FIN 580: Market Microstructure and Trading

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

August 24, 2015

Outline

1. Course Overview
   - Microstructure and Algorithms
2. Course Mechanics
   - Preliminaries
   - Assignments and Grading
3. Background on Capital Markets
   - Basic Trading Concepts
   - Characteristics and Elements of Markets

Microstructure, Old and New

- Paraphrased advice from one of my esteemed Harvard advisors:
  - (2008): "Adam, microstructure is a backwater of finance that stagnated years ago"
  - (2011): "Adam, I think that microstructure is the single most exciting area of finance today"
- High-frequency trading and market microstructure are my areas of research
  - I share the 2011 viewpoint...as will you, by the end of this course

What is Microstructure?

- Typical framework for economic analysis of financial markets:
  - Phase 1: Agents, preferences, endowments, information, etc.
  - Phase 2: ?
  - Phase 3: Equilibrium (allocations, prices, beliefs, profits, etc.)
- "Phase 2" = Trading
  - *Market microstructure* is the study of the trading process and trading mechanisms
  - Microstructure examines how financial markets move from disequilibrium to equilibrium

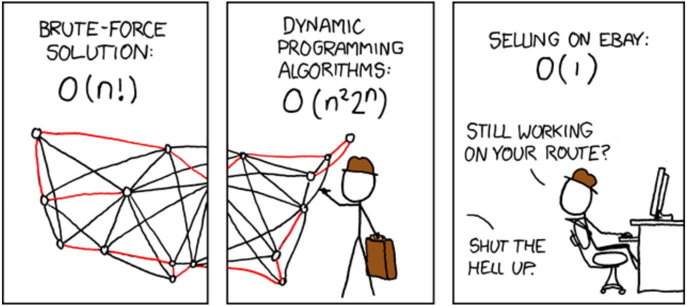
Microstructure Reborn
Algorithmic Trading

- Explosive growth of algorithmic trading has lent new importance microstructure
- "Micro" structural features have recently had **massive** effects
  - Goldman Sachs "programming glitch": losses undisclosed, probably $\approx$ \$100 million
  - Knight Capital lost \$460 million on one "trading error"
  - "Flash Crash" of 2010: in 3 minutes, E-mini and SPY fell 3%
  - High-frequency traders earn hundreds of millions of dollars per year in the E-mini alone

Algorithms, Microstructure, and FIN 580

- This course will emphasize applications of microstructure to algorithmic trading
  - Use economic theory to better understand, utilize, and design trading algorithms

Interlude: Travelling Salesmen



Algorithms, Microstructure, and FIN 580

- This course will emphasize applications of microstructure to algorithmic trading
  - Use economic theory to better understand, utilize, and design trading algorithms
  - Novel microstructure issues related to algorithmic trading
- Microstructure and algorithmic trading affect numerous financial tasks, including:
  - Portfolio management
  - Strategy-feasibility assessment
  - Risk management

Overarching Course Questions

- What do various market participants seek to accomplish? Why?
- What information does market activity transmit? How?
  - Who wants to know certain types information?
  - Who wants to keep that information secret?
- When and how does speed enter the picture?
  - What are the appropriate notions of "speed"?

Logistics

- Lecture/office hours times and locations
- Emergency procedures
- Administrative policies (UIUC 2015-2016 Student Code)
  - Accommodations for students with disabilities
  - Late work, make-up exams
- Enrollment: MSFE students only

Course Materials

- Textbooks and Readings
  - O'Hara
  - Harris
  - Hasbrouck
- Lecture Slides
- YOU MAY NOT RECORD LECTURES!
- Software
  - Matlab

Problem Sets

- Eight problem sets
  - Due dates, submission procedures as per the syllabus
  - Last two "problem sets" will be more open-ended, essentially "projects"
- Problem sets will be almost entirely (Matlab-) programming-based
  - Coordinating on Matlab facilitates modular/interoperable code
  - Still, minimal prior Matlab background assumed
- Solutions posted shortly after each problem set is due
  - Many problem sets will rely on code/scripts from past assignments
  - *Use the posted solutions to ensure that you have working versions of all assigned scripts*

Collaboration on Problem Sets

Collaboration on problem sets is allowed and encouraged (unless explicitly noted otherwise), but:

1. Always list all the other students with whom you collaborated
2. If you and your collaborator(s) have identical solutions, please turn in on a single solution set with all of your names
3. If you collaborate on a problem set, *you are still responsible for understanding exactly what every part of your code does, and why!*

Exams

- Two midterm exams and a final exam
- Mostly programming-based
  - Exams also include some straight-forward questions from assigned readings
- The exams are cumulative, but the second midterm will emphasize material from the problem sets that were due after the first midterm

Exam Format

Because I allow collaboration on problem sets, the exam programming questions are intended to assess each student's individual understanding

- Bring all of the (corrected) code from previous problem sets to each exam
- Exam programming questions will be just *slightly* different from those on past problem sets
- Make some minor modifications to your past code to answer the exam questions, then explain what you changed and why

Grading

- Assignment relative weights
  - Problem sets #1-6: 30%
  - Problem sets #7-8: 20%
  - Midterm #1: 15%
  - Midterm #2: 10%
  - Final exam: 25%
- Grades will be curved to a median of A- (I will use +/- grades)
- Massive underperformance on a midterm will trigger downward adjustment of the associated problem-set scores
  - Refer to syllabus for full details
  - I expect this to be an off-equilibrium path...

Where Does Trading Occur?

- Exchanges
  - Operate systems through which traders interact and find counterparties ("trading facilities")
    - Historically: trading floors and pits
    - Modern: computer servers and matching engines
  - Dictate a standardized structure for the trading process
    - Trading rules and practices
    - Recording and dissemination of market data
- Over-the-counter
  - Bilaterally negotiated transactions (off-exchange)
- Alternative Trading Systems, Internalization, etc.
  - Later in the course...

Exchanges

- This term, we will primarily consider trading on exchanges
  - More algorithm-friendly than OTC trading
  - Similar (if not quite identical) underlying economic principles
- Stock exchanges vs. futures exchanges
  - Listing standards (stocks) vs. in-house clearing (futures)
  - A given stock may trade of several exchanges; a given futures contract usually trades on just one exchange
- (Also, no short-sale complications for futures)

Who Trades on Exchanges?

- Literal answer:
  - Exchange members (proprietary trading)
  - Brokers (agency trading)
    - Prime brokers
    - Retail/discount/online brokers
- Brokers trade on behalf of their clients to earn commissions/fees
  - Why are the clients trading?
  - And why are the exchange members trading?

Motives for Trading

- Information
- Non-informational motives
  - Hedging
  - Liquidity (i.e., needing cash)
  - Gambling
- And at least one more...later in the course!

Introducing Orders

- Orders are simply trading instructions that traders convey to the exchange
- Orders typically specify (at least) the following information:
  - Instrument to be traded
  - Buy vs. sell
  - Quantity
  - Price (as applicable)
  - Trader ID
  - Order ID#
  - Date/time of order placement
  - Special instructions/ order type

Two Key Order-Types

- Market Order
  - Specifies quantity, but not price
  - The specified quantity is traded almost immediately
  - Executes almost immediately, but the prices at which it executes can be arbitrarily bad
    - And a market order always "crosses the bid-ask spread"
    - I.e., a market buy first trades at the best ask, a market sell first trades at the best bid
- Limit Order
  - Specifies quantity *and* price
  - Execution price is predetermined, but the time it takes for the limit order to execute can be infinite

Marketable Limit Orders

- Limit buy (sell) order priced above (below) the best ask (bid)
- Marketable limit orders execute immediately against existing opposite orders until either
  - The full quantity of the marketable limit order executes, or
  - The marketable limit order has executed against and filled all existing opposite orders priced weakly better than the limit price
- Marketable limit orders provide a mechanism to execute trades very quickly without incurring unbounded price risk
- Unless otherwise noted:
  - I will use "marketable limit order" and "market order" interchangeably
  - I will use "limit order" to mean "non-marketable limit order"

Gains and Losses on Orders

Suppose that prices follow a random walk, and the bid-ask spread is stationary

- A trader who places a *limit* order can potentially earn some fraction of the bid-ask spread
  - Placing a limit order gives other traders the opportunity to trade with you exactly when they want to
- A trader who places a *market* order will always lose some fraction of the bid-ask spread (in expectation)
  - Placing a market order lets you trade exactly when you want to

Price(s)

The price of a security isn't a precise concept

- There are multiple, distinct numbers that could be equally valid characterizations of "price" (in some sense)
  - Last sale price
  - Current best ask quote (lowest quoted selling price)
  - Current best bid quote (highest quoted bid price)
  - Mid-point price ($\frac{bestask+bestbid}{2}$)
- best_ask-best_bid = bid-ask spread

Price(?)

- Also possible that none of the choices on the previous slide correspond to anything meaningful
    - Last sale price:
        - Information is too out of date
    - Current best ask (bid) quote
        - Next transaction might be a sell (buy)
        - Quote might be indicative instead of firm
    - Mid-point price ($\frac{bestask+bestbid}{2}$)
        - bid-ask spread may not be symmetric

Liquidity

- "Liquidity" is a vague term that loosely refers to the cost and difficulty of trading
- Some more precise market characteristics:
  - Immediacy (how soon can you trade)
  - Tightness (bid-ask spread)
  - Depth (how large a quantity can be traded near the best bid/ask)
  - Resiliency (how soon does transient price-impact dissipate)
    - More about resiliency later in the course

Summary and Looking Ahead

- Ongoing:
  - Objectives, motives and strategies of market participants
  - Information transmission through markets
  - Implications for, and applications to trading algorithms
- Today:
  - General components of markets
    - Exchanges, traders, orders, prices
  - Basic terminology and characteristics
- Next lecture: A specific market structure (order-driven markets)
  - PS#1 posted afternoon of 8/24, Due Monday, 8/31

Readings

- This lecture:
  - Course syllabus
  - Hasbrouck Ch. 2
  - (Hasbrouck Ch. 1)
  - Harris Ch. 1, Ch. 3, Ch. 4.1-4.4 (pp. 68 - 77)
- Next lecture:
  - Hasbrouck Ch. 3
  - (Hasbrouck Ch. 5, Ch. 6)
  - Harris Ch. 5, Ch. 6

FIN 580 Section MMT
Lecture 2

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

August 26, 2015

Outline

Connecting Buyers and Sellers

- Traders need some mechanism ("execution system") to find counterparties
  - Brokered trading system
  - Quote-driven system
    - Dealer markets (cf. Hasbrouck Ch. 6)
  - Order-driven system
    - Rules for order-precedence and trade-pricing
- Limit-order markets
  - Dominant structure for equities, listed futures, listed options
  - Continuous trading, no *explicit* intermediation (e.g. by a dealer)
- Call auctions (cf. Hasbrouck Ch. 5)
  - Open and close of continuous trading sessions in some markets
  - Also used independently

Recall: Order Types

- Market order
  - Specifies quantity, but not price
  - Immediately executes
- Limit order
  - Specifies quantity *and* price
  - Execution time is indeterminate
- "Marketable" limit order: buy (sell) priced above (below) the best ask (bid)
  - Immediate (but possibly only partial) execution
  - We can think of market orders as a special case of limit orders

The Limit Order Book

- Limit order book (a.k.a "orderbook" a.k.a. "book")
  - Collection of all active limit orders
    - Limit orders in the book called "resting" or "passive" orders
  - Two "sides" of the orderbook: buy side and sell side
- Transactions occur when a new order arrives and executes against resting order(s) in the book
  - Order that initiates trade called the "aggressive" order
    - Hasbrouck uses 'aggressive' differently—ignore
    - (Joel: apologies if you happen to see these slides)

Example Sequence of Events

1. Trader places a new order, suppose it's "buy 100 shares at a price no greater than \$60.00"
2. If there are resting sell orders in the book at prices weakly below \$60.00, the new order executes against them until
    1. The full 100 shares have been bought, or
    2. All resting sell orders at prices weakly below \$60.00 have been filled
3. The new order is entered into the buy side of the book, with its original quantity (100) replaced by the remaining unexecuted quantity
    1. If the new order fully executed (so the remaining unexecuted quantity is zero), it is not entered into the book

Priority

Consider the following book sell-side for the previous example:

Order ID	Price	Quantity
I	58.00	5
II	59.00	3
III	59.00	22
IV	59.00	77
V	60.00	100
VI	61.00	45

Which sell orders would execute?

Priority

Order ID	Price	Quantity	Executes?
I	58.00	5	?
II	59.00	3	?
III	59.00	22	?
IV	59.00	77	?
V	60.00	100	?
VI	61.00	45	NO

- Clearly not order VI (price is too high)
- But what about the rest?

Rules to Determine Priority I

- First consideration: price
  - Sell limit orders with lower prices get priority
  - Buy limit orders with higher prices get priority
- Price priority is pretty universal
  - Think about traders' surpluses: why does price priority help? But what could still go wrong?

Order ID	Price	Quantity	Executes?
I	58.00	5	YES
II	59.00	3	?
III	59.00	22	?
IV	59.00	77	?
V	60.00	100	NO
VI	61.00	45	NO

Rules to Determine Priority II

- Second consideration: typically time
  - At a given price, orders entered earlier execute earlier

Order ID	Price	Quantity	Entry Time	Executes?
I	58.00	5	10:01:33	YES
II	59.00	3	10:54:03	NO
III	59.00	22	9:48:59	YES
IV	59.00	77	10:01:32	PARTIAL (73)
V	60.00	100	9:36:13	NO
VI	61.00	45	10:53:58	NO

Alternative/Additional Determinants of Priority

- Visibility priority:
  - Visible orders get priority over hidden orders at the same price
  - Typically, priority among visible orders at the same price is determined by time (and similarly among hidden orders)
- Size "priority"
  - Typically, size incentives accomplished via pro-rata matching among all resting orders at the same price
- We'll mostly work with pure price-time priority in this course

Time Priority and Multiple Exchanges
Preview

- An important complication with time priority arises for an instrument traded on multiple exchanges
- If we consolidate the limit order books from two spatially separated exchanges, there's no perfect way to decide time priority between two limit orders from different exchanges
  - Not just a technological issue...relativity also gets in the way
- Price priority across different exchanges is (more) feasible, and part of Reg NMS relates to this
  - More on Reg NMS near the end of the term
- Later in the course we'll look at some multiple-exchange issues
  - For now, picture futures markets...

Matching Engine Overview

- For the first problem set, you will build a simple matching engine to process orders from simulated traders
  - When an order arrives, execute any matches against resting orders using price-time priority
  - Enter the unexecuted portion of the new order into the book
  - Only entries and executions at this point; cancellations, modifications, etc. later
- Central task is to maintain a properly organized orderbook
  - Execute correct transactions
  - Update the book to reflect new orders and executions
- Secondary tasks
  - Extract and record the (simulated) market data in real time
  - Analyze various aspects of the recorded market data

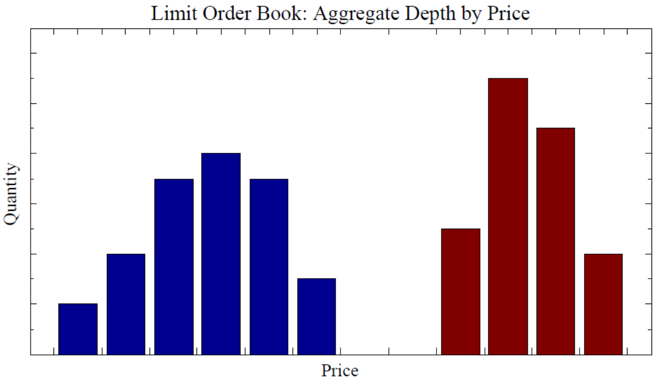
Features of Orders
Real and Simulated

- In real markets, orders are submitted with (at least) the following information:
  - Instrument to be traded
  - Order type/special instructions
  - Direction (buy vs. sell)
  - Quantity
  - Price (as applicable)
  - Trader ID
- Real matching engines tag each order with (at least)
  - Date/time of arrival
  - Order ID

Features of Orders
Real and Simulated

- In our simulated market right now:
  - Instrument to be traded: only one possibility
  - Order type/special instructions: always limit orders, no special instruction
  - Direction (buy vs. sell): NEEDED
  - Quantity: NEEDED
  - Price: NEEDED
  - Trader ID: NEEDED
- Matlab matching engine should tag each order with
  - Time of arrival (discrete time, one order each period)
  - Order ID (use time of arrival)

Sample Depth Summary



Price(s)

The price of a security isn't a precise concept

- There are multiple, distinct numbers that could be equally valid characterizations of "price" (in some sense)
    - Last transaction price
    - Current best ask (lowest quoted selling price)
    - Current best bid (highest quoted bid price)
    - Mid-point price ($\frac{bestask+bestbid}{2}$)
- best_ask-best_bid = bid-ask spread

Price(?)

- Also possible that none of the choices on the previous slide correspond to anything meaningful
  - Last transaction price:
    - Information is too out of date
  - Current best ask (bid)
    - Next aggressive order might be a sell (buy)
    - Depth at best might be tiny
  - Mid-point price ($\frac{bestask+bestbid}{2}$)
    - bid-ask spread may not be symmetric

Liquidity

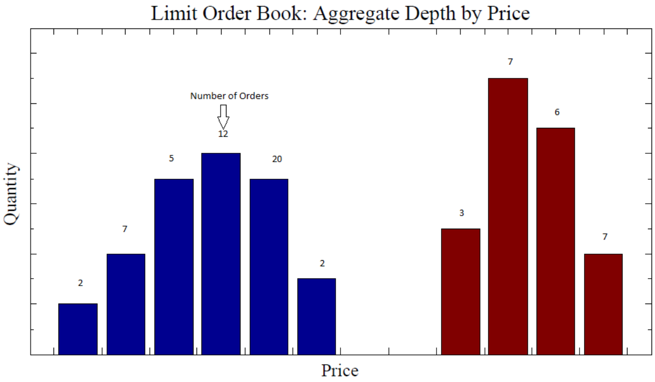
- "Liquidity" is a vague term that loosely refers to the cost and difficulty of trading
- Some more precise market characteristics:
  - Immediacy (how soon can you trade)
  - Tightness (bid-ask spread)
  - Depth (how large a quantity can be traded near the best bid/ask)
  - Resiliency (how soon does transient price-impact dissipate)
    - More about resiliency later in the course
- The ambiguity surrounding price and the ambiguity surrounding liquidity are linked...

Incremental Orderbook Management

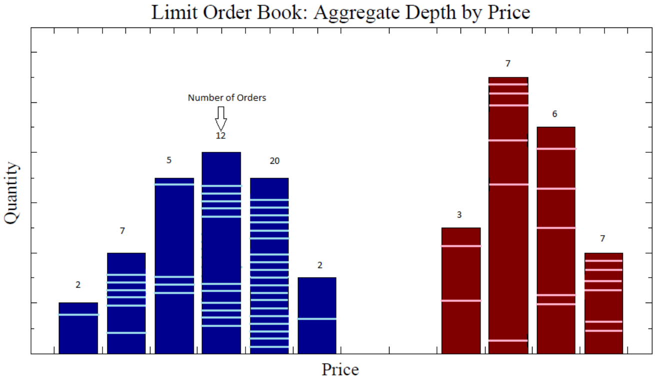
Direct Application of Matching-Engine Expertise

- Faster and more efficient to send/receive messages about *changes* in the book than the entire (slightly new) state of the book
- Maintaining an up-to-date book on a tick-to-tick basis is relevant in industry
  - Basic incremental orderbook management isn't challenging
  - But some related issues *are* very difficult
- E.g., determining your order's location in the queue

Sample Depth Summary with Order Count



Sample Depth with Orders Distinguished
(Rarely Disclosed Directly in Market Data)



Analog Reconstruction

Summary and Looking Ahead

- Today:
  - Limit order book markets operate under fairly simple rules
    - Key challenge #1: managing/organizing large, complicated data in near-real-time
    - Key challenge #2: translating high-dimensional "summary" data into useful information
  - Specific application: incremental LOB management and estimating queue position
- Next lecture: market-making and immediacy

Readings

- This lecture:
  - Hasbrouck Ch. 3
  - (Hasbrouck Ch. 5, Ch. 6)
  - Harris Ch. 5, Ch. 6 [both chapters are slightly out-of-date]
- Next lecture:
  - *O'Hara Ch. 1
  - Harris Ch. 8, Ch. 14

FIN 580 Section MMT
Lecture 3

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

August 31, 2015

Comments Regarding PS#1

- Don't worry
- Historically, PS#1 aplexes $\approx$ 75% of FIN580 students

Comments Regarding PS#2 and Beyond

- Don't worry
- Hereafter, problem sets will focus more on building/analyzing trading algorithms
  - A.K.A., trading "robots"
- Robots aren't really that tough...
  - You just need to know how to deal with them

Dealing with Some (Non-Trading-) Robots

Outline

Designated Market Makers

- Recall: quote-driven markets (dealer markets)
  - Explicitly intermediated by a dealer a.k.a. "market-maker"
    - Market-maker is a counterparty in every trade
  - Facilitates continuous trading opportunities in sparsely populated markets
- Hybrid systems
  - "Designated market-maker" responsible for maintaining continuous trading if other traders' limit orders are insufficient
    - Also, maintain "a fair and orderly market"
  - As compensation for taking on these obligations, DMMs get special perks/privileges
  - DMMs compete against other traders' limit orders, but DMMs still have official special status

De Facto Market Makers

- Pure order-driven systems (limit-order markets)
  - No officially designated market-makers
  - Traders can still follow an intermediation strategy
    - I.e., act as a market-maker
- No special privileges or obligations for such market makers
  - Same core concept: act as an counterparty/intermediary for traders who need instant execution
- Market makers sometimes described as "liquidity providers"
  - Much more useful description: "immediacy providers"

Immediacy

- Market makers supply the (valuable) service of allowing other traders to trade *right now*, rather than later
  - MMs intermediate between buyers and sellers who arrive asynchronously
  - Cf. Grossman & Miller "Liquidity and Market Structure" (JF 1988)
- Particularly clear in pure limit-order markets (no DMMs)
  - Traders choose between placing a marketable or a resting limit order
- Classic notion: MMs maintain a continuous market presence
  - Accomplish via resting limit orders
  - *De facto* MMs needn't truly be present in market all the time
    - "Continuous" relative to infinitely short-lived market orders

Mechanics of Market-Making

- Basic strategy
  - Place resting limit orders on both sides of the market
  - Buy low/sell high; repeat
- At what prices should you place your resting limit orders?
  - This question lies at the heart of successful market-making
- Are there any catches?
  - Inventory (next lecture)
  - Adverse selection (next four lectures after inventory)
  - Fixed costs (this lecture)
- "Getting caught is the mother of invention" (Hunter S. Thompson)

PS#1 Points to Note

- Framework developed in PS#1 lets us examine questions about immediacy in more concrete terms
- Point 1: All background traders are purely non-strategic
  - Submit limit orders with prices and directions drawn at random
  - Choice between placing an aggressive order vs. a passive order entirely random
- Important implications of Point 1:
  - No private information or possibility for adverse selection
  - Fairly steady, balanced inflow of marketable buy and sell orders

PS#1 Points to Note Contd.

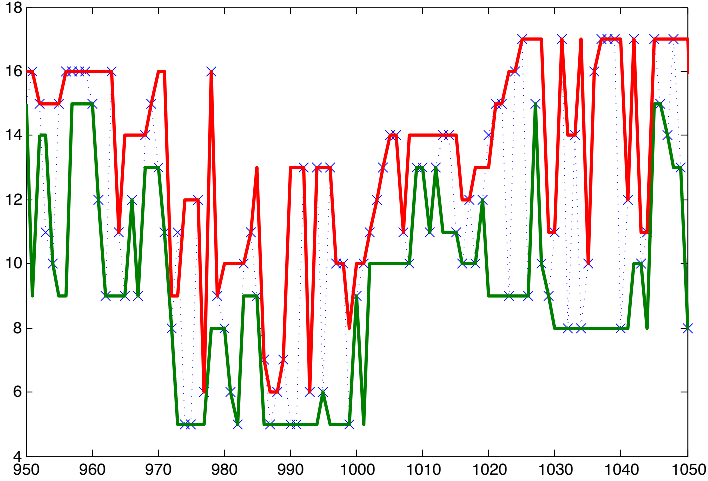
- Point 2: Background traders' order prices are drawn uniformly from a bounded interval
  - Mean-reverting prices
  - Even though inventory may not be stationary [hint], there is a natural fundamental value with only bounded, transient variation around it
  - Hence inventory risk isn't a pressing concern
- This setup basically isolates the valuable opportunity to provide immediacy
  - PS#2 will require some (among other things) analysis of elements related to sharp notions of immediacy
  - Costs and average time-to execution
    - More about this later in the lecture

Preliminaries (Revisiting Lecture 2)

Notions of Price

- Recall from last lecture ambiguities about the concepts of "price" and "liquidity"
- Price
  - Can compute theoretical expected long-run price: $\frac{minprice+maxprice}{2}$
    - Pretty close to simulation results (I got 10.61 last night)
  - How informative are the best bid, best ask, midpoint, last transaction, etc. ?

Best Bid, Best Ask, Last Transaction

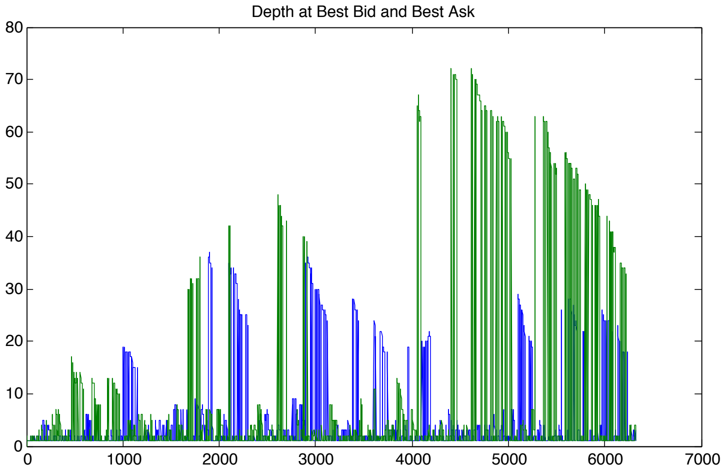


Bid-Ask Spread
At the times marketable orders are placed

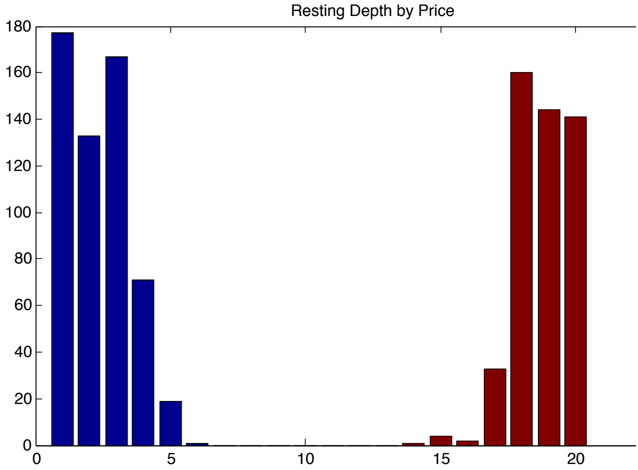
Preliminaries (Revisiting Lecture 2)
Notions of Liquidity

- Liquidity was the other ambiguous characteristic we discussed
- Tightness
  - Previous slide (kind of...)
- Resiliency
  - Since the long-run expected price is constant, all price-impact is transient
  - We'll use this simple setting to take a first serious look at resiliency after midterm #1 material
- Depth

Depth at the Best Bid and Best Ask



Why Such Wild Variation in Depth at the Best?



Market-Making in Our Simulated Market

- Simplest possible approach: only enter resting orders
  - Tends to work here, but only by a fluke
    - Buys (sells) above (below) fundamental value—even done passively—will tend to lose money
    - Modeling quirk: due to pile-up of resting depth at the highest and lowest prices, the median passive buy will be priced below fundamental value
- More realistic approach: only enter resting orders priced favorably to your expectation of fundamental value
  - Trivial in this setting, since FV is known and constant
  - Basic idea generalizes, but forecasting FV becomes the challenge

Market-Making in Our Simulated Market, Contd.

- Latter approach on previous slide is profitable, but not optimal
- Can do better than selecting a price at random
  - Want your passive orders to be profitable, but also want them them to execute in a timely manner
  - I conjecture that buys priced sufficiently low and sells priced sufficiently high in the simulation have an infinite expected time to execution
  - Remains a conjecture at this time. Fun project?

Market-Making in Our Simulated Market, Contd.

- Subject to being profitable in expectation, a passive order that executes is preferable to one that doesn't
- MMs want to be at or near the top of the queue
  - If you're too far from the top, you'll have to wait forever (perhaps literally) before your order executes
- Motivation for "Leap-Frogging"

Time-to-Execution in Real Markets

- Cost and time-to-execution of aggressive/marketable orders well understood
  - Cost: bid-ask spread (actually, a fraction of the spread)
  - Time-to-execution: almost zero
    - Only delayed by message travel-time, matching processing time, etc.

Time-to-Execution in Real Markets

- Relationship between cost/profit and time-to-execution for passive orders is less clear
  - E.g., suppose best bid = 10, best ask = 11, and you place a buy order at price p
    - Obvious tradeoffs with $p = 10$ vs. $p = 11$
    - But what about $p = 9$ vs. $p = 10$?
- Seriously, I don't think anybody has figured out a rigorous answer
- Analysis from pure-dealer markets doesn't carry over well to order-driven markets

Fixed Costs, Profits and Competition

Grossman-Miller (1988) Again

- Internal and external bounds to leapfrogging
  - External: someone else might get there first
  - Internal: need to remain far enough out to be profitable
- With continuous prices and competitive MMs, we would get Bertrand competition, and spreads would be competed down to the zero-profit (i.e., zero *trading* profit) level
- But being a market-maker has some fixed costs
  - Not necessarily huge, but these costs exist
  - With continuous prices and spreads competed down to the zero-profit level, MMs couldn't cover their fixed costs
  - Market failure (no MMs) in this case

Imposing Minimal Price Increments

- Exchanges don't use truly continuous prices
  - Prevents the market failure that occurs if market-making is unprofitable
- Exchanges use discrete prices, with minimal increments, "ticks"
- The zero-profit best bid and best ask will almost surely fall in the interior of a tick interval
  - Hence price granularity ensures a small profit for MMs on each transaction, to cover their fixed costs
- Overall equilibrium is reached as more MMs enter
  - As more MMs enter, each one participates in fewer transactions, until their profits just cover fixed costs
  - (Hypothetically)

A Place Where "Speed" *Might* Matter

- Entry by new MMs leads to a zero-average-profit equilibrium only if all the MMs are identical
  - Otherwise, equilibrium no-entry condition is zero marginal profit
- If you can react in less time than another MM, you might—*might*—get more transactions then her and hence earn positive profits in equilibrium
  - But time priority makes the issue non-obvious
  - Low-latency does not equate to time travel[1]
  - We'll return to related issues repeatedly

[1] See Harris pp. 180

Imposing Minimal Price Increments

- Discrete prices are more important than you expect, even if you condition your expectation on the information on this slide.

Richer Simulation Structure

- For future problem sets, starting with PS#2, there will be a potentially smart trader, "robot_1"
    - You will write algorithms to control robot_1 to accomplish various tasks
    - Some analysis of the outputs will also be part of the assignments
- The simulated environment will still include some background traders
- To examine adverse selection, we will look at several different specifications for background traders' behavior

Preventing Self-Crosses

- Adopt the convention that passive orders get cancelled if they would otherwise execute against an aggressive order from the same trader
- In reality, traders are responsible for preventing self-crosses themselves
  - Messy programming task
  - Not very conceptually difficult
- Avoiding self-crosses is *very important* in real markets for legal/regulatory reasons

What You Can Probably do This Week

- PS#2 isn't due until SEPTEMBER 14
  - But I'll post it Wednesday
  - I'll also post an updated template with a mechanism for "robot_1" to participate in the market against the background traders
- Some of the questions you'll be able to answer using the material covered today
  - Simple market-making algorithms
  - The rest will involve material from the next few lectures
    - Introduce inventory control into your MM algorithms
    - Adverse selection issues...
- The next few lectures will include the material necessary to do the rest of PS#2

Walking Through the Tough Parts of PS#1

- Hardest part of PS#1 is probably computing an individual trader's inventory and P&L paths
- Organizing a trader's transactions by passive vs. aggressive useful for many applications
- Two non-mutually exclusive things will help *ex-post*
  - Try to figure it out yourselves based on the solution code
  - I'll walk through the key ideas and machinery next lecture

Readings

- This lecture:
  - *O'Hara Ch. 1
  - Harris Ch. 8, Ch. 14
- Next Lecture: Inventory considerations
  - *O'Hara Ch. 2
  - Harris Ch. 13
  - PS#1 due; solutions posted
  - PS#2 posted (due 9/14)

FIN 580 Section MMT
Lecture 4

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

September 2, 2015

Administrative Matters

- PS#1 due
- PS#2 posted today, due 9/14

Outline

Richer Simulation Structure

- For future problem sets, starting with PS#2, there will be a potentially smart trader, “robot_1”
  - You will write algorithms to control robot_1 to accomplish various tasks
  - Some analysis of the outputs will also be part of the assignments
- The simulated environment will still include some background traders
- To examine adverse selection, we will look at several different specifications for background traders’ behavior (not on PS#2, though)

Preventing Self-Crosses

- Some students asked about this in the context of PS#1
  - Kind of irrelevant in PS#1, more important when not all trading-robots are identical and interchangable
- Adopt the convention that passive orders get cancelled if they would otherwise execute against an aggressive order from the same trader
- In reality, traders are responsible for preventing self-crosses themselves
  - Messy programming task
  - Not very conceptually difficult
- Avoiding self-crosses is *very important* in real markets for legal/regulatory reasons

Market-Makers
From Previous Lecture

- Market makers
  - Designated
    - Official status, special obligations and privileges
  - *De Facto*
    - Follow intermediation strategies, but not formally required or assisted to
    - Pure limit order markets
    - (Hybrids)

Market-Makers

From Previous Lecture (Contd.)

- Underlying source of MM revenue: supplying immediacy
  - MMs act as always-available counterparties
  - Accommodate urgent trading needs
  - Intermediate between buyers and sellers who need to trade instantly but arrive at different times
- 90% of market-making is showing up (kind of...)

Challenges for Market Makers

- Last time:
  - Place resting orders at prices far enough from expected fundamental value to earn profits
    - In particular, must cover your fixed costs (Cf. Lecture 3)
    - Exchange will typically set tick-size to facilitate this in equilibrium
  - Place resting orders close enough in so that they execute in a timely manner
- Two other central concerns for market makers
  - Avoid adverse selection (next four lectures)
  - Manage inventory risk (this lecture)

Inventory

- Even in the setting with uniform-price background traders, inventory (in shares) is not stationary for naïve market-maker strategies, including:
  - Just placing random passive orders
  - Placing passive orders at the best (random side)
  - Placing passive orders one tick better than the best (random side)
    - Or place at best, if an order priced one tick better would be marketable
- None of these have any attempt at inventory control

Illustration from Random Passive-Order Algo

- Run simulation for 23,322 steps (after burn-in of 3,322)...
  - For context, trading profit = 2741, trading volume = 828

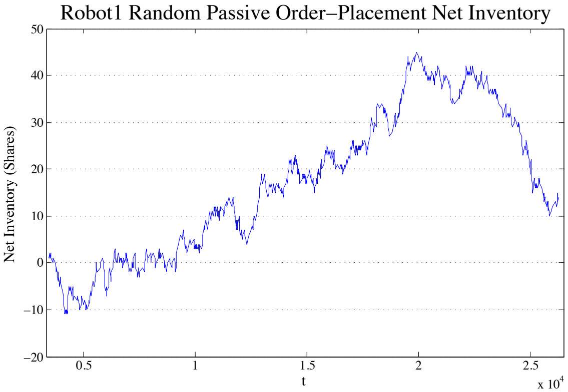


Illustration from Passive-Order at the Best Algo

- Run simulation for 23,322 steps (after burn-in of 3,322)...
  - For context, trading profit = 3594, trading volume = 1842

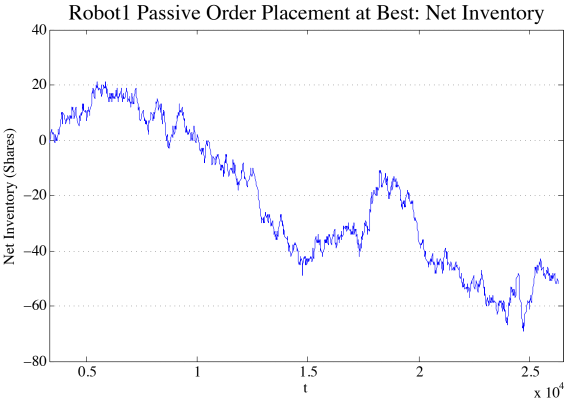
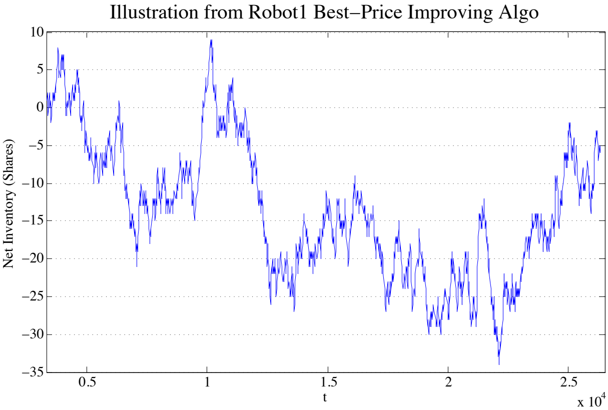


Illustration from Best-Price-Improving Best Algo

- Run simulation for 23,322 steps (after burn-in of 3,322)...
  - For context, trading profit = 2608, trading volume = 1927



Inventory Targets

- Target inventory level
  - MMs may desire a directional position for the same reasons as anyone else
    - Hedging, speculation, etc.
- For our current purposes, we can assume WLOG that inventory target is 0. (WLOG = "Without Loss Of Generality")
  - We're considering a single market (can ignore hedging)
  - Directional bets are conceptually separable from market-making (can ignore speculation)

Inventory Risk

- Adverse selection risk
  - Vital issue, but more related to *changes* in inventory than to *level* of inventory
- "Diversifiable" inventory risk
  - Risk from price movements uncorrelated with level of inventory
  - Familiar idiosyncratic risk of financial assets
    - Though feasibility of diversification varies...
  - Potentially market/systematic risk also (compensated)
- Concerns about solvency, "liquidity" (in the cash-on-hand sense) make excessive inventory risk unattractive
  - Possibly risk-aversion from MMs' preferences/utility functions, too

Theory

- Earliest days of microstructure: dealer inventory was a central object of interest (Cf. O'Hara Ch. 2)
  - Heavy emphasis on link between inventory and spreads
  - Models became *really* technical, still only modest empirical accuracy
  - Dealer markets became less pervasive
- Modern microstructure
  - Mostly consider inventory effects in extreme/specialized contexts
  - But these extreme contexts are extremely important!
    - Cf. Duffie's work
  - General observation: open questions in academic finance often translate to money on the table

Practice

- Avoiding meltdowns as important as ever
  - Internal risk-management in large financial firms
  - Algorithmic trading requires explicit inventory-management elements
- Beyond the remaining pure-dealer markets, character of inventory concerns differs from e.g. Harris' treatment
  - Customer competition in hybrid markets changes things for DMMs
  - Scrutiny of stub-quotes (following Flash Crash)
- Changes in nature of information available to non-DMM market participants
  - Anonymous trading becoming the norm
  - Rich orderflow, cross-security, cross-venue, etc. data widely available

Dealer vs. Hybrid vs. Pure Order

- Customer limit orders prime determinant of spread in most modern markets
  - Exclusive determinant in pure order-driven markets
- One-tick markets
  - Beautiful, and some extremely important ones
  - Individual MMs can't effectively manage inventory by adjusting their respective "quotes"
- Time-to-execution becomes an issue
  - Conceptually similar to pure-dealer case discussed by Harris
  - Technical considerations very different

Time-to-Execution

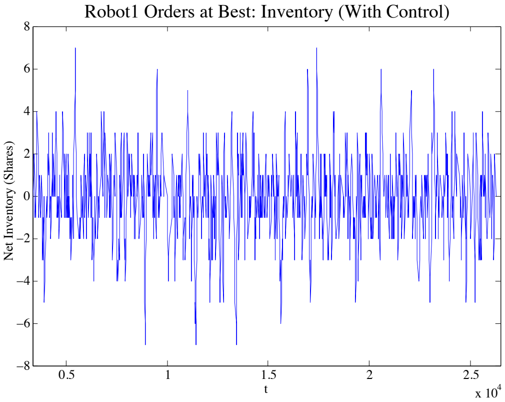
- If transaction prices follow a true random walk, even if nobody places additional resting orders in front of yours:
  - Probability that your resting order will execute eventually = 1
  - Expected time before your resting order executes = ∞
- Clearly a little unrealistic, but even a resting order you place at better than the best may not execute for a long, long time
  - Another task that requires frequent adjustments to your resting orders
  - And sometimes, you may use marketable orders
    - Raises the issue of price-forecasting...later in the term
- This is a hard problem

A Few Specific Inventory Twists for HFTs

- [In the E-Mini and assorted other futures markets, at least]
- Contract "roll" and settlement/expiration
- Special significance of inventory positions at the end of the trading day
  - Calculation of margin requirements
  - Directional positions above a certain size must be disclosed, along with some other trader data
- HFTs seem to change their behavior to unload inventory near contract expiration
  - Also try to end the trading day flat, or at least below the required reporting level
    - Presumably out of consideration to regulators, to save them unnecessary work ;-)

Inventory Control on MM "at Best" Algo

- Run simulation for 23,322 steps (after burn-in of 3,322)...
  - For context, trading profit = 3291, trading volume = 1800



Approaches to Inventory Control

- "Hard" (inventory constraints)
    - Switch to only making a one-sided market
        - I.e., cancel all your resting orders on one side of the book
    - Aggressively trade excess inventory
        - Dollar-value constraints vs. net quantity constraints
- "Soft" (inventory targets)
    - Quote less competitively on one side of the market
    - Only enter new orders on one side of the market
- Approaches are complementary (for MMs)

Aside: Misinformation about High-Frequency MMs

Harris pp. 293 (margin note):

> *"High-frequency dealing is a bit like picking up pennies in front of a steamroller. Sometimes you get in and out quickly, and profit a little. [YES]*
>
> *Sometimes you miss an opportunity or you pass because you have no safe opportunity. [MAYBE]*
>
> *However, if you are not very careful, you get caught and lose everything!" [NO!]*

Aside: Misinformation about High-Frequency MMs (Continued)

- To reiterate about that third statement, NO!
    - Provided you don't let your inventory drift off to infinity
- The "picking up pennies in front of a steamroller" similie was coined in reference to highly leveraged stat arb
    - Usually fixed-income arb
- High-frequency trading is very, *very* different
    - More like flipping pennies which you get when it lands heads, and lose when it ends tails
        - No steamroller
    - If you're smart, the coin is biased towards heads
    - And you flip it thousands of times per day, 250 days per year

Tracking Passive and Aggressive Trades

- Trading Profits, Inventory, Marking-to-Market (As in PS#1)
- Analyzing past market data (specifically, transactions)
  - Not just bookkeeping: evaluate strategies, search for persistent inefficiencies, etc.
  - Serious applications examined after midterm #1
- Two sides to each trade
  - Buyer and seller
  - Passor and aggressor (order-driven markets)
  - Different (and orthogonal) ways to split up transaction data
- Obvious dualities in either case
  - Lots of redundant data

Tracking Passive and Aggressive Trades (Ctd.)

- Initial question: did the transaction occur at the bid price or the ask price?
  - What is the appropriate price concept for us to use for the given transaction
- With balanced round-trips (zero net inventory), trading profits trivial to calculate
  - Otherwise, need to consider changes in inventory value
  - Marking-to-market
    - Not really a very satisfactory solution...
- Second topic: time-to-execution

Matlab Tips

- Colon (:) operator—specify entire column or row
  - E.g., bid_ask_stor_mat(17,:) = $[bestbid_{17}, bestask_{17}]$
  - bid_ask_stor_mat(:,1) =vector of best bids
- Delete (conformable) elements by setting equal to empty brackets [ ]
  - e.g., live_buy_orders_list(:,2) = [ ]

Matlab Tips

- "sort": B = sort(A, dim, mode)
  - Sort each column (dim=1) or row (dim=2) of matrix A
  - 'mode' specifies ascending/descending
  - [B, IX]=sort(A,...) also gives index in A of each element in B
- "sortrows": B = sortrows(A, [column #s])
  - MUCH MORE USEFUL for us than 'sort'
  - Sorts rows lexicographically based on specified columns
  - Each row is preserved

Matlab Tips (Logical Subscripting)

- E.g.
  robot_1_live_buy_indic=(live_buy_orders_list(:,1)==1)
- robot_1_live_buy_indic(k)
  $= \begin{cases} 1 & \textit{if } \text{live\_buy\_orders\_list}(k,1)\text{=1} \\ 0 & \textit{otherwise} \end{cases}$
- robot_1_live_buy_orders_list
  =live_buy_orders_list(robot_1_live_buy_indic,:)

Matlab Tips (Array Operations)

- Array ("element-wise") operations: .* .^ ./
- $\begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix} .* \begin{bmatrix} 7 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 28 \\ 0 \\ 2 \end{bmatrix}$
- Element-wise multiplication with logical subscript arrays
  - Just keep track of when you need to return to "logical" format

Readings

- This lecture:
  - *O'Hara Ch. 2
  - Harris Ch. 13
- Next Lecture: private information and adverse selection
  - Harris Ch. 9, Ch. 10

FIN 580 Section MMT
Lecture 5

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

September 9, 2015

Administrative Matters

- Reminder: PS#2 due Monday, 9/14

Outline

Market-Making: Recap of Previous Lectures

- Source of revenue
  - Immediacy provision
  - MMs accommodate urgent trading needs
- Costs/risks
  - Fixed costs (cf. competition, tick size, etc.)
  - Inventory risk
    - Diversifiable
    - Controllable

The Size of Spreads

- In a competitive environment, fixed costs only affect the bid-ask spread via tick-size
- (Diversifiable) Inventory risk?
  - Part of the story
  - Can manage diversifiable inventory risk through quoted quantity as well as price
    - Usual leap-frogging rationale for tight pricing
  - Take "diversifiable" seriously, consider competitive environment...
- *Adverse selection*

Private Information

"One of these days in your travels, a guy is going to show you a brand-new deck of cards on which the seal is not yet broken. Then this guy is going to offer to bet you that he can make the jack of spades jump out of this brand-new deck of cards and squirt cider in your ear. But, son, do not accept this bet, because as sure as you stand there, you're going to wind up with an ear full of cider."

–Sky Masterson (Marlon Brando, Guys and Dolls)

Private Information

- Placing a market order represents a choice by the aggressor
  - Axiomatically increases the aggressor's subjective expected utility
  - Implications for the passor?
- Two possibilities
  - Positive (private)-sum
    - E.g., urgent demand to hedge
  - Zero (private)-sum
    - Information-based trading

Private Information

- Minimal demand to hedge jack-of-spades cider-loss risk
  - Jack-of-spades bet is purely informational trading
    - Zero (private)-sum
- Notice that "accepting bet" = "acting as aggressor"
  - Can choose not to accept bet
  - (Should only accept the bet if it's your best way to buy cider)
- But suppose you *offer to bet* that nobody can do the jack-of-spades trick
  - You're acting as the passor (MM)
  - Only that one guy will accept your bet

Adverse Selection

- Divide aggressors into "informed" and "uninformed"
  - Interpret as relative to MM (binary classification in absolute terms is unrealistic)
- Informed traders trade in the same direction as expected future price movements
  - I.e., buy before price increases, sell before price decreases
  - Informed traders' counterparties trade in the opposite direction
- The times when informed traders want to trade are the worst times to trade with them
  - "Adverse selection"

Adverse-Selection Risk

- Recall that holding inventory is risky because inventory value depends on the current price
  - More rigorously, inventory value depends on current expectation of future transaction price
  - Losses arise when price-change and net inventory have opposite signs
    - "Adverse" price movements

Adverse-Selection Risk

- When considering diversifiable inventory risk, we assumed MM's inventory-changes uncorrelated with future price-changes
  - Backward-looking inventory management was sufficient
- Need forward-looking mechanisms to manage adverse-selection risk

Immediacy as an Option

- Providing immediacy is equivalent to providing an (American-style) option to trade
  - Aggressors only "excercise" this option when they find doing so to be advantageous
- How do MMs price this option?
  - MMs don't explicitly "sell" the option (they offer it for free)
  - MMs must therefore design the option to satisfy this price constraint
    - I.e., MMs must choose their quotes accordingly

Market-Making with Adverse Selection

- What is the optimal price for MM to place passive orders and/or quotes?
- Simplifying assumption/fact: WLOG, we can focus on the MM-zero-profit prices
  - Round MM-zero-profit prices to the outside tick
  - Competition among MMs, along with price-priority, will drive MM pricing to this

Market-Making with Adverse Selection

- What is the zero-profit price for MM to place passive orders and/or quotes?
  - Expected future price, *conditional upon MM's order executing*
  - E.g., $\mathbb{E}[FV|\textit{aggressive buy}]$

Informativeness of Aggressive Orders

- Suppose that all aggressors are informed
  - An informed aggressor will only buy if his expectation of "fundamental value" exceeds the best ask
    - Symmetric situation for informed selling
  - If an informed aggressor buys, then uninformed trader's expectation is

$$\mathbb{E}[FV|buy] = \mathbb{E}[FV|FV \geq best\ ask] \geq best\ ask$$

Informativeness of Aggressive Orders

- No trading occurs if $\mathbb{E}[FV|buy] > best\ ask$...generic case if all aggressors informed
  - Holds regardless of how the best ask is set
- If only some aggressors are informed, still have

$$\mathbb{E}[FV|buy] > \mathbb{E}[FV]$$

  - I.e., on average, aggressive orders partially informative
- But it is not necessarily true that

$$\mathbb{E}[FV|buy] \geq best\ ask$$

Computing Conditional Expectations

- Easy to find inequalities that the conditional expectations like $\mathbb{E}[FV|buy]$ satisfy
  - No parametric assumptions used in previous slides
- Computing the actual numerical value of $\mathbb{E}[FV|buy]$ is challenging
  - Requires some additional structure, such as parametric assumptions
  - Can also impose structure by treating past data as representative of various future distributions
- Next two lectures: simple framework to model and analyze these conditional expectations

Connections to PS#2

- Expected FV conditional upon aggressive order direction—not covered
- Foundations for understanding adverse selection
  - Execution of mispriced passive orders
  - Execution of correctly priced passive orders
- We will start to look at how to determine a "correct price"
  - Important to understand how (in)correctly priced orders behave

Anonymity

- So far, we have assumed that orders are anonymous
- In later lectures, we will consider statistical methods MMs can use to infer informedness
- The inference task is very different if MMs know their counterparties
    - But nobody wants to seem informed *ex ante*
    - Credible signals, reputation, etc. come into play
- Can use brokers to trade anonymously, even if market is not anonymous

Informed Trading and Price-Discovery

- Informed trading has positive social value
  - Side-effect of informed trading: price discovery/ informative prices
  - If you're uninformed, you don't want to trade with an informed counterparty, but you do want informative prices
- Informed trading $\neq$ insider trading
  - Insider trading $\subset$ informed trading
    - *Strict* subset!

Readings

- This lecture:
  - Harris Ch. 9, Ch. 10
  - Harris Ch. 13
- Next Lecture: adverse selection continued
  - Hasbrouck Ch. 12

FIN 580 Section MMT
Lecture 6

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

September 14, 2015

Administrative Matters

- PS#2 due today, solutions posted this afternoon
- PS#3 posted this afternoon, due 9/21
- Grading distribution, class size, etc.

Outline

"Public Priority"
Cf. Hasbrouck Ch. 6

- Pure dealer market
- Pure limit-order market

Hybrid market (1) .... Hybrid market (2)

- Customer limit orders compete with dealer for "market making" business
- Conflict of interests if dealer also processes all orders (including customer limits)

"Public Priority"

Looking out for the little guy

Market Characteristics

	# Trades Post-Burn-In	Avg. Spread Post-Burn-In
Uniform Passive	3732	4.85
Four Worse	3631	4.82
One Worse	3704	4.44
Best	3766	4.47
One Better	3830	4.25
Four Better	3824	4.06

Profits, Volume, Inventory

Robot_1

	Profit	Volume	Max Inventory Magnitude
Uniform Passive	919	515	27
Four Worse	1134	312	41
One Worse	1546	667	36
Best	1417	771	31
One Better	1000	805	36
Four Better	-300	800	23

Time to Execution

Robot_1

	Mean	Median	Std. Dev	% Orders Executed
Uniform Passive	110	8	297	62%
Four Worse	253	49	729	38%
One Worse	**524**	31	**1183**	79%
Best	95	16	325	93%
One Better	13	4	56	99%
Four Better	12	3	66	99%

Improvements on Naive MM

- Inventory control
  - Approx. same MM profits, reduced max inventory by $\approx 500\%$
- Fundamental value (baby "adverse selection")
  - Increased MM profits by $\approx 80 - 100\%$
- In the words of real M.M. (2013):
  - *"Why be a king...when you can be a god?"*

Private Information and Adverse Selection

Last Lecture

- Market-makers grant potential aggressors a free option
  - Aggressors get to choose when—and *if*—to trade
  - Aggressors only "excercise" this option when they find trading to be advantageous
- Uninformed aggressors
  - Gains from immediacy do not come at the MM's expense
- Informed aggressors
  - Trade at precisely the times it will prove costly for their counterparties
  - Adverse selection

MM Response to Adverse Selection

Last Lecture

- What is the optimal price for MM to place passive orders and/or quotes?
  - Simplifying factors: tick size, competition
- Expected future price, *conditional upon MM's order executing*
- $$\mathbb{E}\left[FV|aggressive\ buy\right] > \mathbb{E}\left[FV\right] > \mathbb{E}\left[FV|aggressive\ sell\right]$$
- These conditional expectations can be hard to compute

Glosten-Milgrom Model

- Glosten-Milgrom model
  - Elegant parametric model of MM determining optimal (zero-profit) spreads
  - Can explicitly calculate expected future price, *conditional upon MM's order executing*
    - Also, can show that this is an equilibrium outcome
- Important model to know at a general level
  - I won't actually work through the model
  - Our analysis/design of MM algorithms will be conceptually equivalent to solving a G-M model

Recall: Uniform-Price Background Traders

- Expected long-run average price, "fundamental value," was constant
- $$\mathbb{E}[FV|aggressive\ buy] = \mathbb{E}[FV] = \mathbb{E}[FV|aggressive\ sell]$$
- MM will tend to profit from placing passive orders near the best prices here
  - MM loses money on passive orders mispriced relative to FV
- On average, few orders entered near the best prices will be mispriced in this context
  - Many passive orders that eventually execute were not mispriced

Pure Random-Walk Prices

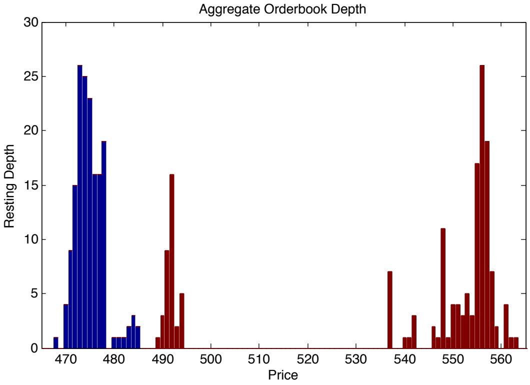
- The uniform-price background-trader setup had FV remain the same for every period
  - $$FV := \frac{maxprice + minprice}{2}$$
  - Can think of prices being drawn uniformly from $[FV - priceflex, FV + priceflex]$, treat interval midpoint as FV
  - Where "price_flex" is a parameter
- Opposite extreme: FV changes every period
  - Draw prices uniformly from $[FV_t - priceflex, FV_t + priceflex]$
  - Assume $FV_t := p_{t-1}$

Pure Random-Walk Prices

- Suppose background traders place orders of randomly chosen direction at prices $p_t \sim U([p_{t-1} - priceflex, p_{t-1} + priceflex])$
  - Set price_flex = 1, $p_0 = 500$
  - t_max=6322, burn_in_period=1322
  - max_quantity=1

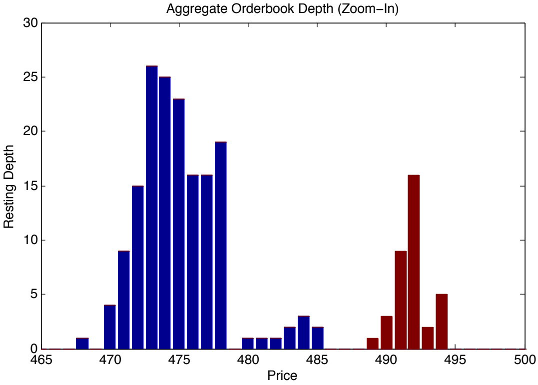
Pure Random-Walk Prices

Aggregate Resting Depth

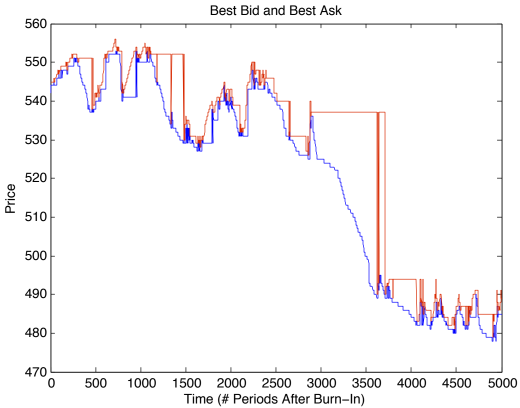


Pure Random-Walk Prices

Aggregate Resting Depth



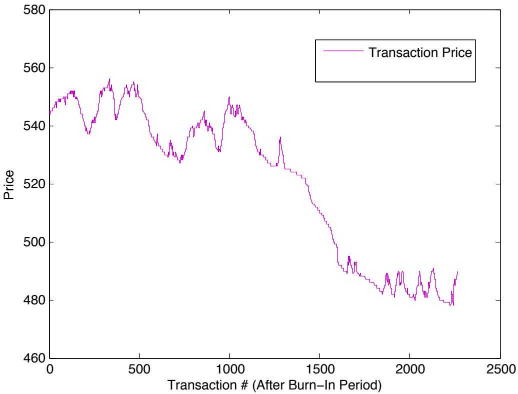
Pure Random-Walk Prices
Best Bids/Asks



Market-Maker Performance

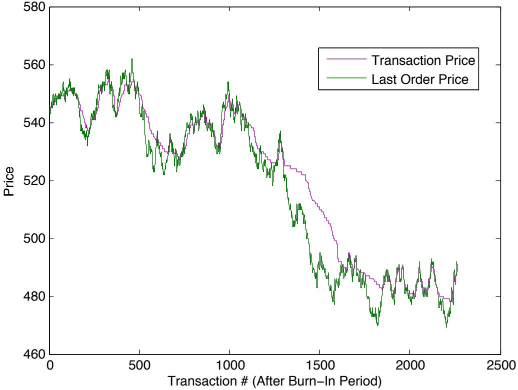
- Consider a market-maker who only places passive orders
  - Assume that the prices of the MM's orders do not reset the distribution from which background traders' prices are drawn
- In the long-run, the MM *will always lose money!*
  - Placing orders at the best, better than the best, worse than the best, etc. doesn't change the result
- What's going on?

Pure Random-Walk Prices
Transaction Prices



Pure Random-Walk Prices

Transaction Prices and Last_Order Prices



Adverse Selection

- Under what circumstances will a resting buy order at price b execute?
- Necessary condition: $p_t \leq b$
  - If $p_t = b$, then purchase price equals expected future price
  - If $p_t < b$, then purchase price *exceeds* expected future price
- Hence

$$\mathbb{E}\left[FV \middle| passive\ buy\ executes\ at\ b\right] < b$$

- Symmetric situation for passive sell orders

Middle Ground

- In the pure-random-walk-price setting, every aggressive order is effectively "informed"
  - This should lead to a no-trade equilibrium outcome
  - Cf. last lecture
- More realistic scenario: only some fraction of orders are informed
  - Contrast with uniform-price setting, where *no* orders are informed

Middle Ground

- Suppose background traders place orders of randomly chosen direction at prices

$$p_t \sim U\left(\left[lastorderprice_{t-1} - priceflex, lastorderprice_{t-1} + priceflex\right]\right)$$

- Assume that

$$lastorderprice_t = \begin{cases} p_t & \text{with probability } \theta \\ lastorderprice_{t-1} & \text{with probability } 1-\theta \end{cases}$$

  - Uniform prices: $\theta = 0$
  - Pure random-walk prices: $\theta = 1$

Closer Look At BGT Behavior

- Order sign
  - buy_sell_robot_j=randi(2);
  - buy_sell_robot_j=2*(buy_sell_robot_j-1.5);
- Order quantity
  - quantity_robot_j=randi(max_quantity);
- Order status
  - alive_indicator_robot_j=1;

Closer Look at BGT Behavior

- Order price possibilities
  - lowest_choice=max((last_order_price-price_flex),min_price);
  - highest_choice=min((last_order_price+price_flex),max_price);
- Order price selection
  - price_robot_j=randi([lowest_choice,highest_choice]);

Closer Look at BGT Behavior

Built into Simulation Template

```
prob_last_order_price_resets=θ;

test_for_last_order_price_choice_j=rand(1);

if test_for_last_order_price_choice_j<prob_last_order_price_resets
        last_order_price=price_robot_j;
end
```

What We Can't Model Yet

- Relationship between the size of the bid-ask spread and adverse selection risk
- The price that's moving around isn't the mid-point price in the usual sense
- (We may revisit this later in the term)

What We Can't Model Yet

- Size of spread doesn't (strongly) affect MM profitability in the current partial-random-walk setting
- Difference between the price of the MM's order and FV doesn't depend on the size of the spread

What We Can't Model Yet (contd.)

- We are using the same distribution to generate both passive orders and aggressive orders
  - If spread is large, there won't be as many aggressive orders
  - Passive orders that reduce the spread become more likely
- Instructive, and possibly representative for some markets
  - Not universally applicable
  - Usually, expect to see large distinctions between aggressive orders and price-improving passive orders
    - Formal theoretical result, also holds up empirically

What We *Can* Model

- Given spread behavior and price process, is it profitable to be a MM?
  - Or is there too much informed trading?
- Depends on MM strategy and price process
  - MM strategy: can estimate cutoff value for θ necessary for MM profitability
  - Can estimate the underlying value of θ, given historical (simulated) data
- On PS#3, covered in next lecture

Aside: "Econophysics"

- Using the same distn to generate both passive and aggressive orders:
  - Econophysicists *love* this
  - Think of it as the microstructure equivalent of string theory...
- Fun fact #1 : my uncle, Rod Clark, independently coined the term "econophysics" in one of his science-fiction stories
  - Turns out that it had been used in some article 4 years earlier
  - But that article wasn't (intended to be) science-fiction
- Fun fact #2: Back in 1999, Rod wrote a sci-fi story in which he imagined/described what we now call "high-frequency trading"
  - Titled "Redshift: Greenstreem". Check it out!

Readings

- This lecture:
  - Hasbrouck Ch. 12
- Next Lecture: adverse selection continued
  - No new readings

FIN 580 Section MMT
Lecture 7

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

September 16/21 2015

Administrative Matters

- PS#3 due WEDNESDAY, 9/23
- Midterm #1 on October 5 (maybe later; working on logistics)

Outline

MM Overview

- Compensation for supplying immediacy
- Competition
  - Zero-(marginal)-profit spreads
  - Tick-size and fixed costs
- Inventory: risks and management
- Adverse selection
  - Trading against informed aggressors
- At a conceptual level, that's most of market-making

What Now?

- Institutional/organizational details
  - E-mini: ask me
  - Everything else: ask Hasbrouck, Harris, Werner, Angel, Glosten, Spatt, Biais, ...
  - Not the focus of this course (or of modern microstructure!)
- Applying theory to develop more sophisticated MM algorithms
  - Still in a simplified, simulated setting
  - But starting to contain all the ingredients of real-world algos

Turning Toys into Tools of the Trade

- Naive MM algos from PS#2
  - Make money when it fall into their laps (uniform-price model)
  - Lose money when they face more than a small amount of adverse selection
- Almost all the orders that every MM places are passive orders near the best bid and best ask
  - Major difference between serious MMs and our toy robots: knowing when to submit those orders and when to cancel them
  - I.e., supply immediacy only when adv. selection risk small

Better Analytic Machinery

Statistical Inference About Algo Performance

- To refine our MM algorithms, need better techniques to analyze their performance
- So far, we've only looked at point estimates of performance stats
  - Need to quantify the precision of those point estimates to compare them
  - Main method: bootstrapping
- Refresher on bootstrapping later in this lecture

Better Analytic Machinery
Choice-Parameter Optimization

- Need techniques to find optimal choice-parameter values
  - Leading example: finding the maximum level of adverse selection at which MM algo still profitable
- Grid search (this is not a CS course)
  - Usually a last resort!
- Economic theory to simplify/structure the region of parameter space to search
  - Bisection, if not something even better
  - Refresher on bisection later in this lecture

Better Analytic Machinery

Statistical Inference About Market Parameters

- Estimate key market parameters from past data
  - For the moment, we will focus on estimating adverse selection risk
- Foundation for expanding the set of conditioning variables used by our algorithms
  - Build a richer state-space than the current {best bid, best ask, robot1 inventory}
  - Choose statistical summaries to keep state-space manageable
- This lecture: estimating the parameter prob_last_order_price_resets $:= \theta$

Parametric Price Process

- Recall that background traders now place orders of randomly chosen direction at prices

$$p_t \sim U\left([lastorderprice_{t-1} - priceflex, lastorderprice_{t-1} + priceflex]\right)$$

- Price process is characterized by *priceflex* and θ:

$$lastorderprice_t = \begin{cases} p_t & with\ probability\ \theta \\ lastorderprice_{t-1} & with\ probability\ 1-\theta \end{cases}$$

Price Process Parameters

- Meaningful economic interpretation for θ
  - Fraction of informed trading (loosely)
- Realistic case: $0 < \theta < 1$
  - Uniform prices: $\theta = 0$
  - Pure random-walk prices: $\theta = 1$

Price Process Parameters

- "*priceflex*" is basically a scale parameter
  - Characterizes the probability distribution of price changes at short horizons
- Bounded support for the price-change distribution is convenient, but not necessary
- Trivial to perfectly estimate *priceflex* in the current model
  - I'll just take *priceflex* to be known
  - More generally, expect to estimate analogues of θ and *priceflex* jointly

Estimating θ
Strategy

- Suppose we observed the sequences $\{lastorderprice_t\}_{t=1}^{t\_max}$ and $\{p_t\}_{t=1}^{t\_max}$
  - Construct the sequences of first-differences

$$\{lastorderprice_t - lastorderprice_{t-1}\}_{t=2}^{t\_max} := \{\Delta lop_t\}_{t=2}^{t\_max}$$

$$\{p_t - p_{t-1}\}_{t=2}^{t\_max} := \{\Delta p_t\}_{t=2}^{t\_max}$$

  - Estimate of θ is

$$\tilde{\theta} = \frac{\sum \mathbb{I}\{\Delta lop_t \neq 0\}}{\sum \mathbb{I}\{\Delta p_t \neq 0\}}$$

  - I.e., fraction of non-zero changes in p that are accompanied by a non-zero change in *lastorderprice*
- But the whole point is that we don't see *lastorderprice*

Estimating θ

- Assume that we *only* observed the sequence $\{p_t\}_{t=1}^{t\_max}$
  - Construct the sequence of first-differences

$$\{p_t - p_{t-1}\}_{t=2}^{t\_max} := \{\Delta p_t\}_{t=2}^{t\_max}$$

  - Can't estimate θ by the fraction of non-zero changes in p_t
  - But we can do something similar
- Look at certain big changes in p_t...

Estimating θ

- To lighten notation, define $c := \textit{priceflex}$
- $\max(\Delta p_t) = 2c$, only attained when
  - $p_{t-1} = \textit{lastorderprice}_{t-2} - c$, and
  - $p_t = \textit{lastorderprice}_{t-1} + c$, and
  - $\textit{lastorderprice}_{t-1} = \textit{lastorderprice}_{t-2}$
  - [Draw a picture to convince yourself]
- If $\Delta p_t = 2c$, then $(\Delta p_{t+1} = c) \implies (\textit{lastorderprice}_t = p_t)$
- Thus

$$\left(\Delta p_t * \Delta p_{t+1} = 2c^2\right) \implies (\Delta \textit{lop}_t \neq 0)$$

Estimating θ

- Can show that

$$\mathbb{P}\left\{\Delta p_t * \Delta p_{t+1} = 2c^2\right\} = \frac{\theta}{2c+1}\mathbb{P}\{\Delta p_t = 2c\}$$

- Hence

$$\implies \theta = (2c+1)\frac{\mathbb{P}\left\{\Delta p_t * \Delta p_{t+1} = 2c^2\right\}}{\mathbb{P}\{\Delta p_t = 2c\}}$$

- Replacing probabilities with sample proportions gives

$$\hat{\theta} = (2c+1)\frac{\sum \mathbb{I}\left\{\Delta p_t * \Delta p_{t+1} = 2c^2\right\}}{\sum \mathbb{I}\{\Delta p_t = 2c\}}$$

Derivation of $\hat{\theta}$
(Missing Details)

$$
\begin{aligned}
&\mathbb{P}\{\Delta p_{t+1} = c | \Delta p_t = 2c\} \\
=&\mathbb{P}\{\Delta p_{t+1} = c | \Delta p_t = 2c \text{ and } \Delta lop_t \neq 0\} \mathbb{P}\{\Delta lop_t \neq 0 | \Delta p_t = 2c\} \\
&+ \mathbb{P}\{\Delta p_{t+1} = c | \Delta p_t = 2c \text{ and } \Delta lop_t = 0\} \mathbb{P}\{\Delta lop_t = 0 | \Delta p_t = 2c\} \\
=&\mathbb{P}\{\Delta p_{t+1} = c | \Delta p_t = 2c \text{ and } \Delta lop_t \neq 0\} * \theta + 0 * (1 - \theta) \\
=&\theta \mathbb{P}\{\Delta p_{t+1} = c | \Delta p_t = 2c \text{ and } \Delta lop_t \neq 0\} \\
=&\theta \mathbb{P}\{\Delta p_{t+1} = c | \Delta lop_t \neq 0\} \\
=&\frac{\theta}{2c+1}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}\left\{\Delta p_t * \Delta p_{t+1} = 2c^2\right\} =& \mathbb{P}\{\Delta p_{t+1} = c | \Delta p_t = 2c\} \mathbb{P}\{\Delta p_t = 2c\} \\
=& \frac{\theta}{2c+1} \mathbb{P}\{\Delta p_t = 2c\}
\end{aligned}
$$

Estimating θ by $\hat{\theta}$

- The estimator $\hat{\theta}$ is consistent, but it's inefficient
  - Cf. method-of-moments vs. maximum likelihood
- We're throwing away lots of information by considering only the definite instances of a permanent price change
  - Could construct a better estimator for θ by using all observations
  - Lots of tedious algebra...

Estimating a Real-World Analogue of θ

- Our basic estimation strategy extends easily to real-world settings
- Intuition: one very large transient price movement is much less probable than a smaller permanent one together with a smaller transient one
- E.g., suppose $X_1, X_2 \sim i.i.d.\ N(0,1)$ and $X_1 + X_2 > 6$
  - $\mathbb{P}\{X_1 > 3 \textit{ and } X_2 > 3\} \approx 10,000,000 \times \mathbb{P}\{X_1 > 6 \textit{ or } X_2 > 6\}$

Two Black Swans

Not Really So Rare

One Black Dodo Bird

Extremely Rare

Estimating a Real-World Analogue of θ

- We are exploiting the tail properties of transient price movements at short horizons
  - Rigorously, exploiting tails that decay faster than exponentially ("super-exponential")
  - Do not need *all* price movements to have super-exponential tails
- If the super-exponential tail condition isn't met, can adapt the estimation to look at non-tail events
- Or we can likely just ignore the heavy-tail problem (seriously)
  - We don't directly care about the sizes of the large price movements, just how often they occur
  - Very different from risk-management

What to do with $\hat{\theta}$

- Given spread behavior and price process, determine if it is profitable to be a MM
  - Depends on MM strategy and price process
  - MM strategy: can estimate cutoff value for θ necessary for MM profitability
- Can now estimate the underlying value of θ from lagged data
- Many possibilities from here
  - E.g. write MM algo that only enters market when it's expected to be profitable

Bootstrapping

- Data: $\{x_t\}_{t=1}^{T}$ assumed *i.i.d.*
- Want to estimate some statistic about x_t
  - Consider $\mathbb{E}[x_t]$ for concreteness
  - (Nonlinear statistics are more relevant in practice)
- Suppose that the distribution of x_t is completely unknown
  - Point estimates are still easy to obtain
  - Confidence intervals, quantiles can be hard

Bootstrapping

- Central Limit Theorem and related asymptotic results can have poor finite-sample performance
  - Small-sample distortions depend on the true distribution of x_t
- Main steps:
  - 1) Use the empirical distribution of x_t to approximate the true distribution
  - 2) Draw new samples from the empirical distribution of x_t and compute point estimates
  - 3) Use empirical distribution of point estimates to approximate their true distribution
- Can get extremely good approximations of the point estimates' distribution

Bisection

- Motivating example:
  - I'm thinking of a number, y, between 0 and 1
  - Every time someone guesses, I'll say if y is lower or higher
  - 7 guesses, mystery prize for a guess within 0.01 of y

Bisection

- Bisection solution:
  - First guess: $g_1 = \frac{1}{2}$
    - Define $side_1 := \begin{cases} -1 & if \;\; y \;\; lower \\ +1 & if \;\; y \;\; higher \end{cases}$
  - Second guess:$g_2 = g_1 + side_1 * \frac{1}{4}$
  - N^{th} guess: $g_N = g_{N-1} + side_{N-1} * 2^{-N}$
- Probabilistic extension...

Looking Forward

- Midterm #1 still only covers lecture material up through 9/23 and through PS#4

FIN 580 Section MMT
Lecture 8

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

September 21/23, 2015

Administrative Matters

- PS#3 due
- PS#4 posted today or tomorrow, due Friday 10/2 (e-mail your solutions to me)
- No class on 9/30
- Midterm #1 on October 5

Outline

Finally

- Key features and basic operation of financial exchanges
  - Primary focus: limit-order markets
- Market-making principles
  - Source of revenue
  - Competition
  - Inventory management
  - Adverse selection
- Now we can start taking the "algorithmic" part seriously

Disclaimer and Stern Warning

- Don't use our stylized algorithms to do real trading in real markets
  - Seriously. Don't do it.
  - This material is for educational purposes only
- The new mechanisms could *hypothetically* generalize to real settings
- Don't use our stylized algorithms to do real trading in real markets

Take-Away Points From PS#3

- Naive MM algos get slaughtered by adverse selection
- Permanent price innovations have unexpected effects on mechanisms that were well-behaved in the uniform-price setting
  - Inventory (both dollar value and number of shares)
  - Quoting at the best vs. better

Take-Away Points From PS#3

- We have a convenient way to parameterize adverse selection risk
  - *"prob_last_order_price_resets"*
  - Much less stylized than it might seem (probability of informed trading...)
  - Permanent and transient price impact from the same order—coming soon
- This representation opens the door to tuning choice parameters, estimation, and more

Take-Away Points From PS#3

- Tuning/optimizing choice parameters
  - Conceptually simple (bootstrap, bisection, backtesting,...)
  - Slippery and temperamental in practice
    - Non-stationarity again
- Importance of theoretical underpinnings for choices
  - Simplify optimization tasks
  - Add robustness
- Uncertainty/precision potentially hard to properly quantify

Next Steps

Statistical Inference About Market Parameters

- Estimate key market parameters from past data
  - Focus on estimating adverse selection risk
  - I.e., prob_last_order_price_resets:= θ)
- Considered one estimation approach last lecture
  - This lecture: start generalizing
  - Consider new issues related to estimating market parameters
- Discuss applications and main directions for extensions

Parametric Price Process
Recall

- Recall that background traders now place orders of randomly chosen direction at prices

$$p_t \sim U\left(\left[lastorderprice_{t-1} - priceflex, lastorderprice_{t-1} + priceflex\right]\right)$$

- Price process is characterized by *priceflex* and θ:

$$lastorderprice_t = \begin{cases} p_t & \textit{with probability } \theta \\ lastorderprice_{t-1} & \textit{with probability } 1-\theta \end{cases}$$

Estimating θ
Strategy from Last Lecture

- Suppose we observed the sequences $\{lastorderprice_t\}_{t=1}^{t\_max}$ and $\{p_t\}_{t=1}^{t\_max}$
  - Construct the sequences of first-differences

$$\{lastorderprice_t - lastorderprice_{t-1}\}_{t=2}^{t\_max} := \{\Delta lop_t\}_{t=2}^{t\_max}$$

$$\{p_t - p_{t-1}\}_{t=2}^{t\_max} := \{\Delta p_t\}_{t=2}^{t\_max}$$

  - Estimate of θ is

$$\tilde{\theta} = \frac{\sum \mathbb{I}\{\Delta lop_t \neq 0\}}{\sum \mathbb{I}\{\Delta p_t \neq 0\}}$$

  - I.e., fraction of non-zero changes in p that are accompanied by a non-zero change in *lastorderprice*
- But the whole point is that we don't see *lastorderprice*

Estimating θ

Strategy from Last Lecture

- Assume that we observed the sequence $\{p_t\}_{t=1}^{t\_max}$
  - Construct the sequence of first-differences

$$\{p_t - p_{t-1}\}_{t=2}^{t\_max} := \{\Delta p_t\}_{t=2}^{t\_max}$$

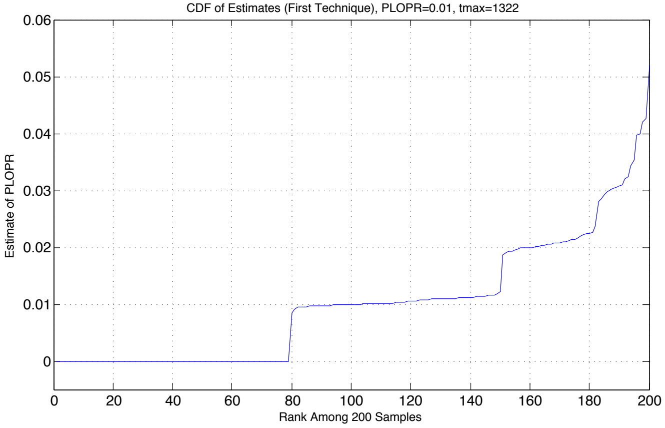
  - To lighten notation, define $c := \mathit{priceflex}$

Estimating θ

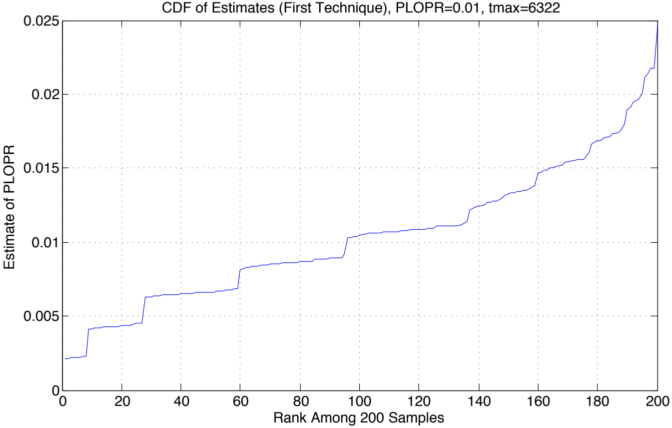
Strategy from Last Lecture

- $\max(\Delta p_t) = 2c$, only attained when
  - $p_{t-1} = lastorderprice_{t-2} - c$,
  - $p_t = lastorderprice_{t-1} + c$, and
  - $lastorderprice_{t-1} = lastorderprice_{t-2}$
- If $\Delta p_t = 2c$, then $(\Delta p_{t+1} = c) \implies (lastorderprice_t = p_t)$
- Counting, algebra give a simple estimator for θ

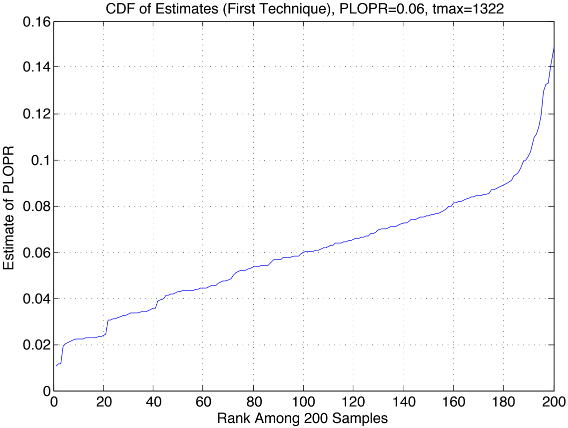
θ Point-Estimate Quantiles (t_max=1322)



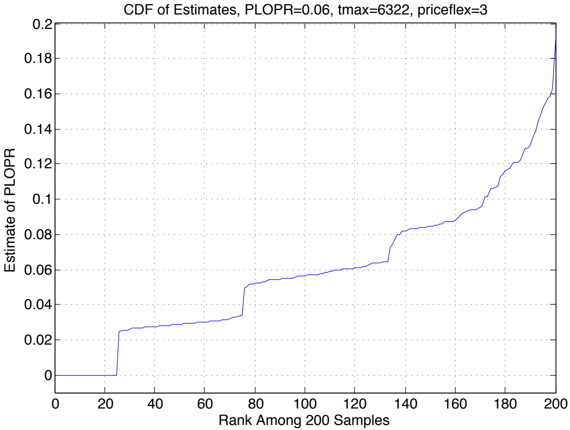
θ Point-Estimate Quantiles (t_max=6322)



θ Point-Estimate Quantiles (PLOPR=0.06)



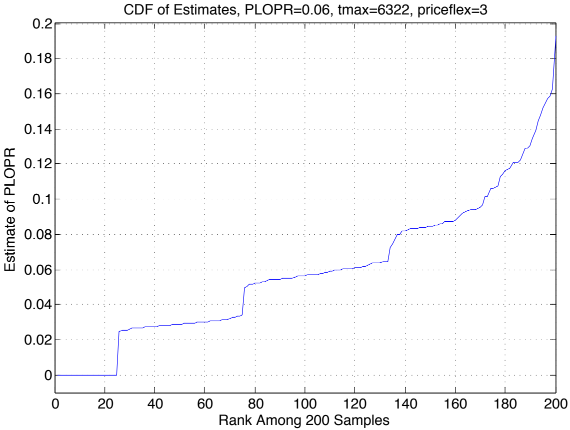
Quantiles (t_max and PLOPR large, but priceflex=3)



Estimating θ Better

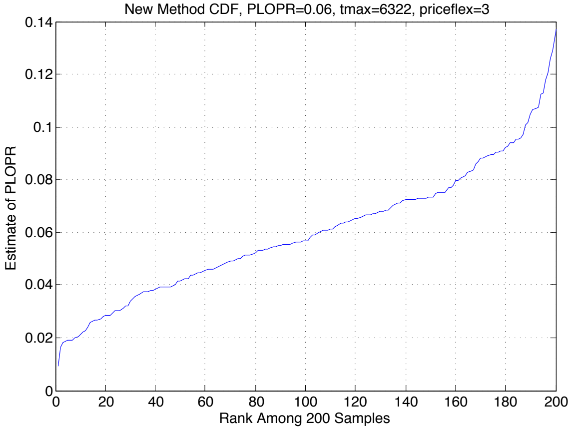
- We are only looking at one specific necessary condition for a permanent price-change to have occurred
- I.e., a maximum price change immediately following, and in the same direction as, another maximum price change
  - When *price_flex*=1, this wasn't restrictive
  - When *price_flex*>1, it is restrictive
- Size of the price-change following a maximum price-change doesn't matter; only direction matters
  - This gives us more observations with which to work

θ Point-Estimate Quantiles

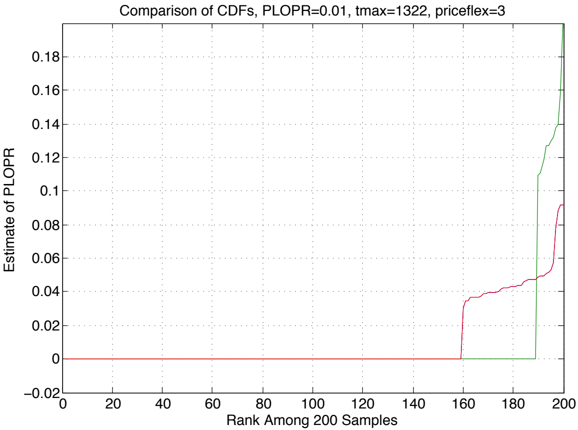


(Same picture as before, just included here for reference)

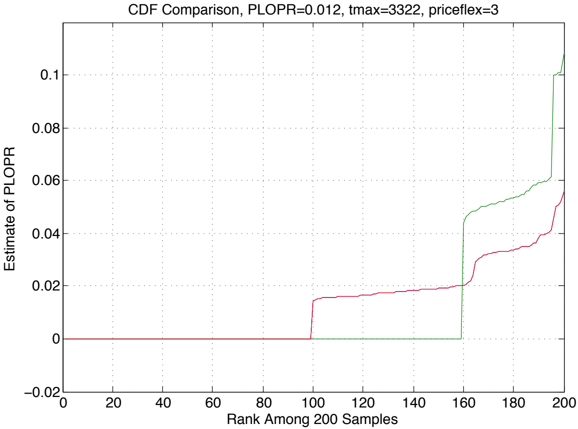
θ Point-Estimate Quantiles



θ Point-Estimate Quantiles



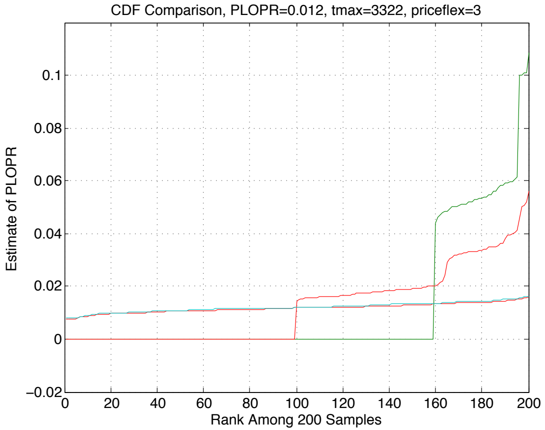
θ Point-Estimate Quantiles



Estimating θ EVEN Better

- Looking only at what happens immediately following a maximum price-change is unnecessarily restrictive
- If two prices differ by more than $2c$, a permanent price-change occurred somewhere in between
- Same intuition as looking at variance scaling over time
- Small t_max, small θ, large $price\_flex$ no longer problematic

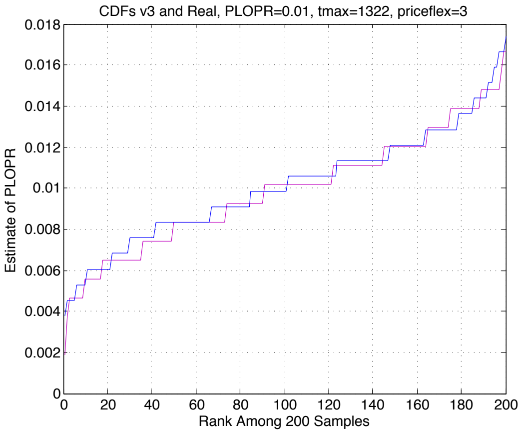
θ Point-Estimate Quantiles



One Catch...

- Not all permanent price changes will cause a difference of more than $2c$
- Analogous to issue that arose with our first two estimators
    - Much harder to correct this bias
    - Can use simulation methods, but proper correction depends on true value of θ
- But this bias tends towards zero as θ gets small
    - So this technique performs best exactly when the first two perform worst

θ Point-Estimate Quantiles



Why We Need Estimators for θ

- As noted, want algorithms that condition decisions on a richer state-space of market conditions
- The new question: why do we need multiple estimators of θ?
  - Routine answer: different estimators are useful for different situations
  - Deeper answer: we want estimators that are well-suited to particular time-scales
- Initial look at the "speed vs. smarts" tradeoff (among other things)

Who Will Get the Prize that's on the Floor?

- Usain Bolt
- Fast

Who Will Get the Prize that's on the Floor?

- Noam Elkies
- Smart

Who Will Get the Prize that's on the Floor?

- $20 bill at the end of a long hallway
    - Edge: Bolt
- $20 bill at the end of a long hallway, inside a safe with a combination that can be determined by solving certain classes of elliptical equations
    - Edge: Elkies
- Disclaimer
    - Noam is actually a very good athlete
    - I have no information about Bolt's intellect; he could be as smart or smarter than Elkies

Many Types of Speed

And Many Types of Smarts

- Later in the term:
  - Latency, processing speed
  - Algorithmic complexity, feasible data scope
  - Data recency (a little bit now)
- Now:
  - Sample-size vs. time left to trade
  - Exiting an unprofitable market ASAP
  - Time-varying underlying market parameters

Some Important Applications and Extensions

- Sample-size vs. time left to trade covered in PS#4
  - More natural and important interpretations than "end of the trading day"
  - (Later this term)

Some Important Applications and Extensions

- Estimate/decision quality as a function of sample size
  - Simplest (monotone) case covered in PS#4
  - Slowly varying market parameters
    - Observations that are too old become misleading
  - Seasonality (esp. intraday)

Further Important Applications and Extensions

- Dynamic entry/exit decision
  - Instead of one-shot, "watch, then decide," repeated "watch, decide, watch, reevaluate, watch, reevaluate,..."
  - Continuous monitoring
  - Think of timescales measured in minutes or probably longer
    - Maybe even days

Further Important Applications and Extensions

- Short-horizon, real-time "entry/exit"
  - Timescales measured in a small number of order/message arrivals, or in milliseconds
  - Not "switch algo on/switch algo off" decisions
    - Instead, "enter/keep/modify/cancel orders" decisions
  - We need some new machinery to tackle this...
- END OF MATERIAL FOR MIDTERM #1

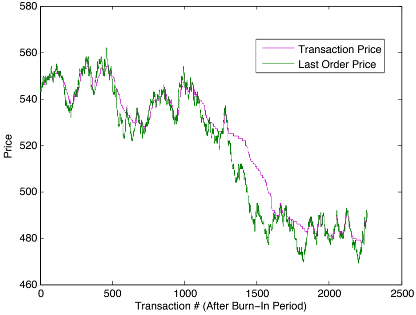
Aggressive Order Flow Characteristics
In the Real World

- Aggressive order direction is persistent
  - E.g., in E-mini market, $\mathbb{P}\{AO_t = buy | AO_{t-1} = buy\} \approx 0.75$
  - Persistence decays very slowly with horizon
    - Have seen numerous streaks of 100+ AOs in a row in the same direction in the E-mini
    - (For comparison, $(0.75)^{100} < \left(\frac{1}{4}\right)^{20} = 2^{-40} \approx 10^{-12}$)
- AO quantities not obviously persistent

Aggressive Order Flow Characteristics
In the Real World

- Signs of AOs are correlated with future permanent price movements
  - Aggressive (e.g.) buy order executes against resting sell orders, potentially raising the best ask mechanically
  - But the best bid also tends to rise after aggressive buy orders
- AOs tend to move the price more than their mechanical effect alone would
- Aggressive order flow carries information about future prices

Aggressive Order Flow Characteristics
In Our Simulations



Aggressive Order Flow Characteristics
In Our Simulations

- Strong persistence in aggressive orders' signs
- Serial correlation in AO sign decays very slowly
- Direction of past aggressive order flow correlated with position of *last_order_price* relative to best bid/ best ask
    - Hence correlated with future permanent price movements
- But does the mechanism in the simulation capture the same structural effects at play in the real world?

Aggressive Order Flow Information
In the Real World

- We assume that informed traders' information eventually becomes public
- Info can be revealed...
  - Directly (e.g., public announcement), or
  - Indirectly, via order-flow features
- In practice, much price-movement cannot be traced to specific public announcements
- Uninformed traders with passive orders (like the MM) assumed to infer information from trading activity

Looking Forward

- Sequences of aggressive orders
- Aggressive order size
- Price-impact of large aggressive orders
    - Permanent
    - Transitory
- Measuring price-impact, trading costs, etc.

FIN 580 Section MMT
Lecture 9

Adam D. Clark-Joseph

University of Illinois Urbana-Champaign

September 28, 2015

Administrative Matters

- PS#4 due Friday, October 2 (e-mail solutions to me)
- No class on Wednesday (9/30)
- Midterm #1 on Monday, October 5 during lecture period

Midterm Logistics

- Covers material up to, and including PS#3
  - Material up to, and including, slide 35 of Lecture 8.
- Held in computer lab in basement of WH (probably WH 24)
- Starts at 11:00AM, sharp. Ends at 12:20pm.
  - I.e., normal lecture period

Midterm Logistics

- Bring your code from all past problem sets on a flash drive
  - Make sure that your code works properly
  - Feel free to use the posted solution code
- Open note (slides from lectures, past problem sets and solutions)
  - NOT open book, not open Hasbrouck-notes
- NO collaboration during the exam
- Main focus: programming from problem sets
  - Small section about assigned readings

Outline

Two Types of Decisions

- Picking an appropriate strategy for the environment of interest
  - "Is it worthwhile to have a convertible (car, not bond) in Urbana?"
  - Should I switch on my MM algorithm in a given market?
    - And w*hen* should I switch it on?
- Picking optimal actions, given a strategy and current conditions
  - "Should I put the top down on my convertible today?"
  - Event X just happened. What should my MM algo do right now?

Different Timescales, Different Decision Types

- Static entry/exit decision
    - Long time horizons
        - Cf. style of car to buy
- Dynamic entry/exit decision
    - Continual monitoring and reevaluation
    - Cf. putting the convertible top up or down
    - Think of timescales measured in minutes to days

Different Timescales, Different Decision Types

- Short-horizon, real-time "entry/exit"
  - Not "switch algo on/switch algo off" decisions
    - Instead, "enter/keep/modify/cancel orders" decisions
  - Timescales measured in a small number of order/message arrivals, or in milliseconds
- This task has a new flavor...

New Data Demands

- Long-term strategic decisions: want to know average conditions
  - Idiosyncratic variation/heterogeneity is undesireable noise
- Real-time tactical decisions: want to know the details of current conditions
  - If it's raining, I want to put my convertible top up
  - Irrelevant at that moment whether rainstorms are rare on average
    - (I hate getting wet.)

New Data Demands

- Using average estimate of θ as the basis for a one-shot entry/exit decision *was* appropriate before
  - Robot1 traded against a random sample of background traders' orders
  - Robot1 treated all aggressive orders in an identical manner, regardless of whether he expected them to be informed
- Treating all aggressive orders in an identical manner is not optimal for an MM
  - Ideally, MMs want to
    - Trade with uninformed counterparties
    - NOT trade with informed counterparties
  - If MM could foresee and avoid informed traders, the avg. fraction of informed trading wouldn't matter

Aggressive Order Flow Characteristics

In the Real World

- Aggressive order direction is persistent
  - E.g., in E-mini market, $\mathbb{P}\{AO_t = buy | AO_{t-1} = buy\} \approx 0.75$
- Persistence decays very slowly with horizon
  - Have seen numerous streaks of 100+ AOs in a row in the same direction in the E-mini
  - For comparison, $(0.75)^{100} < \left(\frac{1}{4}\right)^{20} = 2^{-40} \approx 10^{-12}$
    - (Which would be about once every 30,000 years)
- AO quantities not obviously persistent

Aggressive Order Flow Characteristics
In the Real World

- Signs of AOs are correlated with future permanent price movements
  - Aggressive (e.g.) buy order executes against resting sell orders, potentially raising the best ask mechanically
  - But the best bid also tends to rise after aggressive buy orders
- AOs tend to move the price more than their mechanical effect alone would
- Aggressive order flow carries information about future prices

Aggressive Order Flow Information

In the Real World

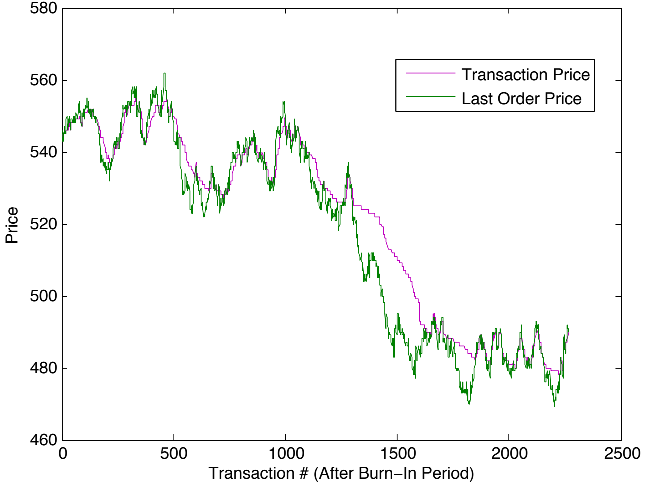
- We assume that informed traders' information eventually becomes public
- Info can be revealed:
  - Directly (e.g., public announcement), or
  - Indirectly, via order-flow features
- In practice, much price-movement cannot be traced to specific public announcements
- Uninformed traders with passive orders, like the MM, assumed to (and appear to) infer information from trading activity

Examining Richer Data

- [This part of the lecture is fairly close to Hasbrouck Ch. 7]
  - [Hasbrouck Ch. 7 is a mix of Glosten-Milgrom and Easley-O'Hara]
- We've considered the informational content of a single aggressive order
  - Expectation of future price updates somewhat in the direction of the AO when AO executes
- Now we'll consider a sequence of aggressive orders
  - In particular, streaks of AOs all in the same direction

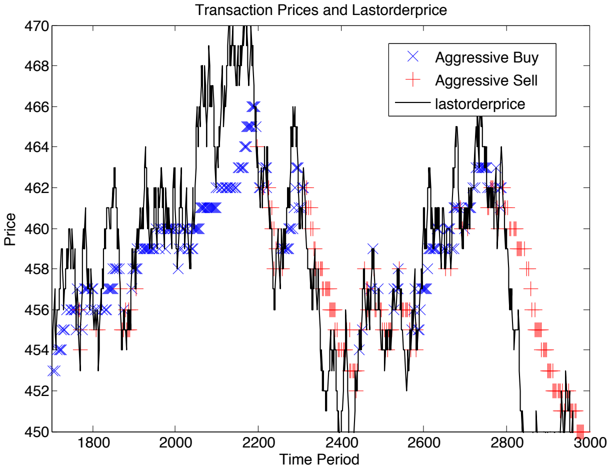
Aggressive Order Flow Characteristics

In Our Simulations



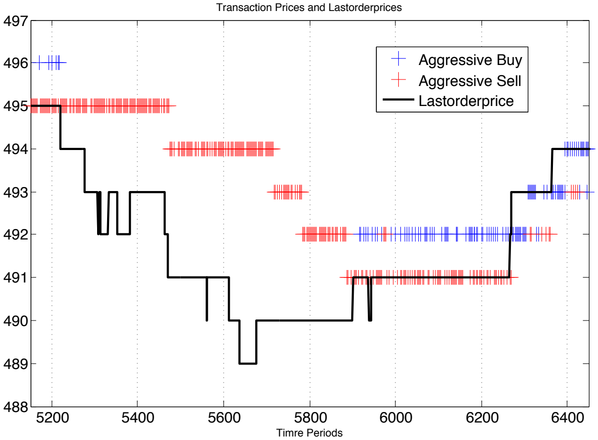
Aggressive Order Flow Characteristics

In Our Simulations



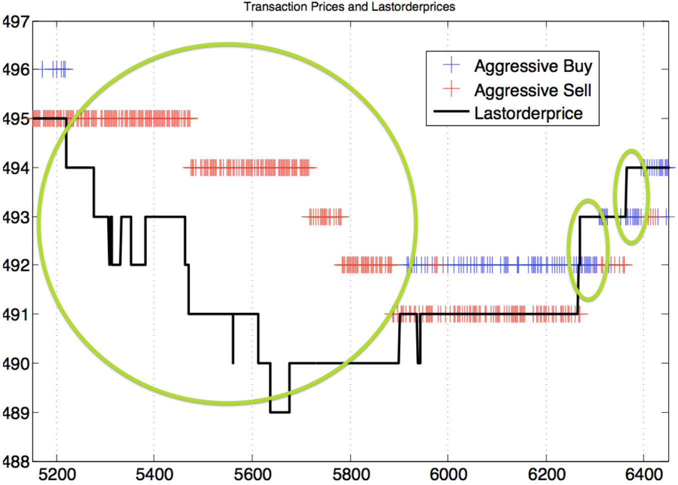
Aggressive Order Flow Characteristics

In Our Simulations



Aggressive Order Flow Characteristics

In Our Simulations



Aggressive Order Flow Characteristics
In Our Simulations

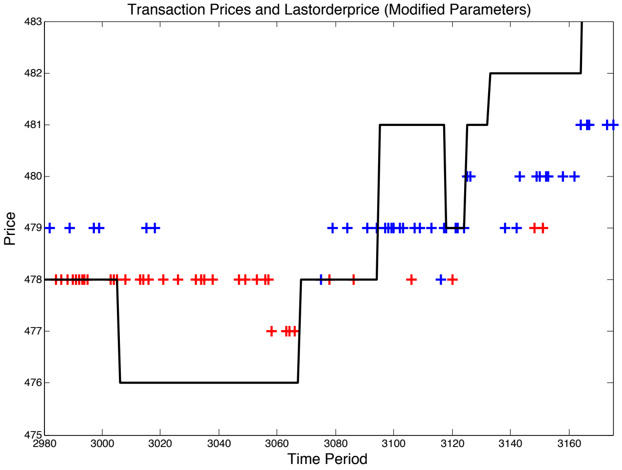
- Strong persistence in aggressive orders' signs
- Serial correlation in AO sign decays very slowly
- Direction of past aggressive order flow correlated with position of *last_order_price* relative to best bid/ best ask
  - Hence correlated with future permanent price movements
- But does the mechanism in the simulation capture the same structural effects at play in the real world?

Partly

- Streaks of aggressive buys (e.g.) arise when *last_order_price* is above the current best ask
  - *last_order_price* has already changed, but orderbook does not yet reflect this change
  - Natural interpretation as informed trading
- Absence of uninformed aggressive trading in the wrong direction at times when aggressive orders in opposite direction are informed
- Feature of real world, as well as of the simulation
- But it's just a mechanical artifact in the simulation for certain parameter values

Aggressive Order Flow Characteristics

In Our Simulations Using Different Parameter Values



Two-Fold Revelation by Aggressive Orders

- Two things we can infer from the arrival of an aggressive buy (e.g.)
  - Expectation of FV is higher than we had thought
  - If the next AO is a buy, it's more likely to be informed
- Non-obvious implication:
  - Even though a subsequent aggressive buy is more likely to be informed, we may learn less from it!

Learning from a Sequence of AOs

- Suppose we observe two aggressive buys in a row
- Why would we expect the second AO to be informed with higher probability if it's a buy?
  - More likely to see two buys in a row when the signs of adjacent AOs are positively correlated
  - Trading on the same piece of private information would induce such correlation
- So if the second AO is telling us something, it's (probably) telling us the same thing as the first AO
  - Assuming we know about the average autocorrelation structure of AO signs, our belief update from the first AO included expected contributions from the conditionally expected future AOs

- Read the Hasbrouck example in Ch. 10 (also make sure to read Hasbrouck Ch. 9)
- Central intuition: our current expectation includes our expectation of future revisions
  - Since AO sign is persistent, we can forecast direction of future AOs
  - Hence can also forecast what we would learn from those
  - We only learn from the surprise component of aggressive order flow

Easley-O'Hara Model

- Endogenize trading decisions of informed and uninformed aggressive traders
  - When (sequentially) to place market orders
  - What size of market order to place
- Strategic considerations now important for the MM's inference problem
- Can obtain a separating equilibrium in which informed traders submit large market orders, and uninformed traders small ones

Clock-Time and Event-Time

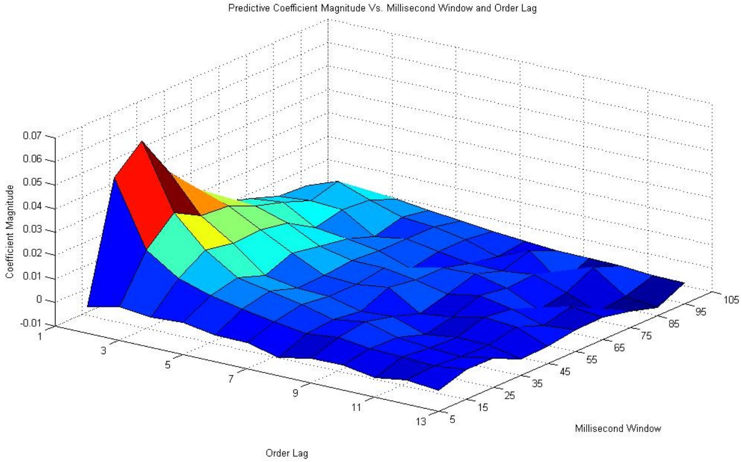
- "Clock-time"—measured by seconds/minutes/hours/etc.
- "Event-time"—measured how many times a specified type of event has occurred
    - E.g., event = aggressive order arrival
- Canonically, we can take aggressive orders to be uniformly spaced in event-time

Clock-Time and Event-Time

- Absence of uninformed aggressive trading in the wrong direction at times when aggressive orders in opposite direction are informed
  - Empirical regularity, but not explicable in terms of our simulations
- One real-world mechanism: AO clustering in clock-time
- Empirically, AO arrival rates change sharply throughout the day
  - Quiet, slow periods, with ocasional bursts of activity
- Streaks of many AOs in the same direction usually don't last long in terms of clock-time...

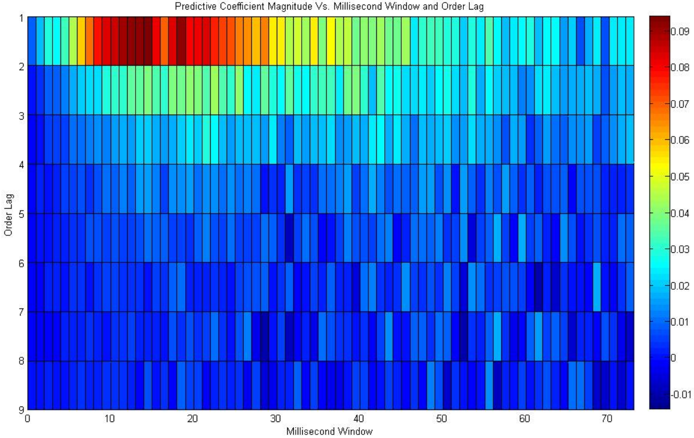
Clock-Time and Event-Time

How much can we learn from AO signs as they grow less recent?



Clock-Time and Event-Time

How much can we learn from AO signs as they grow less recent?



Next Week

- Monday: overview and recap
- Wednesday: No Class
- Monday 10/5: Midterm #1, during lecture time