PRINCETON UNIVERSITY

ORF 474: High Frequency Trading
Spring 2020
Robert Almgren

Lecture 10b

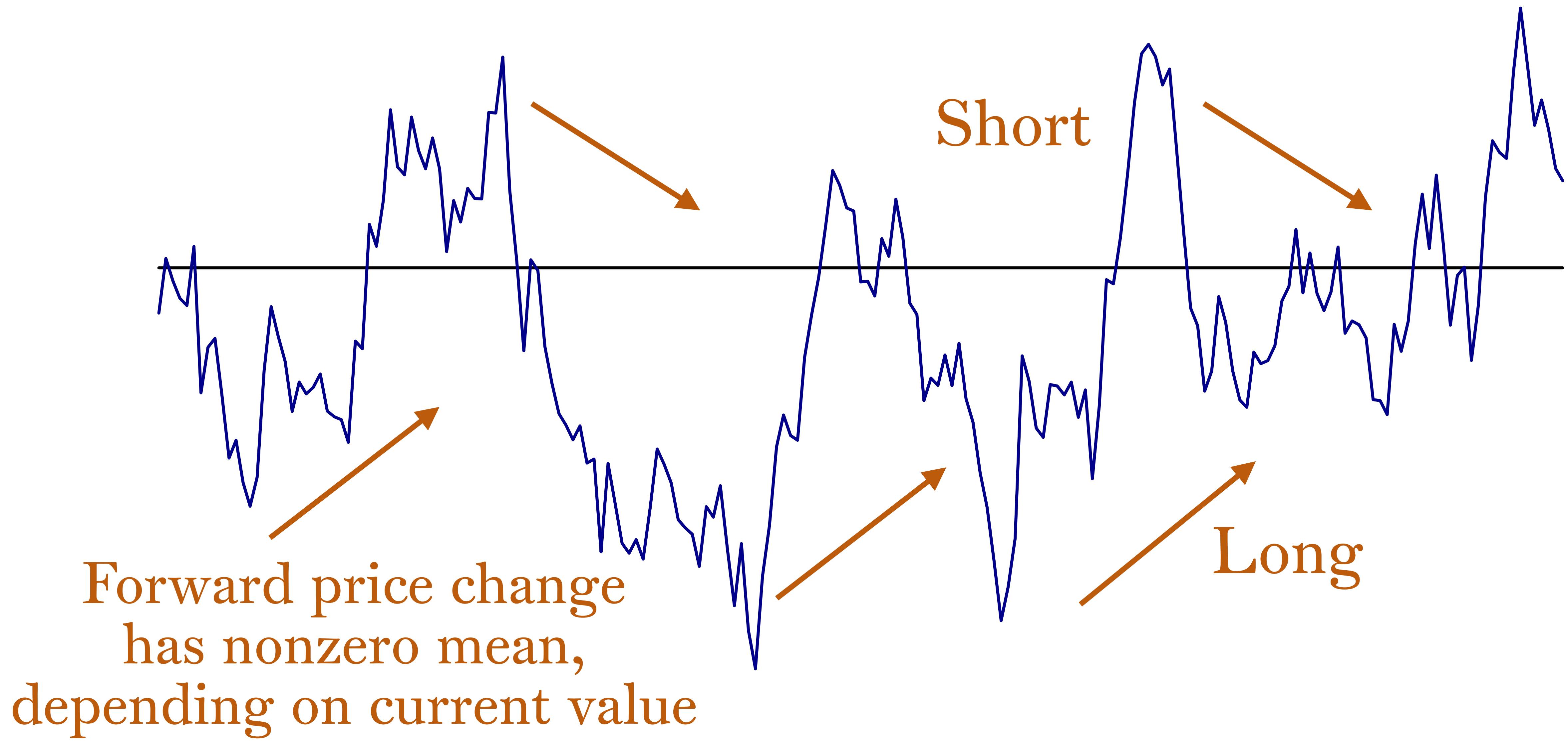
April 15, 2020

Outline

- Outline of reversion trading
- Cautionary example
- Cointegration and reversion
- General trading strategies
- Bubble example (momentum)
- Technical analysis

Main idea

- Identify mean-reverting tradeable process
- Trade to take advantage of reversion

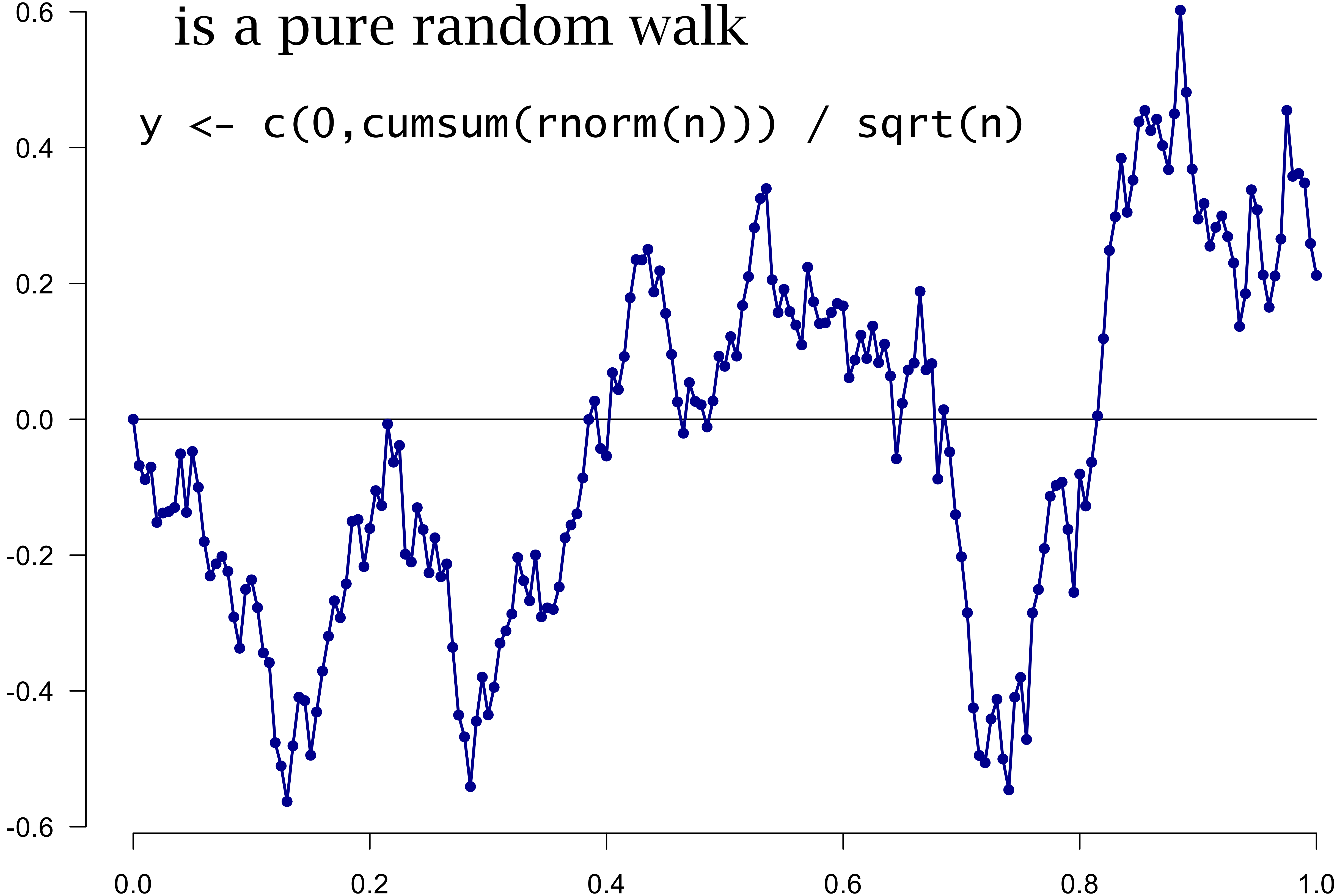


Cautionary example for averaging

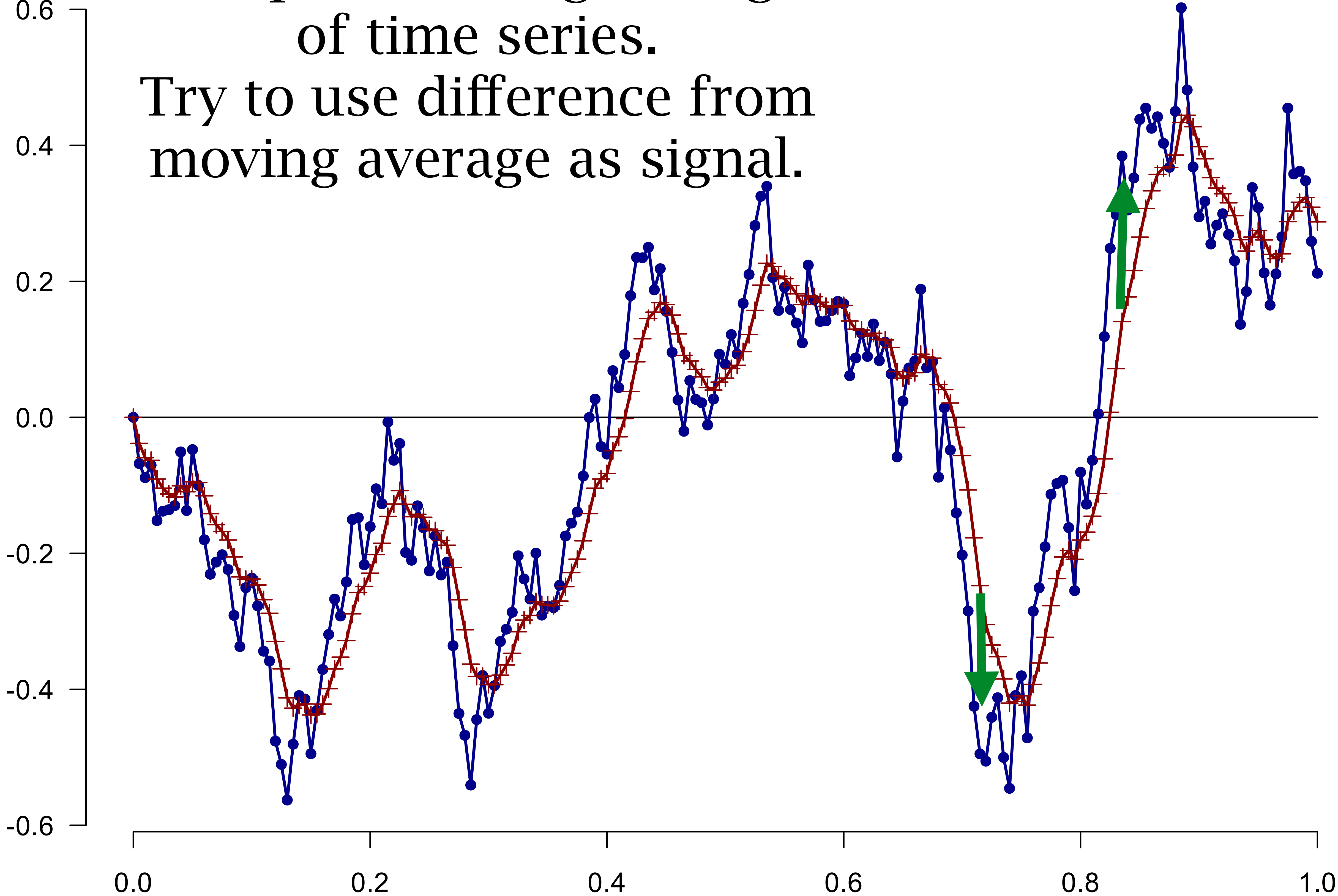
Homework on cointegration:

- Compute moving averages of prices, and covariance
- Take differences relative to moving average
- Use difference as predictor for prices

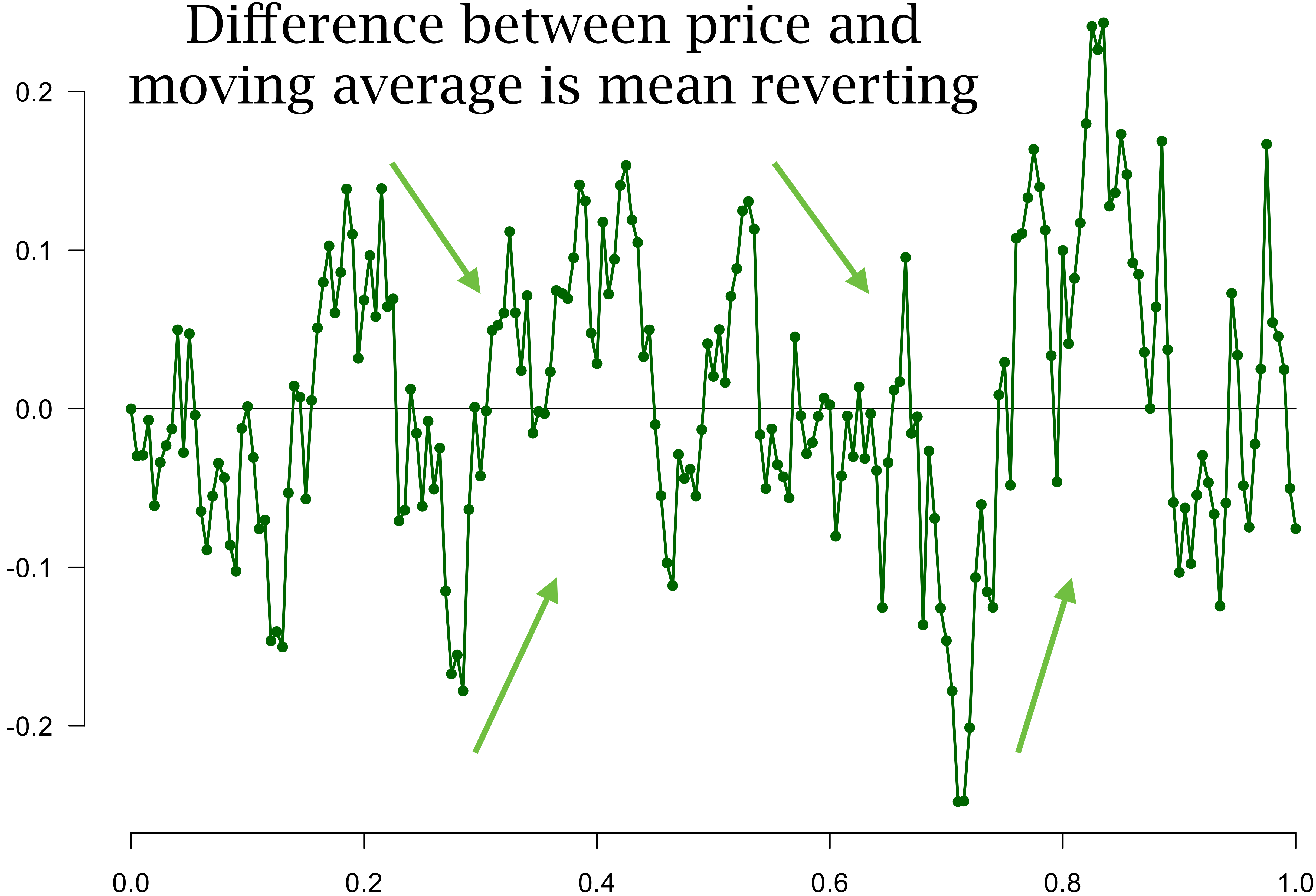
This time series is a pure random walk

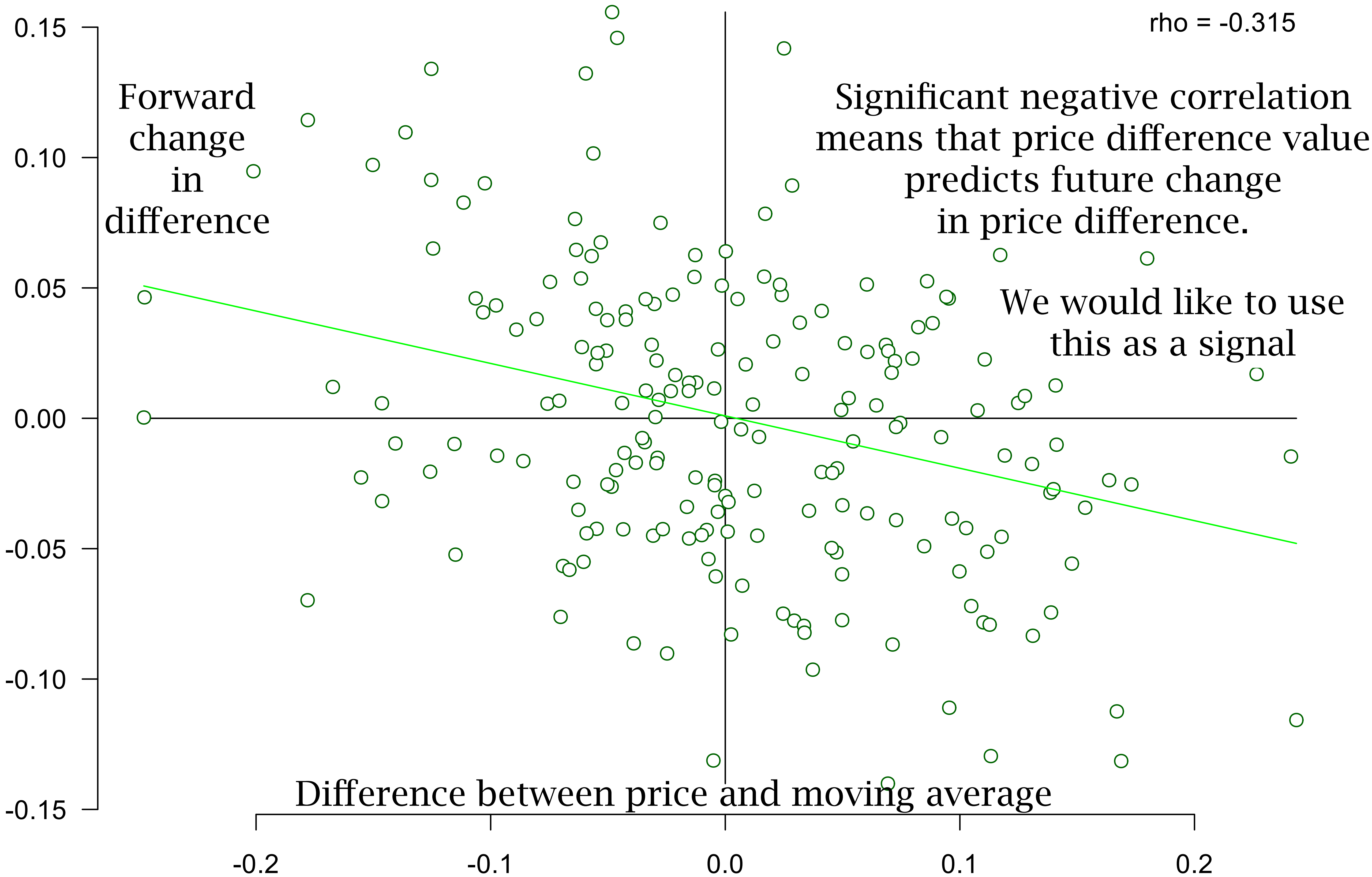


Compute moving average of time series. Try to use difference from moving average as signal.

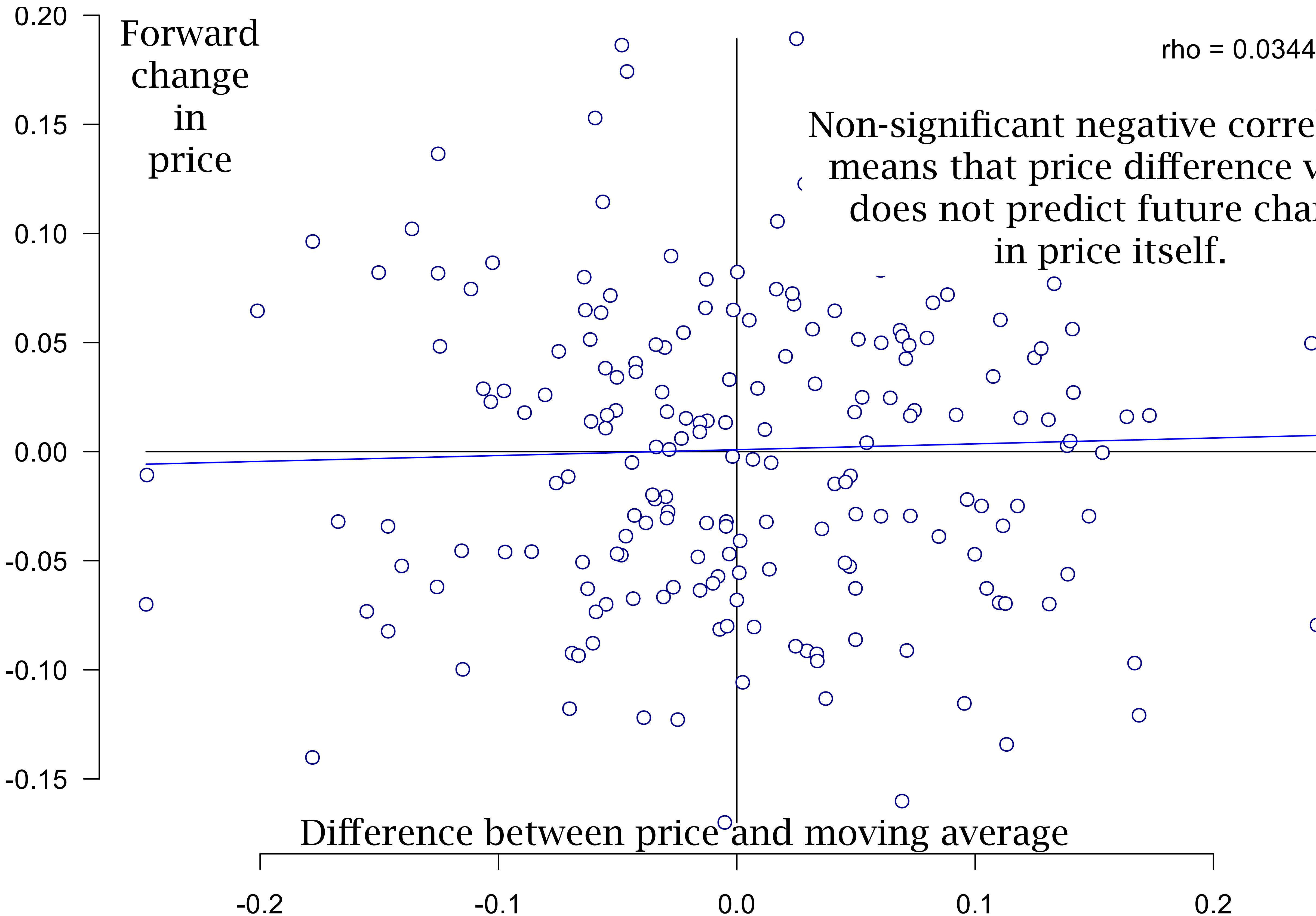


Difference between price and moving average is mean reverting





rho = -0.315
Forward change in difference
Significant negative correlation means that price difference value predicts future change in price difference.
We would like to use this as a signal
Difference between price and moving average
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.2
-0.1
0.0
0.1
0.2


Forward change in price
rho = 0.0344
Non-significant negative correlation means that price difference value does not predict future change in price itself.
Difference between price and moving average
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.2
-0.1
0.0
0.1
0.2

Conclusions from example

- Very easy to fool yourself
- Need to use rolling mean to accomodate variation
- But can confuse

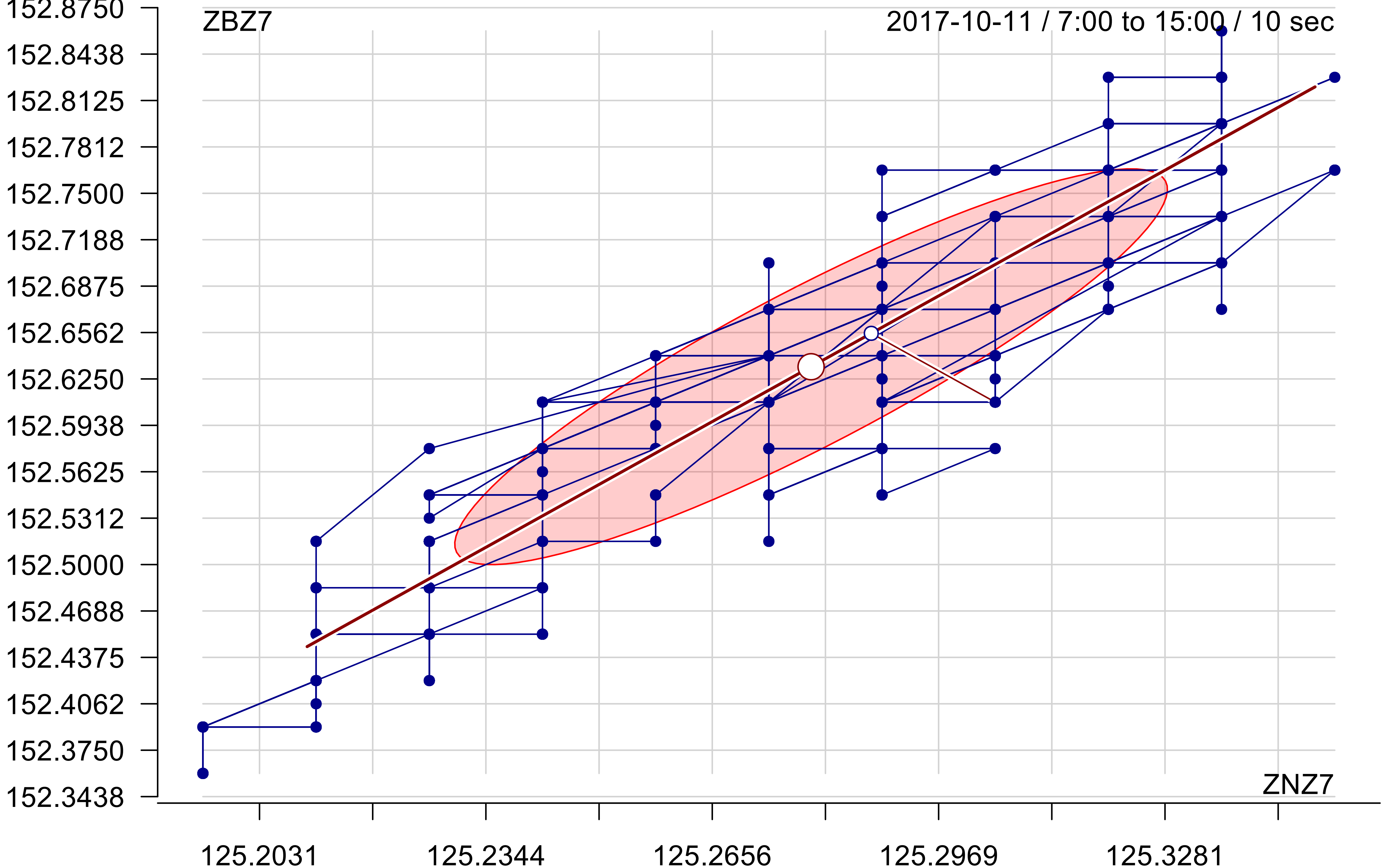
  mean moving to price value, with

  price value moving in a predictable direction
- Always test predictability on original signal

Reversion examples

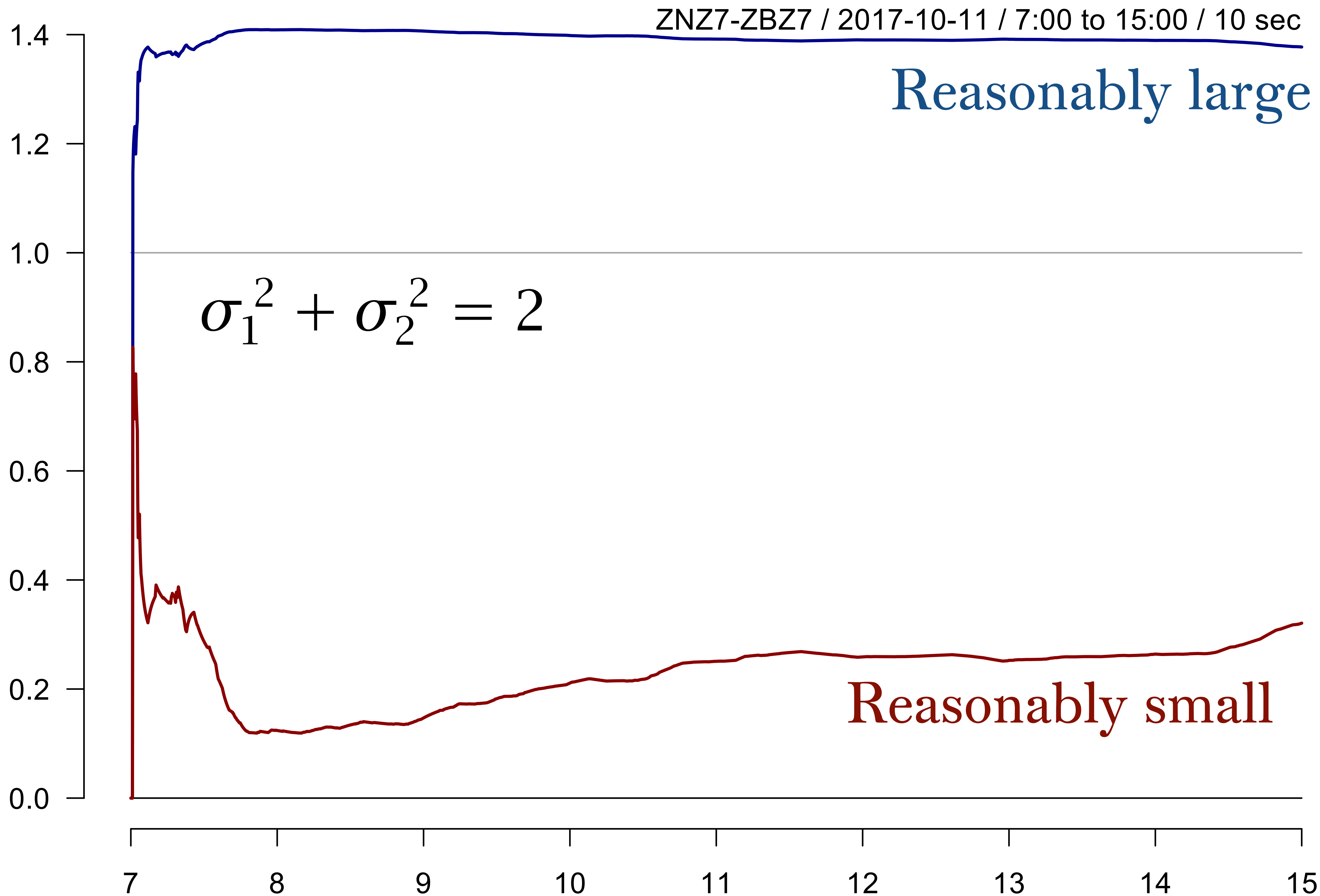
- Pick two products that seem to move together
- Compute PCA of cumulative covariance from start of day (uniform averaging)
- Show
  1. 2-dimensional plot of prices
  2. Eigenvalues of cumulative covariance matrix
  3. Prices and signal forecasts
  4. Correlation of signal with forward return
  5. Dependence of correlation on lag

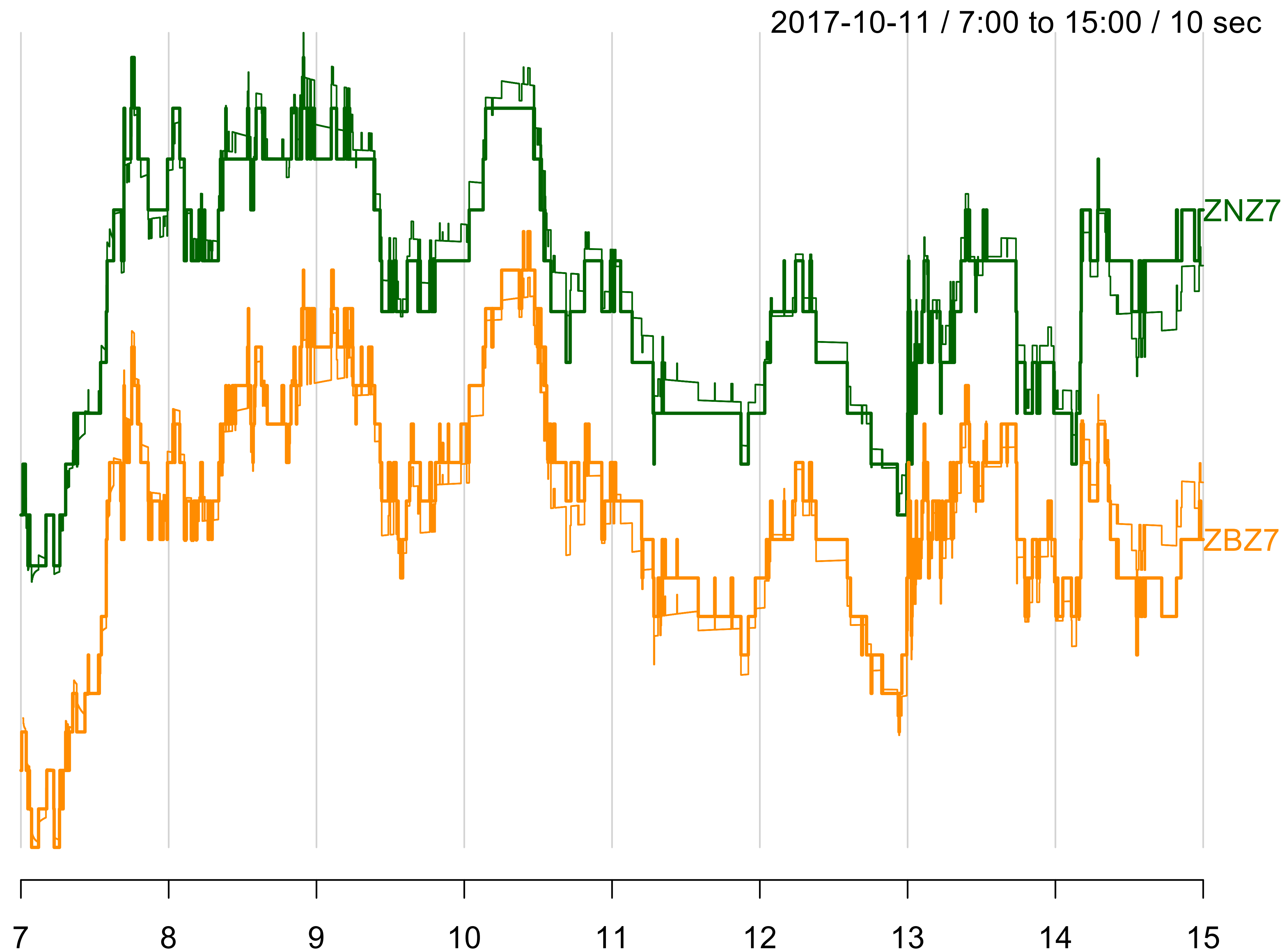
Treasuries



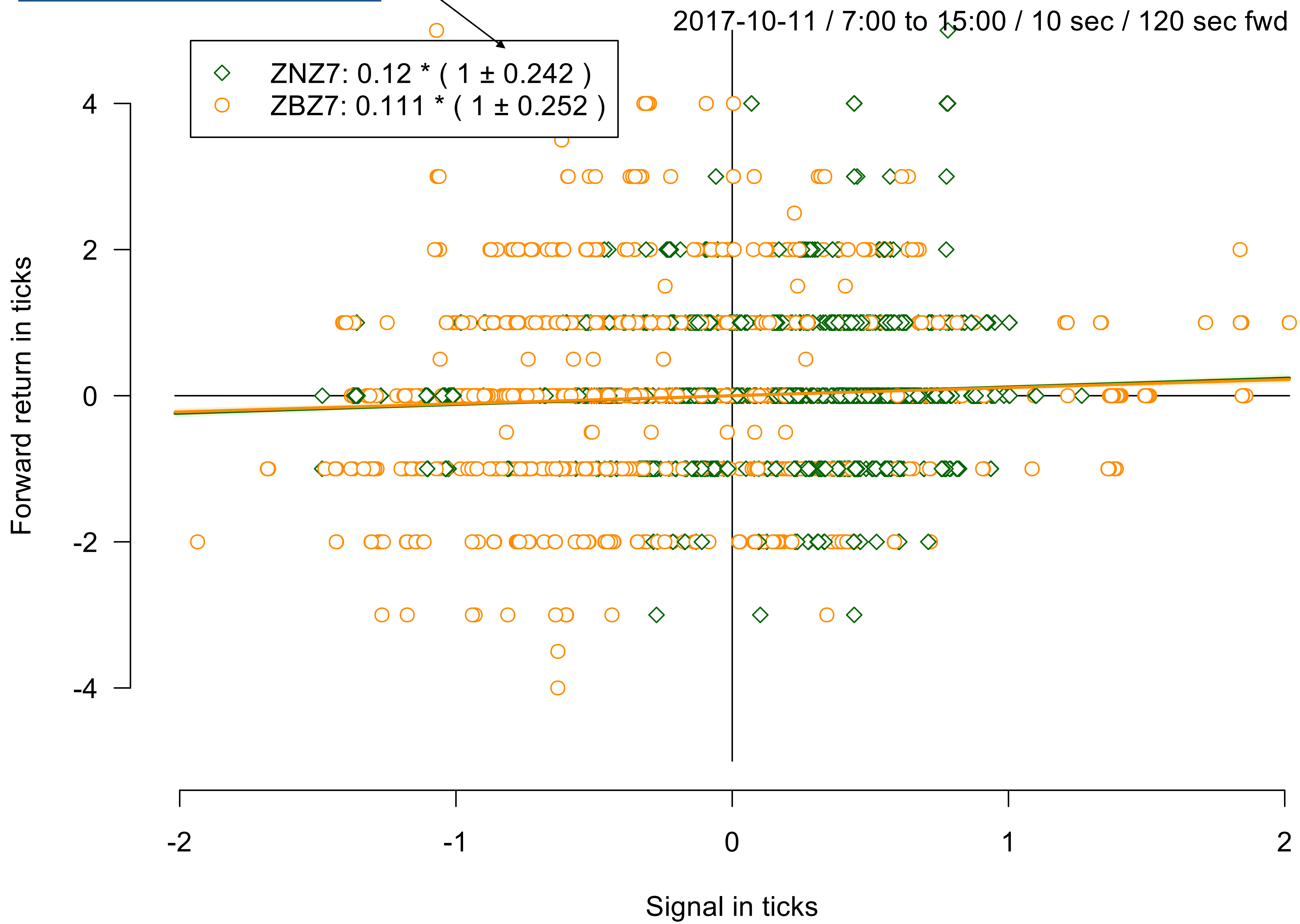
ZBZ7
2017-10-11 / 7:00 to 15:00 / 10 sec
152.8750
152.8438
152.8125
152.7812
152.7500
152.7188
152.6875
152.6562
152.6250
152.5938
152.5625
152.5312
152.5000
152.4688
152.4375
152.4062
152.3750
152.3438
ZNZ7
125.2031
125.2344
125.2656
125.2969
125.3281

Eigenvalues of correlation matrix

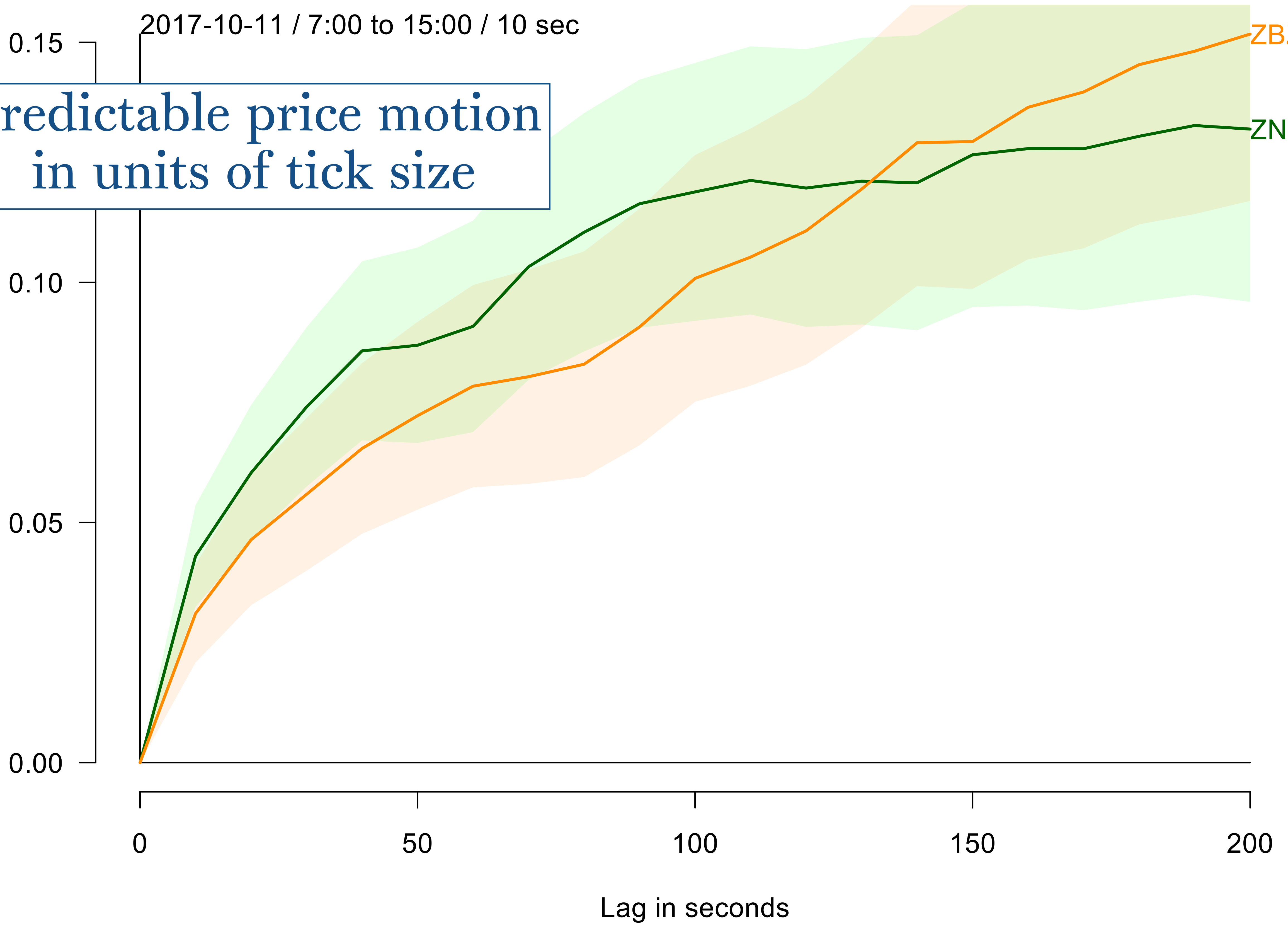



2017-10-11 / 7:00 to 15:00 / 10 sec
ZNZ7
ZBZ7
7
8
9
10
11
12
13
14
15

Std dev of coefficient, not data.
Finance can have high t-stat,
always has low R^2



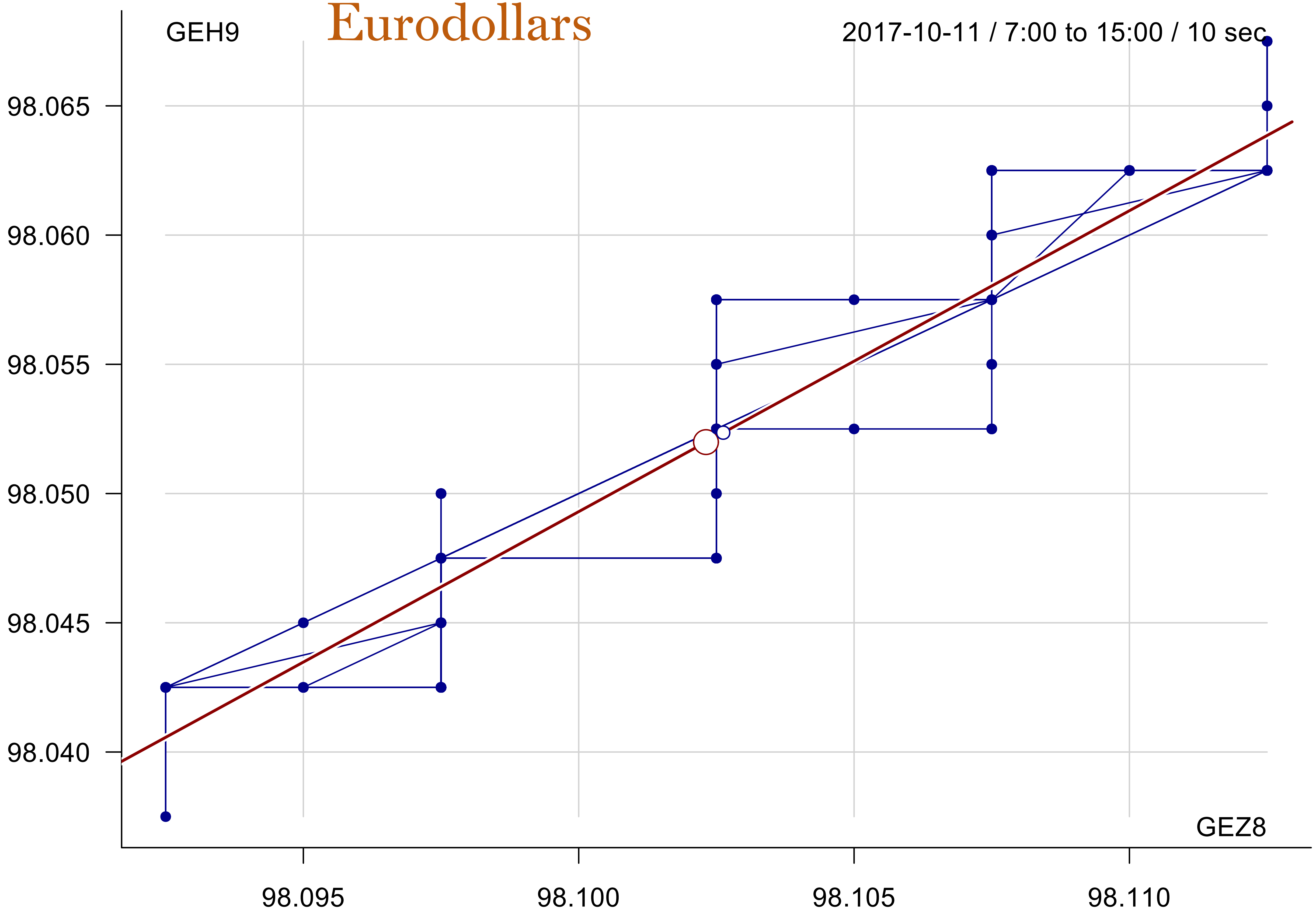
Predictable price motion in units of tick size



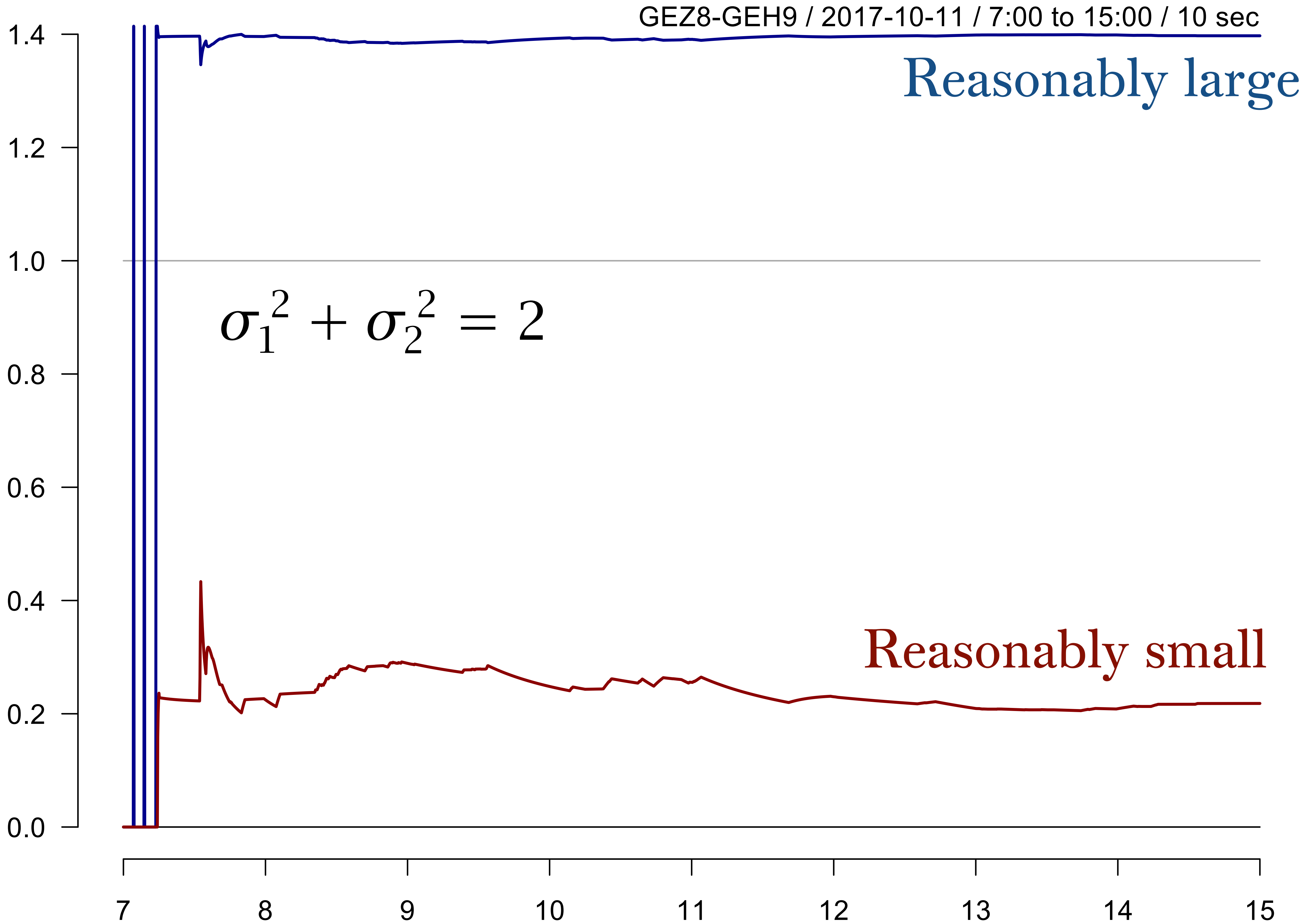
Problems with Treasuries

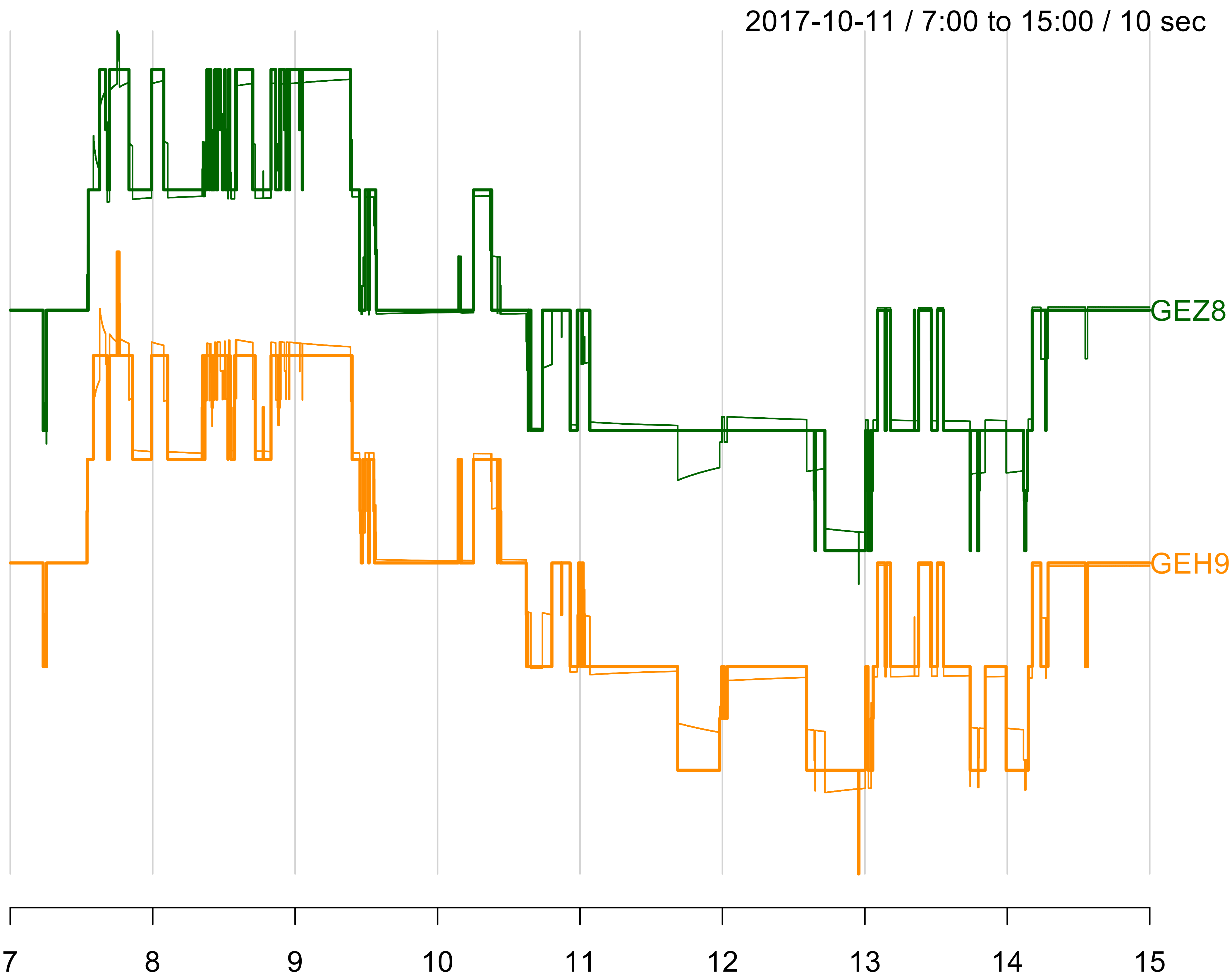
- Signal is weak in Treasuries
- Yield curve dynamics is more than rates-up vs rates-down
- Second principal component likely not mean-reverting
- Higher dimensional models can work better

Eurodollars

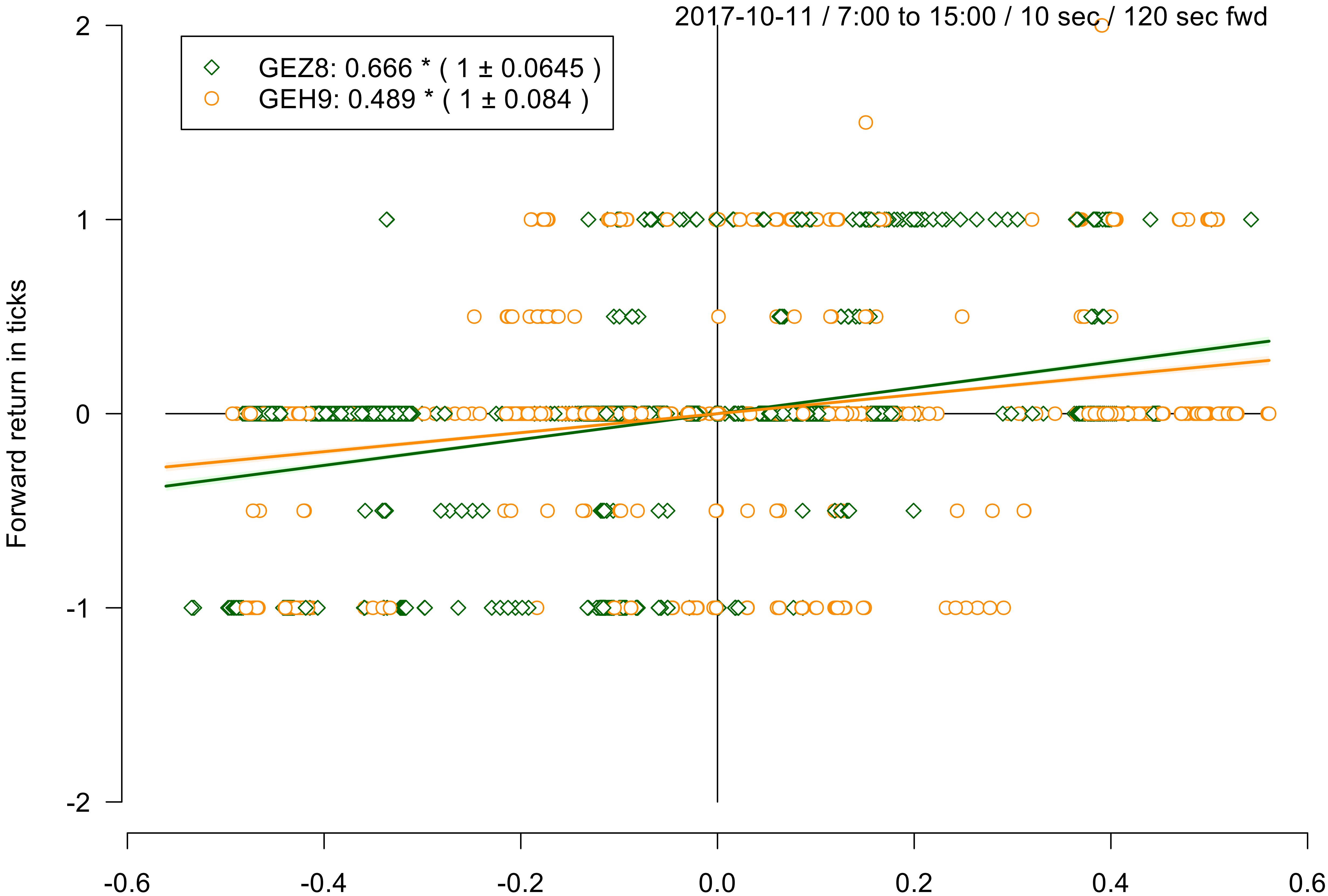


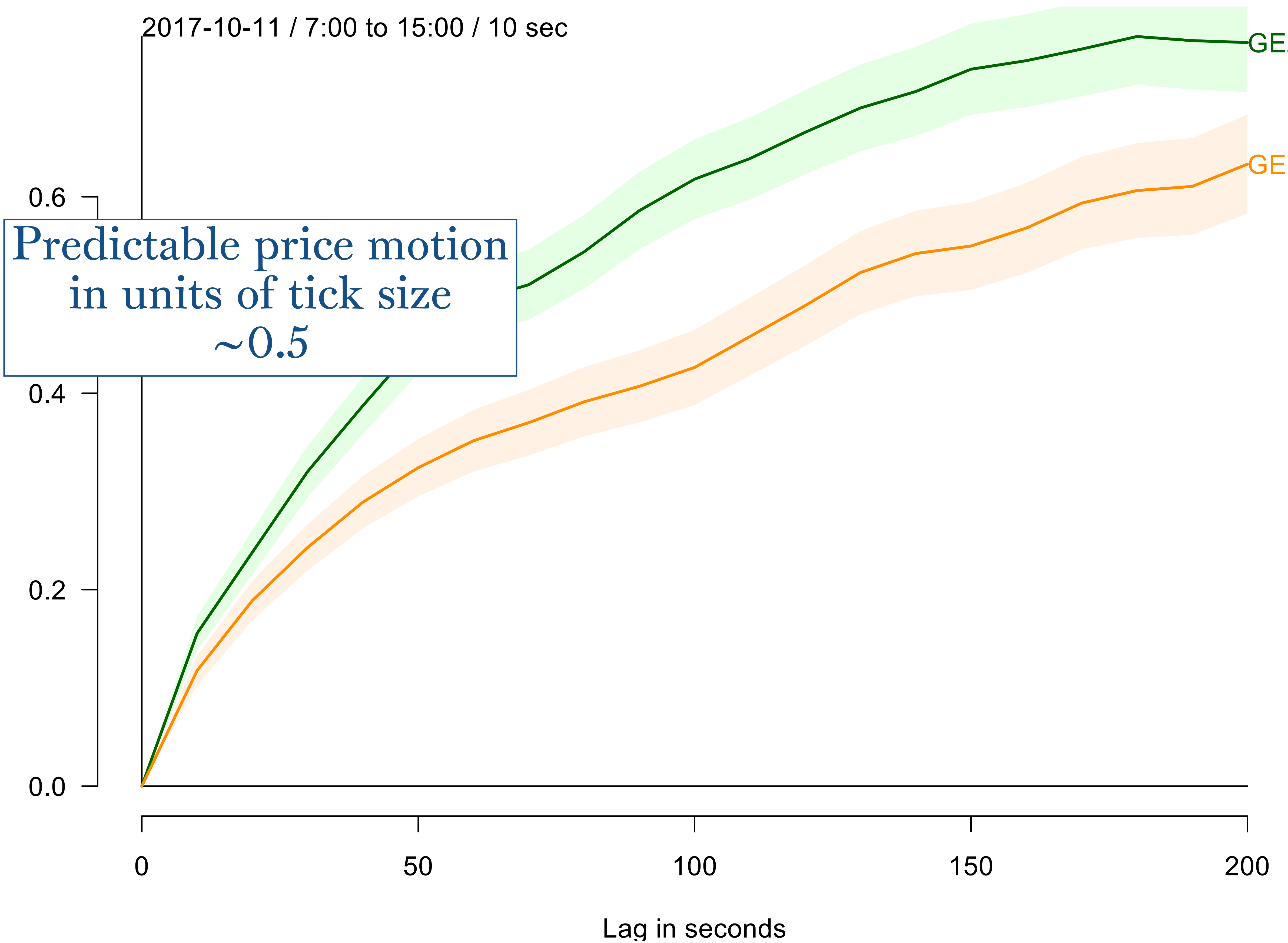
Eigenvalues of correlation matrix





2017-10-11 / 7:00 to 15:00 / 10 sec
GEZ8
GEH9
7
8
9
10
11
12
13
14
15


2017-10-11 / 7:00 to 15:00 / 10 sec / 120 sec fwd
GEZ8: 0.666 * (1 ± 0.0645)
GEH9: 0.489 * (1 ± 0.084)
Forward return in ticks
2
1
0
-1
-2
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6



Predictable price motion
in units of tick size
~0.5

Eurodollar Futures Contract Specs

View Another Product

Quotes | Settlements | Volume | Time & Sales | Contract Specs | Margins | Calendar

Futures | Options

Contract Unit	$2,500 x Contract IMM Index
Price Quotation	Contract IMM Index = 100 minus R R = three-month London interbank offered rate for spot settlement on 3rd Wednesday of contract month. E.g., a price quote of 97.45 signifies a deposit rate of 2.55 percent per annum. One interest rate basis point = 0.01 price points = $25 per contract.
Trading Hours	SUN - FRI: 5:00 p.m. - 4:00 p.m. CT
Minimum Price Fluctuation	Nearest expiring contract month: One quarter of one interest rate basis point = 0.0025 price points = $6.25 per contract. All other contract months: One half of one interest rate basis point = 0.005 price points = $12.50 per contract. The “new” nearest contract begins trading in 0.0025 increments on the same trade date as the last trading day in the expiring “old” nearest contract.
Product Code	CME Globex: GE CME ClearPort: ED Clearing: ED
Listed Contracts	Nearest 40 months (i.e., 10 years) in the March Quarterly cycle (Mar, Jun, Sep, Dec) plus the nearest 4 “serial” months not in the March Quarterly cycle. The new March Quarterly contract month for delivery 10 years hence is listed on the expiration day of the nearby quarterly contract month. For example, if GEZ17 terminates trading on Mon, 18 Dec at 5am CT, then GEZ27 is listed for trading at 5pm CT on Sun, 17 Dec at 5pm CT, for first trade date of Mon, 18 Dec.
Settlement Method	Financially Settled

prc = 100 − LIBOR

40 different maturities

Cash settled
(first one ever)

https://www.cmegroup.com/trading/interest-rates/stir/eurodollar_contract_specifications.html

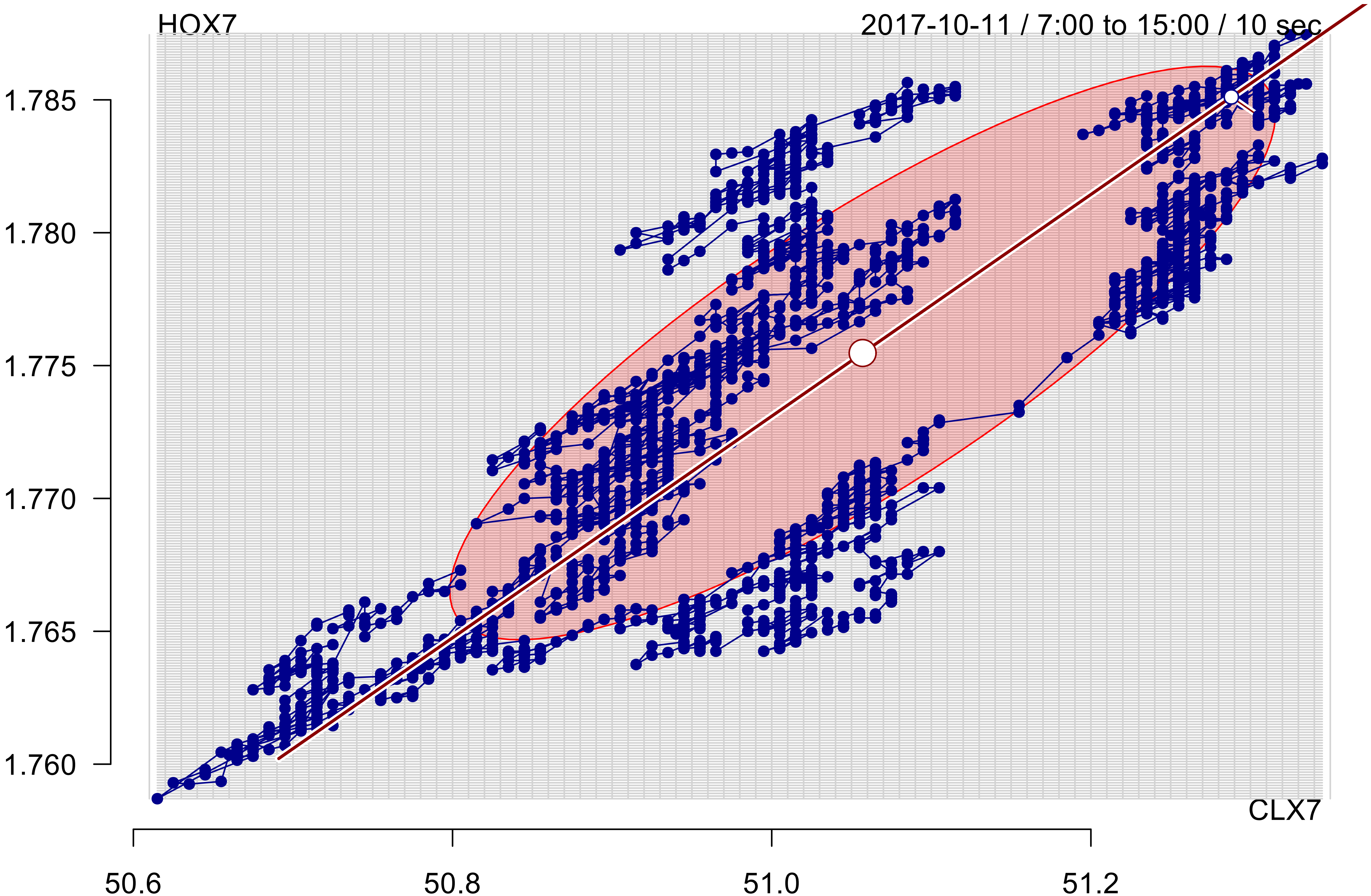
Comments on Eurodollars

- LIBOR up / LIBOR down is most of variance: 2-dimensional model can work well
- Higher dimensional models work even better

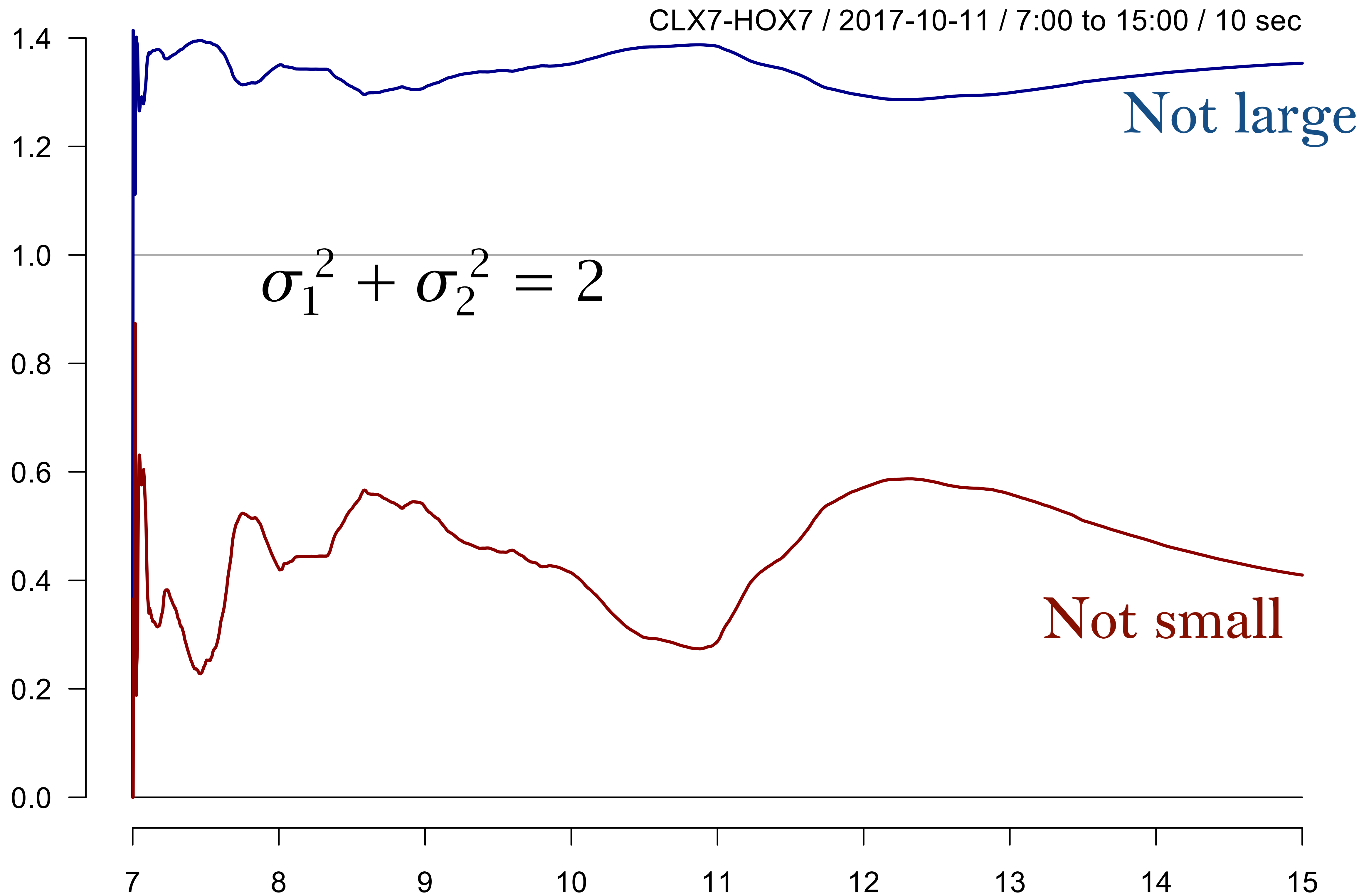
  1 degree of freedom

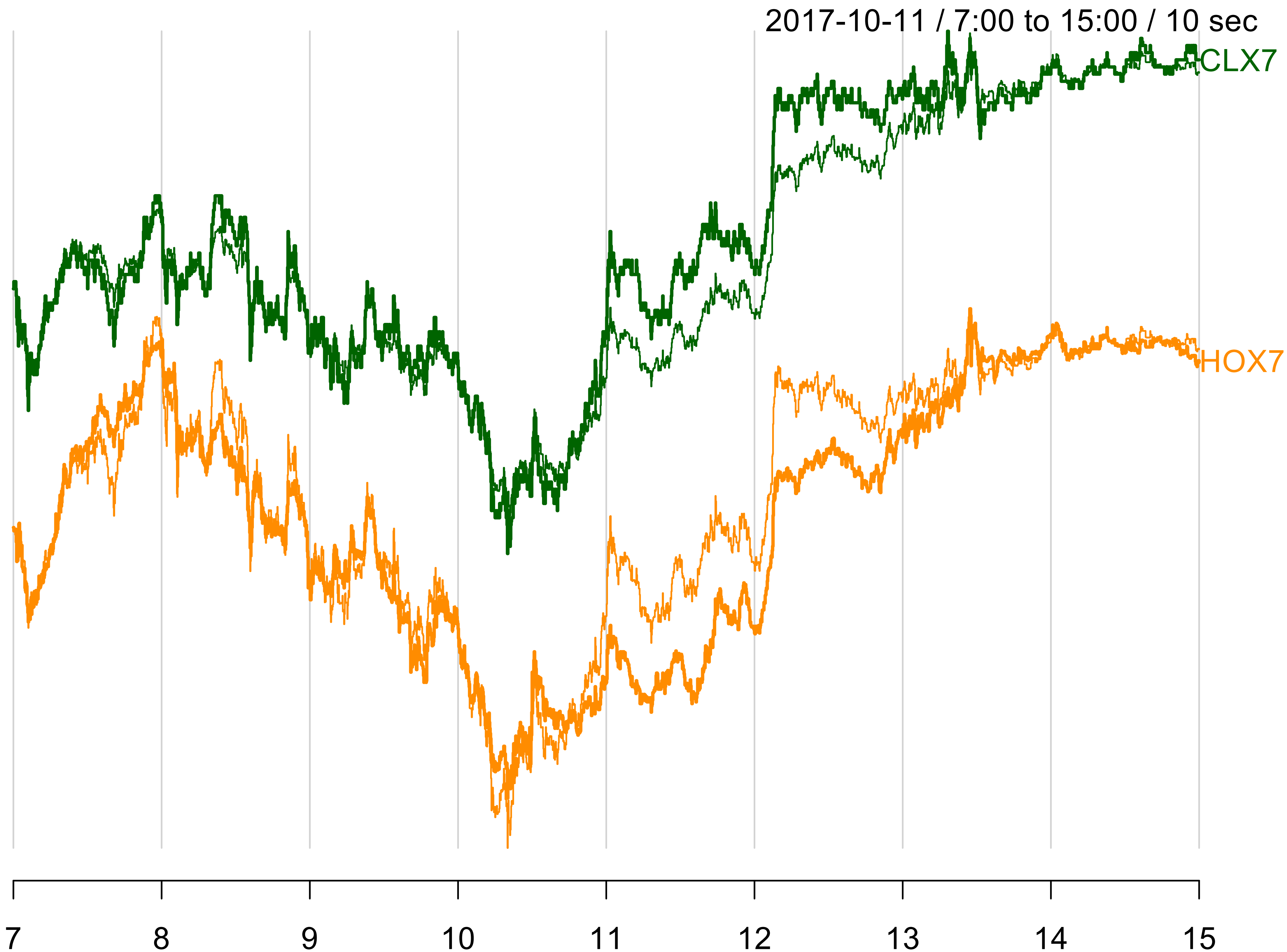
  n-1 reverting directions

Crude Oil vs Heating Oil

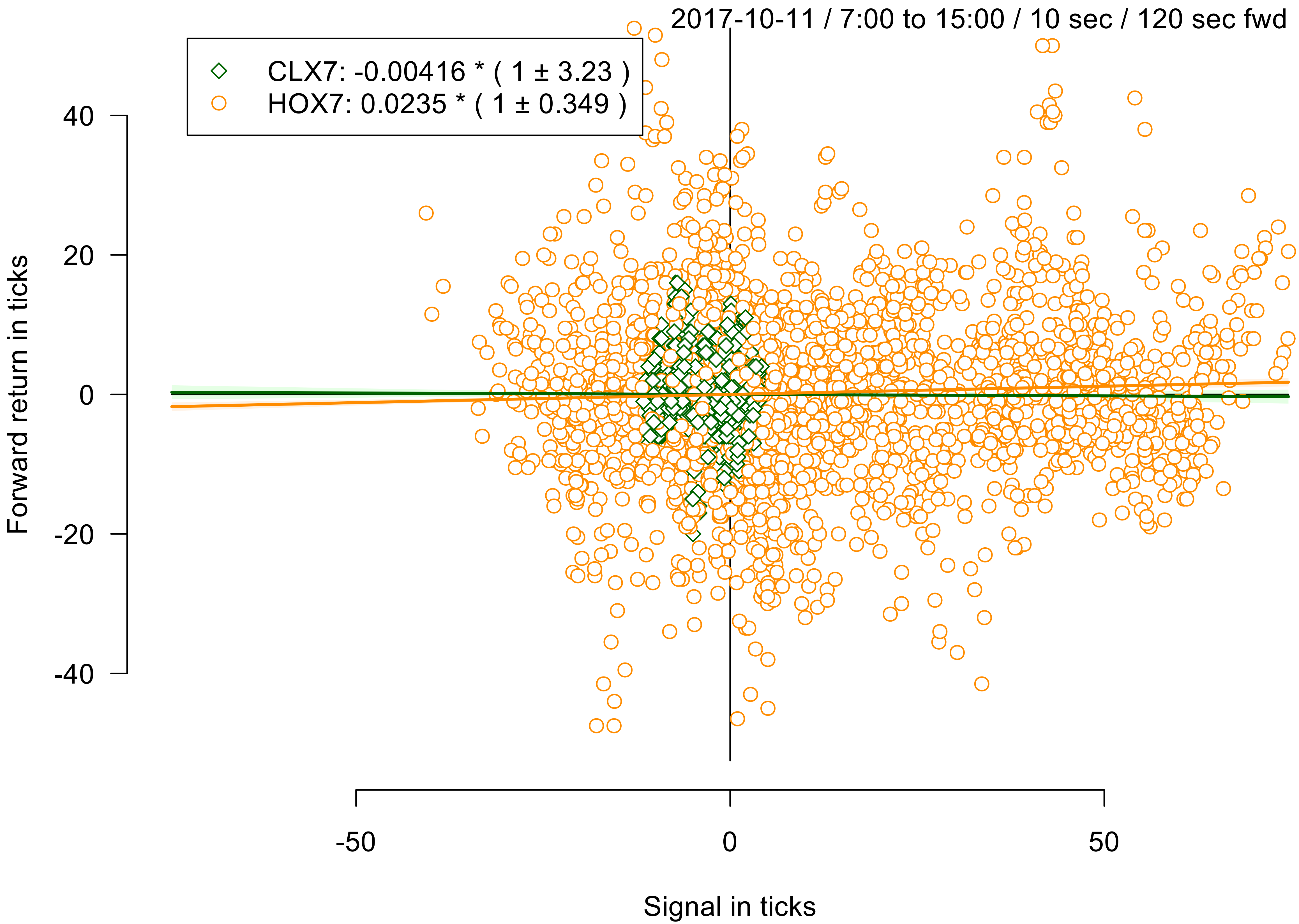


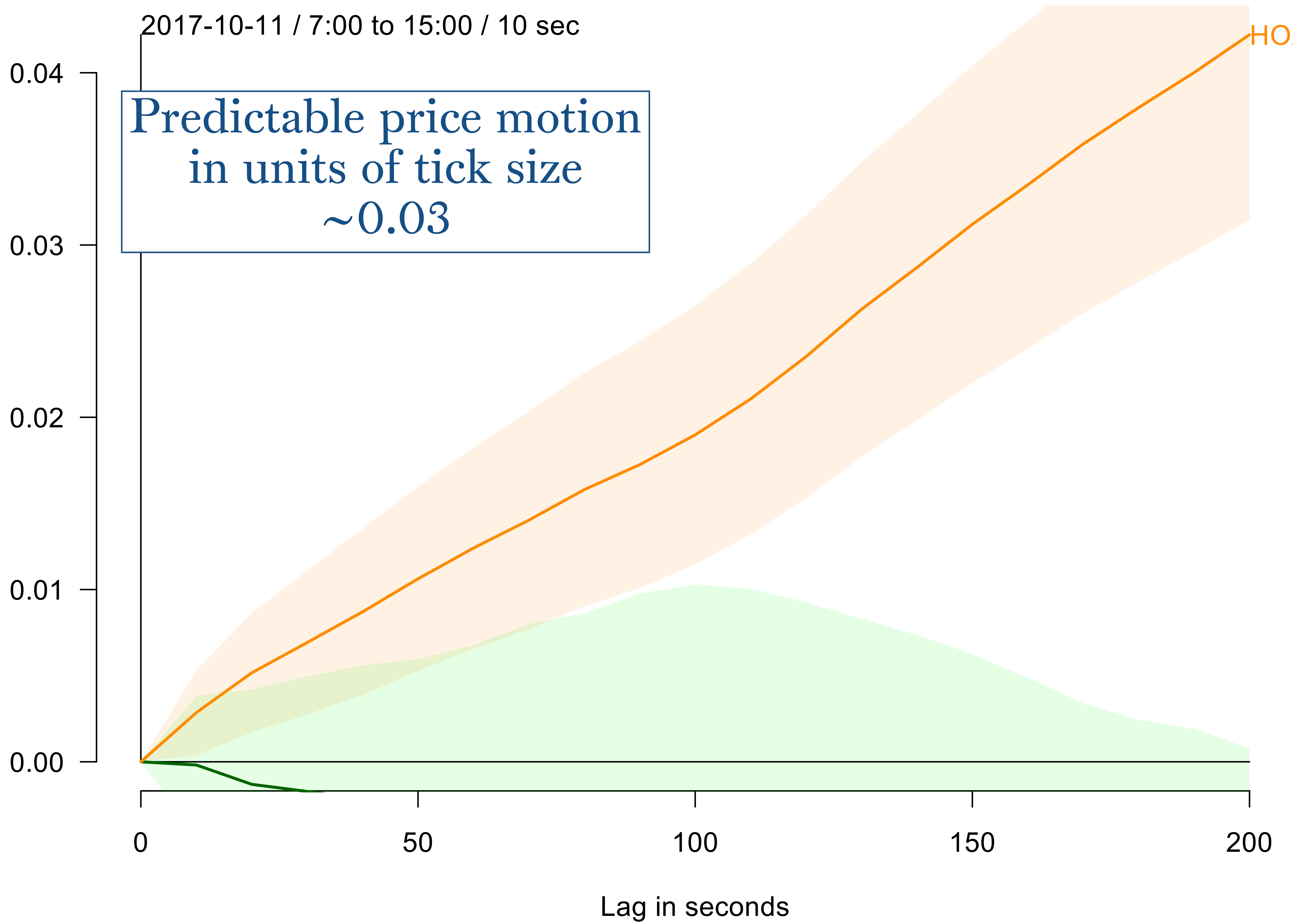
Eigenvalues of correlation matrix





2017-10-11 / 7:00 to 15:00 / 10 sec
CLX7
HOX7
7
8
9
10
11
12
13
14
15


2017-10-11 / 7:00 to 15:00 / 10 sec / 120 sec fwd
CLX7: -0.00416 * (1 ± 3.23)
HOX7: 0.0235 * (1 ± 0.349)
Forward return in ticks
40
20
0
-20
-40
-50
0
50
Signal in ticks

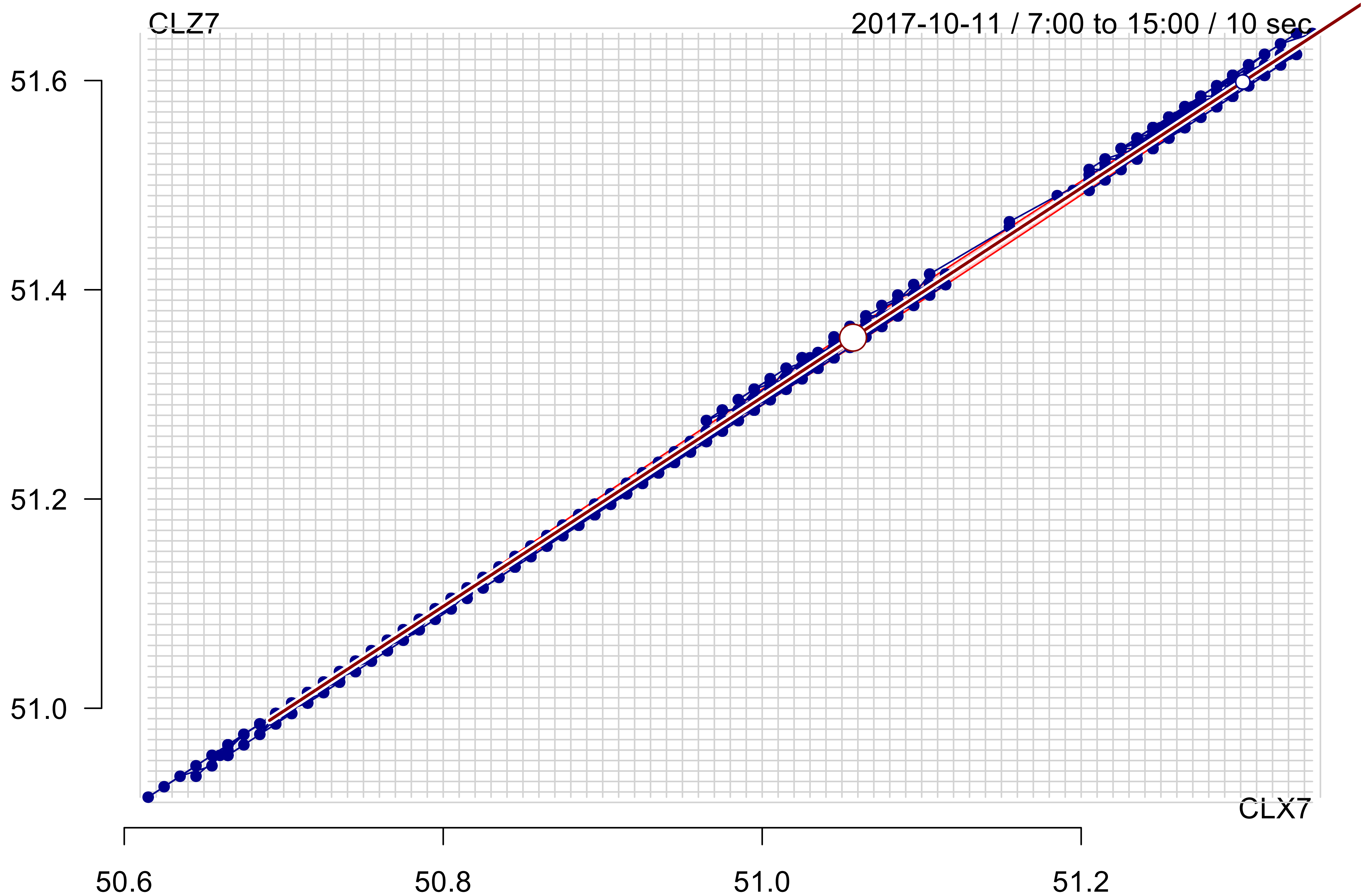


Predictable price motion
in units of tick size
~0.03

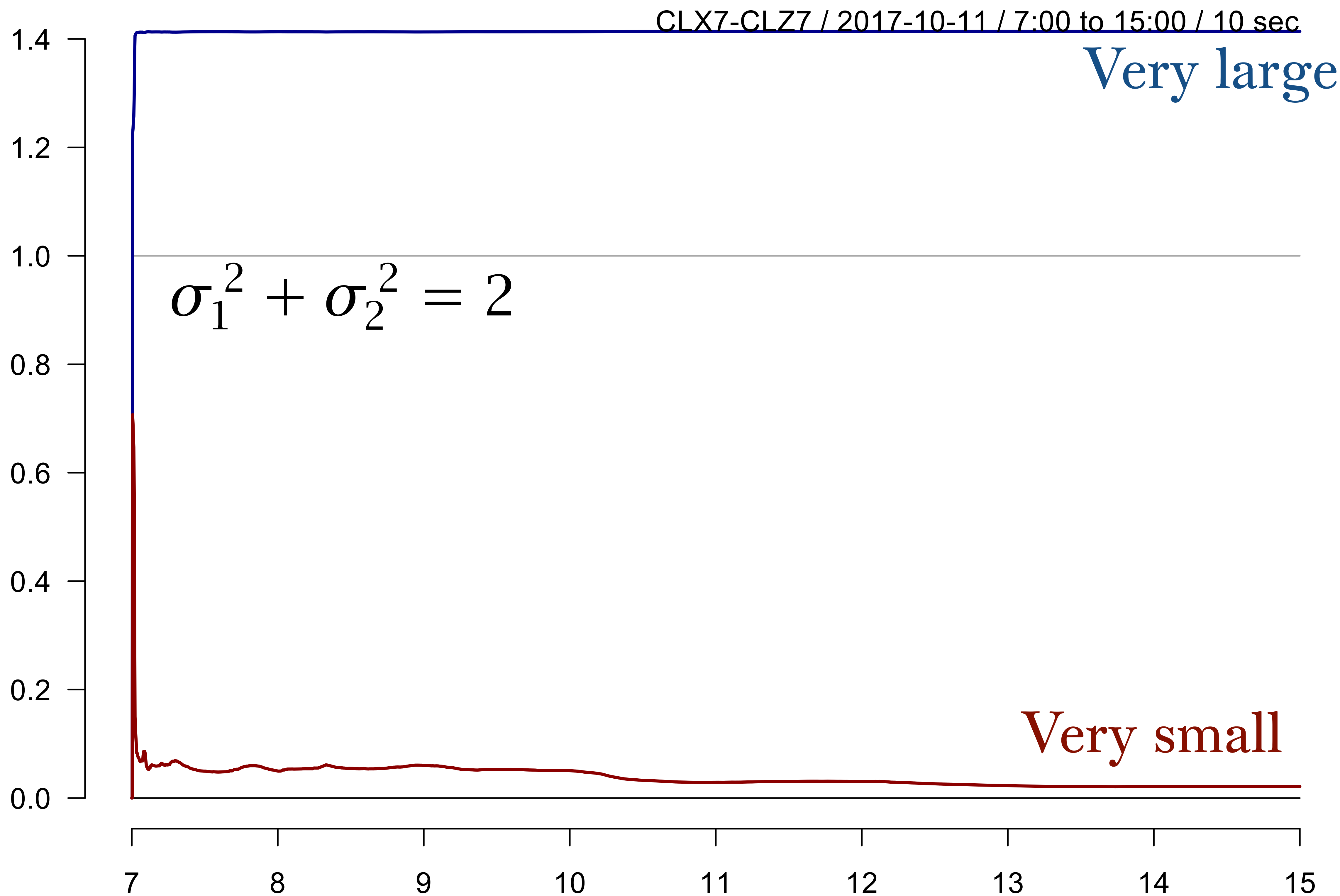
Crude Oil vs Heating Oil

- Poorly cointegrated
- Signal is weak

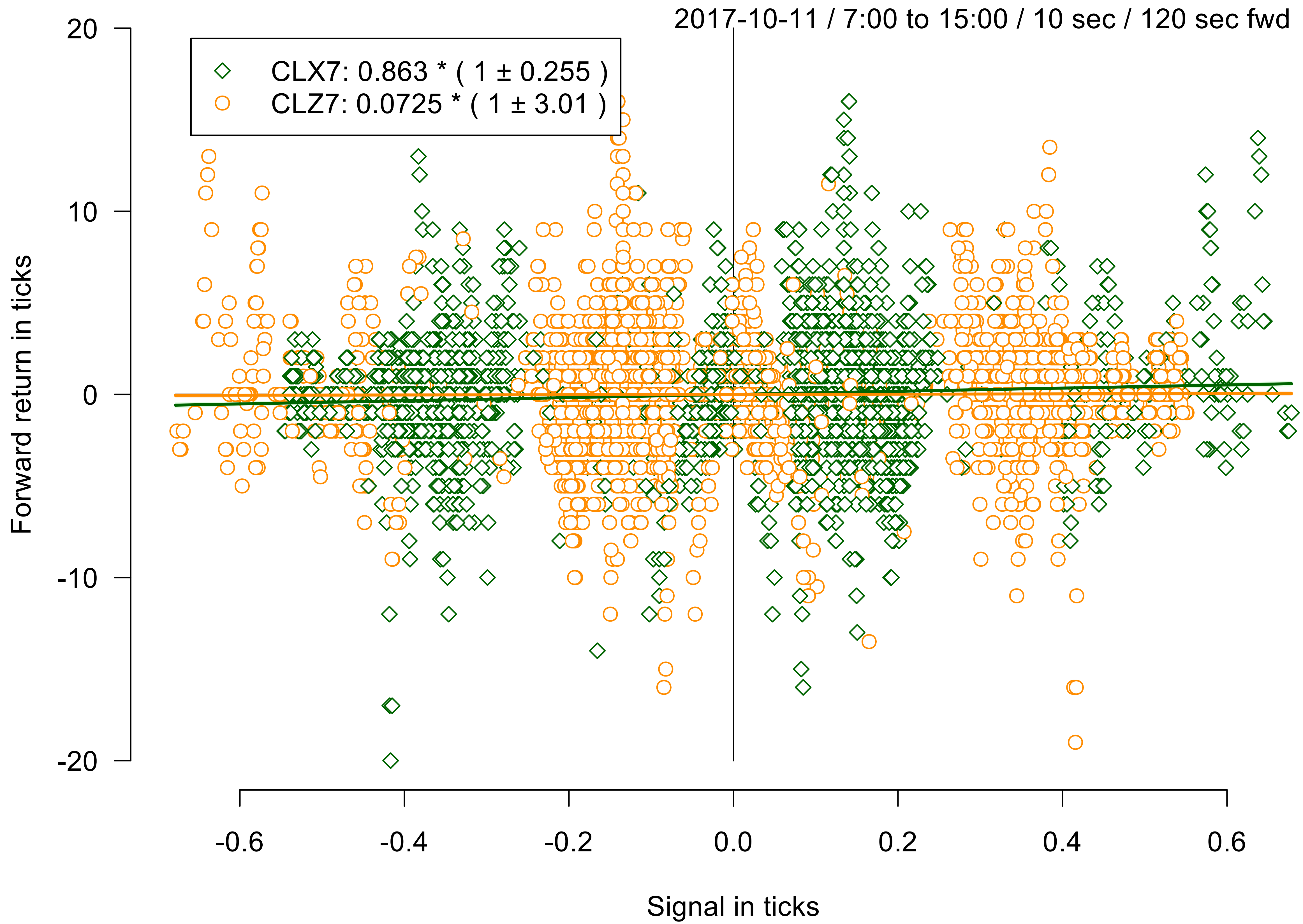
Crude Oil Nov vs Dec

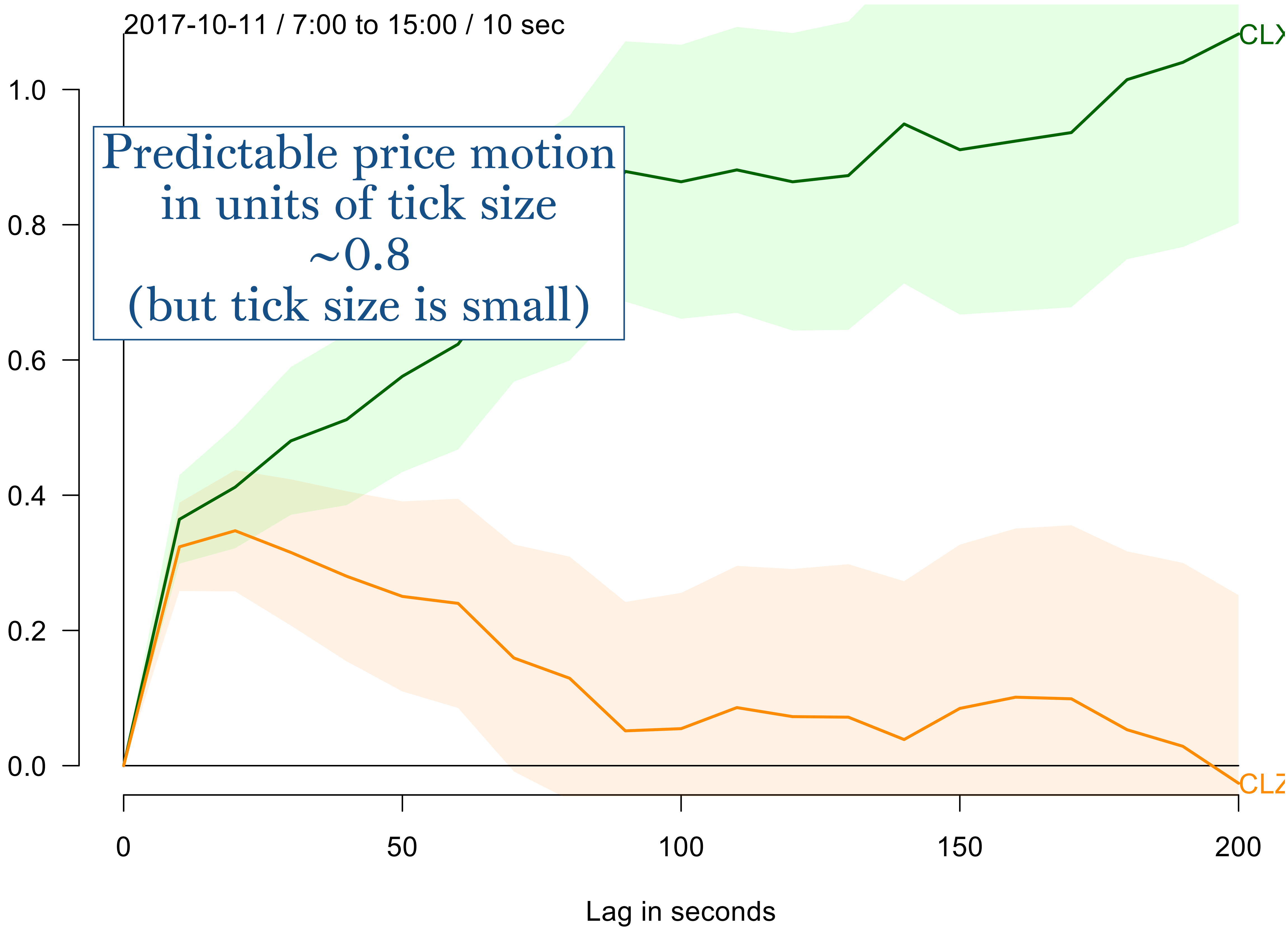


Eigenvalues of correlation matrix



Signal is invisible


2017-10-11 / 7:00 to 15:00 / 10 sec / 120 sec fwd
CLX7: 0.863 * (1 ± 0.255)
CLZ7: 0.0725 * (1 ± 3.01)
Forward return in ticks
20
10
0
-10
-20
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
Signal in ticks

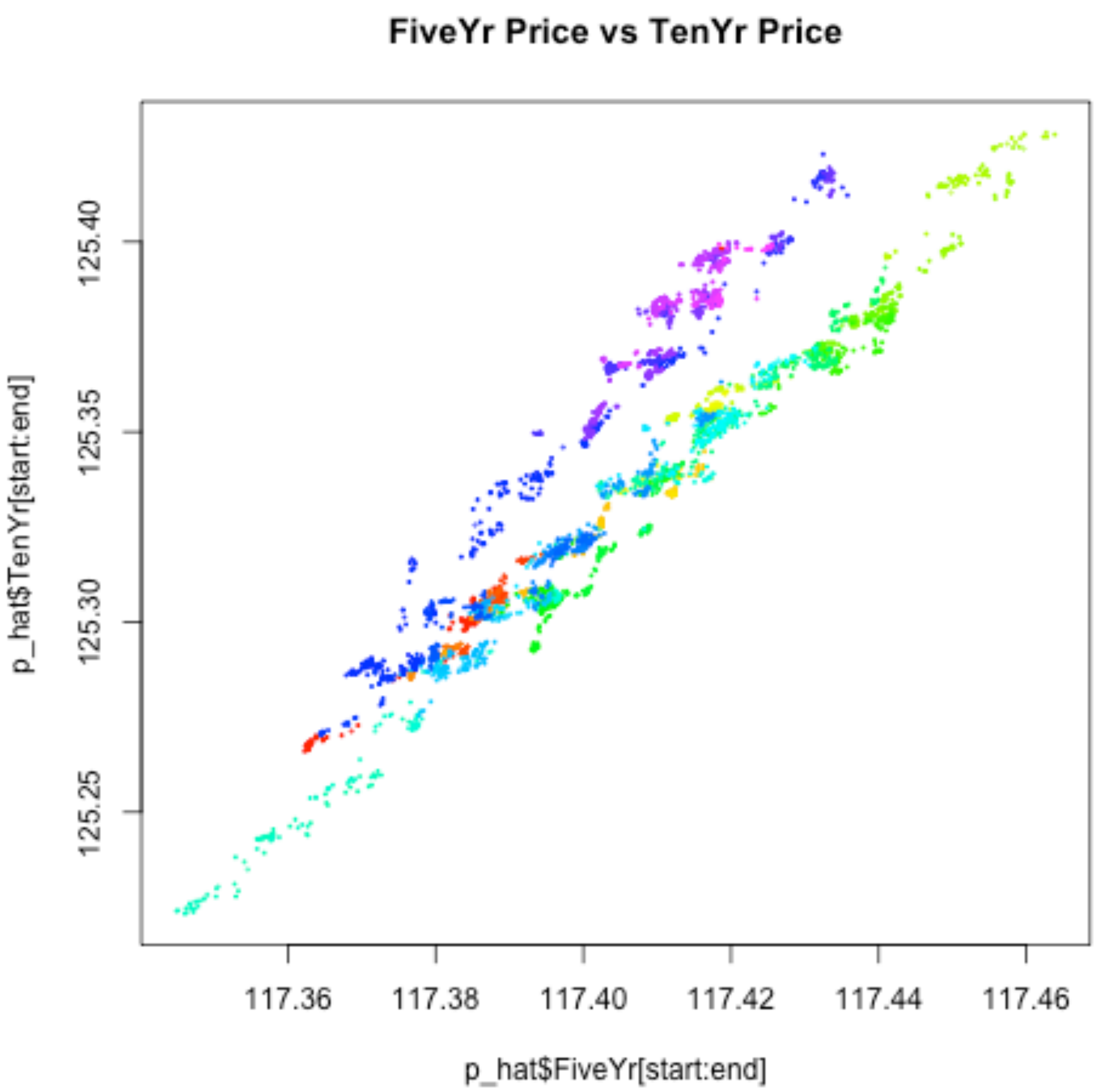


Predictable price motion
in units of tick size
~0.8
(but tick size is small)

Crude Oil Nov vs Dec

- Too tightly coupled to give good signal
- Do not move apart enough to trade

Regime change identification with machine learning

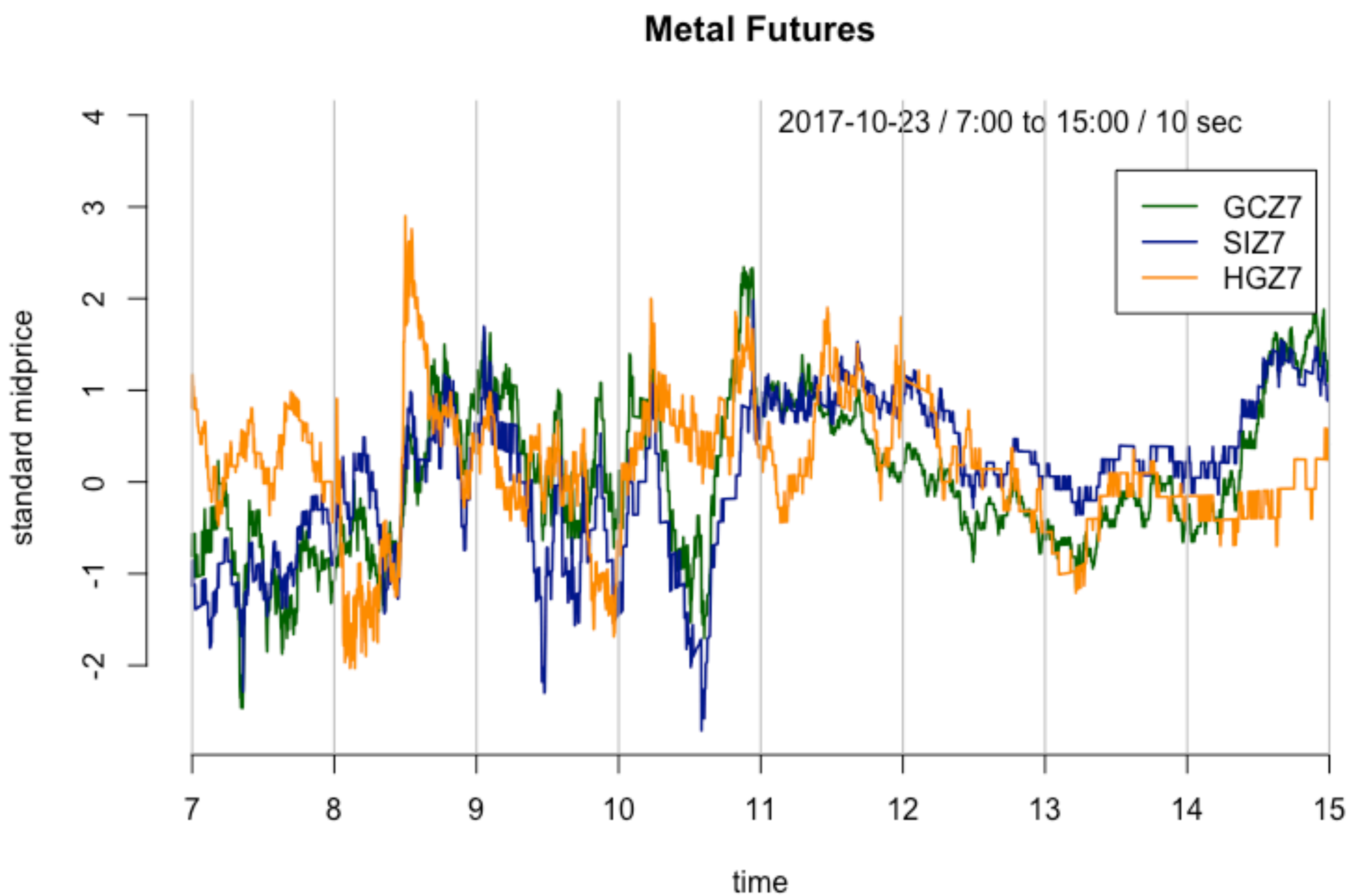


Kyle Xiao, Princeton

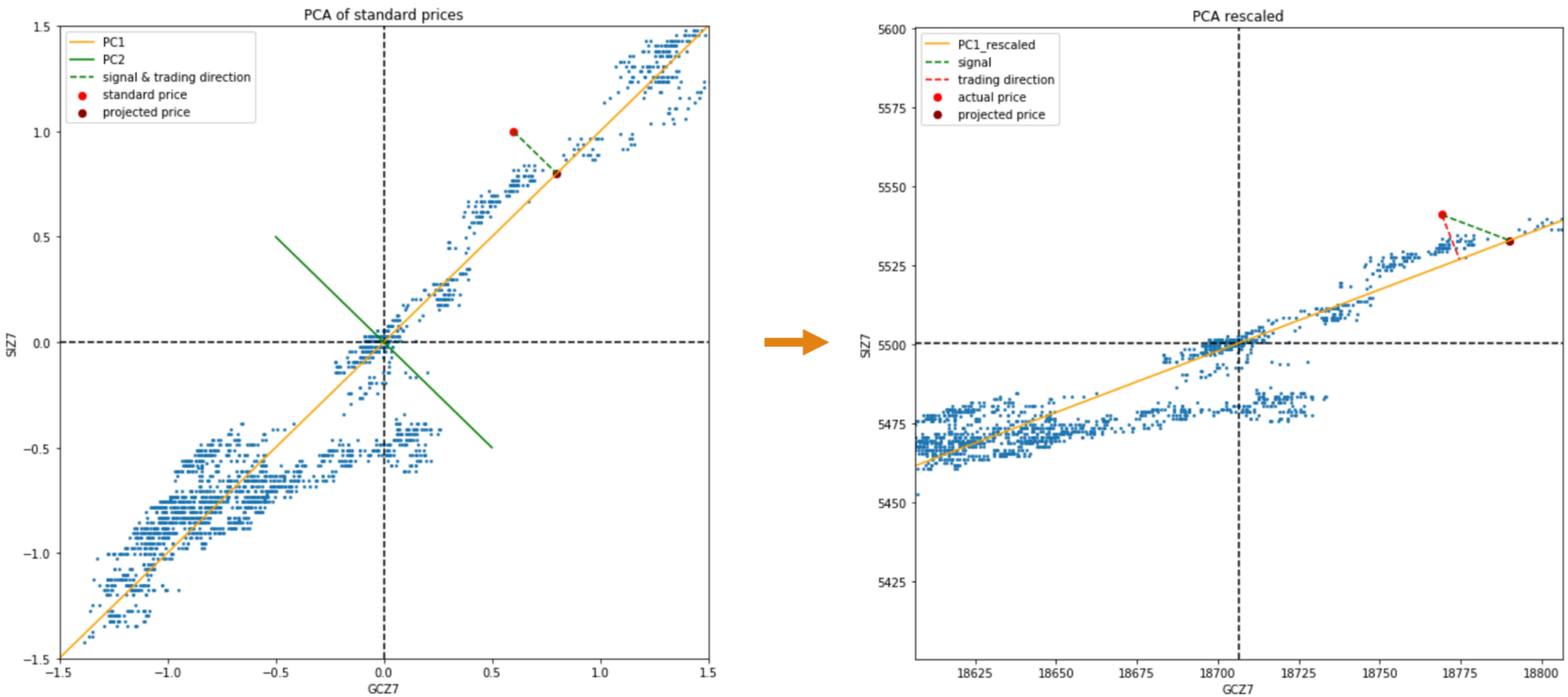
Generation of signal logic

Kayla Zhang,
MS program 2020

Gold/Silver/Copper

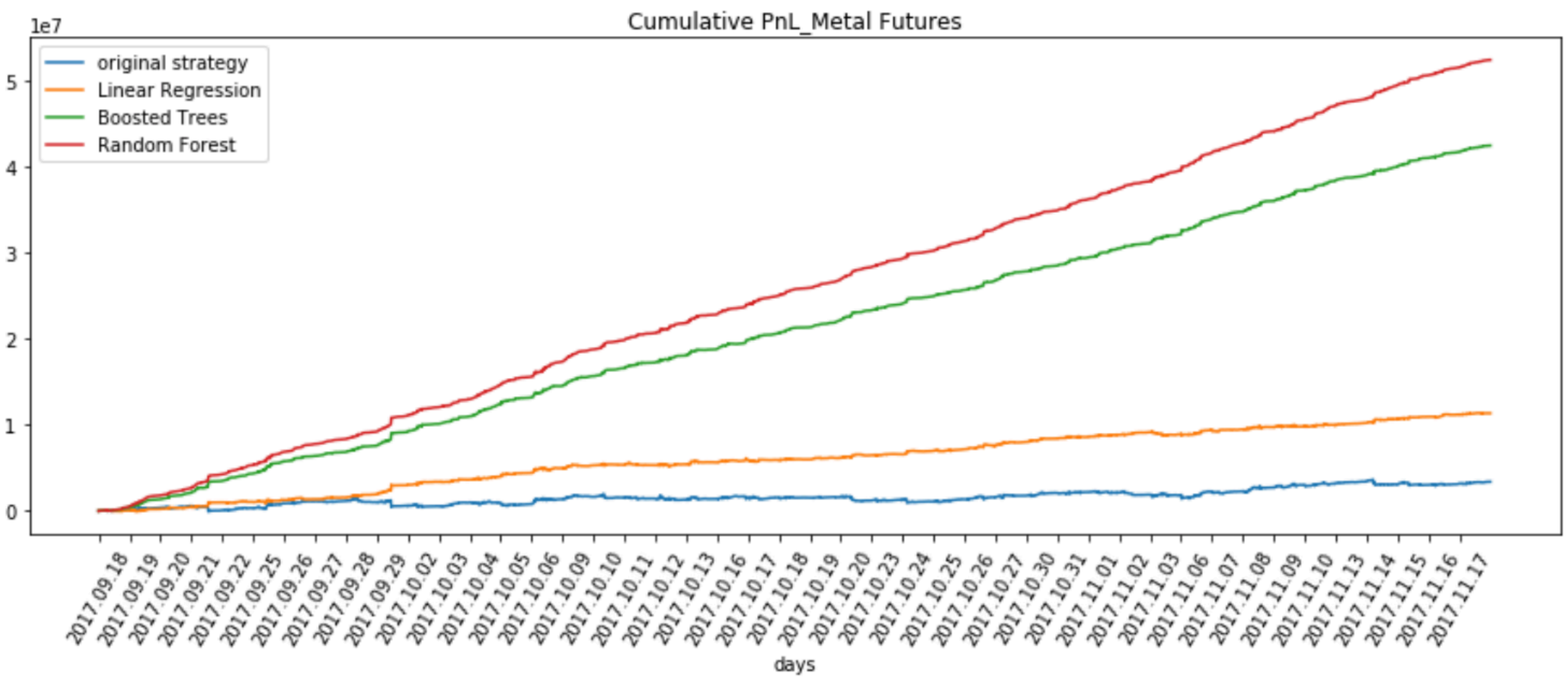


Regression plot



P&L

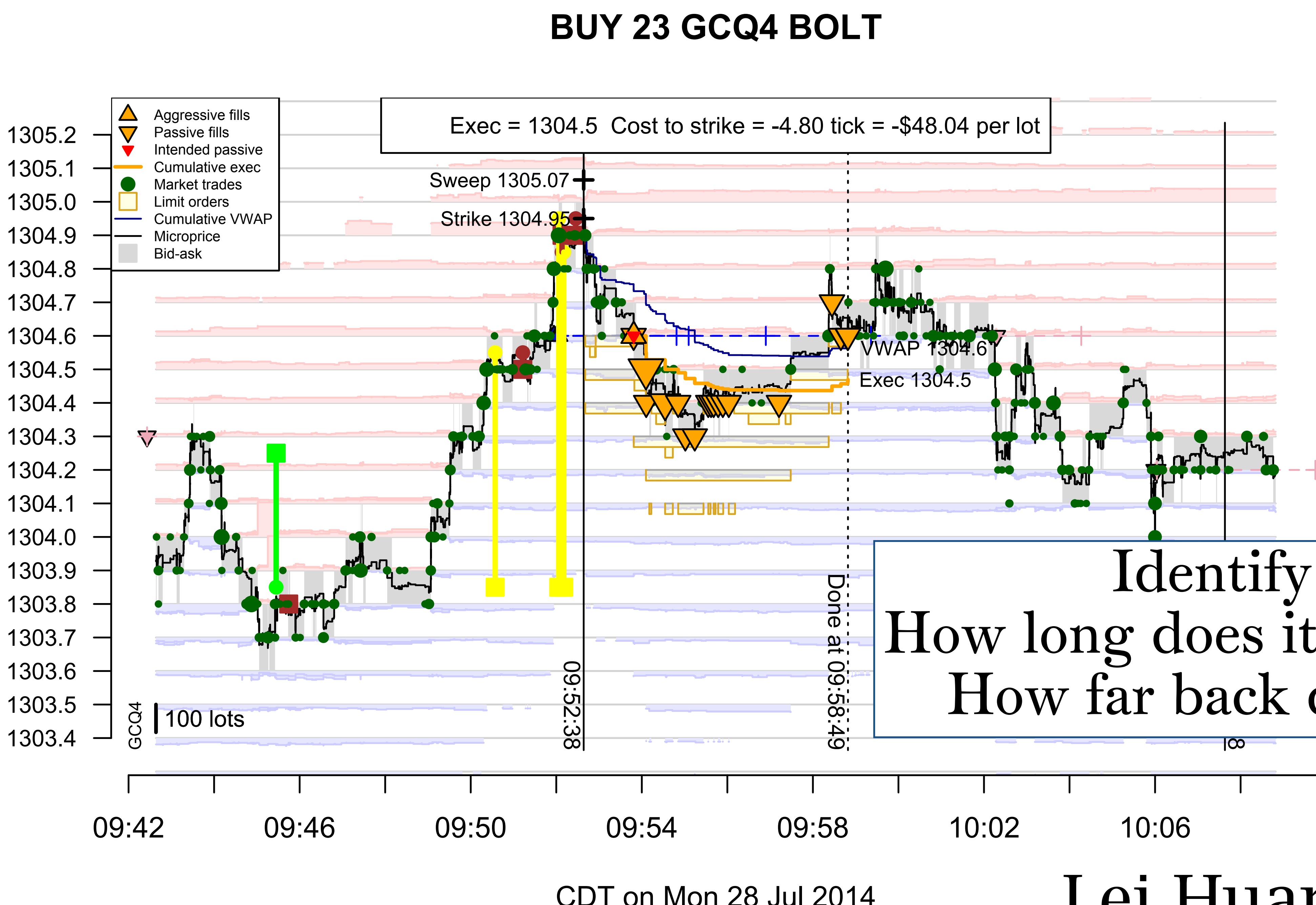
Use machine learning to identify when to trade, when to reduce risk



Other signal generation

- Look for signs of
  - Rebound from rapid price moves
  - Other players in market
  - Differences between assets
- Get idea, invent story, do statistical tests
- Machine learning approaches

Sweep (reversion) signal



Identify sweep
How long does it last? (timeout)
How far back does it come?

Lei Huang, QB

Momentum signal

quantitativebrokers

Shankar Narayanan
January 16, 2018

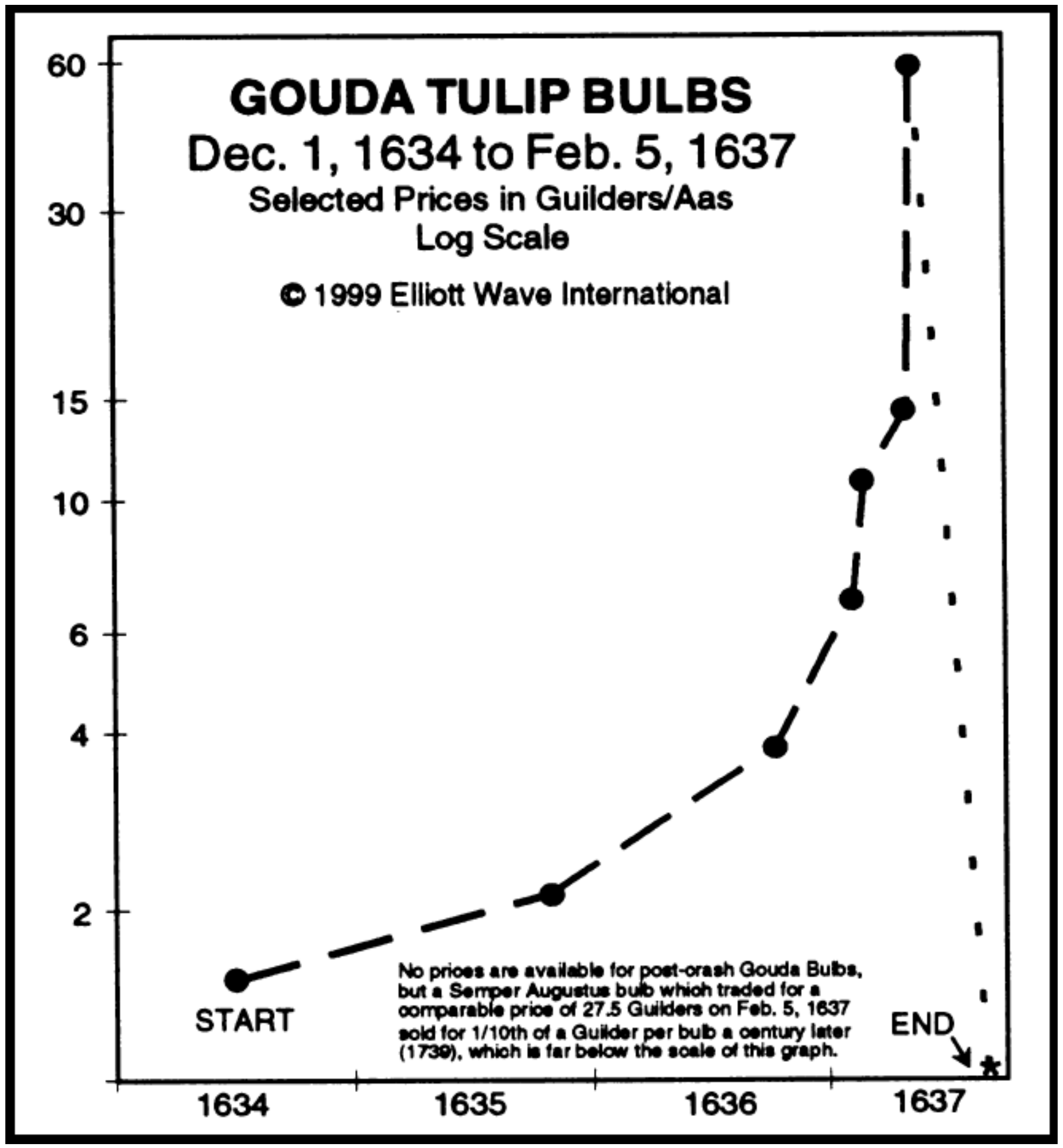
Page 1

Intra-day Price Bubbles

Underlying idea:
There are many econometric techniques that identify signals and forecasts at daily and longer time scales. Apply these at intraday level.

This signal:
Techniques for identifying "bubbles" in asset prices can be used to identify momentum episodes in high frequency intraday trading.

Bubbles



- *"Economic Bubbles"*
  When we talk of bubbles, we tend to think of the mortgage crisis or internet bubble, where as it is an age-old phenomenon

- Also very popular topic; around 500,000 academic citations on Google Scholar

- *Bubbles typically are characterized by several months or even years of sustained growth in price*
  e.g the tech bubble, housing bubble,etc

Shankar Narayanan,
Quantitative Brokers

Real time detection of Bubbles

Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500*

Peter C. B. Phillips
Yale University, University of Auckland,
University of Southampton & Singapore Management University

Shu-Ping Shi
The Australian National University

Jun Yu
Singapore Management University

August 7, 2013

Abstract

Recent work on econometric detection mechanisms has shown the effectiveness of recursive procedures in identifying and dating financial bubbles. These procedures are useful as warning alerts in surveillance strategies conducted by central banks and fiscal regulators with real time data. Use of these methods over long historical periods presents a more serious econometric challenge due to the complexity of the nonlinear structure and break mechanisms that are inherent in multiple bubble phenomena within the same sample period. To meet this challenge the present paper develops a new recursive flexible window method that is better suited for practical implementation with long historical time series. The method is a generalized version of the sup ADF test of Phillips, Wu and Yu (2011, PWY) and delivers a consistent date-stamping strategy for the origination and termination of multiple bubbles. Simulations show that the test significantly improves discriminatory power and leads to distinct power gains when multiple bubbles occur. An empirical application of the methodology is conducted on S&P 500 stock market data over a long historical period from January 1871 to December 2010. The new approach successfully identifies the well-known historical episodes of exuberance and collapse over this period, whereas the strategy of PWY and a related CUSUM dating procedure locate far fewer episodes in the same sample range.

Keywords: Date-stamping strategy; Flexible window; Generalized sup ADF test; Multiple bubbles, Rational bubble; Periodically collapsing bubbles; Sup ADF test;

JEL classification: C15, C22

- Several algorithms have been proposed for real time detection of economic bubbles
- Our algorithm is a **modified version** of the recursive unit root tests (Phillips, Wu and Yu, 2011) and Generalized sup-ADF test (Philips, Shi and Yu, 2015)
- A modified version of a simple unit root test applied to improve the power of real-time detection of bubbles

Shankar Narayanan,
Quantitative Brokers

The Detection of Intra-Day Bubbles

- Test is a generalized version of Augmented-Dickey Fuller test of unit root

- The prototypical model takes the following form:

$$y_t = \rho\,(y_{t-1} - \bar{y}) + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

$$H_0 : \hat{\rho} = 1$$
$$H_1 : \hat{\rho} > 1$$

- When $\hat{\rho} > 1$ the price is believed to be in an explosive state.

Shankar Narayanan,
Quantitative Brokers

Main idea

y_t = prices

AR(1)

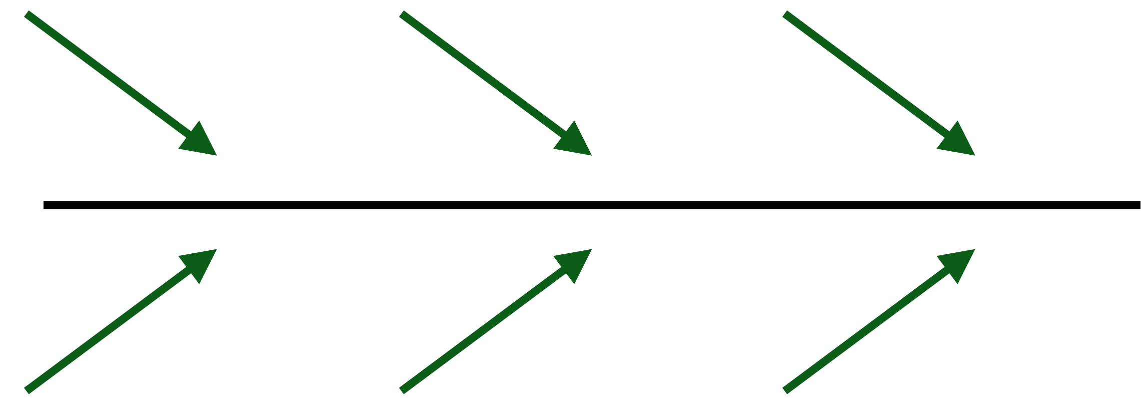
$$y_t = \rho\,(y_{t-1} - \overline{y}) + \epsilon_t$$

or AR(p)

$$y_t = \rho\,(y_{t-1} - \overline{y}) + \delta_1\,\Delta y_{t-1} + \cdots + \delta_{p-1}\,\Delta y_{t-p+1} + \epsilon_t$$

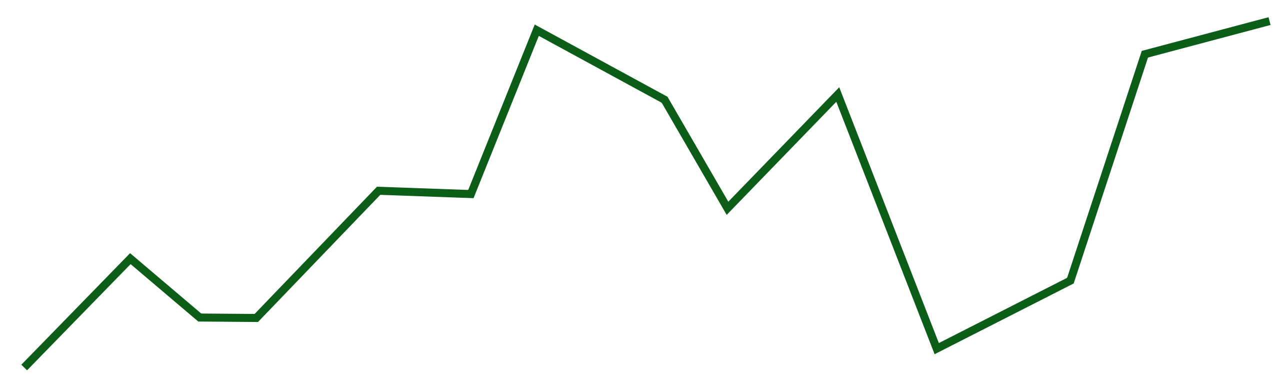
$|\rho| < 1$

mean-reverting



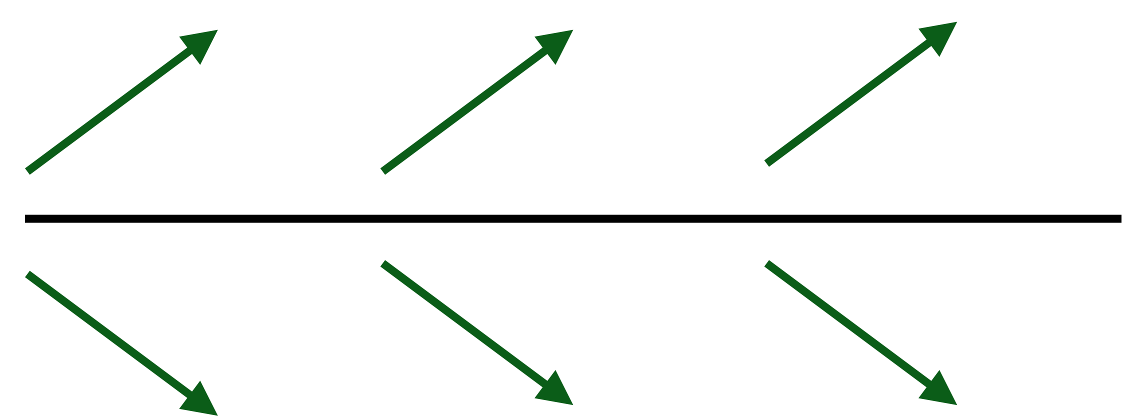
$\rho = 1$

random walk



$|\rho| > 1$

explosive



Possible time series, on finite intervals

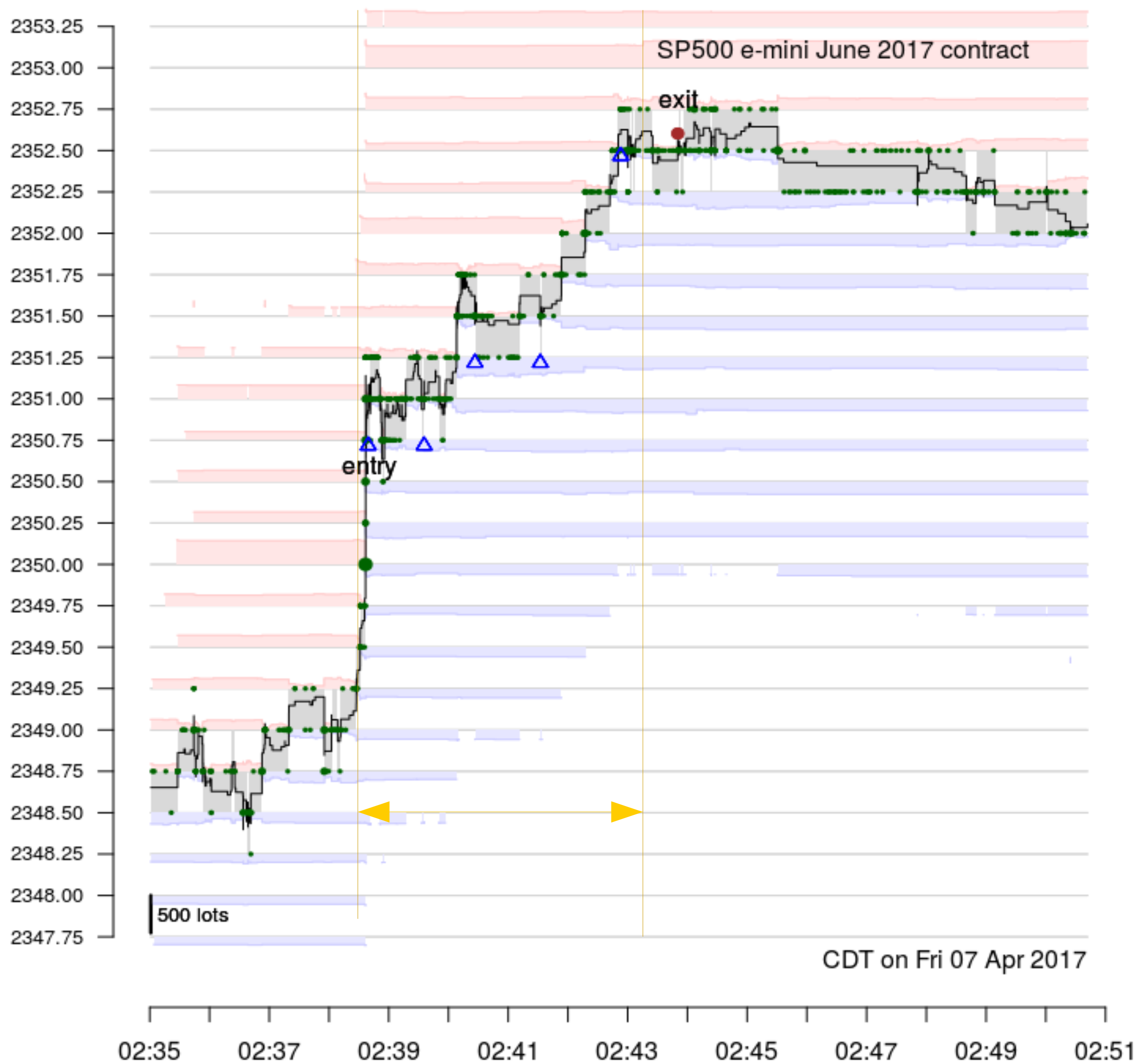
- Mean-reverting

  not realistic (prices have no preferred level)
- Random walk

  only plausible model for overall process
- Explosive

  Can indicate runaway behavior on short times

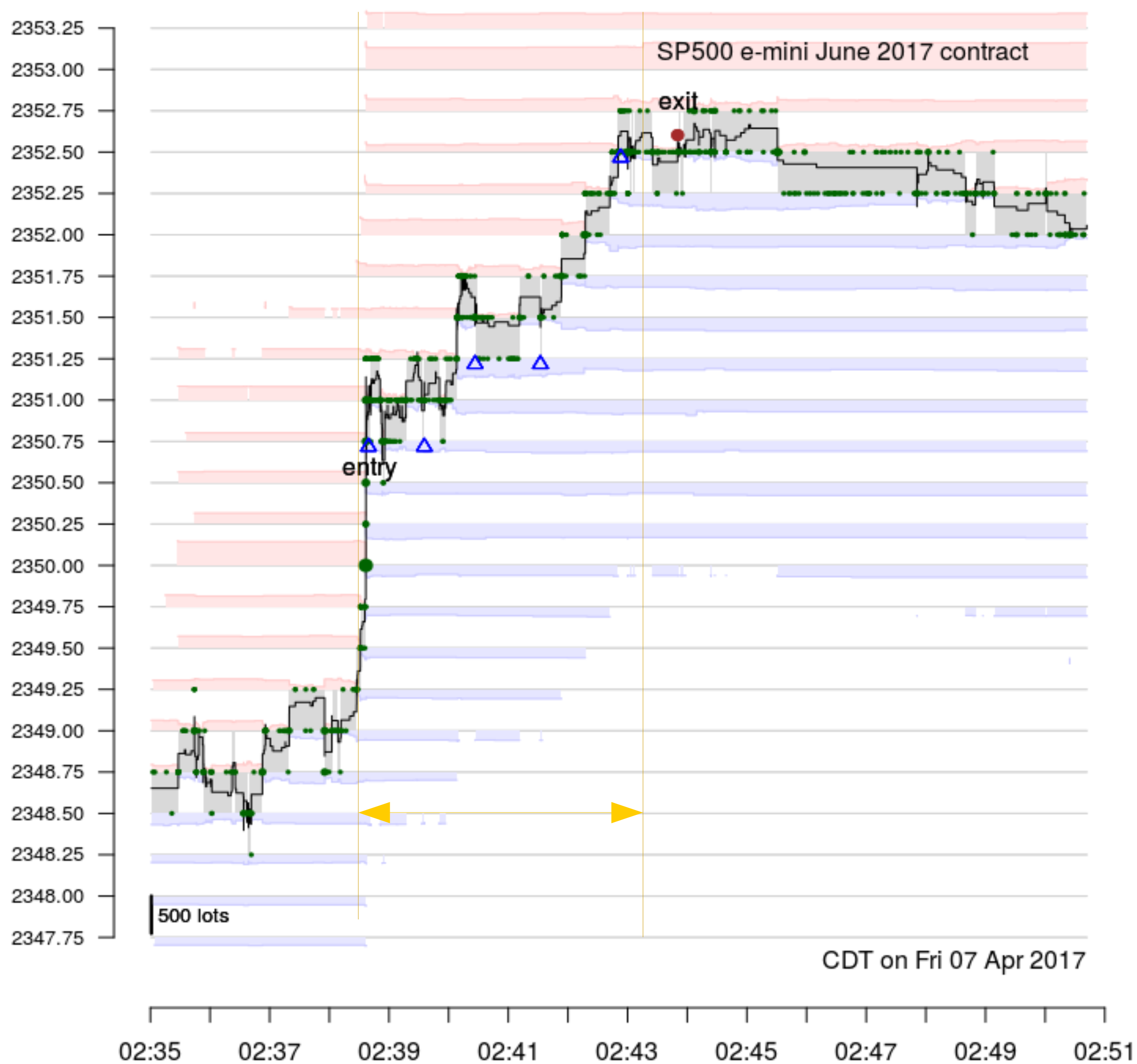
Algorithm tests AR fit on collection of finite intervals, looks for structure across the collection.



Example Buy Signal

- The market was trending up
- Our model correctly identified this and produced a signal about 2 minutes after the rally started (around 2:39 am)
- The signal expired after the price flattened out (around 2:44 am).

Shankar Narayanan,
Quantitative Brokers



Example Buy Signal

- The market was trending up
- Our model correctly identified this and produced a signal about 2 minutes after the rally started (around 2:39 am)
- The signal expired after the price flattened out (around 2:44 am).

Shankar Narayanan,
Quantitative Brokers

Learning to Earn:

An Application of Reinforcement Learning to High-Frequency Trading in the U.S. Futures Market

Kevin Liu

June 2018

Princeton

- **bid-ask spread:** the difference between the bid price and the ask price; representative of the margin of a market maker
- **bid-ask volume misbalance:** the difference between the total bid size and the total ask size
- **percent change in trade price over τ time steps:** current trade price as a proportion of a benchmark trade price executed τ time steps prior
- **percent change in traded volume over τ time steps:** current trade volume as a proportion of a benchmark trade volume executed τ time steps prior

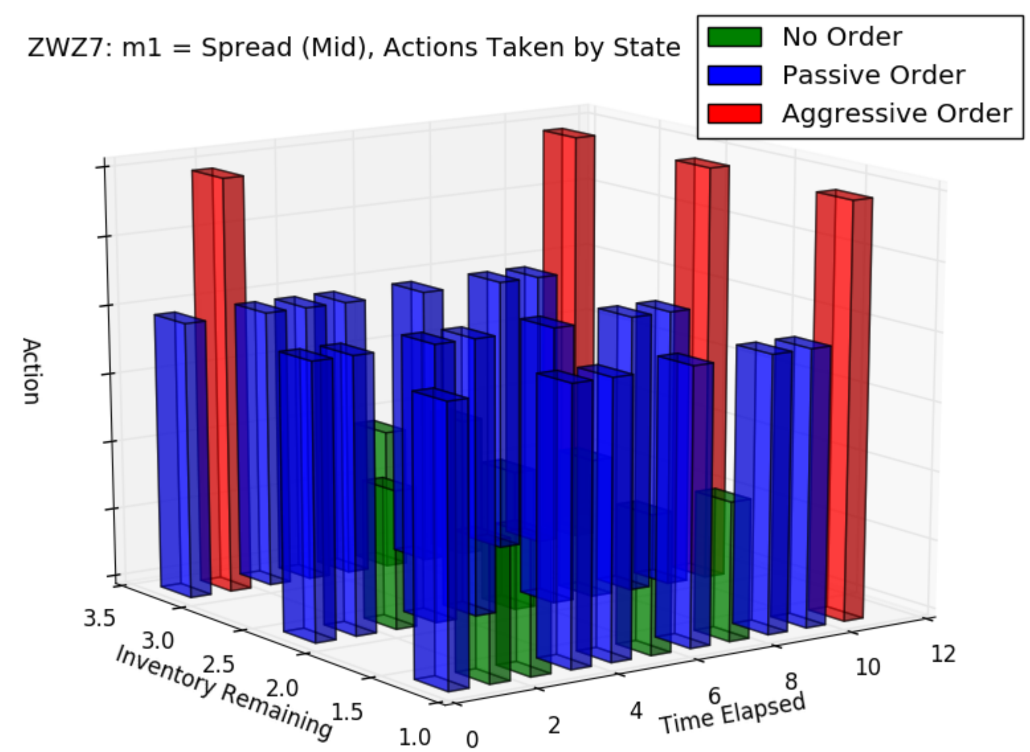


Figure 5.8: Decision rule for one-feature model (with m_1 = spread) at "low" and "medium" imbalance; ZWZ7, 2017

Predicting Changes in the U.S. Treasury Futures Spread During the Roll Period

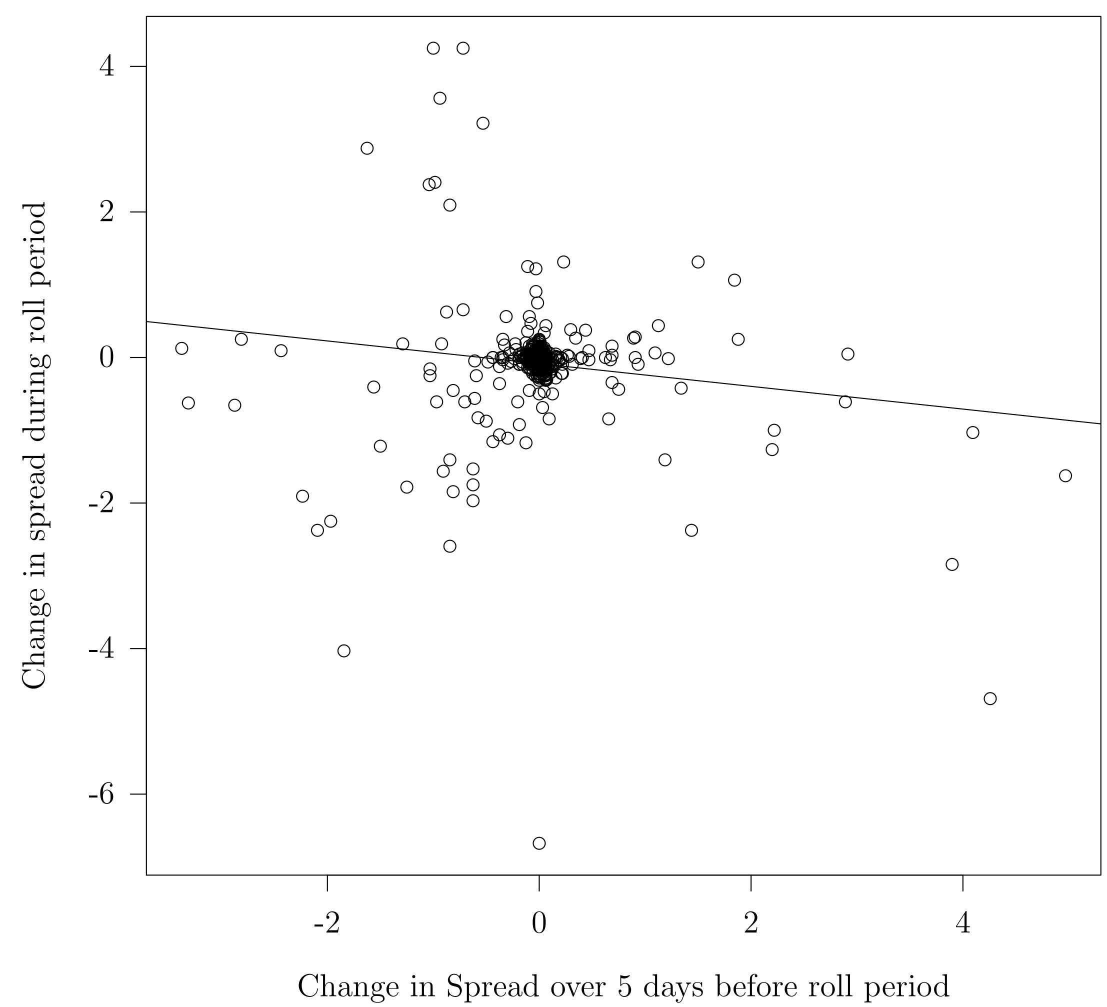
Samuel Russell

June 2018

Princeton

Definition	Formula
xs	Data excluding roll period
l	$length(xs) - 1$
u	$mean(xs)$
std	$\sqrt{((\sum_{n=1}^{l}(xs_n - u)^2))/l}$
Standard Deviation	std
Change in Value over past 5 days	$xs_{l-1} - xs_{l-5}$
u_{10}	$1/10 * (\sum_{n=l-9}^{l}(xs_n))$
std_{10}	$\sqrt{((\sum_{n=l-9}^{l}(xs_n - u_{10})^2))/10}$
(Std. Dev. over past 10 days) / Std. Dev.	std_{10}/std
Exp. Mva. over past 10 days	$\frac{(\sum_{n=l-9}^{l} xs_n * \exp(-.5*(l+1-n)))}{(\sum_{n=l-9}^{l} \exp(-.5*(l+1-n)))}$
Current Value / Mva. past 10 days	$\frac{10*xs_l}{(\sum_{n=l-9}^{l}(xs_n))}$

Reversion in the Spread



Technical Analysis

- Apply simple indicators to price and volume history
- Identify patterns such as

  moving average crossover
  channel breakout
  "head and shoulders"
- Challenge is to filter out high-frequency noise

  look at daily closes

  apply filtering techniques to intraday data
- Conflicts with Efficient Market Hypothesis

  represent belief in momentum or reversion

  not necessarily wrong but one should be skeptical

Example: opening range breakout

- P_t = closing price on day t

Buy on open of day $t+1$ if $\quad P_t - P_{t-1} > \delta$

Sell on open of day $t+1$ if $\quad P_t - P_{t-1} < -\delta$

(Price differences for futures
for equities would use ratios and returns)

more generally:
"channel breakout"

$$M_t = \max\{P_{t-1}, \dots, P_{t-k}\}$$

$$m_t = \min\{P_{t-1}, \dots, P_{t-k}\}$$

Buy on open of day $t+1$ if $\quad P_t - M_t > \delta$

Sell on open of day $t+1$ if $\quad P_t - m_t < -\delta$

Moving average crossovers

$$\bar{P}(t;\Delta) \;=\; \text{Exponential moving average}$$
$$\text{with scale time } \Delta$$

$$R(t) \;=\; \bar{P}(\,t;\,\Delta_{\text{short}}\,) \;-\; \bar{P}(\,t;\,\Delta_{\text{long}}\,)$$

$$\text{Buy if} \quad R(t) > \delta$$
$$\text{Sell if} \quad R(t) < -\delta$$

Bollinger bands

$$B(t) \;=\; \bar{P}(t;\Delta) \;\pm\; k\,\sigma(t)$$

$$\sigma(t) \;=\; \text{some estimate of volatility}$$

Look for breakouts from these bands

Bollinger bands

https://www.bollingerbands.com

Click chart to enlarge

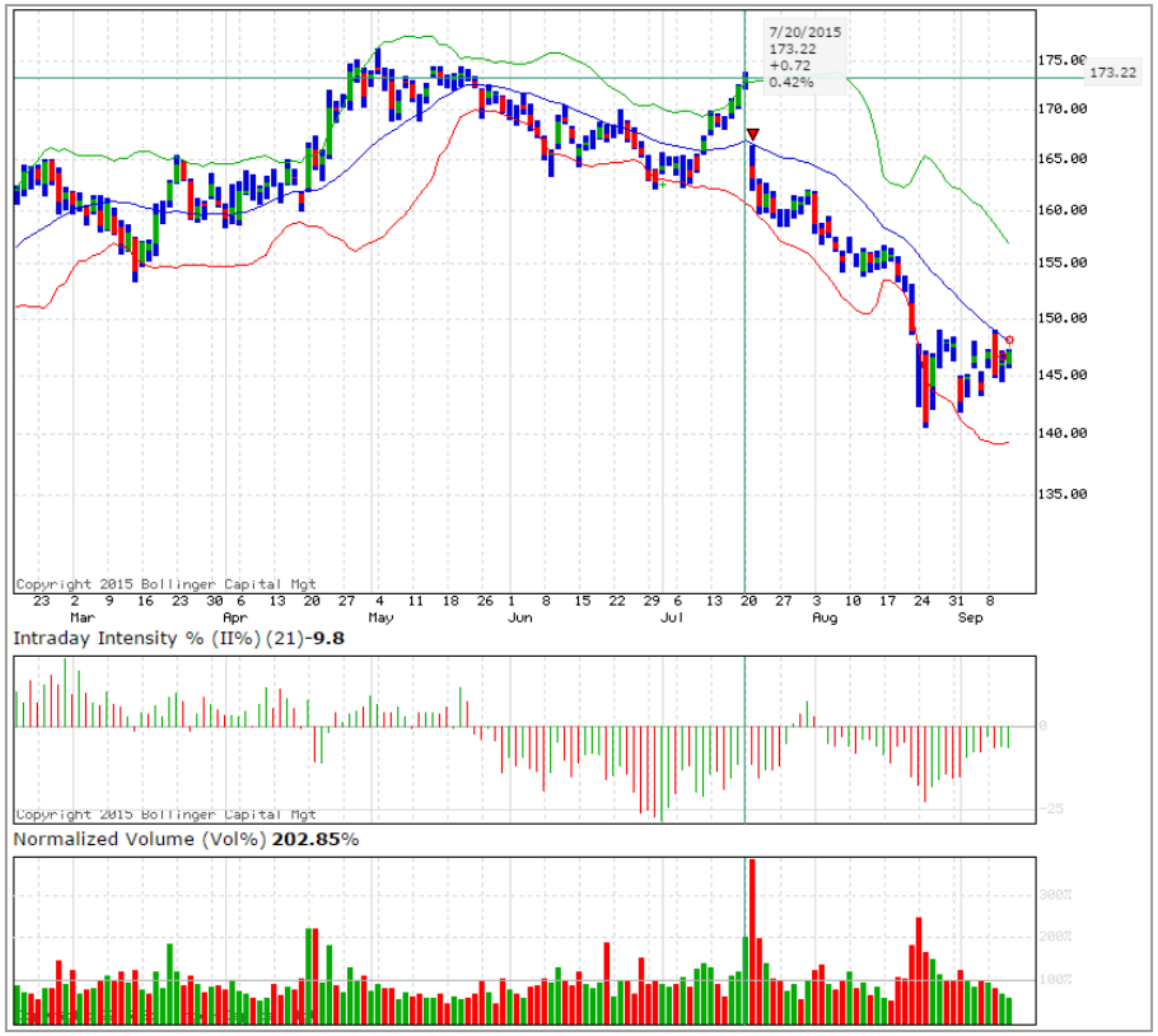


Chart 1: *On 20 July prices tagged the upper Bollinger Band while 21-day Intraday Intensity was deep in negative territory setting up a sell alert. The first down day was the sell signal and entry. The red triangle is a negative PowerShift*

BOLLINGER BANDS

Bollinger Bands are a technical trading tool created by John Bollinger in the early 1980s. They arose from the need for adaptive trading bands and the observation that volatility was dynamic, not static as was widely believed at the time.

Bollinger Bands can be applied in all the financial markets including equities, forex, commodities, and futures. Bollinger Bands can be used in most time frames, from very short-term periods, to hourly, daily, weekly or monthly.

John Bollinger, CFA, CMT
Creator of Bollinger Bands

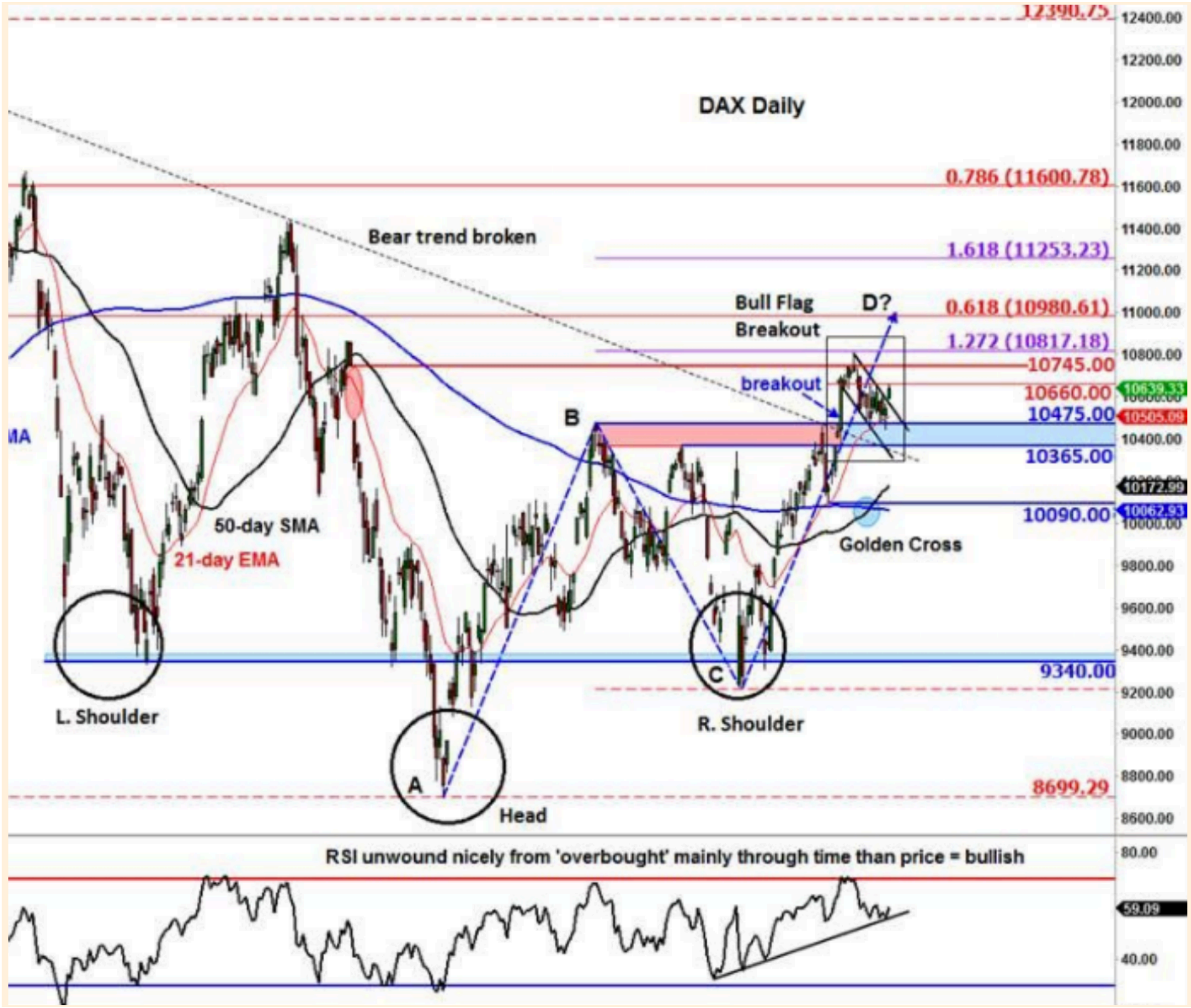
Click chart to enlarge

Chart 2: *A perfect W bottom in Bollinger Band terms set the stage for a sustained uptrend in OHI (weekly bars). Note that %b was much higher on the second retest in October 2013 than it was at the momentum low in August.*

Complex patterns

A Head and Shoulders reversal pattern forms after an uptrend, and its completion marks a trend reversal. The pattern contains three successive peaks with the middle peak (**head**) being the highest and the two outside peaks (**shoulders**) being low and roughly equal. The reaction lows of each peak can be connected to form support, or a **neckline**.

StockCharts
http://stockcharts.com



https://ftalphaville.ft.com/2018/04/19/1524110400000/The-Vomiting-Camel-has-escaped-from-Bitcoin-zoo/

Serious analysis

THE JOURNAL OF FINANCE • VOL. LV, NO. 4 • AUGUST 2000

Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation

ANDREW W. LO, HARRY MAMAYSKY, AND JIANG WANG*

ABSTRACT

Technical analysis, also known as "charting," has been a part of financial practice for many decades, but this discipline has not received the same level of academic scrutiny and acceptance as more traditional approaches such as fundamental analysis. One of the main obstacles is the highly subjective nature of technical analysis—the presence of geometric shapes in historical price charts is often in the eyes of the beholder. In this paper, we propose a systematic and automatic approach to technical pattern recognition using nonparametric kernel regression, and we apply this method to a large number of U.S. stocks from 1962 to 1996 to evaluate the effectiveness of technical analysis. By comparing the unconditional empirical distribution of daily stock returns to the conditional distribution—conditioned on specific technical indicators such as head-and-shoulders or double-bottoms—we find that over the 31-year sample period, several technical indicators do provide incremental information and may have some practical value.

Specifically, our algorithm contains three steps:

1. Define each technical pattern in terms of its geometric properties, for example, local extrema (maxima and minima).
2. Construct a kernel estimator $\hat{m}(\cdot)$ of a given time series of prices so that its extrema can be determined numerically.
3. Analyze $\hat{m}(\cdot)$ for occurrences of each technical pattern.

Apply 2-sided smoothing to price series, and look for patterns of maxima/minima

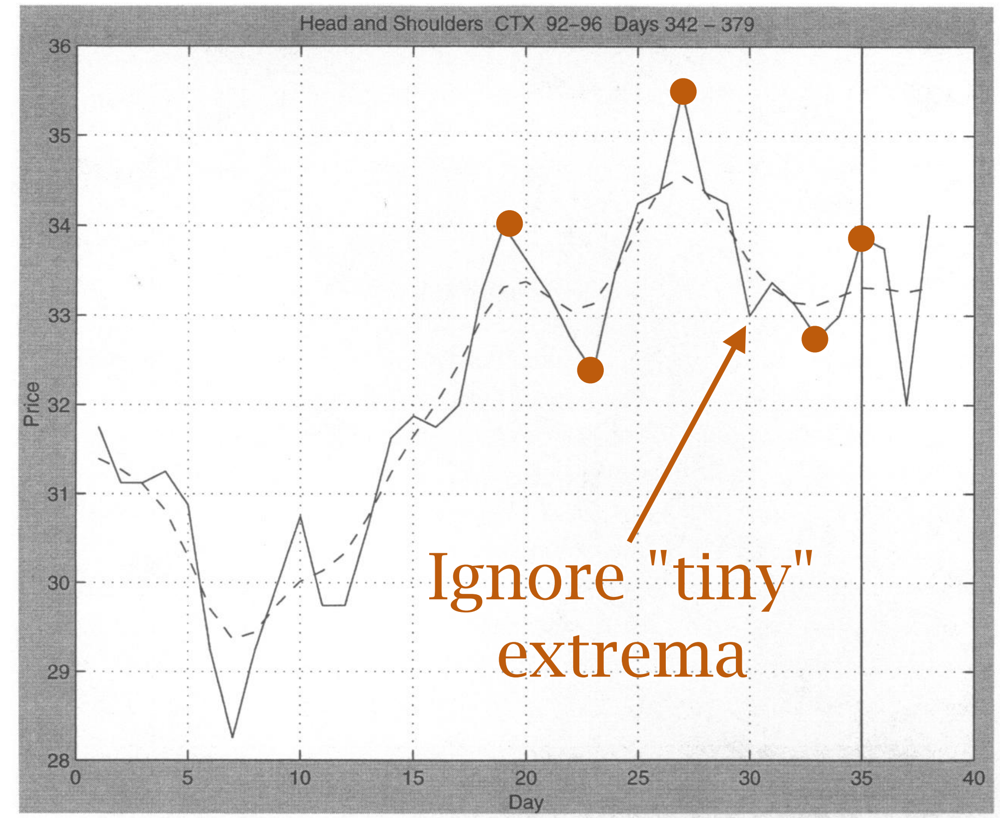
$$\hat{m}_h(\tau) = \frac{\sum_{s=t}^{t+l+d-1} \mathrm{K}_h(\tau - s) P_s}{\sum_{s=t}^{t+l+d-1} \mathrm{K}_h(\tau - s)},$$

Definition 1 (Head-and-Shoulders) Head-and-shoulders (HS) and inverted head-and-shoulders (IHS) patterns are characterized by a sequence of five consecutive local extrema $E_1, \ldots, E_5$ such that

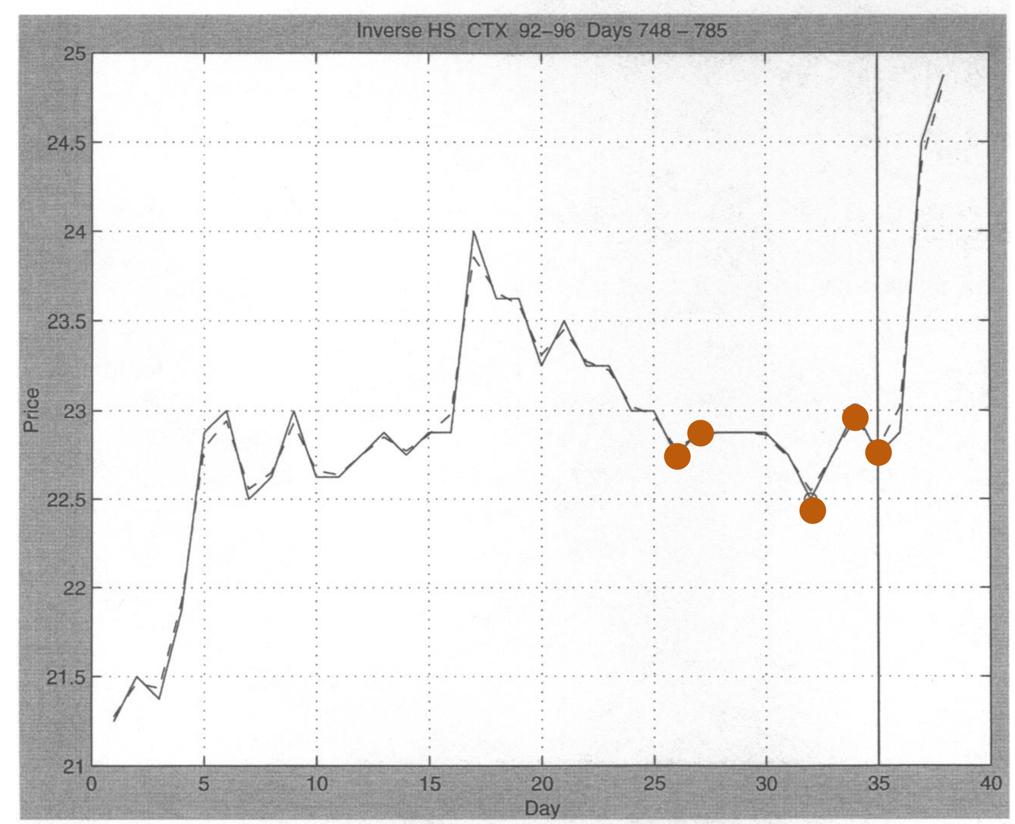
$$\text{HS} \equiv \begin{cases} E_1 \text{ is a maximum} \\ E_3 > E_1,\, E_3 > E_5 \\ E_1 \text{ and } E_5 \text{ are within 1.5 percent of their average} \\ E_2 \text{ and } E_4 \text{ are within 1.5 percent of their average,} \end{cases}$$

$$\text{IHS} \equiv \begin{cases} E_1 \text{ is a minimum} \\ E_3 < E_1,\, E_3 < E_5 \\ E_1 \text{ and } E_5 \text{ are within 1.5 percent of their average} \\ E_2 \text{ and } E_4 \text{ are within 1.5 percent of their average.} \end{cases}$$

Hypothesis: this pattern predicts price reversal



(a) Head-and-Shoulders



(b) Inverse Head-and-Shoulders

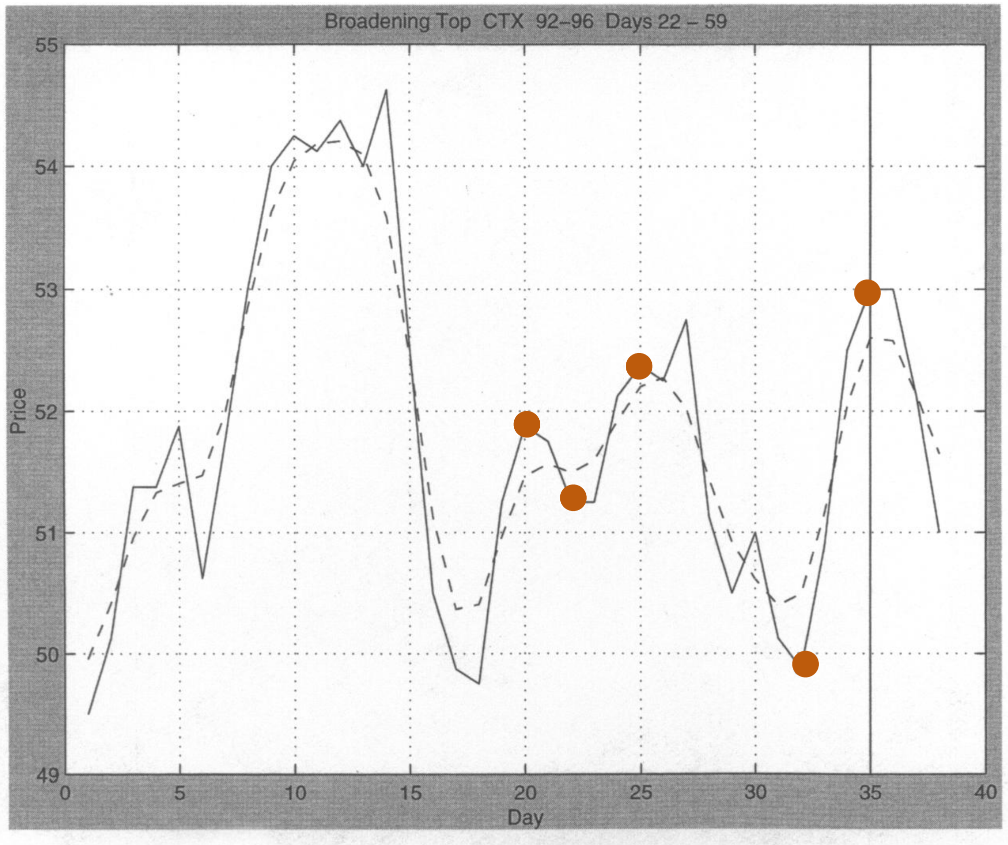
Definition 2 (Broadening) Broadening tops (BTOP) and bottoms (BBOT) are characterized by a sequence of five consecutive local extrema $E_1, \ldots, E_5$ such that

$$\mathrm{BTOP} \equiv \begin{cases} E_1 \text{ is a maximum} \\ E_1 < E_3 < E_5 \\ E_2 > E_4 \end{cases}, \qquad \mathrm{BBOT} \equiv \begin{cases} E_1 \text{ is a minimum} \\ E_1 > E_3 > E_5 \\ E_2 < E_4 \end{cases}.$$

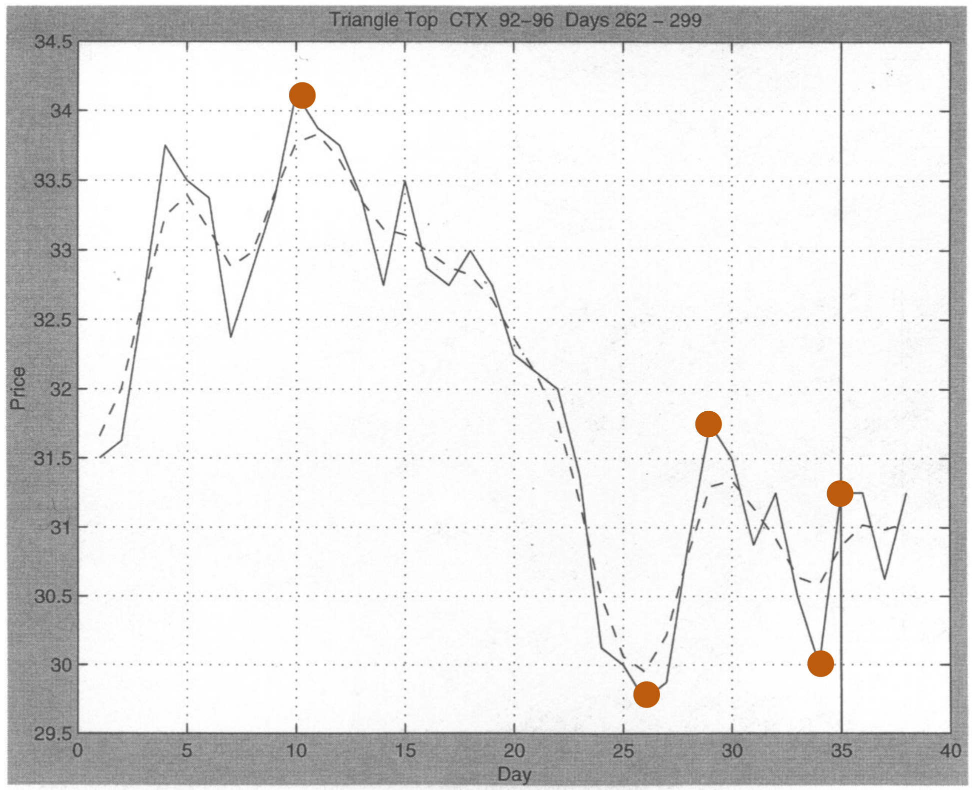
Definitions for triangle and rectangle patterns follow naturally.

Definition 3 (Triangle) Triangle tops (TTOP) and bottoms (TBOT) are characterized by a sequence of five consecutive local extrema $E_1, \ldots, E_5$ such that

$$\mathrm{TTOP} \equiv \begin{cases} E_1 \text{ is a maximum} \\ E_1 > E_3 > E_5 \\ E_2 < E_4 \end{cases}, \qquad \mathrm{TBOT} \equiv \begin{cases} E_1 \text{ is a minimum} \\ E_1 < E_3 < E_5 \\ E_2 > E_4 \end{cases}.$$



(a) Broadening Top



(a) Triangle Top

Other approaches

Patterns in high-frequency FX data: discovery of 12 empirical scaling laws

J. B. GLATTFELDER*†, A. DUPUIS†‡ and R. B. OLSEN†‡

†Olsen Ltd., Seefeldstrasse 233, 8008 Zurich, Switzerland
‡Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, Colchester, Essex, UK

(*Received 27 August 2008; in final form 11 March 2010*)

We have discovered 12 independent new empirical scaling laws in foreign exchange data series that hold for close to three orders of magnitude and across 13 currency exchange rates. Our statistical analysis crucially depends on an event-based approach that measures the relationship between different types of events. The scaling laws give an accurate estimation of the length of the price-curve coastline, which turns out to be surprisingly long. The new laws substantially extend the catalogue of stylized facts and sharply constrain the space of possible theoretical explanations of the market mechanisms.

Quantitative Finance, Vol. 11, No. 4, April 2011, 599–614

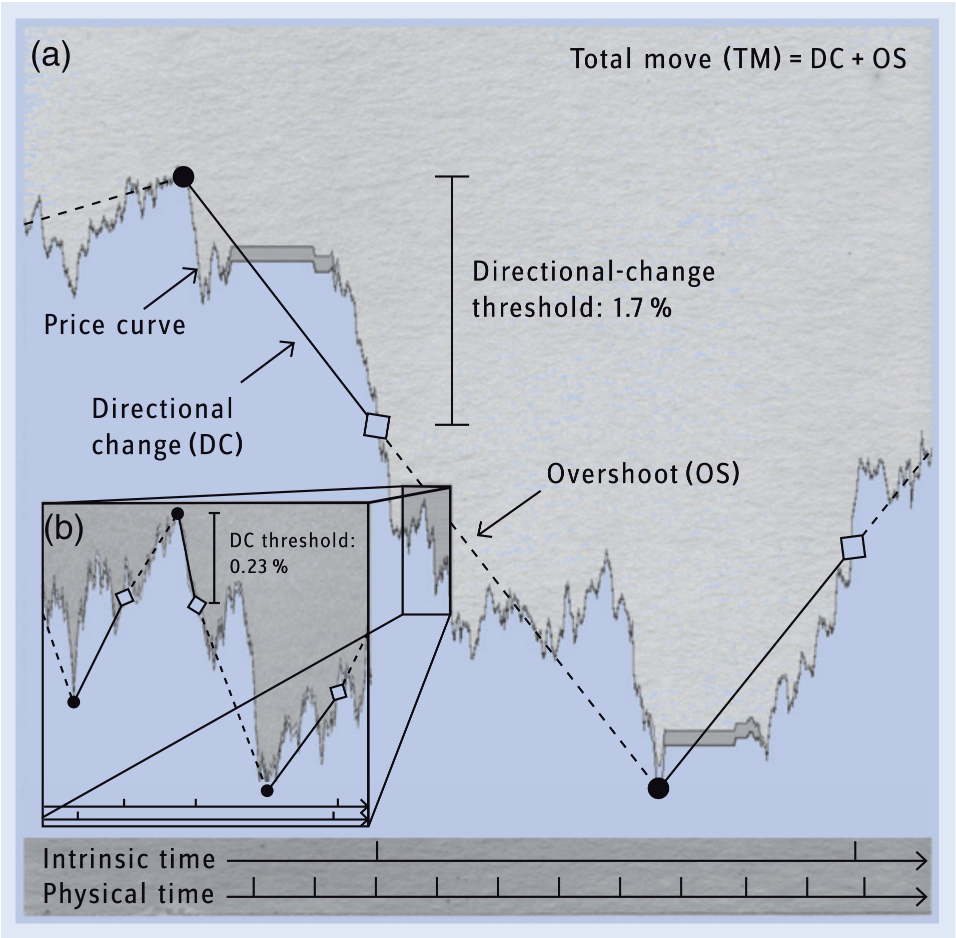


Figure 1. Projection of (a) a two-week and (b) a zoomed-in 36 hour price sample onto a reduced set of so-called directional-change events defined by a threshold: (a) $\Delta x_{\rm dc} = 1.7\%$, (b) $\Delta x_{\rm dc} = 0.23\%$. These directional-change events (diamonds) act as natural dissection points, decomposing a total-price move between two extremal price levels (bullets) into so-called directional-change (solid lines) and overshoot (dashed lines) sections. The directional-change computation is detailed in algorithm **2** of section 3.2. Note the increase of the spread size during the two weekends with no price activity. Time scales depict physical time ticking evenly across different price-curve activity regimes, whereas *intrinsic time* triggers only at directional-change events, independent of the notion of physical time.

Market impact and cost of trading (next time)

- Price prediction may not be useful
  if cannot implement effectively
  - signals are often very weak (fractions of spread)
  - transaction costs can make them unusable
  - need good backtesting or simulation environment
- Two kinds of trading costs:
  - bid-ask spread, passive fill probability
  - market impact from large trades