Dynamics of FX correlations

A minimum spanning tree approach

Andrea Politano

March 13, 2017

Università di Pisa

Table of contents

Introduction

Financial Markets as Complex Systems

Financial Markets (FM)

- **Evolving complex systems** with many interacting entities
- Organized in **hierarchical structures**

Hierarchical arrangement

- Can be found by studying **correlations** between asset returns
- Using a **correlation-based metric** we construct a complete graph whose nodes are the traded assets
- We can extract a **minimum spanning tree (MST)** from which we can identify clusters of elements.

FX market

The foreign exchange market (FX) is a global decentralized market for the trading of currencies.

- The main characteristics of FX market are
  - High liquidity
  - Strong presence of institutional investors
  - Geographical dispersion (OTC market)
  - Continuous operation

Data

The dataset consists of 2969 **daily exchange rates** for $N = 45$ currencies traded in the FX market from Jan 05 2004 to Dec 31 2015 (12 years). *Special drawing right* (XDR) are used as the numeraire.

Europe		America & Oceania	
GBP	British Pounds	CAD	Canadian Dollars
HRK	Croatian Kuna	BRL	Brazilian Reals
CZK	Czech Koruna	ARS	Argentine Pesos
EUR	European Euros	USD	U.S. Dollars
HUF	Hungarian Forint	COP	Colombian Pesos
ISK	Icelandic Krona	JMD	Jamaican Dollars
NOK	Norwegian Kroner	MXN	Mexican Pesos
PLN	Polish Zloty	PEN	Peruvian New Soles
RON	Romanian Leu	CLP	Chilean Pesos
RUB	Russian Ruble	NZD	New Zeland Dollar
SEK	Swedish Krona	FJD	Fijian Dollars
CHF	Swiss Francs	AUD	Austrialan Dollars

Data

Asia		Africa & middle east	
CNY	Chinese Renminbi	DZD	Algerian Dinar
INR	Indian Rupiah	EGP	Egyptian Pound
IDR	Indonesian Rupiah	GHS	Gahanaian Cedis
JPY	Japanese Yen	ILS	Israeli Shekels
MYR	Malaysian Ringgit	ZAR	South Africa Rand
PKR	Pakistani Rupees	TND	Tunisian Dinars
SGD	Singapore Dollars	TRY	Turkish Lira
KRW	South Korea Won		
LKR	Sri Lankan Rupees		
TWD	Taiwanese Dollars		
THB	Thai Bath		
VND	Vietnamese Dong		

Correlations and dynamic asset trees

Return correlations

- Data divided into $M = 237$ two-week stepped windows of width $T = 588$ days.
- Closure ex-rate of the i-th currency at time t by $P_i(\tau)$
- Log-returns given by $r_i(\tau) = \ln\left(\frac{P_i(\tau)}{P_i(\tau-1)}\right)$

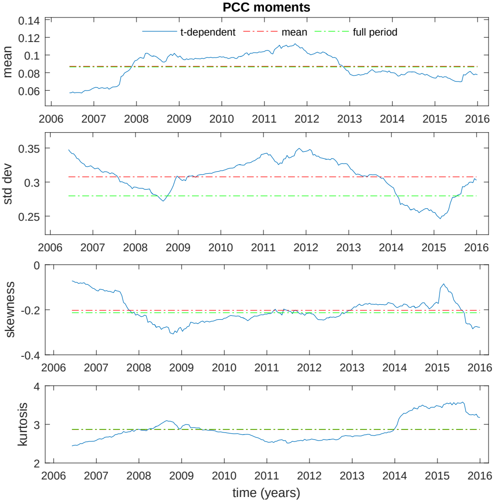
In order to characterize the synchronous time evolution of assets, we use the equal time **correlation coefficient** between asset "i" and "j" defined as

$$\rho_{ij}^t = \frac{\left\langle r_i^t r_j^t \right\rangle - \left\langle r_i^t \right\rangle \left\langle r_j^t \right\rangle}{\sqrt{\left[\left\langle r_i^{t2} \right\rangle - \left\langle r_i^t \right\rangle^2\right] \left[\left\langle r_j^{t2} \right\rangle - \left\langle r_j^t \right\rangle^2\right]}}$$

- These correlation coefficients fulfill the condition $-1 \leq \rho_{ij}^t \leq +1$ and form $M = 237$ $N \times N$ correlation matrices C^t

Correlations moments

Let us first characterize the correlation coefficient distribution by its first four moments



- Increse in the mean from January 2007
- Mean higher than average between 2008 and 2013
- Skewness always smaller than zero
- Kurtosis greater than three after 2014

MST construction

In order to construct the MST we have 3 steps to perform

The MST is a simply connected graph that connects all N nodes of the graph with $N-1$ edges such that the sum of all edge weights $\sum_{d_{ij}^t \in T^t} d_{ij}^t$ is minimum.

MST construction

In order to construct the MST we have 3 steps to perform

1. Calculate the correlation matrix for each time window. This results in $M = 237$ 45×45 matrices C^t.

The MST is a simply connected graph that connects all N nodes of the graph with $N - 1$ edges such that the sum of all edge weights $\sum_{d_{ij}^t \in T^t} d_{ij}^t$ is minimum.

MST construction

In order to construct the MST we have 3 steps to perform

1. Calculate the correlation matrix for each time window. This results in $M = 237$ 45×45 matrices C^t.
2. Derive a set of matrices D^t through the correlation-based metric $d_{ij}(t) = \sqrt{2\left(1 - \rho_{ij}(t)\right)}$ with $0 \leq d_{ij}(t) \leq 2$

The MST is a simply connected graph that connects all N nodes of the graph with $N - 1$ edges such that the sum of all edge weights $\sum_{d_{ij}^t \in T^t} d_{ij}^t$ is minimum.

MST construction

In order to construct the MST we have 3 steps to perform

1. Calculate the correlation matrix for each time window. This results in $M = 237$ 45×45 matrices C^t.
2. Derive a set of matrices D^t through the correlation-based metric $d_{ij}(t) = \sqrt{2\,(1 - \rho_{ij}(t))}$ with $0 \leq d_{ij}(t) \leq 2$
3. Construct a set of MST T^t starting from the matrices D^t

The MST is a simply connected graph that connects all N nodes of the graph with $N - 1$ edges such that the sum of all edge weights $\sum_{d_{ij}^t \in T^t} d_{ij}^t$ is minimum.

Normalied tree length

As a measure of the temporal state of the market we define the normalized tree length as $L(t) = \frac{1}{N-1} \sum_{d_{ij}^t \in T^t} d_{ij}^t$

- Average tree length drop after October 2008
- Standar deviation increase in the same period
- Shrinking of clusters and stretching of the tree

Tree occupation and central vertex

Mean occupation layer

In order to characterize the spread of nodes we introduce the mean occupation layer (MOL)

$$l(t, v_c) = \frac{1}{N-1} \sum_{i=1}^{N} L(v_i^t)$$

- $L(v_i^t)$ denotes the level of the vertex v_i in relation to the central vertex v_c. Three alternative definitions for the **central vertex**

Mean occupation layer

In order to characterize the spread of nodes we introduce the mean occupation layer (MOL)

$$l(t, v_c) = \frac{1}{N-1} \sum_{i=1}^{N} L(v_i^t)$$

- $L(v_i^t)$ denotes the level of the vertex v_i in relation to the central vertex v_c. Three alternative definitions for the **central vertex**

Mean occupation layer

In order to characterize the spread of nodes we introduce the mean occupation layer (MOL)

$$l(t, v_c) = \frac{1}{N-1} \sum_{i=1}^{N} L(v_i^t)$$

- $L(v_i^t)$ denotes the level of the vertex v_i in relation to the central vertex v_c. Three alternative definitions for the **central vertex**
  1. **Vertex degree criterion** – the node with the highest vertex degree

Mean occupation layer

In order to characterize the spread of nodes we introduce the mean occupation layer (MOL)

$$l(t, v_c) = \frac{1}{N-1} \sum_{i=1}^{N} L(v_i^t)$$

- $L(v_i^t)$ denotes the level of the vertex v_i in relation to the central vertex v_c. Three alternative definitions for the **central vertex**
  1. **Vertex degree criterion** – the node with the highest vertex degree
  2. **Weighted vertex degree criterion** – the node with the highest sum of correlation coefficient associated with the incident edges of the vertex.

Mean occupation layer

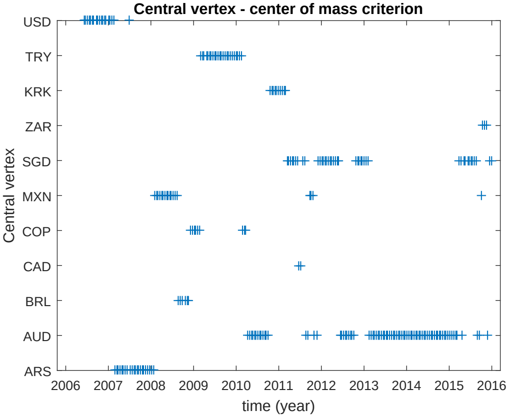
In order to characterize the spread of nodes we introduce the mean occupation layer (MOL)

$$l(t, v_c) = \frac{1}{N-1} \sum_{i=1}^{N} L(v_i^t)$$

- $L(v_i^t)$ denotes the level of the vertex v_i in relation to the central vertex v_c. Three alternative definitions for the **central vertex**
    1. **Vertex degree criterion** – the node with the highest vertex degree
    2. **Weighted vertex degree criterion** – the node with the highest sum of correlation coefficient associated with the incident edges of the vertex.
    3. **Center of mass criterion** – the node that produces the lowest value for mean occupation layer $l(t, v_c)$

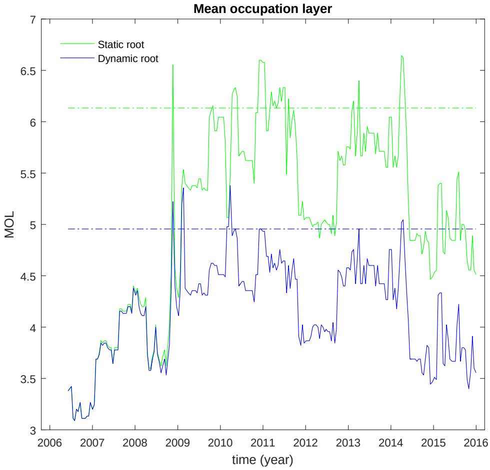
Mean occupation layer

- The first two criteria lead to the U.S. dollar (USD) as the central currency
- The vertex degree criterion leads to AUD dominating 34.6% of the time, followed by SGD at 17.7%, and USD at 10.1%



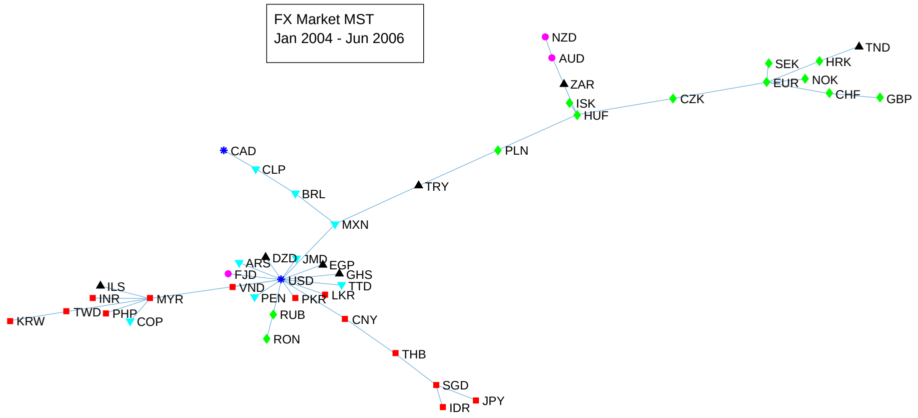
Mean occupation layer

- Increase in the MOL from 2006 to 2008
- Pronounced peak corresponding to October 2008
- MOL higher than average from 2009 to 2014
- High MOL reflect finer market structure, whereas low dips are connected to homogeneous behavior of the system



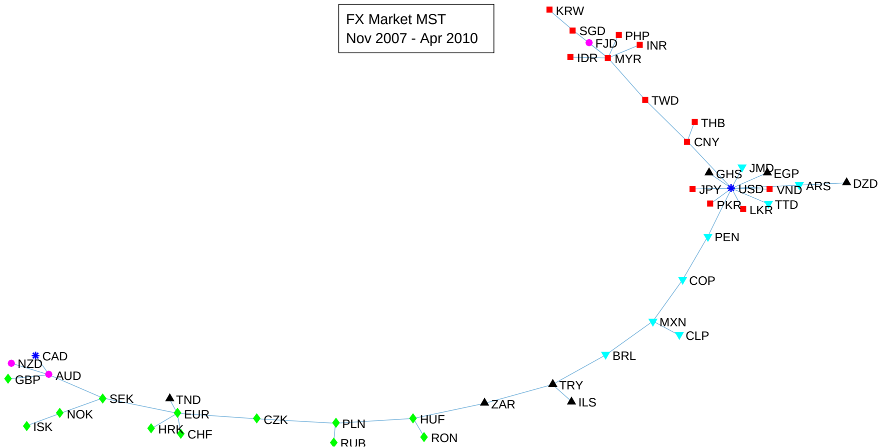
Tree clustering and economic interpretation

MST 2004-2006



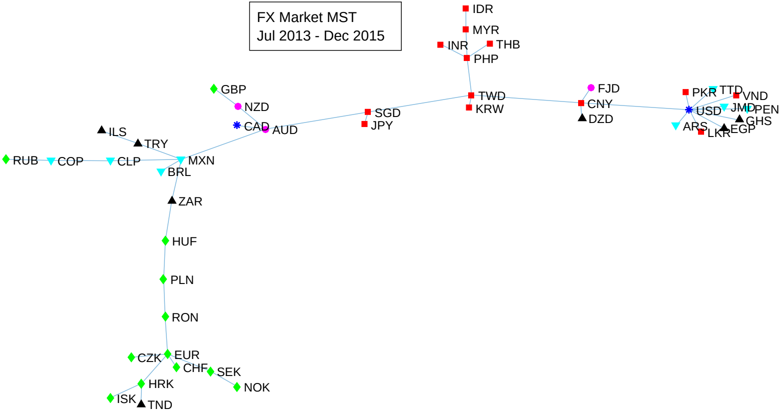
Cluster	Central node	L_{CN}	Cluster	Central node	L_{CN}
International	USD	1.01	Sout-east Asia	MYR	1.11
European	EUR	1.02	South-america	MXN	1.12
Indo-pacific	SGD	1.24	Tree	USD	0.91

MST 2007-2010



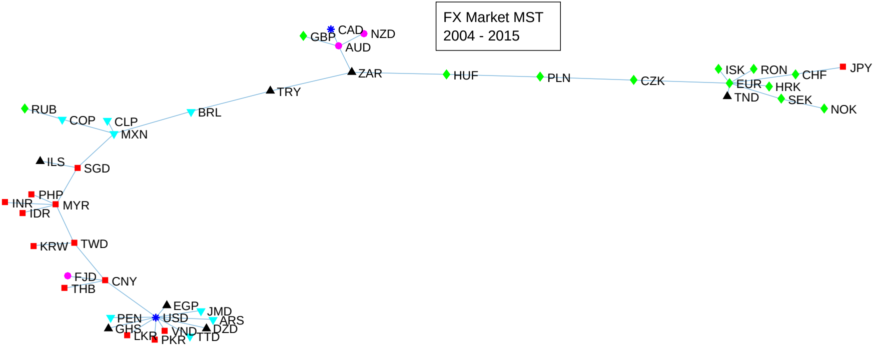
Cluster	Central node	L_{CN}	Cluster	Central node	L_{CN}
International	USD	0.94	Sout-east Asia	MYR	1.09
European	EUR	0.91	South-america	MXN	1.22
Commonwealth	AUD	0.94	Tree	COP	0.87

MST 2013-2015



Cluster	Central node	L_{CN}	Cluster	Central node	L_{CN}
International	USD	0.87	Sout-east Asia	TWD	1.09
European	EUR	0.91	South-america	MXN	1.02
Commonwealth	AUD	1.09	Tree	AUD	0.97

MST full period



Cluster	Central node	L_{CN}	Cluster	Central node	L_{CN}
International	USD	0.90	Sout-east Asia	MYR	1.06
European	EUR	0.95	South-america	MXN	1.09
Commonwealth	AUD	1.00	Tree	MXN	0.94

Asset tree evolution

Survival ratio

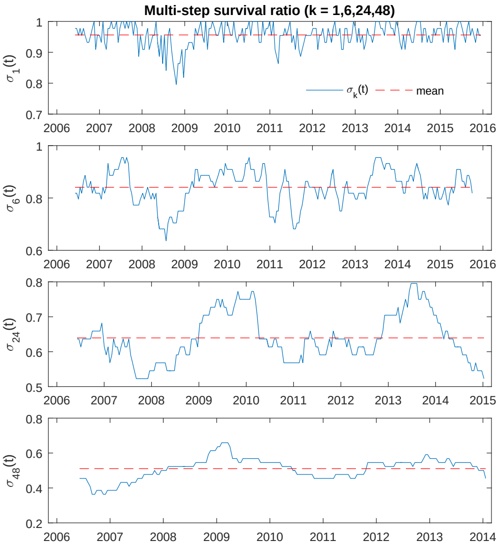
In order to investigate the robustness of asset tree topology, we define the **multi-step survival ratio** of tree edges as

$$\sigma_k(t) = \frac{1}{N-1}|E(t) \cap E(t-1)...E(t-k+1)|$$

- $\sigma(t)$ represents the fraction of edges found common in k consecutive trees at times $t...t-k$
- Under normal circumstances the tree for two consecutive time steps should look very similar
- Some of the differences can reflect real changes in the asset taxonomy

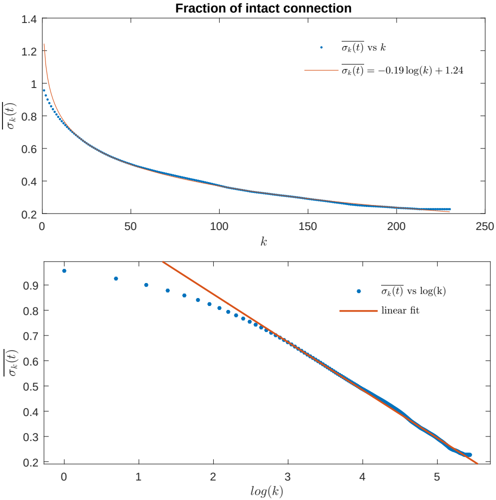
Single-step survival ratio

- $\overline{\sigma_1(t)} = 0.96$. A large majority of links survives from one window to the next
- A prominent dips corresponding to October 2008 indicate a strong tree reconfiguration taking place
- As might be expected, the ratio decreases with increases in step k.
- When the value of k increase the curve gradually becomes more smooth. system

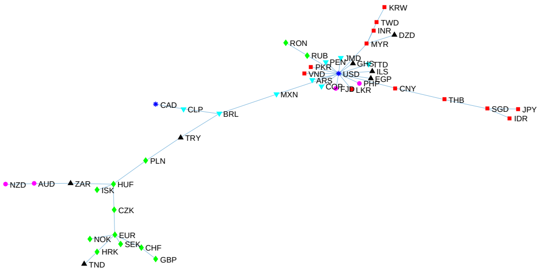


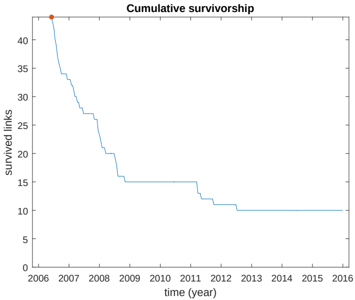
Connection decay

- The connections disappear quite slowly
- A small proportion of links remains intact creating a stable base for construction of the MST
- The existence of islands of stability is of importance for the construction of portfolios

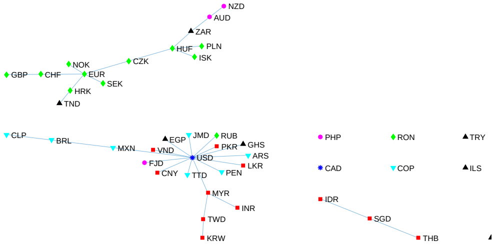


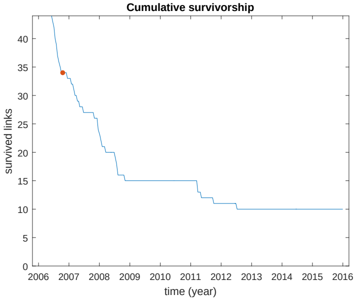
Survived links



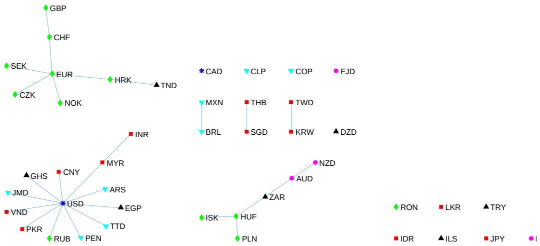


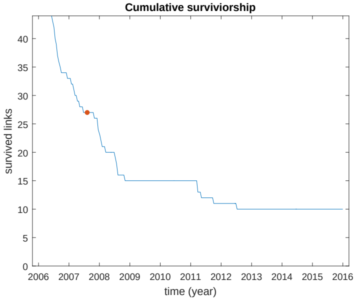
Survived links



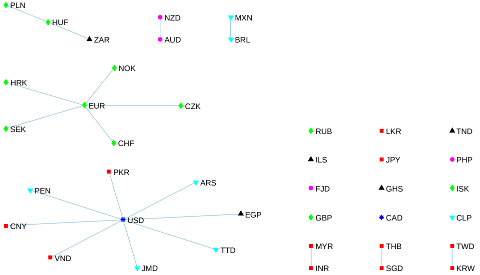


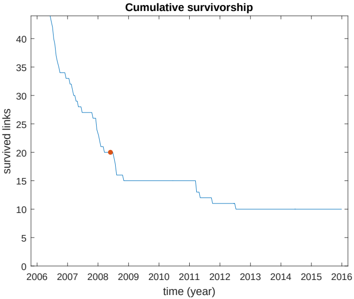
Survived links



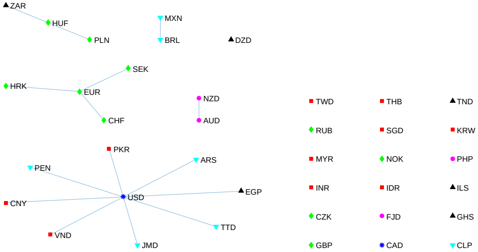


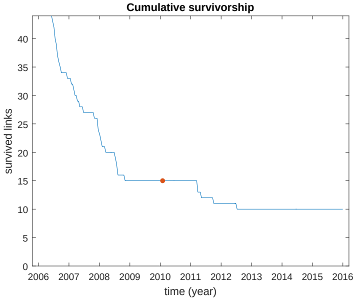
Survived links



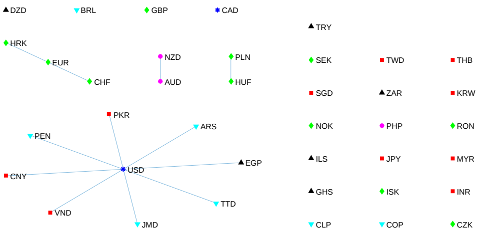


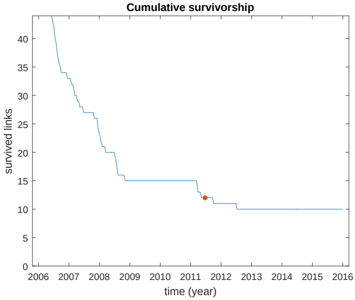
Survived links



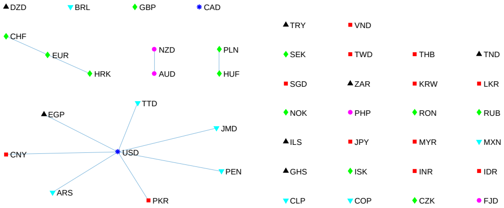


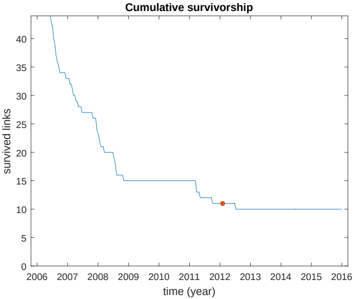
Survived links



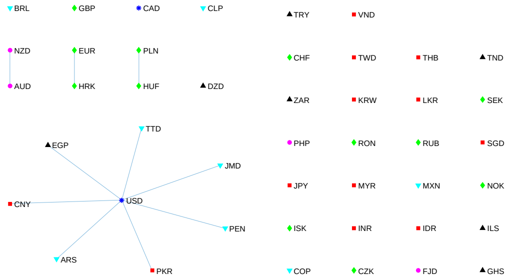


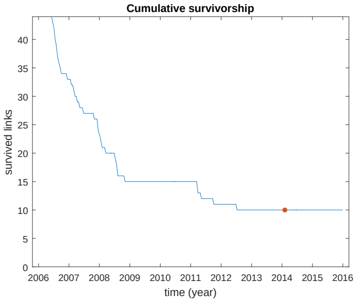
Survived links





Survived links





Summary and conclusions

Summary and conclusions

We have studied the distribution of correlation and the dynamics of asset trees.

- The tree evolves over time and the normalized tree length decreases and remains low during bear markets, thus implying the shrinking of the asset tree particularly strongly during a stock market crisis.
- We have also found that the mean occupation layer fluctuates as a function of time, and experiences an increse at the time of market crisis due to topological changes in the asset tree.
- The US dollar has been confirmed as the central currency of the asset tree

Finally we investigated the robustness of asset tree topology through the multi-step survivor ratio

- We observed a slow decays of $\overline{\sigma_k t}$ as k increase.
- A proportion of links remains intact as k increase, creating a stable base for construction of the MST

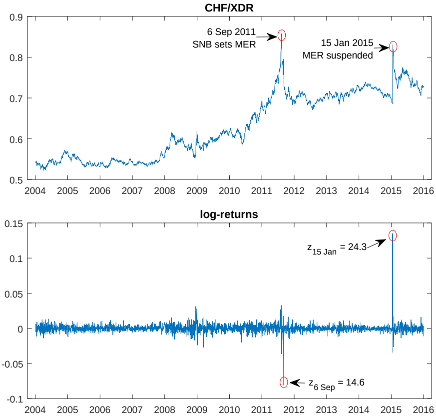
Thank you!

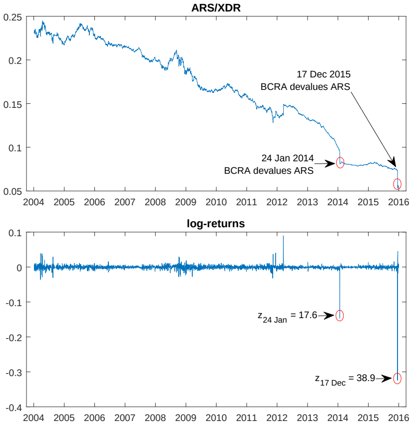
Appendix

Outliers in FX time series

Presence of **outliers** in the time series of FX returns.

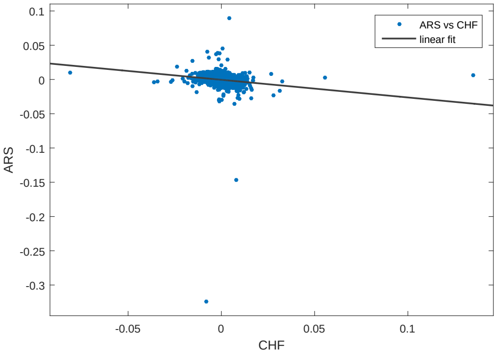
Outliers are due to central banks measures: devaluation, interest rates, setting of minimum exchange rate etc.





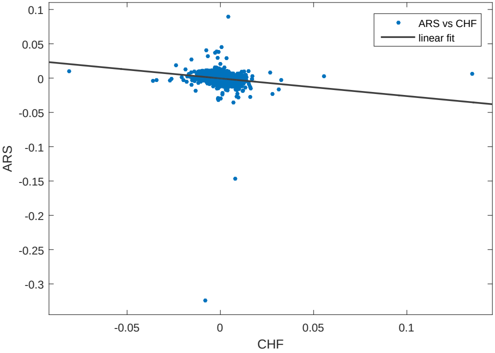
Outliers in FX time series

- PCC sensitive to outliers
  - Finite size breakdown point $B_p = \frac{1}{n}$



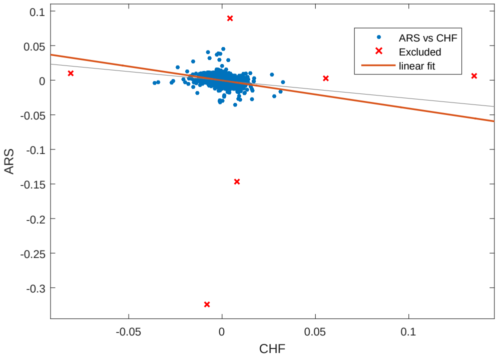
Outliers in FX time series

- PCC sensitive to outliers
  - Finite size breakdown point $B_p = \frac{1}{n}$
- $\rho_{full} = -0.17$



Outliers in FX time series

- PCC sensitive to outliers
  - Finite size breakdown point $B_p = \frac{1}{n}$
- $\rho_{full} = -0.17$
- $\rho_{clean} = -0.37$



Outliers in FX time series

Possible solutions to outliers sensitivity

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.

Outliers in FX time series

Possible solutions to outliers sensitivity

1. Ignoring the problem

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.

Outliers in FX time series

Possible solutions to outliers sensitivity

1. Ignoring the problem
2. Robust counterpart of the PCC

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.

Outliers in FX time series

Possible solutions to outliers sensitivity

1. Ignoring the problem
2. Robust counterpart of the PCC
3. Remove the outliers

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.

Outliers in FX time series

Possible solutions to outliers sensitivity

1. Ignoring the problem
2. Robust counterpart of the PCC
3. Remove the outliers

1. Morally reprehensible

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.

Outliers in FX time series

Possible solutions to outliers sensitivity

1. Ignoring the problem
2. Robust counterpart of the PCC
3. Remove the outliers

1. Morally reprehensible
2. No reason to underweight events

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.

Outliers in FX time series

Possible solutions to outliers sensitivity

1. Ignoring the problem
2. Robust counterpart of the PCC
3. Remove the outliers

1. Morally reprehensible
2. No reason to underweight events
3. Wisest solution in our opinion

We decided to reject events more than 30 MAD away from the median.
0.7% of the data rejected passing from the initial 2969 to 2948.