The PGA Tour and Trading

Peter Cotton

Overview

Quiz: What happens Monday October 7 through Sunday October 13th 2019?

Hello Houston. What do these players have in common?

1. Patrick Reed
2. Jhonny Vegas
3. Chris Stroud
4. Michael Kim
5. Shawn Stefani

Golf and Trading

Hello Houston. What do these players have in common?

1. Patrick Reed
2. Jhonny Vegas
3. Chris Stroud
4. Michael Kim
5. Shawn Stefani

Answer: All are PGA tour players based in Houston.

This is a talk about trading ... sort of ... or

...how to leave your bad decisions for the golf course.

PGA golfer analysis

Professional golf had its Moneyball moment several years ago, largely thanks to NYU Professor Mark Broadie working with Steve Evans from the tour. His work made possible accurate decomposition of performance as shown in Figure 1

Prior to this, the PGA reported largely meaningless statistics such as the average number of putts per round.[1]

[1] For instance, Dustin Johnson would be considered a good putter because he takes fewer putts than most. However this is due to his accuracy in approach shots. In turn, his accuracy in approach shots relates to the fact that he bombs it off the tee and has a shorter club in than other players. Clearly, the old school "putting" statistic is in fact commingling every aspect of the game.

Broadie nudged the tour towards a more logical approach that would typically be referred to as a "value function" in optimal control or reinforcement learning contests. Given any position on the golf hole, we determine the typical number of shots taken by all professionals to finish the hole (circled numbers in Figure 2)

$$\nu(\cdot) = Average\ number\ of\ shots\ to\ finish\ hole \tag{1}$$

Think of $\nu()$ as a function defined for every blade of grass (or water, or sand). Then the "strokes gained" statistics for every shot taken is simply the difference between the starting and ending values of $\nu()$, as demonstrated in Figure 2.

$$strokes\ gained = \nu(before\ shot) - \nu(after\ shot) - 1 \tag{2}$$

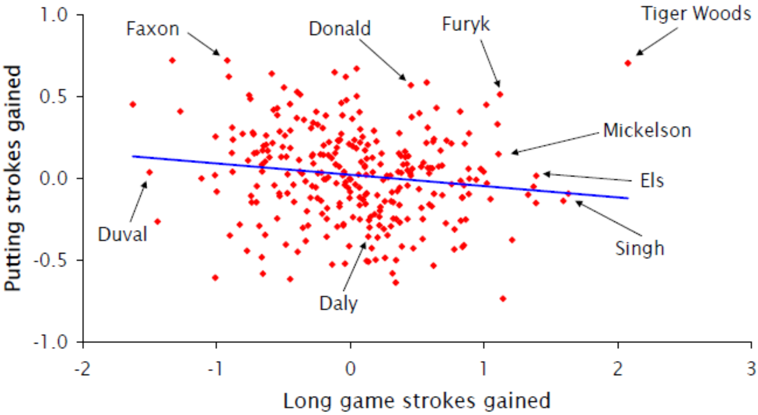

Figure 1: Meaningful breakdown of professional golfer performance

Figure 2: A value function $\nu()$ is assigned to every point on a golf hole, and indicates the typical number of shots required to finish from that point. The putt requires an average of 0.7 shots. The position on the fairway requires 1.8. Relative to average performance, the player has gained 0.1 shots with his approach shot, then 0.3 shots by draining the putt.

SG: TEE-TO-GREEN

Y-T-D-statistics through: **Through Week Ending: 12/19/2016, Dec 18, 2016**

Titleist indicates golf ball usage

RANK THIS WEEK	RANK LAST WEEK	PLAYER NAME	ROUNDS	AVERAGE	SG:OTT	SG:APR	SG:ARG	MEASURED ROUNDS
1	1	Kevin Chappell	10	4.091	1.203	1.888	1.000	1
2	2	Jhonattan Vegas	14	3.044	1.097	1.795	.151	2
3	3	Jim Furyk	6	2.341	.003	2.119	.219	3
4	4	Francesco Molinari	8	2.264	.144	1.866	.254	4
5	5	Anirban Lahiri	12	2.211	.621	1.918	-.328	3

Figure 3: "Strokes gained" permits useful aggregation of various kinds (e.g. by shot type). Shown are breakdowns "off the tee (OTT)", "approach the green (APR)" and "around the green (ARG)".

Completing the analogy

Moral: You cannot have Figure 1 without the work implicit in Figure 2.

PGA Strokes Gained methodology" $\mapsto$ "Ockham"

Position on course $\mapsto$ Trading position.[2]

Strokes gained $\mapsto$ "Mid" P/L + "Ockham inventory reserve" = "dollars gained"

[2] Inventory or projected inventory. Later we can add second order "state" as well.

Assessing flow trading

Trading analysis: Bloomberg

Figure **??** aggregates trader performance in Bloomberg reponses against an automatically generated benchmark.

	MATHIEU VINCENT	SEMARIA IZUZQU	NA MICHAILO	NS GRUEN	IVIER FLAND	RICHARD CATALL	NITIN MAGGON	BASTIAN GREULI	LEAH VIAULT	JUAN DE LE
Algo improvement (per trade)	900.71	721.79	405.71	275.3741	261.5415	220.5625328	114.9137986	88.5963588	84.46993944	58.25012
Algo improvement (total)	174,736.88	47,638.02	76,678.62	31392.65	46815.928	12351.50184	4826.379541	9214.021315	2111.748486	1631.003
Algo improvement - when algo tr	562.85	678.77	272.62	207.4988	76.466536	0	37.26880952	195.3432692	68.786	103.1867
Algo improvement - when algo tr	109,192.43	44,799.02	51,525.70	23654.86	13687.51	0	1565.29	20315.7	1719.65	2889

Figure 4: Value added by algo: by salesperson

	MATHIEU VINCENT	SEMARIA IZUZQU	NA MICHAILO	NS GRUEN	IVIER FLAND	RICHARD CATALL	NITIN MAGGON	BASTIAN GREULI	LEAH VIAULT	JUAN DE LE
Algo improvement (per trade)	900.71	721.79	405.71	275.3741	261.5415	220.5625328	114.9137986	88.5963588	84.46993944	58.25012
Algo improvement (total)	174,736.88	47,638.02	76,678.62	31392.65	46815.928	12351.50184	4826.379541	9214.021315	2111.748486	1631.003
Algo improvement - when algo tr	562.85	678.77	272.62	207.4988	76.466536	0	37.26880952	195.3432692	68.786	103.1867
Algo improvement - when algo tr	109,192.43	44,799.02	51,525.70	23654.86	13687.51	0	1565.29	20315.7	1719.65	2889

Figure 5: Value added by algo: by salesperson

Tentative results

- NA.RETAIL universe (about 11,000 bonds)
- Value added by traders seems to correlate with the degree to which bonds are traded in spread, with about ten dollars per opportunity added. This advantage may disappear when mix filter model is included in this analysis.

- The dataset is subject to various biases. It represents about 10-20 percent of all trades. Nonetheless, the exact same trades are used for algo and trader analysis.

Understanding trader decisions

- While the aggregate result is potentially interesting as a comparison of research algos versus trading, the real interest here lies in the ability of algos to improve trader performance. (c.f. Chess computers circa 1990).

- Trader leaks, trader characterization, and even trader coaching is made possible by looking at back-testing in a new, granular manner. Whereas swapping out algos for traders right now might actually reduce P/L (at least for spready bonds), analysis on a trade by trade basis will likely increase it. See Figure 6

But ... this requires some head scratching. How to analyze one decision in isolation??

	CONSERVATIVE_MARKUP	CONSERVATIVE_MID	CONSERVATIVE_PNL_GAINED	CONSERVATIVE_PNL_LOST	CONSERVATIVE_PNL_MARKUPS	CONSERVATIVE_PNL_TOTAL
0	0.288520556777	105.502542342	0	0	-48.47069609759557	-48.47069609759557
1	0.346402402278	106.363755342	0	0	0	0
2	0.277026773013	105.795936325	0	0	0	0
3	0.182596548624	102.313132677	0	0	7.384638714446794	7.384638714446794
4	0.199491490066	100.050023177	0	0	0	0
5	0.166311085557	100.070072897	0	0	0	0
6	0.205377720127	101.671430633	229.99999999998977	0	0	229.99999999998977
7	0.201943648425	100.340825397	0	322.309542099	0	322.30954209865104
8	0.234141809864	100.6038555	0	81.6182917578	0	81.61829175775779
9	0.233294864203	102.262946143	1039.999999999992	0	0	1039.999999999992
10	0.174421205733	100.91961653	-29.49999999999875	0	0	-29.49999999999875
11	0.220525563499	100.519374962	-15.959999999999468	0	0	-15.959999999999468
12	0.195063874997	100.892398968	0	0	0	0
13	0.174984670185	102.731990219	0	0	0	0
14	0.1437565937	102.909938709	0	0	521.0453655523721	521.0453655523721
15	0.169788575104	102.89290097	0	0	0	0
16	0.123295262829	101.824151739	0	-197.228704836	0	-197.22870483627162
17	0.128738060067	102.413538768	0	0	0	0
18	0.130942323247	101.328259307	-273.74999999999613	0	0	-273.74999999999613
19	0.0728080749593	108.853701707	0	0	-93.33831391410428	-93.33831391410428
20	0.0787238611638	108.755348096	0	0	172.75152979277664	172.75152979277664
21	0.0664652240078	108.491308745	0	0	0	0
22	0.0521450426015	107.619259326	0	0	96.2845731608823	96.2845731608823
23	0.0793810207448	107.661130164	0	0	0	0
24	0.295163628895	109.537930724	0	0	122.46580931790162	122.46580931790162
25	0.228398214536	109.954182977	-1478.999999999999	0	0	-1478.999999999999
26	0.0864065817247	110.17021105	0	0	0	0

	CONSERVATIVE_RESPONSE	COVER_RESPONSE	EXIT_LEVEL	INSIDE_RESPONSE	MATURITY_CAT	PRICE	QUANTITY	REPORTINGPARTYSIDE
0	105.214021785	105.025	105.556	105.025	5	103.663	95000	B
1	106.710157744	0	105.556	106.3	5	106.3	50000	S
2	106.072963098	0	105.375	105.708	5	105.708	10000	S
3	102.130536129	0	103.468	0	5	102.869	1000	B
4	99.850531687	0	100.6	100.38	6	100.38	150000	B
5	99.903761811	0	100.566	100.15	6	100.15	1000000	B
6	101.876808353	100.7	100.266	100.7	6	100.496	100000	S
7	100.542769046	0	100.575	100.776	6	100.776	1000000	S
8	100.83799731	0	100.85	100.9	6	100.9	680000	S
9	102.029651279	102.308	102.53	102.308	6	102.322	500000	B
10	100.745195325	101.031	101.28	101.031	6	101.339	50000	B
11	100.298849398	101.128	101.492	101.128	6	101	12000	B
12	100.697335093	0	101.727	101.633	6	101.633	50000	B
13	102.557005549	0	103.625	103.75	6	103.75	85000	B
14	102.766182115	0	103.595	0	6	104.503	30000	B
15	102.723112395	0	103.125	103.125	6	103.125	25000	B
16	101.700856476	0	102.687	101.58	6	101.58	20000	B
17	102.284800708	0	102.687	102.75	6	102.75	1000000	B
18	101.197316983	101.5	101.5	101.5	6	101.865	75000	B
19	108.780893632	108.38	108.8	108.38	3	108.692	105000	B
20	108.676624235	108.579	108.907	108.579	3	108.763	200000	B
21	108.424843521	0	108.8266	108.73	3	108.73	15000	B
22	107.567114283	107.5	107.713	107.5	3	107.713	66000	B
23	107.581749143	0	107.375	107.625	3	107.625	40000	B
24	109.242767095	109.15	109.846	109.15	3	109.304	200000	B
25	109.725784762	109.865	108.5	109.865	3	109.979	100000	B
26	110.083804469	0	111.075	111.129	4	111.129	100000	B

	SIZE_CAT	SYMBOL_NAME	TICKER	TRADER	TRADED	TRADER_MARKUP	TRADER_RESPONSE	UTCTIME	VENUE
0	1	013817AJ0	AA	tmcintyre	1	0.33954234214323264	105.163	2014-08-01 16:56:28	BLOOMBERG_BOND
1	1	013817AJ0	AA	tmcintyre	0	0.21424465787387703	106.578	2014-08-04 13:44:29	BLOOMBERG_BOND
2	0	013817AJ0	AA	tmcintyre	0	0.6170636746792155	106.413	2014-08-06 19:28:25	BLOOMBERG_BOND
3	0	013817AJ0	AA	tmcintyre	1	-0.5558673228204185	102.869	2015-08-04 14:36:27	BLOOMBERG_BOND
4	2	013817AK7	AA	tmcintyre	0	0.18602317704338134	99.864	2014-08-07 15:17:18	BLOOMBERG_BOND
5	2	013817AK7	AA	tmcintyre	0	0.020072896550416885	100.05	2014-08-12 12:31:07	BLOOMBERG_BOND
6	1	013817AK7	AA	tmcintyre	1	-1.1754306327575108	100.496	2014-08-12 16:31:20	BLOOMBERG_BOND
7	2	013817AK7	AA	tmcintyre	0	0.517174602635194	100.858	2014-08-13 14:10:21	BLOOMBERG_BOND
8	2	013817AK7	AA	tmcintyre	0	0.5811444998287385	101.185	2014-08-13 14:34:22	BLOOMBERG_BOND
9	2	013817AK7	AA	tmcintyre	1	-0.05905385697158749	102.322	2014-09-08 12:58:53	BLOOMBERG_BOND
10	1	013817AK7	AA	tmcintyre	1	-0.4193834696665988	101.339	2014-09-10 16:14:49	BLOOMBERG_BOND
11	0	013817AK7	AA	tmcintyre	1	-1.1056250383127662	101.625	2014-10-01 18:43:14	BLOOMBERG_BOND
12	1	013817AK7	AA	cjai	0	-0.7316010318408956	101.624	2014-10-02 08:20:22	BLOOMBERG_BOND
13	1	013817AK7	AA	tmcintyre	0	-0.22600978092194168	102.958	2014-11-03 13:11:05	BLOOMBERG_BOND
14	1	013817AK7	AA	ddavis	1	-1.5930612914745979	104.503	2014-12-01 08:17:05	BLOOMBERG_BOND
15	1	013817AK7	AA	tmcintyre	0	0.23890097024241186	102.654	2014-12-05 16:56:29	BLOOMBERG_BOND
16	0	013817AK7	AA	tmcintyre	0	-0.6068482613522264	102.431	2014-12-09 17:02:16	BLOOMBERG_BOND
17	2	013817AK7	AA	tmcintyre	0	0.1815387680040743	102.232	2014-12-10 14:31:42	BLOOMBERG_BOND
18	1	013817AK7	AA	tmcintyre	1	-0.5367406932849264	101.865	2014-12-12 13:07:01	BLOOMBERG_BOND
19	2	013817AL5	AA	tmcintyre	1	0.161701707258473	108.692	2014-08-12 12:28:46	BLOOMBERG_BOND
20	2	013817AL5	AA	tmcintyre	1	-0.0076519037325795125	108.763	2014-08-13 13:49:59	BLOOMBERG_BOND
21	0	013817AL5	AA	tmcintyre	0	-0.09569125485148788	108.587	2014-09-03 12:41:42	BLOOMBERG_BOND
22	1	013817AL5	AA	tmcintyre	1	-0.09374067430896105	107.713	2014-12-04 20:01:23	BLOOMBERG_BOND
23	1	013817AL5	AA	tmcintyre	0	0.20713016350306646	107.454	2014-12-05 15:41:05	BLOOMBERG_BOND
24	2	013817AP6	AA	tmcintyre	1	0.23393072423579042	109.304	2014-08-04 13:44:58	BLOOMBERG_BOND
25	1	013817AP6	AA	tmcintyre	1	-0.024817023141494587	109.979	2014-09-11 15:39:53	BLOOMBERG_BOND
26	1	013817AQ4	AA	tmcintyre	0	-0.6057889495431965	110.776	2014-09-03 18:18:37	BLOOMBERG_BOND

Figure 6: Trade by trade analysis

The problem with ad-hoc RFQ analysis

Without theory, only half the story can be told - which is to say none of it.

	Algo trades	Algo misses
Trader trades	size((Trader markup) - (Algo markup))	?
Trader misses	?	0

Table 1: Economic impact of different responses

How theory solves that problem

We assume a trader markup of m and an algo markup of m^*.

	Algo trades	Algo misses
Trade	$s(m - m^*)$	$s(m - \epsilon) - (\nu(x \pm s) - \nu(x))$
Miss	$-s(m^* - \epsilon) + \nu(x \pm s) - \nu(x)$	0

Table 2: Economic impact of different responses

where $\nu(x)$ captures the impact of the change in inventory in a rather subtle manner. It is not merely funding and risk nor anything directly measurable.

Setup

Optimized trading in the context of repeated sealed-bid auctions.

Corporate flow-trading is an example, even though we know the following sequence is often an approximation of reality:

1. Customer inquiry is sent to multiple dealers
2. Dealers respond, not knowing others' responses
3. The customer takes the best price

Practical theoretical result #1: Optimal markups

In a certain stylized class of models a trader with current inventory x should respond to an RFQ of size s by marking up:

$$m = w + \epsilon + \frac{\nu(x-s) - \nu(x)}{s} \tag{3}$$

over the fair market price. This reads straightforwardly:

"markup = market width + adverse selection + marginal inventory cost"

It is a cute formula, by no means obvious, that masks a certain amount of financial and mathematical nuance.[3]

[3] It arises as the solution of an infinite time trading game.

Inventory indifference price

The central concept here is again $\nu(x)$. It is the biggest markdown a *fully rational trader* (i.e. using this theory) will accept if offered the chance to get out of her inventory x immediately.

In order to know the indifference markdown of the fully rational trader who is using this theory, you have to know how she is going to trade in the future (which also depends on $\nu(x)$ naturally, which is what we are trying to compute).

Circular? Yes. Such is life.

Aside: trading outcomes viewed from algo perspective

Substituting in the algo markup, the value added by the trader from the algo perspective is as follows:

	Algo trades	Algo misses
Trader trades	$s(m - m^*)$	$s(m - m^*) + sw$
Trader misses	$-sw$	0

Table 3: Economic impact of different responses from algo perspective

However, we don't necessarily use m^* in the algo as this is a little simplistic. It isn't necessarily the case that the algo makes sw when it trades so we'd usually use the previous formulation instead.

Comparison to direct costs

The trader's inventory indifference price $\nu(x)$ is not to be confused with the direct cost of holding a bond $c(x)$ (though it often is). The latter is viewed as a parameter by which the former is computed. For example:

$$c(x) = c_+ x^+ + c_- |x^-| + c_2 x^2$$

perhaps with parameters like

$$\begin{aligned} c_+ &= 0.045 \\ c_- &= 0.04 \\ c_2 &= \frac{0.1 \times 0.045}{10,000,000} = 4.51e^{-10} \end{aligned}$$

since we have both funding, risk costs and self-imposed equity return hurdles that should be internalized into the algorithm.

Perhaps:

$$c_+ = \frac{return\ on\ equity}{leverage} + funding = \frac{0.30}{8} + 0.07 \sim 4.5\%$$

Direct costs are the managerial lever here, so do with them what you like!

Aside continued: What's wrong with the "common sense" approach using direct costs $c(x)$ in place of $\nu(x)$?

One often hears straight-forward arguments for the cost of taking on a bond, such as mean holding time multiplied by cost and so forth.

These sound reasonable, but ignore the counteracting benefit of larger inventory. So they implicitly overstate the real cost of increasing inventory and, therefore, tend to be too defensive.

So no. You need to use $\nu(x)$ in

$$m = w + \epsilon + \frac{\nu(x-s) - \nu(x)}{s} \tag{4}$$

not $c(x)$.

Aside: Relating direct and indirect inventory cost

Here is why direct costs $c(x)$ are not synonymous with $\nu(x)$.

- In deciding on the true cost of moving from inventory x to $x+s$, an optimal trader will weigh not only the differential in direct costs proportional to $c(x+s) - c(x)$ but also the *differential value of trading opportunities at each inventory level*.

- The larger the inventory, the greater the value to the trader of a trading opportunity because getting out (or rather, the potential to get out) helps more. *This counteracts direct inventory costs*

Aside continued: A financial argument

One way to pursue the connection between $c(x)$ and $\nu(x)$ considers two possible courses of action by a trader. Either:

1. The market maker liquidates her holding and pays $\nu(x)$ to do so. She then waits until the next trade opportunity, makes a trading decision, and then again liquidates her position if necessary.
2. She defers liquidation until after the trading opportunity.

Both courses of action recombine into the same final inventory $x = 0$. So by definition of $\nu(x)$ as the indifference markup, both paths must incur the same costs. This leads us to a functional relation for $\nu(x)$, at least for a few tractable models. I omit the mathematical details.

Aside continued: Another quasi-intuitive angle

More translation from mathematics to English...

1. An optimal trader will accept a differential direct cost $c(x+s) - c(x)$ for the typical time between trading opportunities if she makes this up at the next trading opportunity. But what does she make up exactly?

2. Answer: her (non-myopic) differential gain equals the differential probability of executing a trade (she is more likely to execute at the higher inventory level because she is more aggressive) multiplied by the typical gain from trading (which, remarkably, is invariant to inventory and always equals sw).

Differential trading probabilities

We have arrived at a "trading law"

$$\textit{change in direct cost} = \textit{change in trading probability} \times \textit{size} \times \textit{width} \quad (5)$$

in which the true, non-myopic inventory indifferent cost $\nu(x)$ is buried inside the change in trading probability. Here width is the market width, not our market maker's width.

But $\nu(x)$ is interesting in its own right...

Skew and width from $\nu(x)$

Our magic formula works for both bid and ask. Combining them both we notice the following:

$$\begin{aligned} trader\ mid &= market\ mid - \frac{\overbrace{\nu(x+s) - \nu(x-s)}^{skew}}{2s} \\ trader\ ask - trader\ bid &= \Delta + \frac{\nu(x+s) - 2\nu(x) + \nu(x-s)}{2s} \end{aligned}$$

So we have a new interpretation of $\nu(x)$. The first difference of $\nu(x)$ is skew and the second difference is width.

Internalization

Because $\nu(x)$ is an indifference price for an optimal trader, it can also be used as an internal transfer price. This is true not just for bonds, incidentally, but also for block trades in other asset classes or anywhere we find something resembling a sequence of sealed-bid auctions.

Our profit optimizing market maker will really have two different bid-offer pairs. One for the outside world and one for inside. In practice however, it is up to management and the quants assisting them to determine this indifference price ... since the trader will not do it.[4]

[4] Sometimes policy dictates that internal trades be done at mid. This is not quite fair, and moreover distorts the work of the market maker. In contrast what we are suggesting here is an internal transfer price varying in inventory that does not nudge the outward facing market maker away from optimal behaviour.

Result #2: Inquiry imbalance

Market imbalance impacts optimal skew and width, obviously.

1. Suppose the probability that a client inquiry is a client selling inquiry is p.
2. We re-compute the optimal long term strategy, and compare it to the case $p = \frac{1}{2}$.

The surprising result is that the optimal trading adjustments are very simple formulas that don't depend on $\nu(x)$.

Faced with imbalance parameter p a trader should skew by an *additional* amount

$$\delta = \frac{\log(1-p) - \log(p)}{2h}$$

and widen by an *additional* amount:

$$\gamma = \frac{\log\left(\frac{1}{2}\right) - \frac{\log(1-p)+\log(p)}{2}}{h}$$

compared with the optimal choice of skew and width in the case $p = 1/2$.

The astute reader will notice that $\gamma > 0$ in the expression for the width adjustment given above - in keeping with the intuition that an optimal trader needs to be wider, all else being equal, if customer inquiry is imbalanced.

Result #3: The mid doesn't matter

Say whaaatttttt ??!!!!

Again, referencing a stylized RFQ model with exponential markups we discover that a sufficient statistic is the mean of the best price disseminated by other parties.

Our market maker needs to know the sum of the market's mid and width. But she does not need to estimate these individually.

Remark on the (first order) irrelevance of the mid

If you thought the logical approach to building algos was

1. Build mid models

2. Build skew models on top

then yeah so did I once upon a time. But I think the math is trying to tell me something!

Again, this is only true for a stylized model and one can easily break the coincidence.

Result #4: The rule of three

A ten percent increase in accuracy equates to a thirty percent increase in the monetary value to the market maker of each trading opportunity (holding the strategy constant).

There are no diminishing returns on accuracy

For this to hold, the trader must take their own uncertainty into account - unless it is small compared to the market width. Here's a picture...

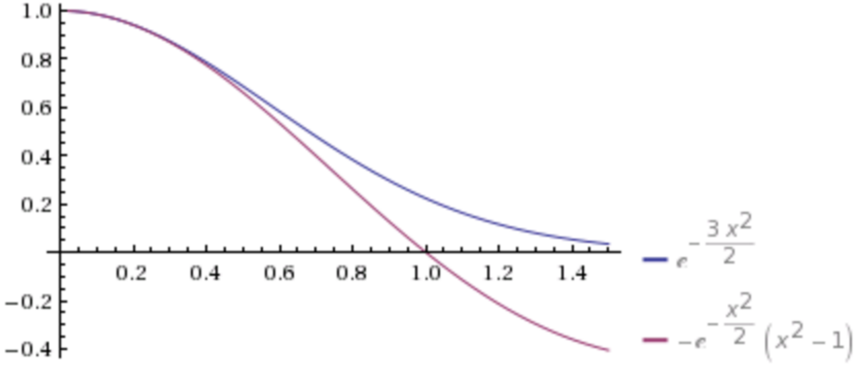

Figure 7: Relative profitability of a humble market maker (blue) and one who assumes their estimates are precisely correct (red)

Result #4 Folk law ... not so much

So don't let anyone tell you that "once you get inside the bid-offer the returns on accuracy are diminishing".

That plausible argument is demonstrably false in this model, and so likely to be a poor guideline in the real world too.

Summary of theoretical work

1. A financial argument for constraints on inventory cost $\nu(x)$.

2. Clarity around how width and skew relate to inventory cost.

In the special case of exponentially distributed competitor markups:

3. A simple way to deal with client enquiry imbalance

4. Analytic expression for width as a function of inventory assuming constant width (and vice versa)

5. A numerical approach improving on both 3 and 4 above.

6. Analytic expression for market making efficiency based on variance in parameter estimates

Practical application

I hope I've convinced you that the inventory indifference price $\nu(x)$ is almost synonymous both with both (stylized) optimal trading and also, to first approximation, the evaluation and improvement of actual trading (not subject to precisely the same simplistic constraints).

A year or so ago I set out to build a data science framework to manage the anticipated complexity of all the various types of business intelligence emanating from this analysis.

Summary of practical work

1. Demonstrated scalable reproducible research using Domino Data labs (now adopted by CIB Data Science). For example the use of 260 32-core 256 Gig machines on Amazon simultaneously to produce the plurality of models entering the trading reports.

2. Created business intelligence pipelines comprising fifty stages of analysis with many intermediate calculations of independent interest - and made it work for 11,000 bonds.

Summary of practical work (cont)

Some recent progress with Domino/Athena inter-operability:

Code: Building on tools written by CORE for their own work that sync code, I'm now able to mock out Athena's test suite for a generic runner, allowing for development of pure mathematical code common to Athena/Domino with exactly the same tests (thus a stronger set of controls than Athena SDLC, not weaker).

Data: With Andrew Stein I've demonstrated REST based data round trip between Athena and the Domino Data labs research environment. This will plug into Pascal's work exposing data domains behind REST.

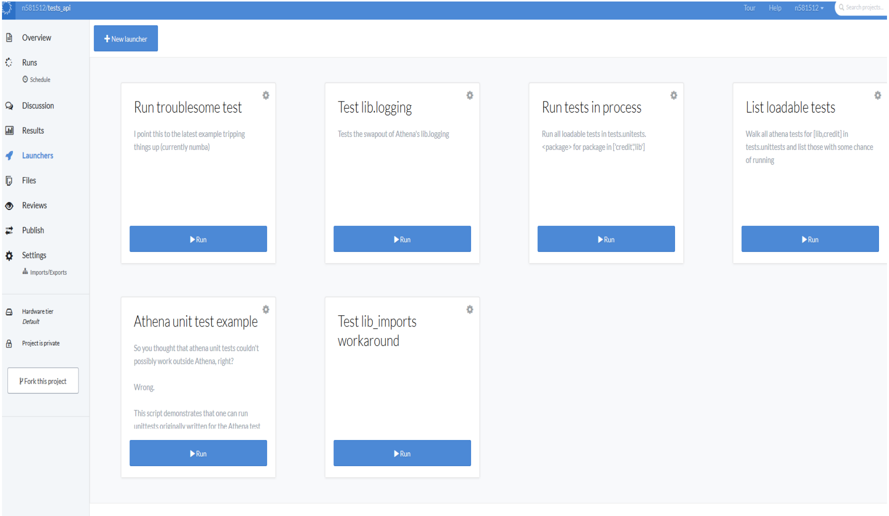

Figure 8: Monitoring tests ex Athena

Looking ahead

Theory: Less stylized evaluation of inventory indifference pricing using techniques that, for the moment, I'm having difficulty explaining to myself (see definition of mathematics).

Practice: More detailed analysis of trading using the same.

```python
import inspect
import unittest2.main

def RunModule():
    """ Swaps out the Athena testfns.RunModule """
    from lib.domino.environment import in_athena
    if not( in_athena()):
        frm = inspect.stack()[1]
        mod = inspect.getmodule( frm[0] )
        unittest2.main( mod, exit=False )
```

Figure 9: A few lines of python that go a long way

Recommendations

Oldies but goodies:

1. Integrate of a third party data science platform (e.g. Domino Data Labs) until such time as CORE, GTI and others can come close to matching the functionality.

2. *Prioritize* well considered, tested approaches to optimizing the flow business over ad-hoc, just-so, or lowest common denominator consensus.

New recommendation

Credit to be an anchor customer for *Roar Data*, a crowd-sourcing company to be incubated by JPM.

Rationale:

1. There are no diminishing returns on accuracy
2. We are constrained in so many ways (regulation, libraries, time, data, model review, controls, head-count, culture and inability to be experts on every possible technique).
3. We are singularly well placed to exploit the world's best predictions (franchise, business and superior understanding of how these translate into optimized trading).

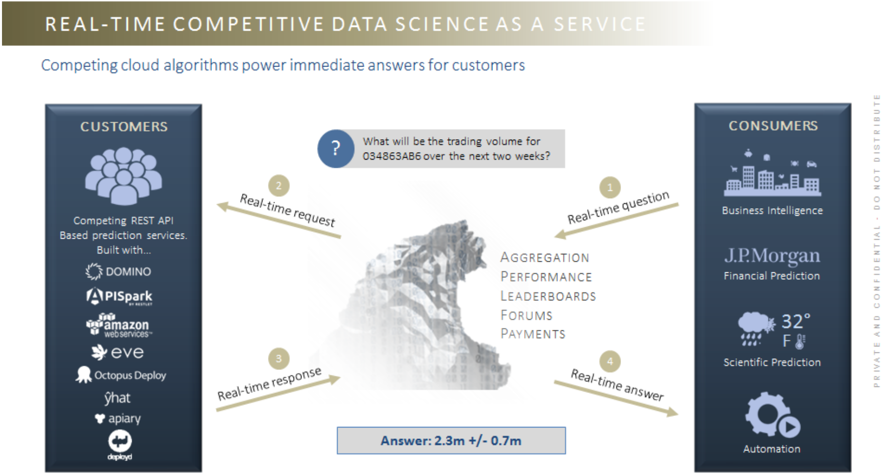

Figure 10: Roar data concept

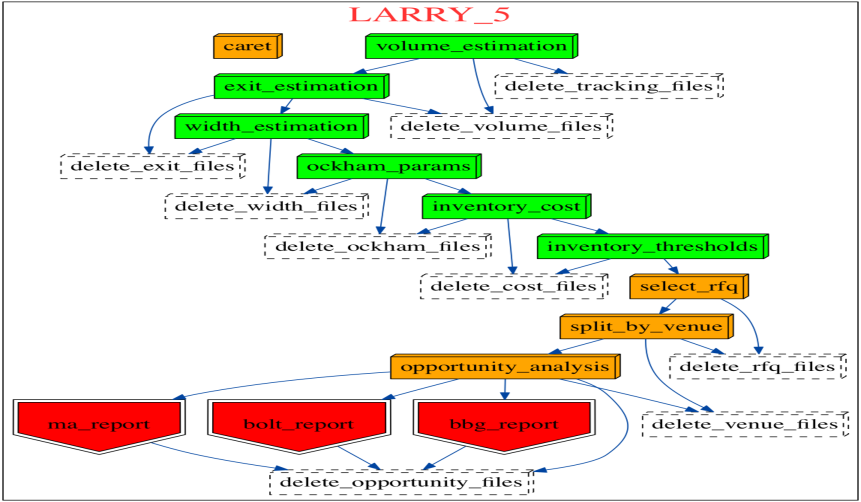

Figure 11: Research pipeline: stage # 5

Questions?