High-Frequency Algorithmic Trading with Momentum and Order Imbalance

My goal is to establish and solve the stochastic optimal control problem that captures the momentum and order imbalance dynamics of the Limit Order Book (LOB). The solution will yield an optimal trading strategy that will permit statistical arbitrage of the underlying stock, which will then be backtested on historical data.

Progress Timeline

DATE	THESIS	STA4505
~~Dec 2014~~	~~Complete CTMC calibration~~	
~~Dec 2014~~	~~Backtest naive strategies based on CTMC~~	
~~Jan-May~~	~~Study stochastic controls: ECE1639, STA4505~~	
~~Jun 5~~	~~Establish models~~	~~Exam Study~~
~~Jun 12~~	~~Establish performance criteria~~	~~Exam Study~~
~~Jun 15~~	~~Derive DPP/DPE~~	~~EXAM~~
~~Jun 19~~	~~Derive DPP/DPE~~	
~~Jun 26~~	~~Derive continuous-time equations~~	
~~Jul 3~~	~~Derive discrete-time equations~~	
~~Jul 10~~	~~Set up MATLAB numerical integration~~	
~~Jul 17~~	~~Integrate functions and plot dynamics~~	~~Integrate and analyze~~
~~Jul 24~~	~~Empty promises; zero fucks given~~	
~~Jul 31~~	~~Code for numerical solutions, calibrations, plots~~	~~Simulate results~~
Aug 7	'wrong' calib method: all IS and OOS backtests	
Aug 14	'right' calib method: all IS and OOS backtests	Project writeup
Aug 21	Compare with Naive methods, dissert. writeup	
Aug 28	Dissertation writeup	Presentation
Aug 31	Submit thesis for review by panel.	

Whiteboard Inspirational Quote of the Week

There is no way to happiness. Happiness is the way.
– Thích Nhất Hạnh

For Our Readers in the Middle East...

Sun 26th: ok bud here's the deal. shit is getting real, time-crunch style, as evidenced by the fact that i installed the Vietnamese TeX language package just to spell that monk's name properly. but here's what we're looking at for the immediate future: UTIAS has to submit a degree recommendation on my behalf to the School of Graduate Studies by Oct 2nd. before then I need to finish a 'comprehensive first draft' that I submit to my two supervisors, Seb obviously being the more important of the two, receive back, make revisions, submit to an external reviewer, receive back, make revisions, submit to the graduate department. all that has to happen first. UTIAS has a recommended deadline of 3 months out for the first draft, 2 months out for the external reviewer, which, given the constraints on time-travel, clearly isn't happening. i shot the Real Big Seb Himself an email just now to run this by him, i'll get his thoughts hopefully soon, but one thing is immediately clear: the more I can compress that timeline above, and perform basic office functions with terrifying efficiency, the more likely is the success. so here we go my man. here's basically the full list of shit I need to do:

- ~~finish coding solutions for the ansatz h for all four cases.~~
- ~~write code for the two ('right' and 'wrong') calibration methods~~
- ~~write script for single-day calib in-sample backtest~~
- write script for single-day calib out-of-sample (+1 week) backtest
- write script for annual calib out-of-sample backtest
- run ALL of those on the entire 2013 calendar year, for each of the 2 calibration methods, for each of the performance criteria, for 4 stocks of varying liquidity. all told: 4 stocks * 4 criteria * 2 methods * 3 backtests * 250 days = 6000 backtests.
- produce plots and tables to summarize results. mimic Seb's four plots to show arbitrary given stock, given day's testing using best calib method, for each of the criteria (another 4 backtests)
- compare the averaged results with the naive strategies from last year
- write-up

Mon 27th: teddy I have been at work for 12h now, during which I have eaten one granola bar and had two coffees. this has been by far the most productive day i've ever had, and yet, absolutely nothing fucking works. in this mad work frenzy i've watched the entire analysis crumble before my eyes and i'm convinced i have to re-do a lot of it.

Tue 28th: what a turnaround day today. in a moment of insight i took the time to read a paper that i suspected might be able to help me out of this jam – and voila, pretty much the first authentic citation of my reserch, and definitely saved the day. anyway, so, a million times more progress than yesterday, from which i conclude the following two things: productivity and progress are not the same thing; and that today's success really could not have been possible without yesterday's insane struggle. it turns out i accidentally followed the fast for tisha b'av - perhaps hashem was channeling some serious cosmic shit into me. regardless, irregardless, i'm delighted to say we're potentially back on a roll here. time to get gymnasty.

Wed 29th: *I'm the pusherrr.* shit is fucking coming together my man! i think what's gonna happen here is i'm gonna postpone coding and generating some of the plots till i start doing some of the report writing... i know roughly the sort sick ass plots i want to include, but at the same time not all of them, and i sort of reckon it'd be a bit more fluid to just start the writing and be like 'oh yeah that'd look real svelte right in here' and generate it, rather than right now try to conjure up all the different things i want. some are obvious and those're already done! yippee. so what this REALLY means ... is that i'll be getting out of here real soon.

Thu 30th: holy fucking shit dude – i've begun assembling the dissertation. it's fucking nuts. first of all i dropped about 4 hours on finagling with the Cambridge University PhD latex template, only to eventually realize i'm better off just writing the code on my own...but i did 'borrow' their cover page format, so i've got myself a nice-looking coverpage, which chiefly amounts to (you probably would've guessed this having read over my CV) having the UTIAS crest in the middle. that aside, i've streamlined my latex code, made it so fucking modular that you'd weep if you saw it, and have spent the last half hour just scrolling up and down on the pdf, up and down, very pleased with myself. obviously 50% of the content needs to be rewritten, being total garbage, but it's a start. the other issue is that i haven't yet written my introduction, or the entire results section, and i'm already over the page limit. *post-script:* okay new plan because somehow having the whole thesis pulled up is just overwhelming. so i'll be shelving that one till that mid-august write time. now we focus on the billion backtests i have to run...

Fri 31st: welp shit. the end of July is upon us - almost as bad as there being a fungus among us. word just in from Sebastian: try to have the first draft done before the last week of August, but can always submit revisions later. so let's shoot for the Monday the 24th. immediately i'm having visions of all the shit that i'd like to slash out of this thesis...namely i think cases 2 and 3 of the continuous time stuff, and if that doesn't mean anything to you then that itself is reason to slash it as you're probably the second most knowledgable of my thesis as it is.

Sat Aug 1st: okay man. many hours have gone into revising the third chapter and i think they're in pretty good mathematical shape. chapters 1 and 2 are basically giant shitshows but i care substantially less about that at the moment, as here's the plan (obviously it involves old faithful): this coming week is basically week 1 of 2 of simulation weeks. i'll let the big bad custom build octocore laptop handle that job, and meanwhile i'll be revising chap 2 and conjuring chap 1 out of thin air. i can see it now. octocore whirring its fan like crazy trying to crunch the numbers, old faithful whirring her fan silly just trying to stay alive. my labmate Bill has commented several times about how much noise my computers make.

The Academic Week in Review

Just some notes to self below, wanted to figure out which tickers to run based on liquidity, but basically it was decided for me because while i wanted to use MMM and AAPL, the data is so fucking messy that after half an hour of tinkering i realized i had not the patience nor desire to go through the whole data cleansing routine.

Avg Vol 3m
x FARO 193,544
GOOG 2,038,000
MMM 2,314,640
SMH 3,926,520
x NTAP 4,378,990
x ORCL 14,457,300
x INTC 31,716,600
AAPL 45,965,900