DEPENDENCE STRUCTURES
FOR MULTIVARIATE HIGH-FREQUENCY
DATA IN FINANCE

Wolfgang Breymann
RiskLab, Department of Mathematics, ETH Zurich, Switzerland

Alexandra Dias and Paul Embrechts
www.math.ethz.ch/~embrechts
Department of Mathematics, ETH Zurich, Switzerland
ACKNOWLEDGEMENT

We thank Olsen Data for having provided us with the high-frequency data used in this study
OVERVIEW

• Motivation

• The data

• Deseasonalisation

• Dependence structure modelling for the whole dataset

• Tail dependence analysis

• Conclusion

For TECHNICAL DETAILS we refer to our PAPER!
MOTIVATION

• The Goal:
  Studying the dependence structure across time scales

• Why?
  – The change of the behavior as a function of the time horizon may contain important information
  – Improves extrapolation from small to large time horizons

• Requires:
  Characterising dependence for horizons from minutes to months

• Here:
  Restriction to high-frequency region (1 hour – 1 day)

• Peculiarities of high-frequency data are taken into account

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
THE DATA

• Tick–by–tick bid and ask quotes


• Collected and filtered by Olsen Data

• Irregularly spaced

• About 10 million data points for a single series

• Regularisation to 5 min. time series by linear interpolation

• Reduction to logarithmic middle prices:

\[ \xi_{\alpha,t} = \frac{\log (p_{\alpha,t}^{Bid} \cdot p_{\alpha,t}^{Ask})}{2} \]
FX PRICES FOR USD/DEM AND USD/JPY

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
DESEASONALISATION OF FINANCIAL DATA

- High frequency financial data present strong seasonalities
- Main periodicities: daily and weekly
- Seasonalities cover more subtle statistical properties
- Affected by Daylight Saving Time (DST)
- Theory of stochastic processes favors time transformation to an activity-based time scale, but:
  - Loss of synchronicity in the multivariate case
- Instead:
  - Volatility weighting based on weekly activity pattern
- Drawback: Aggregation property of returns has to be replaced by a more complicated relationship
HOURLY RETURNS

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
AUTOCORRELATION FUNCTIONS OF ABSOLUTE RETURNS

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
RETURN SCATTER PLOTS

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
MODELLING REQUIREMENTS

• The flexibility of modelling arbitrary patterns that display abrupt volatility changes (→ Japanese lunch break)

• Taking into account slow temporal changes in the habits of the market participants, institutional changes, etc

• Keeping track of Daylight Saving Time (DST) to take into account the one hour displacement between DST and non–DST periods

• The modelling of the geographical decomposition of market activity to take into account local public holidays and other irregularities
THE VOLATILITY PATTERN

- Integrated squared volatility: \( V_t^2 = \sum_{t' \leq t} (v_{t'}[\delta])^2 \)

- Volatility wrt horizon \( \Delta T \):
  \[
  \Delta V_t^2[\Delta T] \equiv V_t^2 - V_{t-\Delta T}^2 = \sum_{i=0}^{n-1} (v_{t-i\delta}[\delta])^2 = (v_t[\Delta T])^2
  \]

- Deseasonalised returns:
  \[
  x_t[\Delta T] = \frac{\xi_t - \xi_{t-\Delta T}}{\sqrt{\Delta V_t^2[\Delta T]}}.
  \]

- \( \delta = 5 \) minutes: elementary time step; \( n = \Delta T/\delta \)

- Aggregation property:
  \[
  x_t[\Delta T] = \frac{x_{t-\Delta T}[\Delta T_1]}{\sqrt{\Delta V_{t-\Delta T}[\Delta T_1]}} \sqrt{\Delta V_{t-\Delta T_2}[\Delta T_1]} + x_t[\Delta T_2] \sqrt{\Delta V_t^2[\Delta T_2]} \]
  \[
  \sqrt{\Delta V_t^2[\Delta T]}
  \]

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
COMPUTING THE VOLATILITY PATTERN

• Decomposition of the volatility:

\[ v_t^2[\delta] = a_t \left( v_\tau^{(d)}[\delta] \right)^2 \]

- with relative market activity factor \( a_t \) and

- volatility averaged over DST period \( d \) conditional to the time in the week, \( \tau = t \mod (1 \text{ week}) \):

\[ \left( v_\tau^{(d)}[\delta] \right)^2 = \frac{1}{N_d} \sum_{i=1}^{N_d} (r_{t_i + \tau}[\delta])^2 \]

• Weekend volatility:

\[ v^{(w)}[\delta] = \left| r_{t_w^{(end)}}[\Delta T_w] \right| \sqrt{\frac{\delta}{\Delta T_w}}, \]

- with weekend length \( \Delta T_w = t_w^{(end)} - t_w^{(start)} \)

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
THE WEEKLY VOLATILITY PATTERN

USD:DEM

USD:JPY

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
ACF OF DESEASONALISED HOURLY ABSOLUTE RETURNS

USD/DEM

USD/DEM and USD/JPY

USD/JPY and USD/DEM

USD/JPY

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
QQ-Plots for USD/DEM

Quantiles of Standard Normal

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
DEPENDENCE STRUCTURE MODELLING FOR THE WHOLE DATASET

- Exploratory analysis (scatter plots)
- Families of copulas across time scales
- Tail coefficient estimates
- Goodness of fit and ellipticality test
SCATTER PLOTS OF DESEASONALISED RETURNS

1 Hour returns

2 Hours returns

4 Hours returns

8 Hours returns

12 Hours returns

1 Day returns

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
COPULA DENSITY OF DESEASONALISED RETURNS

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
PSEUDO OBSERVATIONS WITH NORMAL MARGINS

1 Hour returns

2 Hours returns

4 Hours returns

8 Hours returns

12 Hours returns

1 Day returns

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
FAMILIES OF COPULAS

- Gaussian copula for correlation $\rho$:

$$C_{\rho}^{Ga}(u, v) = \int_{\infty}^{\Phi^{-1}(u)} \int_{\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi \sqrt{1 - \rho^2}} \exp \left\{ -\frac{(s^2 - 2 \rho s t + t^2)}{2(1 - \rho^2)} \right\} ds \, dt$$

- $t$–copula for $\nu$ degrees of freedom and correlation $\rho$:

$$C_{\nu, \rho}^{t}(u, v) = \int_{\infty}^{t_{\nu}^{-1}(u)} \int_{\infty}^{t_{\nu}^{-1}(v)} \frac{1}{2\pi \sqrt{1 - \rho^2}} \left\{ 1 + \frac{(s^2 - 2 \rho s t + t^2)}{\nu(1 - \rho^2)} \right\}^{-(\nu+1)/2} ds \, dt$$

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
FAMILIES OF COPULAS (CONT.)

- Gumbel copula:
  \[ C_{\beta}^{Gu}(u, v) = \exp \left[ - \left\{ (-\log u)^{1/\beta} + (-\log v)^{1/\beta} \right\}^{\beta} \right] \]

- Clayton copula:
  \[ C_{\beta}^{Cl}(u, v) = \max \left[ - \left\{ (-\log u)^{1/\beta} + (-\log v)^{1/\beta} \right\}^{\beta}, 0 \right] \]

- Frank copula:
  \[ C_{\beta}^{Fr}(u, v) = -\frac{1}{\beta} \log \left[ 1 + \frac{(e^{-\beta u} - 1)(e^{-\beta v} - 1)}{e^{-\beta} - 1} \right] \]
COPULA DENSITIES FOR SELECTED COPULAS

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
# Goodness of Fit and Ellipticality Test

<table>
<thead>
<tr>
<th>Time Freq.</th>
<th>Prob. Integral test</th>
<th>P–values for ellipticality test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>p–value</td>
</tr>
<tr>
<td>1 hour</td>
<td>$t$</td>
<td>0</td>
</tr>
<tr>
<td>2 hours</td>
<td>$t$</td>
<td>0</td>
</tr>
<tr>
<td>4 hours</td>
<td>$t$</td>
<td>0.01</td>
</tr>
<tr>
<td>8 hours</td>
<td>$t$</td>
<td>0.27</td>
</tr>
<tr>
<td>12 hours</td>
<td>$t$</td>
<td>0.19</td>
</tr>
<tr>
<td>1 day</td>
<td>$t$</td>
<td>0.74</td>
</tr>
</tbody>
</table>
PSEUDO OBSERVATIONS WITH $t$ MARGINS

1 Hour returns

2 Hours returns

4 Hours returns

8 Hours returns

12 Hours returns

1 Day returns

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
TAIL DEPENDENCE ANALYSIS

- Tail coefficient estimates

- Spectral measure estimation

- Multivariate excesses modelled by copulas
## TAIL COEFFICIENT ESTIMATES

<table>
<thead>
<tr>
<th>Frequency</th>
<th>d.f. $\hat{\nu}$</th>
<th>correl. $\hat{\rho}$</th>
<th>tail dep. coeff. $\hat{\lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 hour</td>
<td>4.339</td>
<td>0.563</td>
<td>0.273</td>
</tr>
<tr>
<td>2 hours</td>
<td>4.269</td>
<td>0.585</td>
<td>0.291</td>
</tr>
<tr>
<td>4 hours</td>
<td>4.282</td>
<td>0.599</td>
<td>0.299</td>
</tr>
<tr>
<td>8 hours</td>
<td>4.833</td>
<td>0.619</td>
<td>0.287</td>
</tr>
<tr>
<td>12 hours</td>
<td>5.438</td>
<td>0.623</td>
<td>0.264</td>
</tr>
<tr>
<td>1 day</td>
<td>5.712</td>
<td>0.624</td>
<td>0.254</td>
</tr>
</tbody>
</table>

Tail coefficient estimates for the DEM and JPY bivariate returns for the different time frequencies considered

$$\lambda = \lim_{\alpha \to 1^-} P(X_2 > F_2^{-1}(\alpha) | X_1 > F_1^{-1}(\alpha))$$
SPECTRAL MEASURE ESTIMATION

• Suppose that the $d$-dimensional random vector $X$ has a regularly varying tail with tail index $\alpha$.

• Limit behaviour of $X$ (vague convergence):

$$\frac{P(\|X\| > tx, X/\|X\| \in \cdot)}{P(\|X\| > t)} \xrightarrow{v} x^{-\alpha}P(\Theta \in \cdot),$$

with $x > 0$, $t \to \infty$, and $\Theta$ random vector on the space $(\mathbb{S}^{d-1}, \mathcal{B}(\mathbb{S}^{d-1}))$.

• Distribution function of $\Theta$ is SPECTRAL DISTRIBUTION of $X$.

• Alternatively:

  $\exists$ measure $\nu$ and positive sequence $(a_n)$, $a_n \to \infty$, such that

  $$nP(a_n^{-1}X \in \cdot) \xrightarrow{v} \nu(\cdot) \text{ for } n \to \infty$$

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
SPECTRAL MEASURE FOR DIFFERENT HORIZONS

1 Hour returns

2 Hours returns

4 Hours returns

8 Hours returns

12 Hours returns

1 Day returns

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
MULTIVARIATE EXCESSES MODELLED BY COPULAS

• Extreme tail dependence copula relative to a threshold $t$:
  \[ C_t(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v)|U \leq t, V \leq t) \]
  with conditional distribution function
  \[ F_t(u) := P(U \leq u|U \leq t, V \leq t), \quad 0 \leq u \leq 1 \]

• Archimedean copulas: \( \exists \) cont., strictly decreasing function \( \psi : [0, 1] \mapsto [0, \infty] \) with \( \psi(1) = 0 \), s.t.
  \[ C(u, v) = \psi^{-1}(-1)(\psi(u) + \psi(v)) \]

• For “sufficiently regular” Archimedean copulas (Juri and Wüthrich (2002)):
  \[ \lim_{t \to 0^+} C_t(u, v) = C_\alpha^{Clayton}(u, v) \]
MULTIVARIATE EXCESSES (CONT:)

- Data:
  1 hour pseudo-returns of DEM and JPY, \((\hat{F}_{1n}(x_{1i}), \hat{F}_{2n}(x_{2i}))\)

- For several thresholds \(t\) for hourly pseudo-returns we modelled
  
  \[
  C_t^{-}(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v) | U \leq t, V \leq t)
  \]
  and
  
  \[
  C_t^{+}(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v) | U \geq t, V \geq t)
  \]

- Thresholds:
  0.03, 0.05, 0.07, 0.1, 0.2, 0.3, 0.7, 0.8, 0.9, 0.93, 0.95, 0.97
BIVARIATE EXCESSES FOR DIFFERENT THRESHOLDS

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
COPULA DENSITY OF BIVARIATE EXCESSES

©2003 (W. Breymann, A. Dias, P. Embrechts ETH-Zürich)
AIC VALUES FOR DIFFERENT THRESHOLDS

G Gumbel
g Gumbel(-X)
C Clayton
c Clayton(-X)
t t-copula
CONCLUSION

• 2-dim. high-frequency data have been deseasonalised by means of volatility weighting

• The dependence structure for 2-dim., high-frequency FX return data has been analysed

• Methods used:

  – Copula modelling

  – Statistical techniques for extremal clustering
CONCLUSION: THE OVERALL PICTURE

• $t$–copulas with successively higher degrees of freedom work best for the whole dataset

• However, $t$–copulas have not enough structure for the shortest time horizons

• Test for ellipticality only rejected for 1 hour and 2 hours returns if the margins are transformed to $t$-distributions with the number of degrees of freedom adjusted to the result of the copula fit

• With the empirical margins, ellipticality rejected for horizons of 8 hours and shorter

• Extreme tails best described by Clayton resp. survival Clayton copula, as predicted by theory
CONCLUSION: FINAL

- This is a first analysis of the bivariate case
- The paper raises a lot of questions

For example:

- Further details on the two-dimensional stylized facts
- What about temporal interdependence
- High-dimensional data, beyond two
- Multivariate deseasonalisation