Asymmetric Time-Scale Dependence Structure: Evidence in International Equity Markets

Jesus Cuauhtemoc Tellez Gaytan, Universidad Autonoma del Carmen, Mexico

Abstract: This paper assesses for asymmetries in the dependence structure between equity markets at different time scales based on a multiresolution decomposition approach using wavelets. Besides broad research has been done in determining the dependence structure pattern among equity markets mostly based on copulas and extreme value theory, it has not considered asymmetries in time scales at all. The daily S&P 500 (US) and IPC (Mexico) returns are decomposed into different resolution levels using the maximal overlap discrete wavelet transform (MODWT) and as filtering function the Symmlet wavelet. Then, at each level to each decomposed time series returns it is fitted the Generalized Pareto Distribution to lower and upper tails. Finally, the dependence structure parameters are estimated across each paired resolution levels using copulas. Empirical results show evidence of asymmetric dependence structure at different time scales and these results could change the way asset allocation is performed when bulls and bears move asymmetrically in different time spans across equity markets.

Keywords: Correlation, wavelet transform, tail dependence, copulas.

1. Introduction

Cont (2001) introduces stylized statistical properties of asset returns which three of them are related to asymmetric phenomena. The first one, gain/loss asymmetry, states that it is equally to observe large drawdowns but not upward movements in stock prices and stock index values. The second property named as aggregational Gaussianity is related to the distribution of returns which approaches a Normal one when increasing time scale. Finally, the third type of asymmetry phenomenon argues that coarse-grained measures of volatility predict fine-scale volatility better than the other way round. The last two types of asymmetries are also argued by Cont and Tankov (2003) as follows:

… prices move essentially by jumps between time scales, they still manifest discontinuous behavior at the scale of months and only after coarse-graining their behavior over longer time scales do we obtain something that resembles Brownian motion.

In particular, asymmetric dependence structure is related to a phenomenon in which markets are more dependent in bear markets than in bull markets. Not only have this type of asymmetry been studied within equity markets as Ang and Chen (2002), and Patton (2004) do, but also across countries and markets which has implied to examining the correlation structure. Hilliard (1979) examines the international equity markets correlation structure of 10 major exchanges mostly
during a financial crisis based on spectral statistics such as autospectrum, coherence, phase angle, and tau. Eun and Shim (1989) estimate a VAR system to analyze co-movements among nine equity markets, whose main finding is that the U.S. stock market is the most influential market in the world.

Lin, Engle and Ito (1994) analyze returns and volatility correlation between New York and Tokyo stock markets based on signal-extraction model with GARCH processes, whose main finding is that cross-market interdependence in returns and volatilities is a bi-directional one. Longin and Solnik (2001) model multivariate distribution tails based on extreme value theory among the United States, the United Kingdom, France, Germany, and Japan equity index returns. The main finding states that correlation increases in bear markets but in bull markets.

The use of copulas to analyzing the correlation structure goes back to Embrechts, McNeil, and Straumann (1998), who present properties and pitfalls about correlation and dependency in risk management. One of its main contributions is to distinguish among different statistical measures when modeling non-elliptical relationships in finance, such as linear correlation, rank correlation and tail dependence. Kim (2005) investigates the differences between the linear measure and alternative copula models, and the different movements between the general dependence and tail dependence among Asian financial markets.

Hu (2006) examines the dependence pattern among S&P 500, FTSE, Nikkei, and Hang Seng stock market indices based on mixed-copula model, in which symmetric and asymmetric dependence structure are captured by the mixture of the Gaussian copula, the Gumbel copula and the Gumbel survival copula. The Gaussian copula is used as a matter of traditional financial modeling, the Gumbel copula to capture positive right tail dependence and Gumbel survival copula to capture left tail dependence. The main finding is that stock market pairs that have lower correlation have almost the same probability to crash together as pairs that have higher correlation coefficients.

The first applications of wavelets theory in economic and financial phenomena analysis are done by Ramsey and Lampart (1998) who examine the relationship between economic data such as personal income-spending, M1 and M2 monetary aggregates. T-bills interest rates and durable, non-durable and consumption goods and services. Aguiar, Azevedo and Soares (2007) analyze the effects of interest rate changes on industrial production, inflation, and M1 and M2 monetary aggregates. Lee (2004) studies price and volatility spillovers between US and Korean stock markets, whose results show that movements in stock returns are mainly caused by short-term fluctuations, and regressing specific resolution levels (or details) D1 and D2 between stock markets, shows that volatility spillovers effects are from U.S. stock markets to the Korean counterparts but not vice versa.\(^1\)

Sharkasi, Ruskin and Crane (2005) investigate the relationships among the U.K., the U.S., Irish, Portuguese, Brazilian, Japanese, and Hong Kong stock markets. They estimate simple regressions and reverse regressions models using row-return series reconstructed from the

---
\(^1\) Detail or crystal \(D_1\) is related to the highest frequencies on a onde-day scale, and cristal \(D_2\) is related to the second highest frequencies on a two-day scale.
wavelet component D1 and the return series rebuilt from the wavelet crystals D1 and D2. The results show that spillovers effects are from the U.K. and Portuguese markets onto the U.S. and Brazilian markets, which influence Asian markets, in turn Japan and Hong Kong impact the European markets. Tellez (2008) examines co-movements among Latin American (LATAM) equity markets based on the maximal overlap discrete wavelet transform, whose results show and overall increasing correlation and the covariance structure points out a certain type of pull-back co-movement by which LATAM equity markets correlation is pulled back-down in short time periods.

This paper aims to assess for asymmetric dependence structure at different time scales based on a hybrid approach which combines copulas and wavelets analysis where: a copula is a function that connects k marginal distributions to recover the joint distribution which allows to model dependency between individual random variables; and wavelets are functions that oscillate (wave) and decay (let) at different number of approximations, which act as filters that allow to capture high-frequency components that occur in short time periods and low-frequency components that occur in long time periods. Firstly, the daily S&P 500 and IPC returns are decomposed into different time scales based on the maximal overlap discrete wavelet transform (MODWT) and as filtering function is used the Symmlet wavelet. Then, at each resolution level or crystal related to a time-scale of each equity index return series, it is fitted the Generalized Pareto Distribution to lower and upper tail distributions. Finally, the dependence structure parameters are estimated using copulas across each equity index details. The results are compared to those of the benchmark model which consists in estimating the global dependence structure.

The paper is organized as follows: section two describes copulas theory which highlights different possible dependence structure patterns that can arise when examining relationships between random variables. Wavelets theory and specifically the multi-resolution decomposition algorithm are presented in section three. Section four describes data and methodology. Empirical results are given in section five and finally conclusions are given in section six.

2. Copulas

A copula is a function that links marginal distributions which allows to restore the joint distribution among random variables and to analyze their degree of association and dependence structure even when modeling non-Gaussian multivariate distributions. The theory of copulas is based on Sklar’s theorem:

**Theorem 1** Sklar’s Theorem Let H be a joint distribution function with marginals F and G. Then there exists a copula C such that for all x, y in \( \mathbb{R} \)

\[
H(x, y) = C(F(x), G(y)).
\]  

\[ (1) \]

---

2 This section is based on Nelsen (1999).
If $F$ and $G$ are continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on $\text{Ran} F \times \text{Ran} G$. Conversely, if $C$ is a copula and $F$ and $G$ are distribution functions, then the function $H$ defined by (1) is a joint distribution function with marginals $F$ and $G$.

**Definition 1** A distribution function is a function $F$ with domain $\mathbb{R}$ such that

1. $F$ is nondecreasing,
2. $F(-\infty) = 0$ and $F(+\infty) = 1$.

**Definition 2** A joint distribution is a function $H$ with domain $\mathbb{R}^2$ such that

1. $H$ is 2-increasing,
2. $H(x,-\infty) = H(-\infty,y) = 0$, and $H(+\infty,+\infty) = 1$.

From Definition 1, if $F$ is strictly increasing, then it has but a single quasi-inverse, which is of course the ordinary inverse, for which it used the customary notation $F^{-1}$. From Theorem 1 if $H$ is the joint distribution function with marginals $F$ and $G$, then it has but a single subcopula $C'$, which is followed by the next corollary

**Corollary 1** Let $H$ be the joint distribution function with marginals $F$ and $G$, and a subcopula $C'$, and let $F^{-1}$ and $G^{-1}$ be quasi-inverse of $F$ and $G$, respectively. Then for any $(u,v)$ in $\text{Dom} C'$,

$$C'(u,v) = H(F^{-1}(u), G^{-1}(v)).$$  \hspace{0.5cm} (2)

The importance of Corollary 1 falls in the process of constructing copula families, where $u$ and $v$ ($u,v \in [0,1]$) are the cumulative distribution functions (CDF) of $X$ and $Y$. A known example is the Gaussian copula,

$$C(u,v) = N_{\rho}(\Phi^{-1}(u), \Phi^{-1}(v)),$$  \hspace{0.5cm} (3)

where is $\Phi$ the normal standard distribution function and $N_{\rho}$ is the bivariate normal standard distribution function with correlation coefficient $\rho$. Also, following Corollary 1, the probability density function (PDF) of a bivariate copula $C(u,v)$ is defined as

$$c(u,v) = \frac{\partial C(u,v)}{\partial u \partial v}.$$  \hspace{0.5cm} (4)

Then, given (4), it is possible to recover the density function $h(x,y)$ of the joint distribution $H(x,y)$ by multiplying the copula density with the marginal densities $f(x)$ and $g(y)$,

$$h(x,y) = c(F(x), G(y)) f(x) g(y),$$  \hspace{0.5cm} (5)
where the copula $C(u,v)$ is modeling the dependence between the random variables $X$ and $Y$. The types of dependence that could arise between $X$ and $Y$, basically in the lower and upper distribution tails are described in the following definition:

**Definition 3** Let $X$ and $Y$ be random variables.

1. $Y$ is left tail decreasing in $X$ [denoted as LTD($Y|X$)], if $P[Y \leq y | X \leq x]$ is a nonincreasing function of $x$ for all $y$.

2. $X$ is left tail decreasing in $Y$ [denoted as LTD($X|Y$)], if $P[X \leq x | Y \leq y]$ is a nonincreasing function of $y$ for all $x$.

3. $Y$ is a right tail increasing in $X$ [denoted as RTI($Y|X$)], if $P[Y > y | X > x]$ is a nondecreasing function of $x$ for all $y$.

4. $X$ is right tail increasing in $Y$ [denoted as RTI($X|Y$)], if $P[X > x | Y > y]$ is a nondecreasing function of $y$ for all $x$.

Finally the following theorem given Definition 3 shows that tail monotonicity is a property of the copula when the random variables are continuous.

**Theorem 2** Let $X$ and $Y$ be continuous random variables with copula $C$. Then

1. LTD($Y|X$) if and only if for any $v$ in $I$, $C(u,v)/u$ is nonincreasing in $u$;

2. LTD($X|Y$) if and only if for any $u$ in $I$, $C(u,v)/v$ is nonincreasing in $v$;

3. RTI($Y|X$) if and only if for any $v$ in $I$, $[1-u-v+C(u,v)]/(1-u)$ is nondecreasing in $v$, or equivalently, if $[v-C(u,v)]/(1-u)$ is nonincreasing in $u$;

4. RTI($X|Y$) if and only if for any $u$ in $I$, $[1-u-v+C(u,v)]/(1-v)$ is nondecreasing in $v$, or equivalently, if $[u-C(u,v)]/(1-v)$ is nonincreasing in $v$.

where $[1-u-v+C(u,v)]$ represents the survival function of the copula $C(u,v)$.

Definition 3 and Theorem 2 describe an important property of copulas which is to modeling tail dependence, then it is said that a copula $C(u,v)$ has left or lower tail dependence if

$$
\lim_{u \to 0} \frac{C(u,v)}{u} = \lambda_l > 0,
$$

and right or upper tail dependence if

$$
\lim_{v \to 1} \frac{C(u,v)}{1-v} = \lambda_u > 0.
$$
\[
\lim_{u \to 1} \frac{C(u,v)}{u} = \lambda_r > 0, \tag{7}
\]

where \(C(u,v)\) is the survival function of the copula \(C(u,v)\) as described in Theorem 2.

3. Wavelets

Wavelets (small waves) are functions of special structure described by basis functions and represented by successive approximation series which allow the decomposition of signals at different scales and resolutions, against the Fourier analysis which only gives frequency components but cannot show their time localization. Wavelets have the property to concentrate their energy along time to allow analysis of temporal, non-stationary and time varying phenomena.

Wavelet analysis has its nearest backgrounds on the Short Time Fourier Transform (STFT), this one based on the Fourier Transform which main purpose is to represent a complex function by a weighted sum of simple functions obtained from a simpler one known as prototype or basis function. The Fourier transform has the characteristic to allow global analysis of a signal, since the terms \(\cos \omega t\) and \(\sin \omega t\) represent global functions, and bus this is meant that the Fourier transform is perfect compactly supported in the frequency domain but not in the time domain,

\[
f(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt. \tag{8}
\]

The STFT works by taking out the required portion of the spectrum and then applying on it the Fourier transform. Nevertheless, the function used to remove the required portion of the signal is a constant window-function which cannot face the problem when signals have very high frequency components with short time spans, and low frequency components with long time spans. The wavelet transform allows time-frequency analysis since it works with a varying-window function: it is possible to increase the window-function radius along time whenever frequencies are decreasing, and decrease it whenever frequencies are increasing.

Figure 1 shows Heisenberg boxes in the frequency-time plane with varying-window functions, where the signal \(f(t)\) is windowed by the function \(\phi(t)\) in the neighbourhood \(t=b\) and the output will be a windowed-function \(f_{\phi}(t) = \phi(t-b)\) on which it is applied the Fourier transform.
Figure 1: Frequency-time plane and the wavelet function.

Given the importance of the window-function described above, the wavelet transform is the convolution of a wavelet function and the original signal, where the former depends on two parameters: 1) translation (localization) parameter, which represents movements of the wavelet function along axis time; 2) dilation parameter, which allows expansion and contraction of the wavelet function, so high and low frequencies of the original signal can be captured. The wavelet transform works in an opposite direction to the STFT, it first decomposes the original signal in frequency bands and then analyzes it over time:

\[ W(a, b) = \frac{1}{\sqrt{a}} \int f(t) \psi^* \left( \frac{t-b}{a} \right) dt, \tag{9} \]

and the original signal can be reconstructed (inverse wavelet transform) as

\[ f(t) = \frac{C_{\psi}}{a^2} \int_{a>0} \int_b W(a, b) \psi^* \left( \frac{t-b}{a} \right) dadb, \tag{10} \]

where \( a > 0 \) and \( b \) are the dilation and localization parameters, respectively; \( \psi \) is the mother wavelet, \( C_{\psi} \) is a constant which depends on \( \psi \), and \( W(a, b) \) is the continuous wavelet transform (CWT). Also, the wavelet transform can be represented as the inner product

\[ W(a, b) = \langle x, \psi_{a,b} \rangle. \tag{11} \]

Since the CWT is a function that depends on two continuous parameters, it may result in redundant information (a varying number of coefficients with few numbers of scales). This problem is solved by discretizing the parameters \( a \) and \( b \) using multiresolution analysis: low-pass and high-pass filters are applied iteratively and then sampled down by two as a cascade. This process results in the Discrete Wavelet Transform (DWT). If \( a \) and \( b \) are represented by \( 2^j \) and \( k2^j \), respectively, then the integral (7) becomes
\[ W(k^{2-j}, 2^{-j}) = 2^{j/2} \int_{-\infty}^{\infty} f(t) \psi(2^j t - k) \, dt, \quad (12) \]

which is approximated as

\[ W(k^{2-j}, 2^{-j}) \approx 2^{j/2} \sum_{n} f(n) \psi(2^j n - k). \quad (13) \]

Given \( a \) and \( b \) discretized, and a function \( f(t) \) in \( L^2(R) \) which can be represented by a sequence of mother and father wavelet functions \( \psi \) and \( \phi \), respectively

\[ \psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k) = 2^{-j/2} \psi\left(\frac{t-2^j k}{2^j}\right), \quad (14) \]

\[ \phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j} t - k) = 2^{-j/2} \phi\left(\frac{t-2^j k}{2^j}\right), \quad (15) \]

the function \( f(t) \) can be written as

\[
\begin{align*}
f(t) &= \sum_{k} s_{J,k} \phi_{j,k}(t) + \sum_{k} d_{j,k} \phi_{j,k}(t) \\
&\quad + \sum_{k} d_{J-1,k} \psi_{J-1,k}(t) + \ldots + \sum_{k} d_{1,k} \psi_{1,k}(t),
\end{align*}
\]

(16)

where the coefficients \( s_{j,k}, d_{j,k}, \ldots, d_{1,k} \) are the wavelet transform coefficients contained in a column vector \( W \) with length \( N \), where the first elements \( N - N/2^J \) are the wavelet coefficients and the last elements \( N/2^J \) are the scaling coefficients. The coefficients \( s_{j,k} \) are the smooth coefficients which represent the underlying smooth behaviour of the signal at the coarse scale \( 2^j \); \( d_{j,k} \) are the detail coefficients named as crystals which represent deviations from the smooth behaviour and describe deviations at the coarse scale, and \( d_{J-1,k}, \ldots, d_{1,k} \), are progressively finer scale deviations.

Therefore, the function \( f(t) \) can be represented in terms of its detail series at different resolutions

\[ D_j(t) = \sum_{k} d_{j,k} \phi_{j,k}(t), \text{ for } j = 1, 2, \ldots, J. \quad (17) \]

and smooth variations
\[ S_j(t) = \sum_k s_{j,k} \phi_{j,k}(t), \]  
which yields  
\[ f(t) = S_j(t) + D_j(t) + D_{j-1}(t) + \ldots + D_1(t), \]

named as multiresolution decomposition (MRD).

The modified version of the DWT is the maximal overlap discrete wavelet transform (MODWT) or non-decreasing DWT, which is a highly redundant nonorthogonal transform resulting in new column vectors \( \tilde{W}_1, \tilde{W}_2, \ldots, \tilde{W}_J \) and \( \tilde{V}_J \); where \( \tilde{W}_j \) contains the MODWT wavelet coefficients associated with changes in \( X \) at a scale \( \lambda_j \Delta(t) \), and \( \tilde{V}_J \) contains the MODWT scaling coefficients associated with changes at scales \( \lambda_{J+1} \Delta(t) \) and higher ones.

Thus, as in (19) it is possible to perform multiresolution analysis of \( f(t) \),

\[ f(t) = \sum_{j=1}^{J} \tilde{D}_j + \tilde{S}_j, \]

where \( \tilde{V}_J \) are the MODWT detail series, and \( \tilde{V}_J \) are the MODWT smooth variations.

Finally the wavelet variance of a signal or time series is the variance of the decomposed time series defined as:

\[ \sigma_X^2 = \frac{1}{N} \sum_{t=0}^{N-1} (X_t - X)^2 = \frac{1}{N} \| W_j \|^2 - X^2 = \frac{1}{N} \sum_{j=1}^{J} \| W_j \|^2 + \frac{1}{N} \| V_j \|^2 - X^2, \]

whose unbiased estimator is defined as

\[ \nu_X^2(\lambda_j) = \frac{1}{2 \lambda_j N_j} \sum_{l=L_j}^{N_j-1} W_{j,l}^2, \]

where \( N_j = N/2^j \), \( N_j = N_j - L_j' \), and \( L_j' = \left( (L-2)(1-2^{-j}) \right) \).
4. Data and Methodology

4.1 Data

The empirical research is performed using daily S&P 500 (U.S.) and IPC (Mexico) prices, which are transformed into daily returns as

\[ \text{Re}_t = \ln \frac{P_t}{P_{t-1}}. \]  

(23)

The whole data consisted of 2,280 observations from January 3rd 2000 up to June 30th 2009. Table 1 shows the descriptive statistics of both equity indices returns under a global approach, in which it is shown that both stock markets have experienced similar maximum and minimum returns even though the IPC has in average outperformed the S&P 500.

Skewness shows that IPC has experienced more positive returns than negative ones, against the S&P 500 which has experienced more negative returns than positive ones. The kurtosis statistic shows that the S&P 500 has experienced fatter tails than the IPC has done, which means that greater extreme (negative) returns can be found in the US stock market than on the Mexican one. Finally, at a significance level of 0.05 the Jarque-Bera test rejects normality in equity indices returns.

<table>
<thead>
<tr>
<th>Table 1: Descriptive Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Kurtosis</td>
</tr>
<tr>
<td>Jarque-Bera (JB)</td>
</tr>
<tr>
<td>Probability (JB)</td>
</tr>
</tbody>
</table>

Though one market can outperform the other, the task is to determine whether both of them are co-moving and the type of co-movement they have been showing. Figure 2 shows at a first glance that both equity indices returns are positively related but in a weaker way since most if the paired observations are concentrated near the origin and also as measured by the linear correlation, this showed to be \( p=0.2185 \).
But what happens with the relationship among outliers or in the tails? It seems that both equity indices returns have shown a stronger relationship in bearish markets than in bullish ones. Then we would expect left tail dependency to be statistical significant rather than a right one, as Hu (2006) finds that ‘stock market pairs with lower with lower correlation have almost the same probability to crash together as pairs that have higher correlation coefficients’. A further step in this paper is to determine left and right tail dependency at different time scales in the frequency-time domain.

### 4.2 Methodology

The S&P 500 and IPC returns are decomposed into different resolution levels or crystals each one related to different time scales. For example, the first resolution level \( j=1 \) is related to a time scale between \( \lambda_1 = 2^{j-1} = 2^0 = 1 \) and 2 days. At this level, the highest frequency components are captured in a time span of one-two days. Then at level \( j=2 \), the second highest frequency components are captured in a time span of \( \lambda_2 = 2^1 = 2 \) and 4 days, and successively up to the highest level which will capture the lowest frequency components or smooth variations which represent the underlying smooth behavior of the time series at the coarse scale 2^1.

Since the filtering or decomposition process implies the convolution of the wavelet function and the original time series, then at each level \( j \) under the MODWT, the wavelet transform coefficients are kept in new column vectors \( \hat{W}_j \) associated with changes in \( X \) at a scale \( \lambda_j \Delta t \). The hybrid approach applied in this paper estimates the dependence structure between the S&P 500 and IPC returns at each resolution level. In this case \( \hat{W}_1^{S&P} \) and \( \hat{W}_1^{IPC} \) are the new column vectors which contain the wavelet transform coefficients in the first level \( j=1 \); \( \hat{W}_2^{S&P} \) and \( \hat{W}_2^{IPC} \)
will be the column vectors containing the wavelet transform coefficients in a second stage, and successively up to $\mathcal{V}^{s&p}_j$ and $\mathcal{V}^{IPC}_j$ which will contain the smooth coefficients.

Then, at each resolution level $D_j$ it is fitted the Generalized Pareto Distribution (GPD) to lower and upper thresholds. The GPD is related to the extreme value theory (EVT) which models the tails of a probability distribution $F$, where the extreme values are the values of a sequence of random variables $X_1, X_2, X_3, \ldots, X_n$ identical and independently distributed which exceeds a threshold $u$. The distribution of exceedances over the threshold $u$ is defined as the conditional probability:

$$F_u(y) = \Pr\{X - u \leq y \mid X < u\} = \frac{F(y + u) - F(u)}{1 - F(u)}, \quad y > 0, \quad (24)$$

if $u$ is large enough, then there exists a positive function $\beta(u)$ such that (24) is well approximated by the Generalized Pareto Distribution:

$$G_{\xi, \beta(u)}(y) = \begin{cases} 1 - \left(1 + \frac{\xi(y)}{\beta(u)}\right)^{-1/\xi} & \forall \xi \neq 0, \\ 1 - \exp\left(-\frac{y}{\beta(u)}\right) & \forall \xi = 0, \end{cases} \quad (25)$$

Where $y \geq 0$ when $\xi \geq 0$; and, $0 \leq y \leq -\beta(u)/\xi$. If $\xi > 0$, it is said that the distribution is in the Fréchet domain, when $\xi < 0$ it is said to be in the Weibull domain, and when $\xi \to \infty$ it is said to be in the Gumbel domain.

Next, the copula $\bar{C}_{\lambda_j}(u^{s&p}_j, v^{IPC}_j)$ to be estimated, is the function that links the tail distributions $u^{s&p}_j$ and $v^{IPC}_j$ of the column vectors $\mathcal{W}^{s&p}_j$ and $\mathcal{W}^{IPC}_j$, respectively, at the time scale $\lambda_j$ or detail $D_j$. For example, $\bar{C}_{\lambda_1}(u^{s&p}_1, v^{IPC}_1)$ is the copula between the tail distributions of the highest frequency components of the S&P 500 and IPC returns contained in the column vectors $\mathcal{W}^{s&p}_1$ and $\mathcal{W}^{IPC}_1$, respectively, in a time frame spanning between $\lambda_j=2^{l-1}=1$ and 2 days.

The copula to be estimated at time scale $\lambda_j$ is said to have lower tail dependence if,

$$\lim_{u^{s&p}_j \to 0} \frac{C_{\lambda_j}(u^{s&p}_j, v^{IPC}_j)}{u^{s&p}_j} = \eta_l > 0, \quad (26)$$

and upper tail dependence if,
\[
\lim_{u_{j,\text{S&P}} \to 1} \frac{C_{\lambda_j} \left( u_{j,\text{S&P}}, v_{j,\text{IPC}} \right)}{1 - u_{j,\text{S&P}}} = \eta_u > 0, \tag{27}
\]

Where \( \tilde{C}_{\lambda_j}^* \) is the survival function of the copula \( C_{\lambda_j} \). The dependence structure coefficients \( \eta_l \) and \( \eta_u \) are estimated following Joe (1997),

\[
\hat{\eta}_l = \lim_{q \to 0^+} \frac{\hat{C}_n(q, q)}{q}, \tag{28}
\]

\[
\hat{\eta}_u = \lim_{q \to 1^-} \frac{1 - 2q + \hat{C}_n(q, q)}{1 - q}, \tag{29}
\]

where \( \hat{C}_n(q, q) \) is the empirical copula defined as

\[
C_n \left( \frac{i}{n}, \frac{j}{n} \right) = \sum_{p=1}^{i} \sum_{q=1}^{j} c_n \left( \frac{p}{n}, \frac{q}{n} \right), \tag{30}
\]

and \( c_n \) is the empirical copula frequency defined as

\[
c_n \left( \frac{i}{n}, \frac{j}{n} \right) = \begin{cases} 
1/n, & \text{if } (x_i, y_j) \text{ is an element within the sample}, \\
0, & \text{otherwise}. 
\end{cases} \tag{31}
\]

5. Empirical Results

5.1 Multiresolution Analysis

Equity indices returns decomposition is based on equation (19) and on its non-decimated version (20). Table shows the S&P 500 and IPC decomposed descriptive statistics in seven resolution levels under the MODWT and the Daubechies LA (8) wavelet or Symmlet (symmetric wavelet) as the filtering function. As shown by skewness, IPC has outperformed the S&P 500 at shorter time spans and both markets have shown the same negative performance at loner ones; as compared to the global approach (see Table 1), this result may explain the existence of more negative returns than positive ones by the S&P 500 when a detailed analysis is performed.
The decomposition results may be used as a *timing indicator* in asset allocation whenever seeking an *optimal* investment horizon. For example, if an investor had invested in the Mexican stock market in time spans above \( \lambda_5 = 2^{5-1} = 16 \) days, he/she would have experienced a poor performance compared to an investment horizon between time spans of one and two days, in which he/she would have well-performed.

As shown by kurtosis where a value greater than three indicates a leptokurtic distribution compared to the Gaussian one, the S&P 500 and IPC decomposed returns show a greater value than three from the lowest detail \( D_1 \) up to the sixth resolution level \( D_6 \); the latter related to a time scale that spans between \( \lambda_6 = 2^{5-1} = 2^3 = 32 \) and 64 days. At very short time horizons spanning between 1 and 4 days, the S&P 500 shows greater values of kurtosis than the IPC does, which means that the S&P 500 has experienced greater extreme values than what the IPC has done. In another way, the IPC is less fat-tailed than the S&P 500. But at higher resolution levels kurtosis value decreases on both decomposed equity indices returns indicating that the empirical probability distribution approaches a Gaussian one. Finally, the Jarque-Bera normality test at a 0.05 significance level indicates that the normality null hypothesis can be rejected at all resolution levels except for the IPC at the seventh level.

The results showed above are described by Cont (2001) as a stylized fact of asset returns related to heavy tails and *aggregational Gaussianity*. The first fact states that unconditional distributions of returns tend to show a tail index value greater than two but less than five, showed by a leptokurtic distribution. The second stylized fact states that whenever time scale increases, asset returns distribution looks more like a Gaussian one. Also Cont and Tankov (1999) describe that ‘prices move essentially by jumps at intraday scales, they still manifest discontinuous behavior at the scale of months and only after coarse-graining their behavior over longer time scales do we
obtain something that resembles a Brownian motion.’ This phenomenon is mainly observed in the IPC behavior and as by Tellez, Vargas and Hernandez (2009) at time scales greater than 32 weeks the IPC starts resembling a Brownian motion as shown by kurtosis values and quantile plots.

Also the correlation structure as measured by the linear relationship between two random variables can be broken down into different resolution levels, which in wavelets theory is known as wavelet correlation. The wavelet correlation is the correlation across paired resolution levels between two time series (or signals), in this case between the S&P 500 and IPC returns, and defined as:

$$\rho_{(S&P500,IPC)}(\lambda_j) = \frac{\gamma_{(S&P500,IPC)}(\lambda_j)}{\nu_{S&P500}(\lambda_j) \nu_{IPC}(\lambda_j)},$$

(32)

Where $\nu^2_{S&P500}(\lambda_j)$ and $\nu^2_{IPC}(\lambda_j)$ are the wavelet variances of the S&P 500 and IPC, respectively, related to the time scale $\lambda_j$ as defined in equation (21); and $\gamma_{(S&P500,IPC)}(\lambda_j)$ is the wavelet covariance.

Table 3 shows the correlation across each period paired resolution levels $j$ between the S&P 500 and IPC returns at seven resolution levels and the smooth detail. It is observed that at the shortest time frame spanning between 1 and 2 days the (linear) correlation between the S&P 500 and IPC returns is almost zero which means that at the highest frequency components or the highest equity indices variations spanning in time horizons between 1 and 2 days, the degree of association is quite weaker. This relationship strengthens when time scale increases and at a coarse-graining scale greater than $\lambda_7 = 2^7 = 128$ days the degree of association becomes the strongest one.3

The previous results show an asymmetric phenomenon in the correlation structure among time scales which can be exploited for risk diversification purposes. At short time frames spanning between one and two days we would expect an efficient portfolio diversification since co-movements as shown by the correlation (implicitly the covariance) is weaker for these time horizons. As time-span increases related at each level to lower frequency components than the previous level, co-movements among the equity indices returns strengthens showing a greater relationship that may cause great losses whenever markets are bearish.

3 See Appendix B for scatter plots.
Thus, an appropriate asset allocation strategy between the U.S. and the Mexican stock markets when considering timing that may diversify risk would be in short horizons with duration between one and at the most four days. A similar research by Ranta (2008) finds that correlations among main world equity indices volatilities are weaker in the lowest time scales which may result in an efficient diversification at short time investment horizons.

5.2 Dependence Structure Estimation

The correlation coefficient measures a linear relationship between random variables and it can be used as a dependency measure whenever their marginal distributions are Normal ones. But, what would be expected to the bivariate relationship under extreme values? Furthermore, does the relationship hold the same at lower and upper tails? Johnson and Wichern (2007) state that the covariance and correlation quantities can be sensitive to outliers and they could indicate association when in fact it occurs a little one. Also, as correlation is a global relationship measure nothing is said about in the tails whenever markets are bullish or bearish. By this is meant, do markets still hold their relationship whenever they crash or boom? Finally, is the tails dependence the same at different time frames?

The previous section showed that the correlation between the S&P500 and IPC is not the same at different time frames and this could be exploited in engineering efficient asset allocation strategies. This section shows the results about the relationship in the tails at different time frames by first fitting the Generalized Pareto Distribution (GDP) to lower and upper tails, and then estimating the empirical copula across the tails between the S&P 500 and IPC decomposed returns at each time scale, where the benchmark model is the empirical copula fitted to the original data.

Panel (a) in Figure 3 shows the marginal distributions of the S&P 500 and IPC original returns after being transformed into uniform distributions where it is observed that the empirical copula resembles more a Normal distribution. The empirical copula estimation can lead to state that the S&P 500 and IPC returns are not dependent in the lower and upper tails. This result is contrasted with Panel (b) in Figure 3 which depicts the dependence structure represented as the (copula) joint bivariate distribution obtained from the uniforms.
It is observed that at the lowest levels of the S&P 500 and IPC returns the probability of one market to crash given that the other market has crashed is almost zero, $P(M_1 = \text{crash} \, | \, M_2 = \text{crashed}) \approx 0$ or $P(M_2 = \text{crash} \, | \, M_1 = \text{crashed}) \approx 0$; inversely, the probability of one market to boom (at the upper levels) given that the other market has boomed is the same almost zero. Then it would be stated that the probability of extreme joint is zero given a GPD fitted to upper and lower tails.

Besides a low probability of extreme joint, there is still a probability of co-movement whenever markets are bullish or bearish (normal periods). Also it is observed the existence of a high probability of one market to boom when the other has crashed and one market to crash when the other has boomed. But the most joint movements are experienced during tranquil periods, i.e. at normal bulls and bears.

Finally, Panels (a) and (b) in Figure 4 present estimates of the lower and upper dependence structure parameter of the paired indices returns computed by equations (28) and (29), respectively. At the lower index, $q \to 0$ from the right, the dependence structure parameter takes lower values; at the upper index, as $q \to 1$ from the left, the dependence structure parameter takes lower values. Since these values showed to be lower around 0.20-0.30, it is confirmed independency in extreme joint.

After the S&P 500 and IPC returns were decomposed into seven resolution levels each one related to a different time scale, each decomposed series were fitted the GPD to the lower and upper tails, and finally estimated the empirical copula at each resolution level. Table 4 in Appendix C shows the estimations of upper and lower tails fitted by the GPD at raw returns and decomposed levels. Positive values mean the tail distribution is in the Fréchet domain, meanwhile negative values mean the tail distribution is in the Weibull domain. As stated by
Longin (1996) a Fréchet distribution suits fat-tailed distributions of returns meanwhile the Weibull distribution suits distribution returns which have no tails, where no tails means that any observations cannot be observed beyond the given threshold.

It can also be observed in Table 4 in Appendix C that as time is aggregated, upper and lower tails approach the Weibull distribution which confirms that as time scale increases the distribution of returns approaches a Gaussian one or identified as the *aggregational Gaussianity* stylized fact. So, in a global approach the distributions of the S&P 500 and IPC tails are better described as fat-tailed ones. But this result is mostly given by the fact that at very short time frames spanning up to 8 days, the tail distributions at each time scale are in the Fréchet domain.

Panel (a) in Figure 5 shows the scatter plot between the marginal distributions at the first resolution level $D_1$, where data resembles an Independent copula more than a Normal one. Panel (b) depicts the dependence structure expressed as the joint bivariate distribution obtained from the uniforms where the probability of extreme joints is almost zero as it happened in the original return series.

Figure 6 confirms a weaker relationship or independence at lower and upper tails in a time frame between 1 and 2 days. Then at very short time frames spanning between 1 and 2 days related to the highest variations of the S&P 500 and IPC values, the probability of one market to crash given that the other has crashed is almost zero as observed in their bivariate joint cumulative probability depicted in Panel (b) in Figure 5. As time-spans get greater, the empirical copula estimated begins to resemble a Gumbel one as shown by the scatter plots between marginal

---

4 See Appendix E for the empirical copula estimations on the remaining resolution levels.
distributions. Besides a weird bivariate joint distributions starting from resolution level $D_4$, the dependence structure parameter value takes its highest values at this level. Specifically at resolution levels $D_4$ and $D_6$ the dependence structure parameter takes values near 0.60. Panel (c) in Figure 11 in Appendix E shows that as $q \to 0$ from the right, the dependence structure parameter approaches its highest value near 1.0 and as $q \to 1$ from the left the dependence structure parameter approaches a value near 0.60.

The above phenomenon could be explained as follows. In a resolution level $D_4$ related to a time frame spanning between $\lambda_4 = 2^{4-1} = 2^3 = 8$ and 16 days, there is a higher probability to occur lower extreme joints between the S&P 500 and IPC than in time frames spanning in shorter days or above 16 days. So the dependence structure would predict greater downside risks in time scales between 8 and 16 days than in any other time scales.

In another way, the S&P 500 and IPC returns are being analyzed with different lenses each lens with a different resolution. The first lens with the maximum resolution is capturing the highest frequencies taking place in time frames spanning between $\lambda_1 = 2^{1-1} = 2^0 = 1$ and 2 days. At this resolution level there is zero probability of occurring extreme co-movements between the S&P 500 and IPC, also the degree of association is weaker between them. As the resolution decreases the time frame increases up to a fourth level in which the dependence structure becomes significant at the lower tails, which means that there is a high probability of one market to crash given that the other market has crashed in a time frame between $\lambda_4 = 2^{4-1} = 2^3 = 8$ and 16 days.

In the standpoint of view of investment portfolio management when considering timing in the process of asset allocation, it would be more efficient a strategy of very short time duration if the strategy is to diversify between the U.S. and Mexican stock exchanges. At very short time durations spanning between $\lambda_1 = 2^{1-1} = 1$ and 2 days, the degree of association between the stock exchanges is weaker than those of long time durations. Also at these short time durations the dependence structure of joints has shown to be of zero probability to occur which means that both stock exchanges are independent in the extremes (no tail dependency). But a strategy of
longer duration would result dangerous since the degree of association increases and there is a high probability of lower extreme joints starting from 8 days.

![Graph](image)

**Figure 6:** S&P 500 and IPC at level $D_1$ (a) Lower and (b) Upper dependence structure.

If the portfolio manager were to assess its position given that he/she has chosen a longer duration and the markets are bearish, then he/she would have experienced a greater loss than in a shorter duration strategy. Furthermore, if one market had crashed then it would have been highly probably that the other market had either crashed. This would result in greater losses since it has been detected a lower extreme dependence between the S&P 500 and IPC at time frames spanning between 8 and 16 days.

The opposite scenario when markets are bullish, if the timing strategy was defined as a longer horizon investment this could result in greater gains. But since the dependence structure has shown independence in upper extreme joints, then if a market has boomed this would not result in a booming in the other market. Then extraordinary returns in one market would not necessary result in extraordinary returns in the portfolio since both markets are not extremely co-moving.

### 6. Conclusions

Wavelets posses the ability to analyze temporal, non-stationarity and time-varying phenomena, phenomena which have characterized financial time series and identified as stylized facts. As contrary to Fourier analysis, wavelet analysis can analyze a time series in a frequency-time domain. The process of analyzing frequencies across different time scales is known as multiresolution analysis in which an original signal is decomposed into different resolution levels using a wavelet as filtering function, and each resolution level related to different time scales.
This paper applied a hybrid approach which combines wavelet and copula theory in such a way to analyze the dependence structure between the S&P 500 and IPC indices at different resolution levels. The analysis was centered in determining extreme joints at different time scales in which firstly the equity indices returns were decomposed into seven resolution levels. Then at each resolution level the series were transformed into uniform distributions and then fitted the Generalized Pareto Distribution to the upper and lower values. Finally it was estimated the empirical copula.

Results show an asymmetric co-movement between the equity markets in which the degree of association is weaker at lower time scales and as the resolution decreases and time scale increases the correlation starts increasing. Also as shown by the kurtosis, the S&P 500 index could be characterized as a more volatile market than the Mexican IPC since the S&P 500 has experienced more extreme values. The Jarque-Bera text showed that both equity indices are not Normal distributed and as time scale increases the empirical distribution approaches a Gaussian one, stylized fact identified as aggregational Gaussianity.

The copula estimation shows extreme independence at lower time scales but as resolution decreases the dependence structure becomes significant at lower extremes, which means a high probability of lower extreme joints that may cause potential losses specifically in investment horizons spanning between 8 and 16 days. Besides a low probability of extreme joints at lower time scales the extreme bivariate joint distribution varies at different time scales. This asymmetric dependence structure at different time frames could result in engineering efficient asset allocation strategies.

The importance of this paper takes place in investment portfolio management whenever a portfolio manager could consider timing in his/her decisions. In case of diversifying a portfolio between the U.S. and Mexican equities it should consider that correlation increases as time-spans increase. At longer time durations there is a high probability of extreme joints basically when markets are bearish. So, an efficient asset allocation would be that one which considers shorter time durations.

An extension of this empirical research would in the way to fitting the tails to known copula families such as the Gumbel or the Student-t distributions. Besides the dependence structure showed low values in the lower and upper tails, this dependence structure should be tested in a strict sense. Also an extension could be done to other equity markets or among individual stocks.
References


Appendix

A: Multiresolution Decomposition of Equity Indices Returns

Figure 7: Multiresolution Decomposition of Mexican IPC Returns based on Symmlet

Figure 8: Multiresolution Decomposition of US S&P 500 Returns based on Symmlet
B: Scatter Plots across S&P 500 and IPC Resolution Levels

Figure 9: Scatter Plots across S&P 500 and IPC Resolution Levels.
C: Estimation Outputs of Upper and Lower Tail Index Parameters

Table 4: Maximum Likelihood Estimation of Generalized Pareto Distribution Parameters.

<table>
<thead>
<tr>
<th></th>
<th>Parameter</th>
<th>U.S. S&amp;P 500</th>
<th></th>
<th>Mexico IPC</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Upper Tail</td>
<td>Lower Tail</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\xi)</td>
<td>(\beta(u))</td>
<td>MLE</td>
<td>(\xi)</td>
</tr>
<tr>
<td>Raw Returns</td>
<td>0.0829</td>
<td>0.0117</td>
<td>335.90</td>
<td>0.1705</td>
<td>0.0103</td>
</tr>
<tr>
<td>(D_1)</td>
<td>0.2422</td>
<td>0.0308</td>
<td>335.90</td>
<td>0.1211</td>
<td>0.0345</td>
</tr>
<tr>
<td>(D_2)</td>
<td>0.1960</td>
<td>0.0162</td>
<td>438.60</td>
<td>0.1262</td>
<td>0.0191</td>
</tr>
<tr>
<td>(D_3)</td>
<td>0.1711</td>
<td>0.0110</td>
<td>500.50</td>
<td>0.0620</td>
<td>0.0127</td>
</tr>
<tr>
<td>(D_4)</td>
<td>-0.0698</td>
<td>0.0083</td>
<td>579.10</td>
<td>0.1486</td>
<td>0.0068</td>
</tr>
<tr>
<td>(D_5)</td>
<td>0.0574</td>
<td>0.0049</td>
<td>((*))</td>
<td>0.1475</td>
<td>0.0041</td>
</tr>
<tr>
<td>(D_6)</td>
<td>-0.4738</td>
<td>0.0070</td>
<td>664.70</td>
<td>0.5778</td>
<td>0.0085</td>
</tr>
<tr>
<td>(D_7)</td>
<td>-0.9463</td>
<td>0.0023</td>
<td>902.30</td>
<td>-0.7956</td>
<td>0.0017</td>
</tr>
<tr>
<td>(S_7)</td>
<td>-0.9231</td>
<td>0.0011</td>
<td>((*))</td>
<td>-1.4526</td>
<td>0.0117</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw Returns</td>
<td>0.1460</td>
<td>0.0105</td>
<td>512.10</td>
<td>0.0249</td>
<td>0.0117</td>
</tr>
<tr>
<td>(D_1)</td>
<td>0.0393</td>
<td>0.0309</td>
<td>365.60</td>
<td>0.1218</td>
<td>0.0267</td>
</tr>
<tr>
<td>(D_2)</td>
<td>0.0654</td>
<td>0.0189</td>
<td>435.30</td>
<td>0.0501</td>
<td>0.0193</td>
</tr>
<tr>
<td>(D_3)</td>
<td>0.1726</td>
<td>0.0148</td>
<td>456.40</td>
<td>0.2102</td>
<td>0.0135</td>
</tr>
<tr>
<td>(D_4)</td>
<td>-0.1058</td>
<td>0.0112</td>
<td>540.10</td>
<td>-0.2011</td>
<td>0.0122</td>
</tr>
<tr>
<td>(D_5)</td>
<td>-0.3035</td>
<td>0.0058</td>
<td>669.30</td>
<td>-0.7641</td>
<td>0.0121</td>
</tr>
<tr>
<td>(D_6)</td>
<td>-0.3797</td>
<td>0.0084</td>
<td>623.40</td>
<td>-0.9367</td>
<td>0.0151</td>
</tr>
<tr>
<td>(D_7)</td>
<td>0.4759</td>
<td>0.0011</td>
<td>802.50</td>
<td>-0.5856</td>
<td>0.0040</td>
</tr>
<tr>
<td>(S_7)</td>
<td>-0.5935</td>
<td>0.0012</td>
<td>((*))</td>
<td>-1.1069</td>
<td>0.0107</td>
</tr>
</tbody>
</table>

(\(*\)) Means the estimation did not converge under the Maximum Likelihood Estimation (MLE) and the tail parameters were estimated under the Log Method of Moments (LMOM).
D: Marginal Distributions and Bivariate Joint Distributions

Figure 10: Fitted Copula GPD at Resolution Levels (left) Marginal distributions and (right) Bivariate joint distribution.
Cont. Figure 10: Fitted Copula GPD at Resolution Levels (left) Marginal distributions and (right) Bivariate joint distribution.
E: Dependence Structure Estimations

Figure 11: Lower and Upper Tail Dependence Structure Estimations based on the GPD.