Wall Street Meets Locust Walk: The Penn-Lehman Automated Trading Project

Michael Kearns
Computer and Information Science
University of Pennsylvania

Economics and Computation

- Electronic Commerce
- Algorithmic Mechanism Design
- Network Economics
- Digital Goods
- Shopbots

- Computational Game Theory
- Combinatorial Auctions
- Information Markets
- Electronic Cheating
- Neuroeconomics

Theory, Algorithms… and Competitions

Meanwhile, in the *Real* World...

- Increased automation in financial markets (exchanges, program trading, tools, strategies)
- Technological, business and regulatory drivers (decimalization, ECNs)
- Increased revelation of market data

We're not going back.

Motivation

- Can we somehow merge
  - the realism, speed and excitement of trading in real securities markets
  - the exploratory, educational and creative nature of trading in virtual markets

The Penn-Lehman Automated Trading (PLAT) Project

- Real-time, automated trading in financial markets
- A realistic, multi-client market simulator using live and historical depth-of-market *data*
- A competitive testbed for trading strategies
- An investigation of order book strategies
- An investigation of AI and machine learning methods for automated trading
- A research and educational partnership between Penn and Lehman Brothers

Market Microstructure and Depth-of-Market

- Consider a typical exchange for some specific security
- Market order: give volume, leave price to "the market"
- Limit order: specify price (away from the market)
- Limit orders are placed in a book on the buy or sell side, ordered by price
- Market orders are matched with the top of the book on the opposing side
- Market orders "guaranteed" transaction but not price; limit orders guaranteed price but not transaction

Here's an example.

Electronic Crossing Networks (ECNs)

- Matching process as old as securities markets
- Often done manually (NYSE specialists)
- Electronic Crossing Networks (ECNs):
  - Automate matching process
  - Now publishing order book data
  - Can be highly liquid
- Examples: Island (14% of NASDAQ volume), Instinet, Archipelago,…
- Consolidation efforts (SuperMontage)
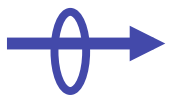- Messaging protocols, APIs (OUCH and ITCH)

The Penn Exchange Simulator (PXS)

- A (conceptually) simple, real-time market simulator exploiting order book data
- Core processes:
  - Data Gathering: frequent (live or historical) polling of any given stock from Island
  - Client Order Processing: acceptance of connections and limit orders from automated clients via API
  - Market Simulation: simulation of matching process in a virtual market merging the Island and client data; computation of profit and loss of clients

refresh | island home | disclaimer | help

GET STOCK

MSFT

MSFT go

Symbol Search

LAST MATCH		TODAY'S ACTIVITY	
Price	24.0700	Orders	52,983
Time	14:57:07.72	Volume	10,243,212

BUY ORDERS		SELL ORDERS	
SHARES	PRICE	SHARES	PRICE
500	24.0620	500	24.0690
6,000	24.0610	500	24.0690
5,000	24.0600	500	24.0700
100	24.0600	200	24.0800
1,100	24.0550	1,981	24.0900
100	24.0500	412	24.0900
5,000	24.0500	3,000	24.0980
200	24.0500	500	24.1000
3,294	24.0500	100	24.1200
1,000	24.0500	2,800	24.1400
3,000	24.0430	5,000	24.1400
100	24.0400	1,000	24.1400
5,503	24.0400	5,000	24.1500
2,100	24.0300	400	24.1600
2,800	24.0300	1,000	24.1700
(412 more)		(694 more)	

As of 14:57:16.178

PXS

BUY	SELL
24.0640	24.0670
24.0630	24.0690
24.0620	24.0690
24.0610	24.0700
24.0605	24.0750
24.0600	24.0800
24.0600	24.0900
24.0550	24.0900
24.0530	24.0980
24.0500	24.0990

Client 1

Client 2

Client 3

Features of PXS

- Simulation merges real market with client orders
- No guesses or models for limit order fills
- Permits investigation of order-book strategies
- Permits high-speed, high volume trading
- Forces real-time performance
- Sandbox for diverse strategies & interaction
- Execute "live" on real-time data or historical

PLAT Project Participants

- Penn Development Team:
  - PXS design, development, maintenance
  - Client strategy design
  - Education of student researchers/users
  - Competition design and execution
- Students:
  - Approximately 30 students designing 14 strategies
  - Many senior projects; many joint CIS-Wharton
  - Several external participants (UTexas and CMU)
- Lehman Brothers:
  - Financial support from Proprietary Trading Group
  - Student mentorship and competition judging
  - Technological and scientific advising

Goals for the Strategies and Competitions

- Attempt to recreate Wall Street "biodiversity"
- Market making, pairs trading, VWAP, block trading, technical strategies, etc.
- Investigate predictive value of order book data
- Investigate application of AI and ML methods
- Create a library of strategies and competition structures for mix and match experimentation

The Competitions

- All had +/- 100,00 share position limits
- All on MSFT trading
- All divided clients into two pools
- December 2002: historical and live data competition
- Order-book face-off
- March 2003: 3-round, 3-week competition
- May 2003:
  - Platinum Platter Competition (PPC 2003)
  - 10 consecutive days of MSFT trading, no rounds
  - 6 scoring criteria balancing profitability and trading "hygiene"
  - Counted towards grades

PPC 2003: The Rules

Criteria Emphasizing Profitability.

- **Daily Profit and Loss.** On a daily basis, 3 points awarded to each client whose end-of-day P&L is highest in its seven-client pool; 2 points to the second highest; 1 point to the third highest. Maximum possible award: 30 ($= 3 \times 10$) points.

- **Overall Consistency of Profit and Loss.** A one-time award of 15 points to any client that has positive cumulative P&L over the 10 trading days of the competition, and also ends at most 3 trading days with negative daily P&L. Maximum possible award: 15 points.

Criteria Emphasizing Robustness, with Weak Profitability Prerequisites.

- **Daily Intraday Position Reversals.** On a daily basis, 2 points awarded to any client that finishes with positive P&L for the trading day, and held share positions in excess of 10,000 shares in both the long and short direction at some point during the day. Maximum possible award: 20 ($= 2 \times 10$) points.

- **Robustness to Market Variation.** An award of 5 points to each client that has positive P&L on any pair of trading days in which the share price rose overall (open to close) on one day of the pair, and fell on the other. For each additional such pair, an additional award of 5 points. Maximum possible award: 25 ($= 5 \times 5$) points, if there are exactly 5 up days and 5 down days for the stock during the 10 competition days.

Criteria Emphasizing Good Trading Practices, with No Profitability Prerequisites.

- **Daily Risk Saturation.** On a daily basis, 2 points awarded to each client that achieves a share position in excess of 50,000 shares (long or short) at some point during the trading day, without exceeding the maximum allowed share position of 100,000 shares. Alternatively, these 2 points will be awarded to clients whose total matched volume of shares for the trading day exceeds a fraction 1/14 (which is half of the per-client average of 1/7) of the total matched volume of all clients. Maximum possible award: 20 ($= 2 \times 10$) points.

- **Daily Position Unwinding.** On a daily basis, any client that is awarded the Risk Saturation points for that day may earn an additional 2 points by ending the trading day with a share position of less than 5,000 shares (long or short). Maximum possible award: 20 ($= 2 \times 10$) points.

Figure 1: Client Scoring Criteria for PPC 2003.

PPC 2003: The Strategies

Strategy Name	Description	Performance
CBR-SOBI	Case-based reasoning applied to the parameters of the SOBI strategy (see text for SOBI description).	First in the blue pool; statistically significant profitability.
MoneyFlow	Predictive strategy using money flow (price movement times volume traded) as a trend indicator.	Second in the blue pool.
OBMM	Market-maker that positions orders in front of the nth orders on both books.	Third in the blue pool.
CReaTiv	"Capitalization on Real Time Volatility" — SOBI modified by recent volatility	Fourth in the blue pool.
OBCrossover	Exponential Moving Average (EMA) crossover strategy moderated by confirmation of order book quartile VWAPs.	Tied for fifth in the blue pool.
OBBreakout	Breakout strategy applied to trend lines on VWAP of buy and sell books.	Tied for fifth in the blue pool.
RaSTa	"Resistance and Support Trading Agent" — Computes support and resistance levels based on peaks in order book volume.	Seventh in the blue pool.
DAMM-STAT	Mixture of a Dynamically Adjusted Market-Maker which calibrates by recent volatility, and a trend-based predictive strategy.	First in the red pool; stellar position management.
Contrarian	Sells on rising prices, buys on falling prices.	Second in the red pool.
OBSigma	Trades based on relative spreads in the buy and sell books, interpreting small standard deviation as a sign of confidence.	Third in the red pool.
OBVol	Simple predictive strategy using total volumes in buy and sell books.	Fourth in the red pool; highest Sharpe ratio and statistically significant profitability.
RapidMM	Market-maker with rapid revision of quotes based on current inside market.	Fifth in the red pool.
CIA	"Central Intelligent Agent" — Predictive strategy applying Boosting to order book snapshots.	Sixth in the red pool.
SimpleTrend	Simple trend prediction strategy.	Seventh in the red pool; statistically significant negative earnings.

Figure 1: PPC 2003 Strategy Descriptions, grouped by competition pool.

Strategy Themes

- Order-book strategies
  - order book imbalance and variations
  - 9 of 14 competition entrants
- Market-making strategies
- Traditional technical strategies:
  - various "breakout" strategies, trend spotting, etc.
  - often with order-book modulation
- Machine learning strategies using order books
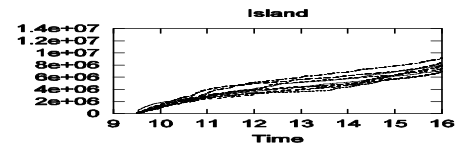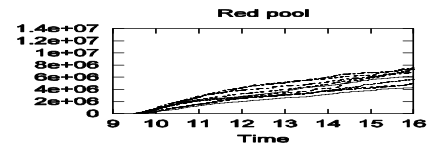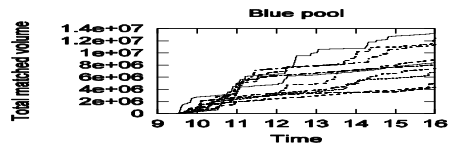 - boosting, case-based learning, support vectors

PPC 2003: The Results

Strat	Pool	Rk	Tot	P	S	U	R	C	V	Av P&L ($)	95% Int	Sharpe
CBR-SOBI	B	1	74	9	18	2	10	15	20	4187	± 3733	0.70
MoneyFlow	B	2	69	15	20	0	4	15	15	2007	± 15692	0.08
OBMM	B	3	46	8	20	0	8	0	10	258	± 7909	0.02
CReaTiv	B	4	42	7	20	0	10	0	5	(2410)	± 6770	(0.22)
OBCrossover	B	5	33	6	6	0	6	0	15	3242	± 4220	0.45
OBBreakOut	B	5	33	10	18	0	0	0	5	3680	± 7963	0.29
RaSTa	B	7	21	5	2	0	4	0	10	1182	± 2441	0.30
DAMM-STAT	R	1	65	6	20	14	10	0	15	685	± 5195	0.08
Contrarian	R	2	55	6	20	2	12	0	15	2022	± 3658	0.34
OBSigma	R	3	54	8	20	6	10	0	10	1649	± 2382	0.43
OBVol	R	4	53	14	0	0	4	15	20	4037	± 1900	1.32
RapidMM	R	5	50	10	20	0	10	0	10	3649	± 9121	0.25
CIA	R	6	30	13	12	0	0	0	5	(1451)	± 9822	(0.09)
SimpleTrend	R	7	27	3	20	2	2	0	0	(24467)	± 17974	(0.84)

Figure 1: Summary of Results for PPC 2003. Column *Tot* shows total points for each client over the 10-day competition. Column *P* shows points earned in the Daily Profit and Loss category. Column *S* shows points earned in the Daily Risk Saturation category. Column *U* shows points earned in the Daily Position Unwinding category. Column *R* shows points earned in the Daily Intraday Position Reversal category. Column *C* shows points earned in the Overall Consistency of Profit and Loss category. Column *V* shows points earned in the Robustness to Market Variation category. Column **Avg P&L** shows average daily Profits and Losses, while column **95% Int** gives the 95% confidence interval around this mean. Column **Sharpe** gives the 10-day Sharpe Ratio for the client. See text for more details.
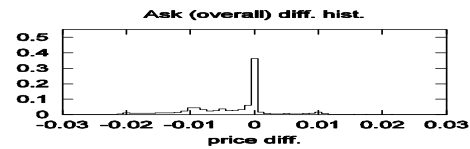
Global Analysis

Total Volume Executed

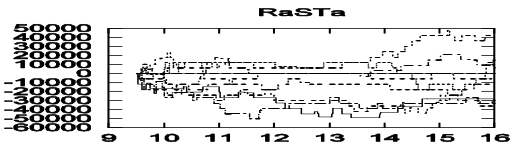# Price and Inside Markets vs. Island
(PXS – Island)
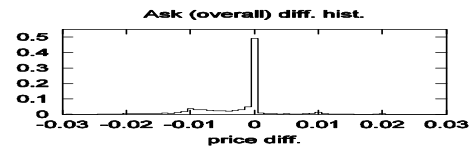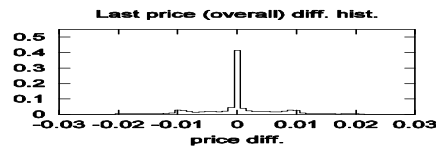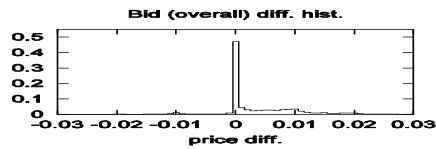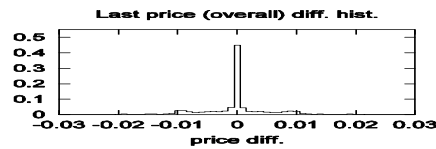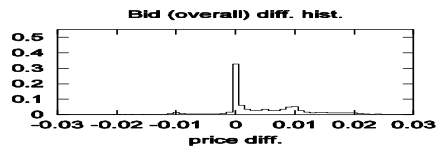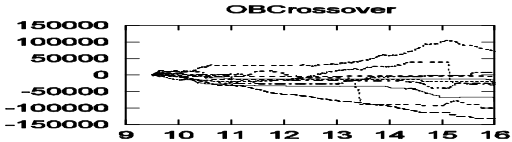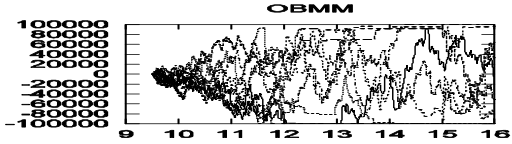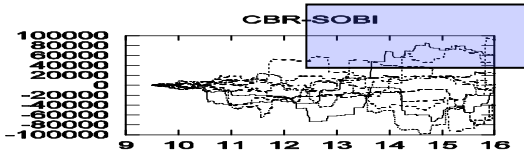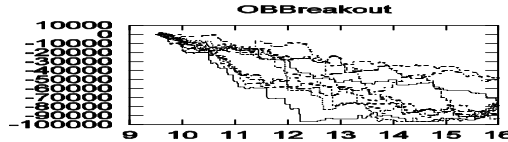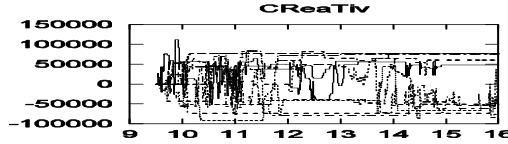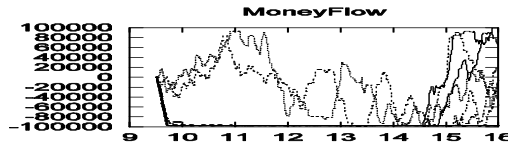
Client Analysis

Blue Pool Client
Share Positions

DAMM-STAT

Contrarian

500K

OBSigma

OBVol

1.7M

RapidMM

CIA

700K

SimpleTrend

60K

2M
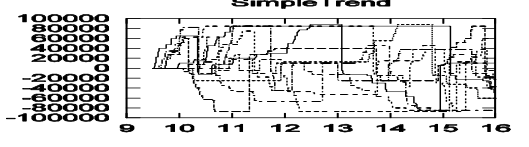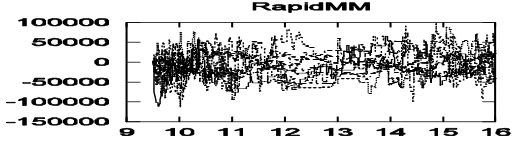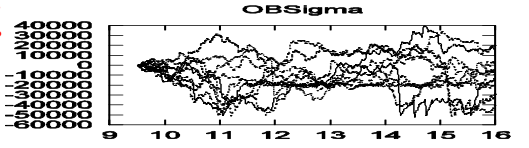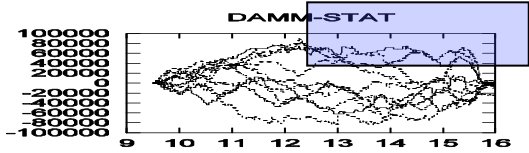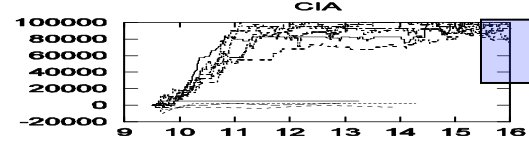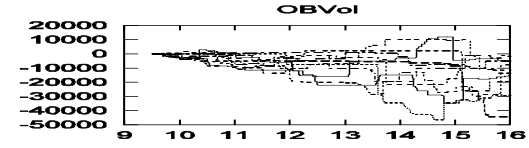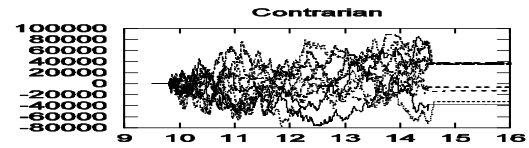
150K

**Red Pool Client
Share Positions**

400K

The SimpleTrend Debacle

SimpleTrend Client Intraday P&L for May 9

SimpleTrend Client Share Position for May 9

Intraday Share Prices for May 9
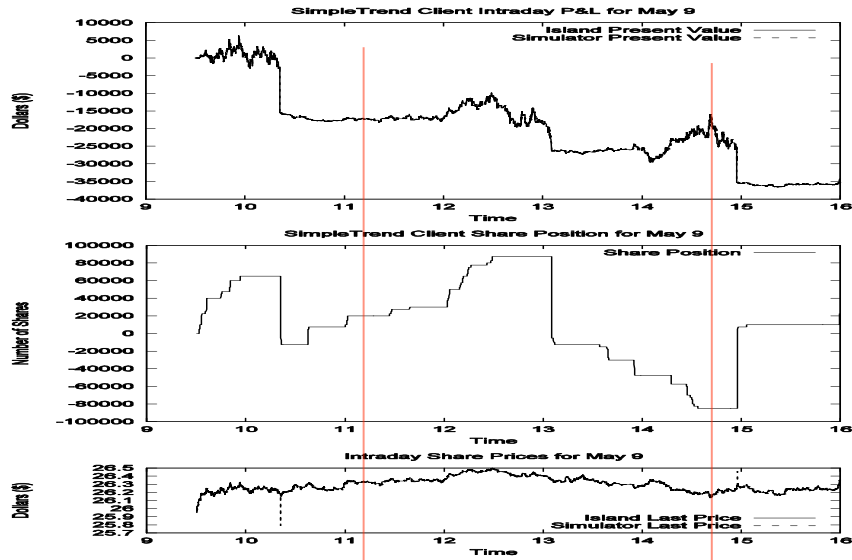
Conclusions

- Electronic Commerce is alive and well
- Penn-Lehman project:
  - exploits advent of order book data
  - mixes the real and virtual
  - educates in a friendly but competitive environment
  - has produced a couple of promising strategy ideas

Collaborators

- Penn:
  - **Luis Ortiz**
  - Berk Kapicioglu
  - Byoungjoon Kim
  - Rashid Tuweiq
  - Elliot Feng

- Lehman Brothers:
  - Andy Ellner
  - Mark Sanborn
  - Colin Rust
  - Amy Papandreou
  - Michael Schulman
  - Michael Bos
  - Michael Bleich

Further Information

- Email: mkearns@cis.upenn.edu
- Web sites:
  - www.cis.upenn.edu/~mkearns
  - www.cis.upenn.edu/~mkearns/projects/plat.html
  - www.cis.upenn.edu/~mkearns/papers/plat.pdf
    (joint paper on the project with Luis Ortiz)
- New participants welcome!