Extreme Co-movements and Extreme Impacts in High Frequency Data in Finance

Zhengjun Zhang and Kazuhiko Shinki

Department of Statistics, University of Wisconsin,
Madison, WI 53706-1532, USA

September 17, 2006

Abstract

Extreme co-movement and extreme impact problems are inherently stochastic control problems, since they will influence the decision taken today and ultimately influence a decision taken in the future. Extreme co-movements among financial assets have been reported in the literature. However, extreme impacts have not been carefully studied yet. In this paper, we use the newly developed methodology to further explore extreme co-movements and extreme impacts in financial market. Particularly, two FX spot rates are studied. Based on the results of our analysis with FX returns, we conclude that there exist extreme co-movements and extreme impacts in FX returns and care has to be taken when we employ portfolio optimization models, especially models without the ability of handling extreme dependencies.

Keywords: high frequency data, exchange rate, extreme dependence.
1 Introduction

In risk management, the extreme quantiles of a single risk factor, or the conditional extreme quantiles of multiple risk factors, are very important quantities. An accurate assessment of these quantities allows for effective handling of potential disasters such as stock market crashes. The difficulty of such a task varies. If the multivariate random variables used to model multiple risk factors are independent, we may only need to study the extreme quantiles of all variables separately. If they are not independent however, the estimation of the conditional extreme quantiles may be very complicated, especially when the joint distribution is not known. Meanwhile, if the multivariate random variables are nearly independent or asymptotically independent in the tails, we may only need to obtain the extreme quantiles for each variable; if they are not, we hope to have some multivariate extreme value distributional functions which can adequately model the extreme observations. Yet, the most important step before building a model is to identify whether the variables of interests are extremely dependent, also known as asymptotically dependent or tail dependent. In financial studies, particularly in this study, extreme co-movement variables and extreme impact variables are extreme dependent variables. They are defined next.

Extreme co-movements in financial market are characterized by the concurrence of different very large financial returns. When large observations have future impacts (for example, in the kth day or in the kth hour) – i.e., large values will be observed in the kth day or the kth hour if a large value has been observed, we say there are extreme impacts in the corresponding assets. Understanding extreme co-movements and extreme impacts among financial assets is very important in risk management and stochastic optimal control. In our data analysis, we found that there exist extreme dependencies in certain time series which were regarded as independent series from previous research results. As a result, previously reported results in the literature may be misleading, and corresponding stochastic models which do not consider extreme dependencies may be problematic. For example, as we introduce later, multivariate normal random variables can not model extreme dependent variables.
There have been considerable efforts on modeling extreme co-movements – for example, Longin and Solnik (2001), Breymann, Dias, and Embrechts (2003), Dias and Embrechts (2003), Zhang (2005), Zhang and Huang (2006), among others. However, results regarding how to test and characterize the extreme dependencies in high frequency data are still very sparse, and the future impacts (extreme impacts) of very large observations at present are barely seen in the literature and in practice. In this paper, we further explore the extreme dependencies in high frequency financial data, particularly two FX spot rates (US dollar against German Mark, US dollar against Japanese Yen) which have been analyzed by Breymann et al (2003), Dias and Embrechts (2003).

This paper aims to provide a methodology which tests extreme dependencies. The methodology can be used for various financial models in risk management and portfolio optimization. We argue that if a model does not consider extreme dependencies when variables are actually extremely dependent, it may overlook some risk factors and underestimate the risk. Hence, the methodology can help to measure the risk more accurately. Furthermore, since extreme dependencies are estimated separately for right and left tails in this paper, the methodology can be used to measure the risk under a certain condition such as financial crisis.

The paper is structured as follows. Section 2 presents definitions and theories. Section 3 presents two FX spot rate data and data transformation through a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model which describes volatilities and a generalized extreme value distribution fitting. In Section 4, we test extreme dependencies for daily data and hourly data. The final section concludes.

2 Extreme dependencies, extreme co-movements, and extreme impacts

For the theory of extreme independence between two random variables, there have been a series of studies in the extreme value literature. It has been shown that the bivariate maxima of a sequence of positively correlated normal random variables are extreme independent as long
as the bivariate normal random variables are not completely correlated – i.e., their correlation is not equal to 1 – (Watson 1954, Sibuya 1960). It has also been shown that the sum of a sequence of random variables and the maximum of the same sequence of random variables are extreme independent under some mild conditions. Some related results are documented in Grady (2000).

One can see that the study of extreme co-movements and extreme impacts is equivalent to the study of extreme dependencies between random variables. In the literature, the extreme dependency between two random variables is characterized by the so-called extreme dependence index.

Sibuya (1960) introduces the asymptotic independence between two bivariate random variables with identical marginal distribution. De Haan and Resnick (1977) extend it to the case of multivariate random variables. In this paper, however, we adopt the term of extreme (in)dependence in connection to the study of extreme co-movements and extreme impacts. The definition of extreme (in)dependence between two random variables is given next.

2.1 Definitions

**Definition 2.1** Two identically distributed random variables $X, Y$ are called extreme independent if

$$\lambda = \lim_{u \to x_F} \Pr(Y > u|X > u) = 0,$$

(2.1)

where $x_F = \sup\{x \in \mathbb{R} : \Pr(X \leq x) < 1\}$. If $\lambda$ exists, it is called the bivariate extreme dependence index which quantifies the amount of dependence of the bivariate upper tails. If $\lambda > 0$, then $(X, Y)$ is called extreme dependent and we say there are extreme co-movements between $X$ and $Y$.

The extreme dependence indexes of certain bivariate random variables are presented in Embrechts, McNeil, and Straumann (2002). Zhang (2003), and Ledford and Tawn (2003) extend the definition of the extreme dependence between two random variables to the lag-$k$ extreme dependence of a sequence of random variables with identical marginal distribution.
The definition of lag-\(k\) extreme dependence for a sequence of random variables is given below.

**Definition 2.2** A sequence of random variables \(\{\xi_i\}\) is called lag-\(k\) extreme dependent if for all \(i\)

\[
\lambda_k = \lim_{u \to x_F} \Pr(\xi_i > u | \xi_{k+i} > u) > 0, \quad \lim_{u \to x_F} \Pr(\xi_i > u | \xi_{i+k+j} > u) = 0, \quad j > 0
\]

(2.2)

where \(x_F = \sup\{x \in \mathbb{R} : \Pr(\xi_1 \leq x) < 1\}\). \(\lambda_k\) is called the lag-\(k\) extreme dependence index. When \(\lambda_k > 0\) exists for some \(k > 0\) and \(\{\xi_i\}\), we say there are extreme impacts within \(\{\xi_i\}\).

The existence of \(\lambda\) and \(\lambda_k\) are assumed in this paper. The problem is how to estimate \(\lambda\) (also \(\lambda_k\)), especially when the joint distributions between any two random variables are unknown. In the following, a way to test the null hypothesis: \(\lambda = 0\) against the alternative: \(\lambda > 0\) is developed, without directly estimating \(\lambda\).

### 2.2 Assumptions

The values of extreme dependence indexes can be estimated using empirical probabilities above certain threshold values, but the estimated values may give misleading results since in general, the estimated values are greater than zero even if the random variables are actually extreme independent. A natural question then arises – i.e., whether there is extreme dependence between two random variables. This can be formulated into the following testing hypothesis problems:

\[H_0: \ X \text{ and } Y \text{ are extreme independent} \quad \leftrightarrow \quad H_1: \ X \text{ and } Y \text{ are extreme dependent},\]

(2.3)

which can also be written as

\[H_0: \ \lambda = 0 \ \text{vs.} \ H_1: \ \lambda > 0,\]

(2.4)

or

\[H_0: \ \lambda_k = 0 \ \text{vs.} \ H_1: \ \lambda_k > 0,\]

(2.5)

for the case of lag-\(k\) extreme dependence. We now study how to test (2.4) and (2.5) in the remaining of this section.
2.3 Transformations

Theoretically, as long as the distribution function is continuous and strictly increasing, it can be transformed into any distribution function which is also continuous and strictly increasing. Moreover, the estimation of the extreme dependence index does not change depending on the distribution. This is due to the following identities:

\[
\Pr(X_1 > u \mid X_2 > u) = \frac{\Pr(X_1 > u, X_2 > u)}{\Pr(X_2 > u)}
\]

\[
= \frac{\Pr(G_1^{-1}[F_1(X_1)] > G_1^{-1}[F_1(u)], G_2^{-1}[F_2(X_2)] > G_2^{-1}[F_2(u)])}{\Pr(G_2^{-1}[F_2(X_2)] > G_2^{-1}[F_2(u)])}
\]

\[
= \Pr\{G_1^{-1}[F_1(X_1)] > G_1^{-1}[F_1(u)] \mid G_2^{-1}[F_2(X_2)] > G_2^{-1}[F_2(u)]\},
\]

where \( F_1(x) \) and \( F_2(x) \) are distribution functions for continuous random variables \( X_1 \) and \( X_2 \) respectively; \( G_1(y) \) and \( G_2(y) \) are two arbitrary distribution functions (All the distribution functions are assumed continuous and strictly increasing).

Based on these identities, without loss of any generality, we can assume that the distribution of each random variable \( (X_1 \) or \( X_2) \) is unit Fréchet distribution which is \( G(x) = \exp(-1/x), \ x > 0 \). To be concrete, at first we estimate \( F_1 \) and \( F_2 \) from their samples of \( X_1 \) and \( X_2 \). After that \( X_1 \) and \( X_2 \) are transformed by the estimated \( F_1 \) and \( F_2 \) and then by the inverse of the unit Fréchet distribution function.

Note that only the distributions over a certain threshold are essential for their extreme behavior. In the literature, methods for exceedance modeling over high thresholds are widely used in the applications of extreme value theory. The theory used goes back to Pickands (1975). The only limit distribution of exceedances over threshold has been shown (Pickands 1975) to be a generalized Pareto distribution (GPD) which has a direct connection to generalized extreme value distribution (GEV). Below we give a brief review. For full details, see Embrechts, Klüppelberg, and Mikosch (1997) and the references therein.

Consider the distribution function of a random variable \( X \) conditionally on exceeding some high threshold \( u \), we have

\[
F_u(y) = P(X \leq u + y \mid X > u) = \frac{F(u + y) - F(u)}{1 - F(u)}, \ y \geq 0.
\]
As \( u \to x_p = \sup \{ x : F(x) < 1 \} \), Pickands (1975) shows that the GPD are the only non-degenerate distribution that approximate \( F_u(y) \) for \( u \) large. The limit distribution of \( F_u(y) \) is given by

\[
G(y; \sigma, \xi) = 1 - (1 + \xi \frac{y}{\sigma})^{-1/\xi}.
\] (2.6)

In the GPD, \( y_+ = \max(y, 0) \), \( \xi \) is the tail index, also called shape parameter, which gives a precise characterization of the shape of the tail of the GPD. For the case of \( \xi > 0 \), it is long-tailed, i.e., \( 1 - G(y) \) decays at rate \( y^{-1/\xi} \) for large \( y \). The case \( \xi < 0 \) corresponds to a distribution function which has a finite upper end point at \(-\sigma/\xi\). The case \( \xi = 0 \) yields the exponential distribution with mean \( \sigma \):

\[
G(y; \sigma, 0) = 1 - \exp(-\frac{y}{\sigma}).
\]

The Pareto, or GPD, and other similar distributions have long been used models for long-tailed processes. After a GPD is fitted to the exceedance values of observations over a high threshold \( u \), the data can be transformed to unit Fréchet scales based on the fitted GPD function. An alternative way to obtain unit Fréchet scales is to fit the generalized extreme value distribution (GEV) (2.7) to local maxima of observations.

\[
H(x) = \exp \left[ - \left( 1 + \frac{\mu - x}{\psi} \right)_+^{-1/\xi} \right],
\] (2.7)

where \( \mu \) is a location parameter, \( \psi > 0 \) is a scale parameter, and \( \xi \) is a shape parameter which is the same parameter as in (2.6).

Pickands (1975) first established the rigorous connection of the GPD with the GEV. The GEV parameter estimations can be obtained from a GPD fitting. Smith (2003) is a good reference about GEV and GPD fitting and model diagnosis. GEV parameters are estimated by the maximum likelihood method in this paper.

### 2.4 The theory

Since Fréchet distribution has a simple form with a positive support, it is convenient to develop a test statistic on it. The gamma test for testing extreme independencies between unit Fréchet random variables is introduced here.
Theorem 2.3 Suppose \( \{V_i\} \) and \( \{W_i\} \), \( i = 1, \ldots, n \) are two series of independent identically distributed random variables with distribution function \( G(x) = \Pr(V \leq x) = \exp\{-1/(u + x)\} \), \( x \geq 0, u > 0 \). Then

\[
\Pr\left\{ \frac{u + W_i}{u + V_i} \leq t \right\} = \begin{cases} \frac{t}{1+t} - \frac{t}{1+t}e^{-(1+t)/u}, & \text{if } 0 < t < 1, \\ \frac{t}{1+t} + \frac{1}{1+t}e^{-(1+t)/u}, & \text{if } t \geq 1, \end{cases} \tag{2.8}
\]

\[
\lim_{n \to \infty} \Pr\{n^{-1}\max_{i \leq n} (u + W_i)/(u + V_i) + 1 \leq x\} = e^{-(1-e^{-1/u})x}. \tag{2.9}
\]

Moreover,

\[
\lim_{n \to \infty} \Pr\{n^{-1}\min_{i \leq n} (u + W_i)/(u + V_i) \leq x\} = 1 - e^{-(1-e^{-1/u})x}. \tag{2.10}
\]

The random variables \( \max_{i \leq n} (u + W_i)/(u + V_i) \) and \( \max_{i \leq n} (u + V_i)/(u + W_i) \) are tail independent – i.e.,

\[
\lim_{n \to \infty} \Pr\{n^{-1}\max_{i \leq n} (u + W_i)/(u + V_i) + 1 \leq x, \ n^{-1}\max_{i \leq n} (u + V_i)/(u + W_i) + 1 \leq y\} = e^{-(1-e^{-1/u})x-(1-e^{-1/u})y}, \tag{2.11}
\]

Furthermore, the random variable

\[
\Gamma_{u,n} = \max\{(u + W_i)/(u + V_i)\} + \max\{(u + V_i)/(u + W_i)\} - 2 \max\{(u + W_i)/(u + V_i)\} \times \max\{(u + V_i)/(u + W_i)\} - 1,
\tag{2.12}
\]

is asymptotically gamma distributed, i.e.

\[
n\Gamma_{u,n} \xrightarrow{\mathcal{L}} \zeta, \tag{2.13}
\]

where \( \zeta \) is Gamma\((2, 1 - e^{-1/u})\) distributed.

A proof of Theorem 2.3 is given in Zhang (2003). Under certain mixing conditions, Zhang (2003) also shows that (2.12) and (2.13) hold when \( V_i \)'s are not independent, and \( W_i \)'s are not independent (and hence \( (V_i, W_i) \)'s are not iid).

Note that (2.8) is not continuous at \( t = 1 \). This is caused by the point mass at \( x = 0 \) in \( G(x) \), and the point mass is closed to one when \( u \) is very large. However, this does not induce
a problem from (2.9) through (2.12), because only $t \geq 1$ part is essential in (2.8). Actually the maximum of $(u + W_i)/(u + V_i)$ over $i = 1, ..., n$ is larger than one when $n$ is large enough for $u$.

$\Gamma_{u,n}$ is termed as tail quotient correlation coefficient for a very large threshold $u$ in Zhang (2003). Notice that $\Gamma_{u,n}$ only involves the values above the threshold, i.e. the tail values of observed sequences (or after transformed to unit Fréchet scales). In practice, equations (2.12) and (2.13) together provide a gamma test statistic which can be used to determine whether extreme dependence between two random variables is significant or not. When $n\Gamma_{u,n} > \zeta_\alpha$, $H_0$ of (2.3) is rejected, where $\zeta_\alpha$ is the upper $\alpha$th percentile of the gamma$(2, 1-e^{-1/u})$ distribution. Using simulated examples, Zhang (2003) shows that the gamma test can efficiently identify the extreme (in)dependencies between various dependence variables. Since the gamma test is simple and practically efficient, in this study, we shall apply the gamma test to the data.

Suppose now the variables have been transformed to have unit Fréchet margins. In empirical study, the limit distribution of exceedances over a threshold for the original random variables may be estimated by using GPD at first, and then it can be transformed into unit Fréchet random variable as shown in Section 2.3. In the case of testing extreme dependence, we let $V_i = (X_i - u) I\{X_i > u\}$, $W_i = (Y_i - u) I\{Y_i > u\}$, where $I(\cdot)$ is the indicator function. Values of the threshold $u$ are taken at 95%th (or higher level) sample percentile. In the case of testing lag-$k$ extreme dependence, we let $Y_i = X_{i+k}$, choose a bivariate subsequence $\{(X_{i_l}, Y_{i_l}), \ l = 1, ..., i = 1, ...\}$ such that $i_l - i_{l-1} > k$, and compute the value of the test statistic for the subsequence. If the observed process is lag-$j$ dependence, the gamma test would reject the $H_0$ of (2.3) when $k = j$, but it would retain the $H_0$ of (2.3) when $k = j + 1$.

In the literature, Ledford and Tawn (1996), Peng (1999), Draisma, Drees, Ferreira, and de Haan (2004), and others studied testing extreme dependence problems in which the null hypothesis is that the two random variables are extreme dependent. Under a class of models, various test statistics have been proposed. The derivation of a test statistic under the null hypothesis of extreme dependence is still an open problem.
So far, we have established a testing procedure for determining whether there exist extreme dependencies between random variables. Once the $H_0$ of (2.3) is rejected, we conclude there exists extreme dependence between two random variables of interest. We now turn to real data analysis.

3 The data and data transformation

Our objectives are to study 1) the extreme dependencies between USD/DEM and USD/YEN returns, and 2) extreme impacts on each of the two FX rates. Hence, in this study, let $X_1, X_2, ...$ be USD/DEM returns (daily or hourly), $Y_1, Y_2, ...$ be USD/YEN returns (daily or hourly), and $\xi_1, \xi_2, ...$ be either of these two series.

3.1 The data

The data was collected by Olson & Associates. The observations cover the period from 27 April 1986 until 25 October 1998. The six different time horizons considered are one, two, four, eight, twelve hour and one day periods. Breymann et al (2003) conducted filtering and deseasonalization on the data.

Extreme dependence indexes (and hence extreme co-movements, extreme impacts) are associated with identically distributed random variables, but asset return series such as a FX spot rate is considered not identically distributed due to the long memory of its volatility. Using the deseasonalized data, Dias and Embrechts (2003) conducted dynamic copula modeling with multivariate GARCH model. Since their results are rather convincing, we standardize the deseasonalized data by GARCH model.

In this paper, we use the deseasonalized one hour and one day data and then fit them to univariate GARCH(1,1) models (Bollerslev, 1986). The deseasonalized data is divided by estimated standard deviations and standardized return data (also called devolatilized data, or GARCH residuals) are obtained. We then study extreme co-movements and extreme impacts within the standardized returns. Note that we have restricted our GARCH model fitting to
Table 1: Estimated values of parameters in GEV fitting using standardized return series and threshold value $u = 1.2$ for Daily data. $N_u$ is the number of exceedances over $u$. Numbers in parentheses are estimated standard deviations.

<table>
<thead>
<tr>
<th>Series</th>
<th>$N_u$</th>
<th>$\mu$ (SE)</th>
<th>$\log \psi$ (SE)</th>
<th>$\xi$ (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>USD/DEM Negative</td>
<td>341</td>
<td>3.305482 (0.145446)</td>
<td>-0.438589 (0.151050)</td>
<td>0.001969 (0.061733)</td>
</tr>
<tr>
<td>USD/DEM Positive</td>
<td>323</td>
<td>3.024353 (0.107579)</td>
<td>-0.667407 (0.119296)</td>
<td>-0.060415 (0.050207)</td>
</tr>
<tr>
<td>USD/YEN Negative</td>
<td>328</td>
<td>3.561064 (0.168287)</td>
<td>-0.297793 (0.154565)</td>
<td>0.011028 (0.063915)</td>
</tr>
<tr>
<td>USD/YEN Positive</td>
<td>297</td>
<td>2.949730 (0.132979)</td>
<td>-0.496941 (0.145713)</td>
<td>0.056199 (0.060298)</td>
</tr>
</tbody>
</table>

GARCH(1,1). Other GARCH type models, such as heavy tailed GARCH models (see Mikosch 2003, Straumann 2005, and others), GARCH models with dependent residual random variables (see Lee and Hansen 1994 etc.), are potential applications.

For a comparison, we plot deseasonalized data against standardized data for USD/DEM, USD/YEN in days and hours – see Figure 1. We also plot USD/DEM deseasonalized data against USD/YEN deseasonalized data, USD/DEM standardized data against USD/YEN standardized data, in days and hours respectively – see Figure 2. The plots for the (non-standardized) deseasonalized data and the standardized data look similar but not identical. Particularly for daily data, most estimated standard deviations fall in the range between 0.9 and 1.2 which results in diagonal patterns which can clearly be seen in the left panel in Figure 1. Hence, using the standardized data is better to avoid misspecification in testing extreme dependencies.

### 3.2 Data transformation

Following Section 2.3, we estimate the distribution of standardized data and transform it into unit Fréchet scales so that the gamma test can be used to test the presences of extreme dependencies. The maximum likelihood estimates of GEV parameters are summarized in Tables 1, 2, 3.
Figure 1: Comparison between deseasonalized data and standardized data. The top plots are for USD/DEM. The bottom plots are for USD/YEN. The left plots are for 1 day data. The right plots are for 1 hour data.
Figure 2: Comparison between USD/DEM and USD/YEN. The top plots are for deseasonalized data. The bottom plots are standardized data. The left plots are for 1 day data. The right plots are for 1 hour data.
For daily data, from Table 1, we see that standardized negative returns of German Mark quoted against US dollar has a fat tail (the estimated shape parameter value is greater than 0), but the standardized positive returns are now short tailed. We also see that both returns of Japanese Yen quoted against US dollar are still fat tailed. Notice that all estimated shape parameter values are different, which in turn suggests asymmetry between negative returns and positive returns. Also notice that the estimated shape parameter values are close to zero, and hence, the standardized daily returns can be thought to have Gaussian tails.

For hourly data, since the data points are very large – about 78230 points. We split the data into 10 sub-series for the convenience of computation. From Tables 2, 3, we see that almost all series are fat tailed, except one USD/YEN standardized positive series having a negative estimated shape parameter value. From the split series, one can see that all estimated scale parameter values, all estimated location parameter values, and all estimated shape parameter values are different though they can be thought approximately close to each other for a particular parameter. Dias and Embrechts (2003) also find there are change points in the series. Therefore, it is reasonable to split the whole series into sub-series. What we have done here is simply and equally splitting the whole sequence into 10 sub intervals.

Notice that daily positive returns of USD/DEM are fitted with a short tailed parameter value, while hourly positive returns of USD/DEM are fitted with long tailed parameter values, we argue that the accumulation of data causes the changing of long tail to short tail. But USD/YEN daily positive returns are still long tailed. In other words, the price movements between two different currencies are significantly different. We explore extreme co-movements and extreme impacts for different returns next.

4 Extreme co-movements and extreme impacts

In this section, we use the gamma test to detect extreme co-movements and extreme impacts in two exchange spot rates. In order to do so, the standardized data is converted into unit Fréchet scale for the observations above certain threshold using the fitted GEV function and
Table 2: Estimated values of parameters in GEV fitting using standardized return series and threshold value \( u = 1.2 \) for USD/DEM Hourly data. \( N_u \) is the number of exceedances over \( u \). Numbers in parentheses are estimated standard deviations.
Table 3: Estimated values of parameters in GEV fitting using standardized return series and threshold value $u = 1.2$ for USD/YEN Hourly data. $N_u$ is the number of exceedances over $u$. Numbers in parentheses are estimated standard deviations.
its connection to the generalized Pareto distribution.

The gamma test is implemented for each certain length of time period, since the extreme
dependence between variables may vary by period. The result is reported as the rejection
rate, that is the number of rejecting null hypotheses over the total number of tests.

4.1 Extreme co-movements of daily data

Usually, the notion of extreme co-movements in the stock market means that large jumps in
returns from two different stocks happen at the same time (in this data analysis, we use the
same day and the same hour). When jumps in returns have an impact in the following trading
days or hours, it is better to consider cross-sectional extreme dependencies or to extend the
meaning of extreme co-movements to that large jumps in returns from two different stocks
happen in a short time period of several consecutive days or hours. We use the extended
meaning and apply the gamma test to all transformed data. In each test, we use threshold
values of 95th percentiles of the data and test level of $\alpha = .05$. The total number of tests
equals the sample size subtracting the maximal length of test windows (in this study we choose
500 consecutive points in each test window). The testing results are summarized in Table 4.

In the table, N. USD/DEM stands for the negative returns of USD/DEM. P. USD/DEM
stands for the positive returns of USD/DEM. N. USD/YEN stands for negative returns of
USD/YEN. P. USD/YEN stands for positive returns of USD/YEN. In the table, the cell of
Row (A, B) and Column (Day 0, Day $i$) means that the return value of asset A on Day 0 and
the return value of asset B on Day $i$ are formed in pairs and used in the test.

From Table 4, we see that there are clearly extreme co-movements between negative
USD/DEM returns and negative USD/YEN returns since the rejection rate of the null hy-
pothesis of extreme independence is 100%. The rejection rate of the null hypothesis of extreme
independence between positive returns is 49%. This number (49%) suggests there were ex-
treme co-movement between positive USD/DEM returns and positive USD/YEN returns,
but extreme co-movements did not appear in all subintervals – i.e., there were change points
that price movements behaved differently. A change point analysis conducted by Dias and
Table 4: This table summarizes the test results by using the gamma test and subset scheme of daily data. The number in each cell is the rejection rate of the null hypothesis of extreme independence.

<table>
<thead>
<tr>
<th>Pairs</th>
<th>(Day 0, Day 0)</th>
<th>(Day 0, Day 1)</th>
<th>(Day 0, Day 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N. USD/DEM, N. USD/YEN)</td>
<td>1.0000</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(N. USD/DEM, P. USD/YEN)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(P. USD/DEM, N. USD/YEN)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(P. USD/DEM, P. USD/YEN)</td>
<td>0.4858</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 5: This table summarizes the test results by using the gamma test and subset scheme of hourly data. The number in each cell is the rejection rate of the null hypothesis of extreme independence.

<table>
<thead>
<tr>
<th>Pairs</th>
<th>(Hour 0, Hour 0)</th>
<th>(Hour 0, Hour 1)</th>
<th>(Hour 0, Hour 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N. USD/DEM, N. USD/YEN)</td>
<td>0.7908</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(N. USD/DEM, P. USD/YEN)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(P. USD/DEM, N. USD/YEN)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(P. USD/DEM, P. USD/YEN)</td>
<td>0.8563</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

4.2 Extreme co-movements of hourly data

Like in the previous section, we use the extended meaning and perform the gamma test to all transformed data. There are totally 2615 tests. In each test, we use threshold values of 95th percentiles of the data and test level of $\alpha = 0.05$. The testing results are summarized in Table 5. In the table, the cell of Row (A, B) and Column (Hour 0, Hour i) means that the return value of asset A on Hour 0 and the return value of asset B on Hour i are formed in pairs and used in the test.

From Table 5, we see that there are clearly extreme co-movements between USD/DEM
hourly price change and USD/YEN hourly price change. Notice that there are more time periods of positive extreme co-movements than time periods of negative extreme co-movements which show an asymmetric price movements in hourly return data. We also see the difference between hourly data and daily data. During market recession, the chance of extreme co-movements will more likely be increasing (from a low rejection rate in hourly data to a high rejection rate in daily data of null hypothesis of extreme independence). During market expansion, however, the chance of extreme co-movements is likely decreasing from hourly movements to daily movements.

We also see that extreme co-movements only occur at the same day or the same hour because all numbers in columns (Day 0, Day 1) and (Day 0, Day 2) are zeros. This phenomenon suggests that a large price change of one spot rate at a particular time doesn’t have an extreme impact on another spot rate at the time points after 1 hour or 1 day. This is interesting. It is suggesting that the cross-sectional extreme impacts only occur in a very short time period, probably a ‘tick by tick’ time period.

4.3 Extreme impacts of daily data

From the previous two sections, we have seen that extreme co-movements between two different exchange currencies were found using the gamma test. We now turn to study extreme impacts within each exchange spot rate sequence. In order to test extreme impacts, we need to project a univariate sequence into a bivariate sequence. A simple way to accomplish this is to let \( X'_i = X_i, Y'_i = X_{i+k}, i = 1, \ldots, n - k \) for a particular lag \( k \).

For each return sequence (negative or positive), the rejection rate of null hypothesis of extreme independence varies from case to case. One can see that identifying extreme impacts is more difficult than identifying extreme co-movements between two different risk factors. Sometimes, the gamma test alone may not be sufficient to suggest a lag-\( k \) extreme dependence. Other statistical quantities together with the test results may help to identify a best lag-\( k \). A natural choice would be the empirical estimation of the extreme dependence index. The

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.0242</td>
<td>0.0415</td>
<td>0.1853</td>
<td>0.0062</td>
<td>0.0261</td>
</tr>
<tr>
<td>2</td>
<td>1.0000</td>
<td>0.0182</td>
<td>0.0647</td>
<td>0.7844</td>
<td>0.0123</td>
<td>0.0437</td>
</tr>
<tr>
<td>3</td>
<td>1.0000</td>
<td>0.0182</td>
<td>0.1313</td>
<td>0.0000</td>
<td>0.0062</td>
<td>0.0092</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>0.0121</td>
<td>0.0419</td>
<td>0.0000</td>
<td>0.0062</td>
<td>0.0113</td>
</tr>
</tbody>
</table>

The tested results are summarized in Tables 6, 7.

In the tables, R.R. denotes the rejection rate of rejecting the null hypothesis of extreme independence. There are totally 671 tests in which each uses 500 data points. For each data set of 500 data points, we compute the empirical estimates of extreme dependence index using the above empirical formula. (M.E.D.I.) stands for the average of estimated extreme dependence indexes. [S.std] stands for the sample standard deviation of estimated extreme dependence indexes.

We see from these two tables, up to lag-4, negative returns have extreme impacts all the time. Extreme impacts within positive returns were found to last about 3 consecutive days.
<table>
<thead>
<tr>
<th>lag</th>
<th>USD/YEN Negative</th>
<th>USD/YEN Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000  0.0189  0.0421</td>
<td>1.0000  0.006  0.076</td>
</tr>
<tr>
<td>2</td>
<td>1.0000  0.0252  0.0975</td>
<td>1.0000  0.0119  0.1017</td>
</tr>
<tr>
<td>3</td>
<td>1.0000  0.0126  0.0445</td>
<td>0.0000  0.0238  0.0133</td>
</tr>
<tr>
<td>4</td>
<td>1.0000  0.0126  0.0333</td>
<td>0.0000  0.0179  0.0105</td>
</tr>
</tbody>
</table>

Table 7: *Lag-k tail independence test for transformed return data.* R.R. denotes the rejection rate of rejecting the null hypothesis of extreme independence. M.E.D.I. means empirical estimation of extreme dependence index. [S.std] stands for the sample standard deviation of estimated extreme dependence indexes.

### 4.4 Extreme impacts of hourly data

Like the previous section, we project each univariate sequence into a bivariate sequence. The testing results are summarized in Tables 8, 9 for lag-\(k=1,2,\ldots,10,96,\ldots,100\). We see there are extreme impacts of hourly returns (negative or positive) for very large lags. The rejection rates are also plotted in Figure 3. From the figure, we see that the rejection rates within the first 30 lags are around 80%. Then the rejection rates are varying in a wide range when lag-\(k\) is increasing.

From the previous section, we have seen that extreme impacts of positive daily data appear to last within the following two consecutive days. From results in Tables 8, 9, we see that extreme impacts of positive hourly data appear to last more than four days (will be higher if larger lags are used in test). This phenomenon suggests that extreme impacts of high frequency data are stronger than extreme impacts of low frequency data. It also suggests great care must be taken when modeling high frequency data.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9409</td>
<td>0.0298</td>
<td>0.0664</td>
<td>0.8625</td>
<td>0.0259</td>
<td>0.0624</td>
</tr>
<tr>
<td>2</td>
<td>0.9485</td>
<td>0.0277</td>
<td>0.0604</td>
<td>0.8729</td>
<td>0.0187</td>
<td>0.0537</td>
</tr>
<tr>
<td>3</td>
<td>0.8498</td>
<td>0.0154</td>
<td>0.0636</td>
<td>0.8601</td>
<td>0.0218</td>
<td>0.0624</td>
</tr>
<tr>
<td>4</td>
<td>0.7268</td>
<td>0.0157</td>
<td>0.0522</td>
<td>0.6782</td>
<td>0.0138</td>
<td>0.0564</td>
</tr>
<tr>
<td>5</td>
<td>0.866</td>
<td>0.0146</td>
<td>0.065</td>
<td>0.7295</td>
<td>0.0144</td>
<td>0.0557</td>
</tr>
<tr>
<td>6</td>
<td>0.7855</td>
<td>0.011</td>
<td>0.0486</td>
<td>0.8159</td>
<td>0.0108</td>
<td>0.0439</td>
</tr>
<tr>
<td>7</td>
<td>0.8356</td>
<td>0.0085</td>
<td>0.0575</td>
<td>0.7273</td>
<td>0.0103</td>
<td>0.0592</td>
</tr>
<tr>
<td>8</td>
<td>0.9547</td>
<td>0.0077</td>
<td>0.063</td>
<td>0.8699</td>
<td>0.011</td>
<td>0.0655</td>
</tr>
<tr>
<td>9</td>
<td>0.8104</td>
<td>0.0108</td>
<td>0.0594</td>
<td>0.7425</td>
<td>0.0054</td>
<td>0.0464</td>
</tr>
<tr>
<td>10</td>
<td>0.7791</td>
<td>0.0087</td>
<td>0.0528</td>
<td>0.755</td>
<td>0.01</td>
<td>0.0548</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>96</td>
<td>0.8653</td>
<td>0.0013</td>
<td>0.0569</td>
<td>1.0000</td>
<td>0.0003</td>
<td>0.135</td>
</tr>
<tr>
<td>97</td>
<td>1.0000</td>
<td>0.0015</td>
<td>0.0866</td>
<td>1.0000</td>
<td>0.001</td>
<td>0.1088</td>
</tr>
<tr>
<td>98</td>
<td>1.0000</td>
<td>0.0005</td>
<td>0.0455</td>
<td>1.0000</td>
<td>0.0013</td>
<td>0.0699</td>
</tr>
<tr>
<td>99</td>
<td>0.0000</td>
<td>0.0015</td>
<td>0.0131</td>
<td>1.0000</td>
<td>0.0005</td>
<td>0.0533</td>
</tr>
<tr>
<td>100</td>
<td>1.0000</td>
<td>0.0008</td>
<td>0.0773</td>
<td>1.0000</td>
<td>0.0008</td>
<td>0.1099</td>
</tr>
</tbody>
</table>

Table 8: Lag-$k$ tail independence test for transformed return data (USD/DEM). M.E.D.I. means empirical estimation of extreme dependence index.
<table>
<thead>
<tr>
<th>lag</th>
<th>Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8911</td>
<td>0.0247</td>
</tr>
<tr>
<td>2</td>
<td>0.8645</td>
<td>0.0263</td>
</tr>
<tr>
<td>3</td>
<td>0.8537</td>
<td>0.0173</td>
</tr>
<tr>
<td>4</td>
<td>0.8871</td>
<td>0.0162</td>
</tr>
<tr>
<td>5</td>
<td>0.8406</td>
<td>0.0142</td>
</tr>
<tr>
<td>6</td>
<td>0.9538</td>
<td>0.0108</td>
</tr>
<tr>
<td>7</td>
<td>0.7673</td>
<td>0.0111</td>
</tr>
<tr>
<td>8</td>
<td>0.847</td>
<td>0.007</td>
</tr>
<tr>
<td>9</td>
<td>0.9058</td>
<td>0.0085</td>
</tr>
<tr>
<td>10</td>
<td>0.9071</td>
<td>0.0057</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>96</td>
<td>1.0000</td>
<td>0.0003</td>
</tr>
<tr>
<td>97</td>
<td>1.0000</td>
<td>0.0008</td>
</tr>
<tr>
<td>98</td>
<td>1.0000</td>
<td>0.0021</td>
</tr>
<tr>
<td>99</td>
<td>1.0000</td>
<td>0.0013</td>
</tr>
<tr>
<td>100</td>
<td>0.609</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

Figure 3: Rejection rates of Lag-1 to lag-100 extreme dependence tests for hourly data.
5 Conclusion

This paper has presented ways of analyzing extreme co-movements and extreme impacts among financial market variables. We focused on two FX spot rates. We found that there exist extreme co-movements and extreme impacts in financial market data.

The methodology proposed in this paper of testing extreme co-movements and extreme impacts among two FX spot rates can be used to assessing extreme co-movements and extreme impacts among other financial returns, such as stock returns. We have proposed procedures of: (1) fitting GARCH(1,1) (can be GARCH(p,q)) model to price returns; (2) obtaining devolatized return data; (3) fitting GPD to exceedance values over a typical high threshold value; (4) transforming the observed values to unit Fréchet scales by fitted GPD (or GEV) functions; (5) performing the gamma test to the transformed data; (6) identifying extreme co-movements and extreme impacts among financial returns. Based on the results of our analysis with FX spot rates, we argue that extreme co-movement and extreme impact problems are inherently stochastic control problems, since they will influence the decision taken today and ultimately influence a decision taken in the future.

For the daily data, there are existing models which are able to model extreme co-movements and extreme impacts among the interest variables. Those models are studied, for example, by Deheuvels (1983), Davis and Resnick (1989), Smith and Weissman (1996), Hall, Peng and Yao (2002), Zhang and Smith (2004a, 2004b), Zhang (2005), Chamú Morales (2005), Zhang and Huang (2006), Heffernan, Tawn and Zhang (2006), among others. The models have been used in stochastic optimal control and portfolio optimizations. For the high frequency data (for example, hourly data), efforts of constructing reasonable models are still needed. Our ultimate goal is to construct portfolio optimization models for high frequency data, especially models which are applicable during the time of market recession or market expansion.
Acknowledgement

The work by the first author was supported in part by NSF Grants DMS-0505528, and DMS-0630210, Wisconsin Alumni Foundation, and Institute for Mathematical Research, ETH Zürich. The authors thank Dr Dias for sharing the data with us.

References


