Final Project

Note: All files and information related to the final project are located in the folders starting with "Final Project*" on Courseworks.

This and related files can be found in the "*Final Project*" folder on CourseWorks. The Final Project groups composition and the Final Project primary markets selection is contained in the file "*Final Project Groups 2023.xlsx*" uploaded into "*Final Project*" folder on CourseWorks.

Before you develop and test the trading strategy, you need to show some basic understanding of the market you will be trading.

Next, at the time series level, you need to run at least both of the statistical Random Walk tests we have studied in the class: Variance Ratio test and the Push-Response test to identify the inefficiencies in those time series. You need to run both tests for both time series and you need to make inferences about the type of inefficiency (trend-following or mean-reversion) present in the time series and its approximate location: range of time-scales of the inefficiency for each of the time series.

Based on the EasyLanguage pseudo-code or the Matlab code that I also provided (files "*main.m*", "*ezread.m*") describing the basic trend-following system "*Channel WithDDControl*", you need to program this system in a high-level language of your choice (preferably: C, C++, java; you can also use Python and Matlab – but you will need to spend time to make the code very efficient; even gnu Fortran or C working under Linux will do the work) and apply it to **two** of the futures markets from the following set:

Number	Blp Ticker	TD Ticker	Description	Data From
1	BO	BO	Soybean oil	Jul-82
2	DX	DX	US dollar index	Jul-89
3	HG	HG	Copper	Dec-89
4	HO	HO	Heating oil	Jan-84
5	JO	JO	FCOJ - Frozen concentrated orange juice	Jul-87
6	JY	JY	Japanese Yen	Mar-77
7	S	SY	Soybeans	Jul-82
8	SB	SB	Sugar #11	Jul-86
9	SF	SF	Swiss Franc	Dec-74
10	TU	TU	US Treasury note - 2-year	Jan-91
11	TY	TY	US Treasury note - 10-year	Jan-83
12	W	WC	Wheat	Jul-82
13	SM	SM	Soybean meal	Jul-82
14	CO	CO	Brent crude oil	Aug-03
15	FC	FC	Feeder cattle	Jan-78
16	DU	BZ	German Schatz - 2-year	Mar-97
17	PL	PL	Platinum	Oct-07
18	XB	XB	Gasoline (RBOB)	Oct-06

19	BTC	BTC	CME Bitcoin Futures	Dec-17
20	KEE	COKE	Coke Futures	May-18
21	DCS	CS	DCE Corn Starch Futures	May-18
22	AK	SOYBN	No. 1 Soybean Futures	May-18
23	AUA	AUG	Gold Futures	May-18
24	CU	CUC	Copper Cathode Futures	May-18
25	VV	CFC	Cotton No.1 Futures	May-18

One **primary** **futures market** I will choose for you out of the well-known Western markets (#1-18) in the table above, and one **secondary market** (you can choose more than one) you must choose yourself, including from the markets in the table above shaded in grey color, such a *"crazy" BTC (Bitcoin)* futures and very interesting Chinese futures markets (#20-25). As you can see, I have prepared for you *six very liquid Chinese commodity futures markets* to choose as a secondary market. *To avoid overlapping, all student groups will have different primary markets,* provided be me. It is OK for groups to have the same secondary market(s). The zipped *"*.csv*" data files containing the OHLC (Open, High, Low, Close) bars with 5-minute resolution from inception until present can be found in the "*Final Project Price Data*" folder, files names carrying the markets' ticker symbols. If you feel your code is fast enough, or you just want to explore more in-depth, you can apply the strategies to 1-min data that you can request from me by email. Conversely, if you were not able to optimize your code to quickly obtain the required results on 5-min data (everyone should be able to do that), and the time is pressing, please, request a 15-min data from me for your markets by email. I will provide them and you will finish your Final Project on time. Please, note that Bloomberg is using the Bloomberg ticker convention, whereas the data files use the provider's ticker convention ("TickData" or "TD", also known as "One Tick"). They typically use the local main exchange symbols as tickers. Also for your reference I have uploaded the Bloomberg descriptions (DES screens), the sample daily OHLC bars chart (GPO screens) and the Contract Table (CT screen), see folder "*Final Project Bloomberg Screens*". You can use them to learn/introduce the markets and to cross-check various contract parameter settings information. All of these markets should present you with trend-following properties to various degree of strength. After you polish your code results on the primary market, each group needs to apply everything developed to **the secondary market**, and repeat all the tested code on it. Please, note that you will not be judged on how much money your strategy makes or on how big is the value of your risk-adjusted return. Instead, you will be judged on how close your results are to the expected ones.

Also for your convenience, I have uploaded into the "*Final Project*" folder on CourseWorks the "Medeleyev Table" of the futures markets that I was discussing during class (the file "*TF Data.xls*", although you may use the Bloomberg DES screens that I referred to earlier), which, among other things, contain: TickData Ticker (column B); Bloomberg Ticker (column C), Description (column D), local currency approximate contract value (column G), local currency Point Value (column H), etc. Some or all of these parameters you may need in your experimentation. The important parameter that you will absolutely need to use is slippage, denoted as "Slpg" in column V. This is the suggested full transaction cost to be used by you if you buy and then sell (or sell and then buy) one futures contract of your chosen market. Every time you do one round-turn transaction like that you will need to subtract that value from the current $P&L (USD Profit and Loss or equity). The data for Chinese futures markets I kept in a separate file because of its uniqueness – they have the third ticker, Chinese one and some other unique properties. The complete file with those futures facts is called "*TrCostsChina 03-07-2019.xls*", also uploaded into "*Final Project*" folder on CourseWorks. The two only values that you need from it is Slpg

from column B, Point Value from column D. Please, do not forget that all numbers expressed in currency local for the market, such as all Chinese futures – in CNY (Chinese Yuan Renminbi). I am sharing with you the Chinese markets, which are amazingly inefficient. There is a lot of trend-following properties in them, you can even "see it" with your eyes by graphing the price series.

All parameters such as tick size, trading hours, and point value should be taken from the Bloomberg DES screens or *TF Data* file and double-checked with us, if in doubt. Please pay attention to the fact that the timestamps in the data files are expressed in the local (exchange) time, i.e., the CME in Chicago is one hour behind EST.

Also, pay attention to the fact that there may be a factor of 100 (or 1/100) difference in the price in the file and the price on Bloomberg: the *TF Data* futures contracts parameters file will help you with this – see the magenta-colored lines.

You will need to first use the in-sample length of 4 years (re-calculate it into the approximate number of 5-min bars to use in each case). I would like you to use **the net profit to max drawdown ratio** as the target function for parameter optimization. I would like to propose that you use the following ranges of parameters (you can adjust those if you have a more or less powerful computer, for example):

ChnLen = 500, 10000, 10 (start at, end at, step with, in 5-mins intervals, 951 in total); and
StpPct = 0.005, 0.10, 0.001 (96 in total).

This corresponds to $(950+1)(95+1) = 91\,296$ grid points in each two-parameter optimization. This optimization can be solved by a direct and full look-through in search for a single maximum (if you can propose and implement alternative search procedures in addition to full search – it will be rewarded while grading). As the number of parameters is small (only two) and the sample space is relatively small (on the order of 10^5), this should be doable on virtually any modern computer. In this experiment, you always have to choose the formally best parameters point. Repeat the optimization procedure by moving your in-sample period from the very end (or start) of your full data sample every quarter (3-months) forward, which you can do approximately, by counting the same number of 5-mins in a quarter, for example. Once you found the optimal parameters for the previous 4 years, use them for the immediately adjacent subsequent quarter (3 months) out-of-sample. Record the strategy equity changes and trade-by-trade table only for the out-of-sample period. Repeat this process iteratively for all remaining quarters, thus producing a fully out-of-sample equity curve and trade-by-trade table. Measure various statistical performance properties of both equity curve and trade-by-trade table: Average Rate of Return, Standard Deviation of Returns, and the Ratio of the two above (known as the Sharpe or Information Ratio), Total Number of Trades, % of Winners or % of Profitable Trades, Return on the Account (**Net Profit to Worst Drawdown**), Average Winner, Average Loser, **Profit Factor**, etc. along with all other characteristics that you will deem useful to gaining insight into the behavior of your market. The quantity and choices of the performance characteristics will make a difference on the grade for your Final Project.

Please, do not forget to keep/save a table with the optimal parameters for each quarter. After you are done with the above experiment you can also solve the in-sample optimization with the whole period being the in-sample length. And calculate its performance characteristics similar to the ones you used above. After you are done, you can compare those to the ones for the out-of-sample experiment and

deduce the decay coefficients for various performance measures to estimate the difference or decay of the out-of-sample vs. the in-sample performance for the performance measures you chose.

Variations of the above for the benefit of better results or more interesting outcomes are welcome as well. If you have any questions, please ask!

Finally, once the code is polished, you need to repeat the above procedure for various choices of the in-sample time T (=1,2,3,…,10 years if possible) and out-of-sample time τ (=1,2,3,4 quarters or even =1,2,3…,12 months).

You will be expected to write all your details in the PowerPoint presentation and present your work (using one or several presenters) during 30-40 mins allocated for that time for each group.

Do not hesitate to ask any question of me or the TA, as well as much of this will be covered during the class.

Good luck!

You will be given 3 weeks to complete the Final Project and to prepare the presentations. I will be available in class every week before presentations at and during our regular lecture time to answer any of your questions.

Final presentations will be during the class time on April 29th, 2023.