

Market Microstructure and Algorithmic Trading

PIMS Summer School 2016
University of Alberta, Edmonton
Lecture 2: July 4, 2016

Robert Almgren



quantitativebrokers

Exchanges

Decide trading rules:

Trading hours

Match algorithm

Implied quoting

Trading halts / circuit breakers

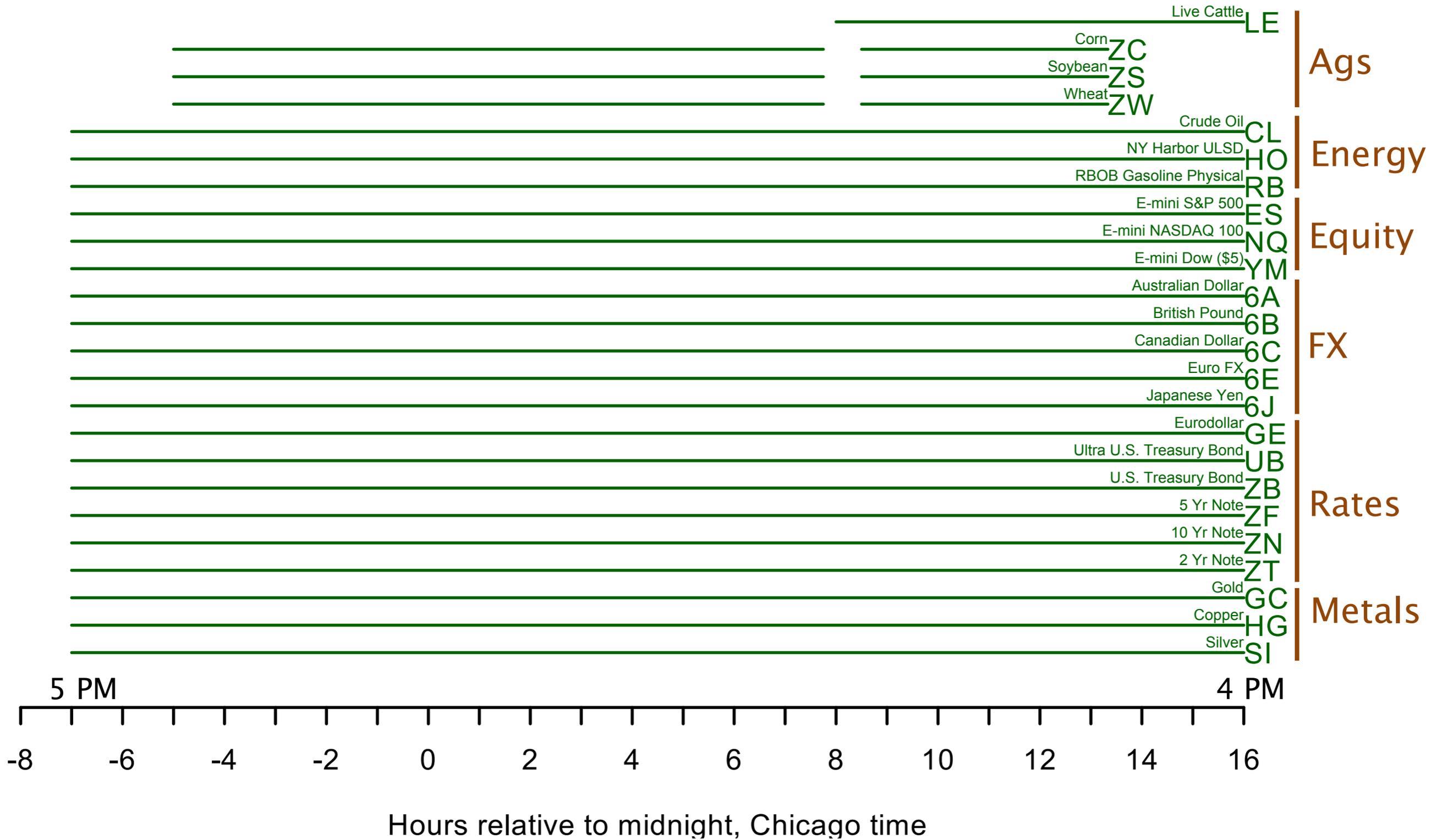
Tick size

exchange decision

effect on trading: reversion

volatility measurement

Trading hours (CME futures)



Trading hours

THE WALL STREET JOURNAL.

CME Group Cuts Grain, Soy Futures Trading to 17½ Hours

By IAN BERRY And JACOB BUNGE

Updated March 5, 2013 12:41 p.m. ET

CHICAGO— [CME Group](#) Inc. will reduce its trading day in grain and soybean futures to 17½ hours from 21 hours, largely undoing a controversial expansion done last year.

The new schedule trims afternoon hours and creates a 45-minute pause in trading in the morning before open-outcry, or "pit," trading begins, exchange officials said Tuesday.

The new hours come in response to criticism that the trading schedule had become bloated, increasing costs for grains traders and creating periods of low-volume, highly volatile trade.

Chicago-based CME's struggle over how to manage one of its oldest markets reflects the delicate balance the world's largest futures exchange operator must strike between keeping key customers happy and fending off competitors in Europe and the U.S.

While longer trading sessions meant less time for big grain merchandisers like [Bunge Ltd.](#) and [Cargill Inc.](#) to square their books between sessions, some smaller grain elevators had embraced a trading day that more closely tracked their grain-buying schedule, according to Mike Hall, president of brokerage [MLH Futures Inc.](#), based in Litchfield, IL.

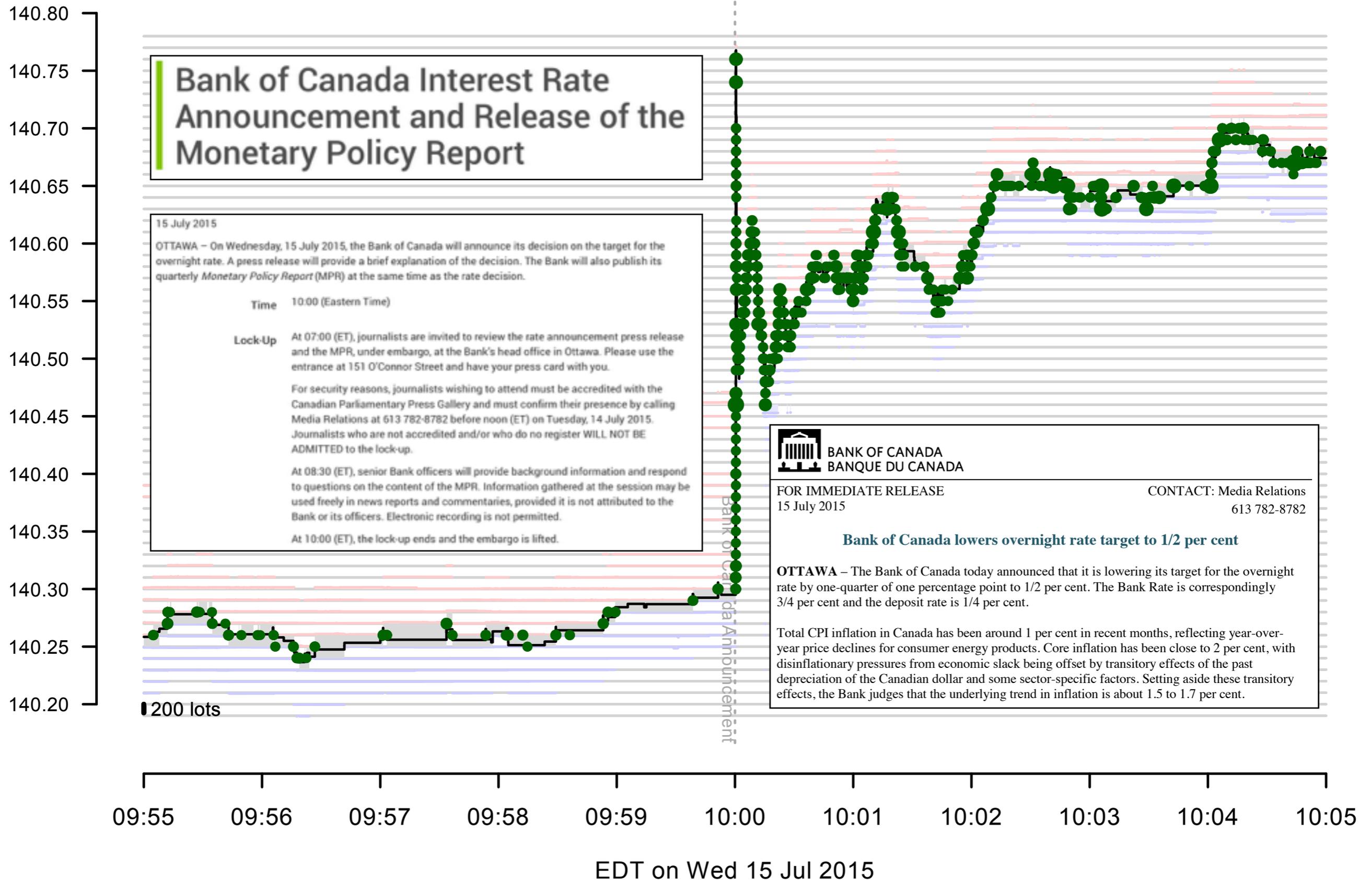
Danny Murphy, president of the American Soybean Association, a trade group representing soybean farmers, said his members' top concern remains the release of key crop reports from the U.S. Department of Agriculture during trading hours. Before the lengthier sessions introduced by CME and ICE, brokers and grain companies had hours to parse the reports before trading began. Since the change, some have complained that releasing the reports amid live trading has triggered wild price swings.

"We still would like to see a pause during trading," said Mr. Murphy, who farms soybeans, corn and wheat near Canton, Miss. CME officials have said they would consider a pause if ICE implements one as well, while ICE has said that it makes more sense to have markets open for those who want to place trades immediately, or look to the futures market for a reference price.



Trading through a crop report

CGBU5: 10-year Canadian note futures



Matching algorithms

How market orders are matched to limit

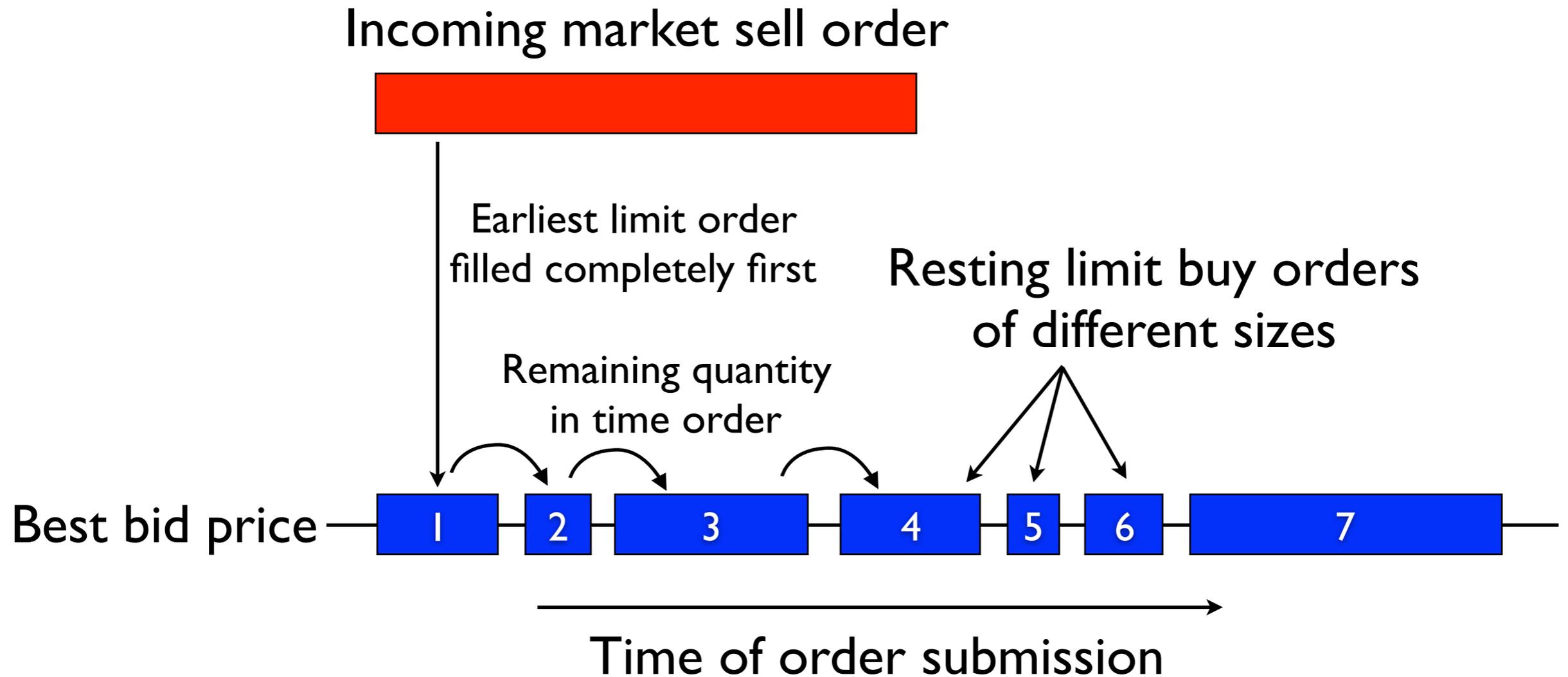
Algorithm fixed by exchange

to attract more volume

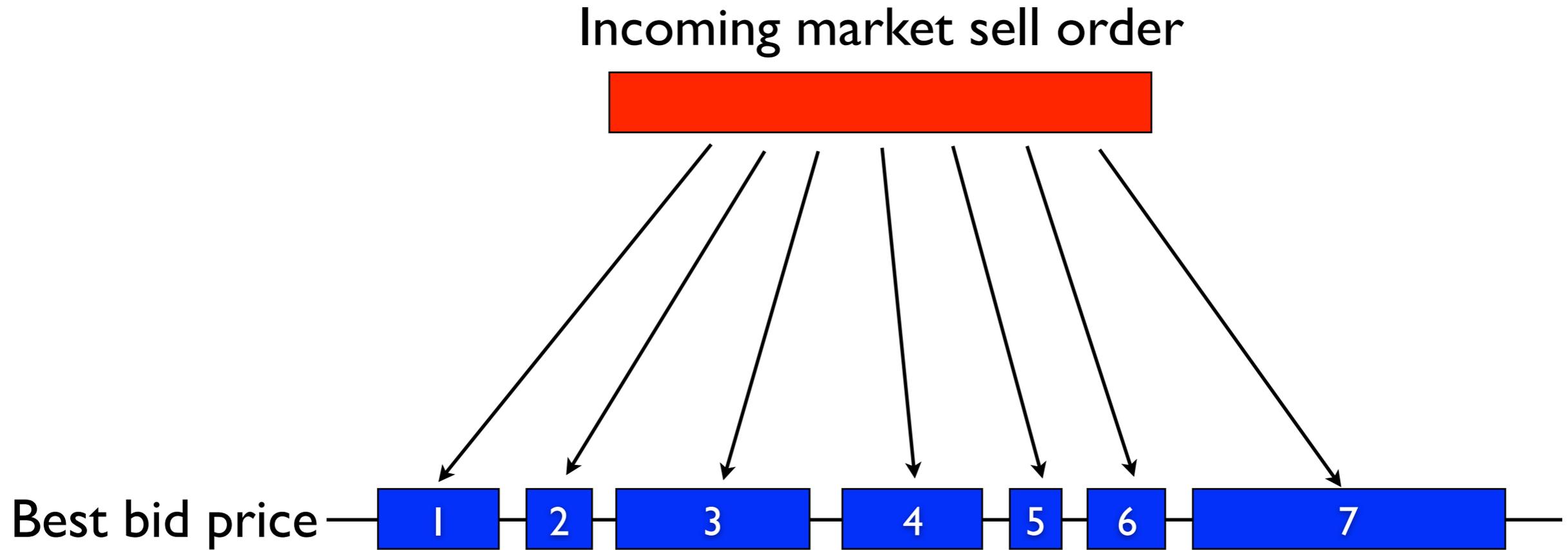
to attract correct mix of participants

etc

First-in first-out (FIFO) order



Pro rata order matching



Incoming volume divided
among *all* resting orders
at best price

Matching Algorithms

This topic explains the different matching algorithms supported on the CME Globex Platform. These matching algorithms support order management functionalities offered to market participants and ensure that each market participant is given the best possible execution at the fairest price.

A matching algorithm is a technique to allocate matched quantities, used when an **aggressor order** matches with one or multiple resting orders. Algorithms apply to both outright and **implied matching**.

Order matching follows three steps:

1. Determine the current prices opposite of the aggressor order, outright and implied prices
2. Determine quantity at best price opposite the aggressor order, based on business rule priorities
3. Allocate resting quantity at best price to trade with aggressor order using the market algorithm

The following are the supported matching algorithms:

Algorithm	tag 1142-MatchAlgorithm Value
Allocation	A
Eurodollar Option	Y
FIFO	F
FIFO with LMM	T
FIFO with Top Order and LMM	S
Pro-Rata	C
Split FIFO and Pro-Rata	K
Threshold Pro-Rata	O
Threshold Pro-Rata with LMM	Q

9 different ways to match market and limit orders, on CME alone

"The FIFO algorithm uses price and time as the only criteria for filling an order. In this algorithm, all orders at the same price level are filled according to time priority; the first order at a price level is the first order matched."

CME Eurodollar matching

1. Orders placed during the “pre-opening” or at the indicative opening price (IOP) will be matched on a price and time priority basis. **Note that implied orders are not taken into consideration, as they are only active during the continuous trading session.**
2. Priority is assigned to an order that betters the market, i.e. a new buy order at 36 betters a 35 bid. Only one order per side of the market (buy side and sell side) can have this TOP order priority. There will be situations where a TOP order doesn't exist for one or both sides of the market (for example, an order betters the market, but is then canceled). There will never be a situation that results in two orders on the same side of the market having TOP order status.
3. Only outright orders can be TOP orders, however the TOP orders of underlying orders that are creating implied orders will be taken into consideration during the matching process so as not to violate the TOP order rule in any market.
4. TOP orders are matched first, regardless of size.
5. After a TOP order is filled, Pro Rata Allocation is applied to the remainder of the resting orders at the applicable price levels until the incoming order is filled.
6. The Pro Rata algorithm allocates fills based upon each resting order's percentage representation of total volume at a given price level. For example, an order that makes up 30% of the total volume resting at a price will receive approximately 30% of all executions that occur at that price. Approximate fill percentages may occur because allocated decimal quantities are always rounded down (i.e. a 10-lot order that receives an allocation of 7.89- lots will be rounded down to 7-lots).
7. The Pro Rata algorithm will only allocate to resting orders that will receive 2 or more contracts.
8. After percentage allocation, all remaining contracts not previously allocated due to rounding considerations are allocated to the remaining orders on a FIFO basis.
 - Outright orders will have priority over implied orders and will be allocated the remaining quantity according to their timestamps.
 - Implied orders will be then allocated by maturity, with the earliest expiration receiving the allocation before the later expiring contracts. If spread contracts have the same expiration (i.e., CONTRACT A-CONTRACT B and CONTRACT A-CONTRACT C), then the quantity will be allocated to the earliest maturing contracts making up that spread (i.e., the CONTRACT A-CONTRACT B will get the allocation before the CONTRACT A-CONTRACT C because the CONTRACT B expires before the CONTRACT C).

Reasons for pro rata matching

Reasons for pro rata matching:

- Historical tradition from pit trading

- Encourage submission of large limit orders

- Allow late entrants to participate

Characteristic of interest rate futures markets

- Eurodollar, Euribor, Treasury calendar spreads

- Short sterling

“Arms race” to oversize limit orders

- limited only by risk of overfilling

- (Jonathan Field & Jeremy Large 2008)

Advantage Futures Chicago

IS PRO RATA AN ACCIDENT WAITING TO HAPPEN?



Special Report by Ginger Szala

The New Year surprise happened fast: the Swiss National Bank announced on January 15 that the Swiss Franc would no longer be capped by the Euro and it jumped 41% in minutes. This black swan event happened just before a U.S. holiday when traders were already winding down for a three-day weekend. Those left trading included a combination of retail players, banks and some prop firms. And the market went into a frenzy.

Fortunately, the FX outright futures at CME utilizes FIFO and not Pro Rata for its fill distribution methodology. If the FX market had used pro rata, the carnage from the Swiss National Bank event could have been even more disastrous.

The fixed income markets at CME utilize pro rata. Had the unexpected rate cut effectuated by the Bank of Canada taken place in the US by our Federal Reserve Board, the dislocation and market chaos created could have been even more extreme. This is because many traders in pro rata markets enter bids and offers far in excess of the quantity they actually expect to buy or sell. They may bid for 1000, hoping to buy a 20 lot when the bid is hit for 100 and the matching algorithm splits up the sale quantity partly based upon a pro rata methodology.

Is there a solution?

The Fed's paper by McPartland studies the NYSE/LIFFE Time Pro Rata algorithm and concludes that pro rata algorithms would be better using a cardinal ranking (rather than an ordinal one) on resting bids and offers "based on the actual length of time that bids and offers have been resting in the order book relative to the time that all of the other orders have been resting in the order book." Durkin states in CME letter to the CFTC that it doesn't use Time-weighted Pro Rata, and it doesn't appear it plans to change that anytime soon.

To be fair, most exchanges use certain combinations of the two methods. But in markets like Eurodollars, which uses the pro rata algorithm, and could see a spike similar to the Swiss Franc with macro economic changes coming into play, problems could ensue. Still, it seems CME Group has taken precautions, providing software to aid trading firms and brokers in pre risk management, requiring kill switches and Message Volume Controls.

<http://www.advantagefutures.com/is-pro-rata-an-accident-waiting-to-happen/>

LIFFE volume pro rata (post-2007)

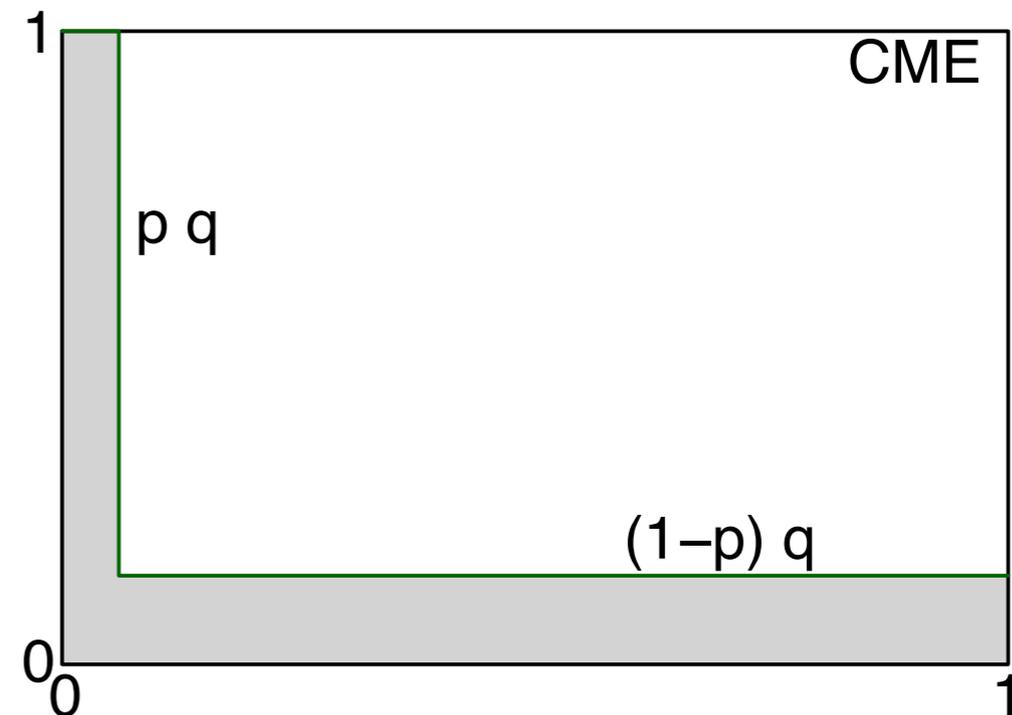
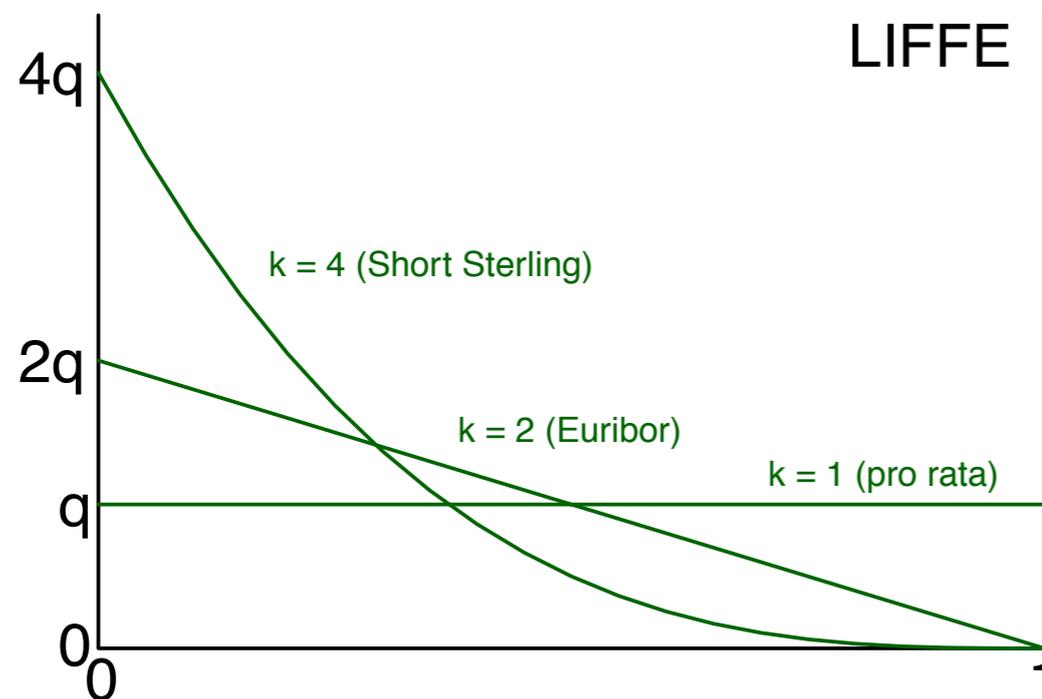
$$f_j = \frac{(V - P_{j-1})^k - (V - P_j)^k}{V^k}.$$

P_j = volume preceding order j

$k=1$: pro rata

$k=\infty$: time priority

2-year Treasury
& some Treasury calendar spreads



Market order size as fraction of resting limit orders

Tick size

CME Group Announces Increase in Minimum Tick Size for 30-Year U.S. Treasury Bond Futures

--Change in Response to Customer Feedback: Aimed to Increase Participation

CHICAGO, July 1, 2009 /PRNewswire-FirstCall via COMTEX News Network/ -- CME Group, the world's largest and most diverse derivatives exchange, today announced plans to increase the minimum trading increment for U.S. Treasury Bond futures to 1/32nd from the current setting of 1/2 of 1/32nd, effective August 30. The change will be applied to all expiration months. The minimum trading increments for futures intermonth and intercommodity spreads as well as options will be unchanged. This contract is listed with, and subject to, the rules and regulations of CBOT.

"There is a renewed focus on the 30-Year contract due to the recent uptick in issuance in the long end of the Treasury curve," said Robin Ross, CME Group Managing Director of [Interest Rate Products](#). "Customer feedback indicates that increasing the tick size will broaden participation from active traders who provide much needed liquidity to this important sector of the Treasury market."

<http://investor.cmegroup.com/investor-relations/releasedetail.cfm?ReleaseID=393362>

**MODIFICATIONS TO THE CONTRACT SPECIFICATIONS FOR THE
THREE-MONTH CANADIAN BANKERS' ACCEPTANCE FUTURES CONTRACT**

MINIMUM PRICE FLUCTUATION

CIRCULAR 037-14
March 19, 2014

I. OVERVIEW

Bourse de Montréal Inc. (the Bourse) hereby proposes to amend the minimum price fluctuation (tick size) on the contract specifications for the Three-Month Canadian Bankers' Acceptance Futures (BAX), such that the minimum price fluctuation for the second, third and fourth quarterly BAX contract months be reduced from 0.01 per \$100 nominal value (a full tick), to 0.005 per \$100 nominal value (a half tick).

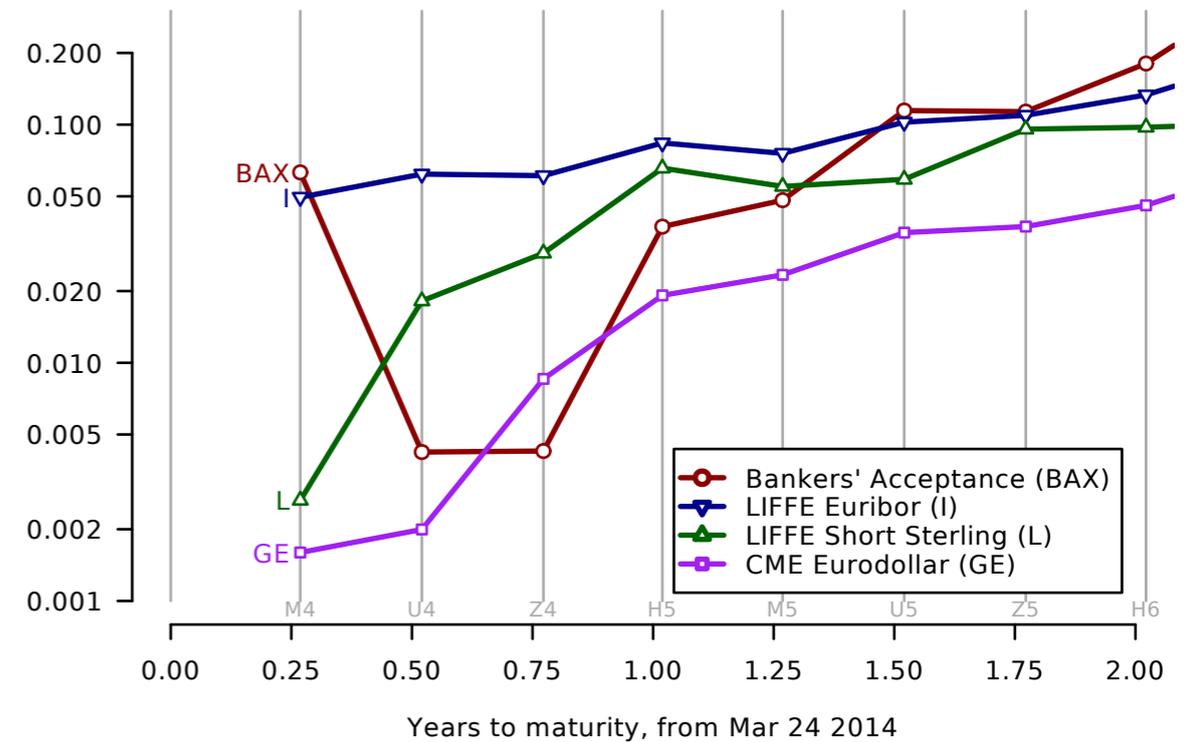
Over the past several years, the Bourse has received repeated requests from end-user participants (such as pension funds, central banks, hedge funds, treasuries and dealers) to extend the minimum price fluctuation of a half tick to at least the six nearest listed contract months, including serials. Domestic and international clients have echoed those sentiments.

Firstly, the feedback received focused primarily on an expected reduction in the cost of trading. Full ticks were deemed too costly to hedge given the current low volatility environment.

Secondly, another benefit of a half tick minimum price fluctuation would be an increase in diversity amongst participants.

Thirdly, half ticks would optimize the BAX in the context of the changing competitive landscape.

Extending the half tick minimum price fluctuation to the next three listed contract months will reduce the profitability per trade for liquidity providers. However, this reduction in profit per trade will be offset by an increase in trading activity. Liquidity providers, in the BAX market, usually place bids and offers in the order book passively waiting for end-users to enter the market and instantaneously get filled. End-user participants, unwilling to pay a full tick for a fill, are simply placing their orders in the book and waiting for a fill to come along, or are looking for fills in competing markets. The result is an order book with a large number of resting orders. A half tick minimum price fluctuation would increase the likelihood of end-user participants lifting offers or hitting bids that liquidity providers have placed in the order book. Therefore, while profitability per trade will decrease, the number of profitable trades will increase.



**Response to Proposed
BAX Price Increment Reduction**

Robert Almgren,* Khalil al Dayri,† Mathieu Rosenbaum,‡ and Shixiang Zhang*

April 18, 2014

On March 19, 2014, the Montréal Exchange issued Circular 000-14, "Request for Comments: Modifications to the contract specifications for the three-month Canadian Bankers' Acceptance Futures Contract" (BAX). The proposal is "... that the minimum price fluctuation for the second, third and fourth quarterly BAX contract months be reduced from 0.01 per \$100 nominal value (a full tick), to 0.005 per \$100 nominal value (a half tick)." We are writing in support of this proposed change.

Robert and Rosenbaum [2011] and Dayri and Rosenbaum [2012] have identified a key parameter in determining whether the minimum price increment of a traded asset is appropriate for the natural dynamics of the price process. This parameter, identified as η , is meaningful for "large-tick" assets and quantifies the aversion to price changes of market participants. It is estimated by measuring the frequency of reversals in trade prices. Values of η near 1/2 are characteristic of markets in which the minimum price increment is small enough that it is not significant for trading. Small values of η are characteristic of markets in which the price increment size is extremely large compared to the intrinsic price dynamics, and represents a substantial constraint on traders.

David Weild, Edward Kim and Lisa Newport September 2012

The trouble with small tick sizes

Larger tick sizes will bring back capital formation, jobs and investor confidence

“That silent whir that you hear on the trading floors of Goldman Sachs, Morgan Stanley and Credit Suisse is the post-apocalyptic sound of an oxygen-deprived, computer-dominated trading floor that has been reengineered to respond to an infestation of tiny ticks.”

Why some large investment banks, large investors and stock exchanges fight for smaller tick sizes, despite their negative impact on the economy

David Weild
Grant Thornton LLP
and former vice chairman of NASDAQ

Tick proliferation and quote flickering damaged the economy

Tick proliferation,¹³ which has led to a loss of economic incentives to make markets, and quote flickering,¹⁴ are the flesh-eating bacteria of the infrastructure needed to support the IPO market and aftermarket.

¹³ Tick proliferation is the decrease in tick sizes.

¹⁴ Quote flickering is measured by the rapid and repeated updates to the National Best Bid and Offer (NBBO).

Beware of the hidden agendas of those who champion smaller tick sizes

http://www.grantthornton.com/staticfiles/GTCom/Public%20companies%20and%20capital%20markets/Trouble_Small_Ticks.pdf

Tick size effects

What is "large tick" or "small tick"?

Nondimensional parameters

Reversion in general

Reversion models: Robert/Rosenbaum (2011)

Nondimensional number

How to compare different products

5000 US stocks

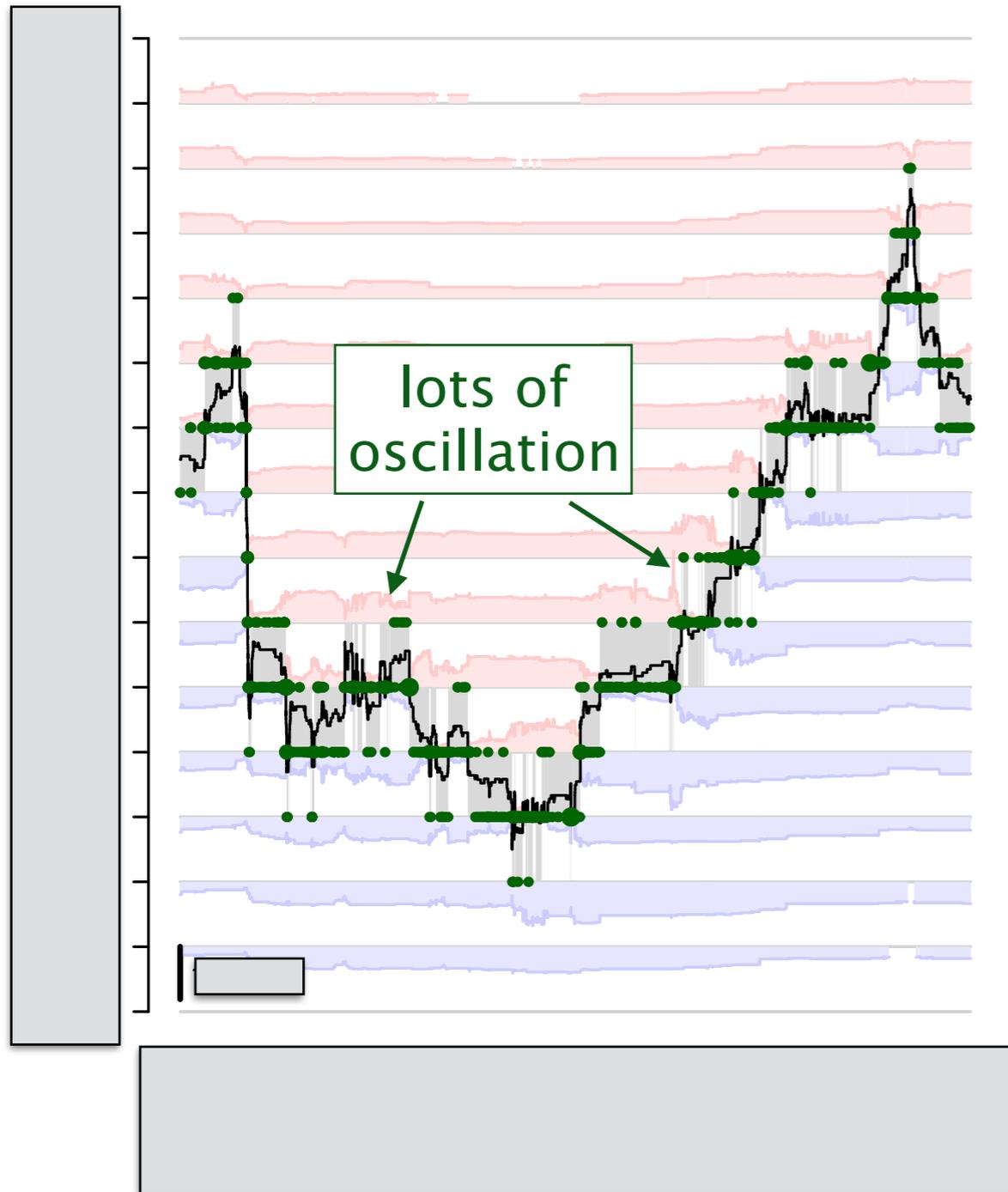
US and international stocks

100's of different futures on CME and worldwide

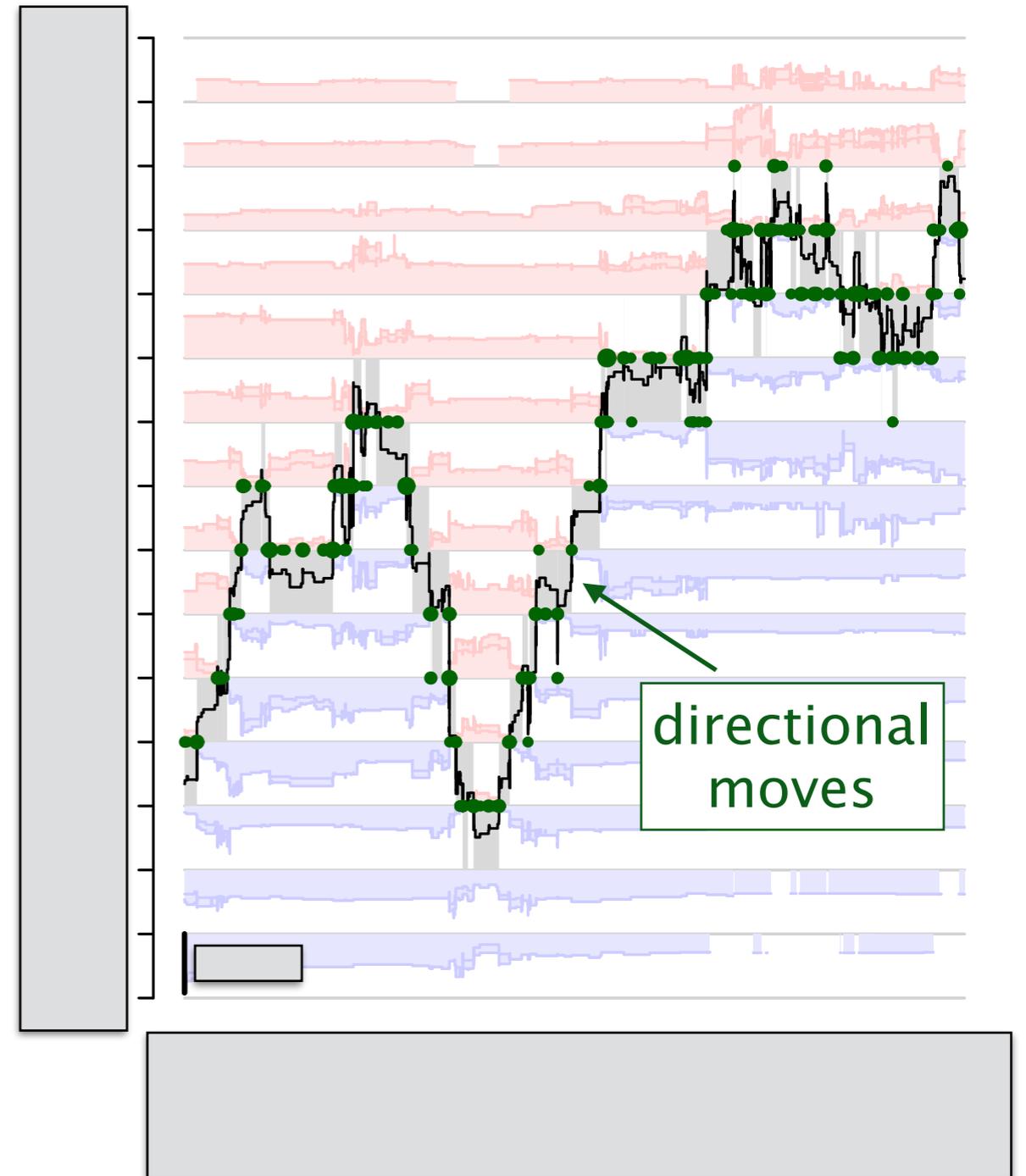
Natural Gas is "different" than 2-year Treasury

Can you tell the difference?

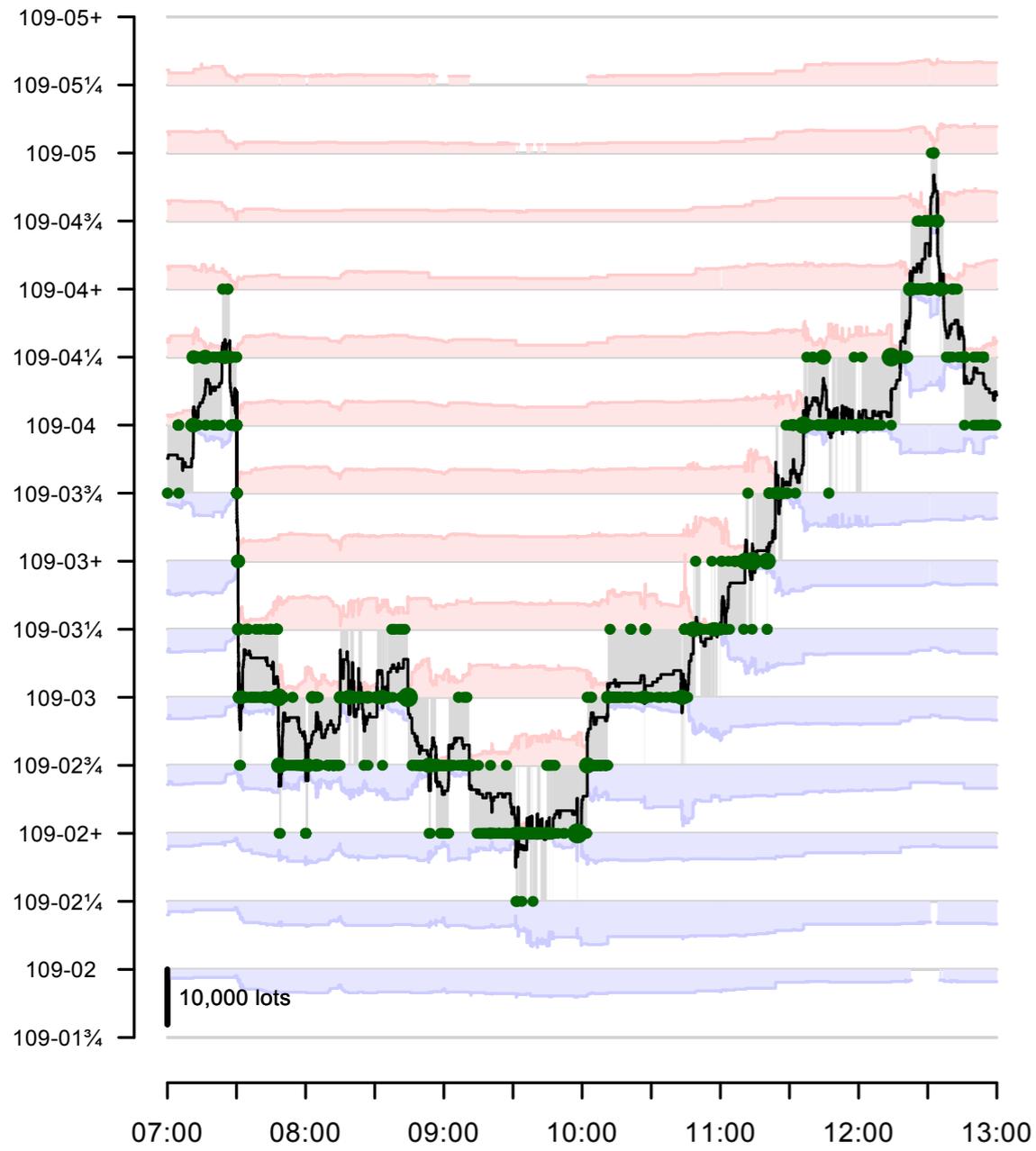
Futures product A



Futures product B



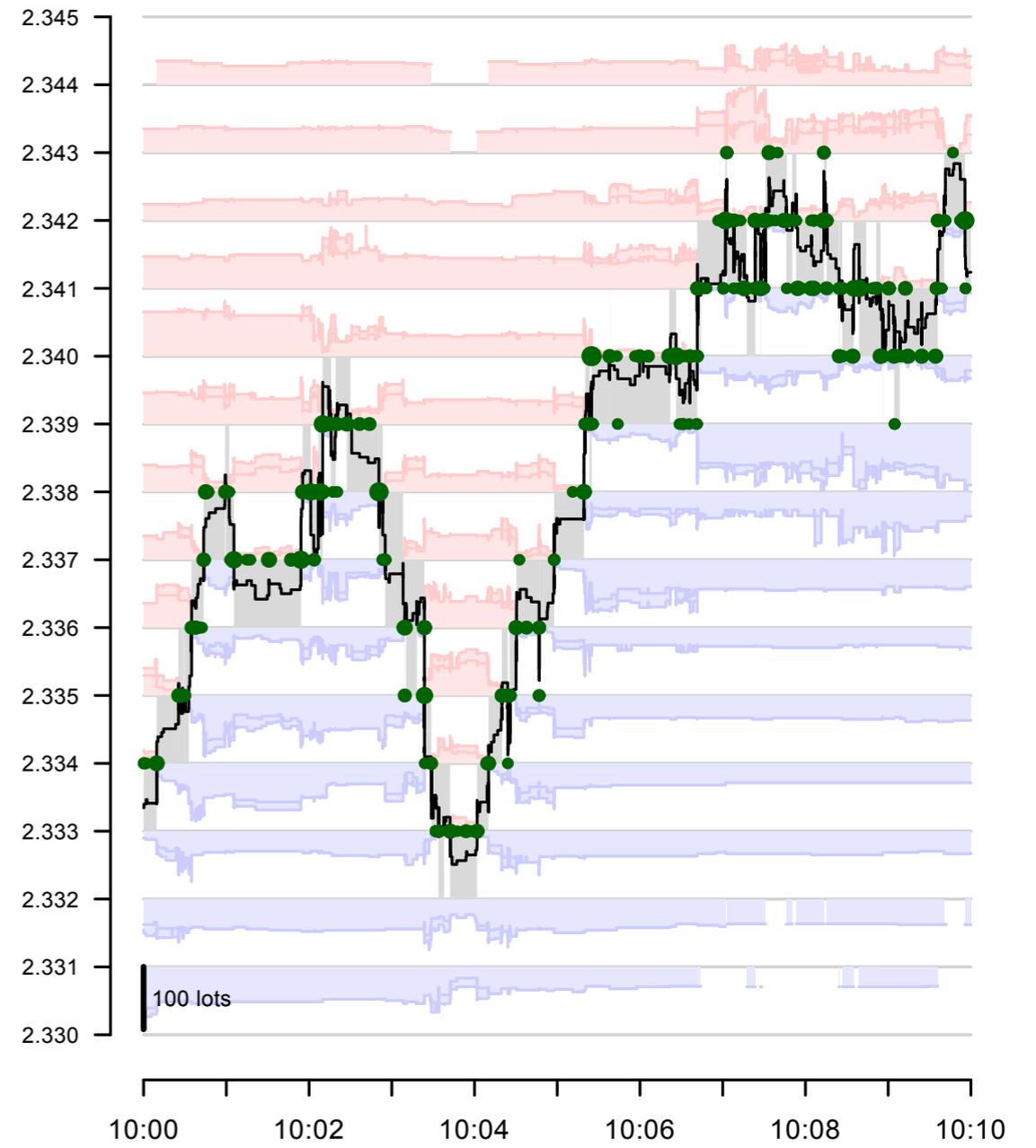
2-year Treasury ZTZ5



CST on Tue 17 Nov 2015

6 hours

Natural Gas NGZ5



CST on Tue 17 Nov 2015

10 minutes

Dimensional parameters

Have units in them

volatility: price change per $\sqrt{\text{time}}$

\$ for stock,
maybe nondim for futures

second, hour, or day

daily volume: shares or lots per day

dimensional

dimensional

Comparison

Stock A: 1MM shares per day

Stock B: 2MM shares per day

Is trading 5,000 shares of stock A in one day the "same" as trading 10,000 shares of stock B in one day?

Is trading 10,000 shares of stock B in one day the "same" as trading 10,000 shares of stock A in two days?

Nondimensionalization

Example: market impact

Trade X shares in time T

Price impact $I = ?$

Nondimensional
exponent and coefficient

$$\frac{I}{\sigma} = \alpha \left(\frac{X}{V} \right)^k$$

Impact as fraction
of daily volatility

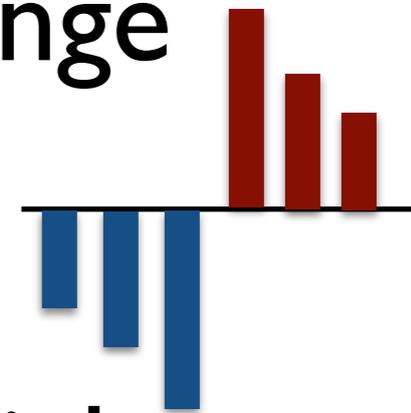
Trade size as fraction
of daily volume

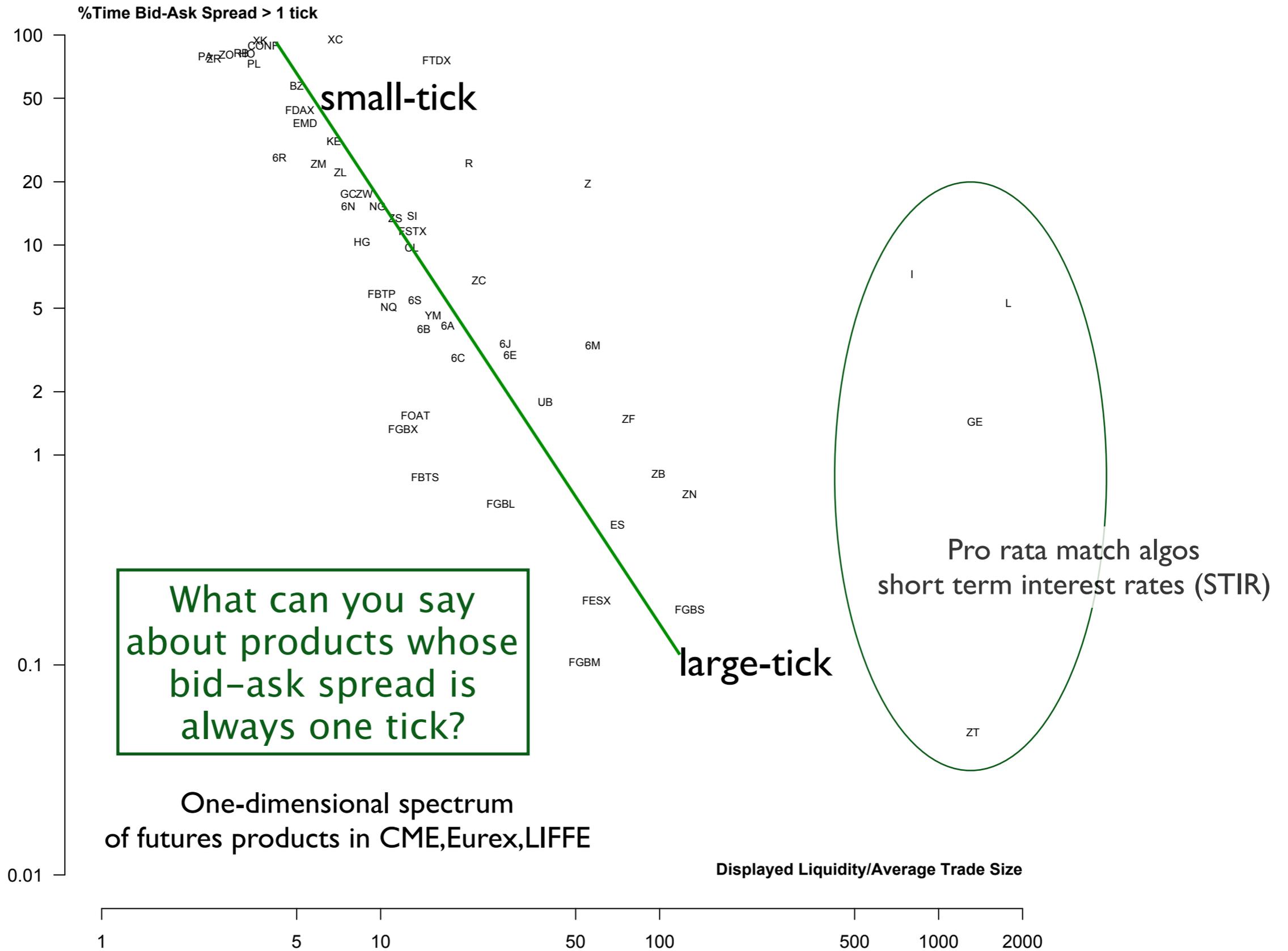
Nondimensionalization brings to same scale

Tick size is nondimensional difference

Nondimensional properties

number of trades before quotes change
shape of order book across levels
number of different limit orders
fraction of time bid-ask spread is 1 tick
number of price changes per long-term change
average quote size / average trade size
Reversion
etc





Reversion

Rapid back-and-forth price moves

Reversion of trade prices

bid-ask bounce

large-tick effects

Reversion of quote midpoint

Roll model (1984)

reversion of trade prices gives effective spread

THE JOURNAL OF FINANCE • VOL. XXXIX, NO. 4 • SEPTEMBER 1984

A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market

RICHARD ROLL*

ABSTRACT

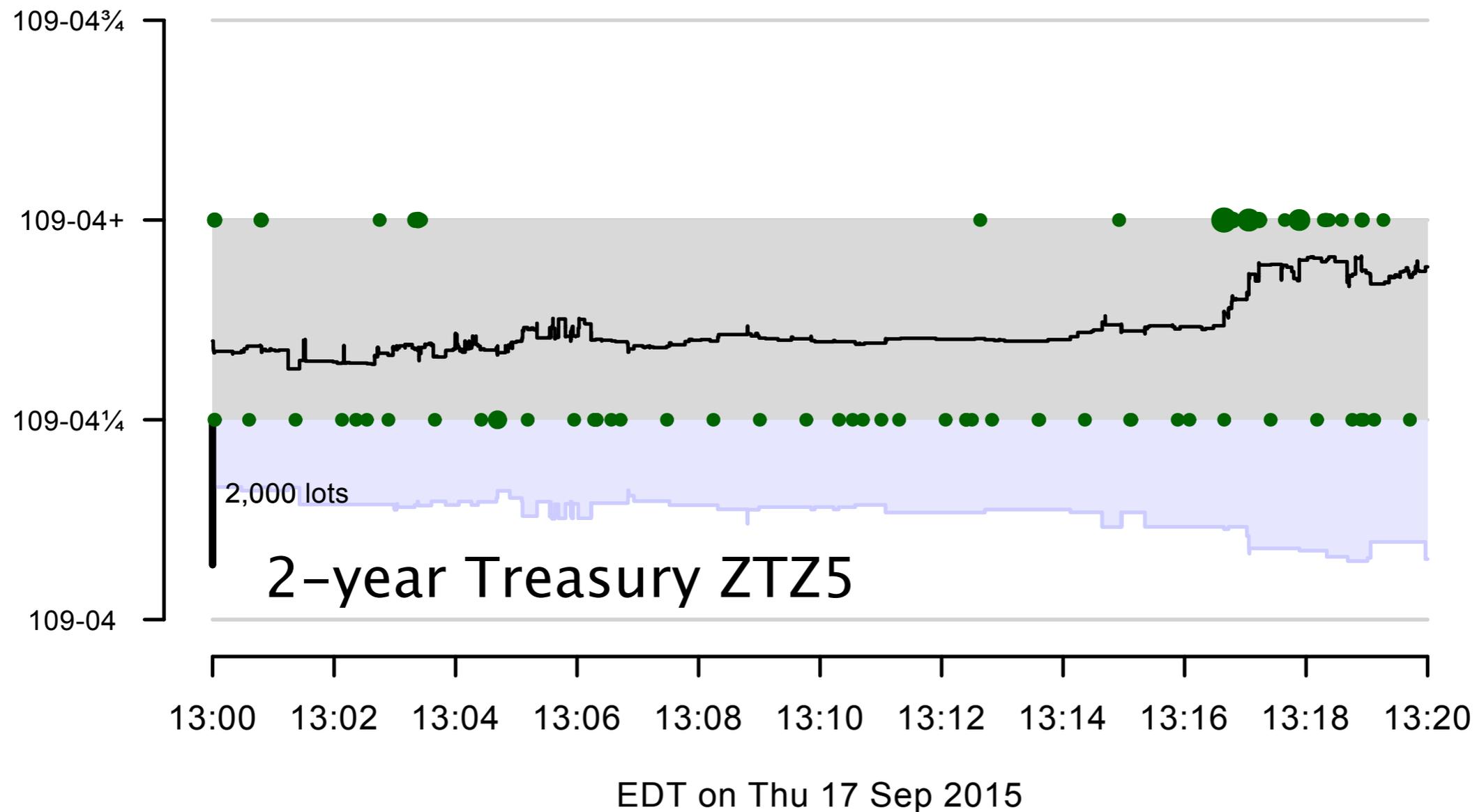
In an efficient market, the fundamental value of a security fluctuates randomly. However, trading costs induce negative serial dependence in successive observed market price changes. In fact, given market efficiency, the effective bid-ask spread can be measured by

$$\text{Spread} = 2\sqrt{-\text{cov}}$$

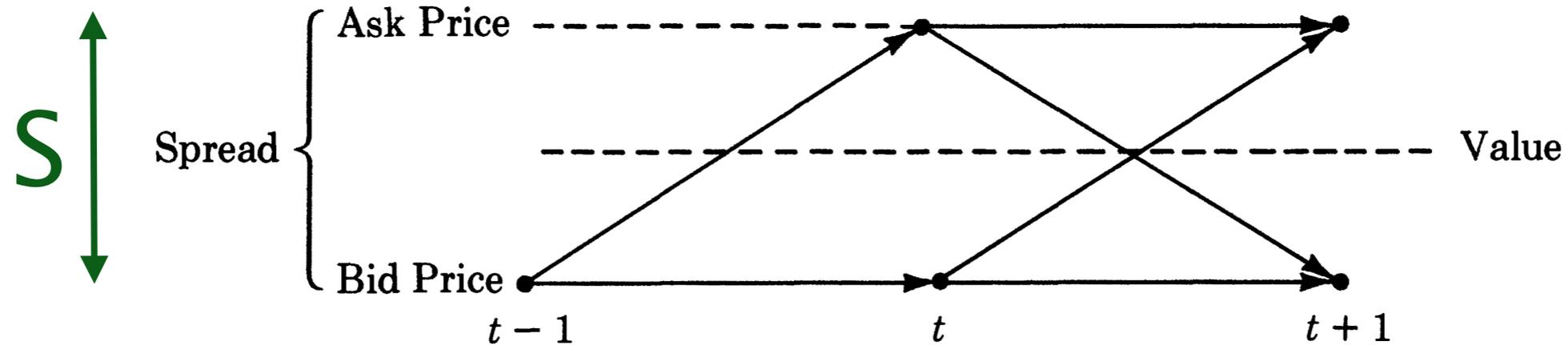
where “cov” is the first-order serial covariance of price changes. This implicit measure of the bid-ask spread is derived formally and is shown empirically to be closely related to firm size.

Roll model

Suppose temporarily that bid/ask do not move



The schematic below illustrates possible paths of observed market price between successive time periods, given that the price at time $t - 1$ was a sale to the market maker, at his bid, and given that no new information arrives in the market.



Each path is equally likely. There is a similar but opposite asymmetric pattern if the price at $t - 1$ happened to be a purchase from the market maker, at his ask price.

4 price sequences, equally likely:
 $(0,0)$, $(0,+S)$, $(+S,0)$, $(+S,-S)$

Covariance $C = \langle \Delta p_i \cdot \Delta p_{i+1} \rangle = -\frac{1}{4} S^2$

Roll formula

$$\text{Effective spread } S = 2\sqrt{-C}$$

Not affected by midpoint motion
if uncorrelated to bid-ask bounce
Independent of observation interval

Roll used daily data

Multiple fills from one market order?

group by time? by second?

Useful when cannot observe actual quotes
(Hasbrouck 2003 futures pit data)

But most of time can observe quote

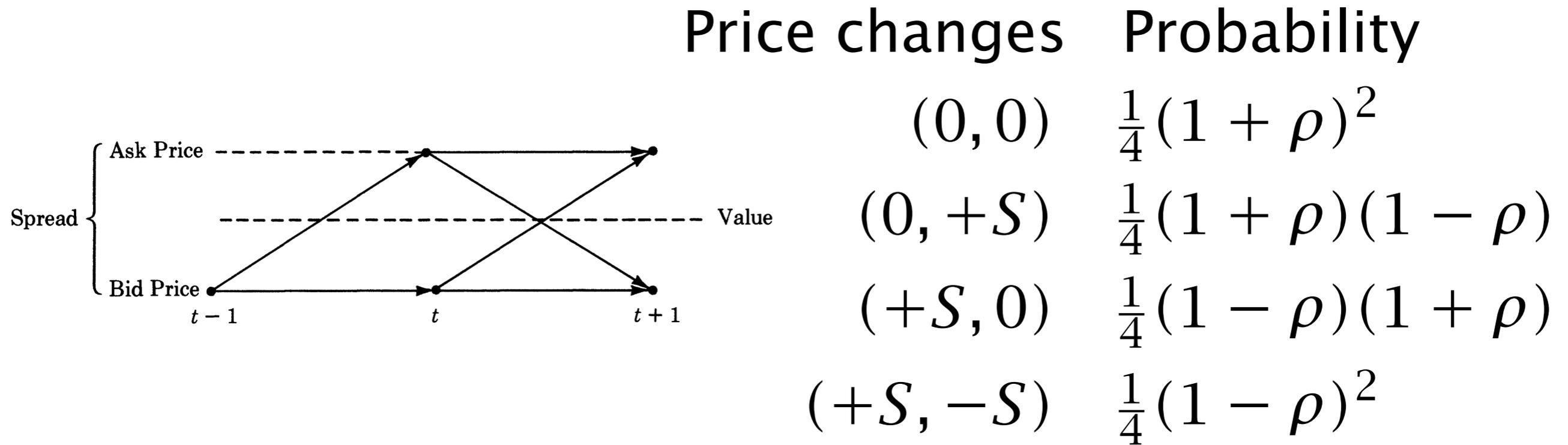
Roll model with serial correlation

$$\mathbb{P}(P_{i+1} = \text{ask} \mid P_i = \text{ask}) = \frac{1}{2} (1 + \rho)$$

$$\mathbb{P}(P_{i+1} = \text{bid} \mid P_i = \text{ask}) = \frac{1}{2} (1 - \rho)$$

ρ is serial correlation of prices
not of price changes

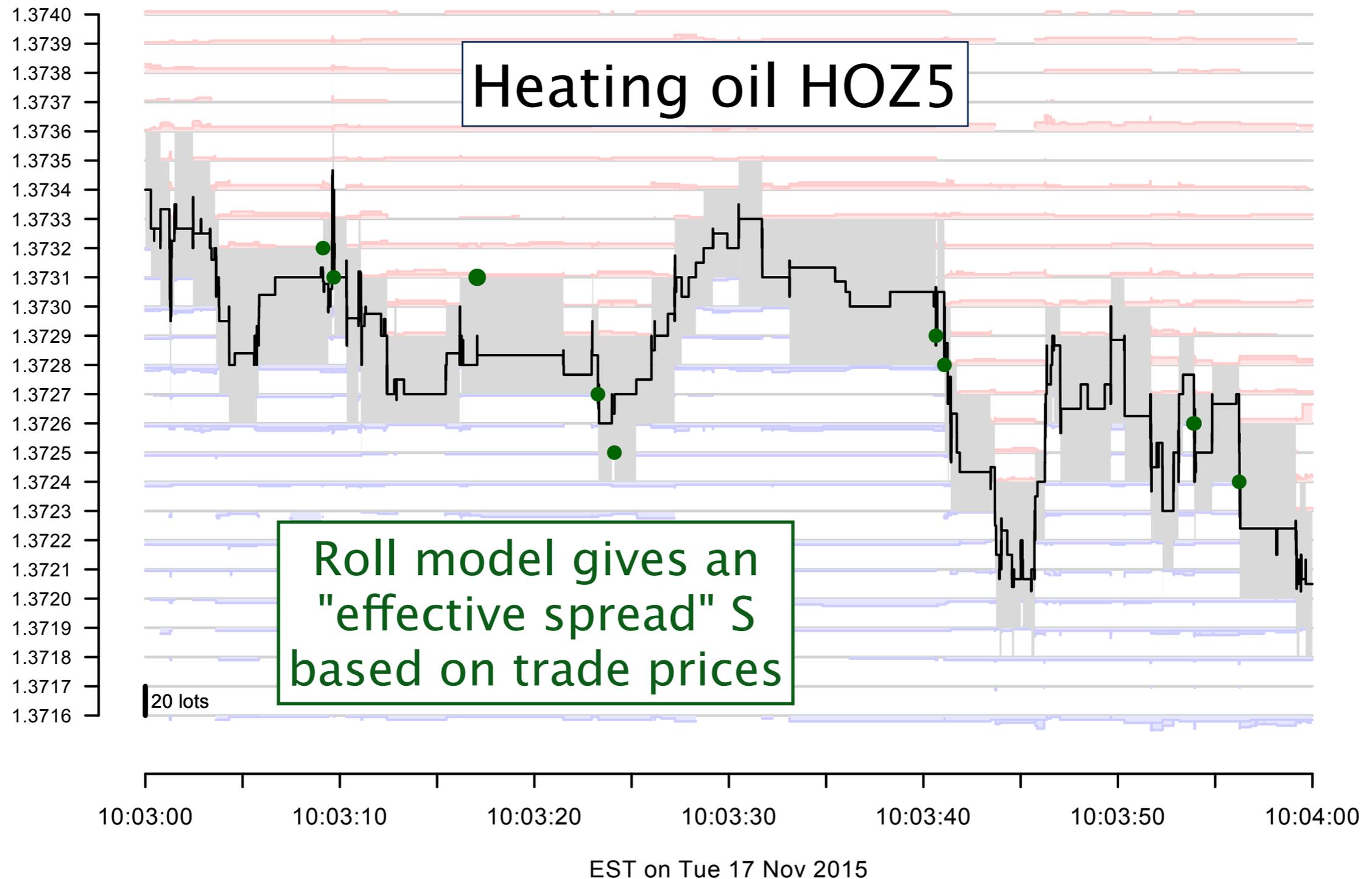
Roll model with serial correlation



$$C = \langle \Delta p_i \cdot \Delta p_{i+1} \rangle = -\frac{1}{4}(1 - \rho)^2 S^2$$

$$S = \frac{2\sqrt{-C}}{1 - \rho}$$

Average spread when not 1 tick



Reversion complicates volatility

Price (or log-price) $X(t)$

$$dX = \sigma(t) dW(t)$$

$$Q(t_L, t_R) = \int_{t_L}^{t_R} \sigma(t)^2 dt$$
$$\approx \sum_{j=1}^N (X(t_j) - X(t_{j-1}))^2$$

if increments are uncorrelated

$$t_0 = t_L, \quad t_N = t_R, \quad t_0 < t_1 < \dots < t_N$$

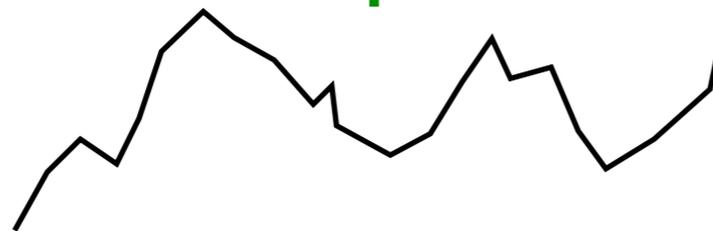
Errors in volatility measure

Discrete calculation

Overestimates if negative serial correlation



Underestimates if positive serial correlation

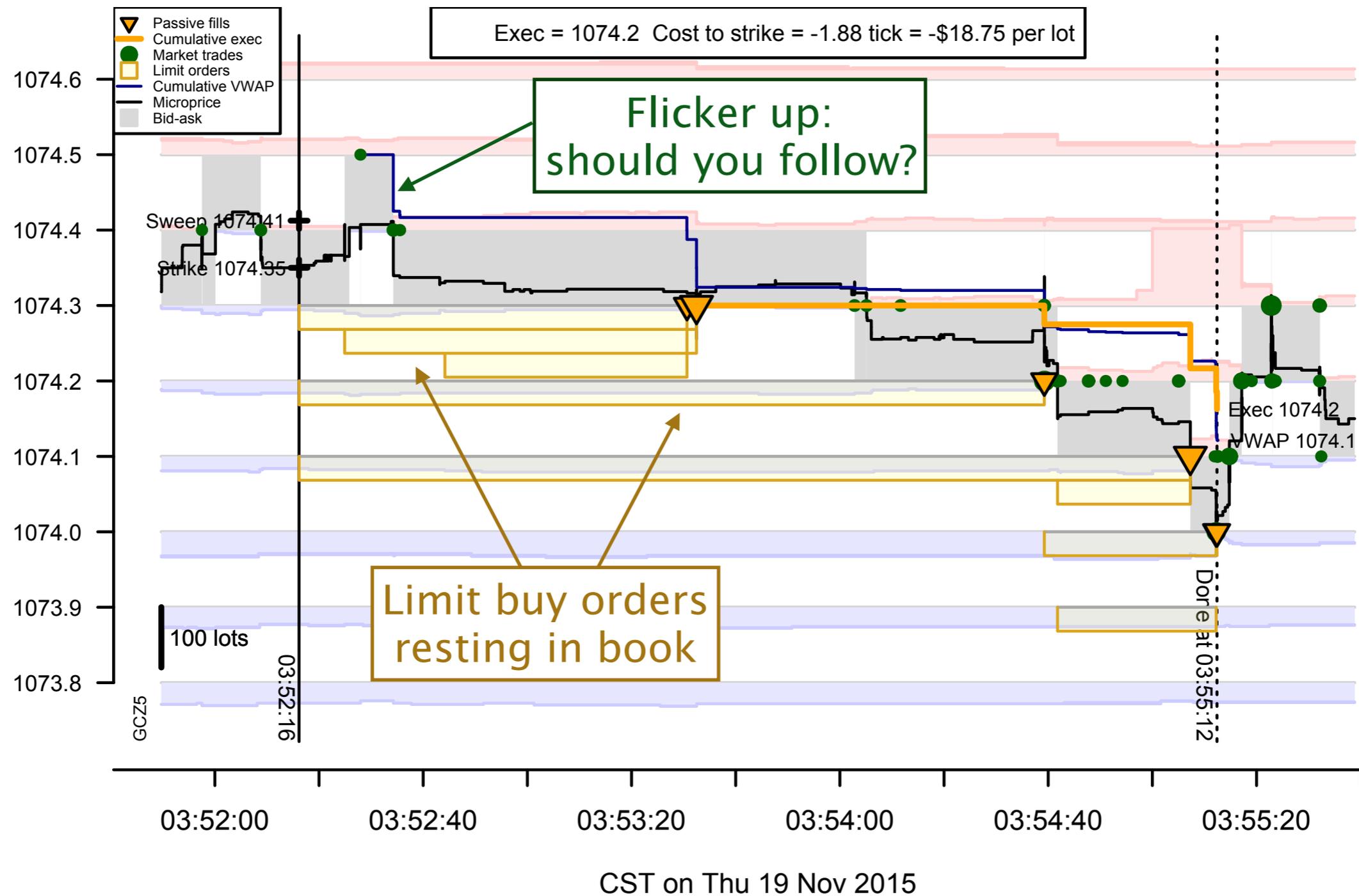


Can use quote midpoints instead of trades

quote midpoints also have reversion

Reversion complicates trading

BUY 8 GCZ5 BOLT



Tick size and reversion

Reversion is easy to measure nondimensionally
How can we measure discreteness (tick size)
When is tick too large or too small?

Uncertainty zone mode

A New Approach for the Dynamics of Ultra-High-Frequency Data: The Model with Uncertainty Zones

CHRISTIAN Y. ROBERT

CREST and ENSAE Paris Tech

MATHIEU ROSENBAUM

CMAF-École Polytechnique Paris UMR CNRS 7641

ABSTRACT

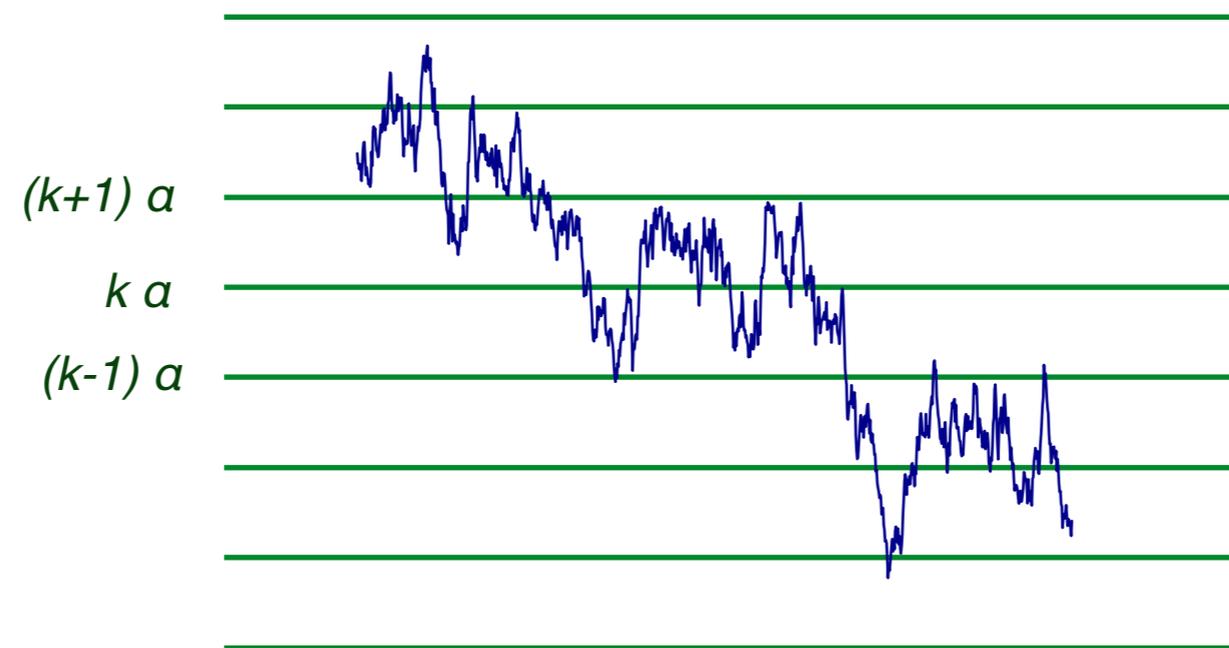
In this paper, we provide a model which accommodates the assumption of a continuous efficient price with the inherent properties of ultra-high-frequency transaction data (price discreteness, irregular temporal spacing, diurnal patterns...). Our approach consists in designing a stochastic mechanism for deriving the transaction prices from the latent efficient price. The main idea behind the model is that, if a transaction occurs at some value on the tick grid and leads to a price change, then the efficient price has been close enough to this value shortly before the transaction. We call uncertainty zones the bands around the mid-tick grid where the efficient price is too far from the tick grid to trigger a price change. In our setting, the width of these uncertainty zones quantifies the aversion to price changes of the market participants. Furthermore, this model enables us to derive approximated values of the efficient price at some random times, which is particularly useful for building statistical procedures. Convincing results are obtained through a simulation study and the use of the model over 10 representative stocks.

"True" underlying price (or log-price) $X(t)$:

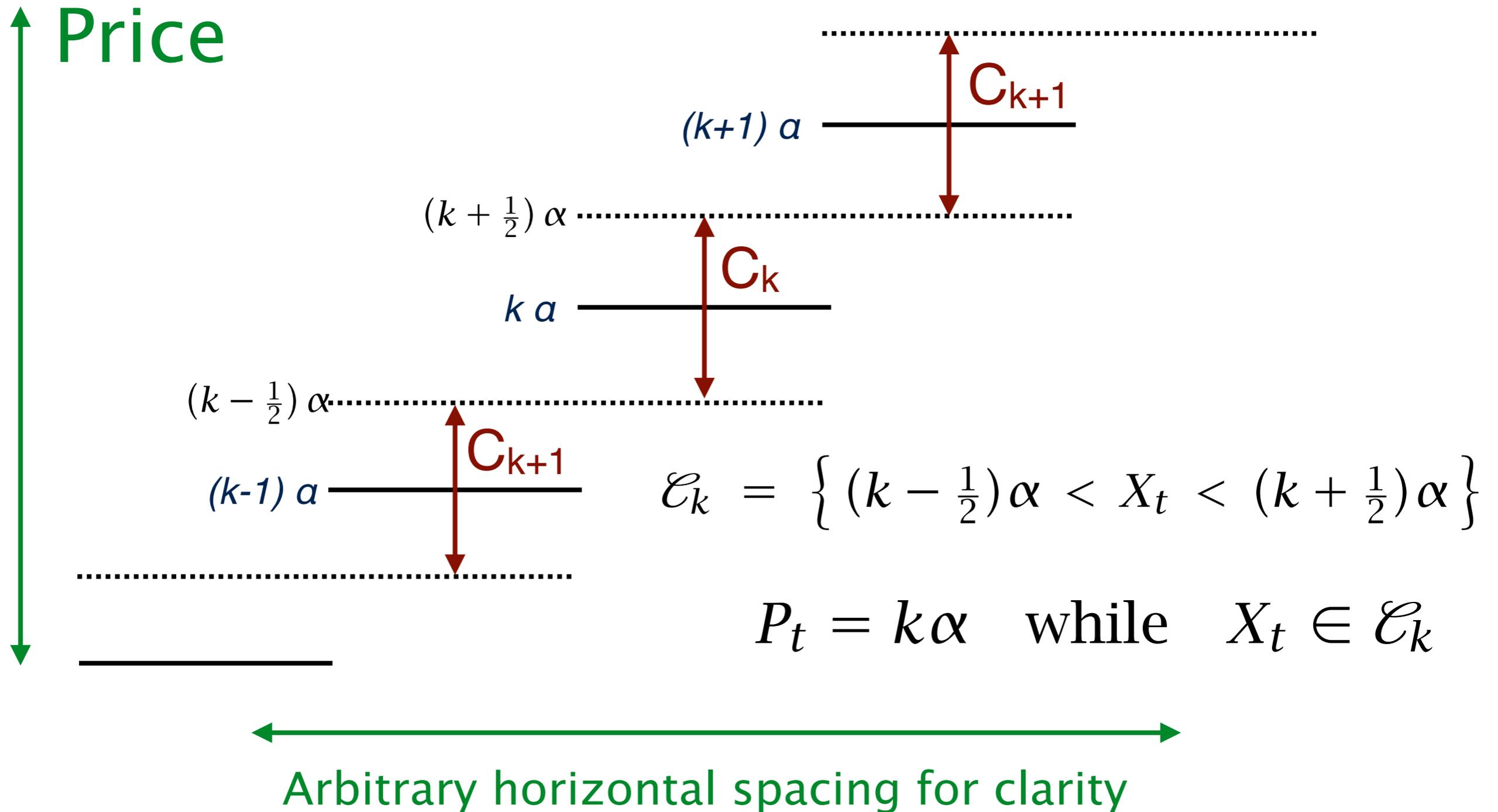
$$dX = \sigma(t) dW(t)$$

How does the continuous process $X(t)$ relate to the discrete grid of spacing a ?

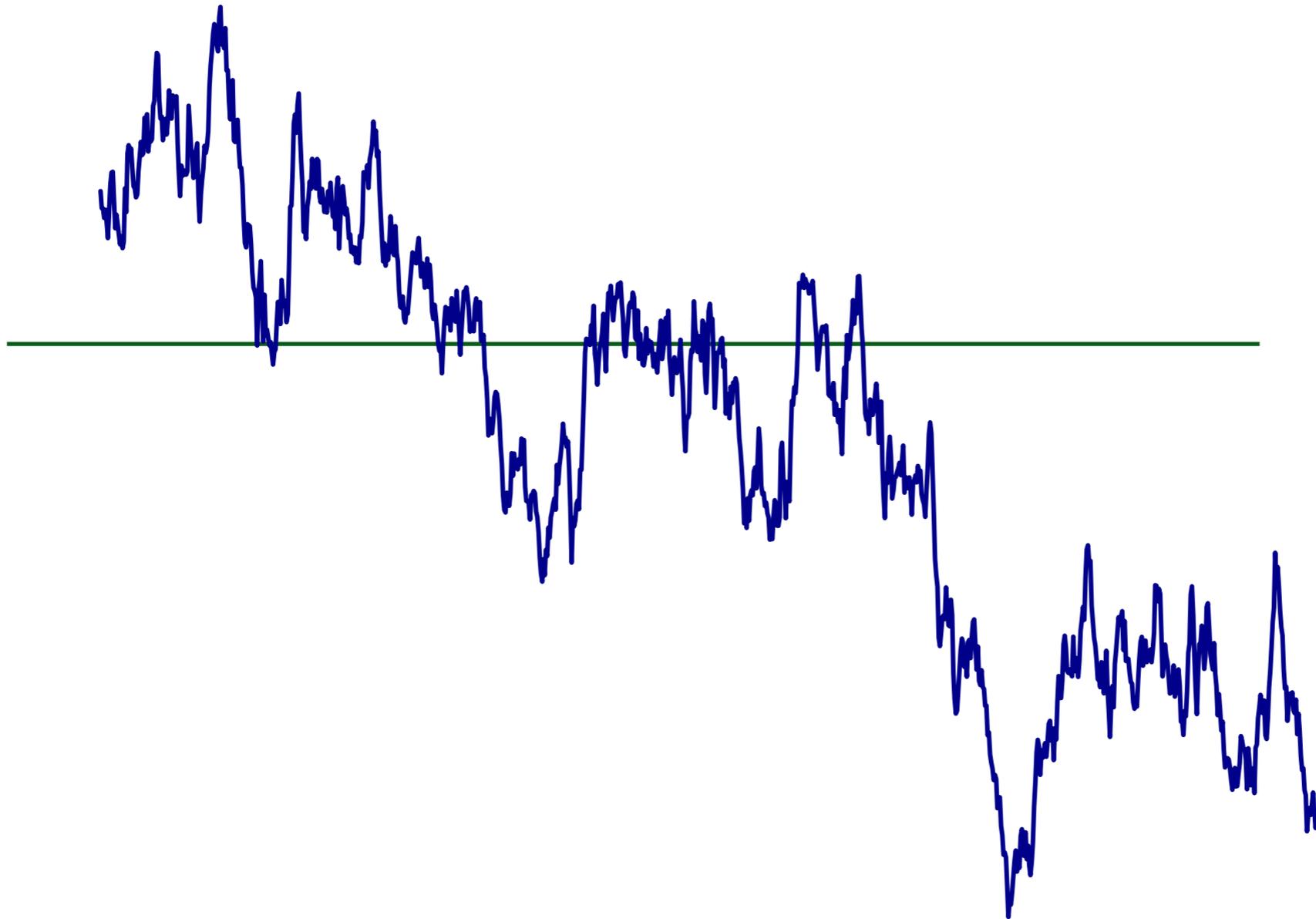
Consider trade prices $P(t)$ for now.



What does not work



Does not work because
infinite fluctuations across boundaries

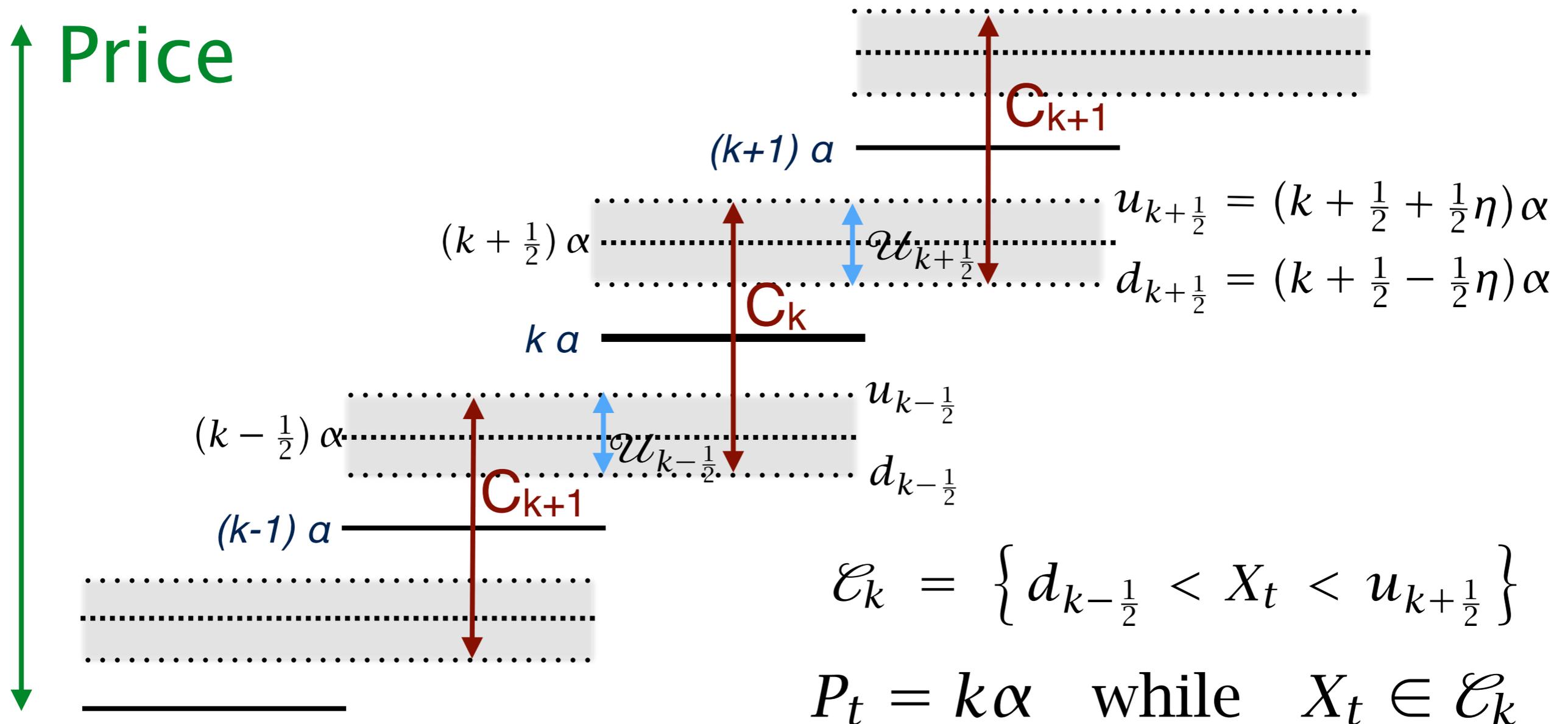


Uncertainty zones

$$0 < \eta < 2$$

This picture $0 < \eta < 1$

Price



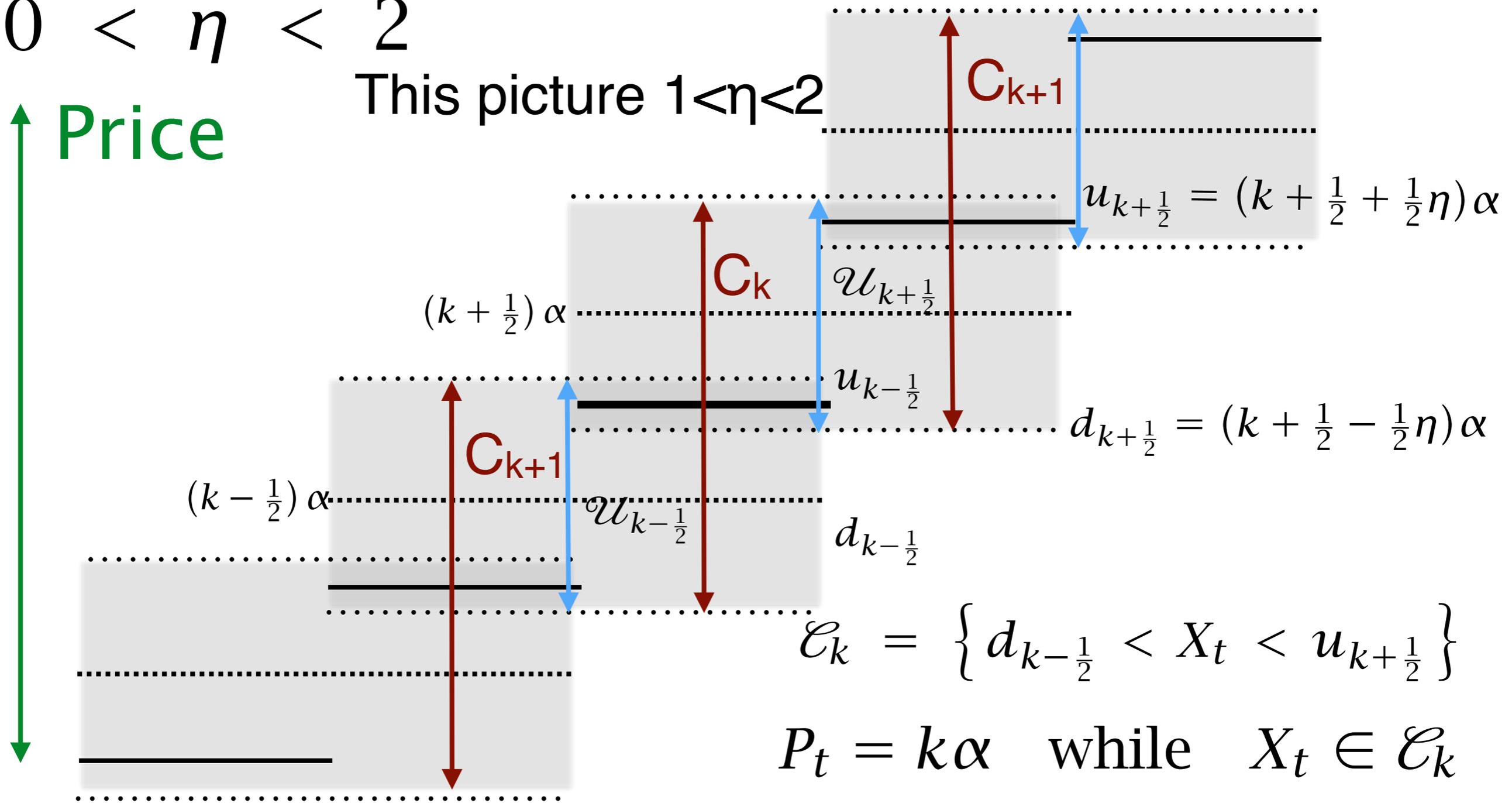
Arbitrary horizontal spacing for clarity

Uncertainty zones

$$0 < \eta < 2$$

This picture $1 < \eta < 2$

Price



Arbitrary horizontal spacing for clarity

Trades P_t occur at price k a while $X_t \in C_k$

At least one trade occurs during each interval

Trade price changes happen when P_t exits C_k

Parameter η measures "stickiness" of price
how far "real" price has to move beyond mid
before trade prices adjust

$\eta < 1$: trade at new price before X_t gets there
tick size is "too large"

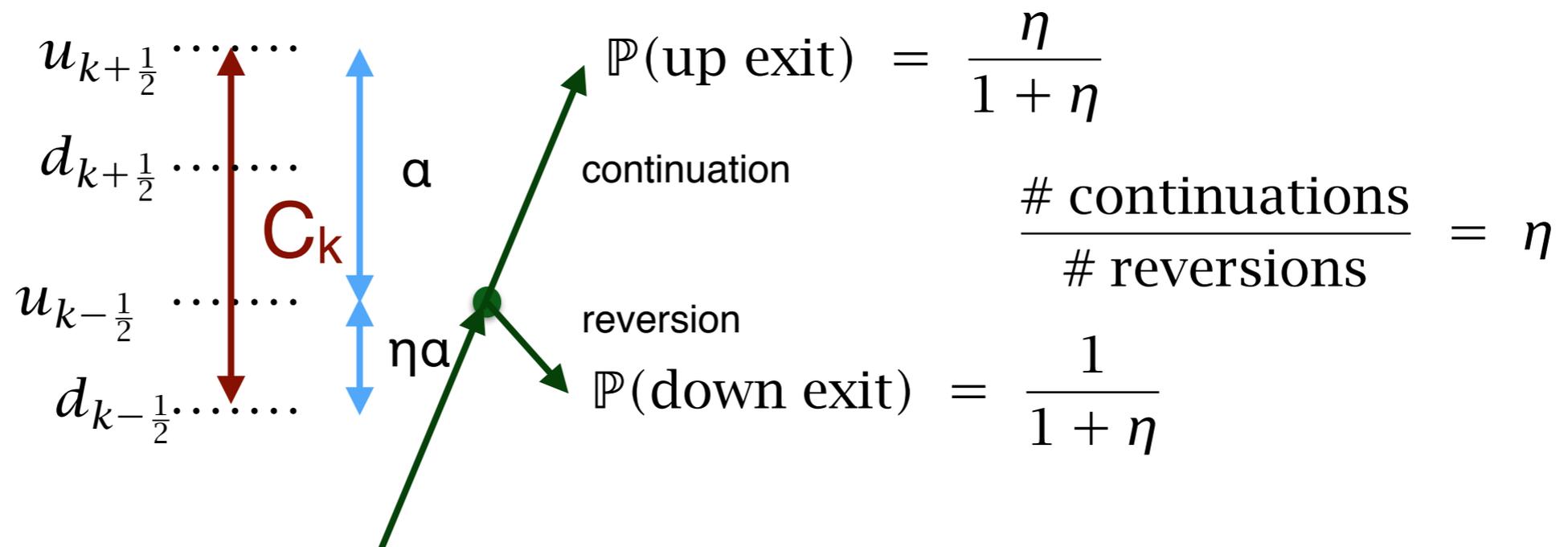
$\eta > 1$: trade at new price after X_t gets there
tick size is "too small"

"Ideal" market would have $\eta = 1$

How to estimate η ?

Observe trades at P_k , previous trades were at P_{k-1}

Which direction will it exit next?



Serial correlation of price changes $\rho = -\frac{1 - \eta}{1 + \eta}$

Estimating volatility

Observe prices $P_0, P_1, P_2, P_3, \dots, P_N$
at times $t_0, t_1, t_2, t_3, \dots, t_N$

Long-term volatility: $\mathbb{E}\left((P_N - P_0)^2\right) = \sigma_L \sqrt{T}$

Short-term volatility: $\mathbb{E}\left((P_j - P_{j-1})^2\right) = \sigma_S \sqrt{t_j - t_{j-1}}$

Relation short- and long-term

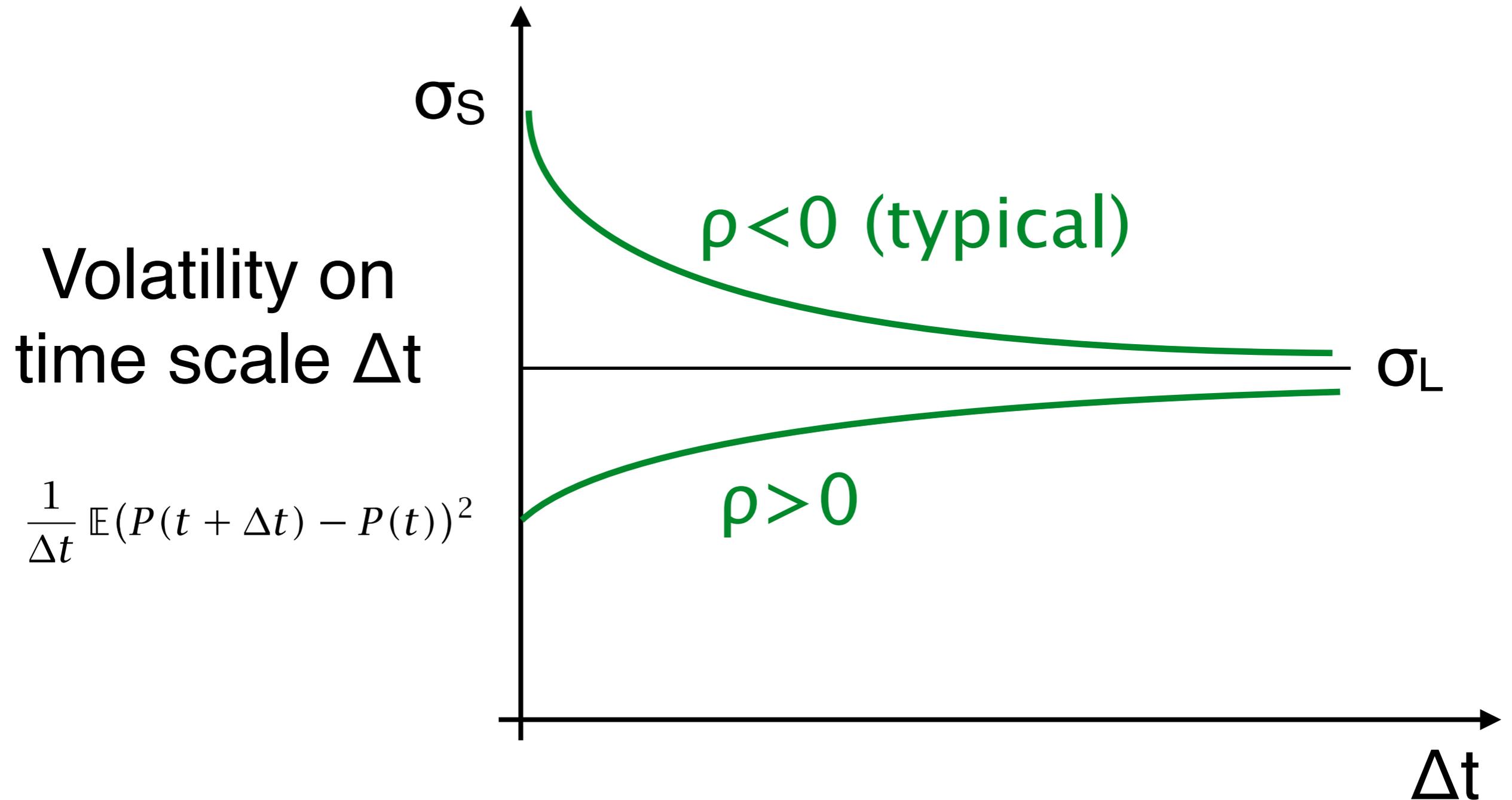
$$P_j - P_{j-1} = \sigma_S \sqrt{t_j - t_{j-1}} \xi_j, \quad \mathbb{E}(\xi_j) = 0, \quad \mathbb{E}(\xi_j^2) = 1$$

$$\begin{aligned} (P_N - P_0)^2 &= \left(\sum_{j=1}^N P_j - P_{j-1} \right)^2 \\ &= \sum_{j=1}^N (P_j - P_{j-1})^2 + 2 \sum_{j=1}^{N-1} \sum_{i=j+1}^N (P_j - P_{j-1})(P_i - P_{i-1}) \end{aligned}$$

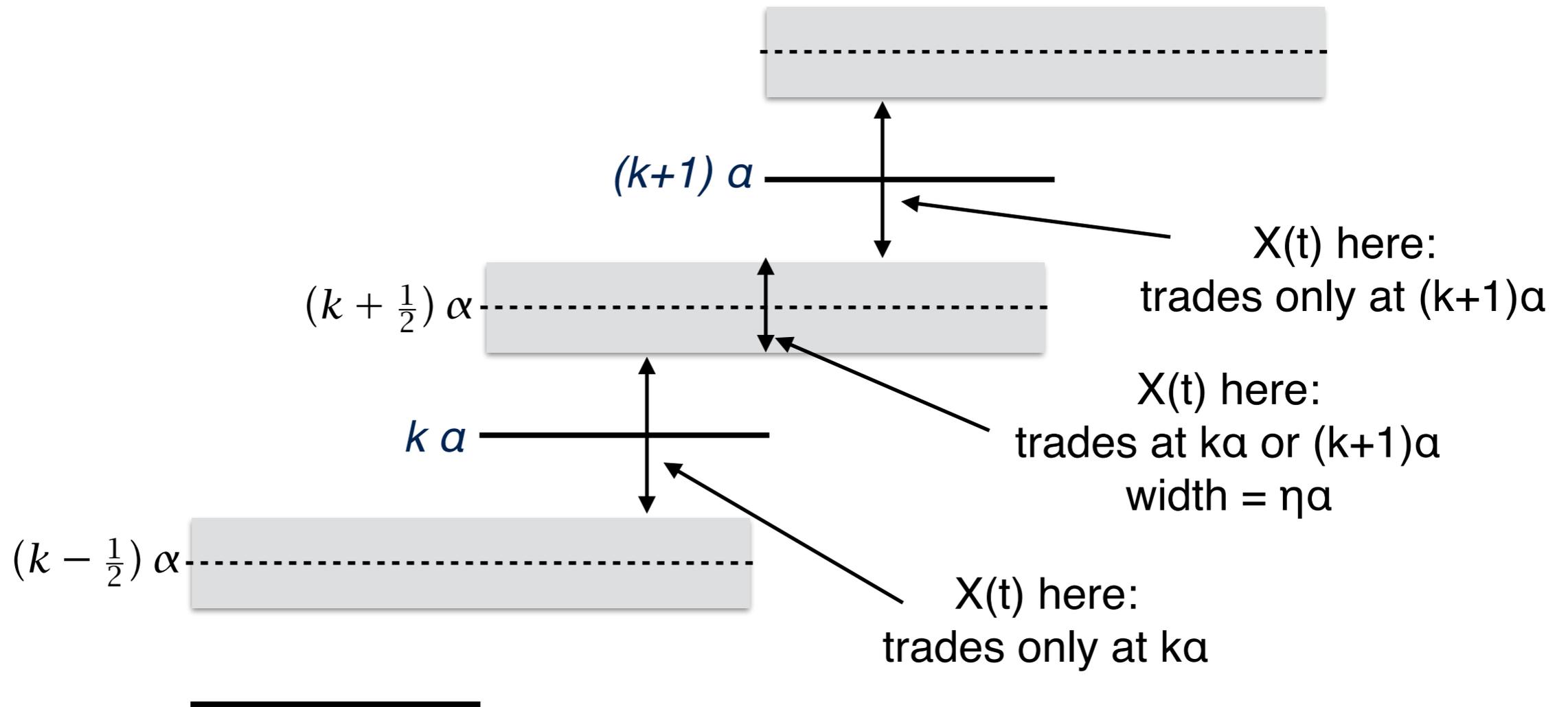
$$\sigma_L^2 = \sigma_S^2 \frac{1 + \rho}{1 - \rho}$$

$$\rho = -\frac{1 - \eta}{1 + \eta} \quad \Longrightarrow \quad \sigma_L^2 = \eta \sigma_S^2$$

"Signature plot"



Effective spread = $\eta \alpha$



Application to algorithms

1. Reversion

- Reversion = likelihood of successive moves in opposite directions

- Trade price reversion:

Bid–ask bounce (Roll model)

Uncertainty zones (Robert/Rosenbaum/Dayri)

- Quote price reversion:

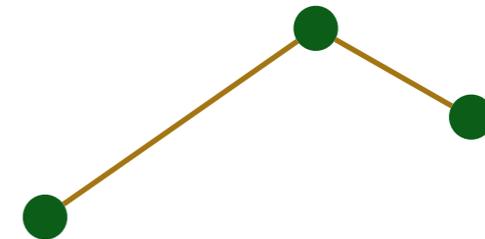
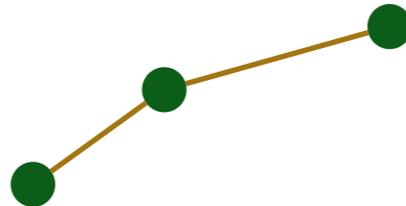
Uncertainty zones/large–tick effects

Multi–tick markets

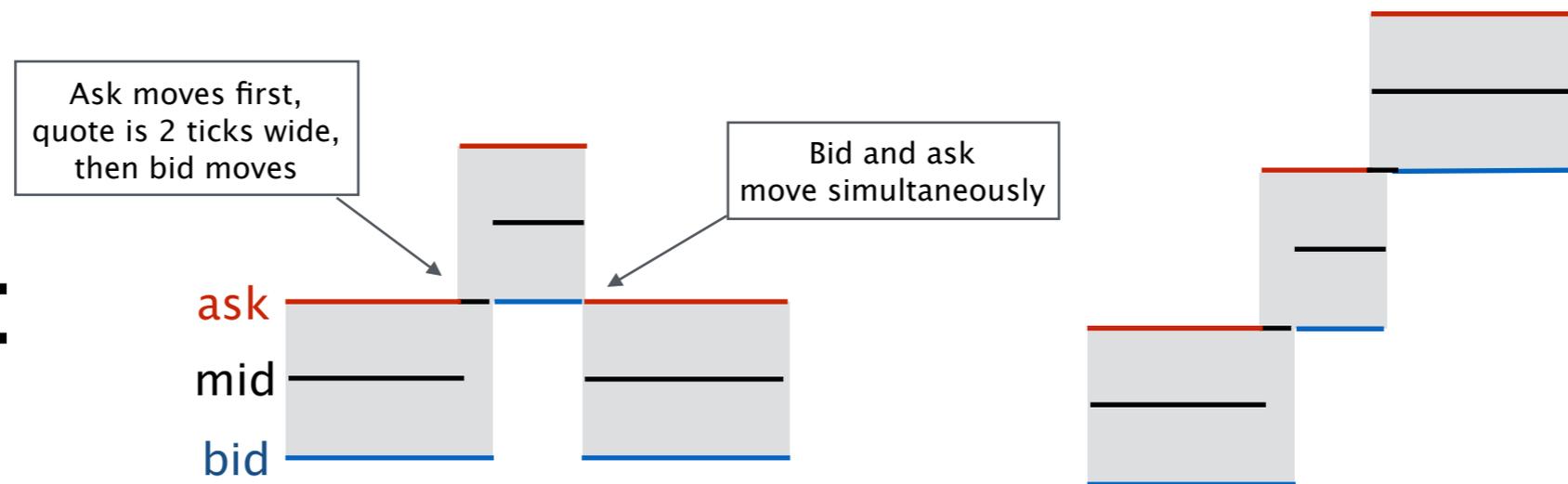
Continuation

Reversal

Trades:



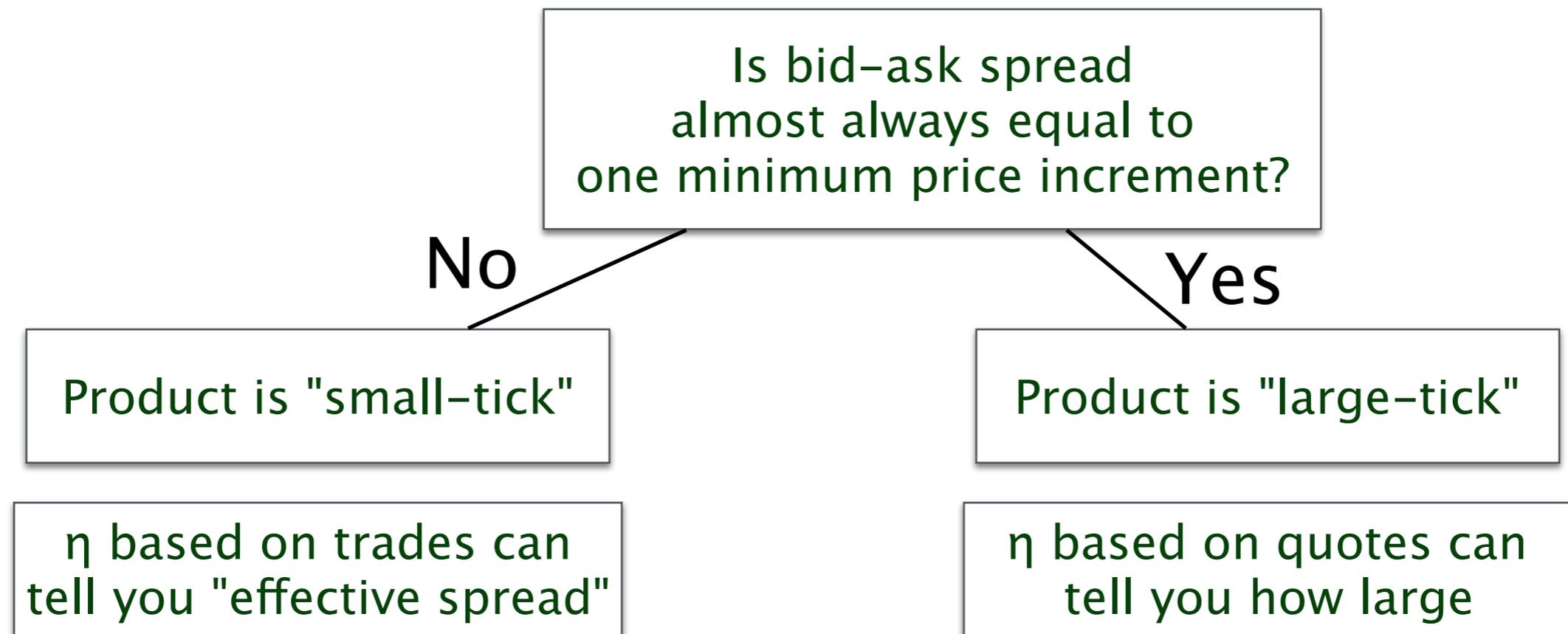
Quotes:



$$\eta = \frac{\# \text{Continuations}}{\# \text{Reversals}}$$

Significance of η : large tick

For a particular product:



Tick size spectrum

Large tick:

bid-ask = minimum tick
quote size \gg trade size
lots of reversion
deep order book

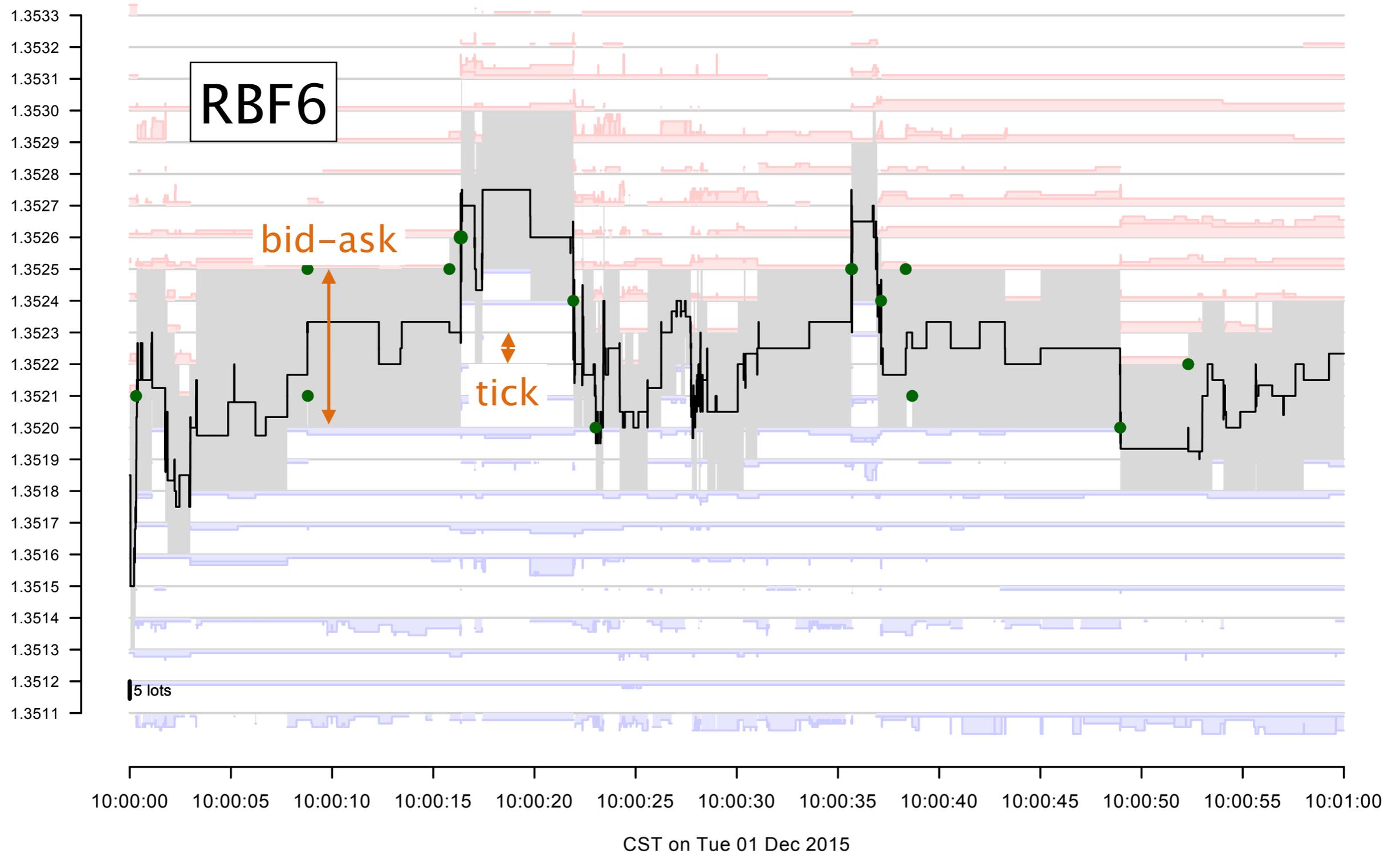


Tick size
(relative to
intrinsic
dynamics)

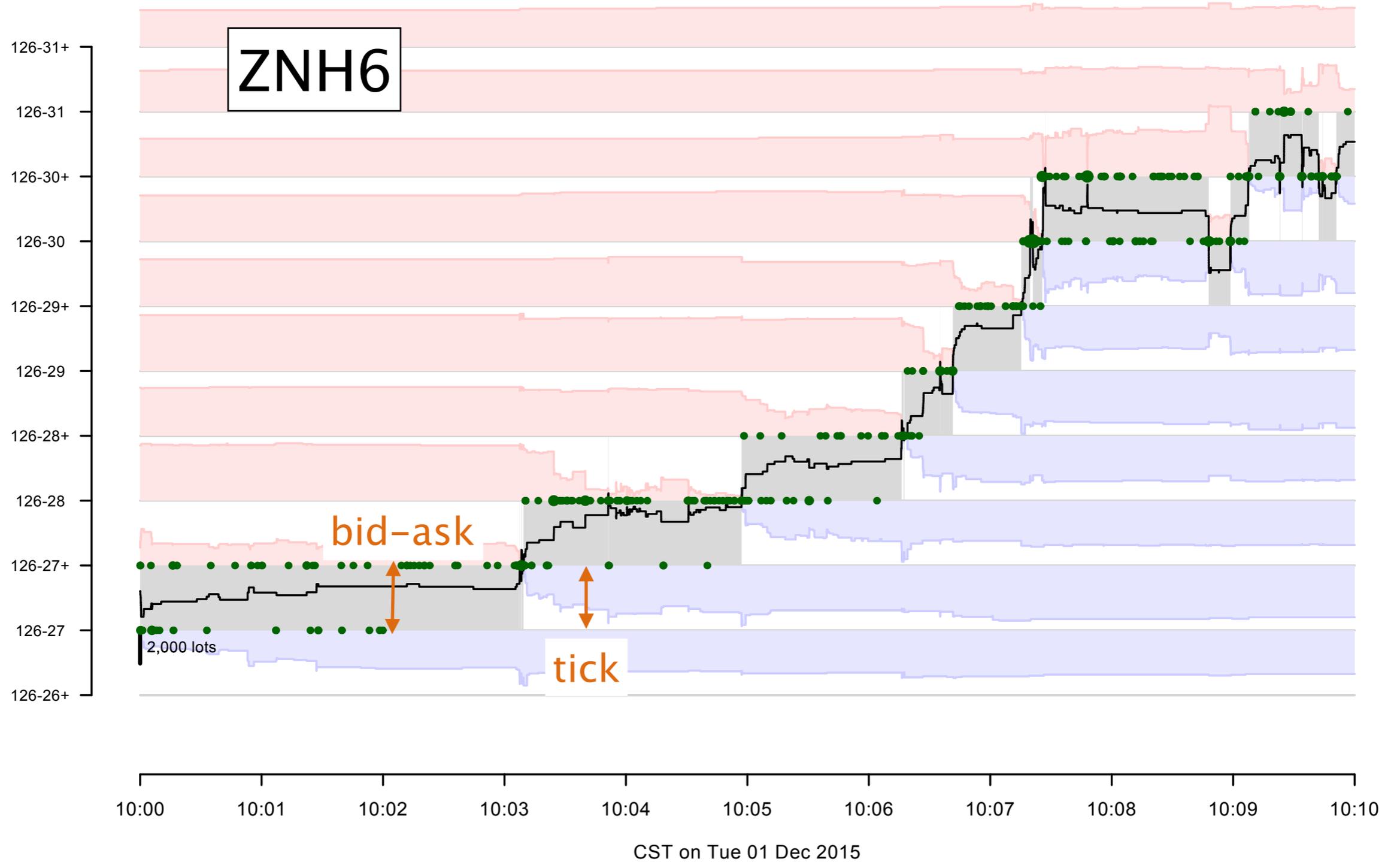
Small tick:

bid-ask $>$ minimum tick
quote size \sim trade size
little reversion
shallow order book

"small-tick" asset: RBOB Gasoline

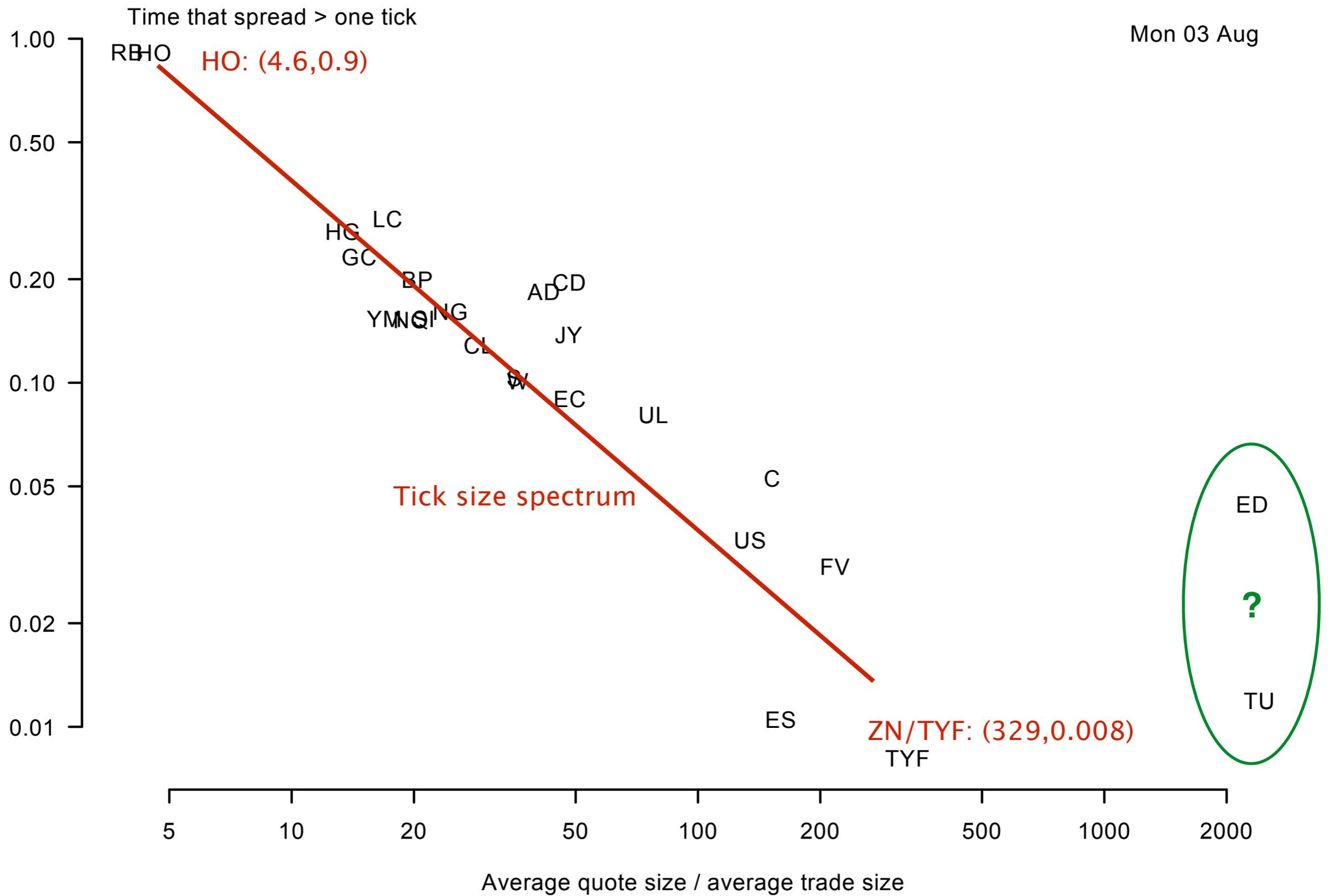


"large-tick" asset: 10-year Treasury

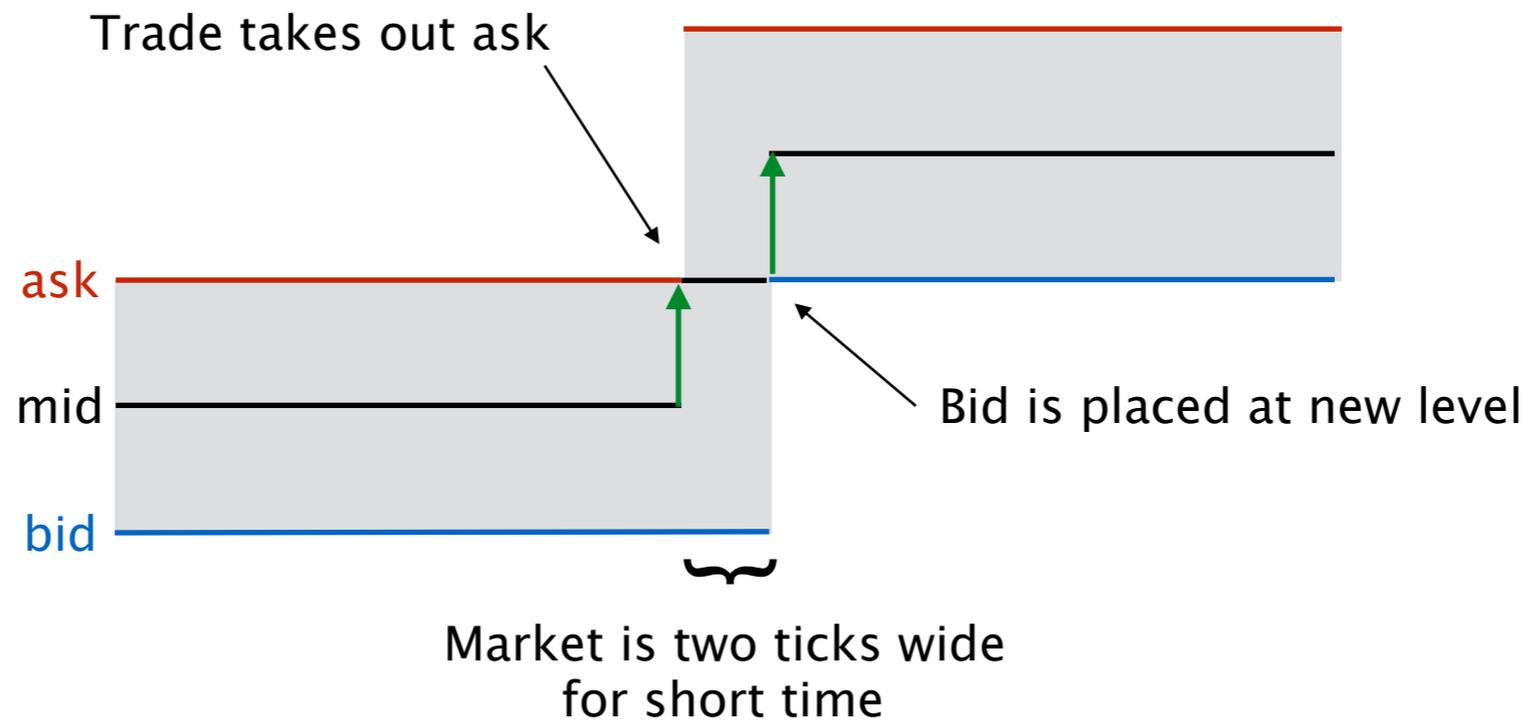


How to characterize?

1. Average quote size / average trade size
2. Fraction of time spread is larger than 1 tick
3. Reversion parameter η based on quote midpoint



Measuring reversion with quote midpoints



**Do not count this as
a continuation**

Spectrum based on reversion parameter (using filtered quote midpoint changes)

