	1.0105
$p = 0.05$	1.0691	1.0002	1.0031	1.0116	1.0031	1.0075
$p = 0.01$	1.0523	1.0000	1.0012	1.0068	1.0012	1.0035

Note: FtQ S/B stands for "flight to quality" from stocks into bonds, for example, as defined in the text. So, S stands for stocks, B for bonds, TB for T-bills and G for gold. E-values stands for the expected amount of co-events conditioned on an extreme percentile given by p.

occurrence probability of extreme co-events. In order to show this, we can again refer to the correlation Figures 3 to 8. The full sample correlation is equal to the last observation of the recursive correlation. In the case of bonds and T-bills the full sample correlation is equal to 43.02%. One can interpret such a correlation as a strong linkage, but this can be rather illusory. If one uses the bivariate normal distribution in order to asses the extreme bonds-T-bills market linkage applying the sample variances and correlation, one would find a conditional co-crash probability of around 0.000125% for the marginal distribution percentile $p = 1/1000$. Using the more accurate EVT estimation of tail behavior we find for the same marginal percentile a conditional co-crash probability of 14.61%. Hence, correlations together with the multivariate normality assumption strongly understate extreme financial market linkages and thereby should not be used for extreme dependence estimation.

5.2.3 Stability of EVT estimates

Here we relax the stationarity assumption of the dependency measures presented in the section before. Results are summarized in Table 5. Panel A1 and B1 show the results for the recursive test (alternative hypothesis = increase in tail dependence) and Panel A2 and B2 give the results for the backward recursive test (alternative hypothesis = decrease in tail dependence).

In Table 5 a clear picture emerges. The recursive test tends to find breaks in the early part of the sample ranging from 1968 until 1980 depending on the asset combination. The only exceptions are the asset combinations stocks with gold and gold with T-bills. A likely explanation is that the gold time series only starts in 1973 so that the break testing procedure only starts in the year 1976. But at that point the break might already have occurred such that the test is unable to pick it up.

Another interesting observation is that in all but one asset combination (stocks and bonds) we find significant breaks for the reverse recursive test all happening after the respective breakpoints for the recursive test. The breaks

Table 5: Bivariate results: Stability test

Panel A1: Recursive test

	Stocks and Bonds		Stocks and T-bills		Stocks and Gold	
	Co-crash	FtQ S/B	Co-crash	FtL S/TB	Co-crash	FtQ S/G
1969	22.3**		28.4**			
	09/25		06/10			
1973				25.35**		
				01/26		
1979		46.0**				
		02/22				

Panel A2: Backward recursive test

	Stocks and Bonds		Stocks and T-bills		Stocks and Gold	
	Co-crash	FtQ S/B	Co-crash	FtL S/TB	Co-crash	FtQ S/G
1983						8.5*
						03/11
1988				13.3**	10.0**	
				05/24/88	02/04/88	
1989			42.7**			
			04/13/89			

Panel B1: Recursive test

	Bonds and T-bills		Bonds and Gold		Gold and T-bills	
	Co-crash	FtL B/TB	Co-crash	FtQ B/G	Co-crash	FtL G/TB
1968	38.97**					
	11/29					
1973		55.5**				
		12/05				
1979				8.8**		
				09/28		
1980			17.1**			
			01/16			

Panel B2: Backward recursive test

	Bonds and T-bills		Bonds and Gold		Gold and T-bills	
	Co-crash	FtL B/TB	Co-crash	FtQ B/G	Co-crash	FtL G/TB
1983	54.6**				50.2**	
	10/12				11/24	
1986			42.9**			
			10/15			
1987				46.0**		
				12/29		
1991		45.8**				408.8**
		06/07				05/30

Note: FtQ S/B stands for "flight to quality" from stocks into bonds, for example, as defined in the text. So, S stands for stocks, B for bonds, TB for T-bills and G for gold. * and ** indicate rejection of the null hypothesis of tail index constancy at the 5% and 1% significance levels, respectively.

in the reverse recursive procedure range from 1983 until 1991 and most of them cluster between 1983 and 1988. So, considered market comovements tend to become stronger in the seventies and again weaker in the eighties. In most of the cases where the recursive test detects a break, the tail index becomes smaller either shortly before the first oil shock in the beginning of the 70s or before the second oil shock and the beginning of Volcker's term at the end of the 70s. In sum, the economically and politically turbulent times surrounding the oil shocks, the Volcker presidency of the FED and extreme market volatility around Black Monday in October 1987 seem to coincide with breaks in the degree of tail dependence. Moreover, most of the asset co-event cases show a tail index behavior that might be described as a U-shape or mean-reverting. As such, the full sample results in Table 4 represent an average across time. Nevertheless, those calculated conditional probabilities provide a good approximation of the true co-dependencies of assets because of the observed mean reverting behavior of the tail indices. A further interesting step could be to split the sample at the observed break points and estimate co-dependencies within every sub-sample. A problem of such an approach, though, constitutes the exact location of the break in the bivariate distributions, because the breaks across co-boom, co-crash, and flight to quality co-events do not always occur at the same point in time. In order to account for this, some kind of common break point estimation is needed but this is beyond the aim of this paper.

Table 6 shows that the results for the squared returns are pretty similar to those of the ordinary returns. Stocks and bonds and especially bonds and T-bills show the highest conditional probability of common extreme volatilities. In the case of stocks and bonds, for example, we estimate a conditional probability of 5.64% that bonds show a volatility among their 0.1% largest ones, given that also stocks' volatility was among their 0.1% largest daily observed volatilities. This constitutes a probability increase by a factor 56 compared to the independence case. For the asset combination bonds and T-bills the same estimated conditional probability even increases with a factor 232. For the other four possible bivariate asset combinations, estimated conditional probabilities for pairs of squared returns are much lower also supporting the results obtained before.

Table 6: Bivariate results: Squared returns

	S & B	S & TB	S & G	B & TB	B & G	G & TB
Tail Index	1.29	1.59	1.74	1.04	1.47	1.65
Optimal m	330	498	462	443	426	499
Con.Pro. in %						
$p = 0.1$	5.64	1.10	0.5	23.20	2.16	2.99
$p = 0.05$	4.61	0.73	0.31	22.52	1.56	2.19
$p = 0.01$	2.86	0.28	0.11	21.01	0.74	1.05
E-Values						
$p = 0.1$	1.029	1.006	1.003	1.131	1.011	1.015
$p = 0.05$	1.024	1.004	1.002	1.127	1.008	1.011
$p = 0.01$	1.015	1.001	1.001	1.117	1.004	1.005

Note: S stands for stocks, B for bonds, TB for T-bills and G for gold. E-values stands for the expected amount of co-events conditioned on an extreme percentile given by p.

The structural break point analysis for the squared returns in Table 7 is generally in line with the results of the ordinary returns. Probabilities of common high volatility tend to increase either before the first or the second oil crises and decrease again at the end of the 80s and the beginning of the 90s. So, on average volatilities' tail indexes tend to co-break with the ordinary returns.

A little word of caution might be appropriate here. In the transformation to the Pareto distribution we implicitly assume stationarity of the univariate return series' tail behavior. We know from Section 5.1.2 that this is not always the case. Nevertheless, we believe that following our approach we are able to distinguish the cases of having only a break in the marginal return distributions or having a break in dependence structure of the marginal distributions. This can also be confirmed by comparing the univariate and bivariate break dates, which do not coincide. If our approach was not able to distinguish both cases break dates would have to coincide. Future research on the theoretical foundations would be very interesting but goes beyond the scope of this paper.

5.2.4 Cross-sectional results

In this subsection we apply the same cross section test for comparing bivariate tail indices across assets as we did for comparison of the univariate indices

Table 7: Bivariate results: Stability test squared returns

Panel A1: Recursive test

	S & B	S & TB	S &	B & TB	B & G	G & TB
1969	22.93 10/08	41.84 02/19				
1974				76.76 07/19		
1978					20.89 08/10	

Panel A2: Backward recursive test

	S & B	S & TB	S &	B & TB	B & G	G & TB
1988			11.84 02/04		76.52 02/01	
1989				126.12 06/26		
1990		28.54 10/12				
1991						455.39 02/28

Note: S stands for stocks, B for bonds, TB for T-bills and G for gold. All breaks are found to be significant at a 5% significance level.

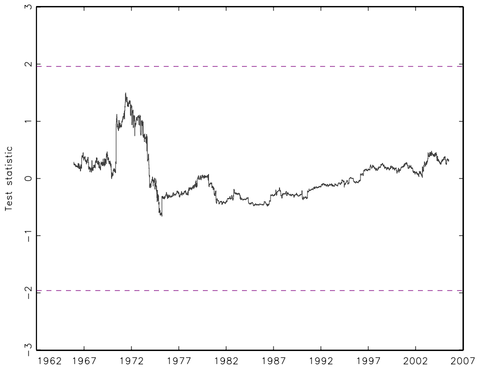
in Table 2 and in Figure 2. Here, we are interested if the co-exceedance tail indices (tail dependence parameters) differ across asset combinations and between co-crashes and/or flight to quality/liquidity. Smaller tail indices here mean that the corresponding co-extreme events are more likely to occur than one with a bigger tail index. In Figures 9 to 15 we show the recursive test statistics as in Equation (14). Figures 9 to 11 show all co-crash combinations, Figures 12 to 14 are for all flight to quality/liquidity combinations and Figure 15 gives the test for all matched cases of co-crashes and flight to quality/liquidity. The figures also show the upper and lower rejection regions at -1.96 and 1.96 corresponding the normal distribution 2.5% levels each. Negative test statistics indicate a smaller tail index for the first pair of assets compared to the second pair. Two examples make the logic clear. In Figure 9 (a) we have the case of the co-crash combination of stocks and bonds with stocks and T-bills. So, we actually compare two bivariate time series' tail indices. In this specific case there does not appear any significant difference between both tail indices through the full sample period. As a second example serves Figure 15 (a) which corresponds to the case where we compare the α of the bivariate series where bonds and T-bills crash together, against the case where we speak of flight to liquidity from bonds to T-bills. Here we see that for most of the time series the tail index for bonds and T-bills co-crashes is significantly smaller (positive test statistics) than the tail index for the flight to liquidity case from bonds into T-bills.

We can draw a couple of interesting conclusions from those figures. First, there are only a few significant differences between bivariate tail indices over the full sample period and all asset combinations. This confirms the general findings in the EVT literature that tail indices usually cannot be found to differ significantly.[16] Second, those cases where we do clearly find significant results always include the asset pair bonds and T-bills. Actually, this was to be expected. The pair bonds and T-bills shows clearly the strongest comoving behavior. Along with this, the probability that bonds crash and T-bills boom (flight to liquidity) at the same time is very unlikely. So, asset combinations including bonds and T-bills will have the tendency towards significant differences in the tail indices, which is what we find. Third, the cases involving bonds and T-bills also seem to be the most volatile in terms of movements of the test statistic. Although results have to be interpreted with caution, a possible explanation is that the bonds and T-bills returns are heavily influenced by changes in the monetary regime. Changes in the Fed's policy will directly move those securities' prices and thereby possibly lead to more changes in their comovement tail indices.

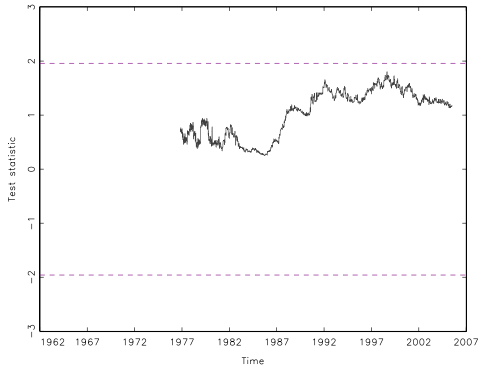
As one good example can serve Figure 10 (a) where we compare the tail indices for co-crashes between stocks and T-bills and bonds and T-bills. Here we see in the 1980s a period with a significantly smaller tail index for bonds and T-bills co-crashes than for stocks and T-bills co-crashes. Again this might be explained by the fact that Paul Volcker was the Fed's chairman and especially bonds markets were characterized by high volatility. In the year 1987 there is a strong change toward insignificance, probably caused by the Black Monday stock market crash and following volatility clearly having

[16] This confirms earlier findings like in Koedijk and Schafgans (1990), Jansen and de Vries (1991) Quintos et al. (2001).

Figure 9: Recursive cross-section test: Co-crashes I



(a) SB/STB

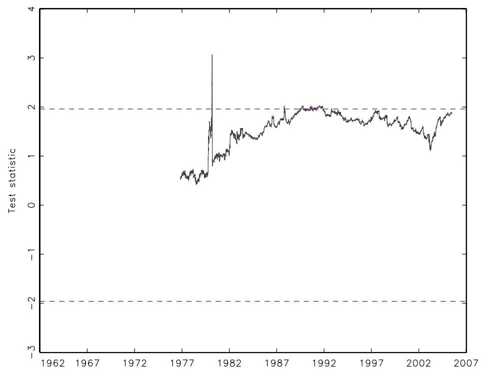


(b) SB/SG

(c) SB/BTB

(d) SB/BG

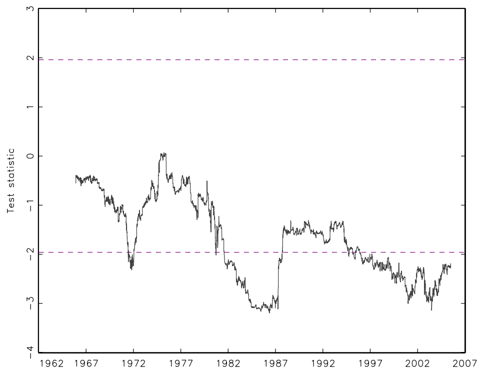
(e) SB/GTB



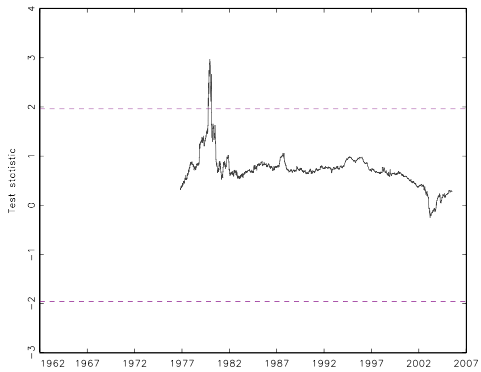
(f) STB/SG

Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 10: Recursive cross-section test: Co-crashes II



(a) STB/BTB



(b) STB/BG

(c) STB/GTB

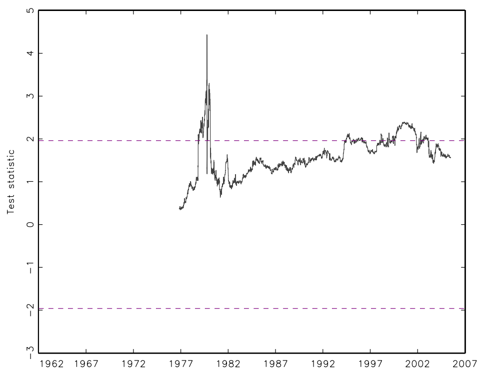
(d) SG/BTB

(e) SG/BG

(f) SG/GTB

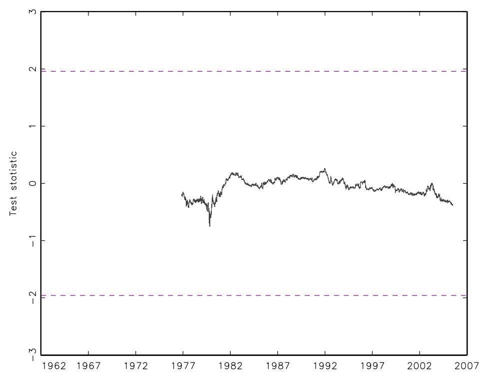
Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 11: Recursive cross-section test: Co-crashes III



(a) BTB/BG

(b) BTB/GTB



(c) BG/GTB

Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

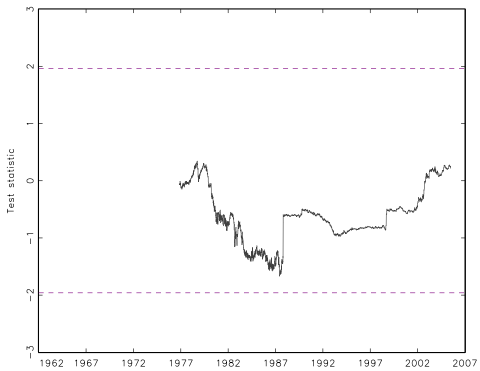
Figure 12: Recursive cross-section test: Flight to quality and liquidity I

(a) SB/STB

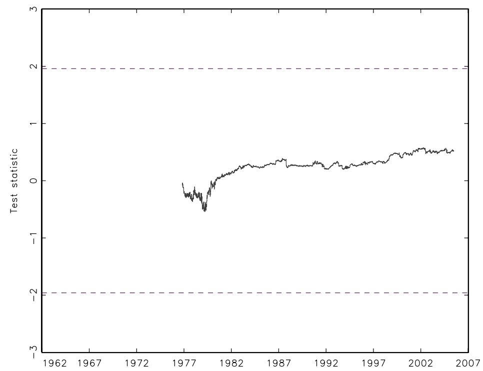
(b) SB/SG

(c) SB/BTB

(d) SB/BG



(e) SB/GTB



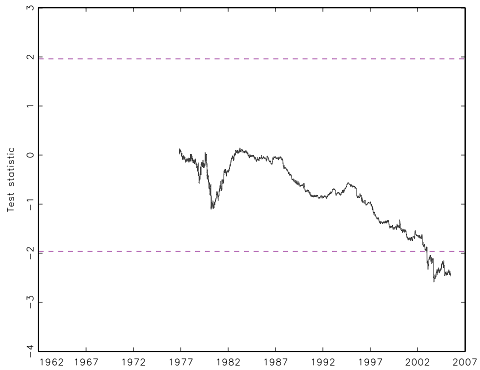
(f) SG/STB

Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

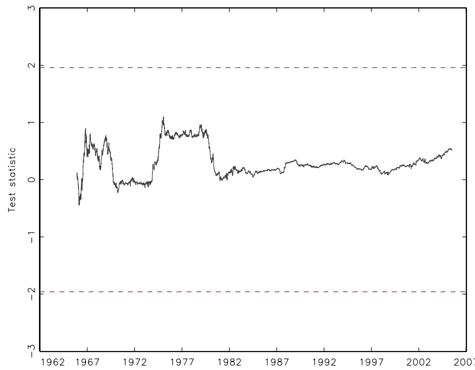
Figure 13: Recursive cross-section test: Flight to quality and liquidity II

(a) SG/BTB

(b) SG/BG

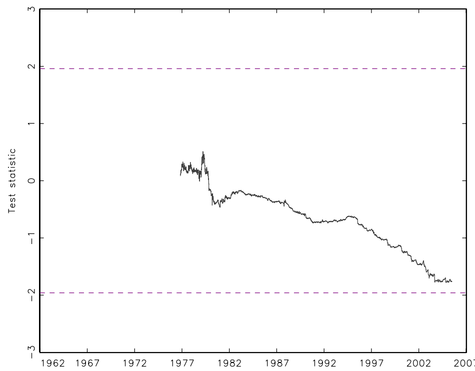


(c) SG/GTB



(d) STB/BTB

(e) STB/BG

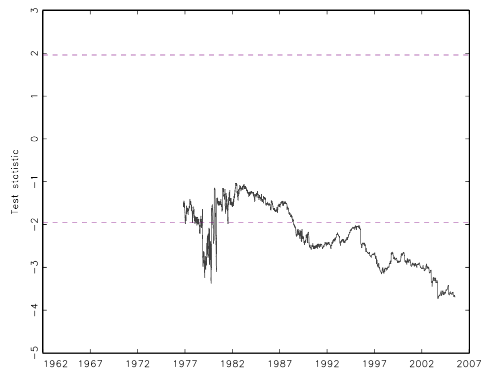


(f) STB/GTB

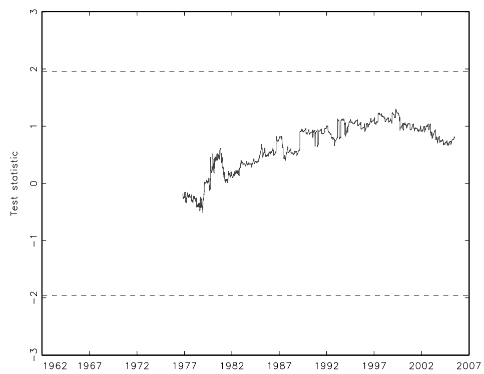
Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 14: Recursive cross-section test: Flight to quality and liquidity III

(a) BTB/BG



(b) BTB/GTB

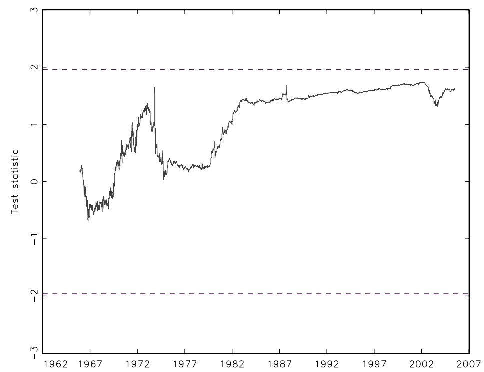


(c) BG/GTB

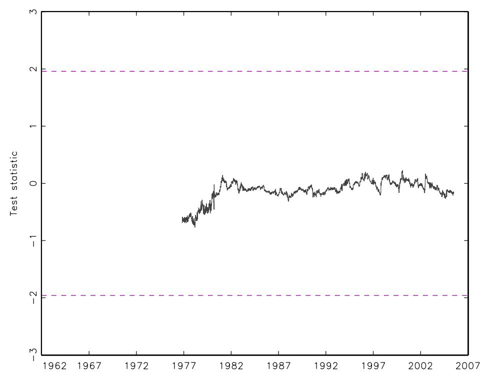
Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

Figure 15: Recursive cross-section test: Co-crashes and flight to quality/liquidity

(a) SB/SB



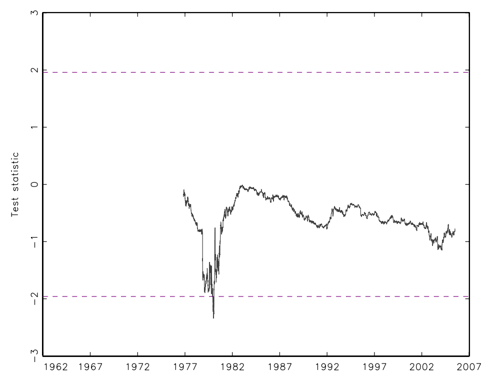
(b) STB/STB



(c) SG/SG

(d) BTB/BTB

(e) BG/BG



(f) GTB/GTB

Note: Test statistics have been calculated using Eq. 14 in a recursive way. Horizontal lines indicate the 2.5% and 97.5% significance levels. The abbreviations are: S stands for stocks, B for bonds, TB for T-bills and G for gold.

an impact on the lower tail index of stocks in that period.

6 Conclusions

In this paper we study the linkages between four different US asset classes (US stocks, government bonds, T-bills, and gold) in times of market turbulence. The linkages were characterized by their asymptotic tail dependence. Studying the likelihood of co-exceedances across asset classes is interesting for investors who have to choose their investment portfolios according to their preferred risk-return combinations. On the other hand, policy makers and supervisory bodies are interested in these linkages because they potentially influence the level of systemic risk in financial markets.

We use a non-parametric multivariate measure to identify the tail dependence of the marginal return distributions and derive estimates for the expected conditional probabilities of return co-exceedances. Such an approach does not rely on a particular probability law for the marginal return distributions and thereby has advantages over the usual conditional correlation measures because wrong parametric assumptions can easily distort extreme-spillover probabilities. We also tested for stability of these linkage measures through time in order to see if there are any periods with stronger or weaker spillover effects.

A preliminary (rolling and recursive) correlation analysis showed similar positive correlations between stocks and bonds as well as stocks and T-bills over time. During periods of market turbulence (oil crises in the 1970s and 1980s, the stock market crash in October 1987 and the terrorist attacks in New York in 2001), correlations typically dropped to below-average values. As expected, T-bills and bonds show a very strong correlation (but decreasing since the second oil shock). Gold turns out to be only mildly correlated with the rest of the assets in the sample set. These time varying correlations already give a first indication that there are possible flight-to-quality and flight-to-liquidity effects (especially during periods of financial uncertainty).

Structural break tests applied on the univariate and multivariate extreme value measures show very similar results. Both cases reveal breaks from thin to fatter tails either before the first or the second oil shock. Shifts back from fat to thinner tails occur during the 1980s. This indicates that asset returns (especially stocks, bonds, and T-bills) showed an increased probability of co-exceedances in the period between the mid 70s and mid 80s. For the squared returns, results are rather similar. Stocks and bonds, and especially bonds and T-bills show the highest conditional probability of common extreme volatilities. Again, breaks seem to occur around economically and politically turbulent times like the two oil shocks, the Volcker FED-presidency, or the stock market crash of 1987.

7 Appendix

We choose the Dow Jones Industrial Average as stock price index and extracted it from Datastream, Inc. 10-year government bond and 3-month T-bill returns were calculated from the corresponding yield to maturity data from the Board of Governors of the Federal Reserve Board.[17] We calculated returns from these yield data according to the methods used, for example, in Campbell et al. (1997, Chapter 10). Gold prices were extracted from www.usagold.com. The stock data are Financial Times Standard & Poors world price indices, whereas the bond data correspond to ten year ("all-traded") government bonds. In order to arrive at the returns for stocks and gold we calculated the first differences of the log of their prices levels. Data on stocks, T-bills, and bonds start on February 2, 1962 whereas gold starts on January 2, 1973. Last observations in the sample are on July 5, 2005. We did not include corporate bond indices, because of our particular interest in the flight to quality phenomenon. The stock and bond returns are not compensated for dividends and coupon payments, respectively.

[17] http:\\www.federalreserve.gov.

References

ANDREWS, D.W.K. (1993), 'Tests for parameter stability and structural change with unknown change point', *Econometrica*, 59, pp. 817–858.

BEIRLANT, J., DIERCKX, G., GOEGEBEUR, Y., AND MATTHYS, G. (1999), 'Tail index estimation and the exponential regression model', *Extremes*, 2(2), pp. 177–200.

CAMPBELL, J.Y., LO, A.W., AND MACKINLAY, A.C. (1997), *The econometrics of financial markets*, Princeton: Princeton University Press.

DE HAAN, L., JANSEN, D.W., KOEDIJK, K., AND DE VRIES, C.G. (1994), 'Safety first portfolio selection, extreme value theory and long run asset risks', in: J. Galombos (Editor), 'Extreme Value Theory and Applications', Dordrecht: Kluwer Academic Publishers, pp. 471–487.

DRAISMA, G., DREES, H., FERREIRA, A., AND DE HAAN, L. (2001), 'Tail dependence in independence', Report 2001-014, Eurandom, Institute for Statistics, Probability, Operations Research, and their Applications.

DREES, H. (2003), 'Extreme quantile estimation for dependent data with applications to finance', *Bernoulli*, 9(4), pp. 617–657.

EICHENGREEN, B., ROSE, A.K., AND WYPLOSZ, C. (1996), 'Contagious currency crises', Working Paper 5681, National Bureau of Economic Research.

EMBRECHTS, P., KLÜPPELBERG, C., AND MIKOSCH, T. (1997), *Modelling extremal events*, Berlin: Springer.

ENGLE, R.F. (1982), 'Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation', *Econometrica*, 50, pp. 987–1007.

ENGLE, R.F. AND KRONER, K.F. (1995), 'Multivariate simultaneous GARCH', *Econometric Theory*, 11, pp. 122–150.

Engle, R.F. and Susmel, R. (1993), 'Common volatility in international equity markets', *Journal of Business and Economic Statistics*, 11(2), pp. 167–176.

Goldie, C. and Smith, R. (1987), 'Slow variation with remainder: Theory and applications', *Quarterly Journal of Mathematics*, 38, pp. 45–71.

Hall, P. (1982), 'On some simple estimates of an exponent of regular variation', *Journal of the Royal Statistical Society B*, 44(1), pp. 37–42.

Hamao, Y., Masulis, R.W., and Ng, V. (1990), 'Correlations in price changes and volatilty across international stock markets', *Review of Financial Studies*, 3, pp. 281–307.

Hartmann, P., Straetmans, S., and de Vries, C. (2004), 'Asset market linkages in crisis periods', *Review of Economics and Statistics*, 86(1), pp. 313–326.

Hartmann, P., Straetmans, S., and de Vries, C. (2005), 'Banking system stability: A cross-Atlantic perspective', NBER Working Paper.

Hill, B. (1975), 'A simple general approach to inference about the tail of a distribution', *The Annals of Statistics*, 3(5), pp. 1163–1173.

Hsing, T. (1991), 'Extremal index estimation for a weekly dependent stationary sequence', *Annals of Statistics*, 21(4), pp. 1547–1569.

Jansen, D. and de Vries, C. (1991), 'On the frequency of large stock returns: Putting booms and busts into perspective', *Review of Economics and Statistics*, 73, pp. 19–24.

Kaminski, G. and Reinhart, C. (2000), 'On crises, contagion and confusion', *Journal of International Economics*, 51(1), pp. 145–168.

King, M. and Wadhwani, S. (1990), 'Transmission of volatility between stock markets', *Review of Financial Studies*, 3, pp. 5–33.

Koedijk, K. and Schafgans, M. (1973), 'Random difference equations and renewal theory for products of random matrices', *Acta Mathematica*, 131, pp. 207–248.

Koedijk, K. and Schafgans, M. (1990), 'The tail index of exchange rate returns', *Journal of International Economics*, 29, pp. 93–108.

Ledford, A. and Tawn, J. (1996), 'Statistics for near independence in multivariate extreme values', *Biometrika*, 83(1), pp. 169–187.

Lin, W., Engle, R.F., and Ito, T. (1994), 'Do bulls and bears move across borders? International transmission of stock returns and volatility', *Review of Financial Studies*, 7(3), pp. 507–538.

Longin, F. and Solnik, B. (2001), 'Extreme correlation of international equity markets', *Journal of Finance*, 56, pp. 649–676.

Mallaris, A.G. and Urrutia, J.L. (1992), 'The international crash of October 1987: Causality tests', *Journal of Financial and Quantitative Analysis*, 27, pp. 353–363.

Morgenstern, D. (1956), 'Einfache Beispiele zweidimensionaler Verteilungen', *Mitteilungsblatt fur Mathematische Statistik*, 8, pp. 234–235.

Poon, S.H., Rockinger, M., and Tawn, J. (2004), 'Extreme value dependence in financial markets: Diagnostics, models, and financial implications', *Review of Financial Studies*, 17(2), pp. 581–610.

Quandt, R. (1960), 'Test of the hypothesis that a linear regression system obeys two separate regimes', *Journal of the American Statistical Association*, 55, pp. 324–330.

Quintos, C., Fan, Z., and Phillips, P. (2001), 'Structural change tests in tail behaviour and the Asian crisis', *Review of Financial Studies*, 68, pp. 633–663.

Resnick, S. and Stărică, C. (1998), 'Tail index estimation for dependent data', *The Annals of Applied Probabilty*, 8(4), pp. 1156–1183.

Sachs, J., Tornell, A., and Velasco, A. (1996), 'Financial crisis in emerging markets: the lessons from 1995', *Brookings Papers on Economic Activity*, 1, pp. 147–198.

Straetmans, S., Verschoor, W., and Wolff, C. (2006), 'Extreme US stock market fluctuations in the wake of 9/11', *Journal of Applied Econometrics*, forthcoming.