

Database Announcement

- Public access database of daily time-weighted quoted spreads, dollar-weighted effective spreads, and dollar-weighted realized spreads and price impact at 1 and 5 second horizons.
- All common stocks and ETFs
- Daily data, based on TAQ daily quote, NBBO, and trade file. Beta version of database at
 - <https://research.wpcarey.asu.edu/investment-engineering/spread-statistics/>
- Updated overnight with one day lag



The Anatomy of Trading Algorithms

Tyler Beason, ASU
Sunil Wahal, ASU

April 2020

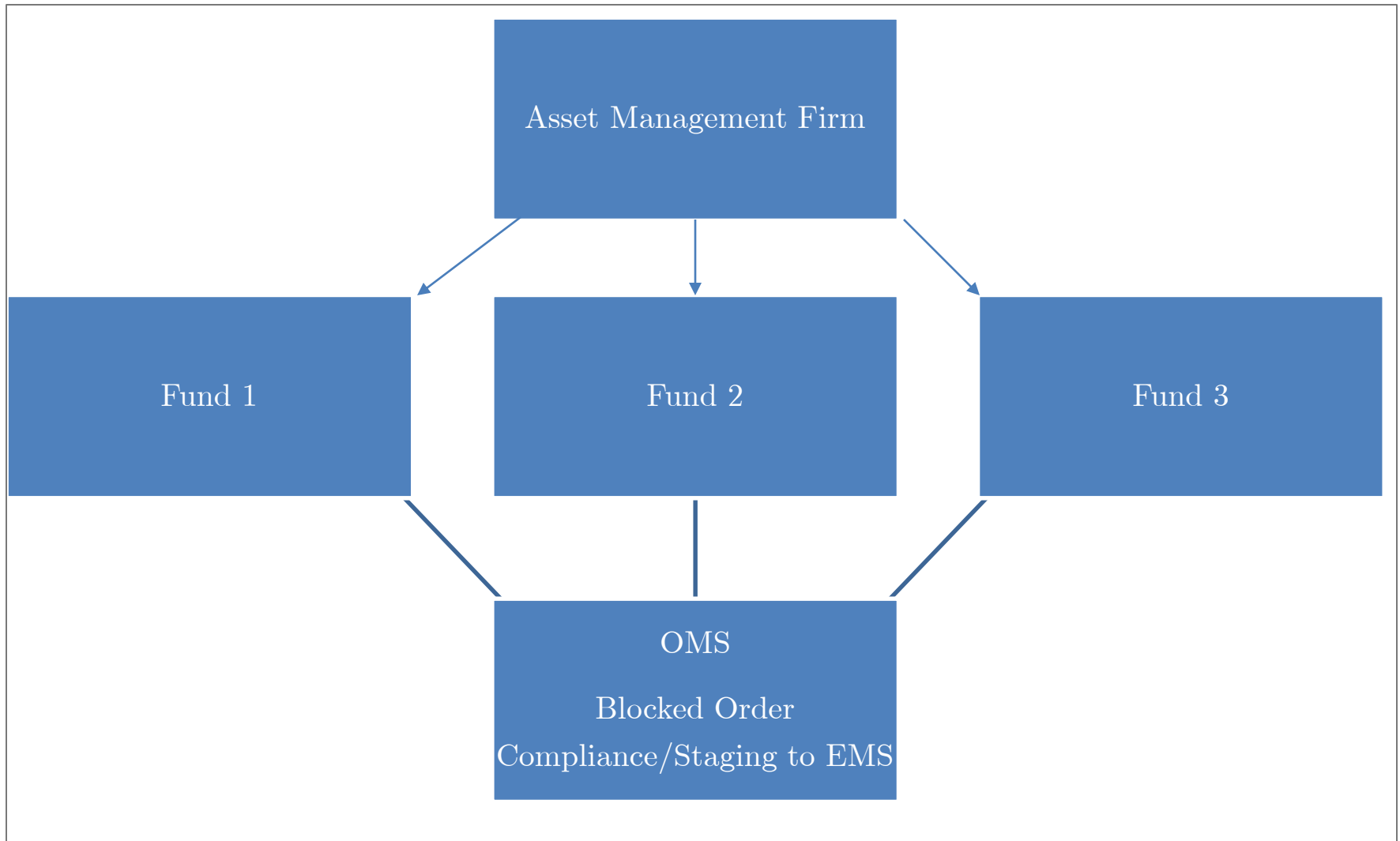
Background

- Classical models are outdated
- New models focus on HFT, ignoring buy-side
- Why?
 - No straightforward intuition (exception, LWY)
 - No data
 - Institutions don't want to release the data
 - Brokerage firms don't want to release the data
 - Ancerno-type data are essentially useless (end of day trading tickets)
 - Single-manager datasets: embedded endogenous choices, bespoke trading, generalizability issues

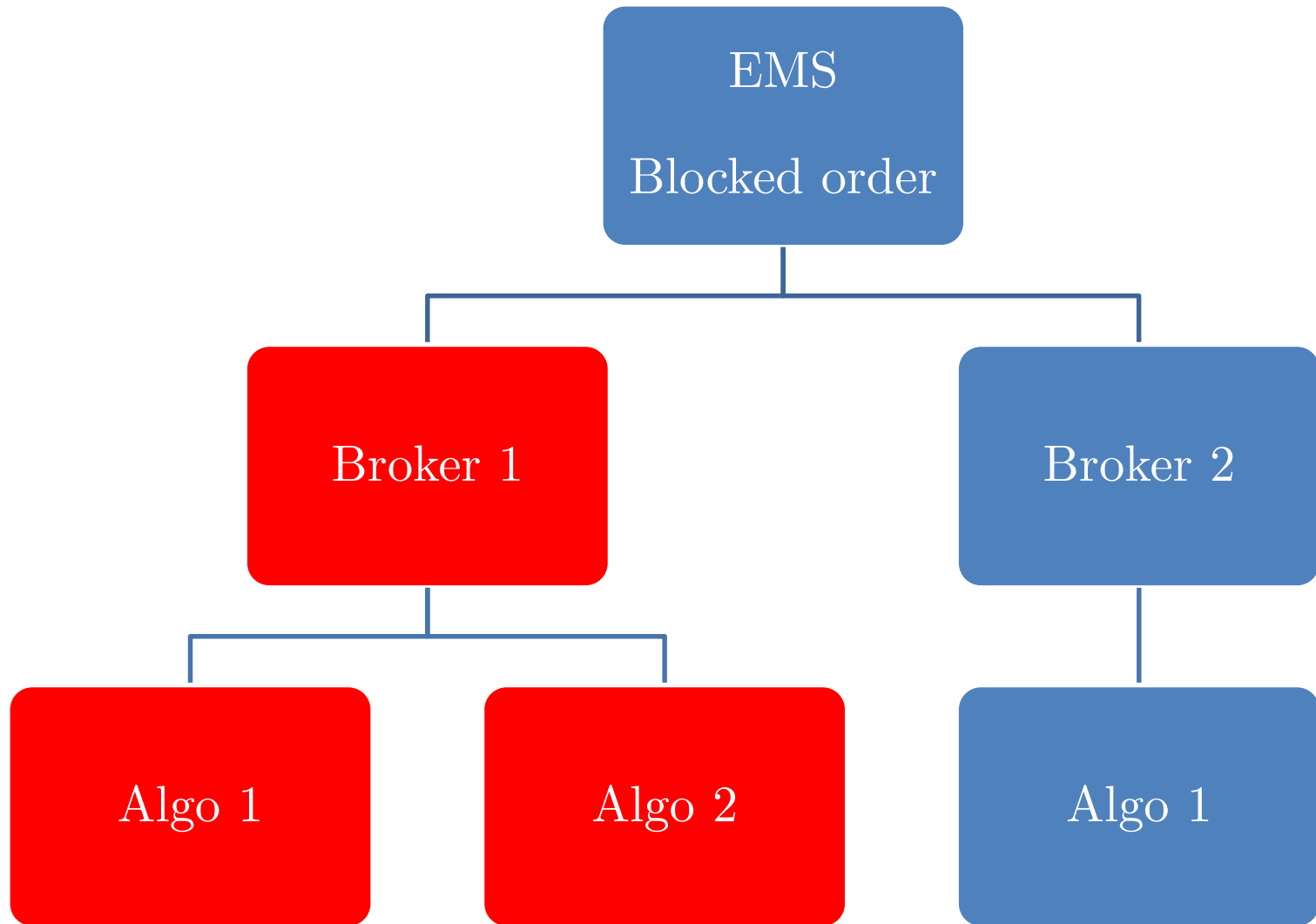
Objectives

1. Anatomy
 - To inform price formation
 - Anatomy informs diagnosis, and if necessary, cure
 - Gross anatomy (Parent)
 - Cytology (Child)
 - Histology (Runs)
2. To study the tradeoff between likelihood of execution and transaction costs
 - Fundamental issue
 - Building block of order choice models and price formation

Trading Process: Level 1 (OMS)



Trading Process: Level 2 (EMS)



Data overview

- Large provider of algorithmic trading services
 - Diverse client range, from buy-side long-only managers to multi-asset hedge funds. 961 clients, 2012-2016
- Four standardized, non-bespoke single-stock algorithms.
 - Use direct exchange feeds, often white-labelled
 - All time stamps in milliseconds with FIX tags
- \$675 billion in demand, trading over 5,000 securities
- Parent -> child -> fill
- 2.3 million parent orders generate
 - 300 million child submissions
 - \$2.1 trillion in notional volume
 - \$388 billion in traded volume

Data elements

Parent

- Client ID
- Algo type
- ID, Symbol, side (B, S, SS)
- Start/End time
- Order Quantity
- Price and/or volume constraints

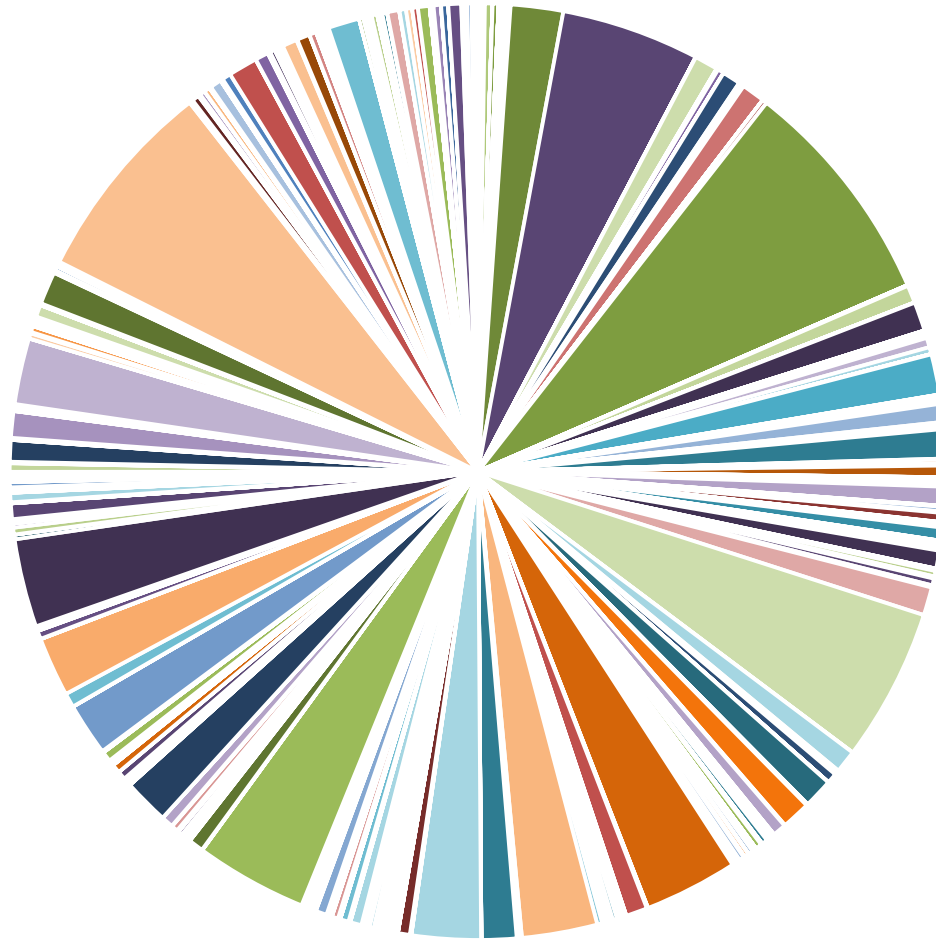
Child (all, not just filled)

- ID, link to parent
- Time (send, receive)
- Order type (M, L, P)
 - Subtype (limit prices)
- Display instruction
- Exec Instr. (FIX 18)
- TIF (FIX 59)
- Venue
- Fill (Price, Quantity)
- Add/Take/Route (FIX 851)
- Fee/Rebate

Gross Anatomy: Parent Orders

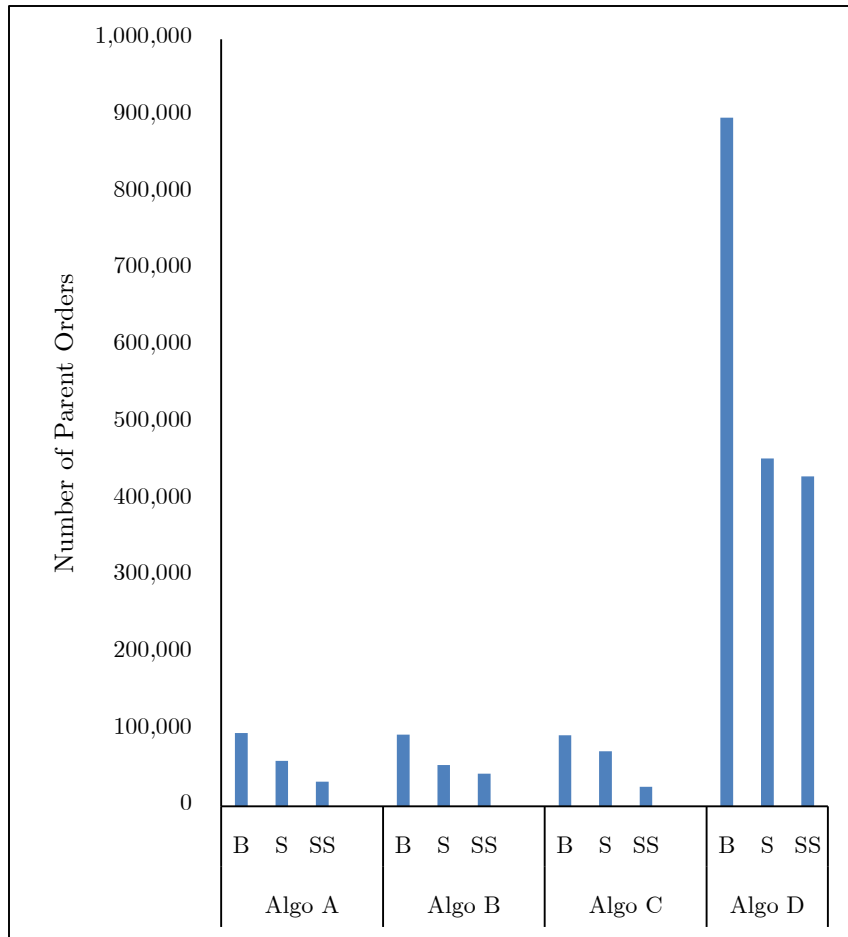
Dollar distribution of algo usage

Distribution of Algorithm Volume by Institutions

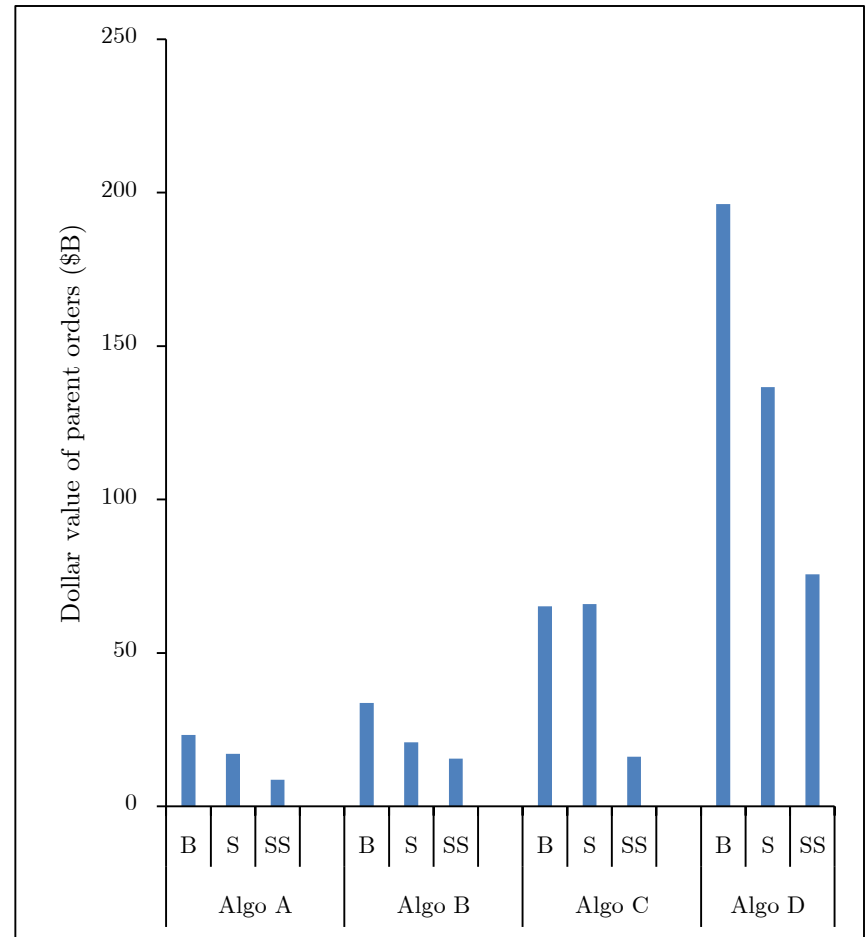


Distribution of parent orders

By Number

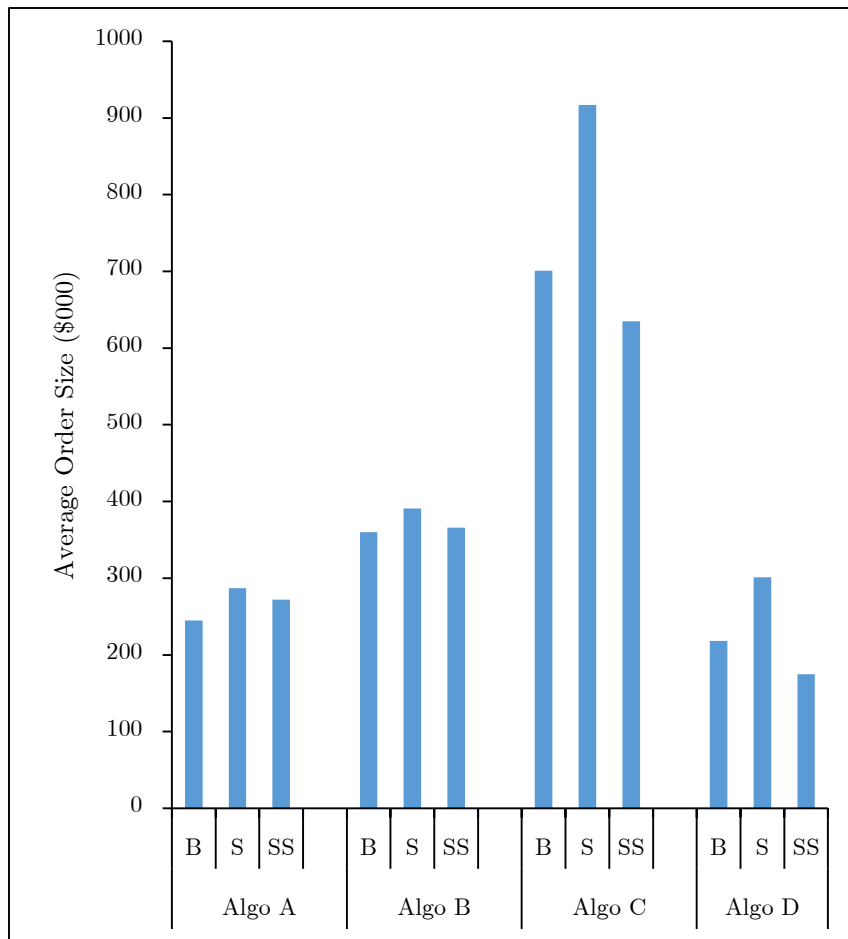


By Dollar Volume (\$B)

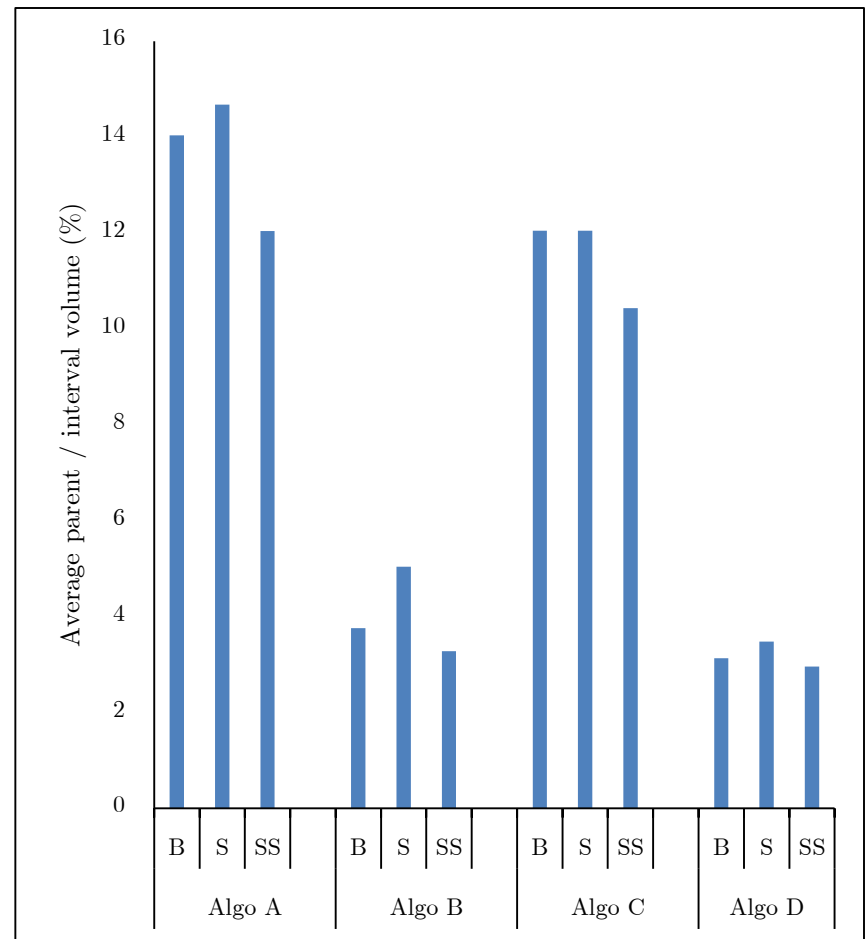


Average parent size

Parent size (\$000)

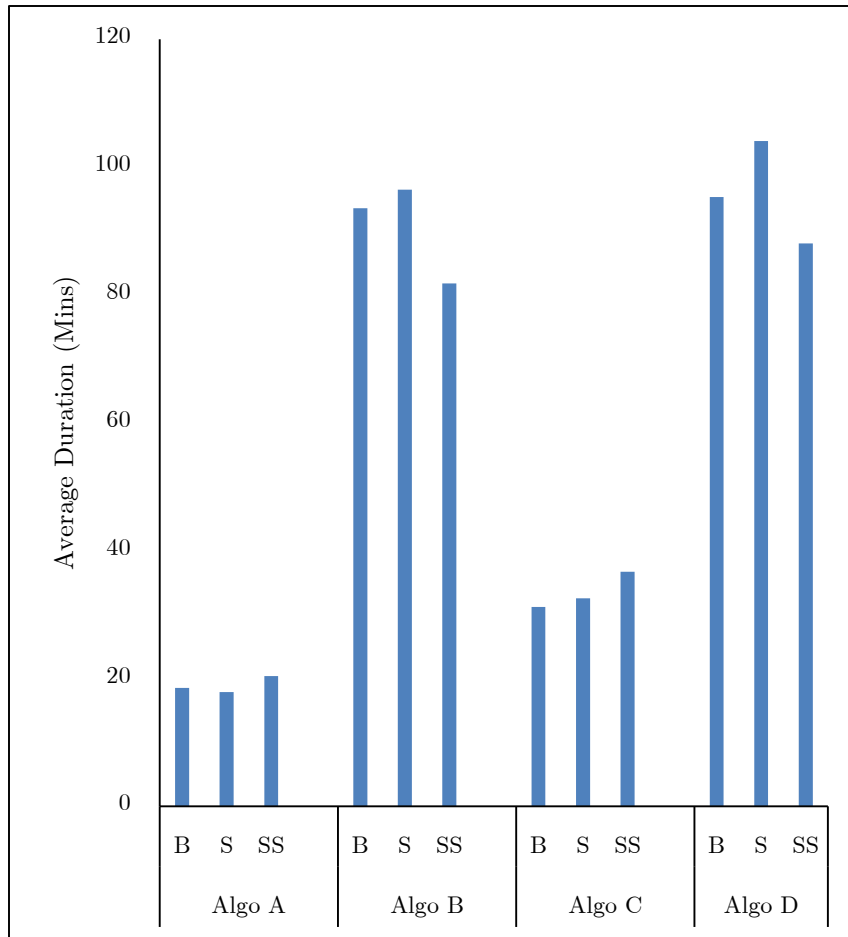


Parent / interval volume

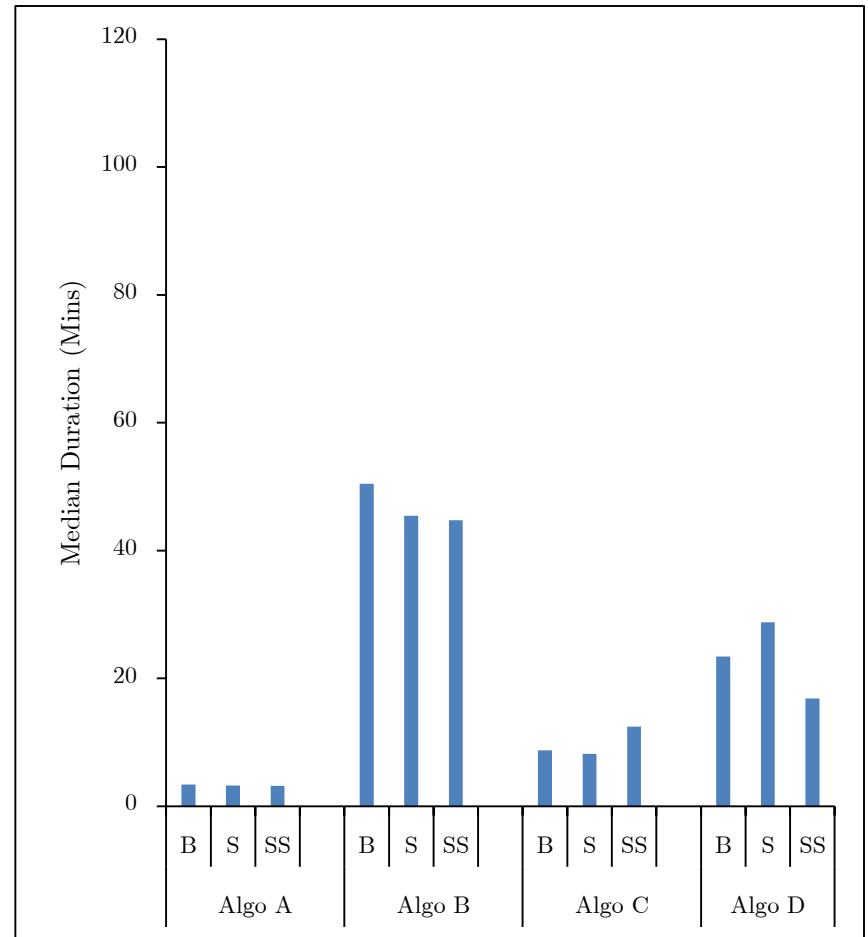


Parent duration

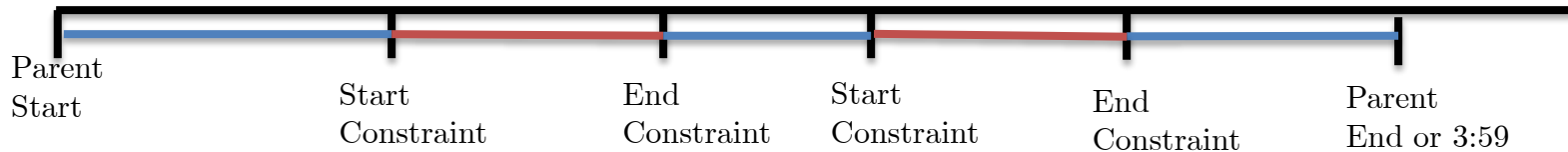
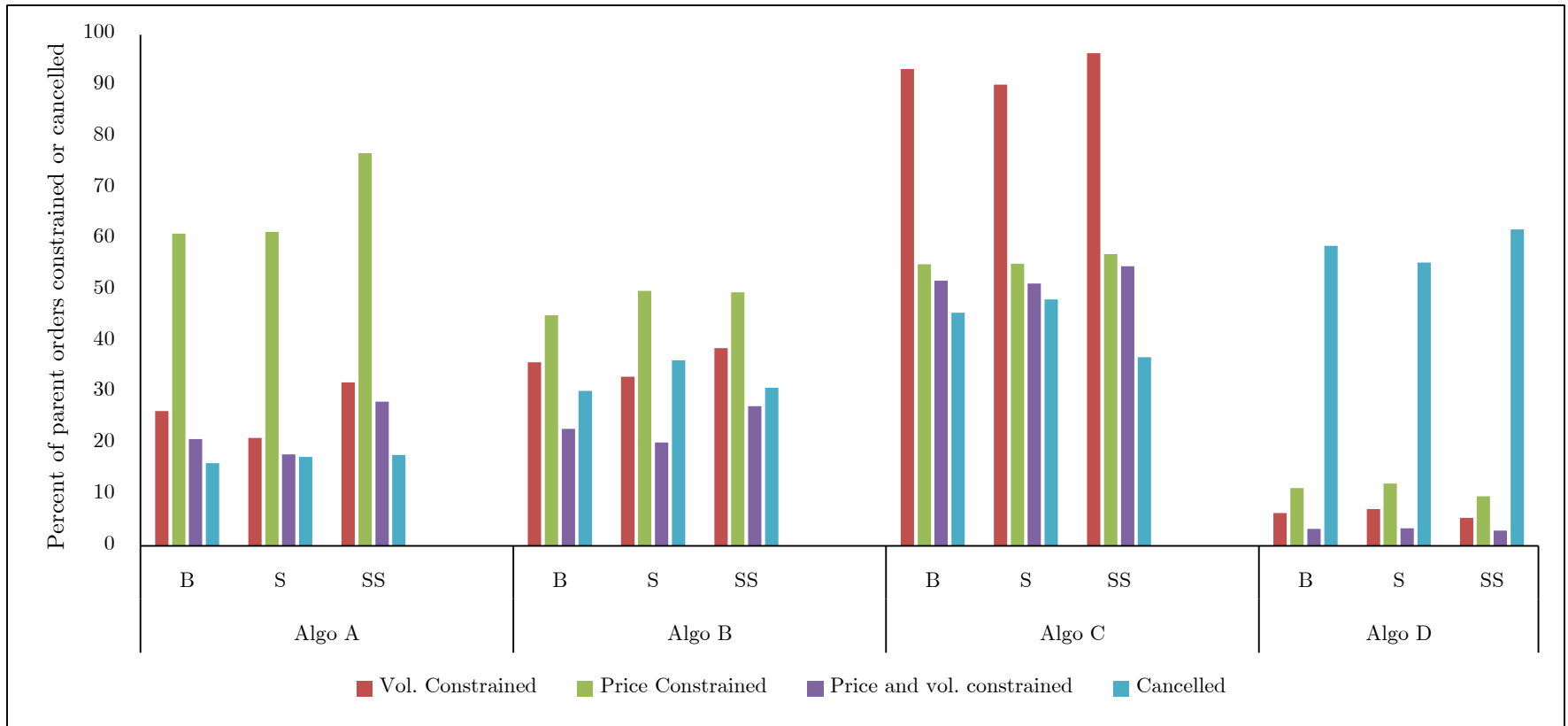
Mean



Median

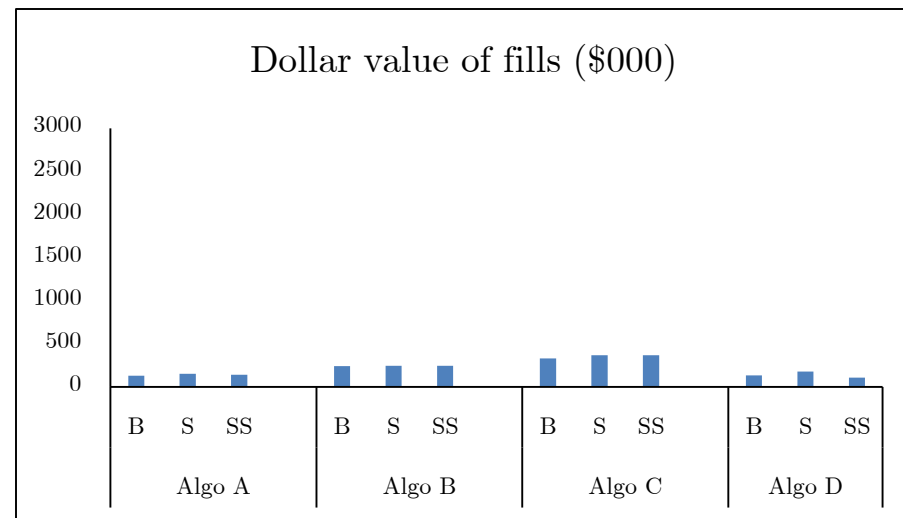
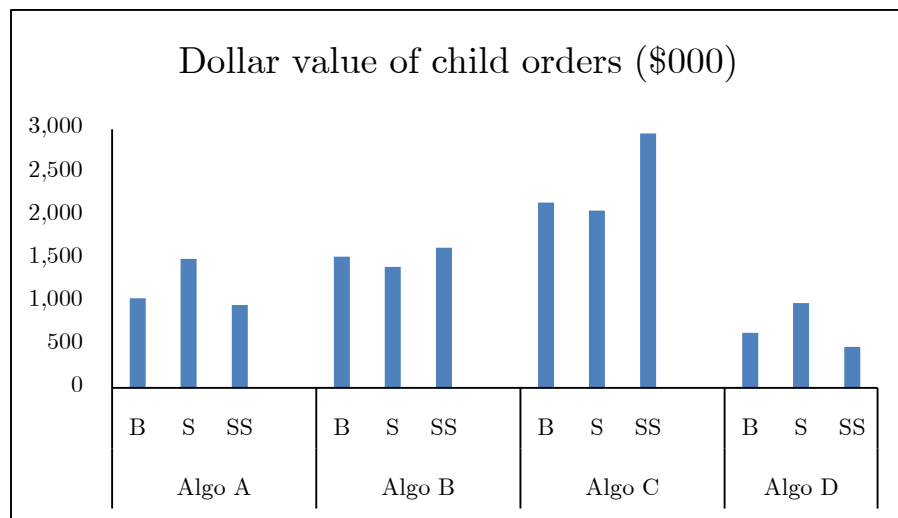
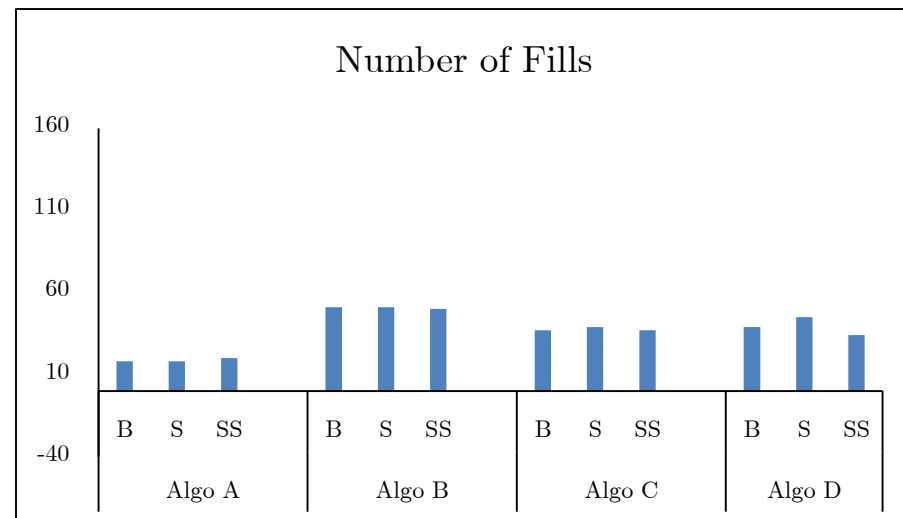
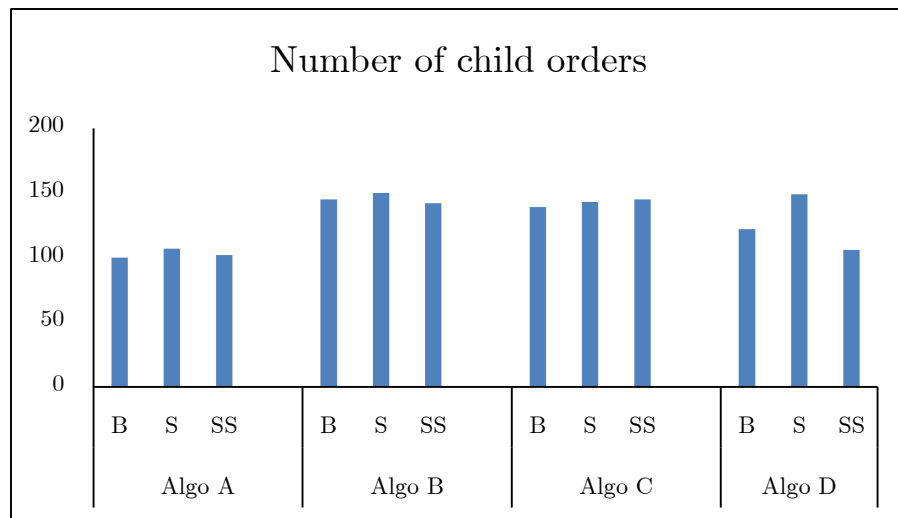


Parent constraints & cancellations



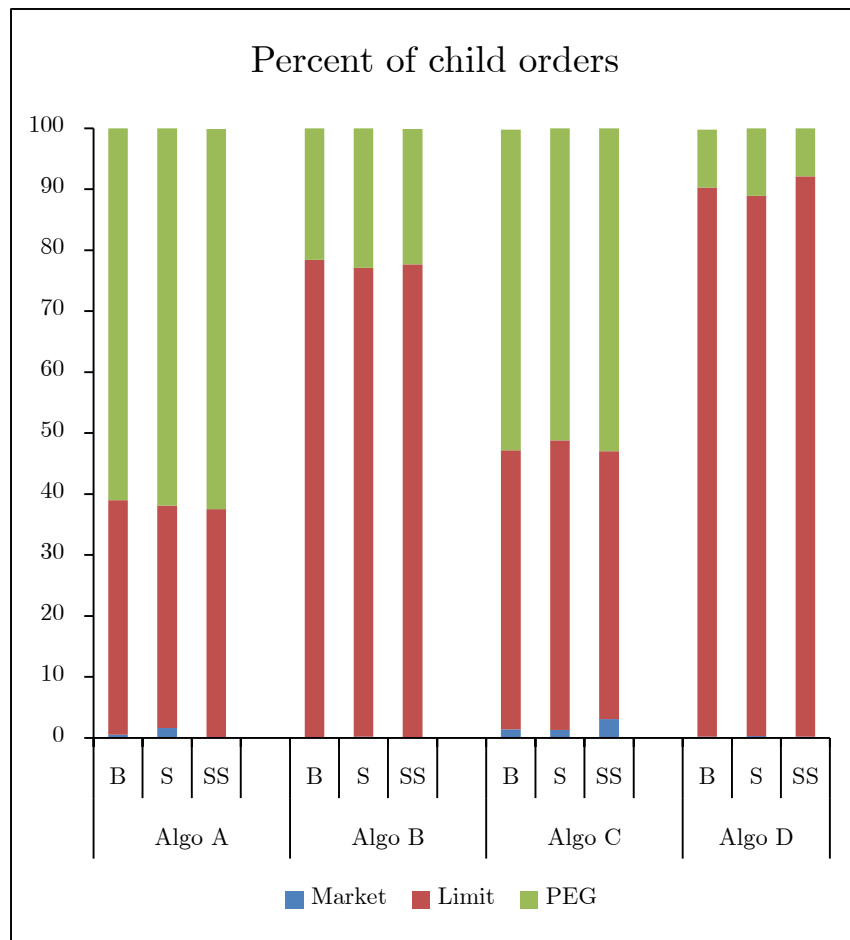
Cytology: Child Orders

Child orders and fills

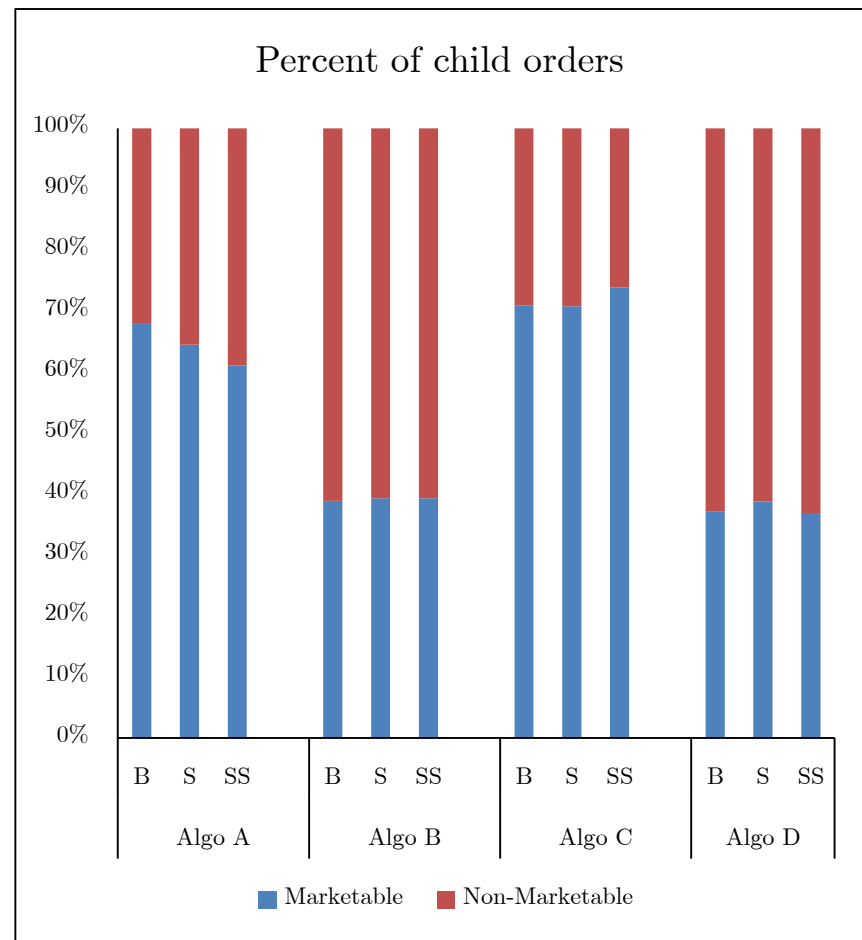


Child order characteristics

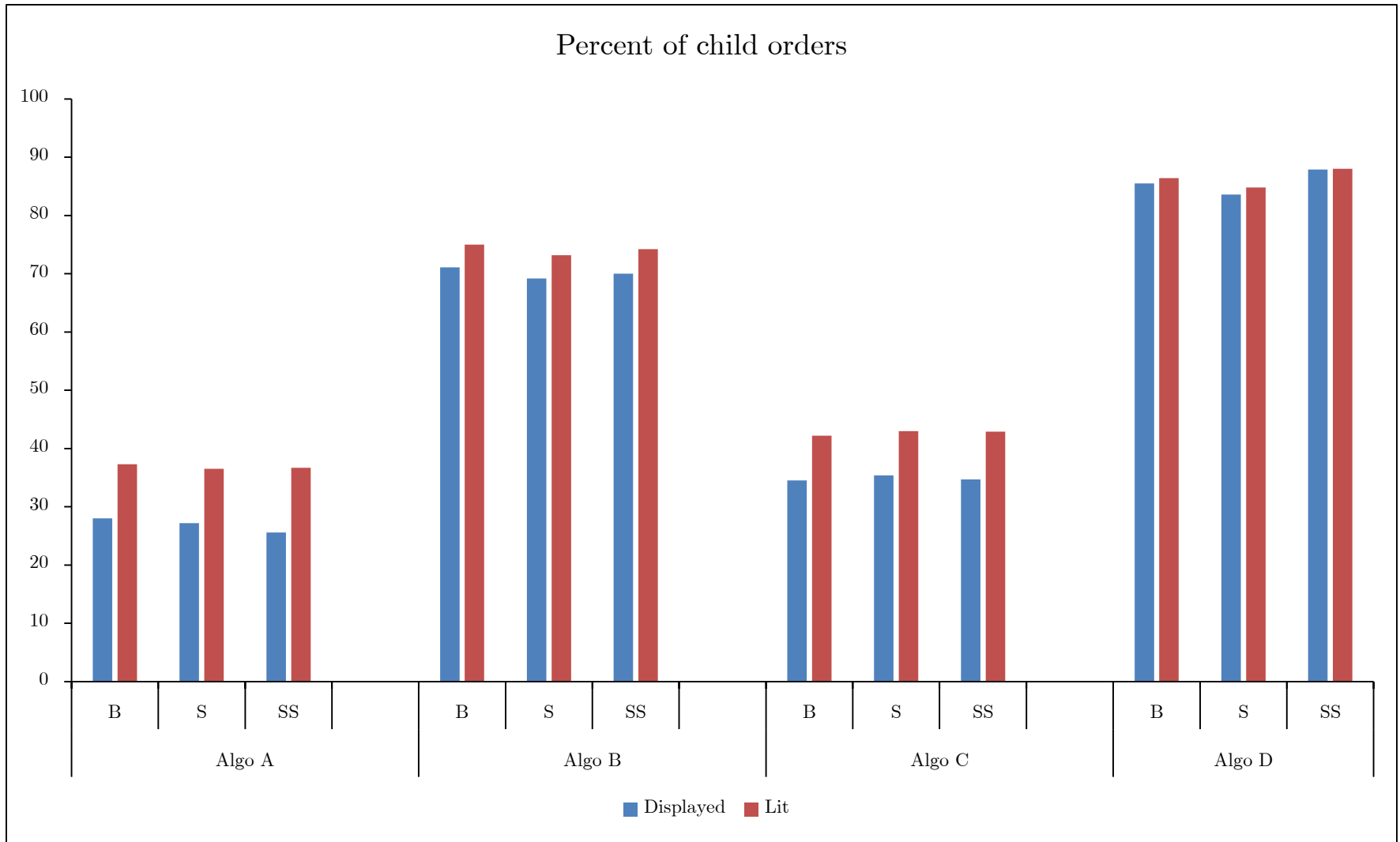
Order type



Marketability



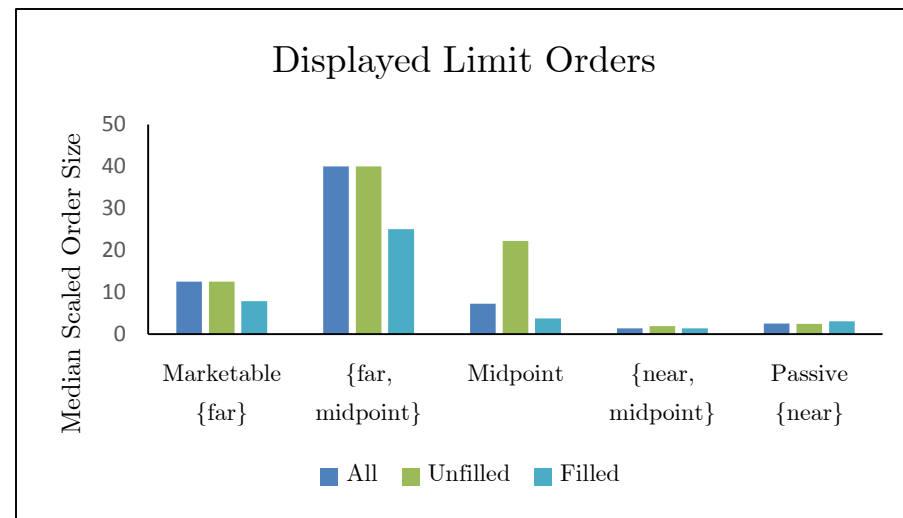
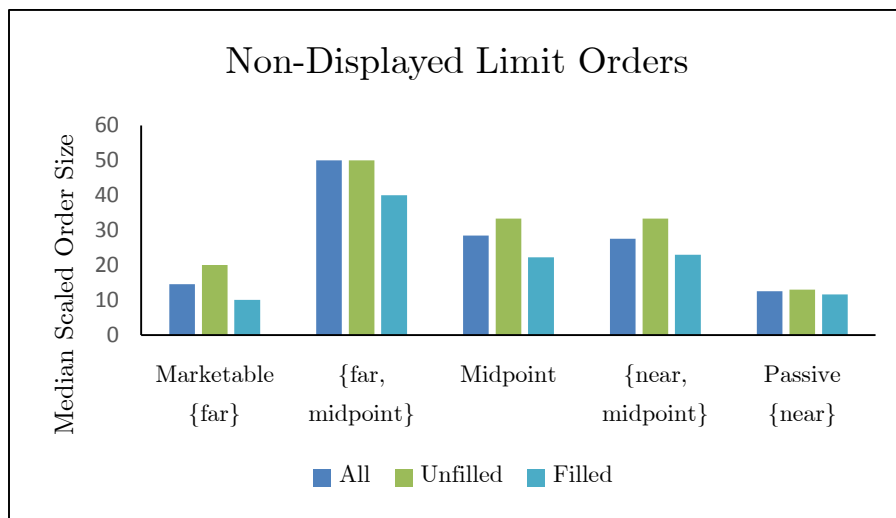
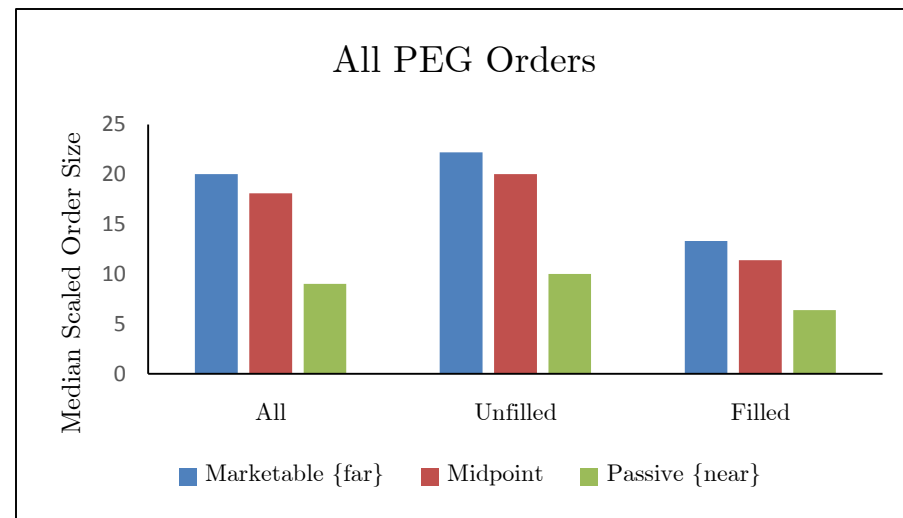
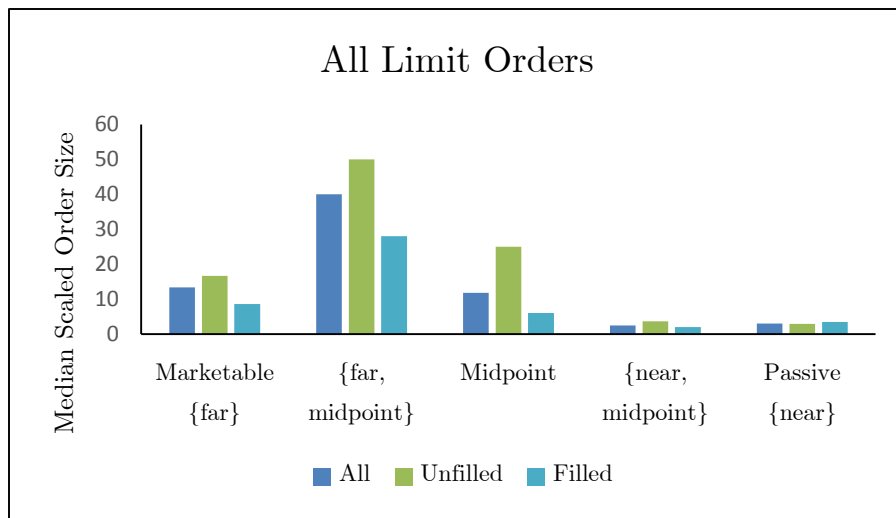
Display and Venue



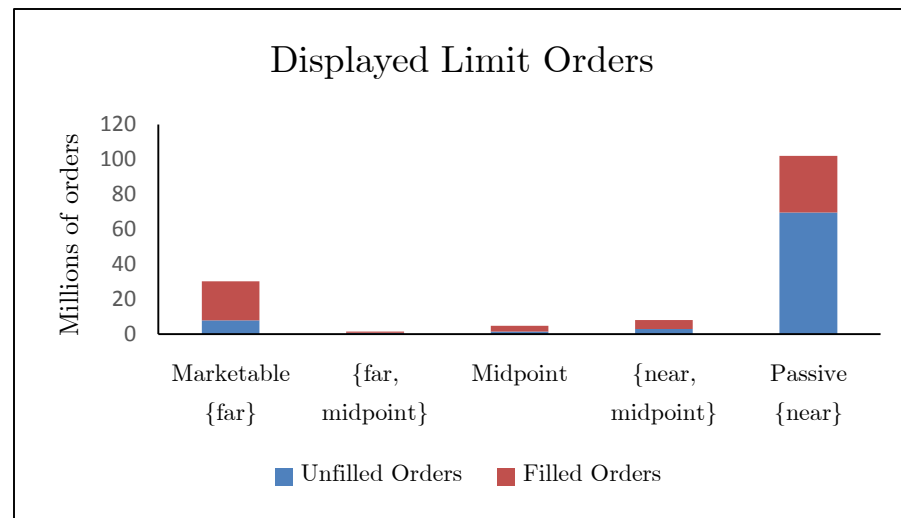
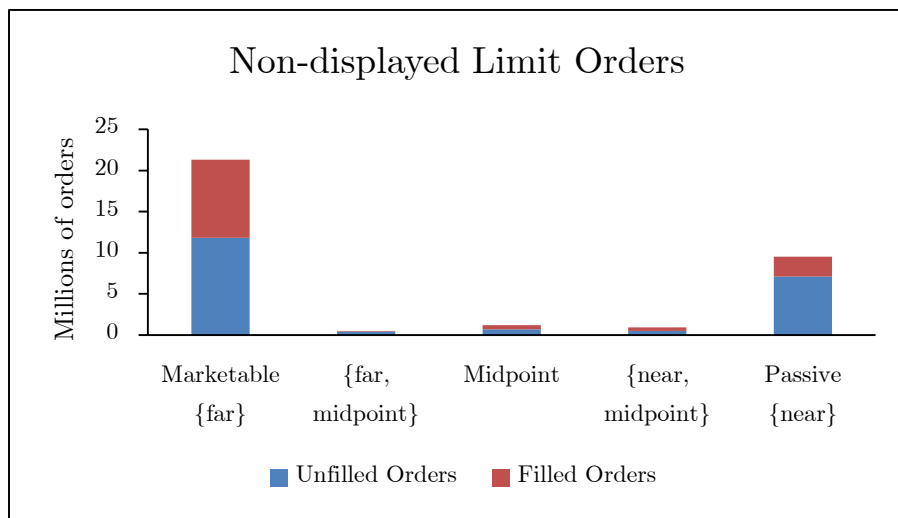
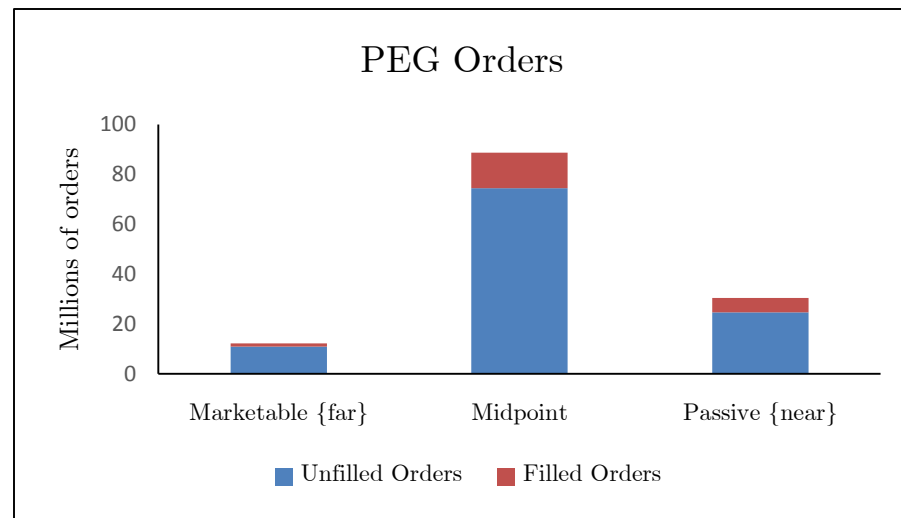
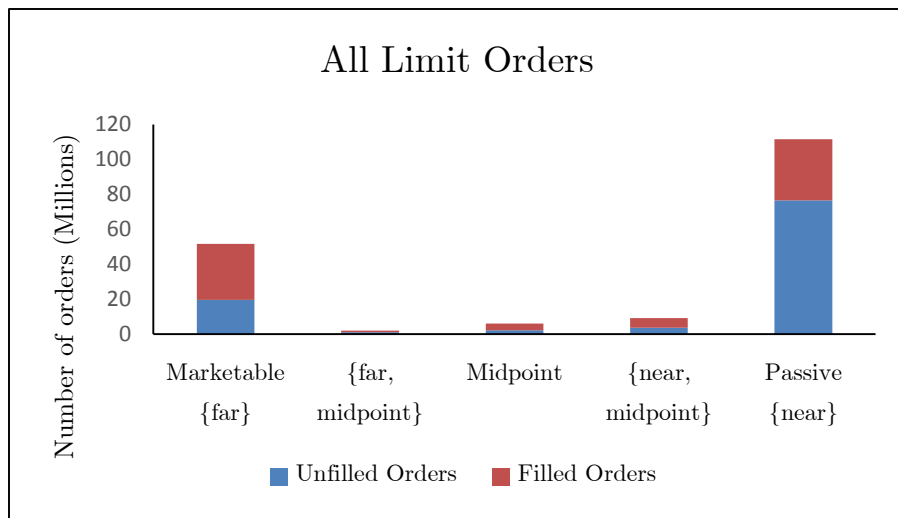
Choices and Outcomes

1. Choices
 - a) Price
 - b) Size
 - c) Display
2. Outcomes: Execution Risk vs Transaction Costs
 - a) Fill Rates
 - b) Time to fill
 - c) Cost

Choices



Outcome: Fill Rates



Outcome: time to fill

- For ex post filled orders:
 - $\text{Time-to-fill}(\text{Passive}) \gg \text{Time-to-fill}(\text{marketable})$
- For marketable but ex post unfilled orders:
 - $\text{Time-to-fill} \sim \text{Time-to-cancel}$
- For passive orders, time to fill is about $\frac{1}{2}$ of time to cancel
- This is useful, but
 - Cancellations affect the conditional distribution of time-to-fill, i.e. censoring.
 - And, execution risk is a function of endogenous choice variables and random market conditions
- Accelerated Failure Time (AFT) models (Lo, MacKinlay, Zhang (2002)). Model T (the life of the order).

Accelerated Failure Time models

	Limit orders			PEG orders	
Intercept (Marketable)	1.599	1.603	1.601	4.330	4.388
{far, midpoint}	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)
Midpoint	1.628	1.610	1.628	-	-
{near, midpoint}	(0.003)	(0.003)	(0.003)	-	-
Passive	3.489	3.566	3.490	-0.459	-0.457
Scaled order size	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Book Asymmetry	5.954	6.050	5.955	-	-
Buy indicator	(0.001)	(0.001)	(0.001)	-	-
Display order indicator	8.420	8.511	8.420	3.705	3.708
Lit indicator	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Log (market cap)	-0.036	-0.030	-0.040	0.010	0.010
Log (volume)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R _{-5,0}	-	-	-0.010	-	-0.010
Scale parameter (σ)	-	-	(0.000)	-	(0.000)
Shape parameter (ν)	-	-	-0.003	-	-0.101
N (censored)	-0.774	-	(0.001)	-	(0.001)
N (uncensored)	(0.000)	-	(0.001)	-	-
	-0.078	-0.076	-0.079	-0.033	-0.033
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
	-0.186	-0.186	-0.187	-0.185	-0.186
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	-0.035	-0.035	-0.036	0.002	0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	3.416	3.422	3.416	7.136	7.136
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
	-0.786	-0.788	-0.786	-3.141	-3.140
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)

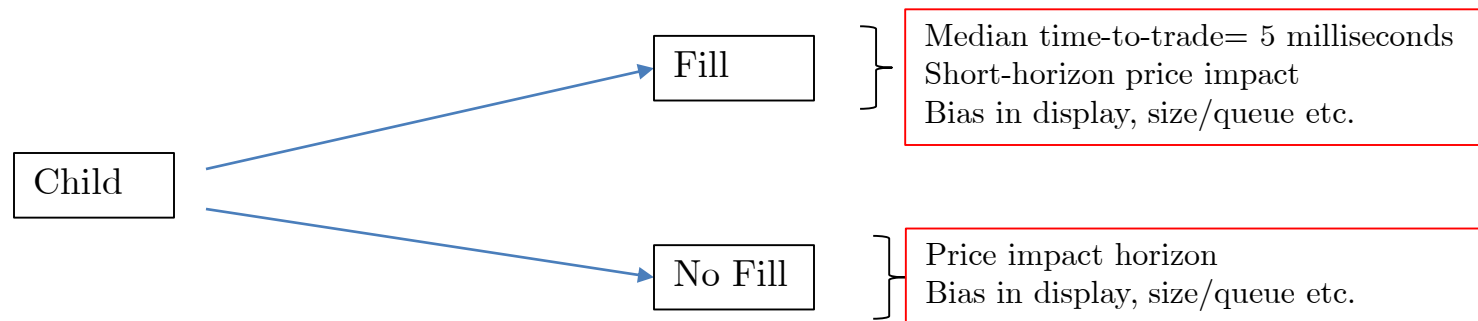
Change in price aggressiveness from marketable to passive orders increases uncensored time-to-fill from 1 ms to 450 seconds

Other results:
 - Effects of order size and book asymmetry are small
 - Trading in large stocks, higher volume and higher volatility shorten time to fill

Displaying order halves the time to fill

Price Movements

- Canonical theory: **trades** incorporate private information and move prices
- Does expression of **trading interest** move prices?
- Passive quotes **supposedly** liquidity providing and do not contain private information
 - Brogaard, Hendershott, Riordan (2019)



Measurement

- Child price impact

$$cpi_{jt\tau} = q_{jt} (m_{j,t+\tau} - m_{jt}) / m_{jt}$$

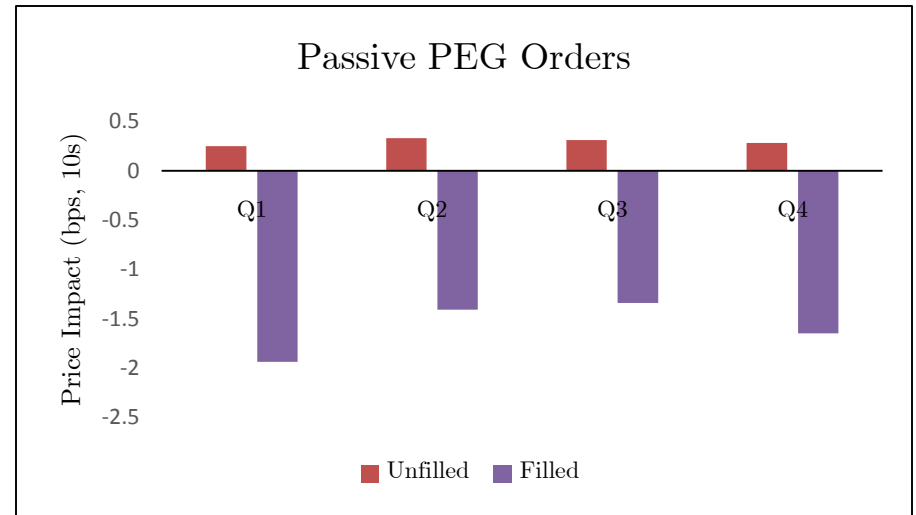
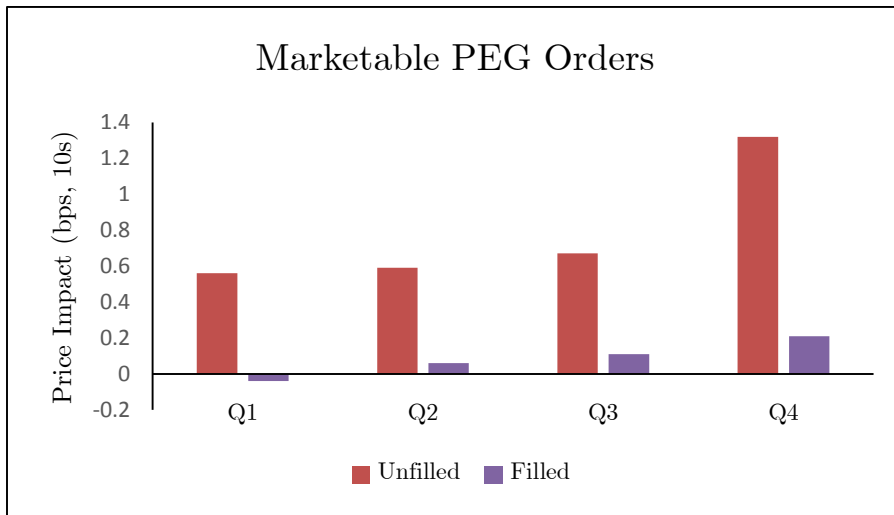
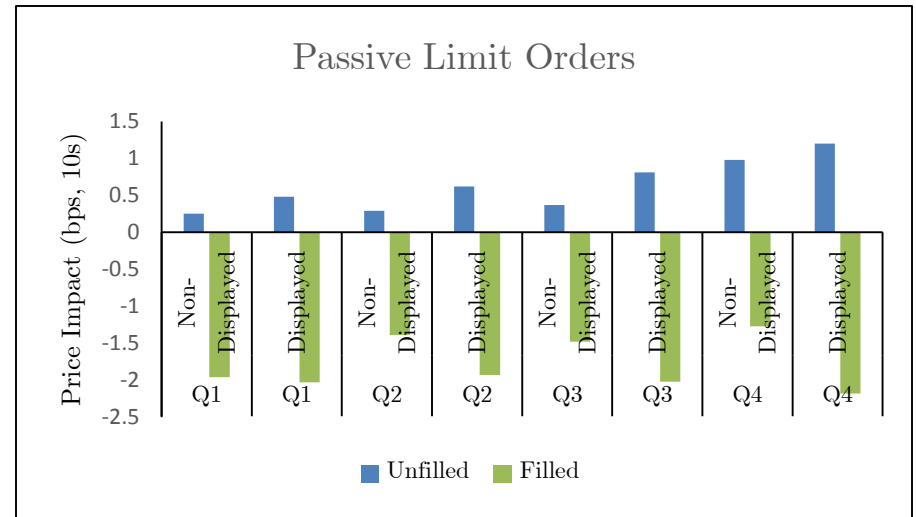
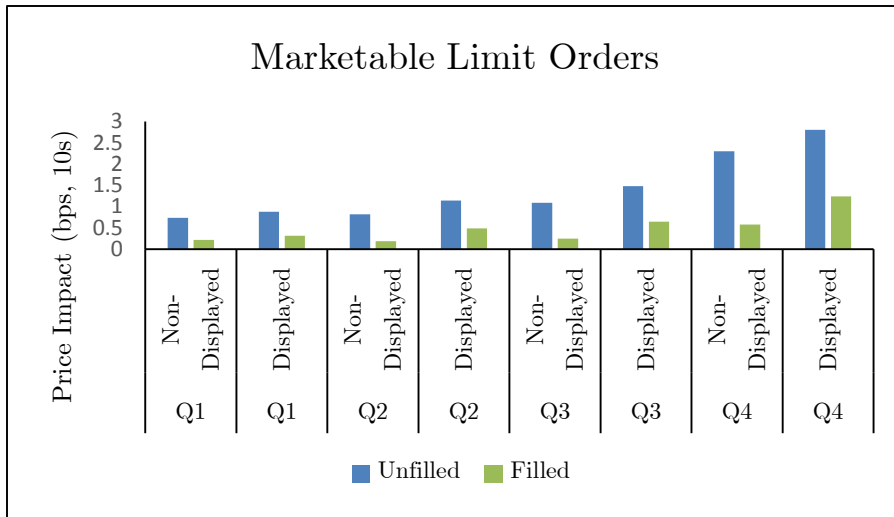
- For child submissions, t is submission time.
- For fills, t is trade time.
- Horizon (τ) issues as in Conrad and Wahal (2020).
 - $\tau = 100$ ms, 500 ms, 1 second, 5 second, 10 seconds post submission/fill

Price Impact

Submission Price	Unfilled Child Limit Orders Price Impact in bps					Filled Child Limit Orders Price Impact in bps				
	100ms	500ms	1 s	5 s	10 s	100ms	500ms	1 s	5 s	10 s
Aggressive (Far)	1.19	1.25	1.28	1.66	2.03	0.41	0.41	0.43	0.64	0.80
{Far, Midpoint}	0.57	0.67	0.71	1.25	1.72	-0.12	-0.15	-0.16	0.09	0.26
Midpoint	0.30	0.33	0.36	0.73	1.06	-0.43	-0.49	-0.55	-0.75	-0.83
{Near, Midpoint}	0.67	0.77	0.85	1.43	1.97	-0.91	-1.08	-1.27	-1.75	-2.02
Primary (Near)	0.08	0.14	0.18	0.51	0.84	-1.04	-1.23	-1.44	-1.82	-2.01

Submission Price	Unfilled Child PEG Orders Price Impact in bps					Filled Child PEG Orders Price Impact in bps				
	100ms	500ms	1 s	5 s	10 s	100ms	500ms	1 s	5 s	10 s
Aggressive (Far)	0.25	0.51	0.54	0.82	1.06	-0.06	-0.04	-0.05	0.05	0.14
P at Midpoint	0.28	0.40	0.42	0.59	0.76	-0.21	-0.25	-0.31	-0.49	-0.60
Primary (Near)	0.04	0.06	0.07	0.17	0.31	-0.79	0.89	-1.00	-1.29	-1.45

Price Impact: Triple Sort



Regressions: Limit Orders

	Unfilled		Filled	
Intercept	1.215	1.360	0.388	0.490
(Marketable)	(0.017)	(0.028)	(0.008)	(0.009)
{far, midpoint}	-0.634	-0.555	-0.646	-0.573
	0.028)	(0.031)	(0.028)	(0.031)
Midpoint	-0.894	-0.912	-1.404	-1.404
	(0.019)	(0.020)	(0.050)	(0.032)
{near, midpoint}	-0.355	-0.403	-2.372	-2.407
	(0.025)	(0.025)	(0.025)	(0.026)
Passive {near}	-1.110	-1.220	-2.458	-2.502
	(0.016)	(0.020)	(0.008)	(0.026)
Display	0.170	0.224	0.217	0.226
	(0.010)	(0.018)	(0.008)	(0.007)
Scaled size	0.099	-	0.151	-
	(0.004)		(0.006)	
Book Asymmetry*Buy	-	0.002	-	0.001
		(0.000)		(0.000)
Display * Scaled size	0.005	-	-0.037	-
	(0.001)		(0.007)	
 Ret_{-5,0} 	0.104	0.117	-0.024	-0.017
	(0.001)	(0.014)	(0.001)	(0.001)
Buy	0.026	0.004	0.006	0.001
	(0.012)	(0.014)	0.007)	(0.008)
Average N	63,815	63,815	45,125	45,125
Average adj-R²	0.103	0.076	0.110	0.104

Regressions: PEG Orders

	Unfilled		Filled	
Intercept (Marketable) {far, midpoint}	0.581 (0.022)	0.688 (0.026)	0.093 (0.015)	0.110 (0.016)
	-	-	-	-
Midpoint {near, midpoint}	-0.160 (0.025)	-0.199 (0.026)	-0.557 (0.015)	-0.549 (0.016)
	-	-	-	-
Passive {near}	-0.508 (0.022)	-0.585 (0.025)	-1.357 (0.019)	-1.660 (0.019)
Scaled size	0.060 (0.003)	-	0.038 (0.005)	-
Book Asymmetry*Buy	-	0.003 (0.000)	-	0.001 (0.000)
 Ret_{-5,0} 	0.058 (0.001)	0.063 (0.001)	-0.042 (0.002)	-0.041 (0.002)
Buy	0.011 (0.009)	-0.002 (0.010)	0.006 (0.008)	0.015 (0.010)
Average N	66,478	66,478	12,228	12,228
Average adj-R²	0.038	0.029	0.037	0.027

Histology: Strategic Runs

Strategic Runs

- The sequence of child orders is not random (à la Hasbrouck and Saar (2013))
 - Common goal
 - Shared codebase
- Run: a sequence of child orders from the same parent in the same price aggressiveness category
- Collapse categories into:
 - Marketable
 - Inside
 - Passive
- Parents that seek at least 1 basis point of ADV and with at least 50 child orders

Description & Transition Matrices

	All Algos	Algo A	Algo B	Algo C	Algo D
# parents	812,132	50,582	82,903	87,751	590,896
Runs/parent	63.09	53.59	51.65	41.23	68.76
Child per run	8.84	11.92	7.45	8.16	8.87
Run duration	566.99	102.04	452.81	158.58	683.46
Run volume	1515.26	5503.08	3024.40	3063.60	732.23
Percent runs					
Passive	45.58	40.58	42.78	42.00	46.94
Inside spread	16.04	16.34	19.98	10.94	16.21
Marketable	38.37	43.06	37.23	47.04	36.84

	All Algos			Algo A			Algo B			Algo C			Algo D		
	P_t	I_t	M_t	P_t	I_t	M_t	P_t	I_t	M_t	P_t	I_t	M_t	P_t	I_t	M_t
P_{t-1}	-	13.9	31.2	-	10.3	33.0	-	17.0	27.2	-	3.0	41.1	-	14.6	30.7
I_{t-1}	12.9	-	5.2	9.2	-	7.1	13.7	-	6.6	2.8	-	6.9	13.9	-	4.8
M_{t-1}	29.3	7.4	-	31.3	9.1	-	25.7	9.8	-	38.4	7.8	-	28.9	7.0	-

Logistic Regressions

- Does the price aggressiveness of the run depend on
 - Whether the prior run received a fill (Fill_{t-1})
 - “Cost” of prior run: cumulative signed midpoint return (SRet_{t-1})
- Dependent variables: $\text{Prob}(\text{Marketable})$, $\text{Prob}(\text{Passive})$

	Algo A			Algo B			Algo C			Algo D		
	Pr(M _t)	Pr(M _t)	Pr(P _t)	Pr(M _t)	Pr(M _t)	Pr(P _t)	Pr(M _t)	Pr(M _t)	Pr(P _t)	Pr(M _t)	Pr(M _t)	Pr(P _t)
	P	I	M	P	I	M	P	I	M	P	I	M
Run _{t-1}												
Fill _{t-1}	0.52	-0.07	0.06	0.02	-0.37	0.09	0.20	-0.02	0.36	-0.04	-1.35	0.65
	(0.00)	(0.01)	(0.05)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
	[7.6]	[1.7]	[0.9]	[0.5]	[-8.3]	[1.7]	[1.1]	[-0.3]	[5.1]	[0.8]	[26.9]	[11.2]
SRet _{t-}	0.01	-0.02	-0.01	0.01	-0.01	-0.00	0.01	-0.01	-0.01	0.00	-0.01	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	[0.1]	[0.4]	[-0.1]	[0.0]	[0.1]	[0.0]	[.04]	[0.0]	[-0.1]	[0.0]	[0.2]	[0.0]

Conclusions

- Central tradeoff: trading versus incurring trading costs. Thus the use of algorithms.
- To do so, they generate hundreds of child orders
- Child orders employ price, time, display priority rules to navigate the tradeoff
- Passive orders have much larger execution risk but still incur substantial price impact
- Marketable orders do not guarantee execution and generate even larger price impact