Topics in Market Microstructure Models, Summer 2015

Lecture 1: Market mechanisms and zero intelligence models of the order book

Tai-Ho Wang (王 太 和)

and

Aims of this course

- To be able to read and understand market microstructure literature.
  - In particular, to be familiar with well-known microstructure models.
- To be familiar with empirical characteristics of order books and time series, understanding their implications for modeling and trading.

Outline of Lecture 1

- Overview of the course
- Market structure
  - Order types
  - Market types
  - Trading venues
- A brief introduction to programming in R
- Zero intelligence models

Why study market microstructure?

- To understand better how to trade: Specifically, how to design algorithms.
  - Market microstructure theory could be viewed as the theory underlying algorithmic trading.
- To be able to talk to informed clients.
- To understand the process of price formation.

Philosophical overview

- Two extreme views of the market:
  - Completely random
    - The zero intelligence model associated with econophysics.
    - Use of statistical physics techniques.
  - Completely rational
    - Strategic trader models associated with efficient markets theory.
    - Game theoretic solutions.
- Reality is somewhere in-between.

Order types

Two basic order types:

- Market orders
  - Always executed if there is sufficient quantity available.
- Limit orders
  - May be executed only if limit price reached (note "may" not "will"!).
  - Typically priorities are price first, then time.

There are many variations such as:

- Fill or kill: Must be entirely filled immediately or not at all
- Immediate or cancel: Fill immediately; remaining quantity is canceled
- All or nothing: Must be entirely filled or not at all when the price is reached
- Marketable limit buy order: Buy at limit price but no higher. Remaining quantity becomes best bid.

Hidden orders

- Users may post hidden limit orders
- Hidden orders have lower priority than visible orders at the same price level.
- Hidden orders are often placed between the visible bid and ask prices so market order submitters may sometimes get a price that is better than the visible best quote.

Market types

- *Quote-driven* markets
  - have market makers responsible for continuously quoting two-way prices
  - For example ISE.
- Liquidity in order-driven markets is generated by customer order flow.
  - Limit orders in the book replace market-maker quotes.
- Hybrid markets such as the NYSE are mixtures of both.
  - For example, specialists are still responsible for ensuring two-way quotes in times of stress.

Auction types

Two basic types:

- Batch auctions or *call markets*
  - For example, opening and closing auctions on NYSE.
  - Trading takes place only when the market is called.
    - Prices are "called out" until a price is found that maximizes the quantity traded.

- Continuous auction or *continuous markets*
  - For example trading during the day on NYSE.
  - Trades can take place whenever the market is open.
  - In a continuous market, the limit order book effectively defines supply and demand curves for the asset being traded.

NYSE opening auction example

Table 1: Example: 9:30am ET: Previous closing price of $17.00

Entry Time	Buy Limit Orders	Sell Limit Orders	Order Price	Matchable	Imbalance	Indicative Price
8:30	1,000		$19.00	0	0	-
9:00		1,000	Market	1,000	0	$19.00
9:05	1,000		$20.00	1,000	0	$20.00
9:25		1,000	$18.00	2,000	0	$18.00

- The match price is the price that maximizes the volume that can be executed.
- There are complicated rules for tie-breaking that vary between exchanges.

Continuous trading

- In the continuous trading session, orders are submitted and canceled.
- The set of orders with signs, sizes and prices is known as the *limit order book*.
- We proceed to give a few snapshots of the BATS order book.
    - BATS is the third largest US exchange after NYSE and Nasdaq. It has the reputation of being the fastest exchange.

Trading venues

In the US, trading takes place on

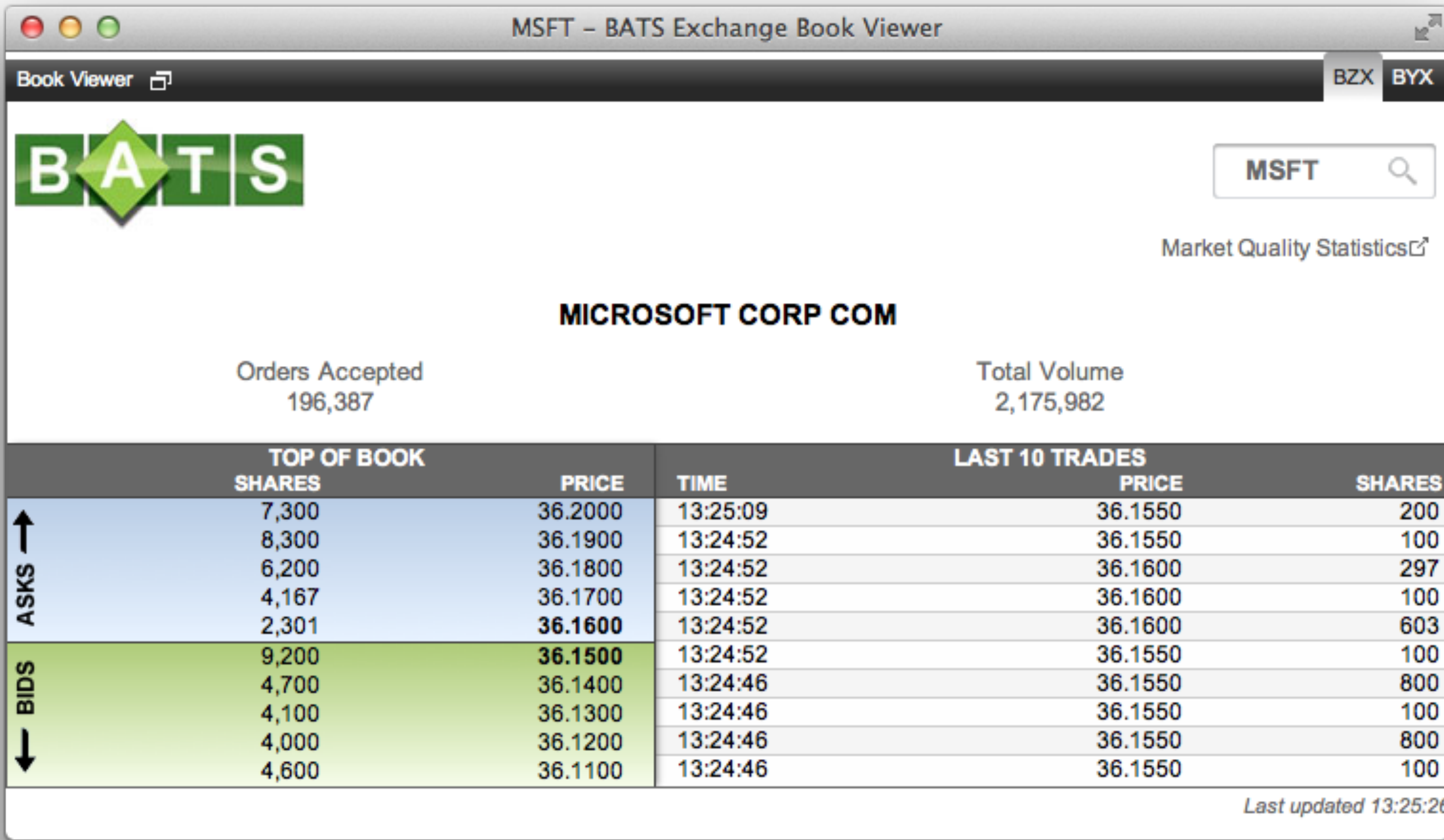
- Organized exchanges such as NYSE and Nasdaq.
    - Trading can be electronic via an ECN (Electronic Communication Network).
    - Prices are formed through an auction process.
- Crossing networks such as POSIT
    - The POSIT algorithm optimizes the maximum possible number of buy and sell orders that match or “cross”.
    - There are currently eleven scheduled intraday crosses every business day.
    - POSIT prices trades at the midpoint of the best bid and offer for each security at the time of the cross, based on information provided directly to the system by a third-party data vendor.
- Dark pools with derivative pricing rules
    - Banks' internal crossing networks such as SIGMA X, Crossfinder, MLXN
    - Other dark pools such as GETMatched (Getco), BIDS, Level etc.

This excerpt taken from the ITG 10-K filed Mar 1, 2007.

POSIT

ITG POSIT crossing destinations, including POSIT Match, POSIT Now, and BLOCKalert, give buyers and sellers opportunities to match equity orders with complete confidentiality, no market impact, and the cost savings of midpoint pricing. POSIT offers unique value for traders with active, quantitative, and passive trading styles. POSIT provides access to rich, diverse liquidity, is useful for all trading styles and is especially valuable for trading small, illiquid names. All POSIT products cross at the midpoint of the National Best Bid or Offer (“NBBO”).

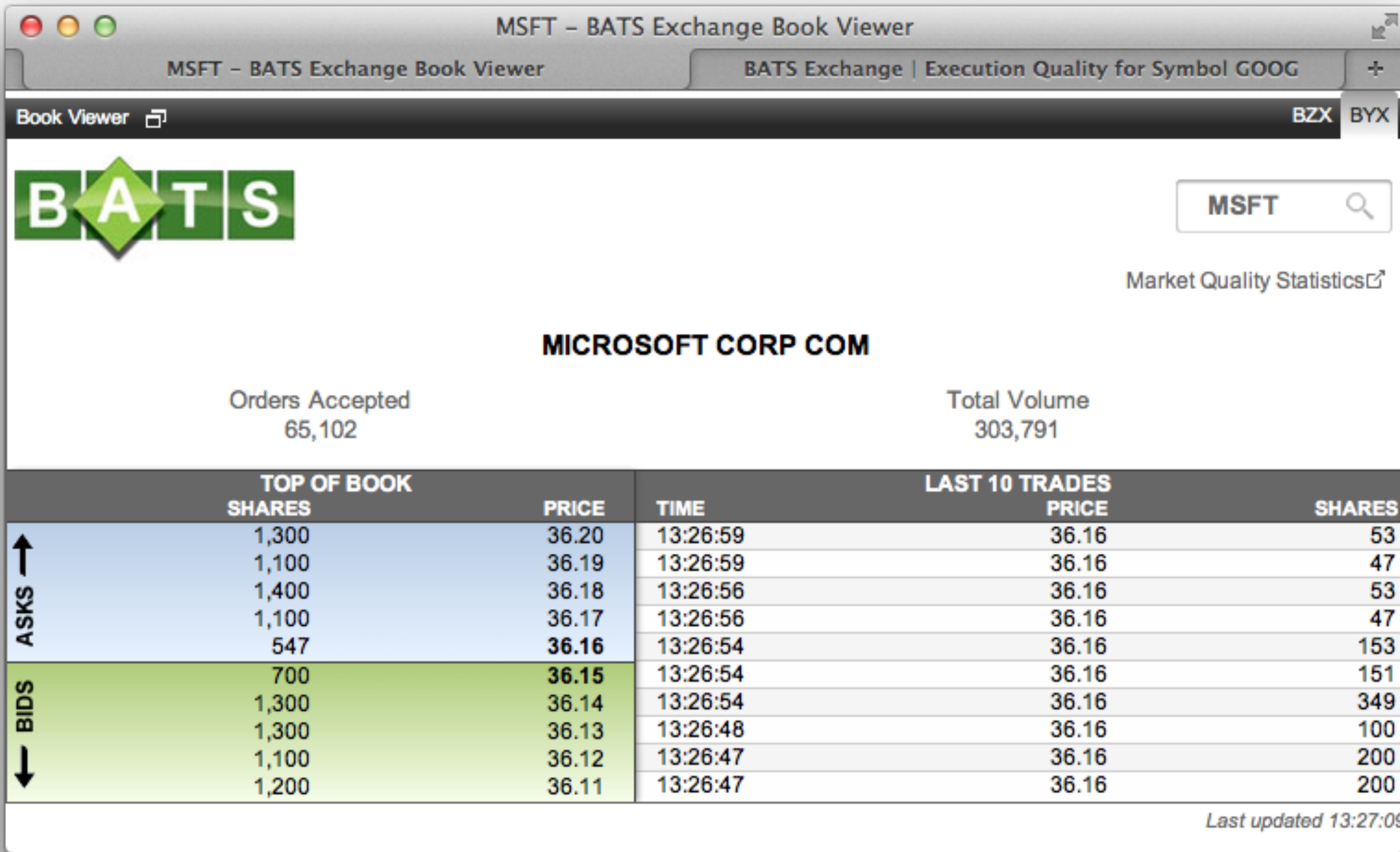
BATS (BZX) order book for MSFT



	TOP OF BOOK SHARES	PRICE
ASKS	7,300	36.2000
	8,300	36.1900
	6,200	36.1800
	4,167	36.1700
	2,301	**36.1600**
BIDS	9,200	**36.1500**
	4,700	36.1400
	4,100	36.1300
	4,000	36.1200
	4,600	36.1100

LAST 10 TRADES TIME	PRICE	SHARES
13:25:09	36.1550	200
13:24:52	36.1550	100
13:24:52	36.1600	297
13:24:52	36.1600	100
13:24:52	36.1600	603
13:24:52	36.1550	100
13:24:46	36.1550	800
13:24:46	36.1550	100
13:24:46	36.1550	800
13:24:46	36.1550	100

BATS (BYX) order book for MSFT



MICROSOFT CORP COM

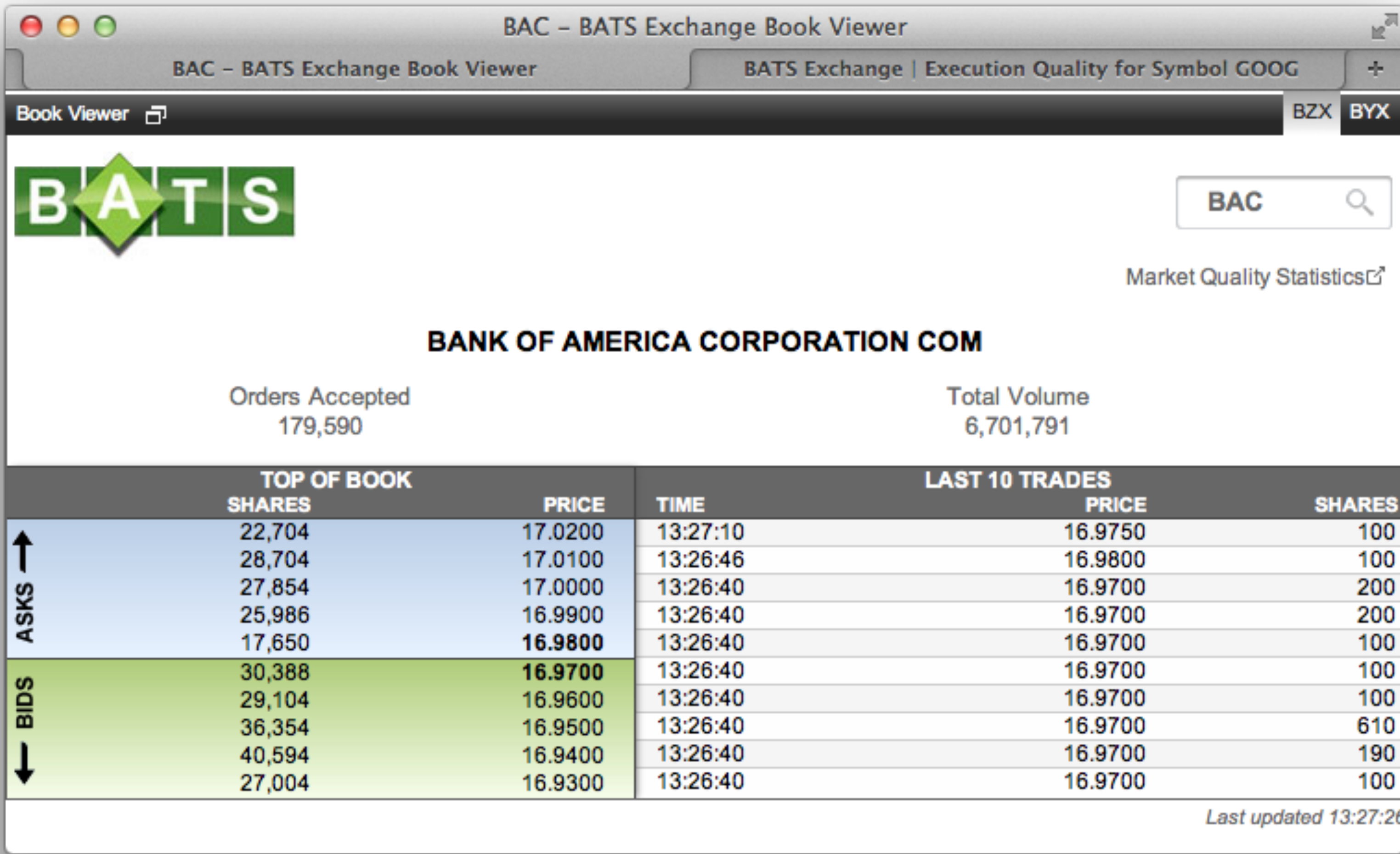
Orders Accepted
65,102

Total Volume
303,791

	TOP OF BOOK		LAST 10 TRADES		
	SHARES	PRICE	TIME	PRICE	SHARES
ASKS	1,300	36.20	13:26:59	36.16	53
	1,100	36.19	13:26:59	36.16	47
	1,400	36.18	13:26:56	36.16	53
	1,100	36.17	13:26:56	36.16	47
	547	**36.16**	13:26:54	36.16	153
BIDS	700	**36.15**	13:26:54	36.16	151
	1,300	36.14	13:26:54	36.16	349
	1,300	36.13	13:26:48	36.16	100
	1,100	36.12	13:26:47	36.16	200
	1,200	36.11	13:26:47	36.16	200

Last updated 13:27:09

BATS (BZX) order book for BAC



BAC – BATS Exchange Book Viewer

BAC – BATS Exchange Book Viewer | BATS Exchange | Execution Quality for Symbol GOOG

Book Viewer | BZX | BYX

BATS

BAC

Market Quality Statistics

BANK OF AMERICA CORPORATION COM

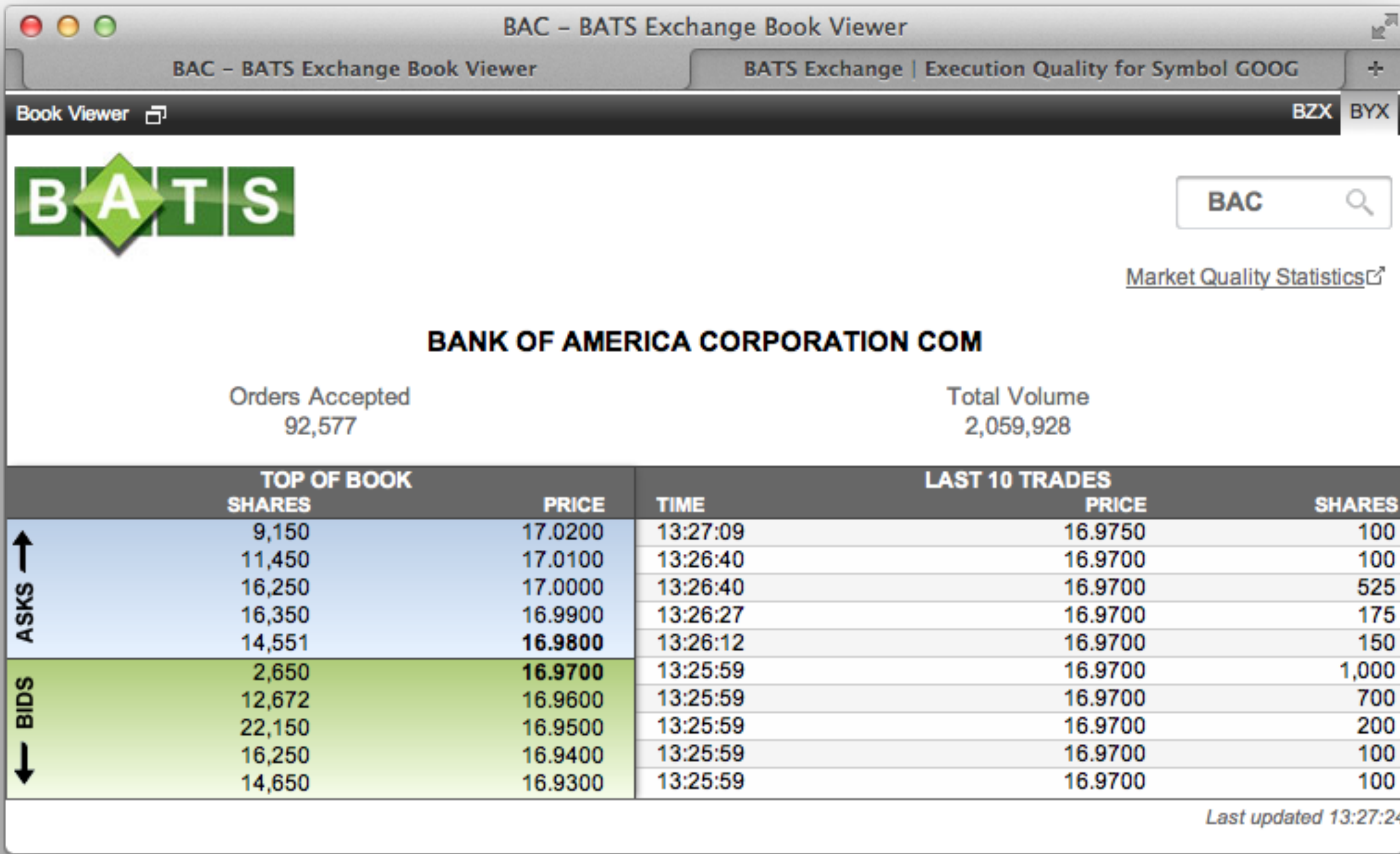
Orders Accepted 179,590

Total Volume 6,701,791

	TOP OF BOOK		LAST 10 TRADES		
	SHARES	PRICE	TIME	PRICE	SHARES
ASKS	22,704	17.0200	13:27:10	16.9750	100
	28,704	17.0100	13:26:46	16.9800	100
	27,854	17.0000	13:26:40	16.9700	200
	25,986	16.9900	13:26:40	16.9700	200
	17,650	**16.9800**	13:26:40	16.9700	100
BIDS	30,388	**16.9700**	13:26:40	16.9700	100
	29,104	16.9600	13:26:40	16.9700	100
	36,354	16.9500	13:26:40	16.9700	610
	40,594	16.9400	13:26:40	16.9700	190
	27,004	16.9300	13:26:40	16.9700	100

Last updated 13:27:26

BATS (BYX) order book for BAC



BAC – BATS Exchange Book Viewer

BAC – BATS Exchange Book Viewer | BATS Exchange | Execution Quality for Symbol GOOG

Book Viewer — BZX | BYX

BATS

BAC

Market Quality Statistics

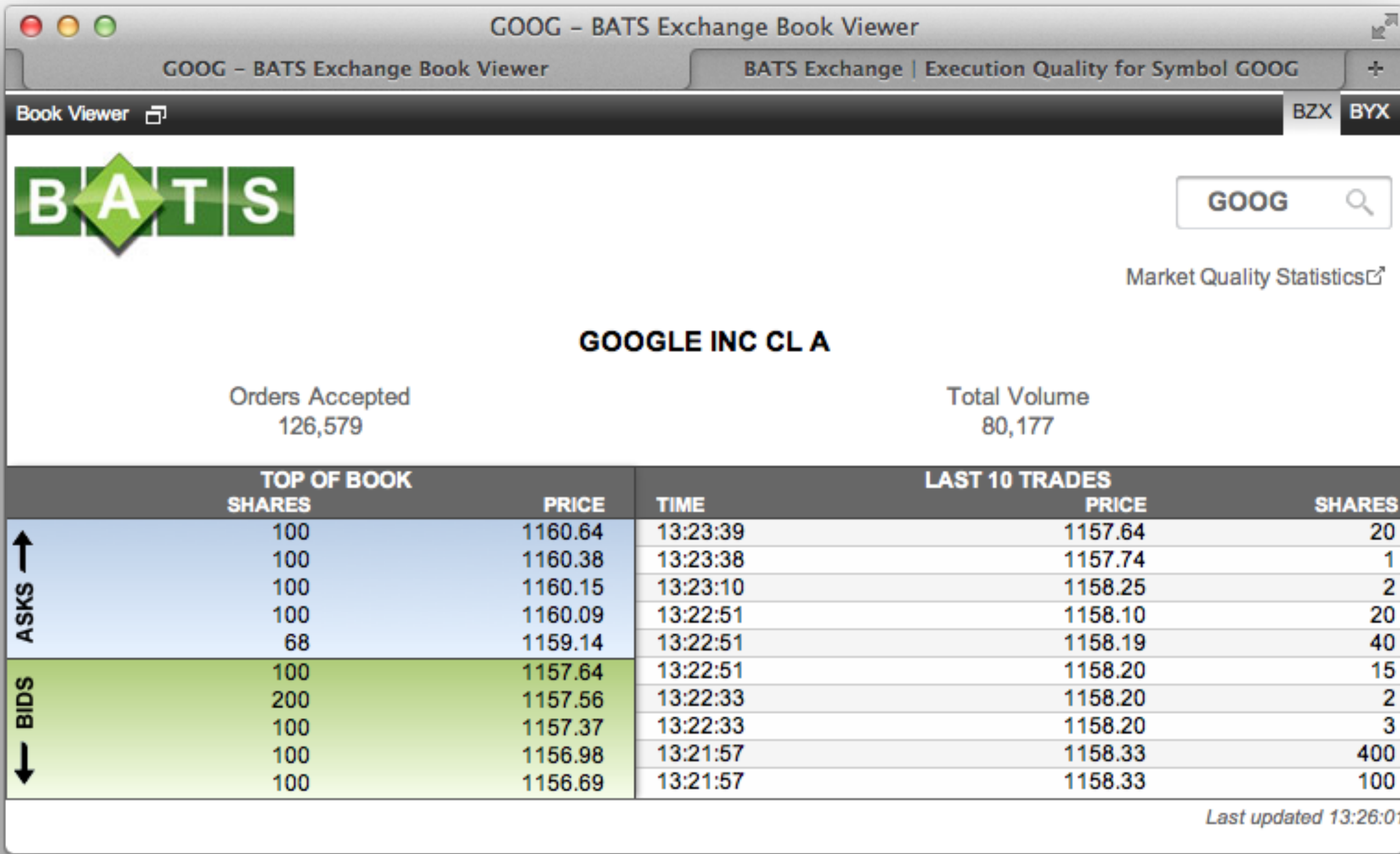
BANK OF AMERICA CORPORATION COM

Orders Accepted	Total Volume
92,577	2,059,928

	TOP OF BOOK		LAST 10 TRADES		
	SHARES	PRICE	TIME	PRICE	SHARES
ASKS	9,150	17.0200	13:27:09	16.9750	100
ASKS	11,450	17.0100	13:26:40	16.9700	100
ASKS	16,250	17.0000	13:26:40	16.9700	525
ASKS	16,350	16.9900	13:26:27	16.9700	175
ASKS	14,551	**16.9800**	13:26:12	16.9700	150
BIDS	2,650	**16.9700**	13:25:59	16.9700	1,000
BIDS	12,672	16.9600	13:25:59	16.9700	700
BIDS	22,150	16.9500	13:25:59	16.9700	200
BIDS	16,250	16.9400	13:25:59	16.9700	100
BIDS	14,650	16.9300	13:25:59	16.9700	100

Last updated 13:27:24

BATS (BZX) order book for GOOG



GOOG – BATS Exchange Book Viewer

GOOG – BATS Exchange Book Viewer | BATS Exchange | Execution Quality for Symbol GOOG

Book Viewer | BZX | BYX

BATS

GOOG

Market Quality Statistics

GOOGLE INC CL A

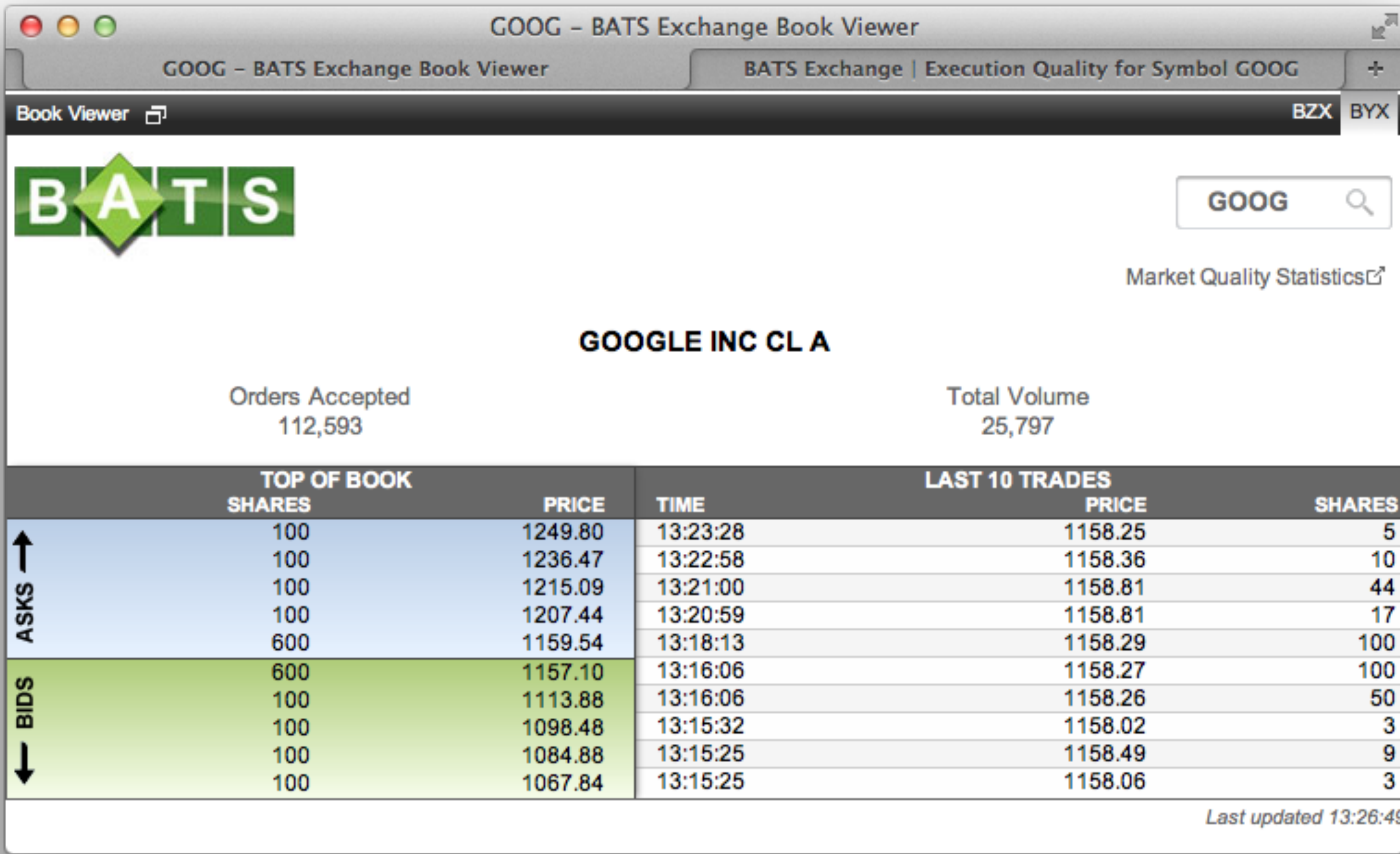
Orders Accepted 126,579

Total Volume 80,177

TOP OF BOOK			LAST 10 TRADES		
	SHARES	PRICE	TIME	PRICE	SHARES
ASKS	100	1160.64	13:23:39	1157.64	20
ASKS	100	1160.38	13:23:38	1157.74	1
ASKS	100	1160.15	13:23:10	1158.25	2
ASKS	100	1160.09	13:22:51	1158.10	20
ASKS	68	1159.14	13:22:51	1158.19	40
BIDS	100	1157.64	13:22:51	1158.20	15
BIDS	200	1157.56	13:22:33	1158.20	2
BIDS	100	1157.37	13:22:33	1158.20	3
BIDS	100	1156.98	13:21:57	1158.33	400
BIDS	100	1156.69	13:21:57	1158.33	100

Last updated 13:26:01

BATS (BYX) order book for GOOG



BZX and BYX

- BZX gives a rebate to limit order submitters (a.k.a. liquidity providers)
    - There is an incentive to place limit orders so the size at best quote is large.
    - Market orders are small.
- BYX gives a rebate to market order submitters (a.k.a. liquidity takers)
    - Queue sizes are smaller.
    - Market orders are bigger.
- Because of different fee structures on different exchanges, the displayed spread is not the true spread.

BZX fee schedule

BATS BZX Exchange Fee Schedule

Effective January 2, 2014

The following is the Schedule of Fees (pursuant to Rule 15.1(a) and (c)) for BATS Exchange, Inc. ("BZX Exchange" or "BZX"). The Schedule of Fees is divided into Equities Pricing, Options Pricing and Physical Connection Charges.

Equities Pricing:

All references to "per share" mean "per share executed." References to "adding" and "removing" liquidity mean adding liquidity to or removing liquidity from the BZX Exchange order book.

Fees for Accessing Liquidity for All Securities Priced $1.00 or Above

$0.0030 charge per share that removes liquidity

Displayed Liquidity Rebates for All Securities Priced $1.00 or Above

Standard rebate*:* $0.0020 rebate per share for adding displayed liquidity for all Members not qualifying for a tiered rebate, as set forth below

Enhanced rebates*:* Tiered rebates per share for adding displayed liquidity for qualifying Members based on ADAV[1] (added liquidity only) or ADV[1] (added and removed liquidity):

Volume Tier	***Member's ADAV is equal to or greater than average TCV[2] of:***		***Member's ADV is equal to or greater than average TCV of:***	***Rebate per share***
Tier 1	0.10%	or	0.25%	($0.0025)
Tier 2	0.20%	or	0.50%	($0.0028)
Tier 3	0.30%	or	0.75%	($0.0029)
Tier 4	0.50%	or	1.00%	($0.0030)
Tier 5	0.75%	or	1.40%	($0.0031)
Tier 6	1.00%	or	1.75%	($0.0032)

BYX fee schedule

BATS BYX Exchange Fee Schedule
Effective December 9, 2013

The following is the Schedule of Fees (pursuant to Rule 15.1(a) and (c)) for BATS Y-Exchange, Inc. ("BYX Exchange" or "BYX"). All references to "per share" mean "per share executed."

Rebates for Accessing Liquidity for All Securities Priced $1.00 or Above

$0.0003 rebate per share that removes liquidity from the BYX Exchange order book for executions by Members that have an ADV[1] equal to or greater than 0.4% of average TCV[2]
$0.0002 rebate per share that removes liquidity from the BYX Exchange order book for executions by Members that have an ADV[1] equal to or greater than 0.2% but less than 0.4% of average TCV[2]
$0.0001 rebate per share that removes liquidity from the BYX Exchange order book for executions by Members that do not qualify for an enhanced rebate based on average TCV[2] as set forth above

Liquidity Fees for All Securities Priced $1.00 or Above

No charge for adding displayed liquidity to the BYX Exchange order book that sets the national best bid or offer ("NBBO") for Members who have an ADV[1] equal to or greater than 0.4% of average TCV[2]
$0.0001 charge per share for adding displayed liquidity to the BYX Exchange order book for all other executions by Members who have an ADV[1] equal to or greater than 0.4% of average TCV[2]
$0.0001 charge per share for adding displayed liquidity to the BYX Exchange order book that sets the NBBO for Members who have an ADV[1] equal to or greater than 0.2% but less than 0.4% of average TCV[2]
$0.0002 charge per share for adding displayed liquidity to the BYX Exchange order book for all other executions by Members who have an ADV[1] equal to or greater than 0.2% but less than 0.4% of average TCV[2]
$0.0003 charge per share for adding displayed liquidity to the BYX Exchange order book for executions by Members that do not qualify for a reduced charge as set forth above
$0.0010 charge per share that adds non-displayed[3] (hidden) liquidity to the BYX Exchange order book
$0.0030 charge per share for non-displayed[3] (hidden) orders or orders subject to price sliding that add liquidity to the BYX Exchange order book and receive price improvement when executed

Large and small tick stocks

- Stocks where the tick size is large relative to the price movement per trade are *large tick* stocks. Stocks where the tick size is small relative to the price movement per trade are *small tick* stocks.
  - BAC is a large tick stock.
  - GOOG is a small tick stock.
  - MSFT is a large tick stock.
- A large tick stock has a spread of one tick nearly all of the time.
- Conversely, the spread of a small tick stock is typically greater than one tick (one cent in the US).

The SFGK zero-intelligence model

In this model due to Smith, Farmer, Gillemot and Krishnamurthy:

- Limit orders can be placed at any integer price level p where $-\infty < p < \infty$.
  - If worried about negative prices, think of these as being logarithms of the actual price.
- Limit sell orders may be placed at any level greater than the best bid $B(t)$ at time t and limit buy orders at any level less than the best offer $A(t)$.
- In particular, just as in real markets, limit orders may be placed inside the spread (if the current spread is greater than one tick).
- Market orders arrive randomly at rate μ.
- Limit orders (per price level) arrive at rate α.
- A proportion δ of existing limit orders is canceled.
- All market orders and limit orders are for one share.
- All the order flows are modeled as Poisson processes.

Analysis of the SFGK order book model

- The SFGK model depends on so few parameters that, even though it is capable of generating complex behavior, it remains amenable to analytical investigation.
- For example, it is easy to see that the asymptotic book depth far away from the best quote must be given by α/δ.

The argument goes as follows:

Asymptotic book depth

- In the steady-state, orders arriving into the book must balance orders leaving the book.
- Far away from the best quote, the probability of a limit order leaving the book as the result of an execution against a market order is very small.
- Thus we need only consider new orders arriving balancing existing orders leaving due to cancelation.
  - At a given price, for a fixed time interval Δ, the expected number of orders arriving is $\alpha\Delta$ and the expected number of orders canceled is $\delta d\Delta$ where d is the number of shares at the book at that price level.
  - For these to balance, we must have $\alpha\Delta = \delta d\Delta$, or equivalently,
$$d = \frac{\alpha}{\delta}.$$

Complex interactions

- Although the model is simple to describe, its behavior can be rather complex.
- For example, suppose there is one share at the best bid when the spread is two ticks.
    - If a new buy order is placed inside the spread one tick below the best offer $A(t)$, the best bid $B(t)$ increases and the spread $s(t)$ decreases to one tick.
    - If a market sell order arrives, the book will revert to its initial state.
- Now suppose that the same events occur but in the reverse order.
    - The market sell order increases the spread to three ticks and a new limit buy order is placed one tick below the best offer.
    - The spread decreases to one tick.
- The resulting book shape and even the mid-quote are quite different.

Distribution of order depth

- In the SFGK model, the distribution of order depth is a function of distance to the best quote.
    - Conditional on no change in the best quotes, order depth would be i.i.d at any given level of the book deeper than the best quote.
    - In fact, as shown by SFGK, this distribution is Poisson with mean α/δ.
    - However, as the order book evolves, the best quote prices do move, and each level in the book retains some memory of having been "visited" by the market price at some prior time.
- Strictly speaking, order depth distributions are therefore conditional on the entire history of the order book and in particular on the distance to the best quote.
- The further away a given level is from the best quote, the closer the order depth distribution will be to Poisson.

Dimensional analysis

- Dimensional analysis is a favorite tool of physicists used to guess the relationships between model outputs and parameters.
    - Effectively using information about scaling relationships.
- The dimensions of the parameters of the SFGK model are as follows:

Parameter	Dimensions
μ	shares/T
α	shares/(ticks T)
δ	1/T

- In the following we use these parameter dimensions to guess the functional form of relationships between measurable quantities of interest and the input parameters.

Inverse table

- For convenience, we provide (unique) expressions for dimensions in terms of input parameters.
- Inverting the previous table gives

Dimensions	Parameter combinations
shares	μ / δ
T	$1 / \delta$
ticks	μ / α

Asymptotic book depth

- The asymptotic depth d of the book far away from mid-market has the dimensions of $shares$ per $tick$. The only combination of parameters with these dimensions is α/δ.
- As we saw earlier, the exact result is in fact just α/δ so the guess in this case is exactly correct.

The spread

- The spread has dimensions of $ticks$ and the only combination of parameters with these dimensions is μ/α.
- Arguing more directly, inside the spread, limit orders are mostly removed by market orders.
  - The total rate of arrival of limit orders is $2\,\alpha\,s$ where s is the steady-state spread because as explained earlier, buy and sell orders may be placed at all levels at the spread or better.
  - The rate of removal of limit orders is just the rate of market orders μ.
- We deduce that (roughly at least)

$$s \sim \frac{\mu}{2\,\alpha}$$

The slope of the order book

- This quantity has dimensions of $shares$ per $tick$ per $tick$ (the derivative with respect to ticks of the limit order density).
- The only combination of parameters that has these dimensions is

$$\frac{\alpha^2}{\mu\,\delta}.$$

Volatility

- The variance of the times series of trade prices generated in this model must have dimensions $ticks^2/T$.
- Dimensional arguments imply that variance per unit time must therefore be proportional to

$$\frac{\mu^2\,\delta}{\alpha^2}$$

- Thus volatility must scale as

$$\frac{\mu\,\sqrt{\delta}}{\alpha}$$

Average quote size

- The average quote size Q should be given roughly by the cumulative steady-state inventory $s/2$ ticks away from mid-market where s is the average spread.
- Thus

$$Q \sim s^2\,\frac{\alpha^2}{\mu\,\delta} \sim \frac{\mu^2}{\alpha^2}\,\frac{\alpha^2}{\mu\,\delta} = \frac{\mu}{\delta}$$

  as we would expect because once again, μ/δ is the only combination of parameters with dimensions of $shares$.

Characteristic share quantity

The only quantity that can be formed from the model parameters with the dimensions of $shares$ is the characteristic share quantity

$$N_c = \frac{\mu}{2\,\delta}$$

Market impact

The most familiar market impact model is $\sigma\sqrt{L}$ where σ is volatility the liquidity L is fraction of average daily volume. In this model, for a given number of shares n,

$$L = \frac{n}{\mu}$$

and as shown above

$$\sigma \sim \frac{\mu\sqrt{\delta}}{\alpha}$$

Re-expressing that model of market impact in terms of the parameters of the SFGK model gives

$$I(n) \sim \frac{\mu\sqrt{\delta}}{\alpha}\sqrt{\frac{n}{\mu}} = \frac{1}{\alpha}\sqrt{\mu\,\delta\,n}$$

which has the correct dimensions of $ticks$.

Market impact

- In general, no matter what the form of market impact, for a given number of shares n the impact must have dimensions of $ticks$.
- The only combination of parameters with dimensions of $ticks$ is μ/α.
- We deduce that the general form of market impact must be

  $$I(n) \sim \frac{\mu}{\alpha} f\left(\frac{n}{N_c}\right) = \frac{\mu}{\alpha} f\left(\frac{2\,\delta\,n}{\mu}\right)$$

  for some function $f(.)$.
- Clearly, the square-root model of market impact is consistent with this.

Realism of SFGK model

- In real markets, order depth falls off rapidly away from the best quote.

- We would therefore not expect the behavior of the SFGK model deep in the order book to accurately mimic a real market.

- However, close to the best quote, the SFGK model *does* appear to capture some salient aspects of real markets.

- For example:
    - A model in a similar style was shown by Bollerslev, Domowitz and Wang to accurately predict the distribution of bid-offer spreads in the foreign exchange market.
    - Farmer, Patelli and Zovko estimated μ, α and δ from London Stock Exchange SETS order book data, comparing average actual bid-offer spreads and volatilities against SFGK model predictions and finding agreement to be quite good.

- As we shall see, there are inevitably some aspects of real markets that zero-intelligence models cannot mimic, notably those associated with strategic behavior of traders.

Implementation of the SFGK model

- It is impossible to simulate order arrivals and cancelations at integer price levels from $-\infty$ to $+\infty$.

- So consider only order arrivals and cancelations in a moving band of width L centered around the current best quotes.
    - L should be chosen conservatively so as to ensure minimal edge effects.
    - Within the band, the arrival rate of limit orders is α, cancelation rate is δ times outstanding shares.
    - Outside the band, orders may neither arrive nor be canceled.

Notation

Table 2: Order book events

Event	Notation
Limit buy order	LB
Market buy order	MB
Cancel buy order	CB
Market sell order	MS
Limit sell order	LS
Cancel sell order	CS

We denote the probability of an event by $\mathbb{P}_{event}$

SFGK pseudo-code

- Compute the best bid $B(t)$ and best offer $A(t)$.
- Compute the number n_B of shares on the bid side of the book from level $A(t) - 1$ to level $A(t) - L$.
- Compute the number n_A of shares on the offered side of the book from level $B(t) + 1$ to level $B(t) + L$.
- Draw a new event according to the relative probabilities $\{\mathbb{P}_{MB}, \mathbb{P}_{MS}, \mathbb{P}_{LB}, \mathbb{P}_{LS}, \mathbb{P}_{CS}, \mathbb{P}_{CB}\} \propto \{\mu/2, \mu/2, L\,\alpha, L\,\alpha, \delta\, n_A, \delta\, n_B\}$.
  - If the selected event is a limit order, draw the relative price level from $\{1, 2, \ldots, L\}$.
  - If the selected event is a cancelation, select randomly which order within the band to cancel.
- Update the order book and increment t.

Ouputs of simulation

This simulation procedure yields a time series of order book data from which we extract quantities of interest such as spread, mid-quote etc.

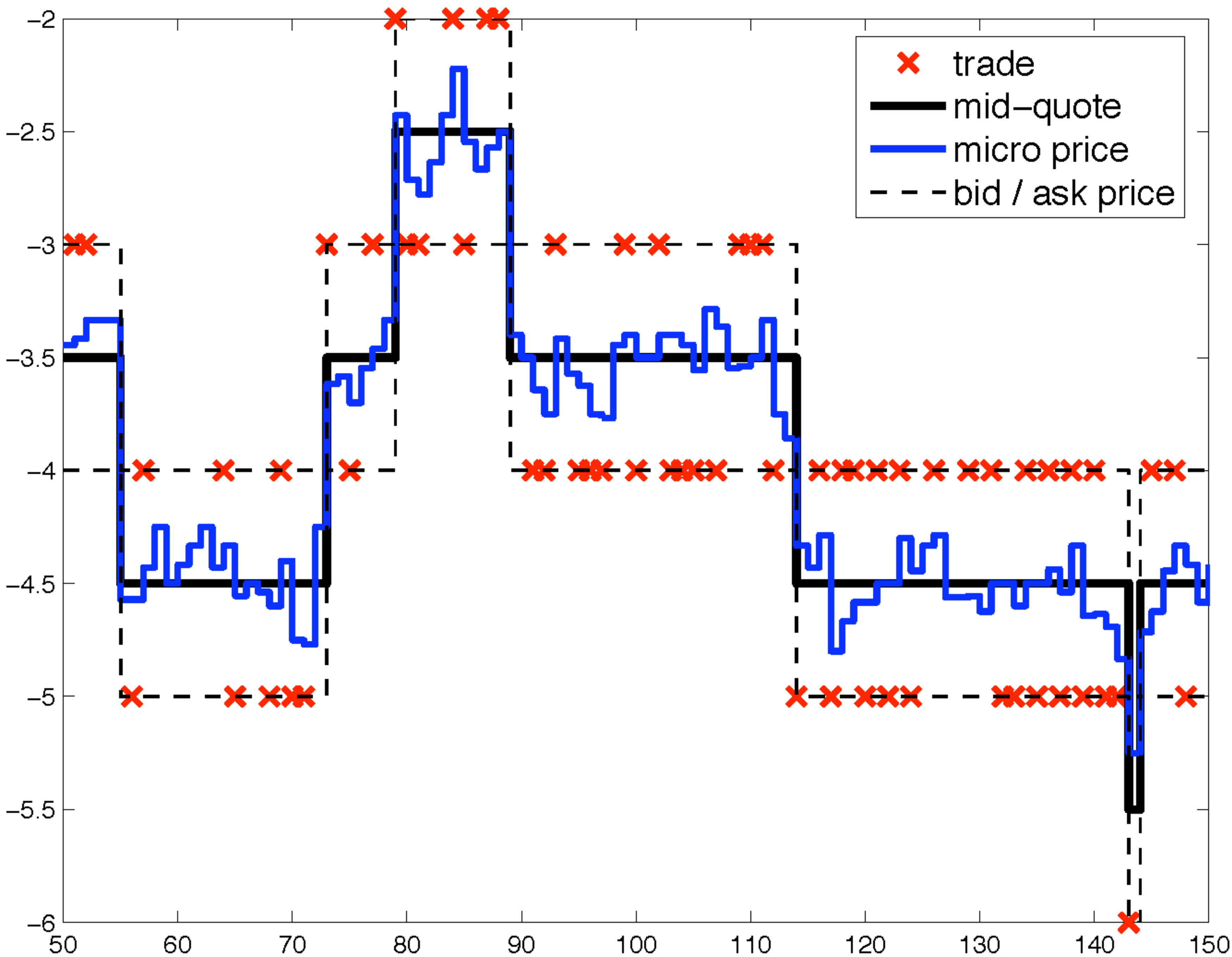


Figure 1: A typical sample path

Introduction to R

There are many places to learn about R. I recommend:

- *Introductory Statistics with R* by Peter Dalgaard
- *R for Beginners* by Emmanuel Paradis available at http://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf (http://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf)
- *R for MATLAB users* available at http://mathesaurus.sourceforge.net/octave-r.html (http://mathesaurus.sourceforge.net/octave-r.html)

R implementation of ZI simulation

In [1]:

```
%load_ext rpy2.ipython
```

In [2]:

```
%%R
download.file(url="http://mfe.baruch.cuny.edu/wp-content/uploads/2015/01/ziSet
up.zip", destfile="ziSetup.zip")
unzip(zipfile="ziSetup.zip")
```

```
Error in download.file(url = "http://mfe.baruch.cuny.edu/wp-conten
t/uploads/2015/01/ziSetup.zip",  :
  cannot open URL 'http://mfe.baruch.cuny.edu/wp-content/uploads/2
015/01/ziSetup.zip'
```

In [3]:

```
%%R
source("ziSetup.R")
```

What inside *ziSetup.R*

In [ ]:

```
%%R
#Book setup
L <- 30 #Set number of price levels to be included in iterations

# Generate initial book
LL <- 1000 #Total number of levels in buy and sell books

# Initialize book with asymptotic depth of 5 shares
initializeBook5 <- function()
{
  Price <<- -LL:LL
  # Book shape is set to equal long-term average from simulation
  buySize <<- c(rep(5,LL-8),5,4,4,3,3,2,2,1,rep(0,LL+1))
  sellSize <<- c(rep(0,LL),0,1,2,2,3,3,4,4,5,rep(5,LL-8))
  book <<- data.frame(Price, buySize, sellSize )
  if(logging==T){eventLog <<- as.data.frame(matrix(0,nrow=numEvents,ncol=2))
  colnames(eventLog)<<-c("Type","Price")
  count <<- 0
  eventType <<- c("LB","LS","CB","CS","MB","MS")
  eventDescr <<- NA}
}
```

In [ ]:

```
%%R
#Various utility functions
bestOffer <- function(){min(book$Price[book$sellSize>0])}
bestBid <- function(){max(book$Price[book$buySize>0])}
spread <- function(){bestOffer()-bestBid()}
mid <- function(){(bestOffer()+bestBid())/2}

#Functions to find mid-market
bidPosn<-function()length(book$buySize[book$Price<=bestBid()])
askPosn<-function()length(book$sellSize[book$Price<=bestOffer()])
midPosn<-function(){floor((bidPosn()+askPosn())/2)}

#Display center of book
go <- function(){book[(midPosn()-20):(midPosn()+20),]}

#Display book shape
bookShape<-function(band){c(book$buySize[midPosn()+(-band:0)],book$sellSize[mi
dPosn()+1:band])}
bookPlot<-function(band){
  plot((-band:band),bookShape(band),
       col="red",type="l",xlab="Price",ylab="Quantity")
}
```

In [ ]:

```
%%R
#Choose from L whole numbers in (1,...,L) with uniform probability
pick <- function(m){sample(1:m,1)}

# Switch logging on
logging <- T

#Buy limit order
limitBuyOrder <- function(price=NA){
  if (is.na(price))
  {prx <<- (bestOffer()-pick(L))}
  else prx <<-price
  if(logging==T){eventLog[count,]<<- c("LB",prx)}
  book$buySize[book$Price==prx]<<-book$buySize[book$Price==prx]+1}

#Sell limit order
limitSellOrder <- function(price=NA){
  if (is.na(price))
  {prx <<- (bestBid()+pick(L))}
  else prx <<-price
  if(logging==T){eventLog[count,] <<- c("LS",prx)}
  book$sellSize[book$Price==prx]<<-book$sellSize[book$Price==prx]+1}

#Cancel buy order
cancelBuyOrder<-function(price=NA){
  q<-pick(nb)
  tmp <- cumsum(rev(book$buySize))  #Cumulative buy size from 0
  posn <- length(tmp[tmp>=q]) #gives position in list where cumulative size >q
  prx <<- book$Price[posn]
```

```
  if (!is.na(price)) {prx <<-price}
  if(logging==T){eventLog[count,]<<- c("CB",prx)}
  book$buySize[posn]<<-book$buySize[posn]-1}


#Cancel sell order
cancelSellOrder<-function(price=NA){
  q<-pick(ns)
  tmp <- cumsum(book$sellSize)  #Cumulative sell size from 0
  posn <- length(tmp[tmp<q])+1
  prx <<- book$Price[posn]
  if (!is.na(price)) {prx <<-price}
  if(logging==T){eventLog[count,]<<- c("CS",prx)}
  book$sellSize[posn]<<-book$sellSize[posn]-1}

#Market buy order
marketBuyOrder <- function(){
  prx <<- bestOffer()
  if(logging==T){eventLog[count,]<<- c("MB",prx)}
  book$sellSize[book$Price==prx]<<-book$sellSize[book$Price==prx]-1}

#Market sell order
marketSellOrder <- function(){
  prx <<- bestBid()
  if(logging==T){eventLog[count,]<<- c("MS",prx)}
  book$buySize[book$Price==prx]<<-book$buySize[book$Price==prx]-1}
```

In [4]:

```
%%R
#Generate an event and update the buy and sell books
#Note that limit orders may be placed inside the spread
generateEvent <- function()
    {
    nb <<- sum(book$buySize[book$Price>=(bestOffer()-L)]); # Number of cancela
ble buy orders
    ns <<- sum(book$sellSize[book$Price<=(bestBid()+L)]); # Number of cancelab
le sell orders
    eventRate <- nb*delta + ns*delta + mu + 2*L*alpha;
    probEvent <- c(L*alpha,L*alpha,nb*delta,ns*delta,mu/2,mu/2)/eventRate;
    m <- sample(1:6, 1, replace = TRUE, probEvent); #Choose event type
    switch(m,
                limitBuyOrder(),
                limitSellOrder(),
                cancelBuyOrder(),
                cancelSellOrder(),
                marketBuyOrder(),
                marketSellOrder()
                );
    }
```

Average book shape in the ZI simulation

In [7]:

```
%%R
# Figure 2: Average book shape (This take some time to run!)

logging <- F # Very important for speed!

alpha <- 1
mu <- 10
delta <- 1/5
initializeBook5()
# Burn in for 1000 events
for(count in 1:100){
  generateEvent()
}
numEvents <- 100000 # Average over 100,000 events
avgBookShape <- bookShape(20)/numEvents
for(count in 2:numEvents){
  generateEvent()
  avgBookShape <- avgBookShape+bookShape(20)/numEvents
}

plot(-20:20,avgBookShape,main=NA,xlab="Relative price",ylab="# Shares", col="r
ed", type="b")
```

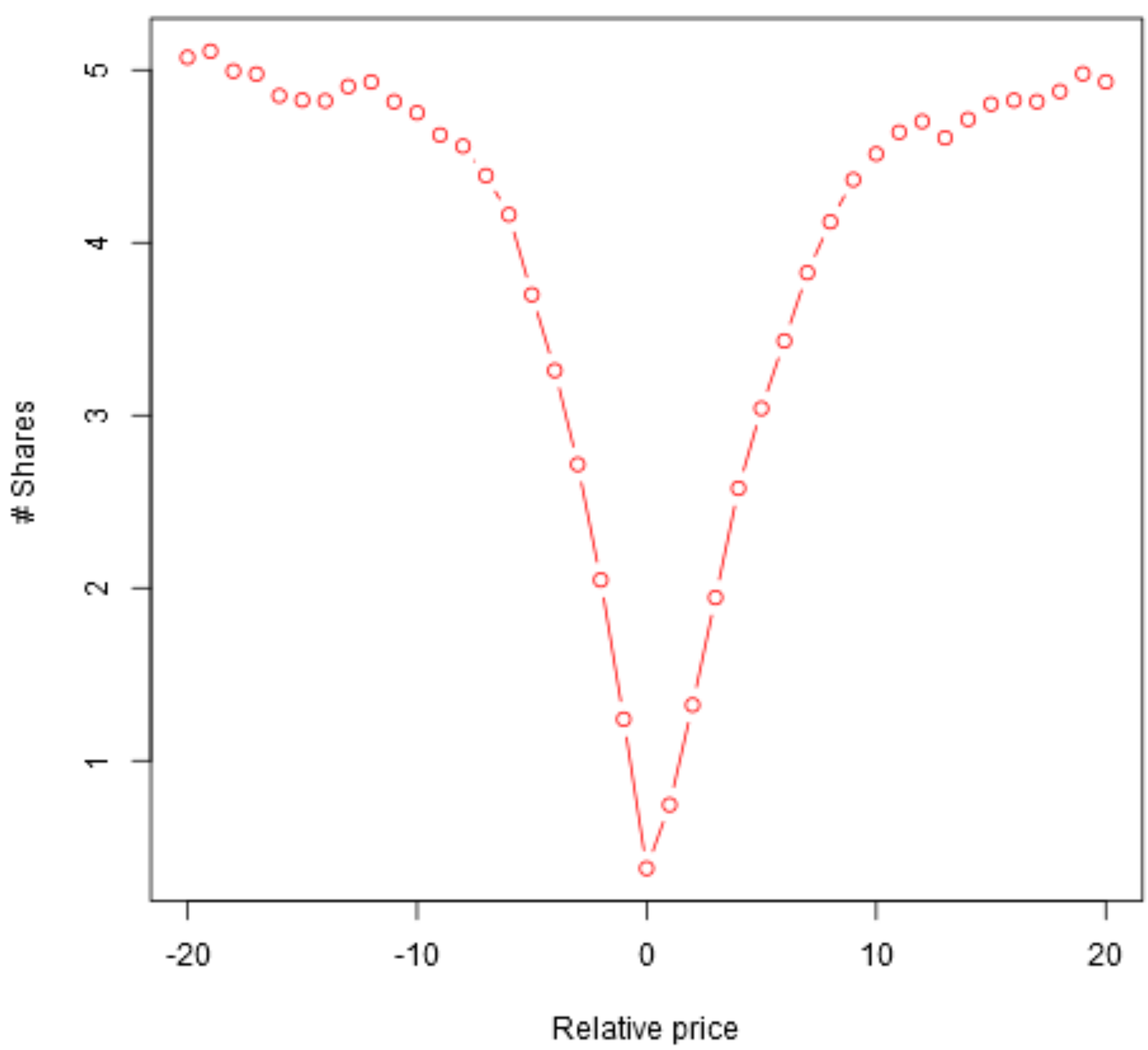


Figure 2: Average book shape over 100,000 events with $\alpha = 1$, $\mu = 10$, $\delta = 1/5$. Note that asymptotic book depth is $\alpha/\delta = 5$.

Price signal in the ZI simulation

- Even in the ZI model, the shape of the order book allows prediction of price movements.
  - Traders really would need to have zero intelligence not to condition on book shape!

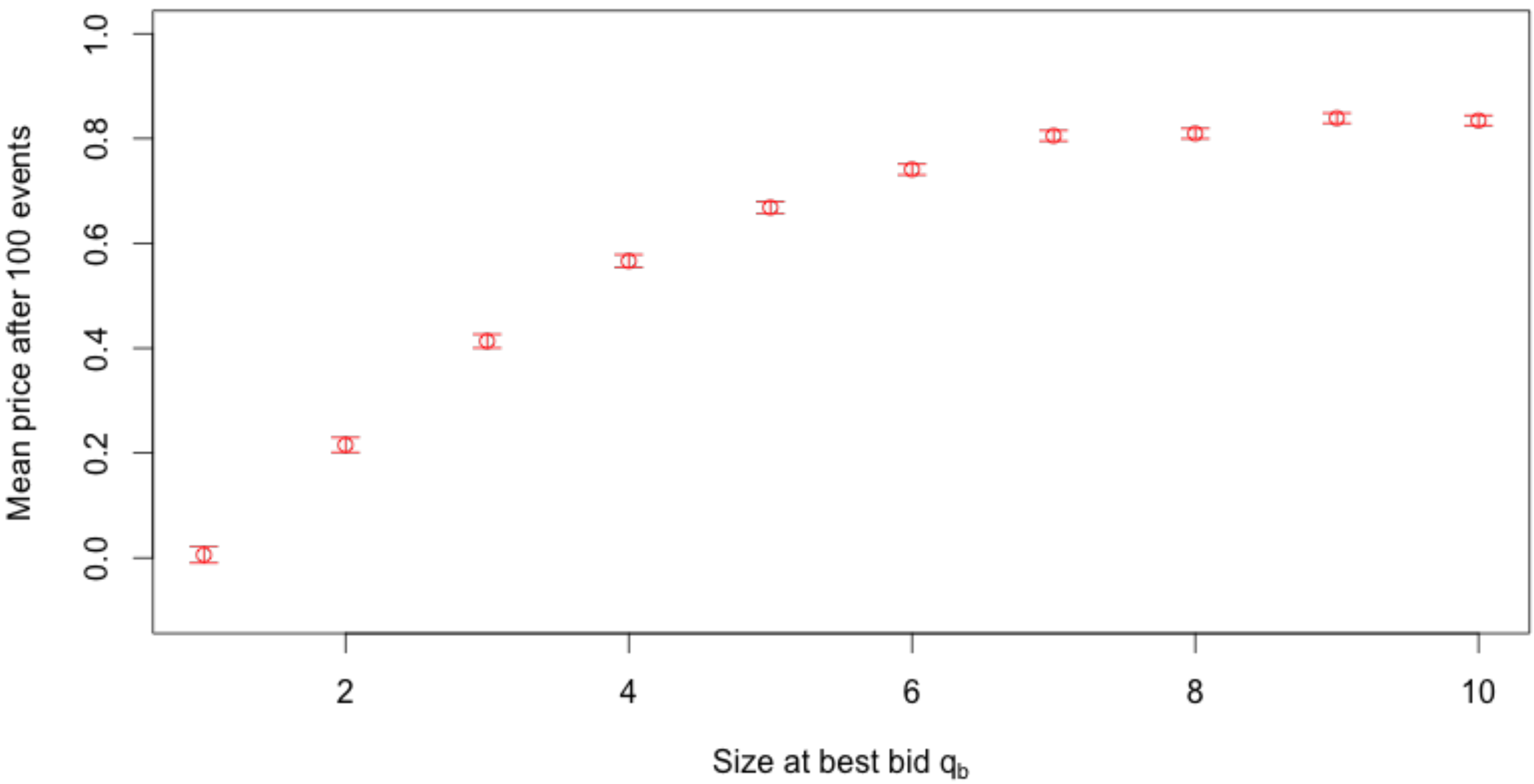


Figure 3: With one share at best offer, future price change vs size at best bid.

Further developments in ZI models

- [Cont, Stoikov and Talreja][5] analyze a generalized SFGK zero-intelligence (ZI) model as a Markov chain.
- They show how to compute conditional probabilities using Laplace transform techniques. For example:
  - The probability that the mid-price increases at its next move.
    - Depends on the order book imbalance.
  - The probability of executing an order before the mid-price moves.
  - The probability of making the spread.
    - Of interest to market makers. Cancel limit orders if price is going to move!
  - [Cont, Stoikov and Talreja][5] choose a limit order arrival ("rain") distribution and a cancelation distribution (by Δ) so as to match observed average book shapes for a Japanese stock.

Cont and de Larrard

[Cont and de Larrard][3] consider the following caricature of the ZI model, appropriate for large tick stocks:

- Only quantities at the best bid and the best ask (q_a and q_b respectively) are considered.
- All order sizes are for one unit. Cancelations are not proportional to existing quantity.
- q_a and q_b are independent.
- The price moves down if $q_b \downarrow 0$ and up if $q_a \downarrow 0$.
- After a price change, q_a and q_b are picked from a stationary distribution of queue sizes, independent of history (which could for example be empirically determined).

[Cont and de Larrard][3] derive closed-form asymptotic expressions for the following quantities:

- Distribution of duration until the next price move.
- Probability of upward move in the price.
- Autocorrelation of price changes.
- Price volatility.

Probability of price move up

Cont and de Larrard derive the following asymptotic expression assuming a balanced book (a book where the arrival rate of limit orders balances the arrival rates of market orders and cancelations):

Proposition (Proposition 2 of Cont & de Larrard)

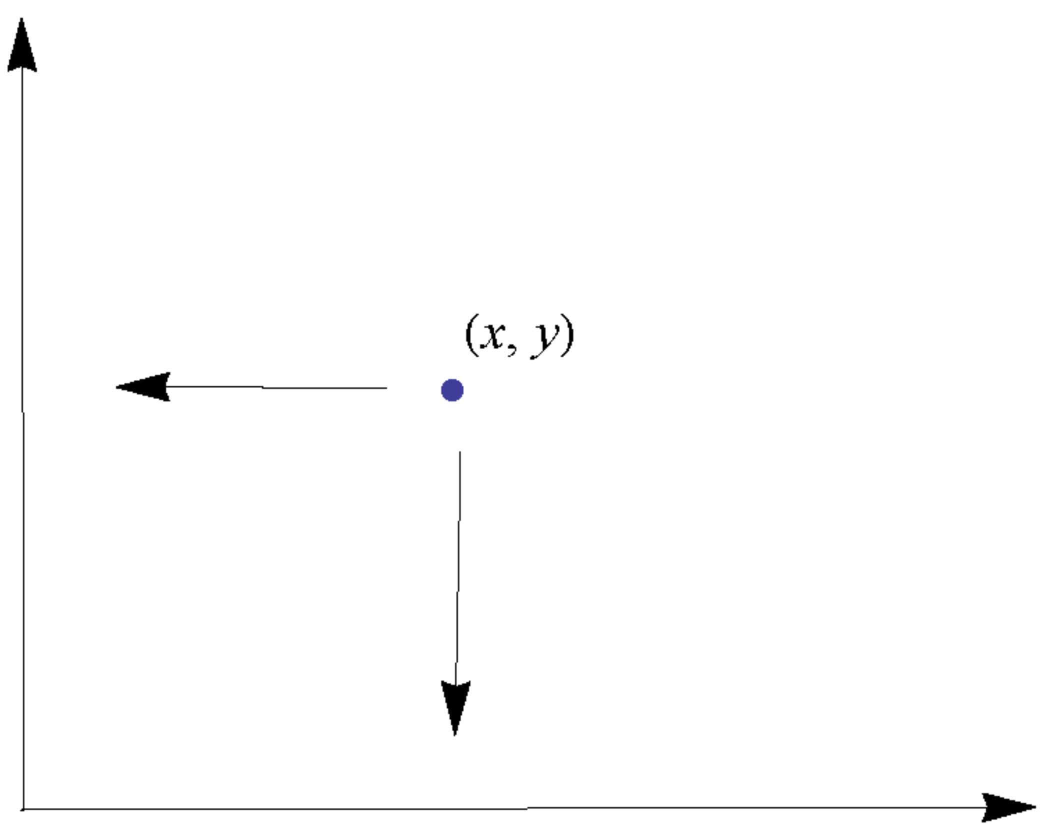
The probability $\phi(n, p)$ that the next price is an increase, conditioned on having n orders at the bid and p orders at the ask is:

$$\phi(n,p) = \frac{1}{\pi} \int_0^\pi \left(2 - \cos t - \sqrt{(2-\cos t)^2 - 1}\right)^p \frac{\sin n\,t\, \cos \frac{t}{2}}{\sin \frac{t}{2}}\, dt$$

- This leads to a natural expression for microprice in large tick markets!

Probability that next move is up

Let $\phi(x, y)$ be the probability that the next move is up given a bid size of x and an ask size of y. Then the boundary conditions for (1) are $\phi(0, y) = 0;\ \phi(x, 0) = 1$.



Asymptotic vs Cont and de Larrard formula

We set $y = 1$ and plot $\phi(x, 1)$ vs x.

Out[48]=

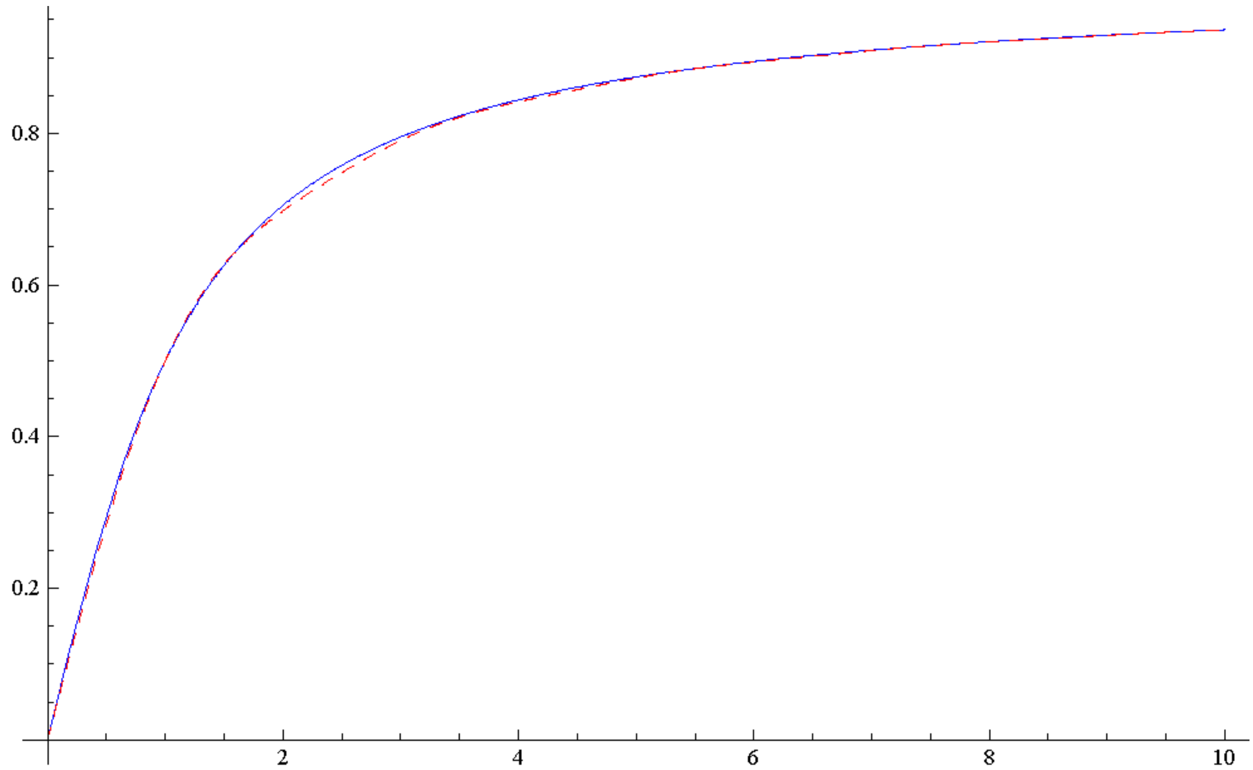


Figure 4: The blue line is simply $\frac{2}{\pi}\,\arctan\,\frac{x}{y}$; the red dashed line is the Cont and de Larrard formula $\frac{1}{\pi}\,\int_0^\pi\,dt\,\left(2-\cos t-\sqrt{(2-\cos t)^2-1}\right)^y\,\frac{\sin x\,t\;\cos\,\frac{t}{2}}{\sin\,\frac{t}{2}}$

Microprice

Now that we have the probability of an up-tick, we may define a quantity which we may consider as an estimate of the efficient price.

Denote the spread by s and the mid-price by $M_t = (B_t + A_t)/2$. Then we may define the * microprice* S_t as

(2)

$$\begin{aligned} S_t &= \text{mid-price} + \text{half spread} \times \text{expected price move} \\ &= M_t + \frac{s}{2}\,\{2\,\phi(q_b, q_a) - 1\} \\ &\approx M_t + \frac{s}{2}\left\{\frac{4}{\pi}\,\arctan\,\frac{q_b}{q_a} - 1\right\}. \end{aligned}$$

- If $q_b \ll q_a$, $S_t \approx M_t - s/2 = B_t$.
- If $q_b \gg q_a$, $S_t \approx M_t + s/2 = A_t$.

Order book imbalance

[Lipton, Pesavento and Sotiropoulos][6] define the *order book imbalance*

$$I = \frac{q_b - q_a}{q_b + q_a}.$$

In terms of I, the microprice (2) becomes

$$\begin{aligned} S_t &\approx M_t + \frac{s}{2}\left\{\frac{4}{\pi}\arctan\frac{q_b}{q_a} - 1\right\} \\ &= M_t + \frac{s}{2}\,\frac{4}{\pi}\arctan\frac{q_b - q_a}{q_b + q_a} \\ &= M_t + \frac{s}{2}\,\frac{4}{\pi}\arctan I. \end{aligned}$$

- We see that book imbalance seems to be the right variable in the sense that S_t is linear in I for small I.
  - Moreover, I is anti-symmetric in q_b and q_a.

Order book imbalance empirically

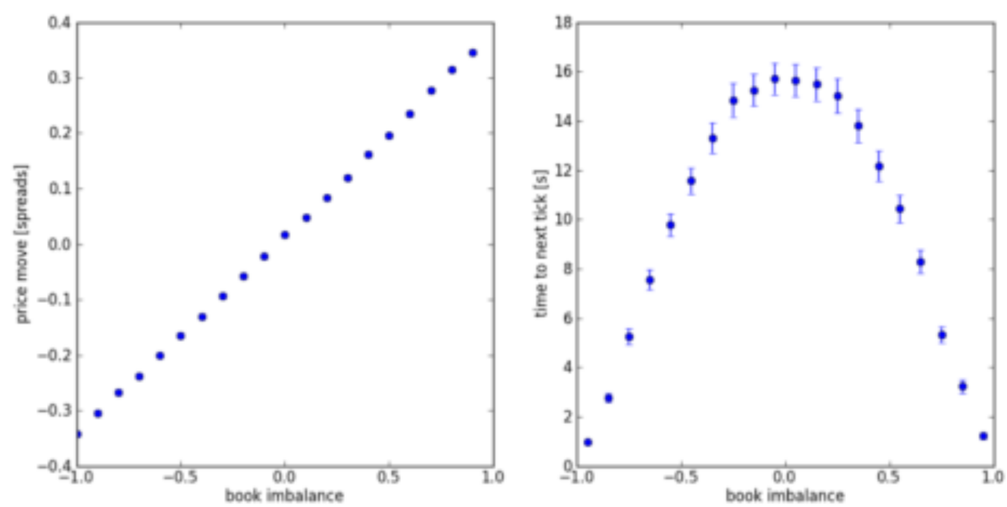


Figure 1: Average mid price move normalised by the bid-ask spread and waiting time until the next mid price move as a function of book imbalance. The data display clearly the non-martingale nature of prices at the short time scales considered here. In the case shown, the average price move can be up to a third of the spread in a highly imbalanced book. The data shown are obtained by averaging the mid price moves of VOD.L for all trading days in the first quarter of 2012. The data are bucketed according the initial imbalance in the book and averaged over the entire trading day. Error bars are calculated by treating the daily averages of price variations as independent data points. The daily average variations display error bars so small that they are hardly visible in the scale of the plot above.

Figure 5: Empirical average price move and expected time to next tick from [Lipton, Pesavento and Sotiropoulos][6]

Graph of microprice vs order book imbalance

In [6]:

```
%%R
# Figure 6: Microprice vs order book imbalance
mPrice <- function(imbalance){4/pi*atan(imbalance)}
curve(mPrice(x),from=-1,to=1,col="red",xlab="Imbalance (I)",ylab="Microprice/H
alf-spread")
abline(h=0,lty=2)
abline(v=0,lty=2)
```

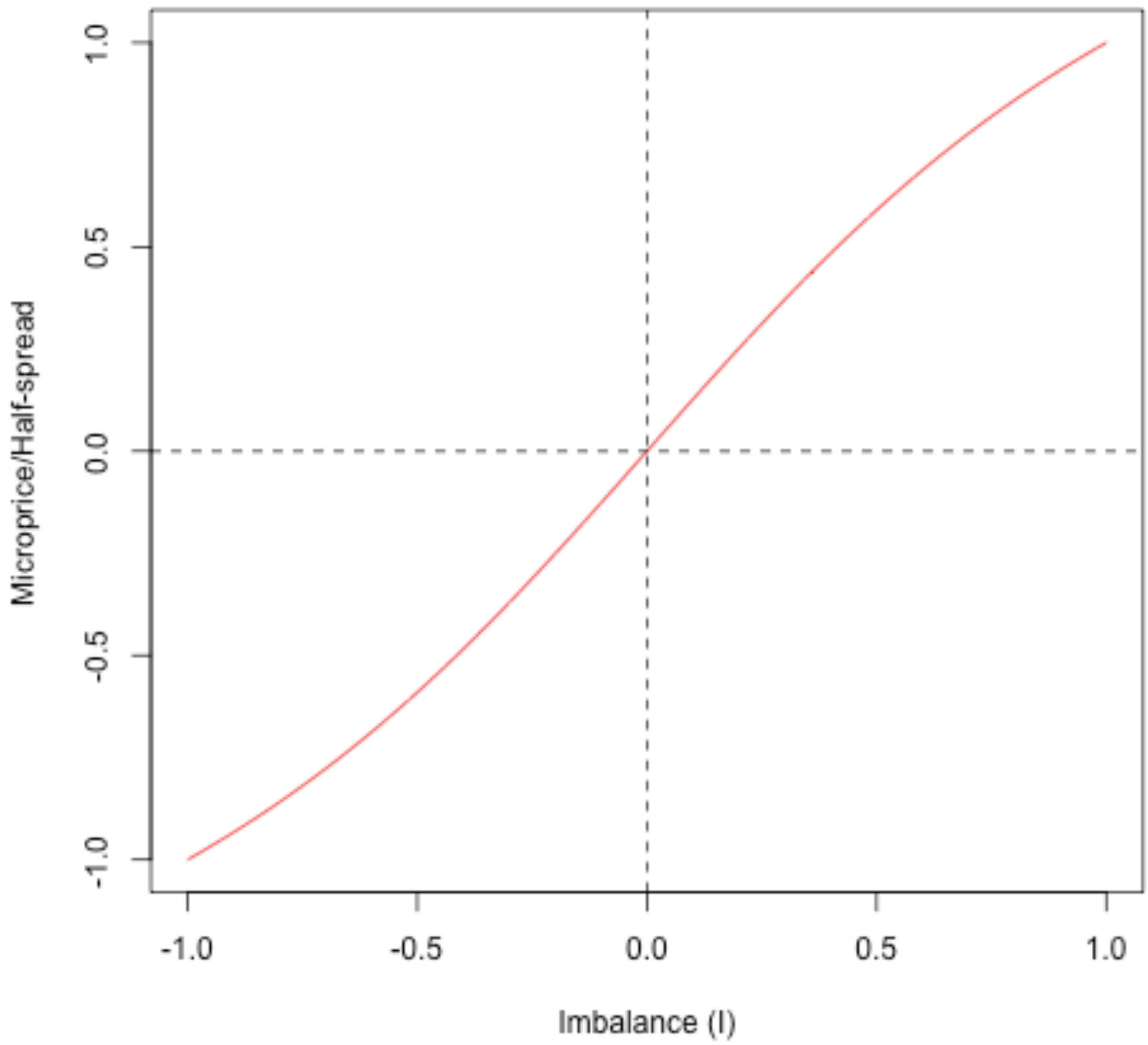


Figure 6: Microprice as a function of the order book imbalance I

- It looks nearly linear!

Price signal in the ZI simulation vs book imbalance

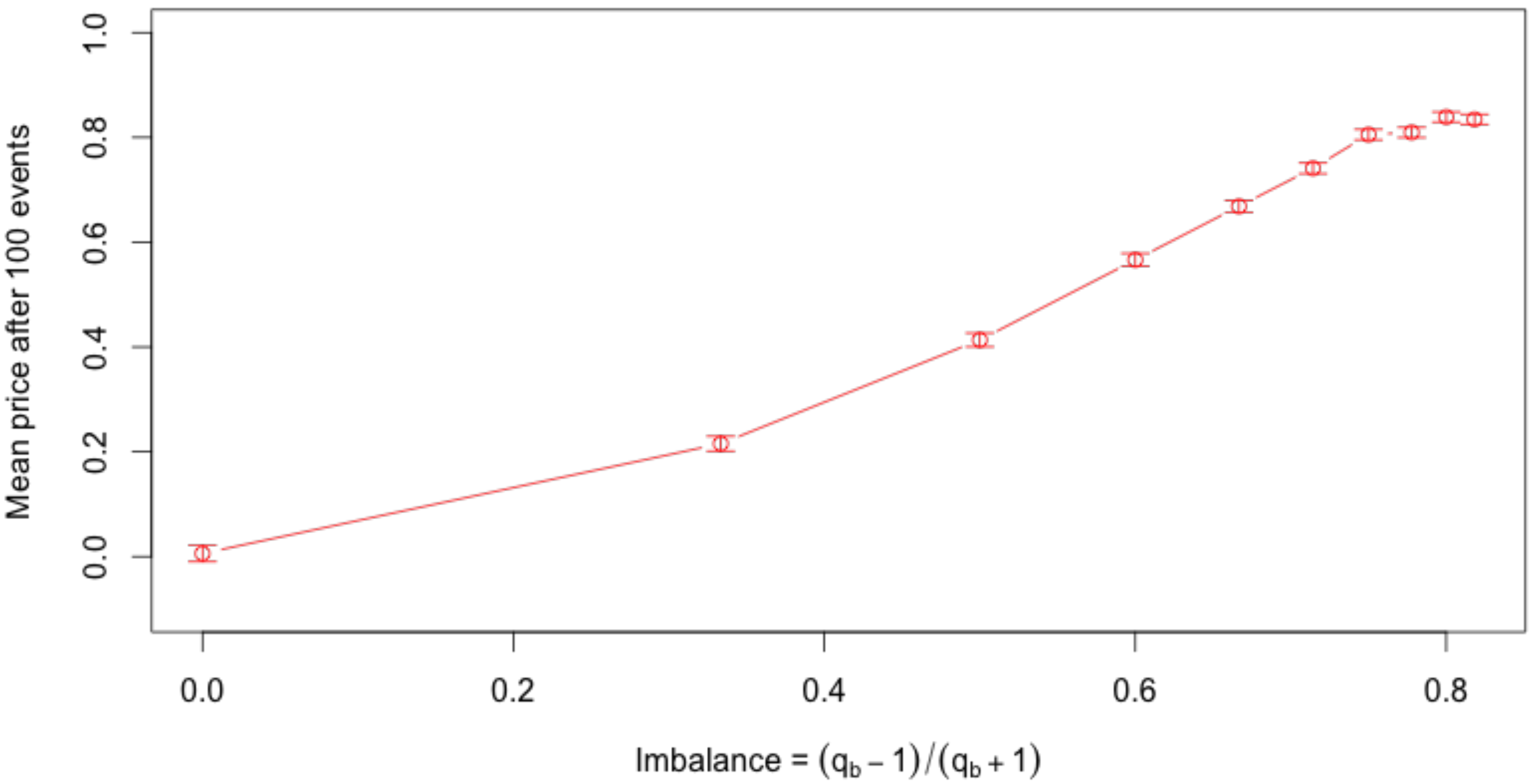


Figure 7: With one share at best offer, future price change vs book imbalance

Price signal in the ZI simulation vs microprice

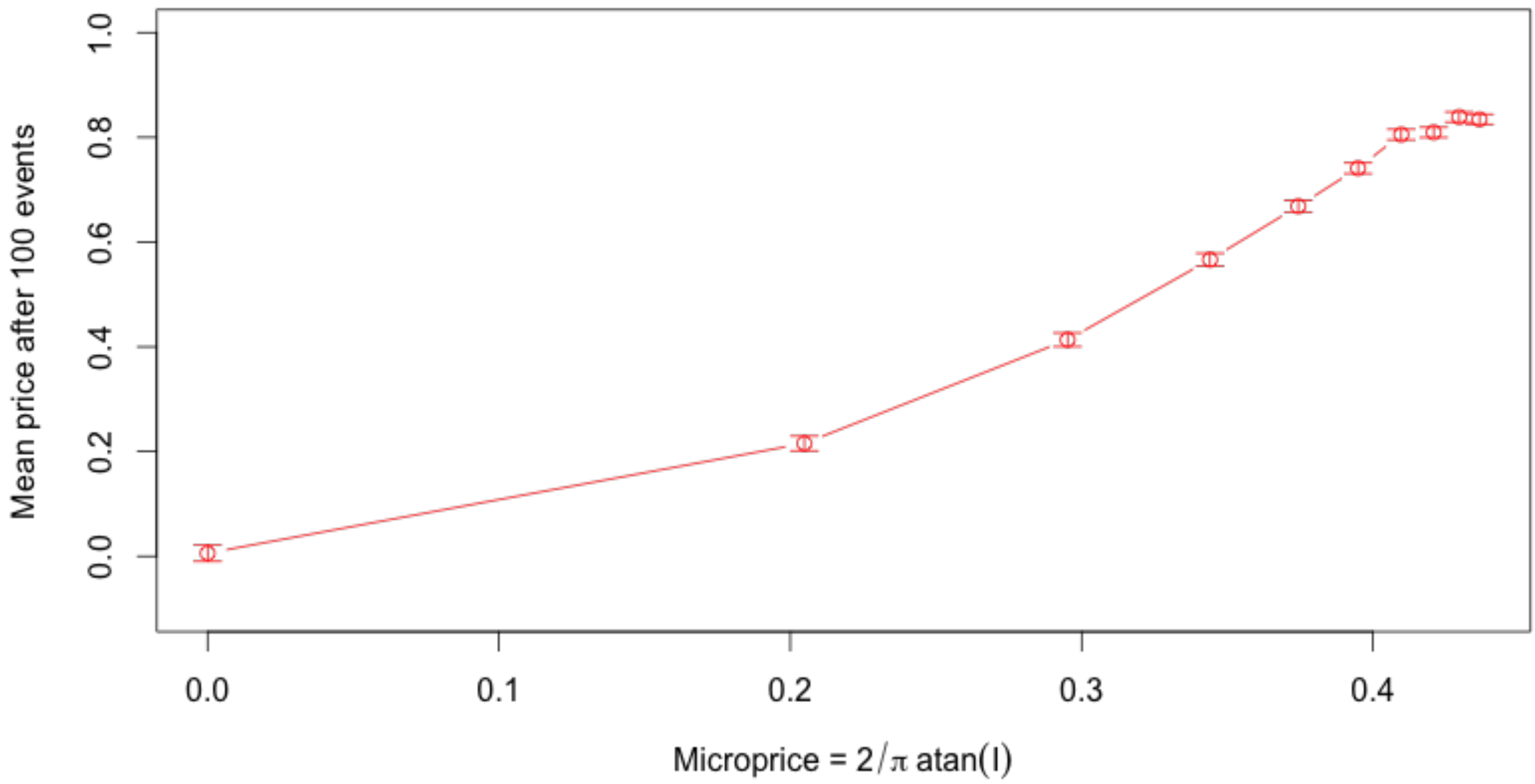


Figure 8: With one share at best offer, future price change vs microprice

References

1. ^ Frédéric Abergel and Aymen Jedidi, A Mathematical Approach to Order Book Modeling, * International Journal of Theoretical and Applied Finance*, **16**(5) 1350025-1350040 (2013).
2. ^ Jean-Philippe Bouchaud, J. Doyne Farmer, and Fabrizio Lillo, How Markets Slowly Digest Changes in Supply and Demand, in * Handbook of Financial Markets: Dynamics and Evolution* 57-156. (2009) available at http://tuvalu.santafe.edu/%7Ejdf/papers/MarketsSlowlyDigest.pdf: Sections 2.2 and 2.9.3.
3. ^ Rama Cont and Adrien de Larrard, Price dynamics in a Markovian limit order market, * SIAM J. Finan. Math.*, ** 4**(1) 1-25 (2013).
4. ^ Rama Cont and Adrien de Larrard, Order book dynamics in liquid markets: limit theorems and diffusion approximations, * SSRN*, (2011).
5. ^ Rama Cont, Sasha Stoikov and Rishi Talreja, A stochastic model for order book dynamics, * Operations Research*, **58** 549-563 (2010).
6. ^ Alexander Lipton, Umberto Pesavento, and Michael G Sotiropoulos, Trade arrival dynamics and quote imbalance in a limit order book, * arXiv* (2013).