

Model based on Transformations, and Cutoffs based on the macromodel

The purpose of this analysis is to evaluate the performance of portfolios constructed using stocks ranked according to transformed BARRA factor exposures. The performance will be analyzed in xBSL space.

[Related document: US Long Model \[Dec 25, 2022\]](#)

SUMMARY

Inputs: Barra fundamental factors

Barra fundamental factor exposures are typically normalized, but if they are not, they must be normalized before analysis.

Process:

Transformations: Transform factor data key transformers are:

- Moving averages. 3M, 6M, 12M , 24M , 36M , 60M
- Moving Difference: 3M, 6M, 12M , 24M , 36M , 60M

Transformation of the raw exposure data involves either calculating moving averages or moving differences.

Normalization window (excluded for this report)

- 3 years and 5 years (More details to be added)

Cutoffs: On cross sectional data for a certain date 6 buckets based on the scores for each factor

- Cutoff intervals are:
 - > 1.2 standard deviation (sd)
 - $> 0.7 \leq 1.2$ sd
 - $> 0 \leq .7$ sd (similarly for the negative side)

The quantile approach to creating portfolios involves the following steps

1. Calculate the moving average scores for each stock each month.
2. Rank the stocks based on their scores.
3. Divide the ranked stocks into equal-sized portfolios called quantiles.
4. The composition of the quantiles will change each month based on the rankings of the stocks.

Each month, the moving average scores are calculated and the stocks are ranked. They are then divided into portfolios called quantiles, resulting in a different composition of the quantiles each month.

In the quantile approach, 10 portfolios of stocks, called quantiles, are created based on the rank assigned to each stock based on its moving average score. The stock with the highest moving average value will be part of the top quantile. The moving average is calculated on a continuous period, taking into account any gaps in data that are filled with NA values.

METHODOLOGY: Dynamic Portfolios

This approach is called dynamic because the portfolio composition changes each month based on the moving average scores. This allows the portfolio to be responsive to changes in the market.

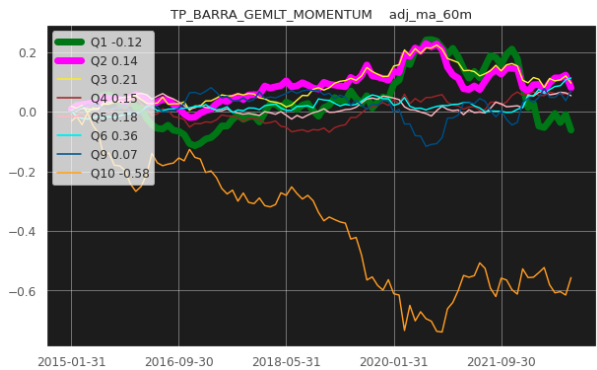
Updated: 2022-Dec-14

Start date 2014-12-31. End Date - 2022-10-31

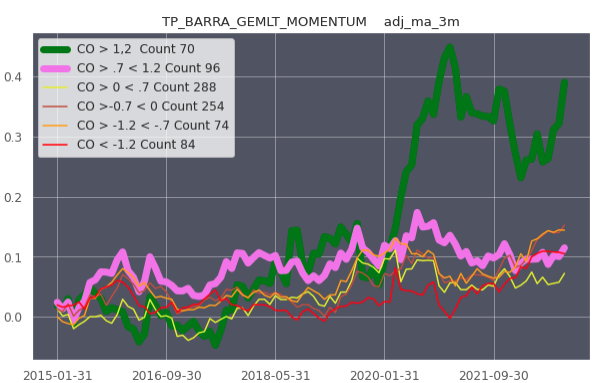
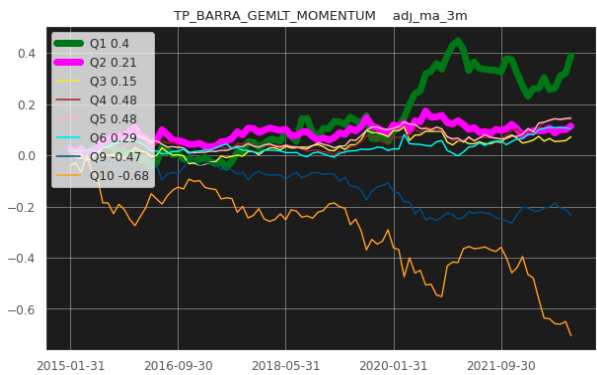
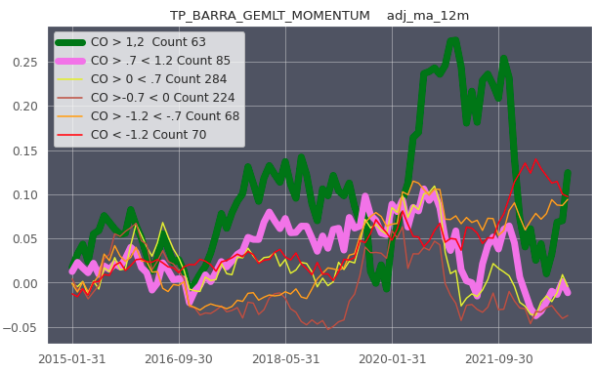
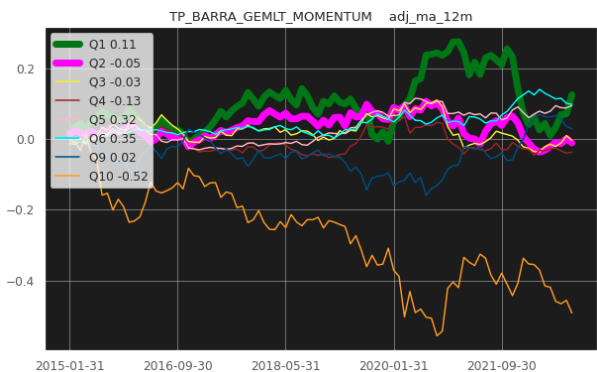
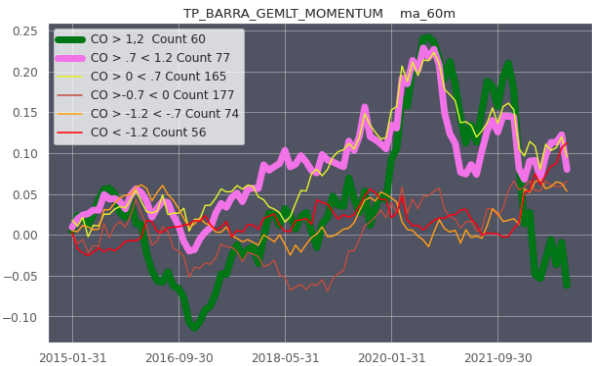
1. Calculate ex-BSL returns for the next month to determine the returns of each quantile.
2. In non-strict mode, do not completely drop the moving average if the rolling window contains NA values.
3. If the exposure is NA, remove it from the cutoffs, resulting in around 60+ stocks in each quantile.
4. Use exposure cutoffs based on z-scores, resulting in a higher number of stocks in each cutoff.
5. Group stocks by their exposures and divide them into equal portions to create the quantiles.
6. View the top and bottom exposure scores in the first quantile on the spread chart.

MOMENTUM

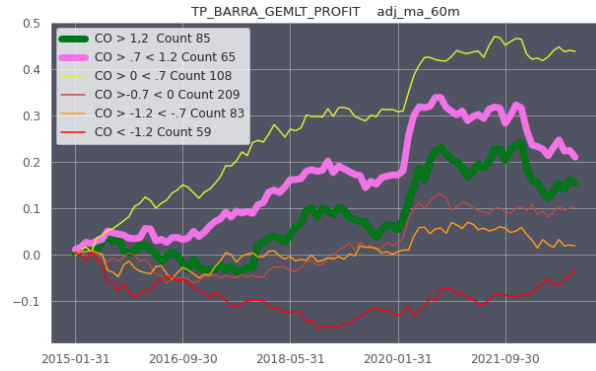
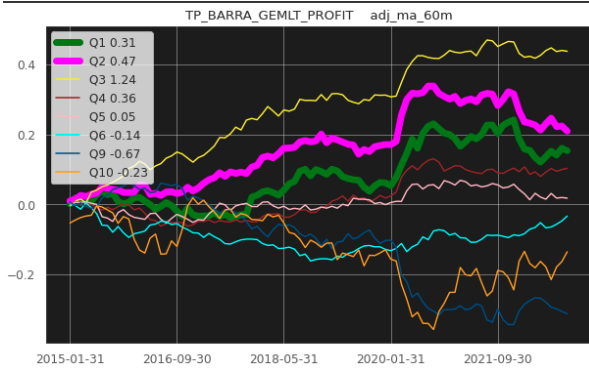
To view a larger image, hold the edge of the image and drag it to resize.



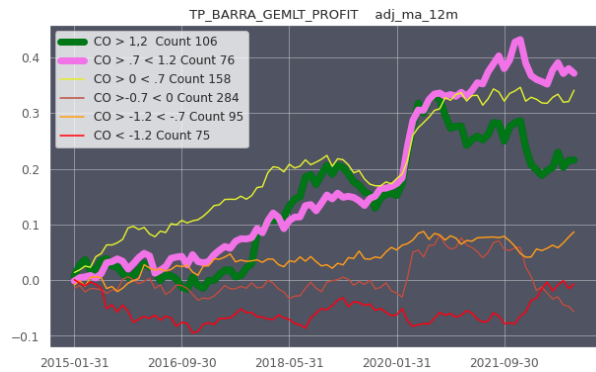
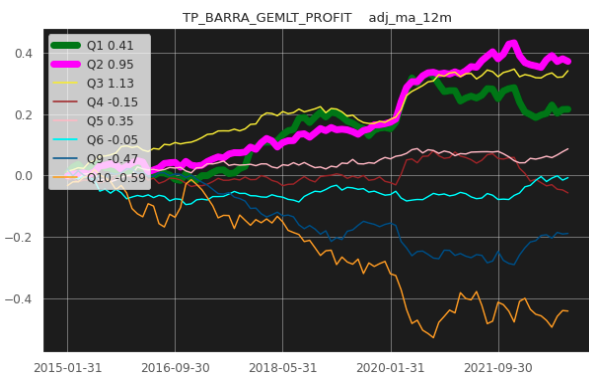
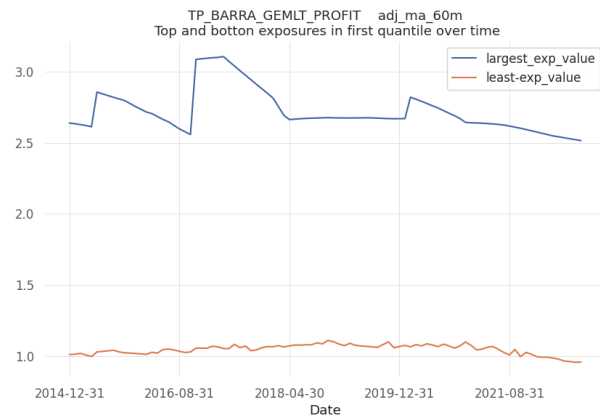
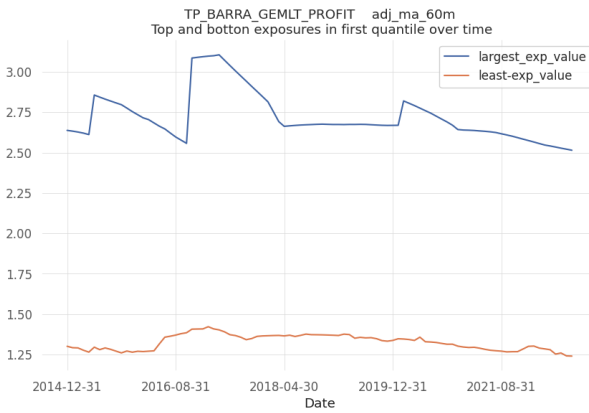
Show IR metrics under the chart. Just for 1st quantile



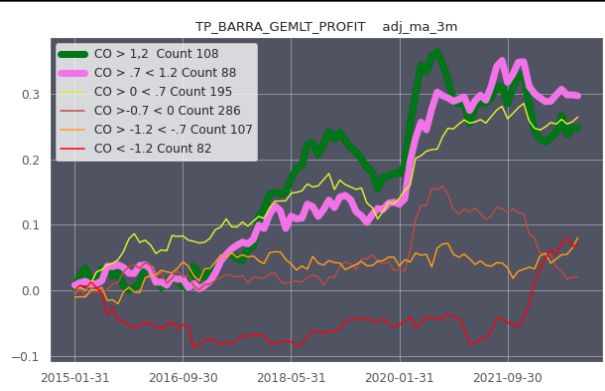
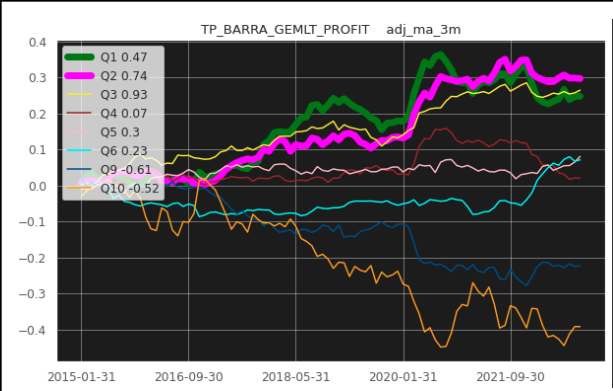
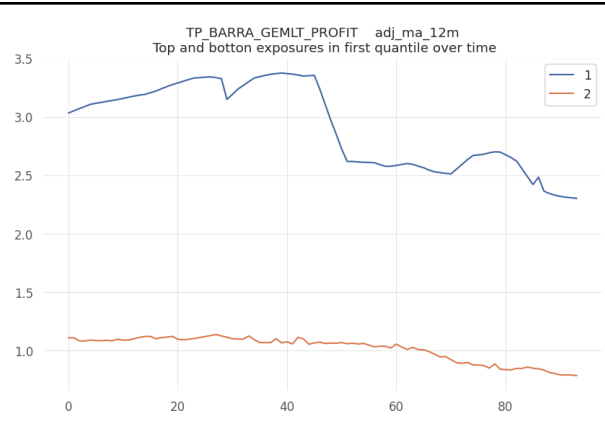
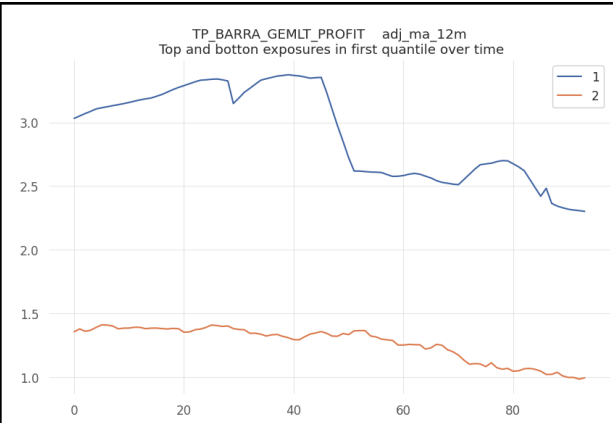
PROFIT



This chart displays the value of the highest and lowest exposure values for a factor at each time period for the first quantile portfolio.

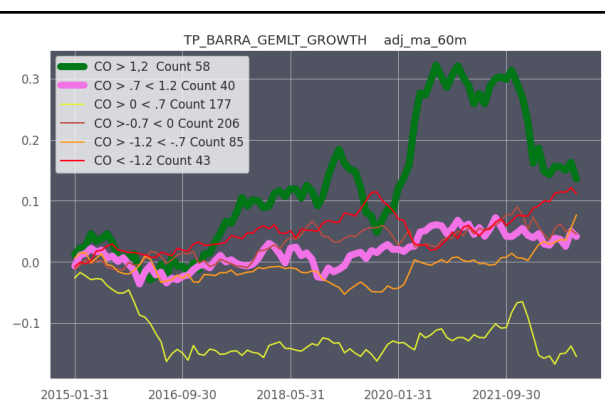
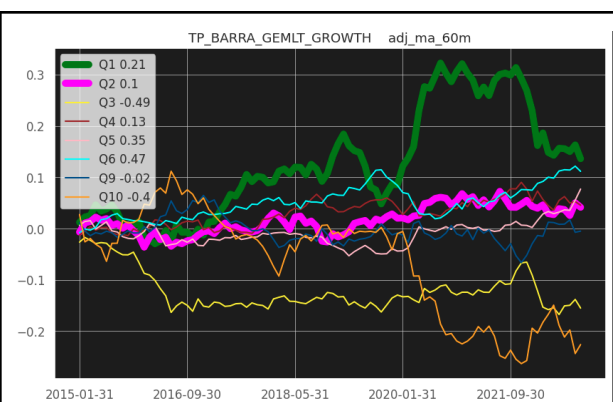


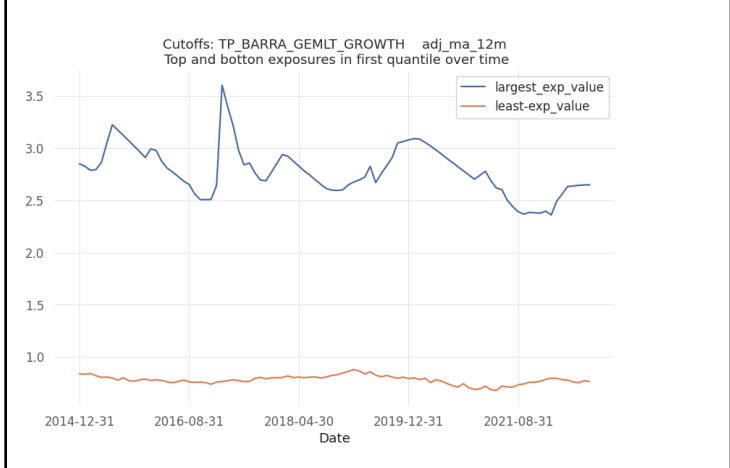
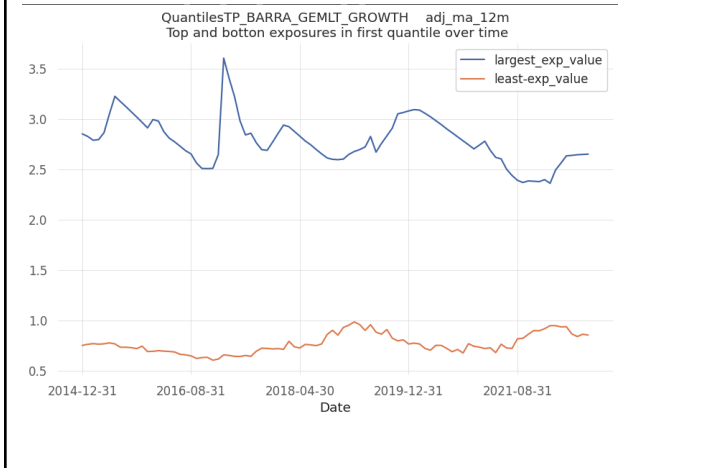
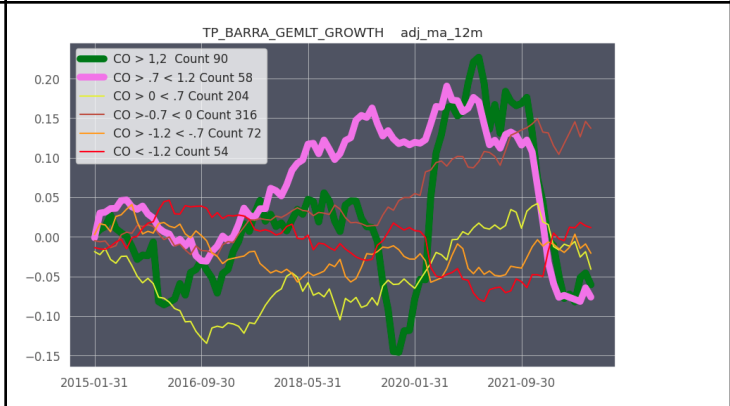
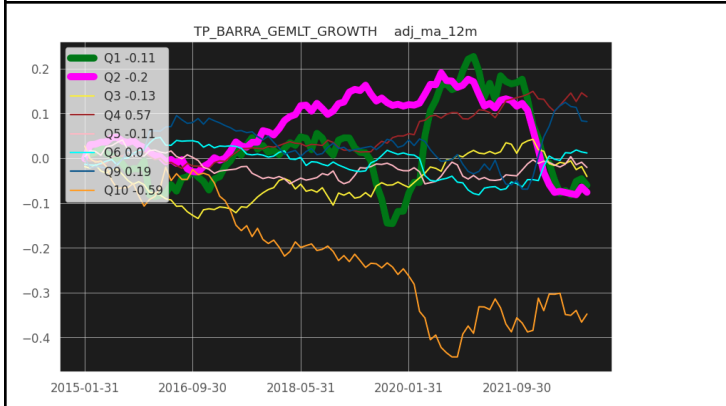
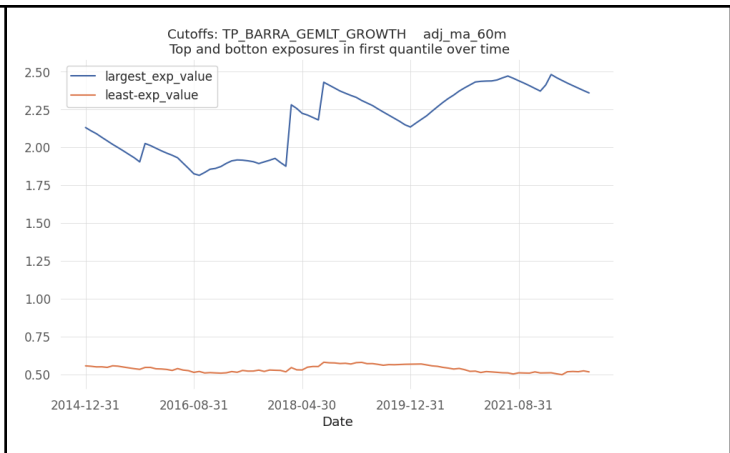
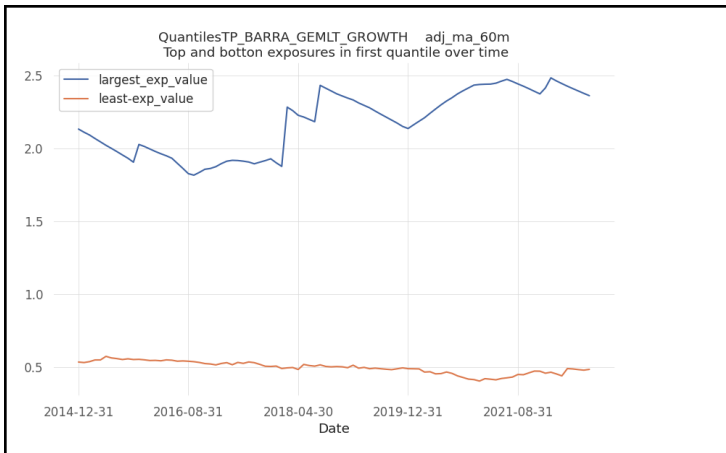
Any diagnostic checks??

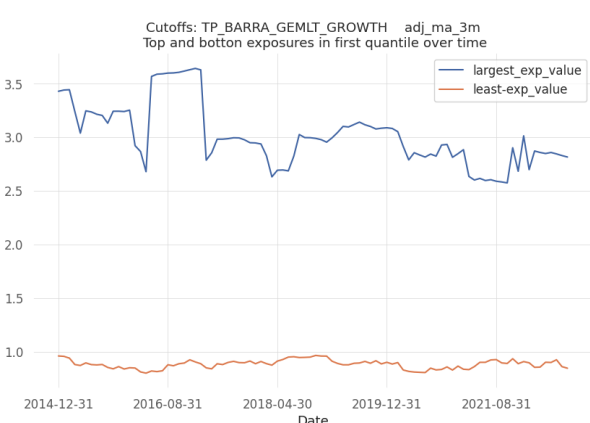
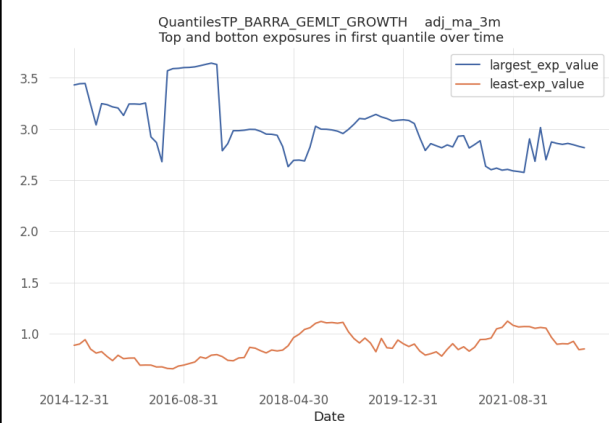
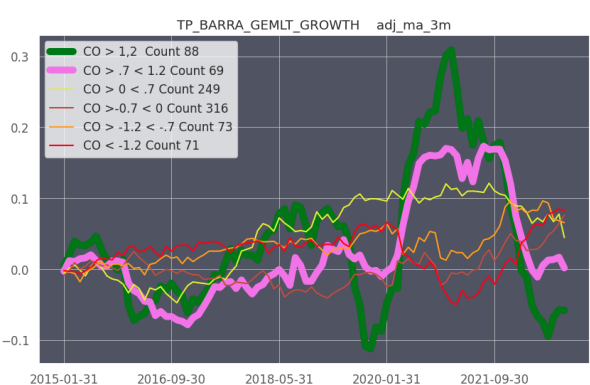
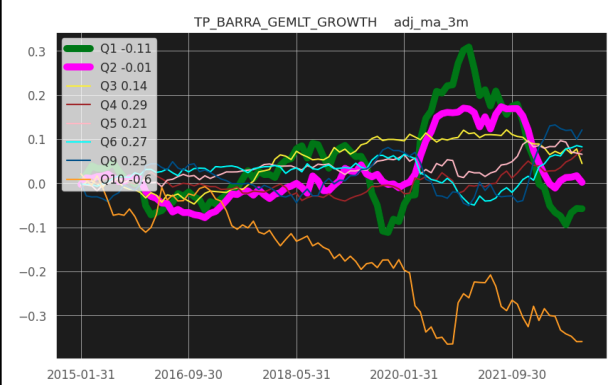


Negative cutoffs trending up. diagnostics?
Debug dates.

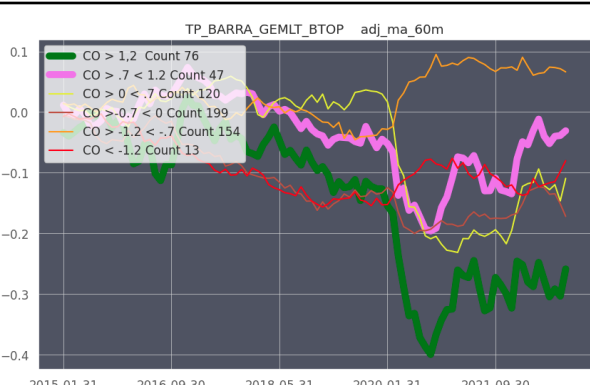
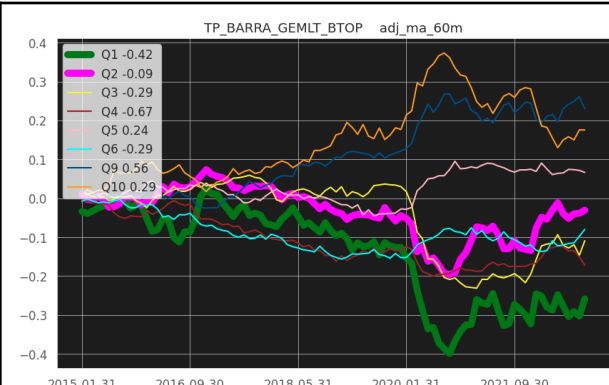
GROWTH

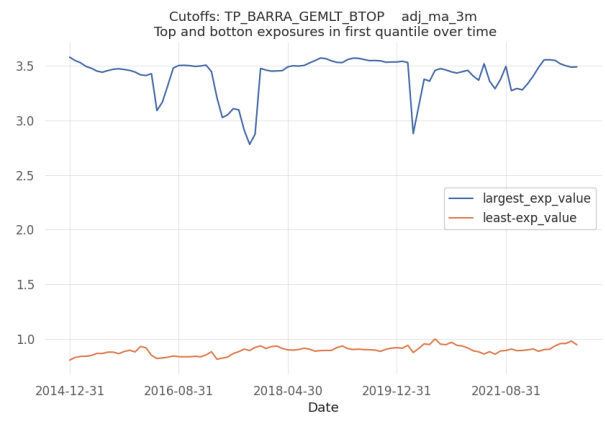
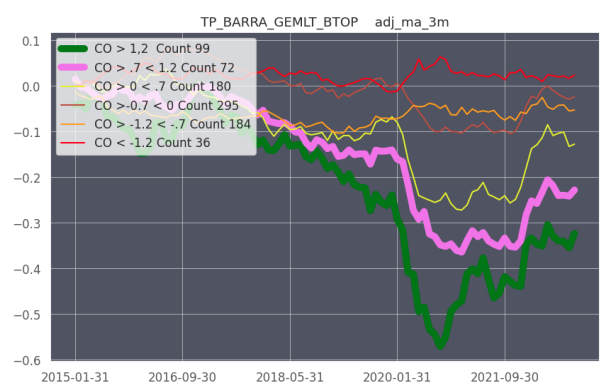
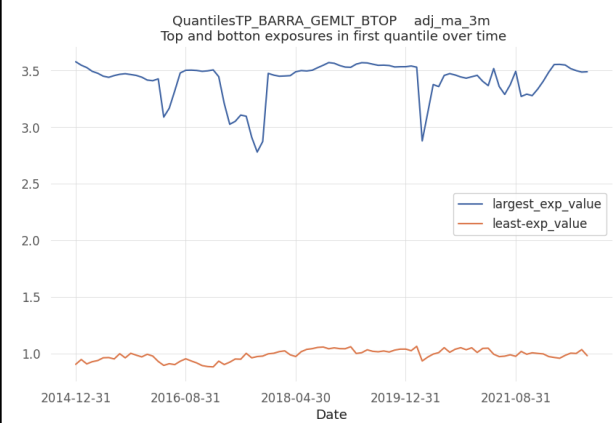
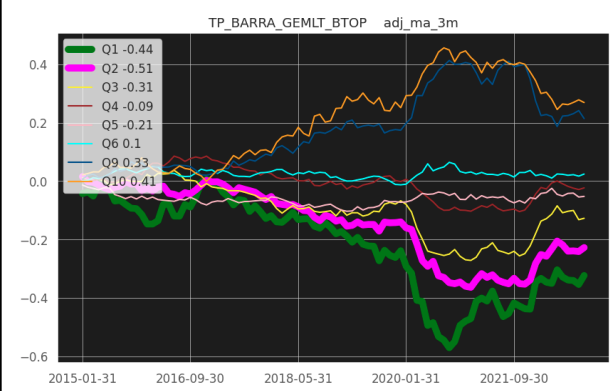
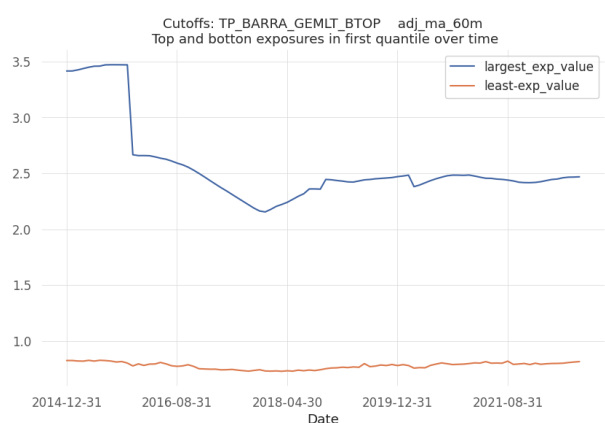
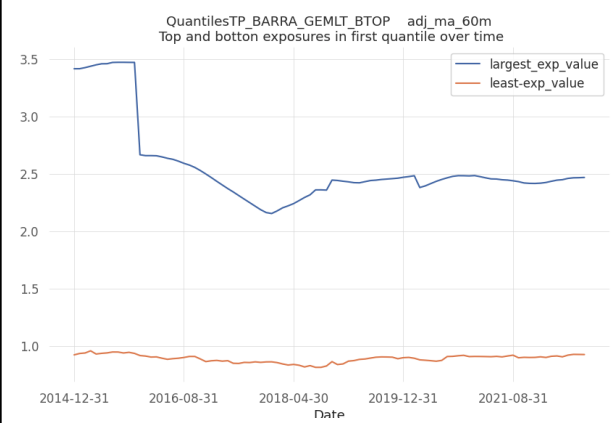




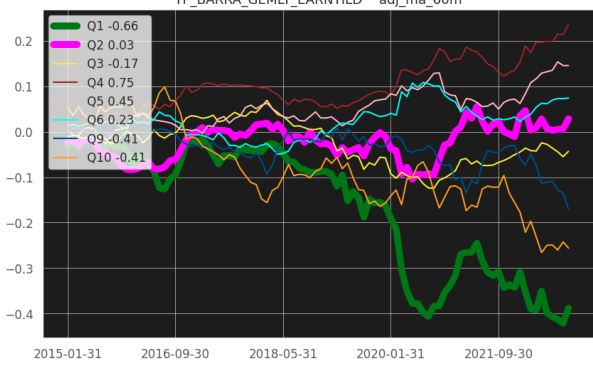
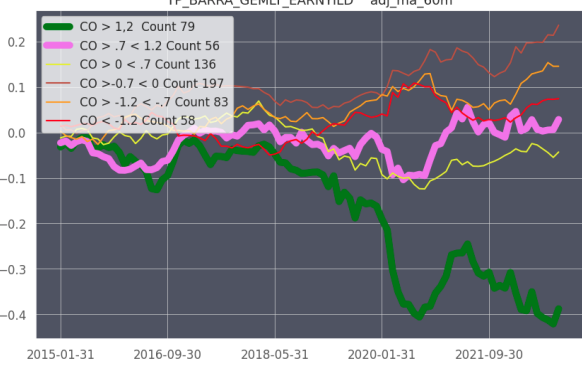


BTOP





EARN YIELD

	
Rerun with z-score instead of raw score for spread chart	
Average of the top quantile z-scores vs Average of bottom z-scores of bottom quantile. Remember to take z-score instead of raw score.	

By year end:

- > 1. Why are negative cutoffs trending up?
- > 2.Tile vs cutoffs

Two universes:**Dec 14, 2021**

US (Start Date: 2014-12-31?)

NA (North America) (1,3,6,9,12,2y,3y,5y Moving Averages)

Summary chart with information ratio. Pick the best 2 and show them in this table format

18 factors + 2 combo factors (important and distances of the combo).

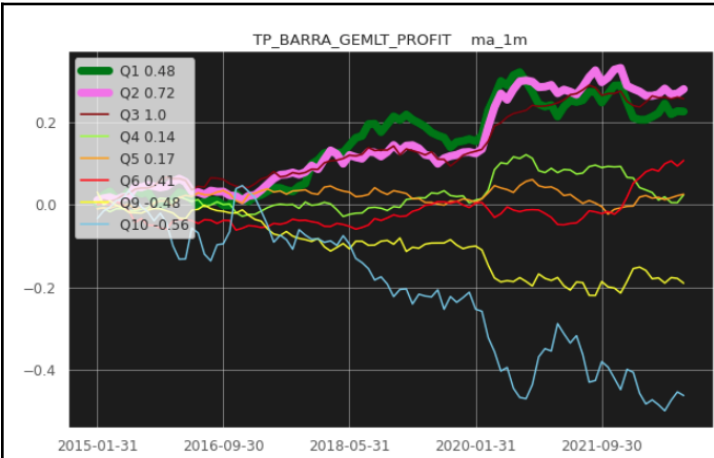
Compare the performance metrics for both approaches

Archive

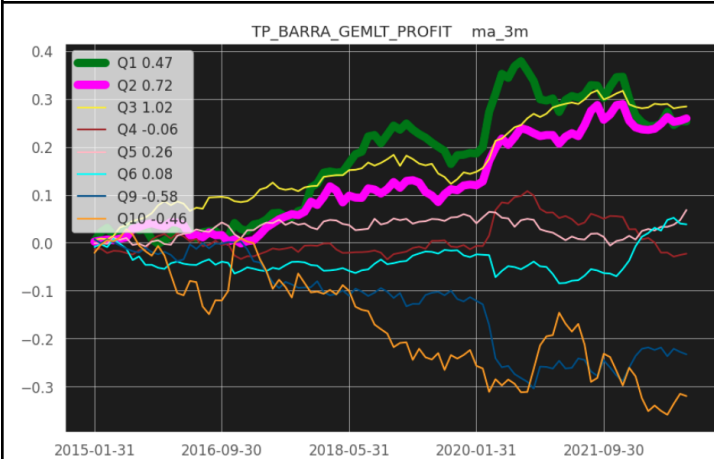
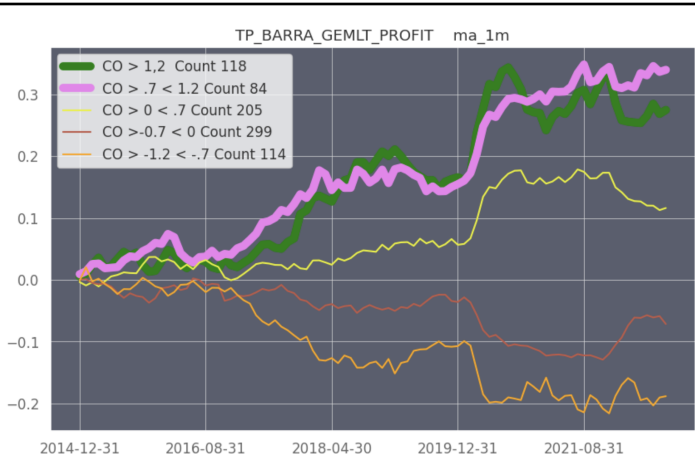
The MA approach was strict in this version.
QUANTILES vS CUTOFFS

[Draft Code](#)

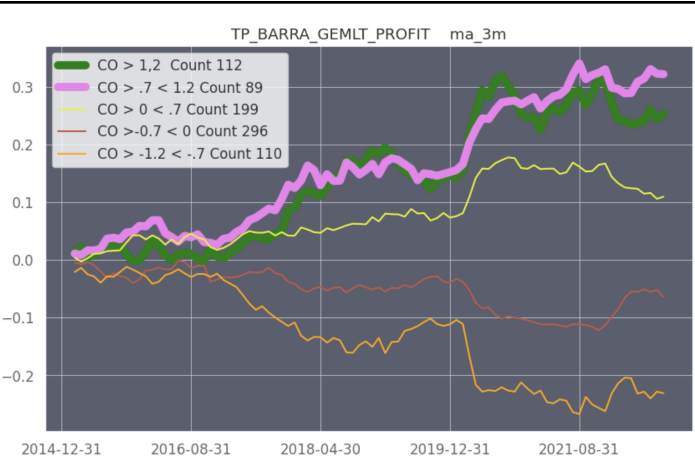
PROFIT



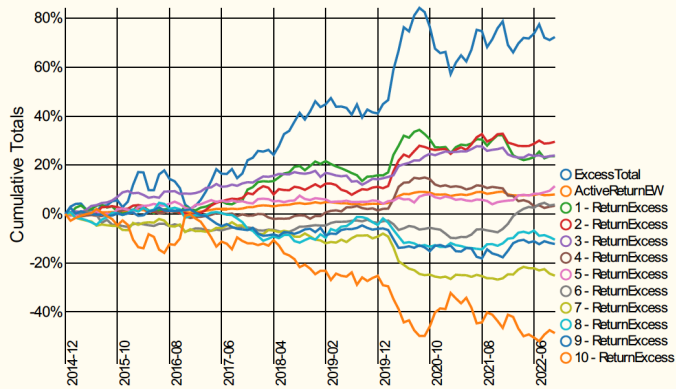
[Link to spreadsheet](#) with returns for all quantiles for 1m MA for PROFIT factor
Information ratio:



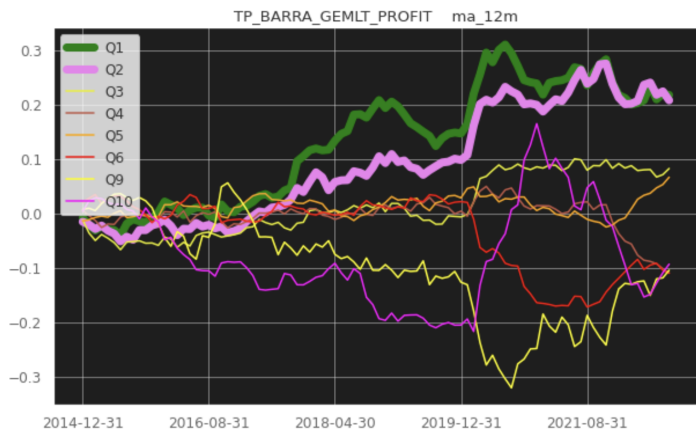
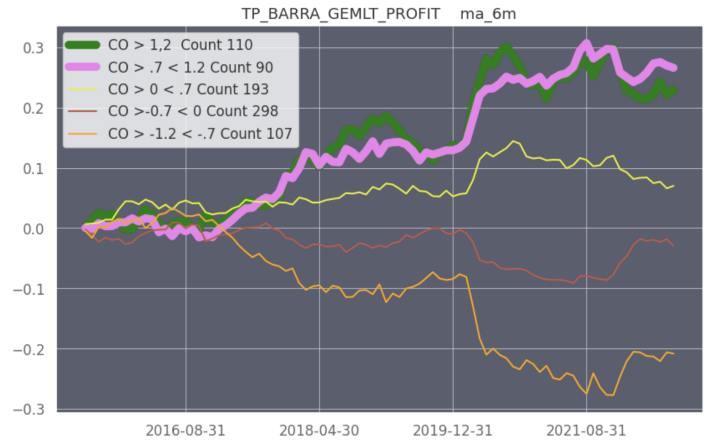
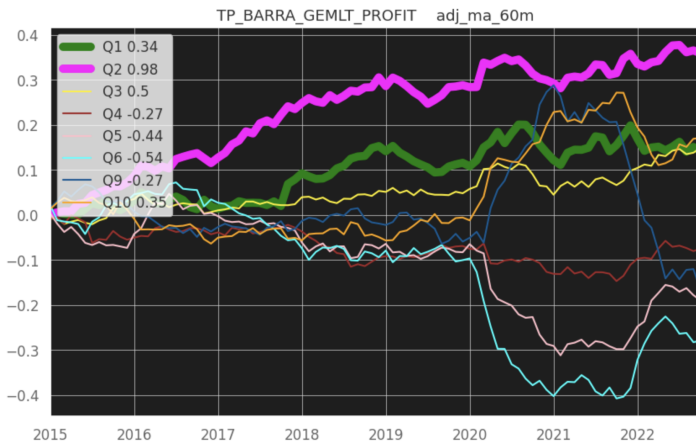
<Matches RK>

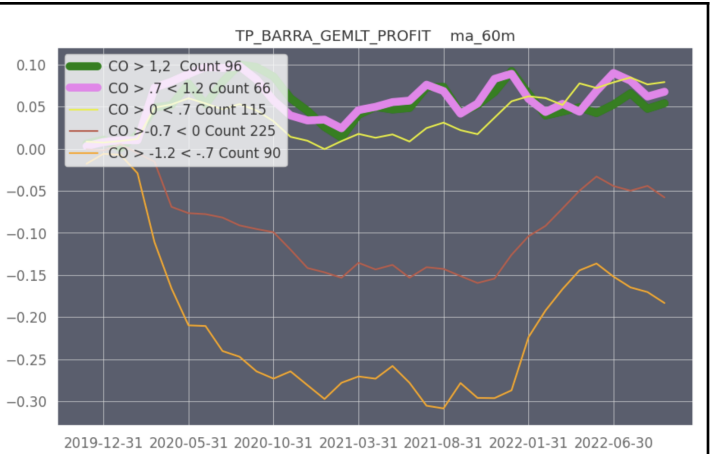
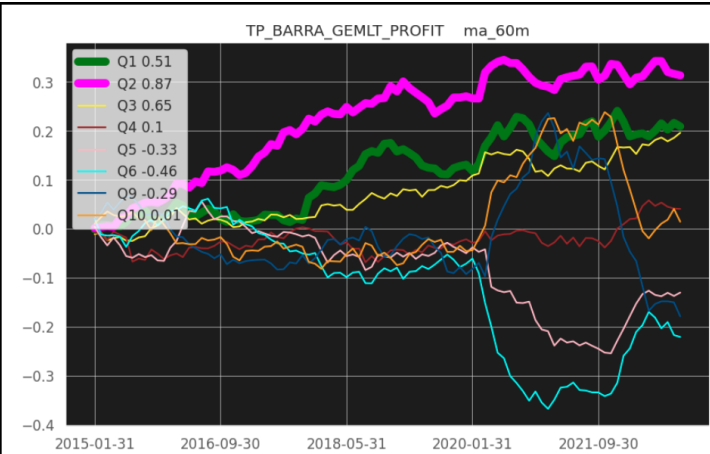


Cumulative Returns

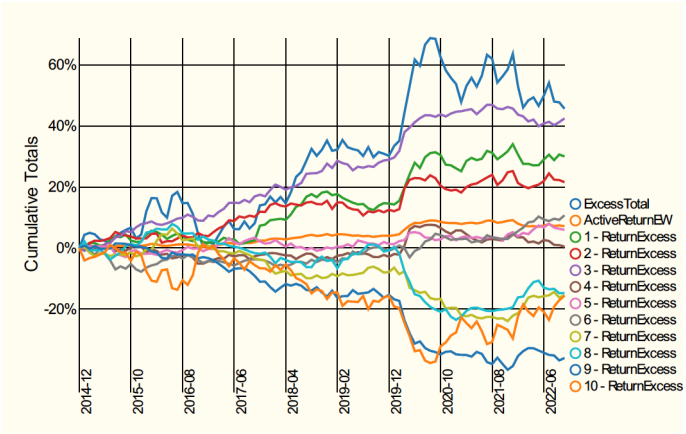


Verifying this will 100% cutoff for Moving average with NA



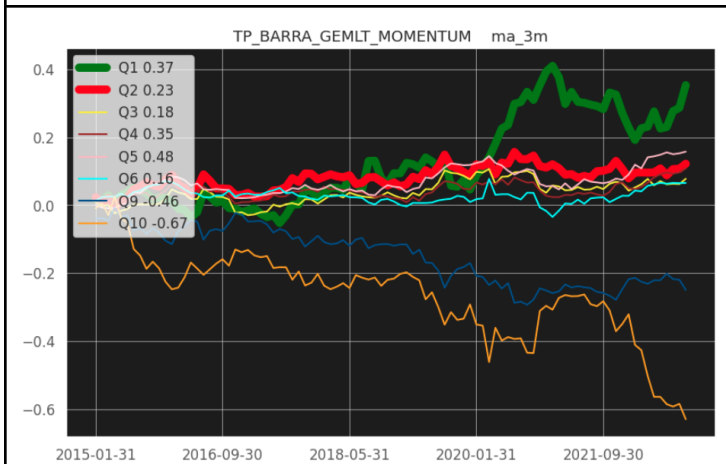
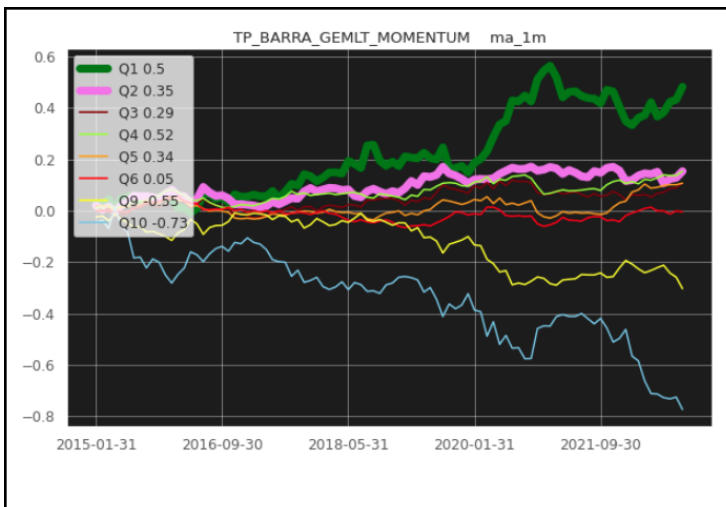


[100% strict. If the window contains even one NA, the mean is NA].
 <Matches>

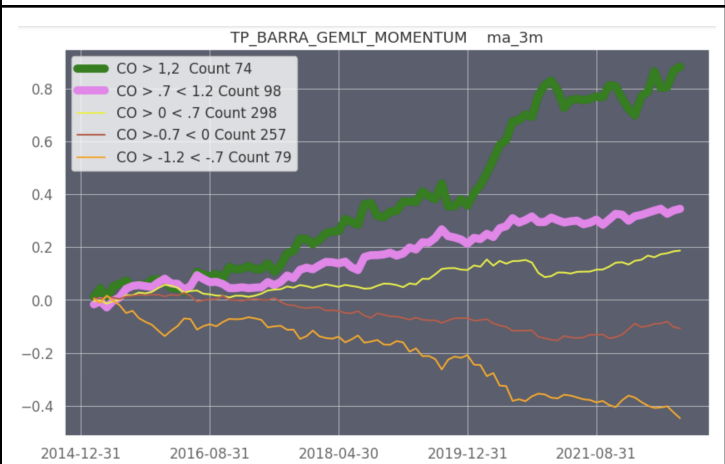
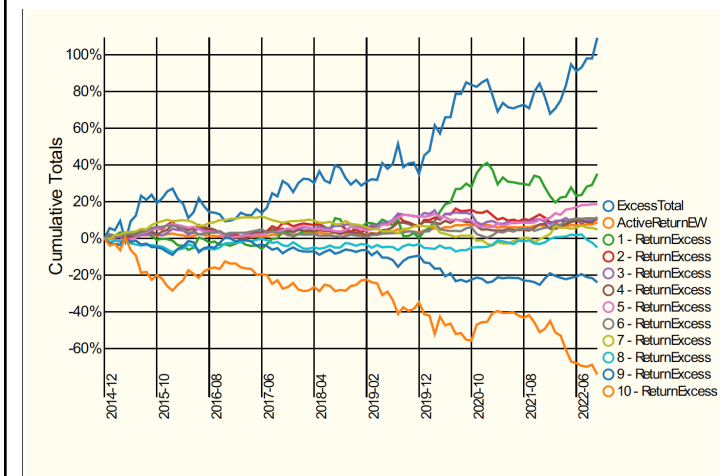


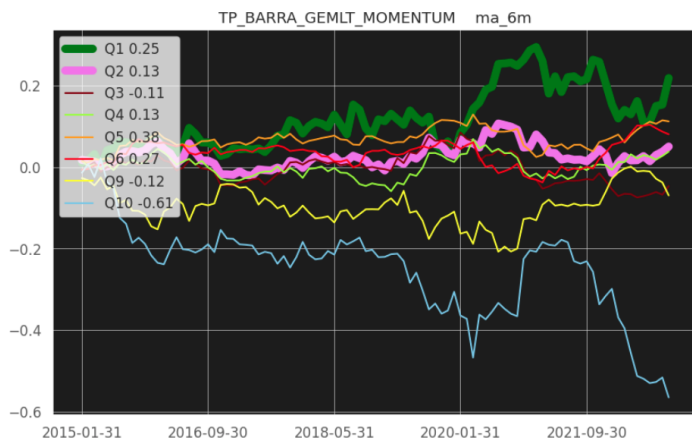
MOMENTUM

QUANTILES[With 1 month fwd excess returns]	CUTOFFS[use 1 M Fwd returns]

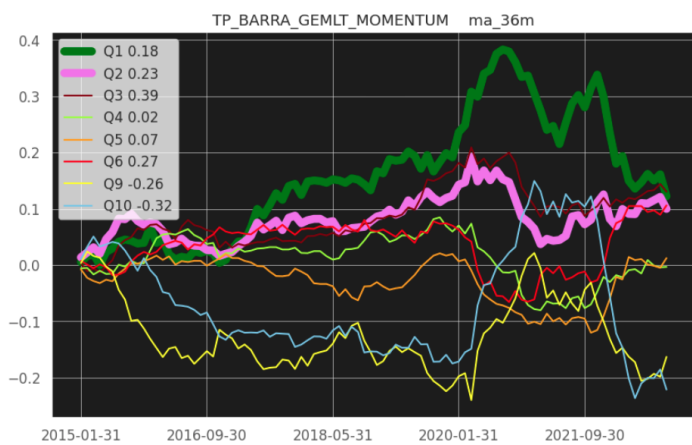
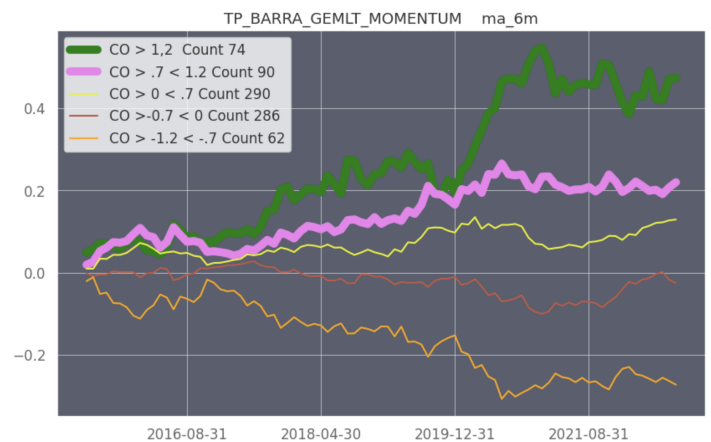
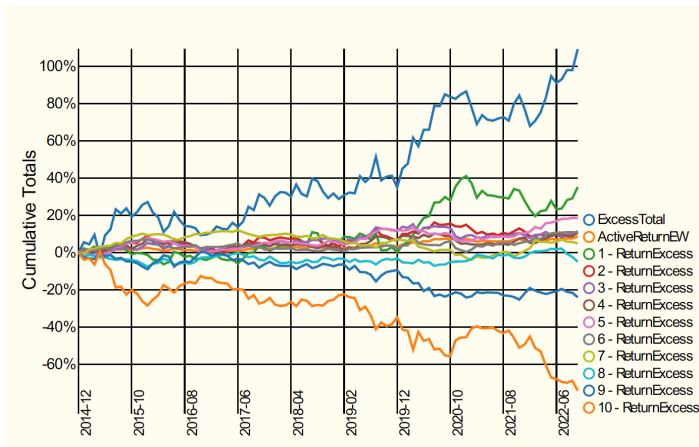


<Matches with RK>

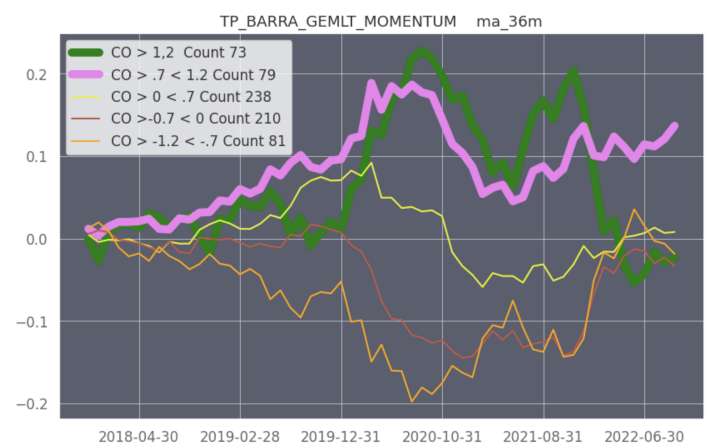




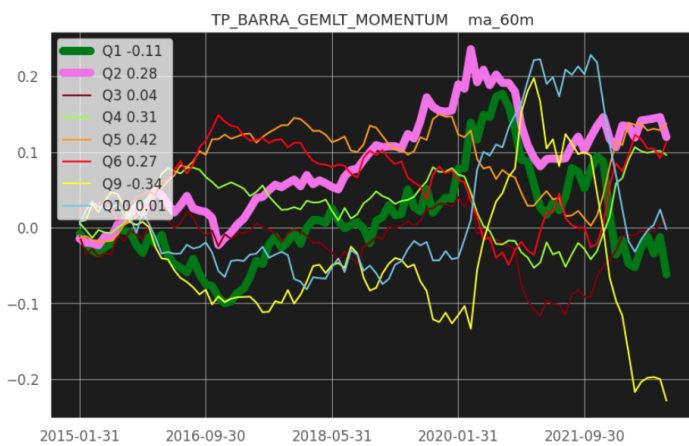
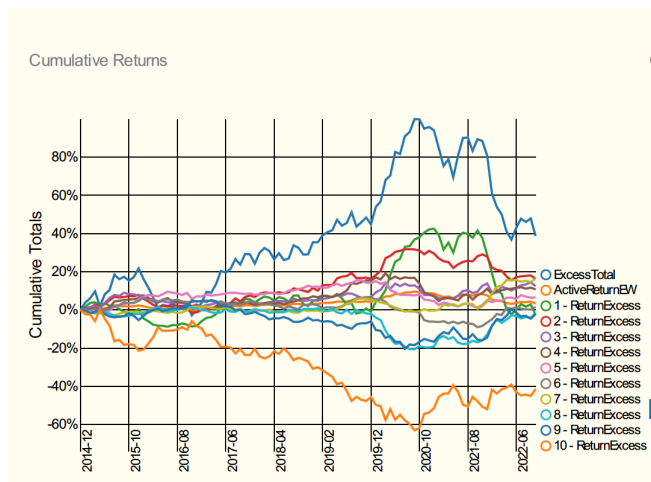
<Matches with RK>



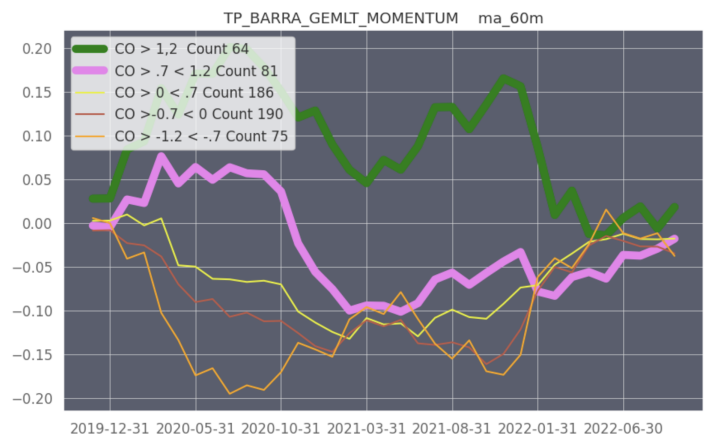
<Matches with RK>



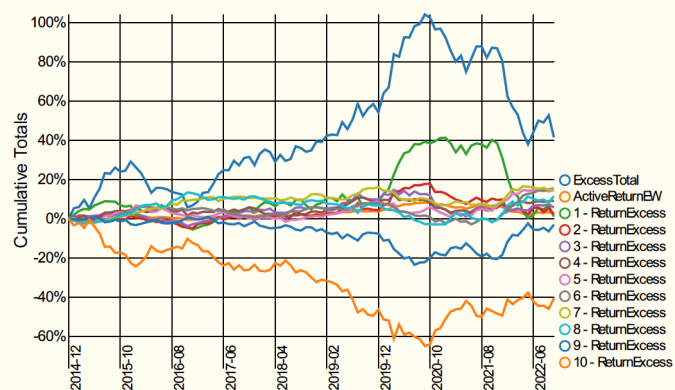
Are the constituents in quantile vs cutoffs similar?
Ask question



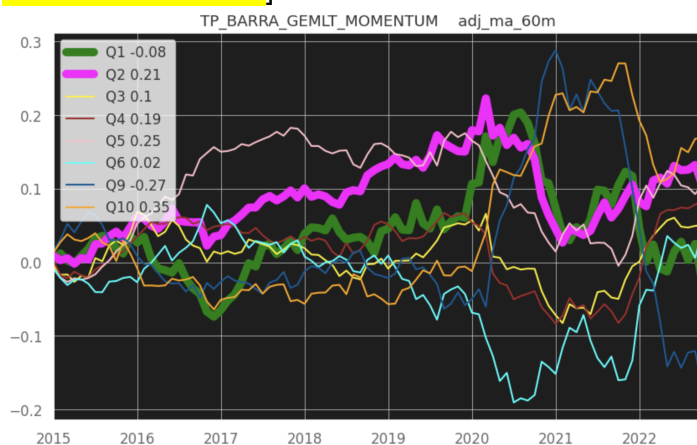
<Does not with RK>
Reason: Missing data for 3 months



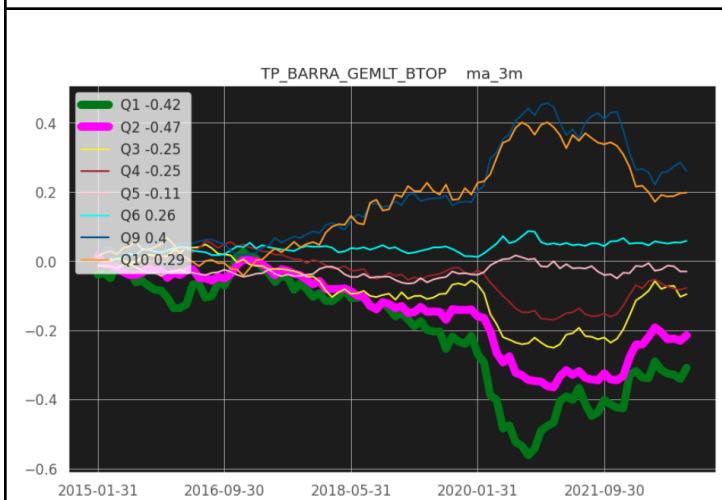
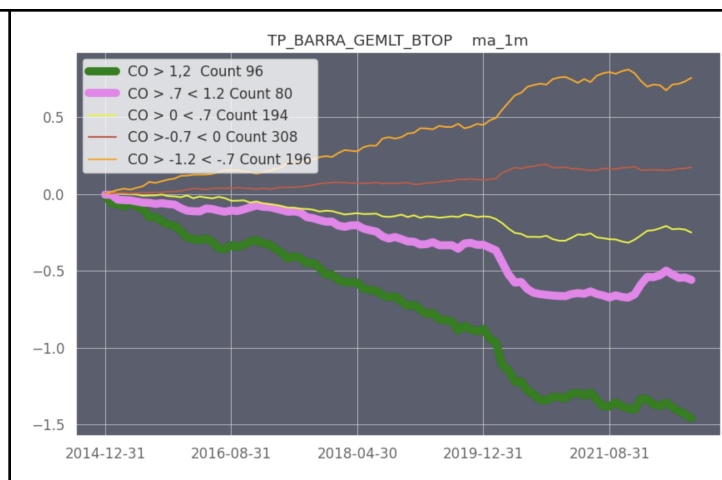
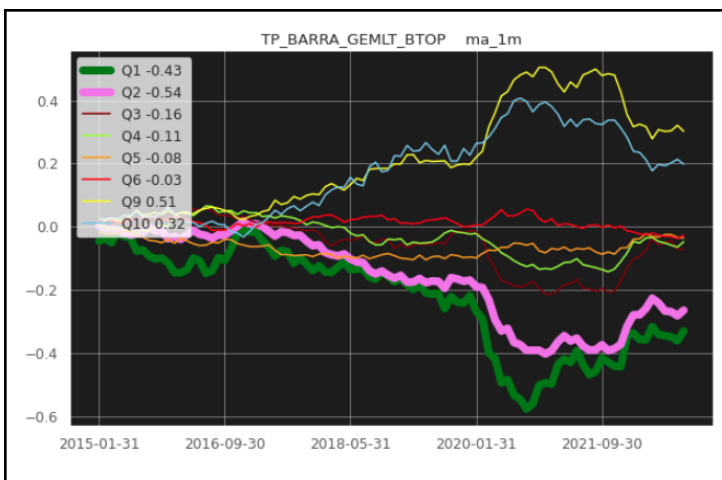
Cumulative Returns



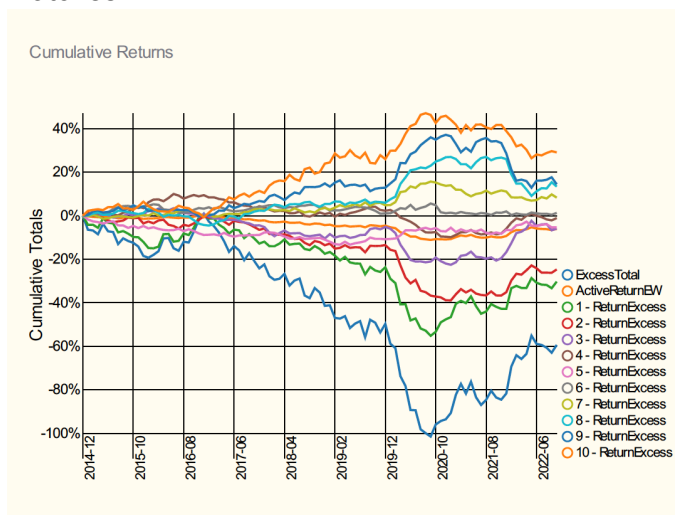
With adjustments for 100% strict. NA [.e if a rolling window contains NA, return NA, otherwise rolling mean. Dec 7th 2022]

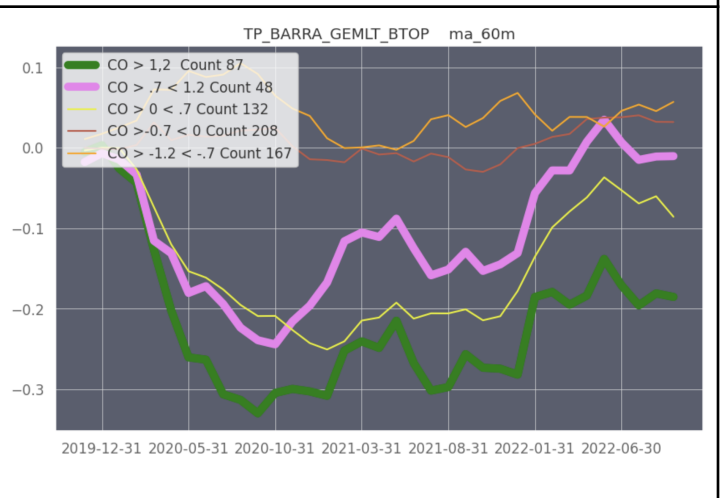
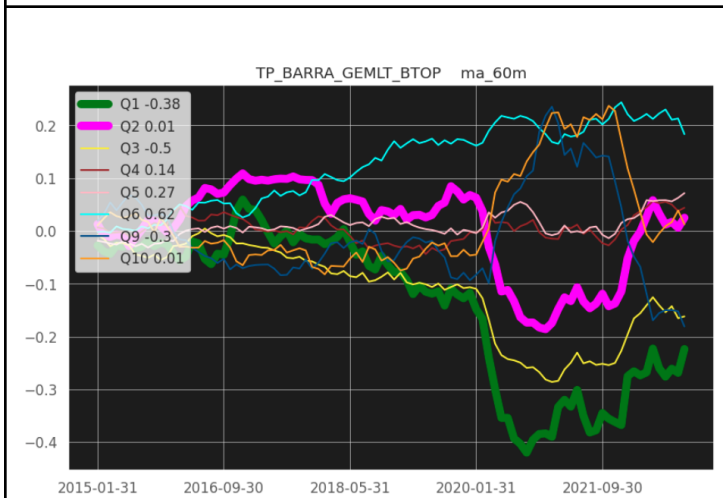
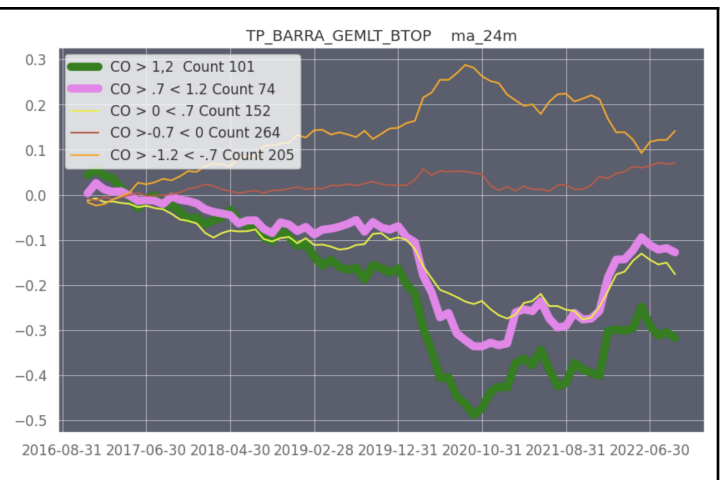
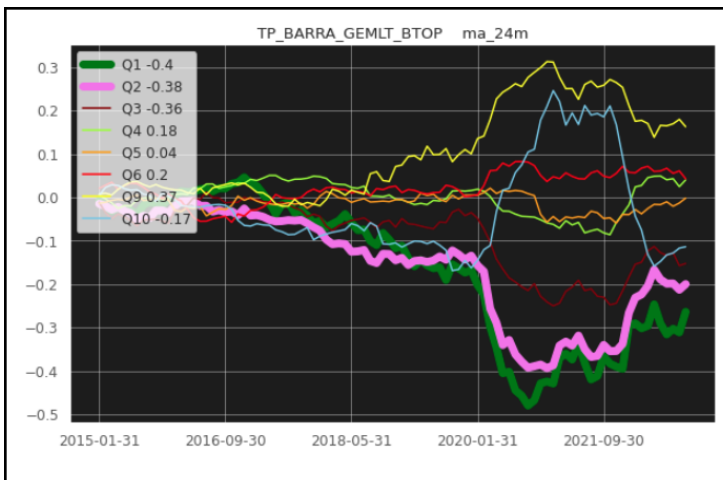


BTOP



<<Matches RK>>

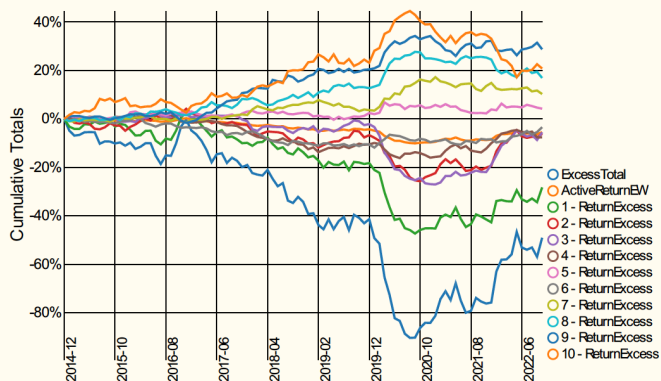




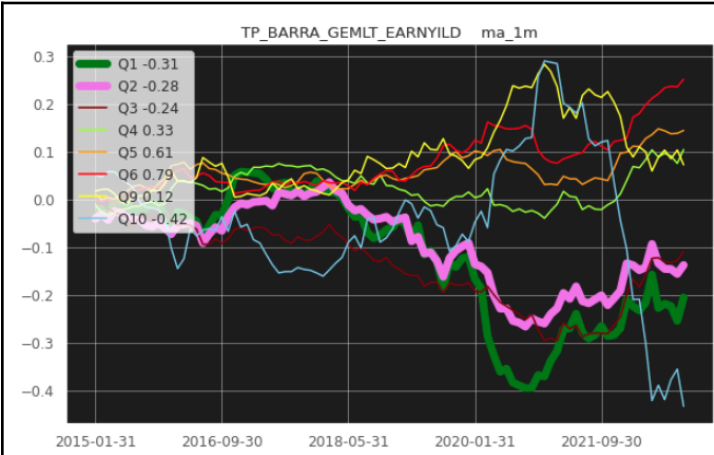
<Top quantile match>

<Check after adding missing data for 3 months>>

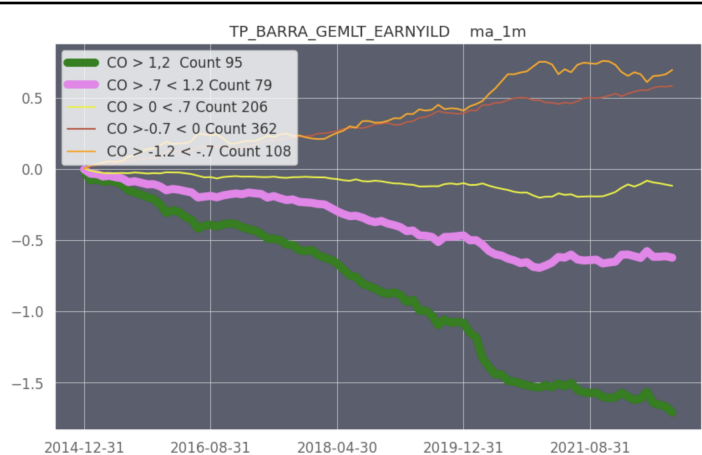
Cumulative Returns



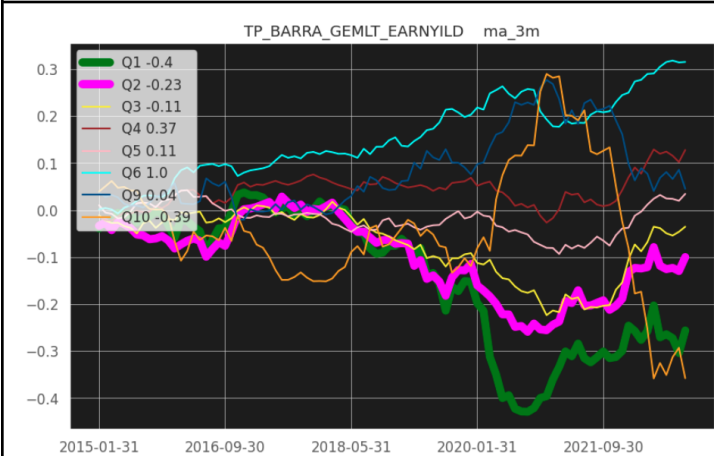
EARN YIELD



Within the green decile. Bad/good performers.

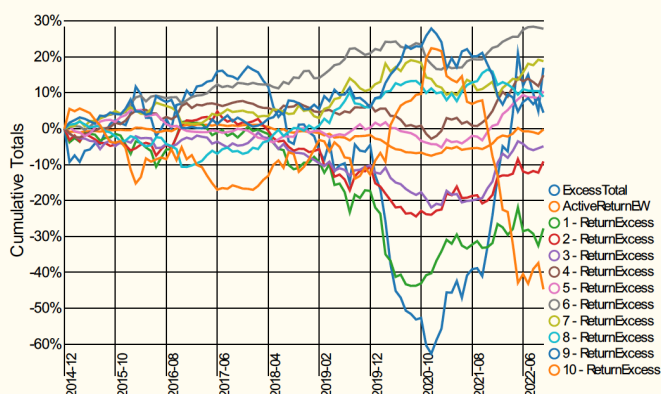


Verify.

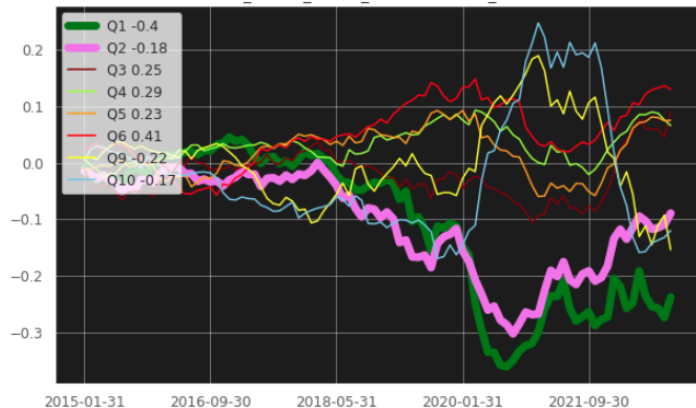


<<Matches with RK>>

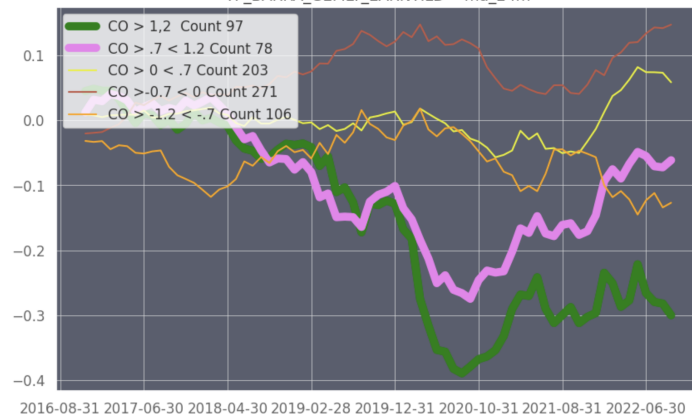
Cumulative Returns



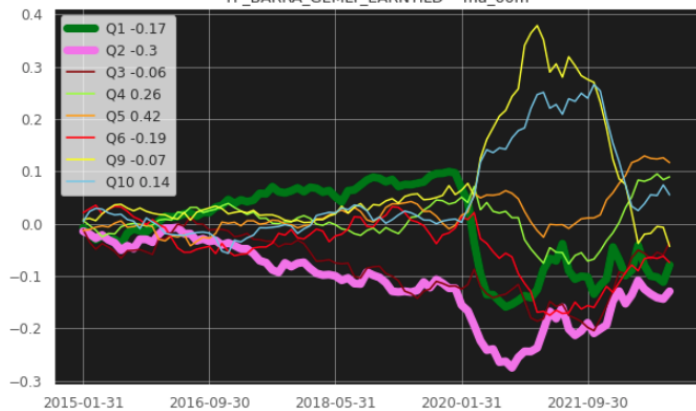
TP_BARRA_GEMLT_EARNYILD ma_24m



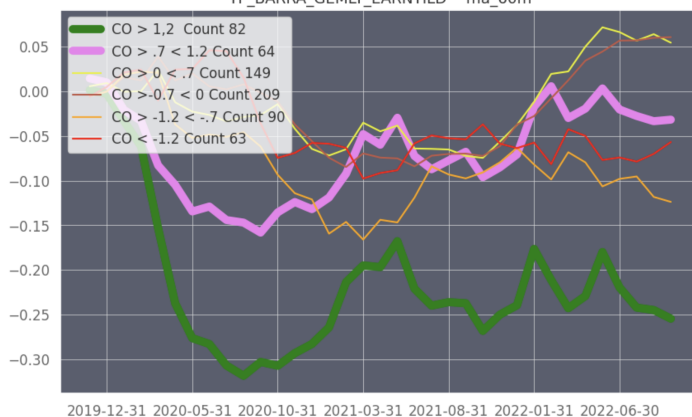
TP_BARRA_GEMLT_EARNYILD ma_24m



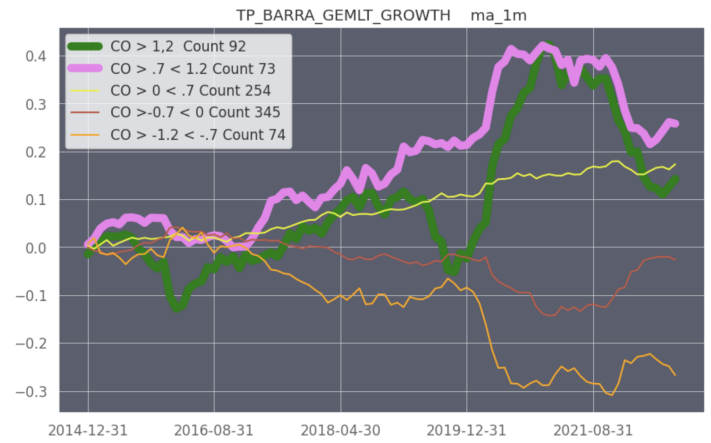
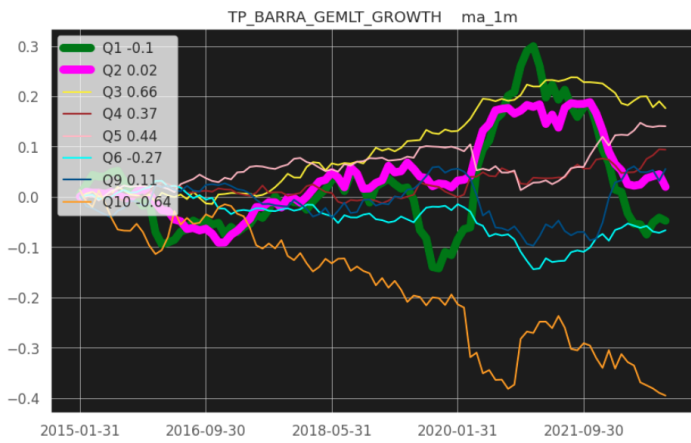
TP_BARRA_GEMLT_EARNYILD ma_60m



TP_BARRA_GEMLT_EARNYILD ma_60m

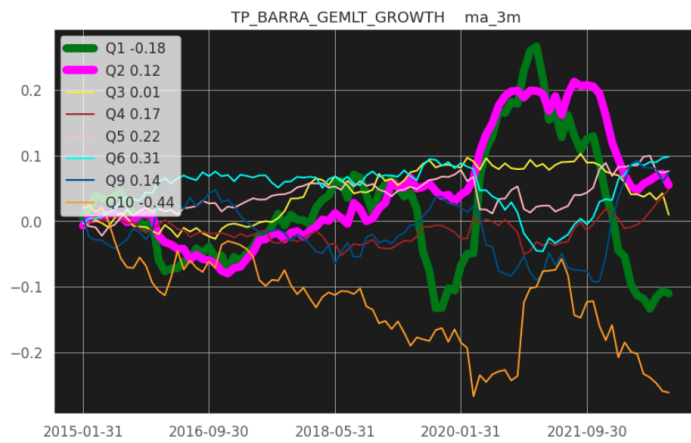


GROWTH

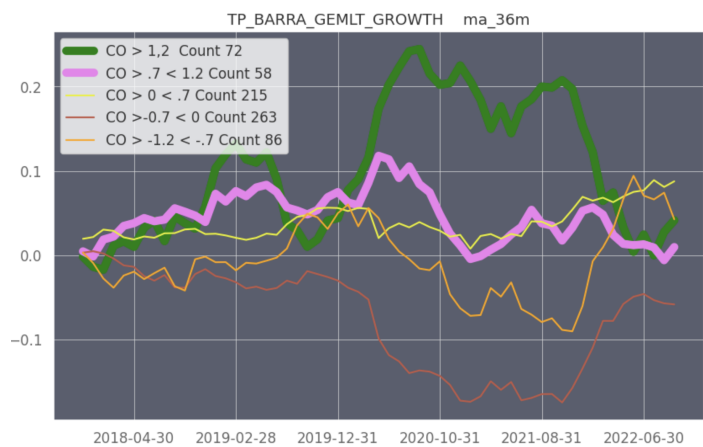
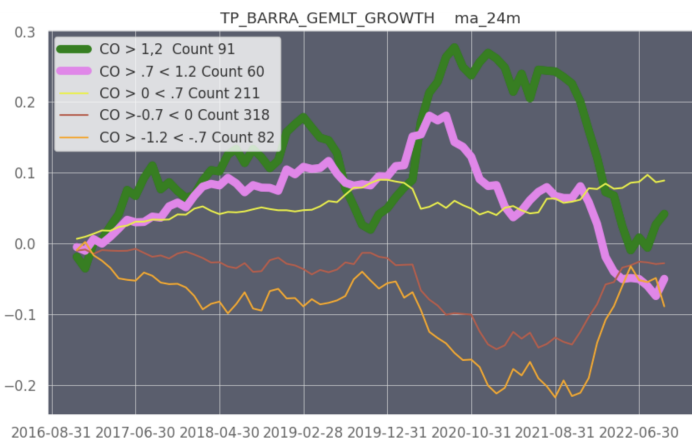
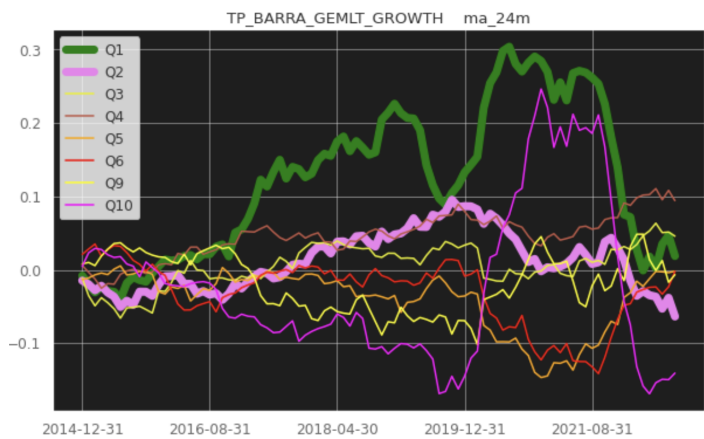
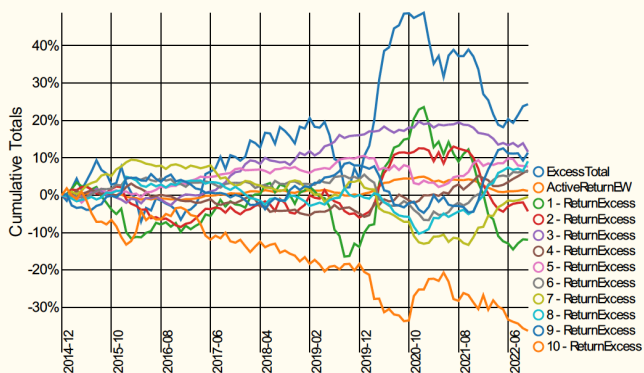


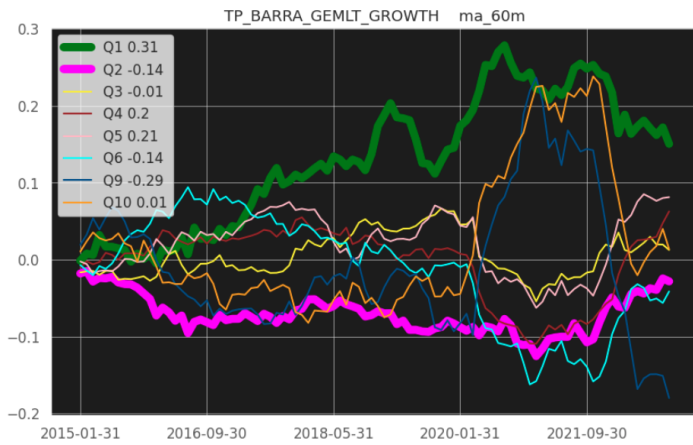
Try: Split top tile. ?Sample size. Deviation of bottom stock in top tile? Width of deviation top stock vs bottom stock over time.

<<Matches RK>>



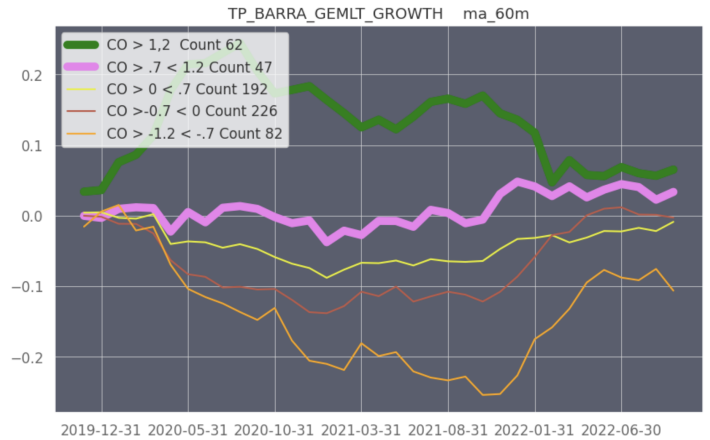
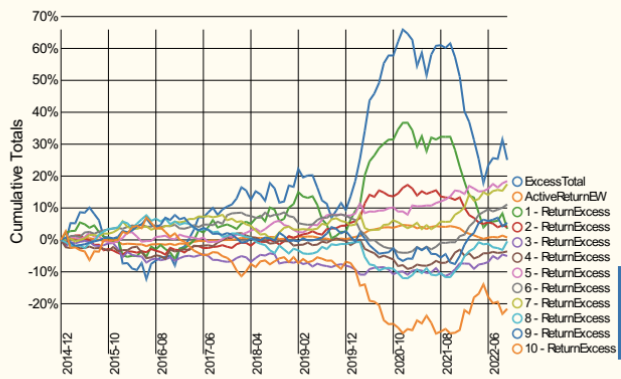
Cumulative Returns





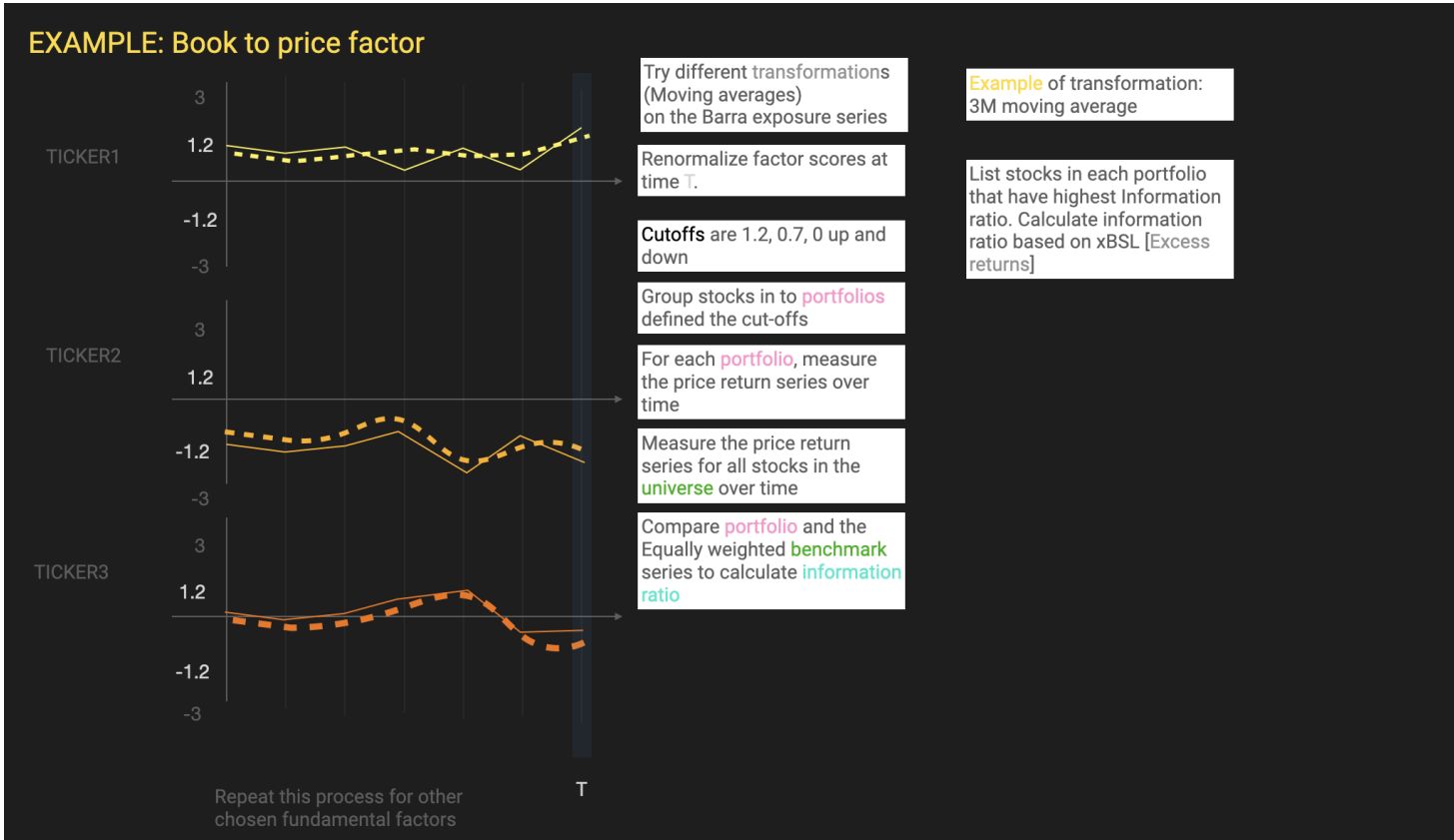
< Does not match RK>
 CCheck after adding missing data>

Cumulative Returns



2019-12-31 2020-05-31 2020-10-31 2021-03-31 2021-08-31 2022-01-31 2022-06-30

METHODOLOGY: Static Portfolios



Transformations

	Moving average (ma)						
	Moving Difference (md)						
	ma_1m	ma_3m	ma_6m	ma_12m	ma_24m	ma_36m	ma_60m
	md_3m		md_6m	md_12m	md_24m	md_36m	md_60m
	ma_12m_60m		ma_24m_60m	ma_36m_60m			
Earning Yield	Create portfolios Quantiles [same # of holdings] Cutoffs. Analyze relative performance of these portfolios						
Book To Price							
Momentum							
Price							
Growth							

Notebook:
[Link](#)

Files Location

My Drive > ... > SAVED_DATA > OCT31 Name TP_BARRA_GEMLT_PROFIT_oct31_cross_section.csv TP_BARRA_GEMLT_MOMENTUM_oct31_cross_section.csv TP_BARRA_GEMLT_GROWTH_oct31_cross_section.csv TP_BARRA_GEMLT_EARNYILD_oct31_cross_section.csv TP_BARRA_GEMLT_BTOP_oct31_cross_section.csv Transformations score for all stocks Cross sectional transformation scores	Performance charts and Constituents list of each portfolio Quantiles Searchable spreadsheet Cutoffs Searchable spreadsheet
---	---

Spreadsheet with Performance Metrics

[Best performing portfolios](#)

for all 900+ stocks. Report run on June 30th data. Run on Oct 31st, 2022 Link		
---	--	--

Top 2 transformations for each feature

Feature: Profit

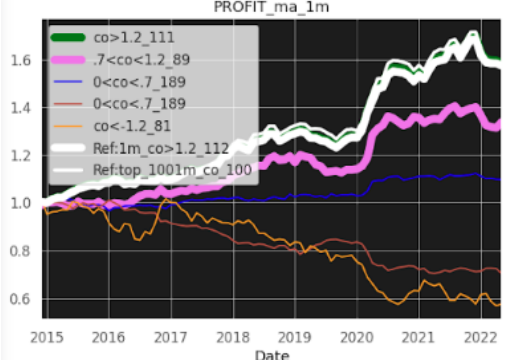
Transformation: PROFIT_ma_1m

These charts show the performance of the quantiles/cutoffs relative to the equal weighted benchmark (all stocks portfolio)

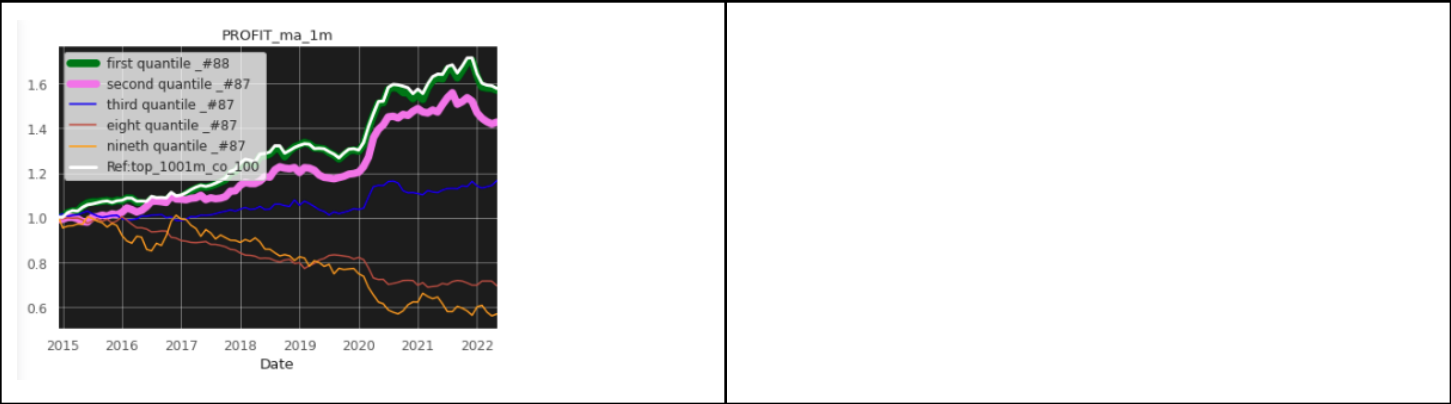
The tag PROFIT_ma_1m_co>1.2 means, the feature analyzed is PROFIT, transformation is: rolling moving average 1 month period, c0>1.2 means, the portfolios are created based on cutoffs and in this case the portfolio includes stocks whose cutoff for moving average of the feature on the backtest data (june 2022) is over 1.2.

The **highlighted value** is the information ratio for this portfolio

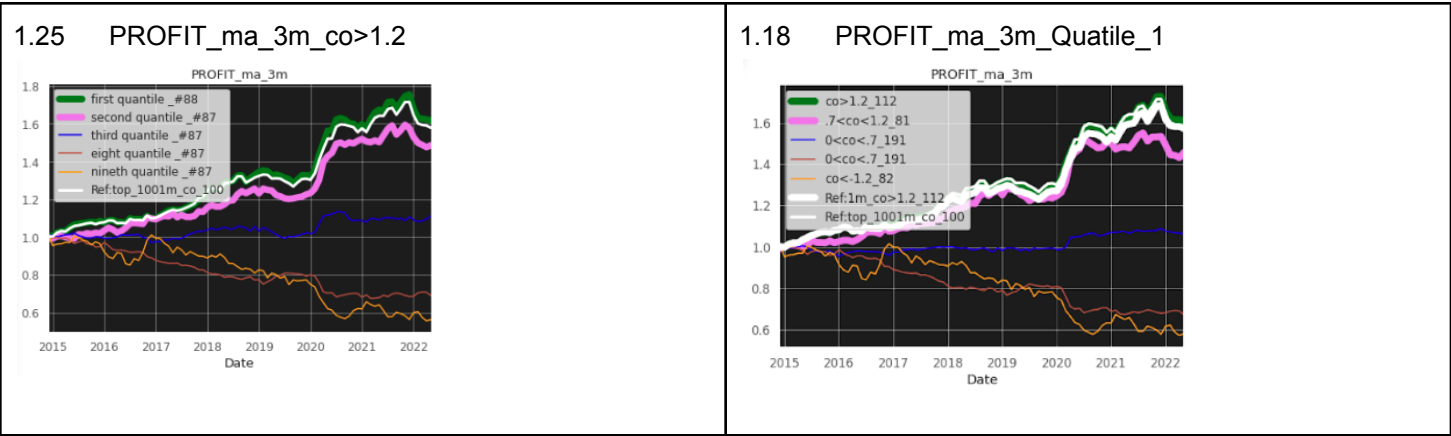
Best performing transformation and cutoff

1.24 PROFIT_ma_1m_co>1.2 	White line is overlaid to compare with performance without transformation for a feature.
---	---

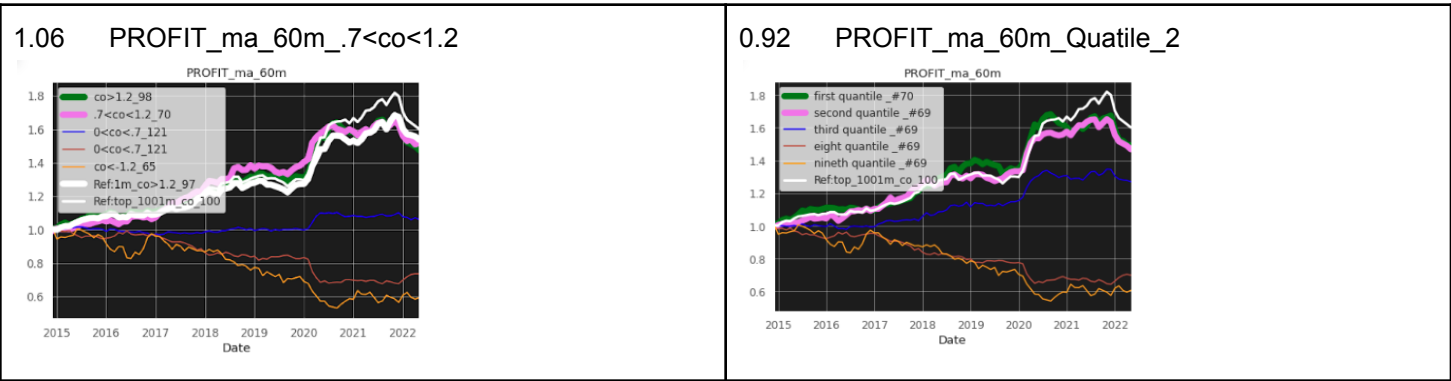
1.14 PROFIT_ma_1m_Quatile_1	Second best performing transformation and cutoff
-----------------------------	--



Transformation: PROFIT_ma_3m



Transformation:PROFIT_ma_60m

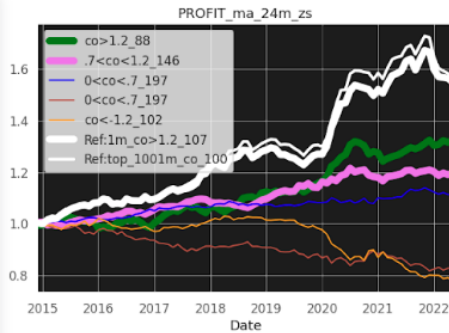


Transformation: PROFIT_ma_24m_zs

0.99
PROFIT_ma_24m_zs_Quatile_1|PROFIT_ma_24m_zs



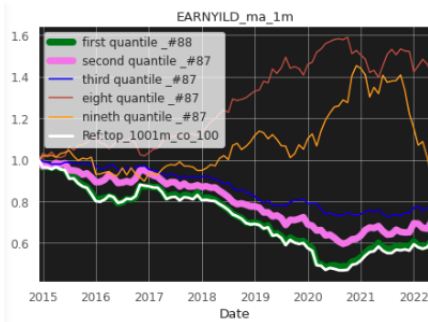
.92
PROFIT_ma_24m_zs_co>1.2



|
Feature: Earning Yield

EARNYILD_ma_1m

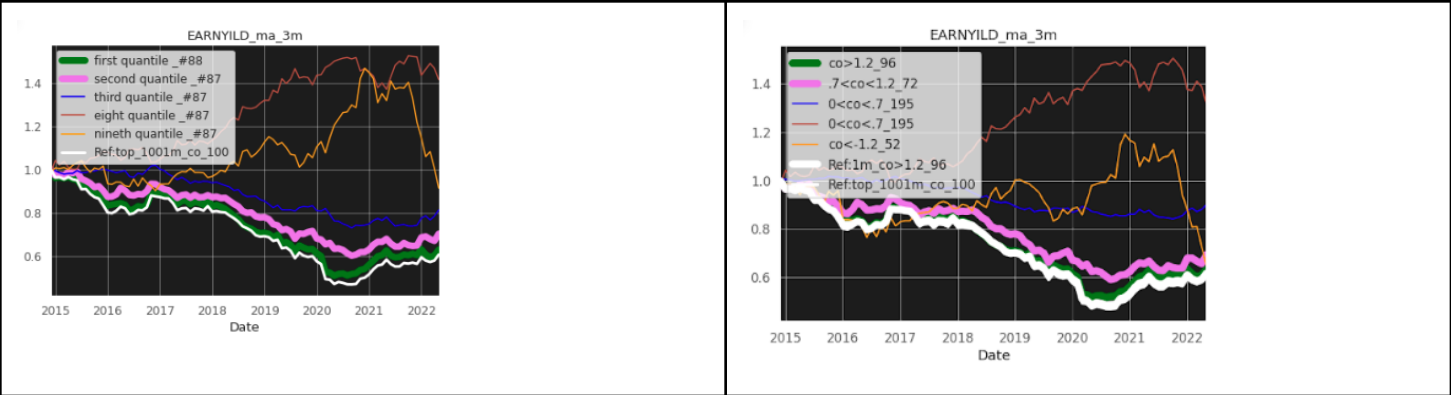
0.6
EARNYILD_ma_1m_Quatile_8



EARNYILD_ma_3m

0.63
EARNYILD_ma_3m_Quatile_8

0.47
EARNYILD_ma_3m_-.1.2<co<-.7

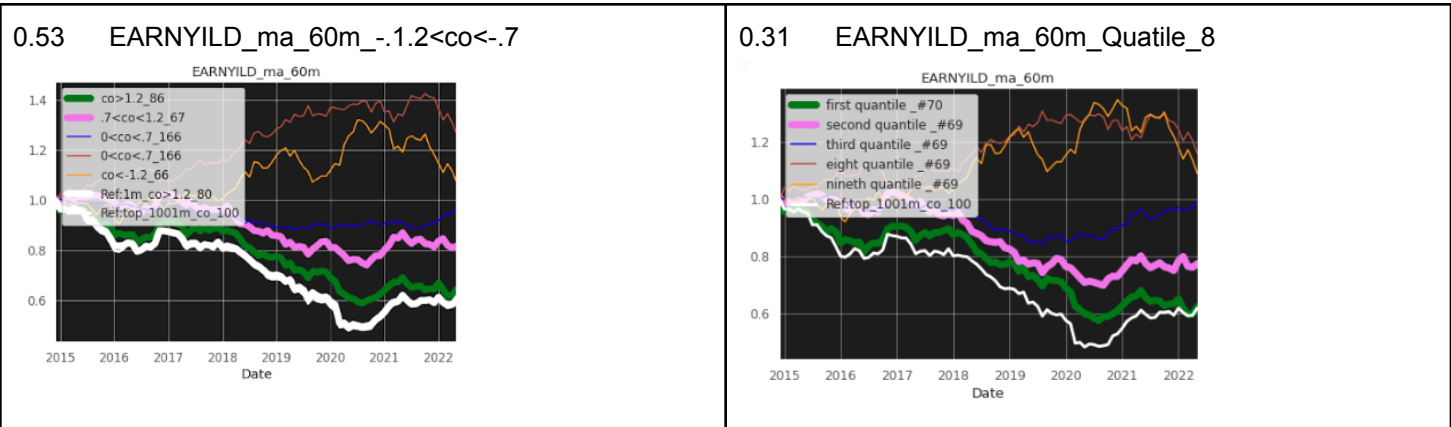


EARNYILD_ma_24m_zs

0.64 EARNYILD_ma_24m_zs_-.1.2<co<-.7|EARNYILD_ma_24m_zs

0.37 EARNYILD_ma_24m_zs_Quatile_8|EARNYILD_ma_24m_zs

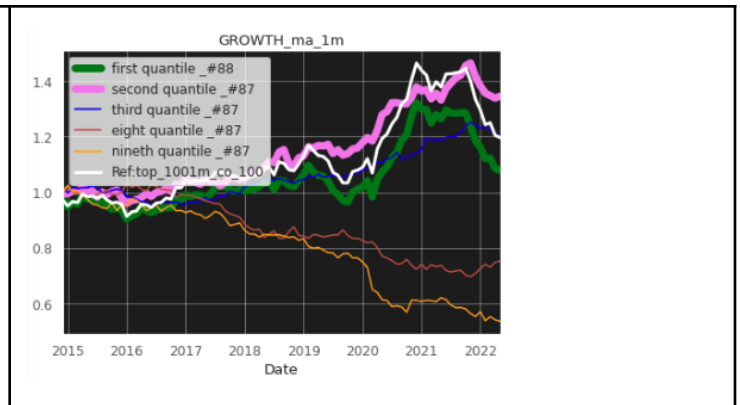
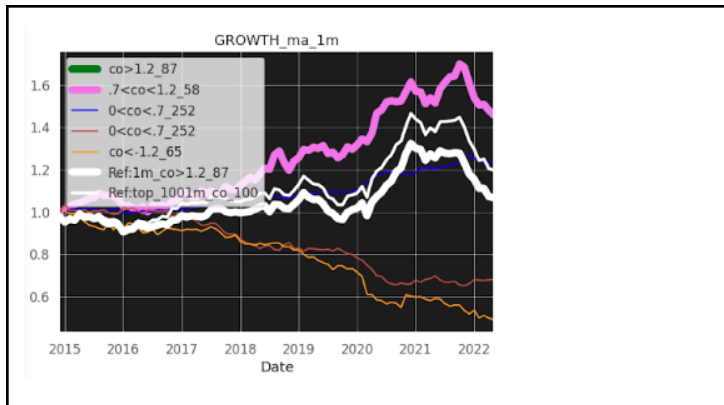
EARNYILD_ma_60m



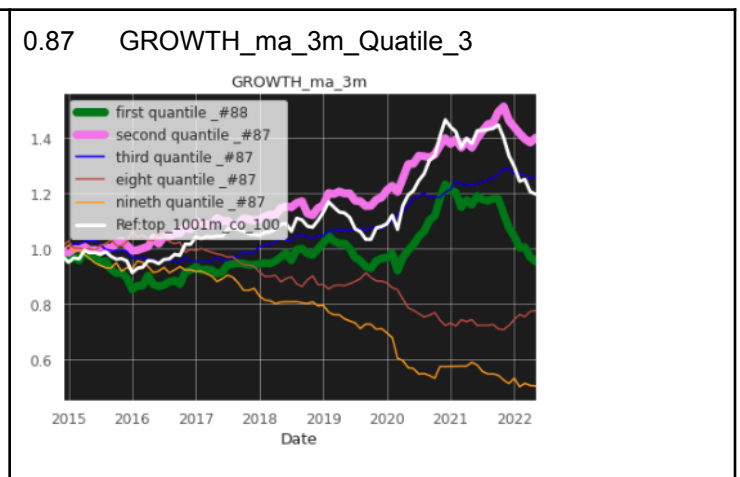
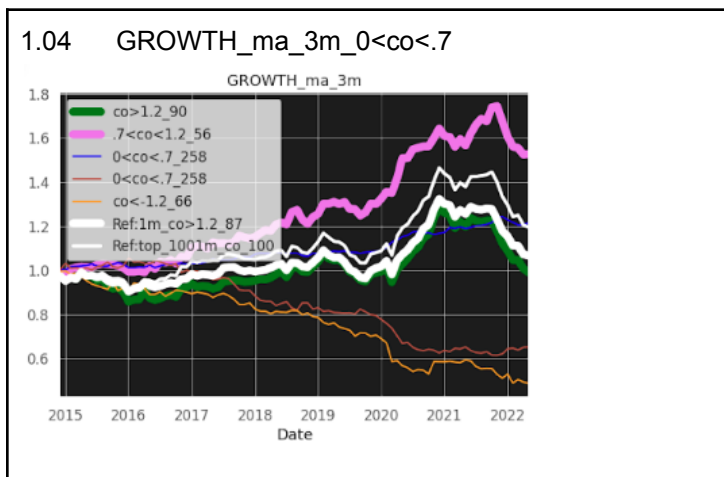
Growth

GROWTH_ma_1m

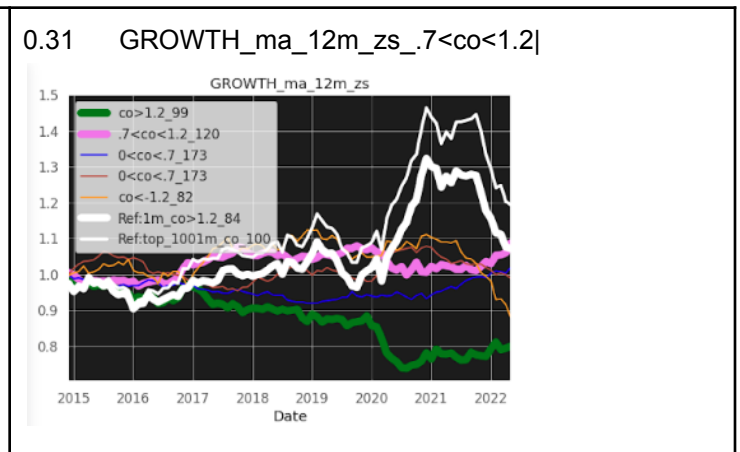
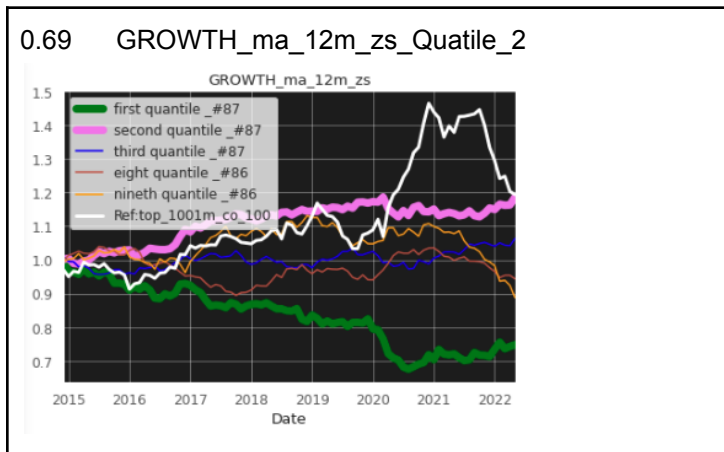
1.12 GROWTH_ma_1m_0<co<.7	0.69 GROWTH_ma_1m_Quatile_2
------------------------------	--------------------------------



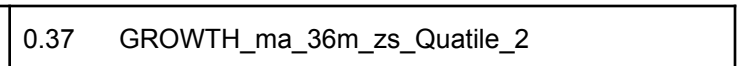
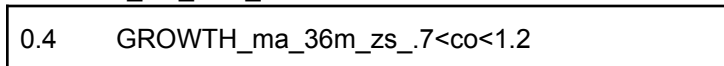
GROWTH_ma_3m

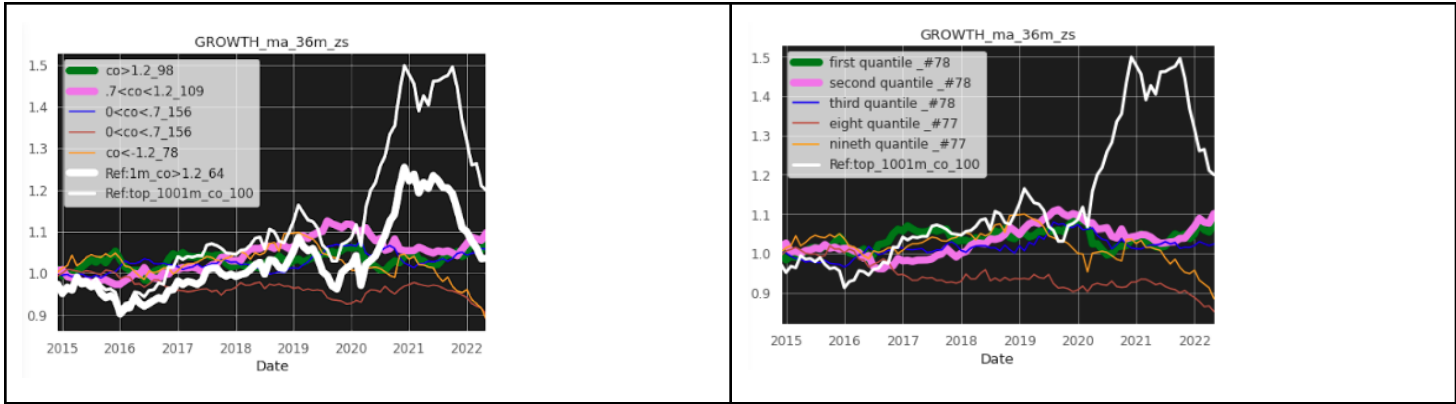


GROWTH_ma_12m_zs

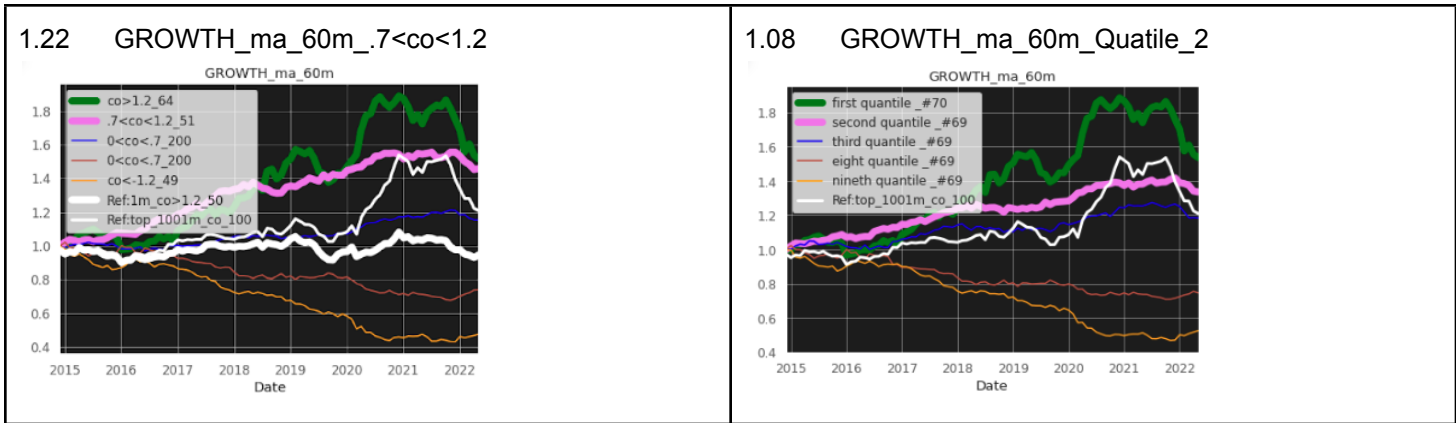


GROWTH_ma_36m_zs



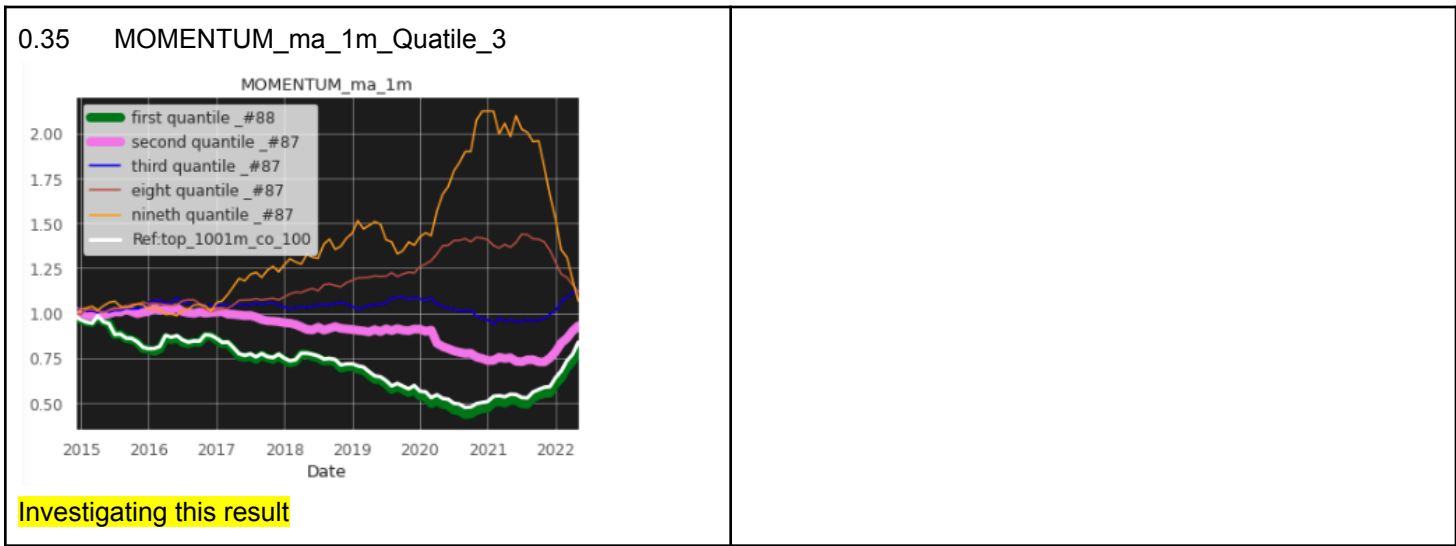


GROWTH_ma_60m



Momentum

MOMENTUM_ma_1m



- Most of the Constituents are energy stocks - Cumulative performance of each constituent is shown here -	
--	--

MOMENTUM_ma_3m

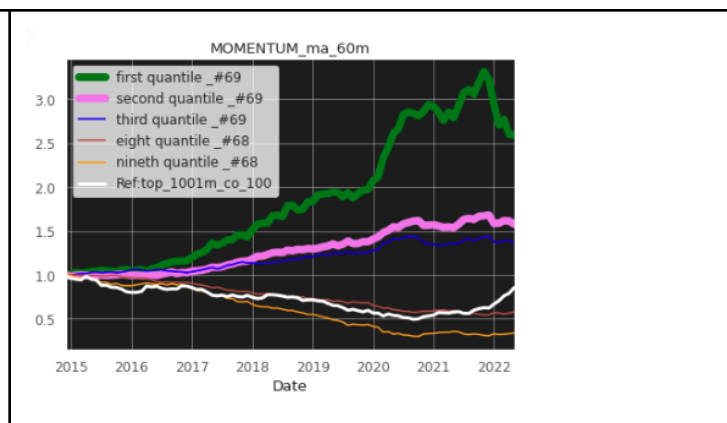
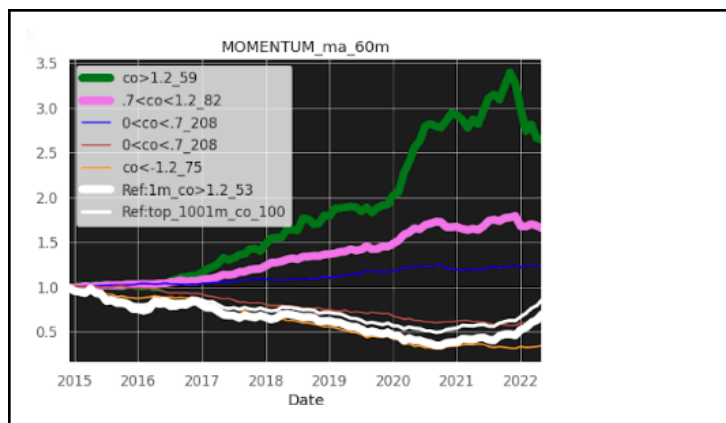
0.32 MOMENTUM_ma_3m_0<co<.7 Investigating this result - Most of the Constituents are energy stocks -	MOMENTUM_ma_3m - Cumulative performance of each constituent is shown here
---	--

MOMENTUM_ma_24m_zs

0.5 MOMENTUM_ma_24m_zs_Quatile_2	0.29 MOMENTUM_ma_24m_zs_.7<co<1.2
--	---

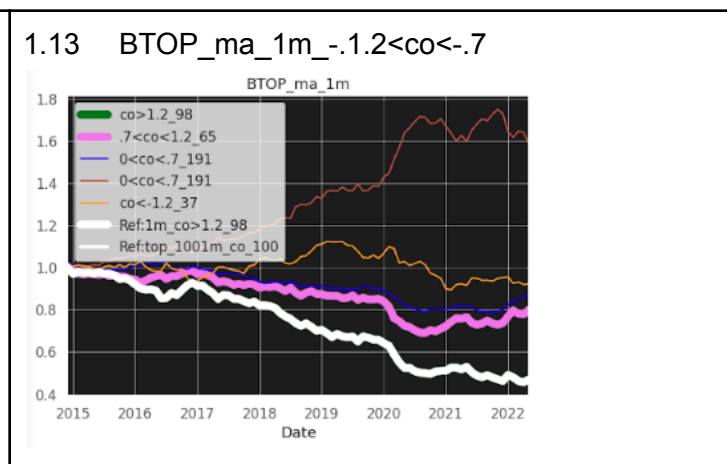
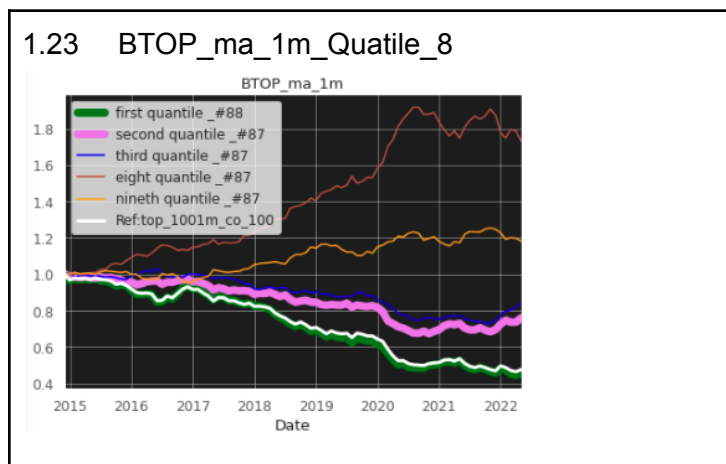
MOMENTUM_ma_60m

1.35 MOMENTUM_ma_60m_.7<co<1.2	1.26 MOMENTUM_ma_60m_Quatile_2
-----------------------------------	-----------------------------------

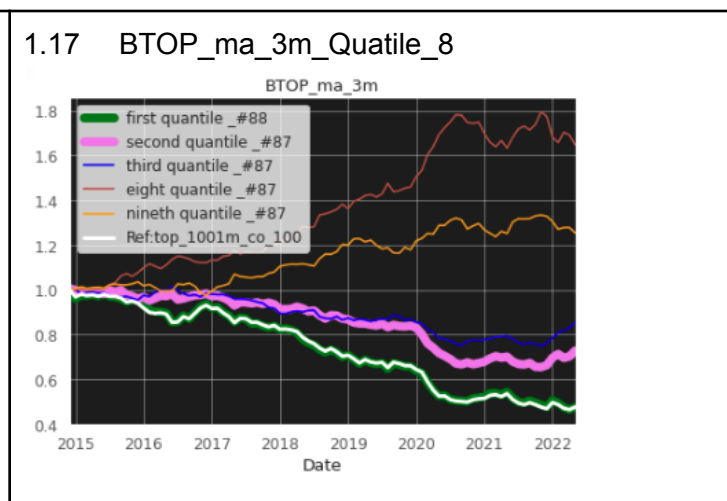
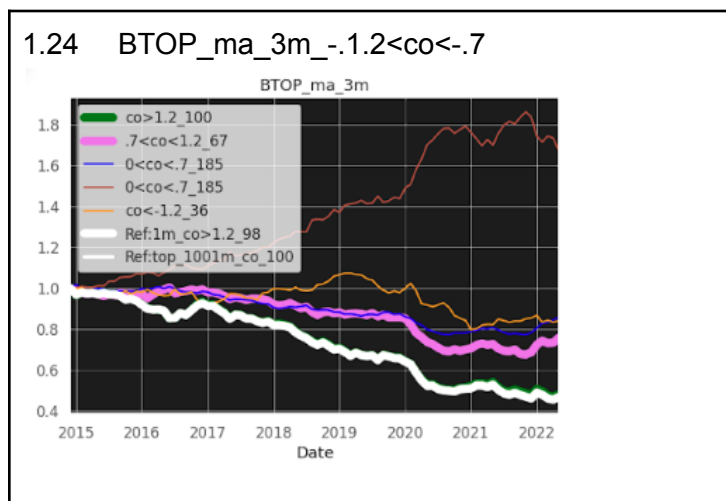


Book To Price

BTOP_ma_1m



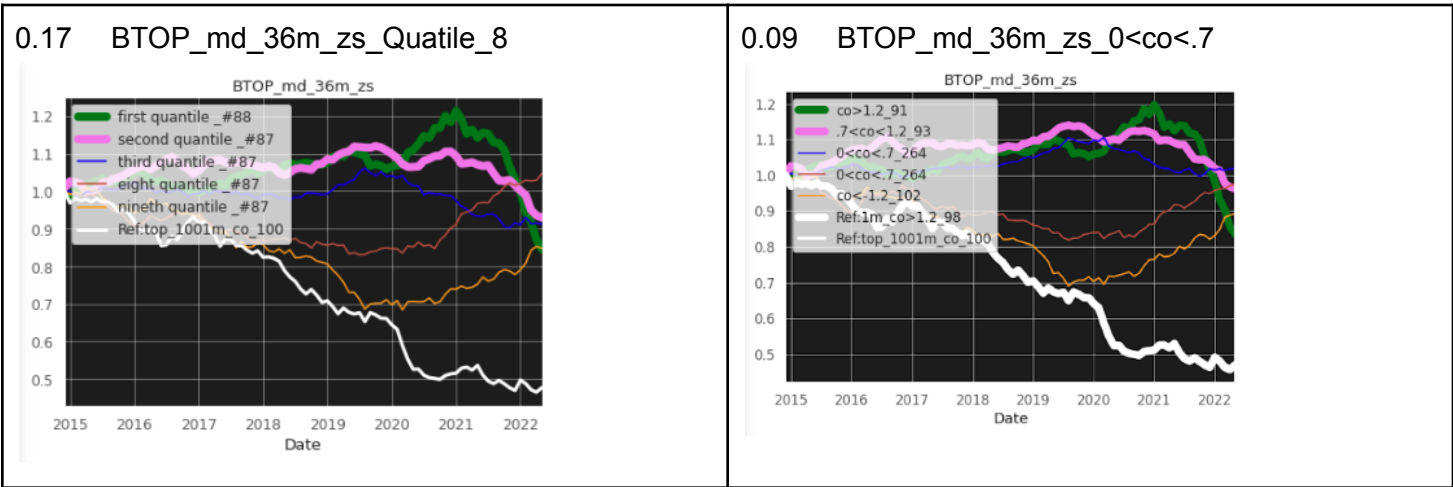
BTOP_ma_3m



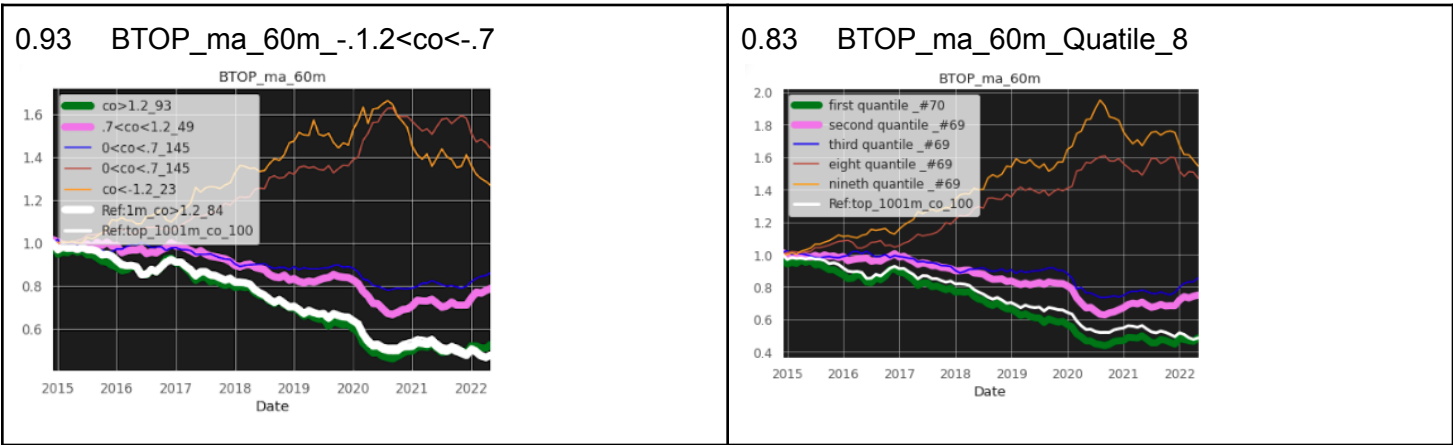
BTOP_md_24m_zs

0.42	BTOP_md_24m_zs_0<co<.7	-0.86	BTOP_md_24m_zs_.7<co<1.2
------	------------------------	-------	--------------------------

BTOP_md_36m_zs



BTOP_ma_60m



Tag columns includes
feature, transformation, cutoff for cross sectional score of transformed exposure

Chart column in the table below provides a link to the trend chart of transformed factors and their cutoffs. Transformation could be moving averages or moving differences.

Table column gives a link to holdings in the portfolio created with cutoffs

[Folder for charts](#)

[Folder with Constituents](#)

Code: [Link](#) Q3:QUANTILES_CUTOFFS_OCT12__BASED_ON_MACRO_MODEL

Task:

Other transformations

Log transform

12MA/5YR MA[z-scores level for cross section]

[Exposures trend for all combinations of transformations and cutoffs](#)

[Portfolios for all combinations of transformations and cutoffs](#)

[Information ratios of all combinations- Tab Oct 12th](#)

Tag description:

Factor name_ transformation _ date of report

Transformations:

- Ma = Moving average of exposures time series
- Md -Moving difference of exposure time series
- Mazs = Z-score of the Ma series
- Mdzs = Z-score of the Md series

Green fill = Portfolio with score > cutoff compared with equal weighted benchmark

NEXT WEEK

- Plot 1M exposures without transformation in all charts as reference
- Number of stocks next to each transformation

- Z-score for each stocks in the portfolio report

NEXT

- Best 5 transformations for each feature.
- Combine them equal weights

CODE

Google Colab Notebook: [Q3:QUANTILES_CUTOFFS_OCT12_BASED_ON_MACRO_MODEL](#)

1

Data location

Add the monthly facts csv file to the designated folder to update the data

2

Data structure

Used Barra Factors

PROFIT

MOMENTUM

GROWTH

BTOP

DATA_PRICE

Other Barra Factors

UNEXPINFLATION_MNTHLY

TRMSTRUCSLOPE_MNTHLY

INDPROD_MNTHLY

HOUSSTRTMOM_MNTHLY

EXCHVALUSD_MNTHLY

CREDRSKPREM_MNTHLY

CMOIL_MNTHLY

SIZE

RESVOL

MIDCAP

LIQUIDTY

LEVERAGE

EARNYILD

EARNVAR

EARNQLTY

DIVYILD

BETA

DATA_EXBETASIZE

LIQUIDITY_RETURN

3

Data preparation

gvKey adjustments

Long format to wide format

Ensure that the key ends with 01

Pivot by Data and gvkey to perform cross sectional normalization

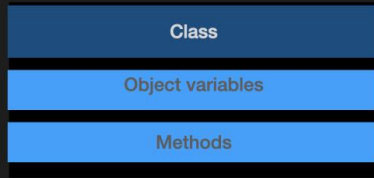
df_exp_pvt

Date	gvkey	TP_BARRA_GEMLT_EARNYILD	TP_BARRA_GEMLT_BTOP
2004-06-30	101301	-1.195	-0.218
	107501	0.526	1.310
	107801	-0.012	-0.426

Code Location: [Genus github](#)

Raw data Location: [Pavan Google Drive location](#)

4 Create data structure



Creating an object instance from this class helps in Reusing code, debugging and performing scenario analysis with changing parameter settings

5 Loop

Foreach feature, perform these steps

```
myObj.df_exp_normalize(list_cols)
```

For each time period (each month), perform a normalization of data

```
df_to_loop = myObj.df_exp_pvt_zscores.pivot_table(index=['gvkey', 'Date'])
```

5.2 Prepare data to index by key and date

```
ipdb> df_to_loop
  gvkey  Date  TP_BARRA_GEMLT_BTOP  TP_BARRA_GEMLT_EARNYILD \
101301 2004-06-30      -0.536756      -1.470895
      2004-07-31      -1.413732
      2004-08-31      -1.247989
      2004-09-30      -0.780565
      2004-10-31      -0.923826
...
33996501 2022-02-28      -0.886130      -1.294497
      2022-03-31      -1.353572
      2022-04-30      -1.329981
      2022-05-31      -1.365364
      2022-06-30      -1.348147
```

5.3

For each gvkey series, perform transformations. Transformations are Moving averages or Moving differences over different time horizons. 3 month moving average and score of that moving average are shown in this example

```
for x in gvkey_index:
    in_df = df_to_loop.loc[x]
    in_df['gvkey'] = str(x)
    in_df['Date'] = in_df.index
    # in_df['roll_mean'] = in_df[[feat]].rolling(n_window).mean()
    in_df['ma_3m'] = in_df[[feat]].rolling(3).mean()
    in_df['ma_3m_zs'] = normalize_data(in_df[[feat]].rolling(3).
    in_df['ma_6m'] = in_df[[feat]].rolling(6).mean()
    in_df['ma_6m_zs'] = normalize_data(in_df[[feat]].rolling(6).
```

Saved cross sectional data: [Google drive location](#)

6 For each ticker, save the transformation data for last available date

```
for x in gvkey_index:
    in_df = df_to_loop.loc[x]
    frames.append(in_df.iloc[-1])
```

Tickers might not have the same end date. To Address this issue, keep track of the last date data for each ticker. Save This cross sectional data (key = ticker, columns = all transformations, date = last available date) it to folder

7 For each transformation, for chosen backtest date, create portfolio based on cutoffs

```
transformations = ['ma_3m_zs', 'ma_6m_zs', 'ma_12m_zs', 'ma_24m_zs', 'ma_36m_zs', 'ma_60m_zs']
for transformation in transformations:

    df_cross_section_t = df_cross_section[df_cross_section['Date'] == "2022-06-30"]
    df_cross_section_t['cross_sec_norm'] = normalize_data(df_cross_section_t[transformation])
```

Normalize cross sectional data.
Create portfolio based on the chosen cutoff.
Get portfolio returns. Compare these returns with all portfolio benchmark

8 Calculate metrics. Save results to spreadsheet

```
port_z_ret = get_portfolio_returns_to_plot(gvlist_cutoff_l, myObj.price_df)
print(get_metrics_for_quantile(port_z_ret, port_all_ret ))
```

Information ratio is the key performance metric
To compare portfolio returns vs the benchmark

[Spreadsheet Reports location:](#)

Methods