

Assignment 1

The “value” factor is one of the most researched portfolio strategies that has also been widely implemented in industry. The strategy has staunch defenders and vigorous critics and before committing to launch a product in this space your team is tasked with creating a comprehensive report on the pros and cons of a retail investment product based on this investment idea and its interaction with other factors.

0. Introduction and background

After decades of attractive returns, the value factor’s performance has waned significantly since the 2008 Global Financial Crisis. A number of practitioners and academics have thus started to question the ability of the value factor to deliver in the future:

- [Does value investing still work?](#)
- [Is value investing back from the dead?](#)
- [Reports of value’s death may be greatly exaggerated](#) (summarized [here](#))

However, there is evidence that the value factor performs well during periods of high inflation, and so it could be time for a triumphant comeback for value, particularly if combined with other complementary factors:

- [Finding value in value](#)
- [Understanding the Performance of the Equity Value Factor](#)

The CIO (Chief Investment Officer) of your company is exploring the possibility of launching a product in this space but, before committing to this course of action she would like the quant research team to **prepare a short, but comprehensive, analysis of the value investment factor and the potential to combine it with the low volatility factor.**

Your report should have the following components:

1. A presentation of the **fundamental intuition** behind the value investment factor;
2. A **quantile analysis** of the predictive power of the value factor as implemented using three different measures:
 - a. **EPQ** = the E/P ratio where E is the last quarter EPS
 - b. **EP12** = the E/P ratio where E is the last 12 months EPS
 - c. **BTM** = the book-to-market ratio;
3. A presentation of the **results of backtesting** the “best” value strategy (measure) from the quantile analysis over the 2005–2020 period;
4. A proposal of some “**clever**” **way to combine** the best performing value factor with the low volatility factor [**Black Belt**];
5. **Optimization of the chosen combination** in terms of rebalancing frequency and number of shares bought.

The deliverables for this assignment are:

- A Jupyter Notebook with your code;
- A PDF file with a short written report.

There is **no formal length requirement** for the report. My suggestion is to aim at anywhere **between 7-8 pages** of “main text” including tables and pictures. You may also add an appendix if you want to add more tables, etc. Please be sure that all the necessary information is in the main body of the report.

1. Presentation of the Strategy

This should be a presentation of the **economic intuition behind** the value investment strategy only. After reading this part the reader should understand:

- What fundamental economic concepts underpin this strategy;
- Why this strategy should work (Isn't it a fundamental violation of market efficiency?).

This part will be mainly descriptive (although, if you want, you may add a table and/or a graph if you need it to make a point). The **intended audience** is the Chief Investment Officer of your firm, so you can assume a high level of sophistication and a good understanding of finance.

As a reference material for this part, you can refer to the preparatory material of the subject and any further articles/resources that you can find.

The length of this part should be **one page of text plus optional graphs and tables**. For this part only you can take graphs and tables from other sources (with proper citation). Please limit the number of graphs and tables in this part. It should be mostly about providing the economic intuition in your own words.

2. Quantile Analysis

In the second part of your assignment, you should present a detailed statistical analysis of the **predictive power of the value factor** implemented via each of the three possible measures (EPQ, EP12 and BTM). Specifically, you should perform a Quantile Analysis for each measure.

In doing this analysis you should use all the available data and you should show **how the predictive power changes over time**. After reading this part the reader should understand:

- **How strong** the predictive power of the three value factors are over the entire sample;
- Whether this predictive power has **changed over time**.

Of particular interest to your boss is the fact that many traditional factors **have weakened in time** and are nowadays significantly less informative than in the past. Has the value really strategy weakened over time?

For the **length of the quantile analysis**, you should aim at **around one-two pages**. If you feel the need to add many more tables and graphs, consider the possibility to relegate some of them to the appendix, leaving only your “main narrative” in the body of the report.

3. Backtesting

For the third part of your report you should proceed with the “best” performing value factor (out of EPQ, EP12 or BTM) from you quantile analysis in part two. You should make it clear what measure you consider “best” and why.

You should provide a detailed report of the **performance of a long-only strategy** that invests in stocks that rank well according to the best performing value factor. Your strategy should:

- Invest in the top 250 stocks;
- Be tested on data between 2005 and 2020;
- Rebalance monthly;
- Assume 0.2% roundtrip transaction costs.

Your report should include:

1. A table with the main diagnostics;
2. A graph with the performance of the strategy and the benchmark;
3. A bar graph with the total active return of the strategy for every year;
4. Graphs with the industry composition of the portfolio and the benchmark.

The results should **be properly commented in your report**. After reading this part, the reader should know:

- Whether the strategy performed well or not over this period;
- Whether the ability of the strategy to “beat the market” is constant across this period or has changed in time;

- Whether the strategy has any strong industry bias compared to the benchmark.¹

The length of this part should be around **two pages**, including comments and **graphs/tables**.

4. Combination with Low-Volatility Factor

As suggested in Andrew Ang's recent article, [Finding Value in Value](#), there may be benefit in **combining** the traditional value investment factor with the low-volatility factor to provide some downside resiliency (since low volatility has tended to lose less during down markets).

In the fourth part of your report you should **propose a “clever” way to combine** the best performing value factor (used in part three) with the low-volatility factor provided—Total Volatility in the TVOL.zip data file.

In your report should include:

- A **description of the method** used to combine the factors and the economic/statistical rationale behind it;
- A **brief summary** of the performance of a backtest of the combined factor with the same assumptions as in part three above. Specifically, it should include:
 - A table with the main diagnostics;
 - A graph with the performance of the combined strategy, the value-only strategy, and the benchmark.

ATTENTION

This part is the **“black-belt”** component of the assignment in the sense that we have not discussed in detail how to combine factors in the subject. Therefore, you will need to research methods of combining factors and how to implement this in python.

This section is worth 20% of the total Assignment mark. If you do not want to attempt this component then **simply take the average of the two factor values** as the combined factor to perform your backtesting. A maximum of 10% is then available for this section (for the backtesting results of the simple average factor value).

The length of this part should be around **one-two pages** including both your comments and **graphs and tables**. As with the other parts of the assignments **we will mainly mark the PDF report**, so all your results have to be presented and commented there.

¹ Industry bias means that the strategy is far more (or less) invested in some industry (for example tech or financials) compared to the market.

5. Strategy Optimization

In the main backtesting you have used monthly rebalancing and invested in the top 250 stocks. These parameters have been suggested by your boss based on her “years of experience”. In the fifth part of your assignment, you need to **optimize the strategy** (analysed in part four) by finding the best values for the key parameters. Specifically, you want to test:

- Rebalancing every 1, 3, 12 months;
- Investing in the top 100, 250, 1000 stocks.

You should test **all these combinations** and record the **Information Ratio** of the strategy. The result of this part should be a table with all the combinations of parameters and the related information ratio. You should briefly comment the result.

After reading this part the reader should know what combination of parameters produces the best performance.

Notes:

1. When performing the backtests for strategy optimization, you should use the same time period and transaction cost assumption as in part three;
2. If you can, try not to repeat the backtesting function nine times, instead use for-loops.

For the length of this part, you should aim at **around one page**. If you feel the need to add many more tables and graphs, consider the possibility to relegate some of them to the appendix, leaving only your “main narrative” in the body of the report.

6. Data

Together with the assignment you will find the following data files:

- **EPQ.zip**, **EP12.zip**, and **BTM.zip** contain monthly data on three different “value” measures (EPQ = E/P ratio where E is the last quarter EPS, EP12 = E/P ratio where E is the last 12 months EPS, and BTM = book-to-market ratio) for a sample of US stocks. The factors are already normalized with a higher numerical value corresponding to a better “value” (a stronger signal);
- **TVOL.zip** contains monthly data of a low-volatility factor (Total Volatility) for a sample of US stocks. Again, the factor is already normalized and defined as “increasing in quality”, meaning that the higher the number, the lower the risk measure;
- **Prices.zip** contains the monthly (adjusted) prices for US companies from 1980.

7. Timeline

The content of the assignment is covered in different weeks of the subject. Here is a rough estimation of when the material (and code) necessary for each session is covered:

Assignment Part	Week	Notebooks / Concepts
1. Presentation of the Strategy	6 and 7	None. Concepts of Investment Factor, Risk Premium and Market Anomaly
2. Quantile Analysis	7	08 – Quantile Analysis
3. Backtesting	8	09 – Full Backtesting
4. Factor Combination	6–8	09 – Full Backtesting and your own coding/research skills
5. Strategy Optimization	6–8	09 – Full Backtesting

8. Submission

You will submit:

1. A PDF file with your report (**A1_XXXXX.pdf**);
2. A Jupyter Notebook with all your code (**A1_XXXXX.ipynb**).

Your PDF report should contain a professionally-looking **cover page** (no specific format required) with your name and Student ID.

Tables in the report should be edited and not simply cut and paste images from the Notebook. The graphs coming from the notebooks should be exported as image files and then inserted in the report and not “screenshotted and pasted”.

The files will be **submitted electronically** using the electronic drobox in Canvas.

8.1 Structure of the Jupyter Notebook

1. You should start with an empty notebook and add all the cells that you need;
2. Your python notebook should contain your **name and student ID** at the top in the initial markdown cell;
3. The first code cell should contain all the **import statements** (including the *apmodule*). No import statements should be located anywhere else in the file;

4. The second code cell should contain all the **user-generated functions**. You should not report here the functions from our own *apmodule*, but use them from the module;
5. The third code cell should contain the initial loading of all the files;
6. The notebook can import **any of the libraries used** in the subject including the *apmodule* library. If you want to import a library that we have not used in class, please check with me beforehand;
7. The file can only load and use the **datafiles provided with the assignment** or other datafiles used in the subject. No external data can be used;
8. Your code should be **well commented** using markdown cells and # comments in the code cells;
9. You should use markdown cells to **help the reader navigate the file**, basically explaining what you are doing in the code cells (or at least the most relevant ones);
10. You should choose **which results to present** in the notebook in order to help the reader without overwhelming. For example, if you simply modify a column in a DataFrame it may not be necessary to show the result on screen;
11. I need to be able to **run your notebook in one go** from top to bottom, so before submitting make sure that all your code runs properly. Do not worry about the running time of your notebook. You do not need to optimize for speed;
12. You can **copy portions of the notebooks** created in class, but please **do not follow** their structure. The notebooks used in class were designed for teaching purposes and do not reflect the sequence of objectives of your notebook. Not everything we have done in class fits with this research project.

You should consider the notebook as the **natural complement to the PDF report**. Imagine that your reader will go through the report and then, from time to time have a look at the corresponding section of the notebook to understand some technical aspect of your model. This implies that:

- All the important information is in the report. You will not receive marks for required elements of the assignment that are in the code but not in the report.
- The notebook should not contain significant elements (for example the quantile analysis using N quintiles) if the results of such analysis are not commented in the report.

9. Marking Guide

Component	Value	Below Expectations 0 - 5	Meets Expectations 6 - 8	Exceeds Expectations 9 - 10
Presentation of the Strategy	10	<ul style="list-style-type: none"> The presentation of the factor portrays a lack of understanding of its economic fundamentals The analysis of the determinants of the factor is confused or insufficient 	<ul style="list-style-type: none"> The presentation of the factor is complete and based on coherent financial and economic reasoning 	<ul style="list-style-type: none"> The presentation is rich and engaging The language is professional and proficient The presentation conveys a deep understanding of the factor and its determinants
Quantile Analysis	10	<ul style="list-style-type: none"> There are serious methodological errors in the analysis The analysis is incomplete The results are not well presented The results are not well commented 	<ul style="list-style-type: none"> There are no serious methodological errors, and the analysis is complete The presentation of the results is sufficient 	<ul style="list-style-type: none"> The analysis is rich, and the problem is analysed from multiple points of view (complementary techniques, multiple time horizons, ...) presenting a comprehensive picture The results are presented in a professional manner providing enough information without overwhelming the reader
Backtesting	10	<ul style="list-style-type: none"> There are serious methodological errors in the analysis or the analysis is incomplete 		<ul style="list-style-type: none"> The analysis is complete and functional
Backtesting: Results	10	<ul style="list-style-type: none"> The results are not well presented and commented 	<ul style="list-style-type: none"> The presentation of the results is sufficient 	<ul style="list-style-type: none"> The results are presented in a professional manner providing enough information without overwhelming the reader
Combining Factors: Methodology	10	<ul style="list-style-type: none"> A simple average is used to combine factors 	<ul style="list-style-type: none"> The presentation of the method used is complete and based on coherent reasoning 	<ul style="list-style-type: none"> The presentation conveys a deep understanding of issues at hand and the pros and cons of different methods
Combining Factors: Backtest Results	10	<ul style="list-style-type: none"> There are serious methodological errors in the analysis or the analysis is incomplete The results are not well presented and commented 	<ul style="list-style-type: none"> The presentation of the results is sufficient 	<ul style="list-style-type: none"> The analysis is complete and functional The results are presented in a professional manner providing enough information without overwhelming the reader

[Continues in next page]

Component	Value	Below Expectations 0 - 5	Meets Expectations 6 - 8	Exceeds Expectations 9 - 10
Optimization	10	<ul style="list-style-type: none"> There are serious methodological errors in the analysis, or the analysis is incomplete The results are not well presented and commented 	<ul style="list-style-type: none"> There are no serious methodological errors, and the analysis is complete The presentation of the results is sufficient 	<ul style="list-style-type: none"> The optimization is well structured with for-loops and without useless repetition of code The results are presented in a professional manner providing enough information without overwhelming the reader
Notebook: Structure	10	<ul style="list-style-type: none"> The notebook does not run properly The structure of the notebook is confusing and unhelpful 	<ul style="list-style-type: none"> The notebook runs properly The notebook is well structured 	<ul style="list-style-type: none"> The notebook is very well designed The coding is elegant and makes good use of functions and loops to avoid repetitions The variables are properly named and easy to follow. The choice of which intermediate results to present is effective to help the reader
Notebook: Comments	10	<ul style="list-style-type: none"> The notebook is not commented, or the comments are insufficient and/or unclear 	<ul style="list-style-type: none"> The comments are sufficiently to allow the reader to properly follow the process 	<ul style="list-style-type: none"> The comments are rich and well-made, allowing the reader to follow both the financial logic and the code structure The comments make good use of Markdown to improve readability
Quality of the Report	10	<ul style="list-style-type: none"> There are frequent spelling and punctuation errors. The language is often non correct and does not convey financial concepts in an effective way There is not a common graphic style and colour palette Graphs and tables are badly formatted 	<ul style="list-style-type: none"> There are only minor spelling and punctuation errors The language is correct, and the financial terminology properly used There is evidence of editing to make the feel professional 	<ul style="list-style-type: none"> The document has a truly professional tone and feel The narrative structure is easy to follow. The language is very effective, and the key ideas presented in a succinct but clear way Tables and graphs are effective in communicating the key results