"Every why hath a wherefore."
- William Shakespeare

Causality in Finance

Ioana Boier

Quant Minds International, November 3, 2020

Disclaimer

The views and opinions expressed in this presentation are solely my own and do not reflect those of my employer.

Agenda

A Guided Tour of Causality with Applications to Finance:

- What is it?
- Why should we consider it?
- How to detect and measure it?
- How to use it (in Finance)?

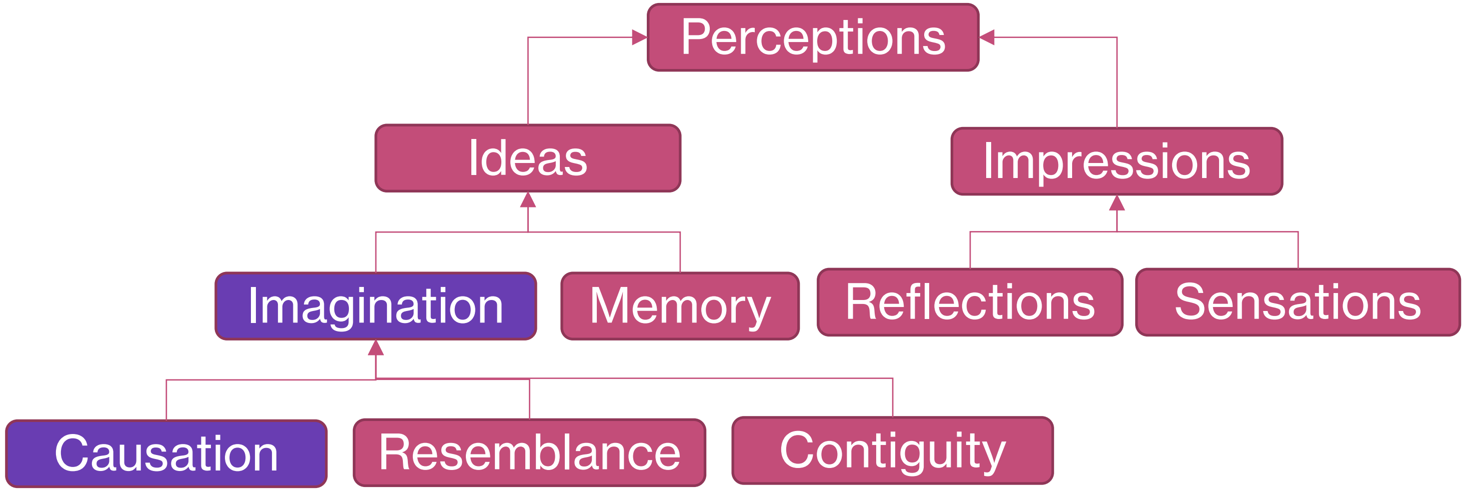
What Is Causality?

- The relationship between cause and effect.
  – *Wikipedia*

- The prerequisite and result of imagination.
  – *David Hume*

- The science of taking “WHY” seriously.
  – *Judea Pearl*

- If an *improbable coincidence* has occurred, there must exist a common cause.
  – *Hans Reichenbach*



Source: D. Hume, Treatise on Human Nature, 1739

Why Causality?

We live in an era of Big Data and Data Science

- Data is the “new oil” -----> “data economy”
- Data can tell us what happens, but not why something happens

  “Number of deaths of corona virus doubles every two days in Europe, every 3 days in the US, and every 10 days in South Korea.”
- Explain to justify and to trust

Correlations are a main staple of financial modeling. When they break, disaster may ensue.

- Diversification: the more correlated the investments, the greater the need for diversification
- Approximation: basis of regression (use the movements of certain market observables as a proxies for others)
- Decomposition: basis of factor analysis and cross-curve risk & PNL breakdowns
- Statistical arbitrage: identify closely correlated assets; wait until their time series diverge; short “winners”, buy “losers”; reverse & repeat
- ...

Why Causality?

1. A good decision made for the wrong reasons is still a wrong decision

2. Lack of correlation $\not\Rightarrow$ lack of causation

3. Finance (and Life): non-linear and directional

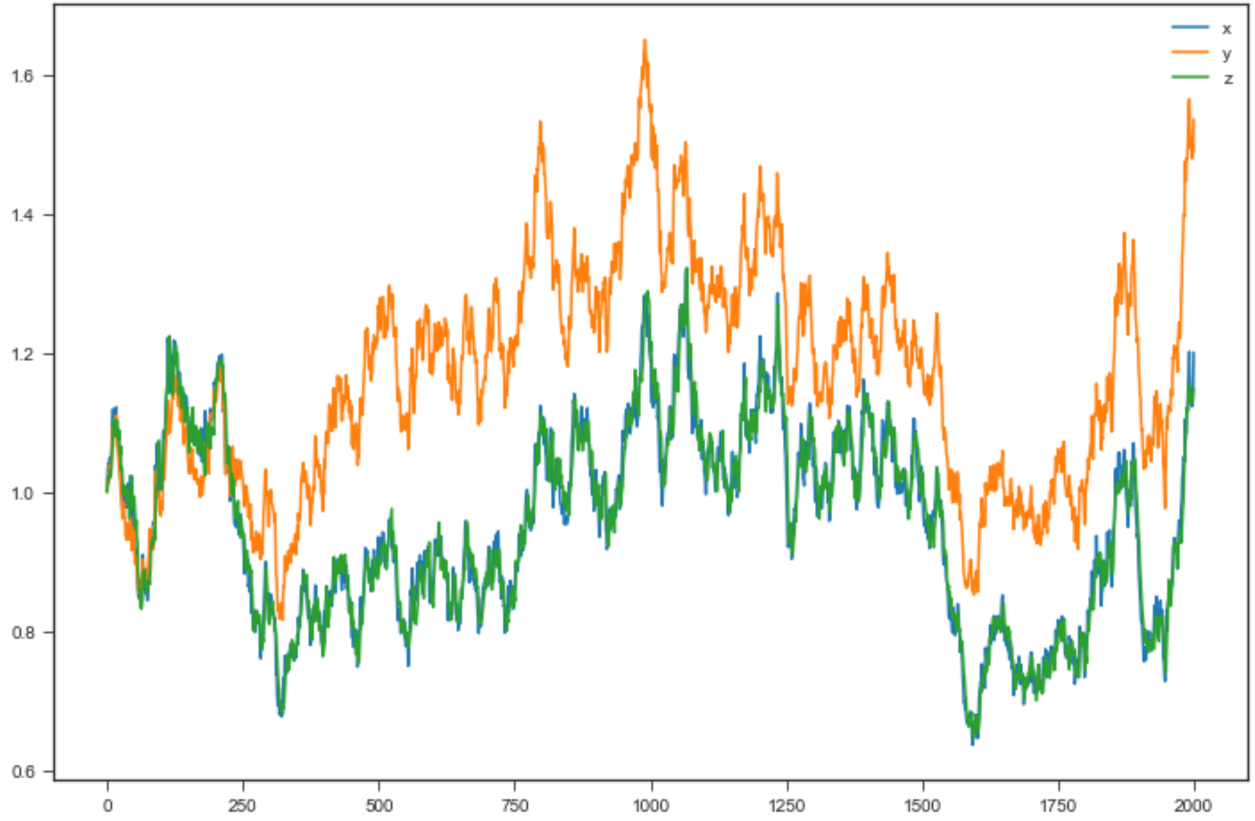
Lack of Correlation $\nRightarrow$ Lack of Causation

$$\frac{dX_t}{X_t} = \sigma_{X,t} dW_t$$

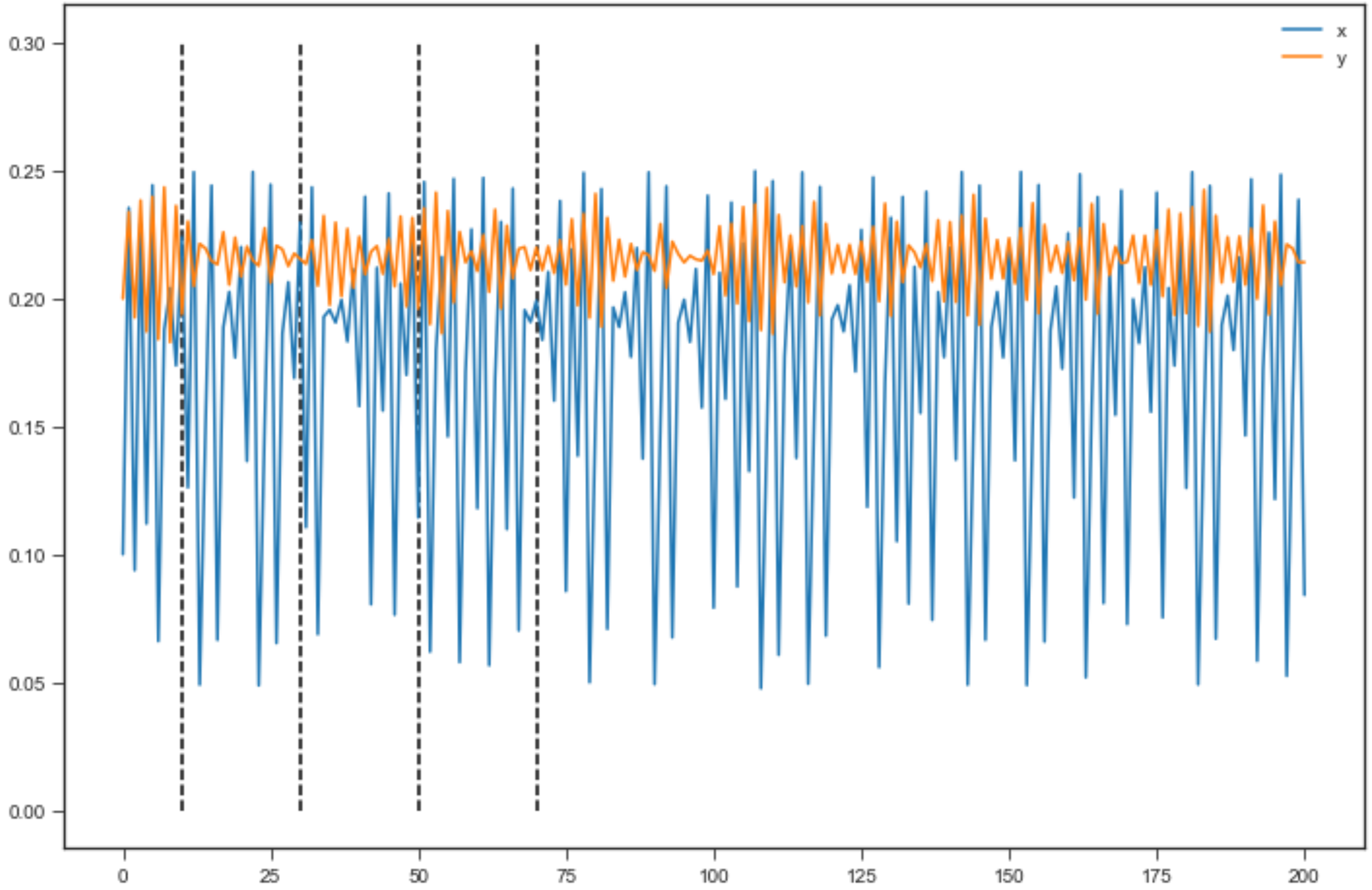
$$\frac{dY_t}{Y_t} = \sigma_{X,t} \rho dW_t + \sigma_{Y,t} \sqrt{1-\rho^2} dU_t$$

$$\frac{dZ_t}{Z_t} = \theta(\beta X_t - Z_t + \mu)dt + \sigma_{Z,t} dU_t$$

$< dX_t, dY_t > = 0.93, \; < dX_t, dZ_t > = 0.02$



Correlation & Non-Linearity

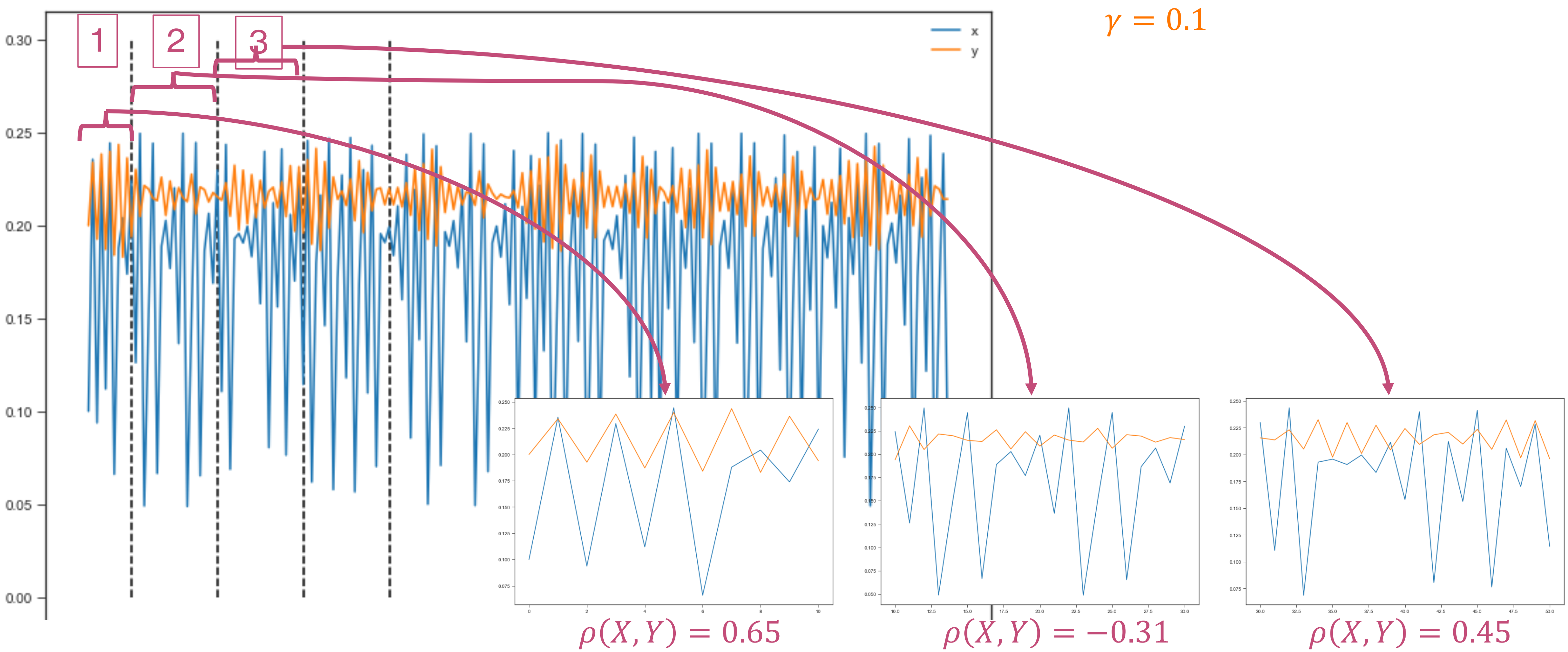


$$X_{t+1} = aX_t(1 - aX_t)$$

$$Y_{t+1} = bY_t(1 - bY_t - \gamma X_t)$$

Influence of X_t on Y_t controlled via parameter γ

Correlation & Non-Linearity



Methods

Randomized trials

- Randomly assign participants into experimental and control group(s)
- Experimental group receives intervention, control group does not
- The only expected difference between the control and experimental groups is the outcome variable being studied
- Drawbacks: expensive, time-consuming; works with averages; ethical issues

Observational data

- Observe without interfering, but how to get counterfactual outcomes?
- *Confounder*: affects both the independent and dependent variables → spurious correlations
- Theoretical underpinnings – identifiability conditions
- Statistical
- Information-theoretical
- State space (topological)
- Causal Bayesian networks
- ...

Statistical Methods

"For two simultaneously measured signals, if we can predict the first signal better by using the past information from the second one than by using the information without it, then we call the second signal causal to the first one." **- Wiener, 1956**

Lead-Lag

Intertemporal Cross-Correlation

- Linear (LICC)
- Non-linear (NICC)

$$NICC_{x,y} = \frac{\sum_t R_y(t+\tau)^2 - \sum_t (R_y(t+\tau) - \boldsymbol{f}(R_x(t)))^2}{\sum_t R_y(t+\tau)^2}$$

Joint Evolution

Cointegration

- Linear
- Non-linear

$$z_t = \boldsymbol{f}(y_t, x_t)$$

Regression

Granger Causality

- Linear
- Non-linear

$$y_t = \boldsymbol{f}_y(y_{t-1}, \ldots, y_{t-k}, x_{t-1}, \ldots, x_{t-j}; 1) + \varepsilon_t$$

Information-Theoretical Methods

Transfer entropy: amount of directed (time-asymmetric) transfer of information between two random processes *X* and *Y*

- Amount of information is measured using *Shannon's entropy*
- Generalizes Granger Causality (TE for jointly Gaussian ~ LGC)
- It usually requires more samples for accurate estimation than GC
- Variants: Partial TE, Symbolic TE, PSTE

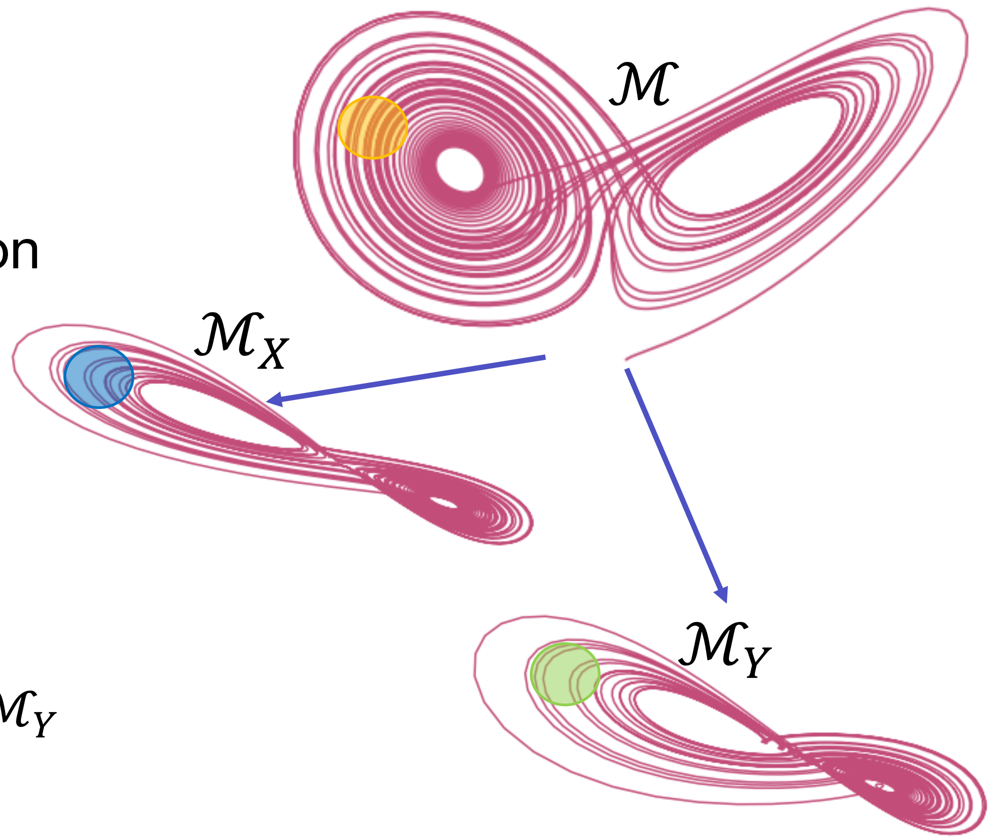
Mutual Information: amount of information shared by *X* and *Y*

- *MIME:* Mutual Information from Mixed Embedding
- Variant: *Partial MIME*

State Space Methods

Non-separable dynamical systems:

- Acyclic, separable model too simplistic
- Economy, climate, brain, etc. : bi-directional information flow and cyclic interaction
- *Cause-effect* becomes "*directed influence*": $X \leftrightarrows Y$



Topological causality:

- Takens' Theorem: $\mathcal{M}$ topologically equivalent to $\mathcal{M}_X$, $\mathcal{M}_Y$
- Convergent Cross Mapping (CCM)

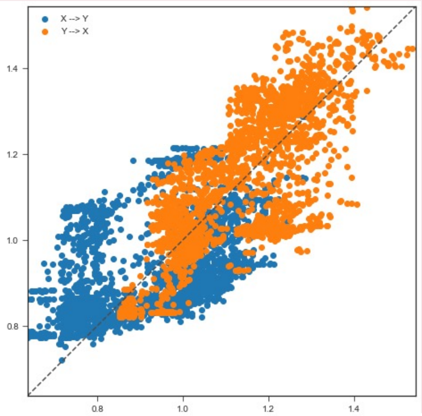
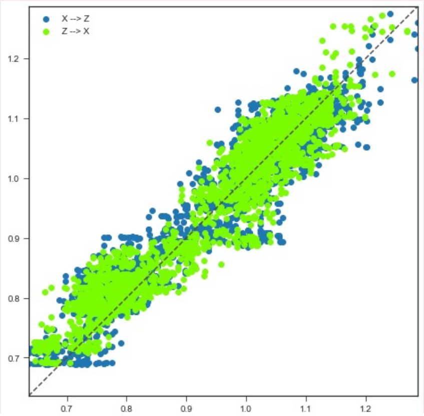
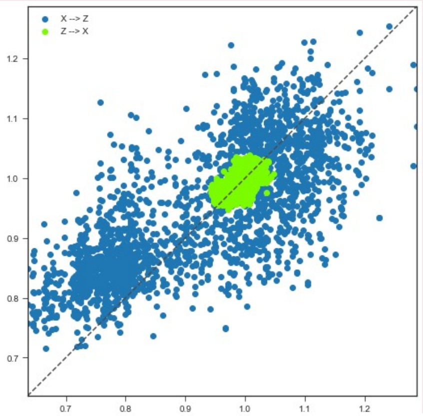
Causality In Practice

Not straightforward to measure, many methods, “open” results

1. Comparisons (magnitude / direction) => score
2. Visualization: causality networks
3. Network analysis:
   - Dependency / influence structure
   - Evolution over time
   - Predictive capabilities

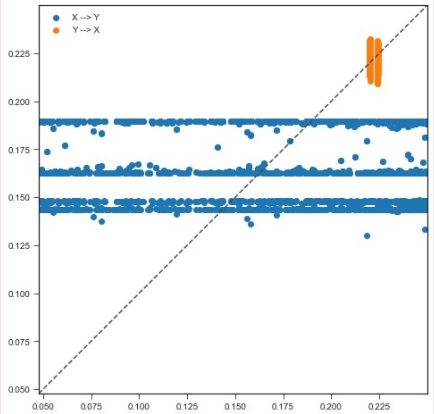
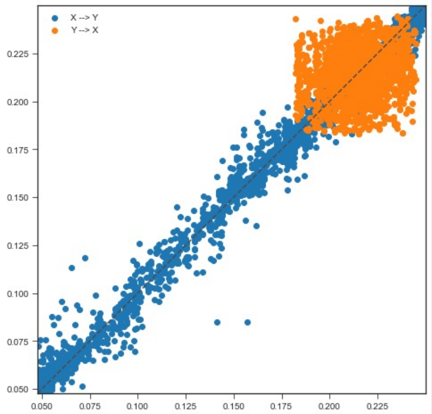
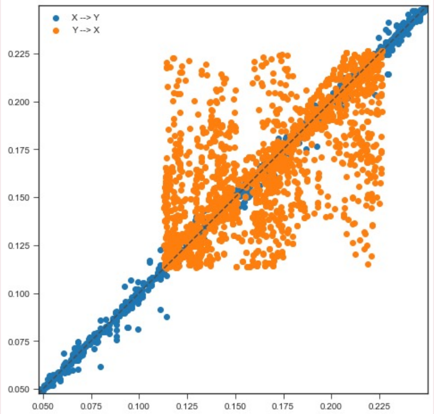
Causality In Practice

$$\frac{dX_t}{X_t} = \sigma_{X,t} dW_t \qquad \frac{dY_t}{Y_t} = \sigma_{X,t}\rho dW_t + \sigma_{Y,t}\sqrt{1-\rho^2}dU_t \qquad \frac{dZ_t}{Z_t} = \theta(\beta X_t - Z_t + \mu)dt + \sigma_{Z,t}dU_t$$

<table>
<tr><th>Synthetic example</th><th>X, Y</th><th>X, Z (β=1.0)</th><th>X, Z (β=0.1)</th></tr>
<tr><td>Granger (Chi^2)</td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>1.0000</td><td>0.7455</td></tr><tr><td>X</td><td>0.8266</td><td>1.0000</td></tr></table></td>
<td><table><tr><th></th><th>Z--></th><th>X--></th></tr><tr><td>Z</td><td>1.00</td><td>0.0</td></tr><tr><td>X</td><td>0.32</td><td>1.0</td></tr></table></td>
<td><table><tr><th></th><th>Z--></th><th>X--></th></tr><tr><td>Z</td><td>1.0000</td><td>0.0</td></tr><tr><td>X</td><td>0.2374</td><td>1.0</td></tr></table></td>
</tr>
<tr><td>CCM</td><td></td><td></td><td></td></tr>
</table>

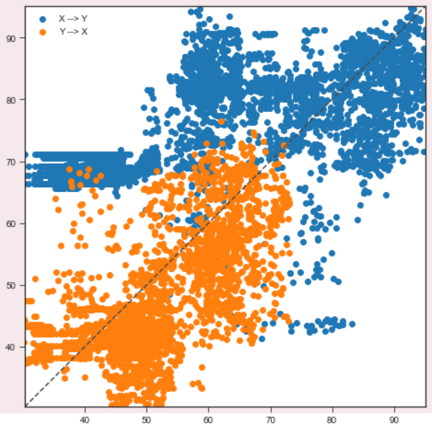
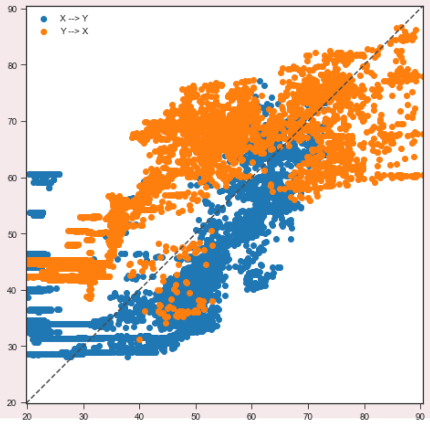
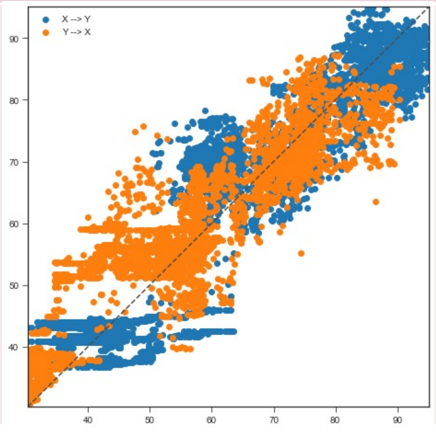
Causality In Practice

$X_{t+1} = aX_t(1 - aX_t)$ $Y_{t+1} = bY_t(1 - bY_t - \gamma X_t)$

<table>
<tr><th>Synthetic example</th><th>X, Y(γ=0.0)</th><th>X, Y (γ=0.1)</th><th>X, Y (γ=0.8)</th></tr>
<tr><td>Granger (Chi^2)</td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>0.0000</td><td>0.0</td></tr><tr><td>X</td><td>0.5999</td><td>1.0</td></tr></table></td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>1.0</td><td>0.0</td></tr><tr><td>X</td><td>0.0</td><td>1.0</td></tr></table></td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>1.0</td><td>0.0</td></tr><tr><td>X</td><td>0.0</td><td>1.0</td></tr></table></td></tr>
<tr><td>CCM</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

Causality In Practice

XOM, COP XLE

<table>
<tr><th>Synthetic example</th><th>X=XOM, Y=COP</th><th>X=COP, Y=XLE</th><th>X=XOM, Y=XLE</th></tr>
<tr><td>Granger (Chi^2)</td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>1.0</td><td>0.0</td></tr><tr><td>X</td><td>0.0</td><td>1.0</td></tr></table></td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>1.0000</td><td>0.0344</td></tr><tr><td>X</td><td>0.1375</td><td>1.0000</td></tr></table></td>
<td><table><tr><th></th><th>Y--></th><th>X--></th></tr><tr><td>Y</td><td>1.0</td><td>0.0</td></tr><tr><td>X</td><td>0.0</td><td>1.0</td></tr></table></td></tr>
<tr><td>CCM</td>
<td></td>
<td></td>
<td></td></tr>
</table>

Causality In Finance

Motivation:

- Better understanding of complex, non-linear interactions
- Impact analysis (not just correlation, not just linear)

Challenges:

- Process (autocorrelation, nonlinearity, different time scales, heteroskedasticity)
- Data (relevant variables/latent variables, time subsampling/aggregation, noise, selection bias)
- Computation (sample size, dimensionality, uncertainty estimation)

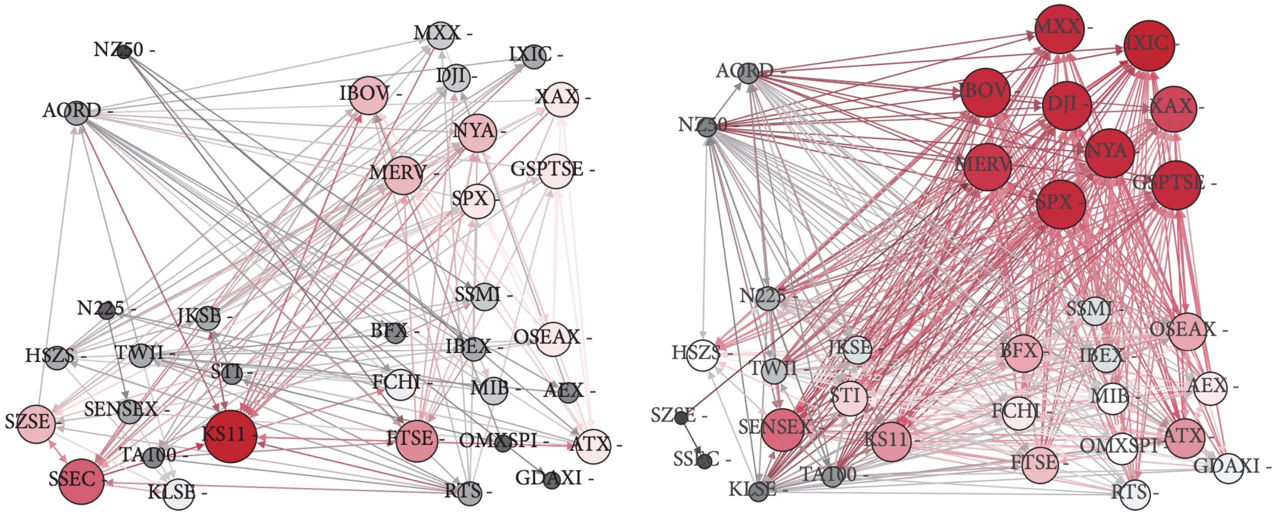
Example use cases:

- Market structure
- Dimensionality reduction / filtering
- Prediction
- ...

Market Structure

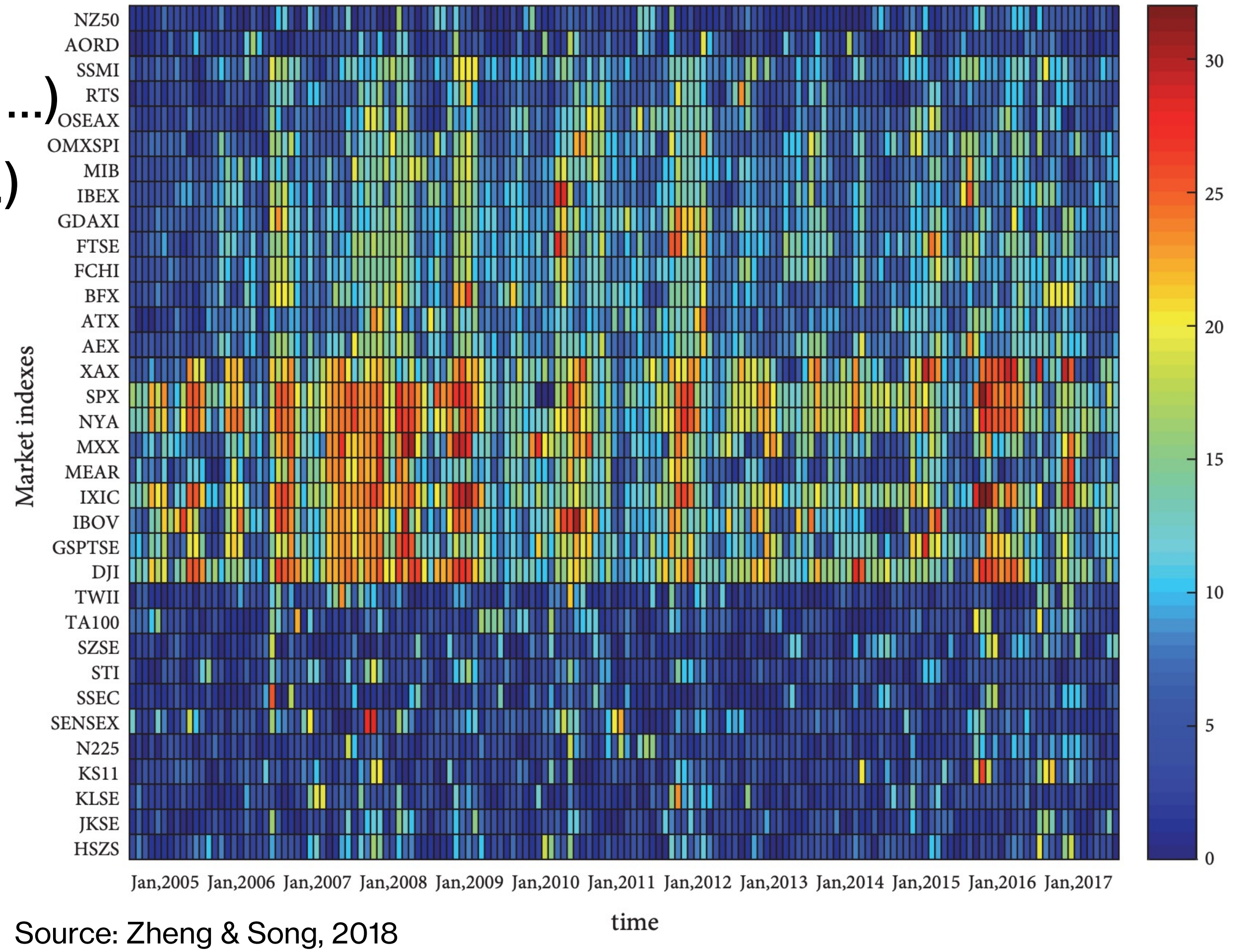
Connection models

- Between assets (gold, oil, stocks, bonds, …)
- Between economies (US, Asia, Europe, …)
- Snapshot in time vs. evolution



(a) April 2006 (minimum out-degree)

(b) September 2007 (maximum out-degree)



Source: Zheng & Song, 2018

Dimensionality Reduction / Filtering

Dimensionality reduction

- Large universe of assets ~ many possible connections
- Prune interactions of interest based on strength of causality links

Example: pairs trading (stocks)

Strategy: identify mean-reverting equilibrium relationship between pairs

- Cointegration filter – necessary condition
- Granger Causality filter – leader/follower pattern

Time Series Prediction

- State-of-the-art prediction methods are correlation-based
- Can causal relations improve prediction?

Source: Xu et al., 2020

Dataset		Exchange-Rate			Energy			Nasdaq		
Methods	Metrics	horizon 5	horizon 10	horizon 15	horizon 5	horizon 10	horizon 15	horizon 5	horizon 10	horizon 15
VAR	MAE	0.0065	0.0093	0.0116	3.1628	4.2154	5.1539	0.1706	0.2667	0.3909
	RAE	0.0188	0.0270	0.0339	0.0545	0.0727	0.0889	0.0011	0.0018	0.0026
	CORR	0.9619	0.9470	0.9318	0.9106	0.8482	0.7919	0.9911	0.9273	0.5528
CNN-AR	MAE	0.0063	0.0085	0.0106	2.4286	2.9499	3.5719	0.2110	0.2650	0.2663
	RAE	0.0182	0.0249	0.0303	0.0419	0.0509	0.0616	0.0014	0.0017	0.0017
	CORR	0.9638	0.9490	0.9372	0.9159	0.8618	0.8150	0.9920	0.9919	0.9860
RNN-GRU	MAE	0.0066	0.0092	0.0122	2.7306	3.0590	3.7150	0.2245	0.2313	0.2700
	RAE	0.0192	0.0268	0.0355	0.0471	0.0528	0.0641	0.0015	0.0015	0.0018
	CORR	0.9630	0.9491	0.9323	0.9167	0.8624	0.8106	0.9930	0.9901	0.9877
MULTIHEAD ATT	MAE	0.0078	0.0101	0.0119	2.6155	3.2763	3.8457	0.2218	0.2446	0.3177
	RAE	0.0227	0.0294	0.0347	0.0451	0.0565	0.0663	0.0014	0.0017	0.0027
	CORR	0.9630	0.9500	0.9376	0.9178	0.8574	0.8106	0.9945	0.9915	0.9857
LSTNET	MAE	0.0063	0.0085	0.0107	2.2813	3.0951	3.4979	0.1708	0.2511	0.2603
	RAE	0.0184	0.0247	0.0311	0.0393	0.0534	0.0603	0.0011	0.0016	0.0017
	CORR	0.9639	0.9490	0.9373	0.9190	0.8640	0.8216	0.9940	0.9902	0.9872
MLCNN	MAE	0.0065	0.0094	0.0107	2.4529	3.4381	3.7557	0.1301	0.2054	0.2375
	RAE	0.0189	0.0274	0.0312	0.0423	0.0593	0.0648	0.0009	0.0013	0.0016
	CORR	0.9693	**0.9559**	**0.9511**	0.9212	0.8603	0.8121	0.9965	0.9931	0.9898
TEGNN-NTE	MAE	0.0076	0.0093	0.0113	2.1753	2.8731	3.4122	0.1601	0.2174	0.2490
	RAE	0.0221	0.0290	0.0315	0.0369	0.0475	0.0588	0.0010	0.0014	0.0016
	CORR	0.9660	0.9531	0.9425	0.9210	0.8587	0.8167	0.9942	0.9907	0.9879
TEGNN-NCNN	MAE	0.0074	0.0096	0.0118	2.2346	2.7488	3.5229	0.1884	0.4454	0.3342
	RAE	0.0240	0.0350	0.0325	0.0575	0.0574	0.0673	0.0012	0.0029	0.0022
	CORR	0.9634	0.9518	0.9398	0.9196	0.8608	0.8121	0.9937	0.9909	0.9856
TEGNN	MAE	**0.0060**	**0.0083**	**0.0104**	**2.0454**	**2.7242**	**3.3232**	0.1549	0.1897	0.2358
	RAE	**0.0176**	**0.0243**	**0.0302**	**0.0358**	**0.0470**	**0.0573**	0.0010	0.0012	0.0015
	CORR	**0.9694**	0.9548	0.9438	**0.9267**	**0.8673**	**0.8221**	0.9951	0.9922	0.9887
TEGIN	MAE	0.0065	0.0089	0.0108	2.1768	2.8097	3.3572	**0.1174**	**0.1664**	**0.2043**
	RAE	0.0188	0.0259	0.0315	0.0375	0.0485	0.0579	**0.0008**	**0.0011**	**0.0013**
	CORR	0.9690	0.9551	0.9441	0.9204	0.8615	0.8131	**0.9968**	**0.9937**	**0.9907**

Using the Right Approach

Depends on:

- Data and its statistical properties
- Knowledge or assumptions about generating processes

Conclusions

- Causal relationships are an integral part of life
- We need better tools to model them in all areas, including Finance
- A path to AGI without causality is unlikely
- Causal modeling is the thread that connects many fields: statistics, machine learning, sequential experimentation, optimization, decision theory, game theory, auction design, computational cognitive psychology, etc.

References

1. D. Hume, *A Treatise on Human Nature*, 1739-1740
2. Judea Pearl, *The Book of Why*, 2018
3. C. Granger, Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 1969
4. B. M. Damghani et. al, *The Misleading Value of Measured Correlation*, Wilmott Magazine, 2012
5. F. Takens, *Detecting Strange Attractors in Turbulence*, Lecture Notes in Mathematics, 1981
6. G. Sugihara et al., *Detecting Causality in Complex Ecosystems*, Science, 2012
7. Q. Zheng and L. Song, *Dynamic Contagion of Systemic Risks on Global Main Equity Markets Based on Granger Causality Networks*, Discrete Dynamics in Nature and Society, 2018
8. H. Xu et al., *Multivariate Time Series Forecasting with Transfer Entropy Graph*, preprint, 2020