Yield Curve PCA in Appropriate Hedging & Relative Value Strategy

Group 3:

Duan Zhao dz2539
Rain Wei xw2762
Aiting Chen ac5429
Yue Wuhaoxiu wy2438
Jung Henry hj2627
Liu, Andrew apl2144
Nunez, Omar omn2108

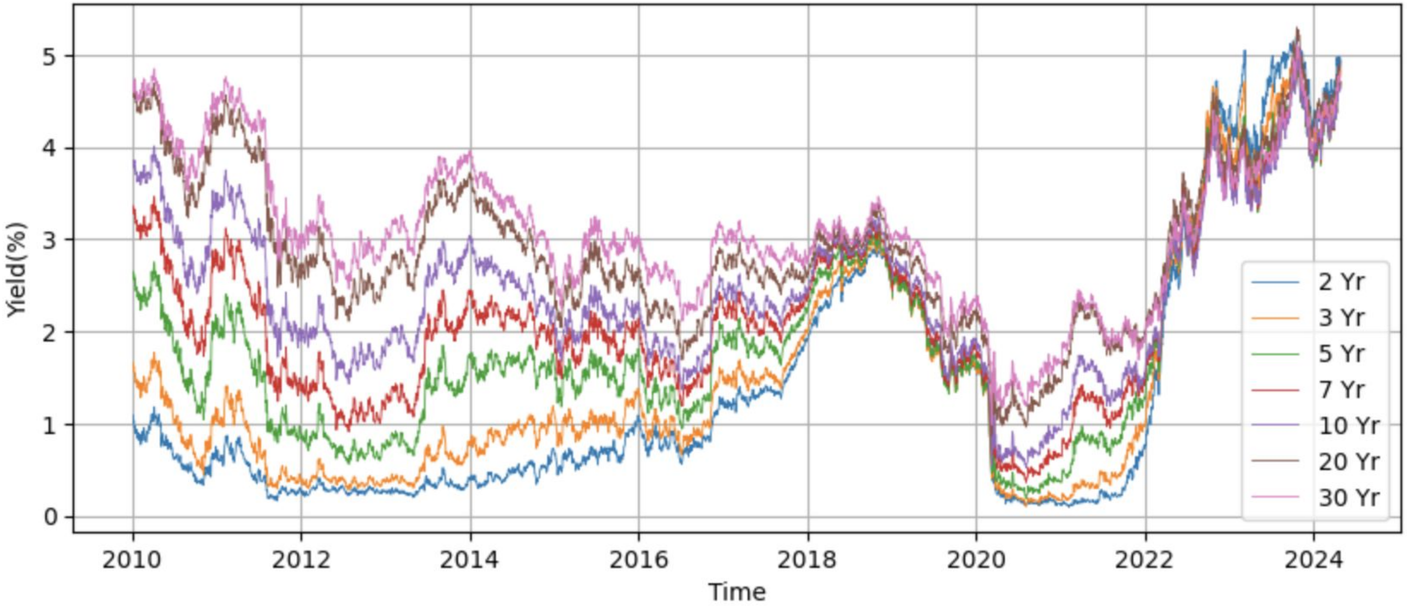
1.Background

- Assumption: the market is driven by a collection of linearly uncorrelated factors.
- One can construct portfolios to hedge against or be exposed to any factor
- PCA can help surface the driving forces of market mechanisms

- Main objectives of our study:
    - Identify and analyze the market mechanisms of yield curve variations
    - Use the decomposed yield curves to realize **appropriate hedging**
    - Use the decomposed yield curves to design **relative value strategies**

2.Data

- 2010-2024 US Treasury Yields (Daily) [tenor > 2 yr]
- Split into train (2010 - 2019) and test (2020 - 2024.4)



PART I – Identify and analyze the market mechanisms of yield curve variations

3. PCA Methodology & Training

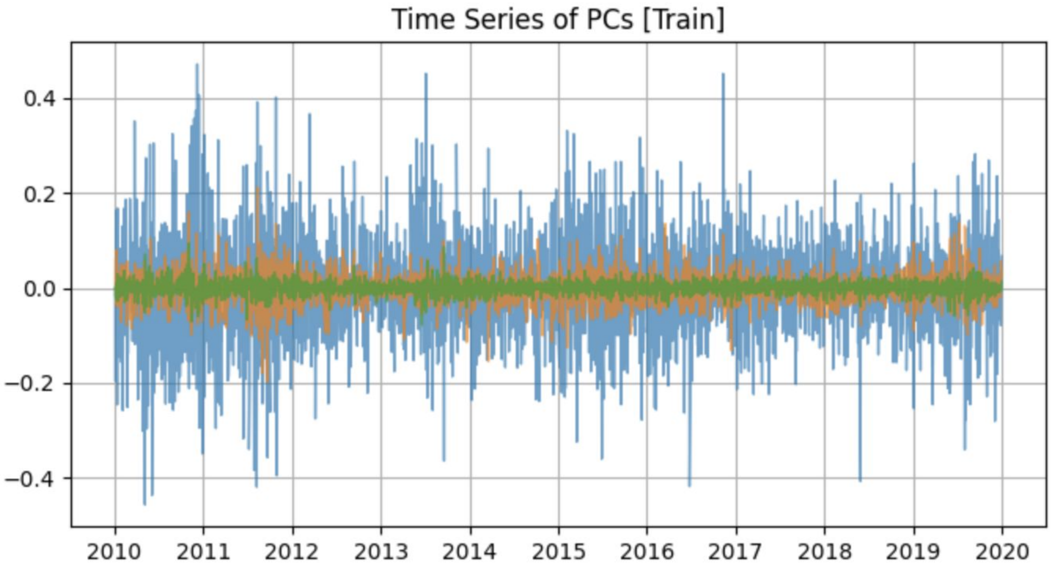
- Assume that the yield curve at time t, $y_i^t \ (i = 1, \ldots, n)$ follows a k-factor linear model

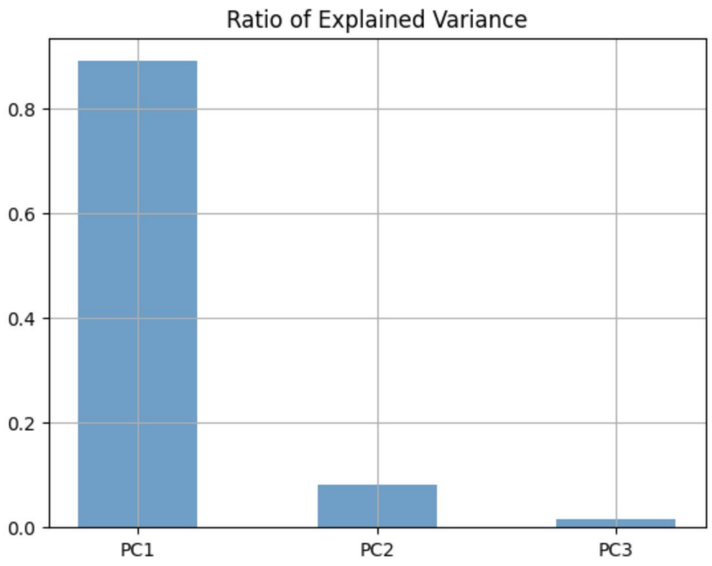
$$\begin{pmatrix} y_1^t \\ \vdots \\ y_n^t \end{pmatrix} = \sum_{i=1}^{k} \alpha_i^t \cdot \begin{pmatrix} f_{i1} \\ \vdots \\ f_{in} \end{pmatrix} + \begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix}$$

- Use PCA to find the uncorrelated linear factors
  - The yield curve is not stationary
  - Use PCA on the yield change

$$\Delta Y^t = \sum_{j=1}^{k} \begin{pmatrix} f_{j1}^t \\ \vdots \\ f_{jn}^t \end{pmatrix} \Delta\alpha_j^t + \begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix}$$

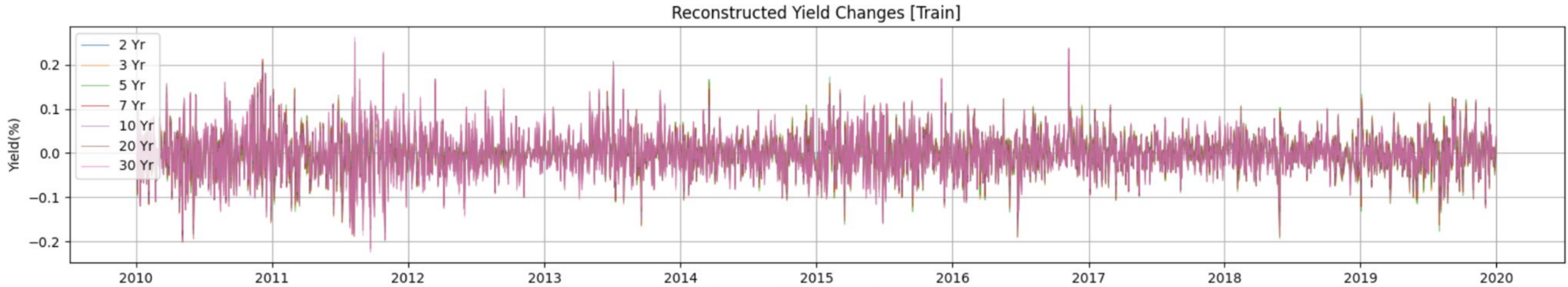
PCA on Yield Curve

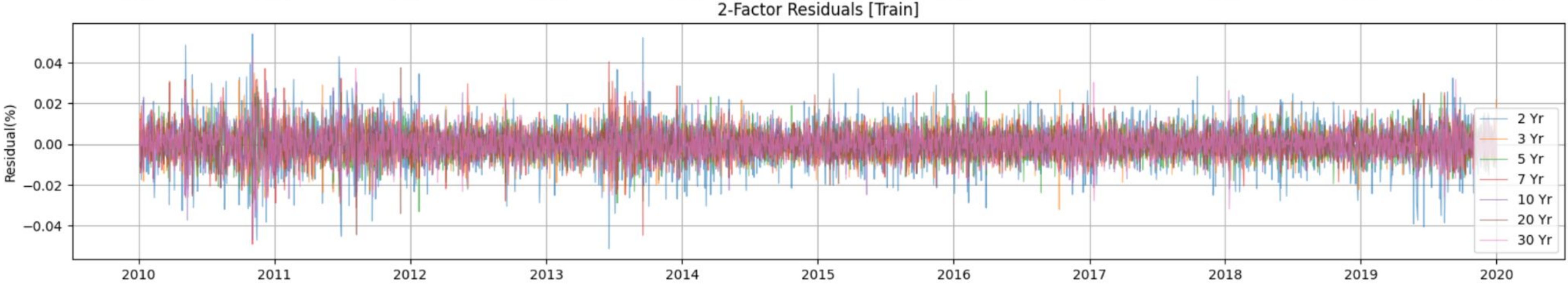




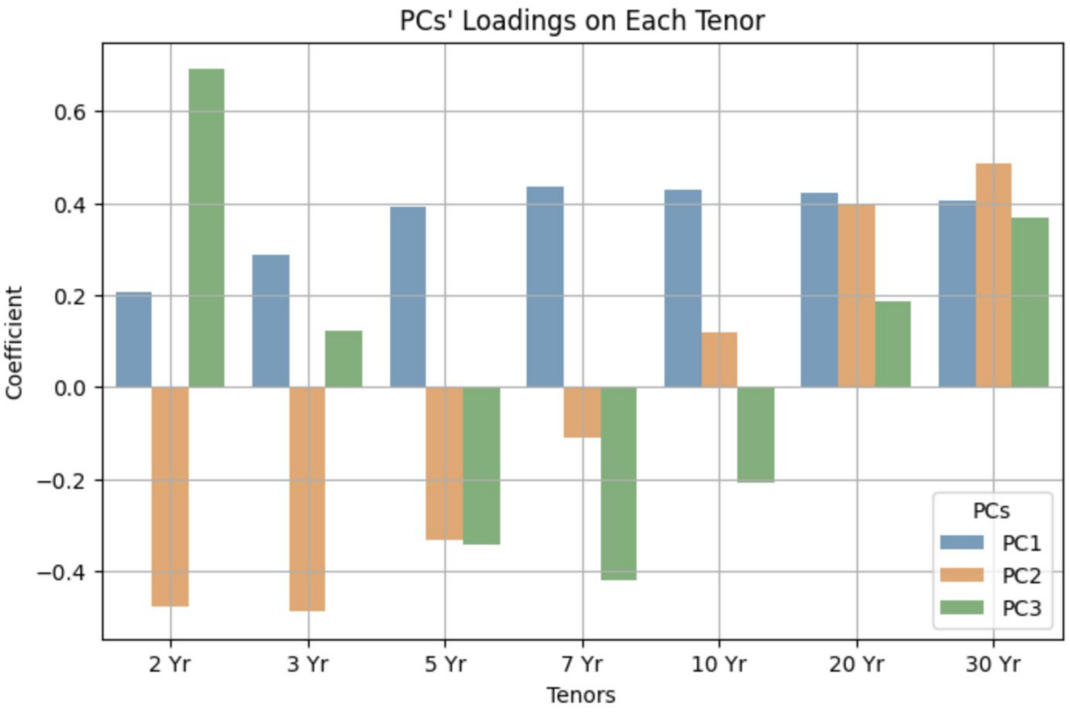
Reconstruction result

```
The R-square of 2 Yr yield returns = 86.55%
The R-square of 3 Yr yield returns = 96.19%
The R-square of 5 Yr yield returns = 97.56%
The R-square of 7 Yr yield returns = 97.48%
The R-square of 10 Yr yield returns = 98.28%
The R-square of 20 Yr yield returns = 98.83%
The R-square of 30 Yr yield returns = 97.92%
```





Interpret the PCs



- PC1- directional dynamics
- PC2 - slope of curve
- PC3 - curvature dynamics

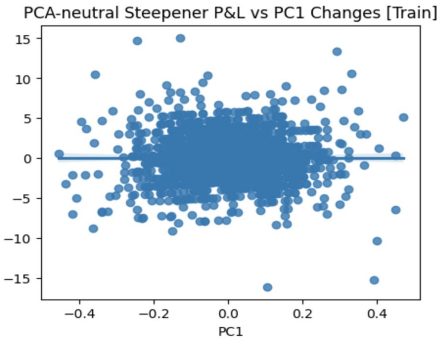
PCs’s corr with external factors (economically)

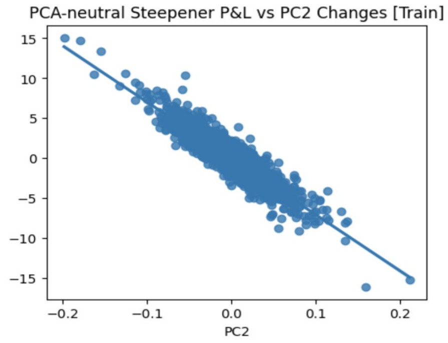
```
PC1
Corr with VIX: -0.3650026119899479
Corr with WTI Crude Oil: 0.26099225922444125
Corr with EUR/USD FX Spot: 0.05980298839797024
--------------------------------------------------
PC2
Corr with VIX: -0.07217738785054835
Corr with WTI Crude Oil: 0.11195188834890238
Corr with EUR/USD FX Spot: 0.010613005197714728
--------------------------------------------------
PC3
Corr with VIX: -0.0586021885276794
Corr with WTI Crude Oil: 0.04646615222927119
Corr with EUR/USD FX Spot: 0.04776221508379735
--------------------------------------------------
```

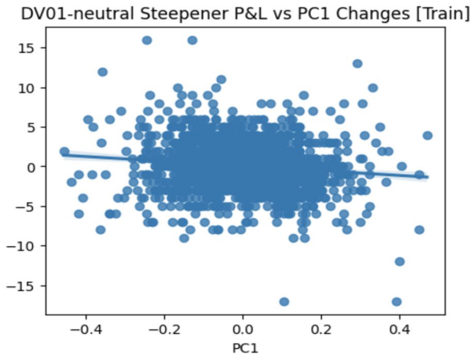
PART II - Use the decomposed yield curves to devise hedging and relative value strategies

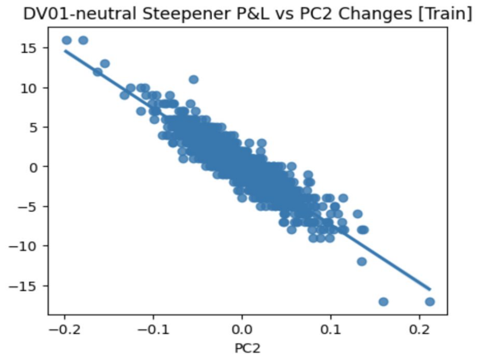
4.1 Appropriate Hedging - Building Curve-neutral Steepener

- By hedging the PC1, one can create a steepening position.
    - The return of the portfolio is only affected by the non-directional steepness and higher-order factors
- 5Y-20Y Steepener
    - The PCA-neutral steepener has almost zero exposure to PC1 changes, while the DV01-neutral Steepener's hedge is not clear.



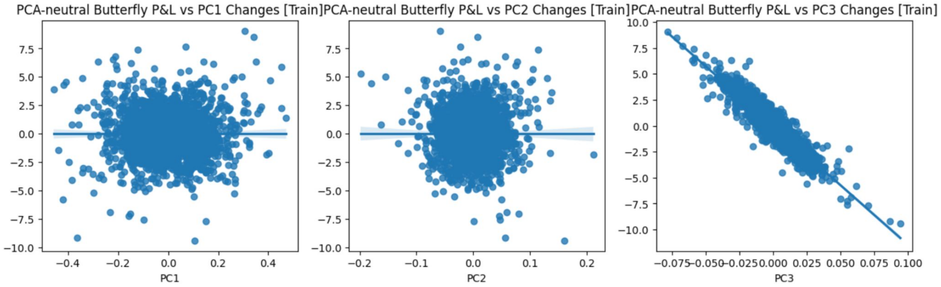


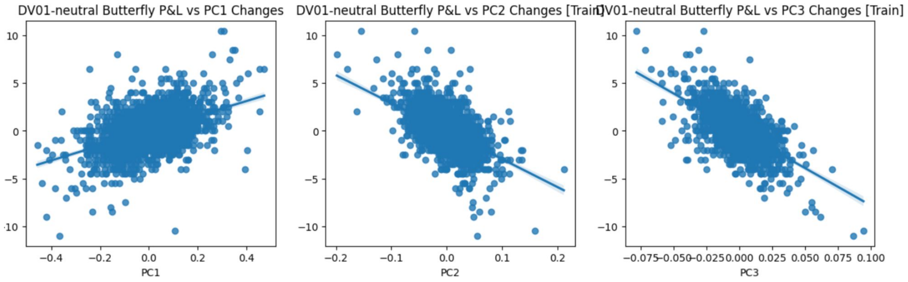




4.2 Appropriate Hedging - Building Curve-neutral Butterfly

- By hedging the PC1 and PC2, one can create a butterfly position.
  - The return of the portfolio is only affected by the net curvature and higher-order factors
- 2Y-5Y-20Y Butterfly
  - The PCA-neutral butterfly has almost zero exposure to PC1 and PC2 changes, while the DV01-neutral Butterfly's hedge is not clear.





5 Strategy

Goal:

1 Immune to PC1 and PC2 changes (based on 2-factor residual)

2 Profitable (acceptable Sharpe ratio)

Methodology:

Use residual operator to translate 2-factor residual autocorrelation into portfolio returns.

5.1 Lemma

f and residual(f) are vectors of the same dimension

Fisrt we define residual operator $residual(\cdot)$ as: $\boxed{\forall f, residual(f) = f - \Sigma_{i=1}^{k}(f \cdot PC_i)PC_i.}$ k=2

Then we can show that, for two d-dimensional multivariate time series $\{r^{(t)}\}$ and $\{f^{(t)}\}$,
$\Sigma_{j=1}^{d}(r_j^{(t+1)} \cdot residual(f_j^{(t)})) = \Sigma_{j=1}^{d}(residual(r_j^{(t+1)}) \cdot f_j^{(t)})$.

The proof is tedious calculations!

Specifically, let's say we have d key tenors on the yield curve (d=7 in our case).

Let:

$f^{(t)}$ be any d-by-1 feature vector on day t,

r_t be the d-by-1 yield change vector on day t,

Moreover, on day t+1, we set our position (in DV01s) on each tenor as the d-by-1 vector $residual(f^{(t)})$,

5.1 Lemma(Ctd.)

Daily change of the j-th tenor yield on day t+1

we can then interpret:

DV01 position for the j-th tenor on day t+1

$LHS = \Sigma_{j=1}^{d} r_j^{(t+1)} \cdot residual(f_j^{(t)})$ as the portfolio's P&L on day t+1. Remember the summation is over each tenor j.

$RHS = \Sigma_{j=1}^{d} residual(r_j^{(t+1)}) \cdot f_j^{(t)}$ as the P&L in the residual space.

This is what we choose to use

Difference between the **realized** and **reconstructed** daily change of the j-th tenor yield on day t+1

Therefore, to make LHS large (i.e. good P&L), what we should do is:

Pick a good feature $f^{(t)}$ **that can make RHS large!!!**

What properties should such $f^{(t)}$ have?
Good correlation with the next day return residual.

At least $sign(f_j^{(t)}) = sign(residual(r_j^{(t+1)}))$!

5.1 Lemma(Ctd.)

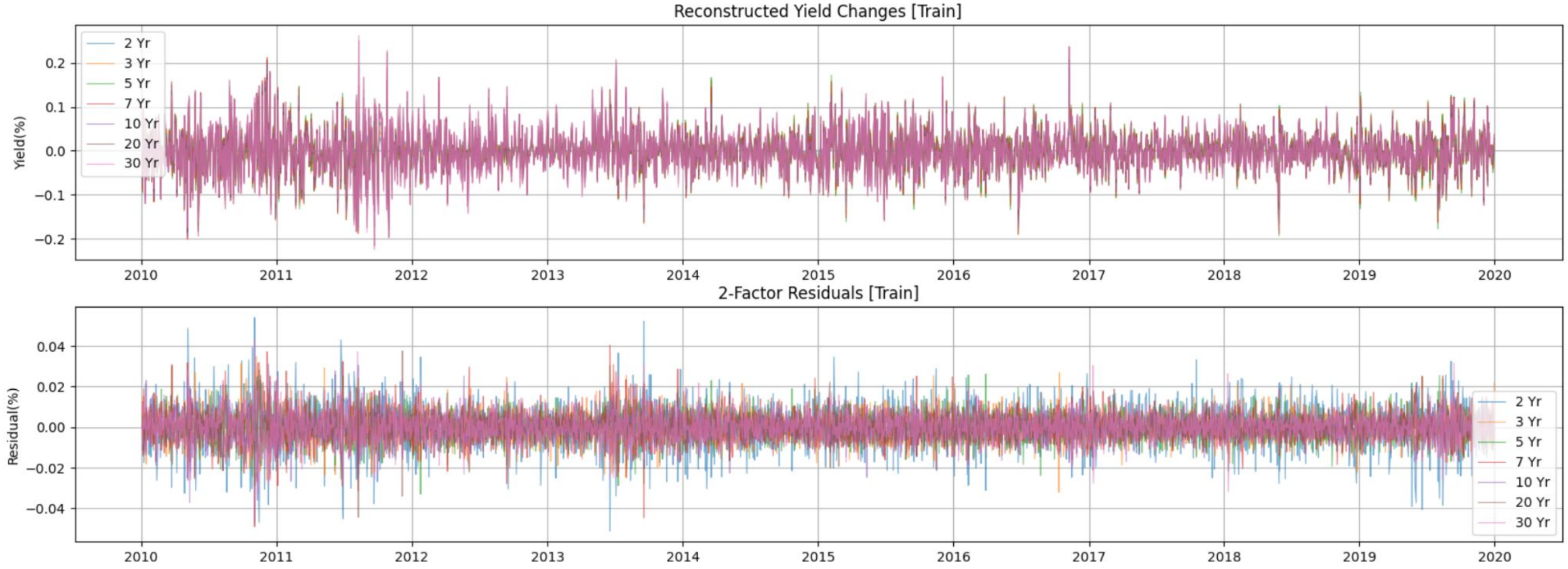
Now, we know that, **to have good portfolio P&L**, we need to **find some feature f** that has good prediction power of the **next day's residual** of yield changes!

This transformation **from portfolio space to residual space** enables us to work **implicitly** on portfolio performance!

5.2 Search for good feature f

2-factor residuals is not correlated with many external factors.

A naive idea: use historical residuals to predict next day's residual!



Clear lag-1 autocorrelation!

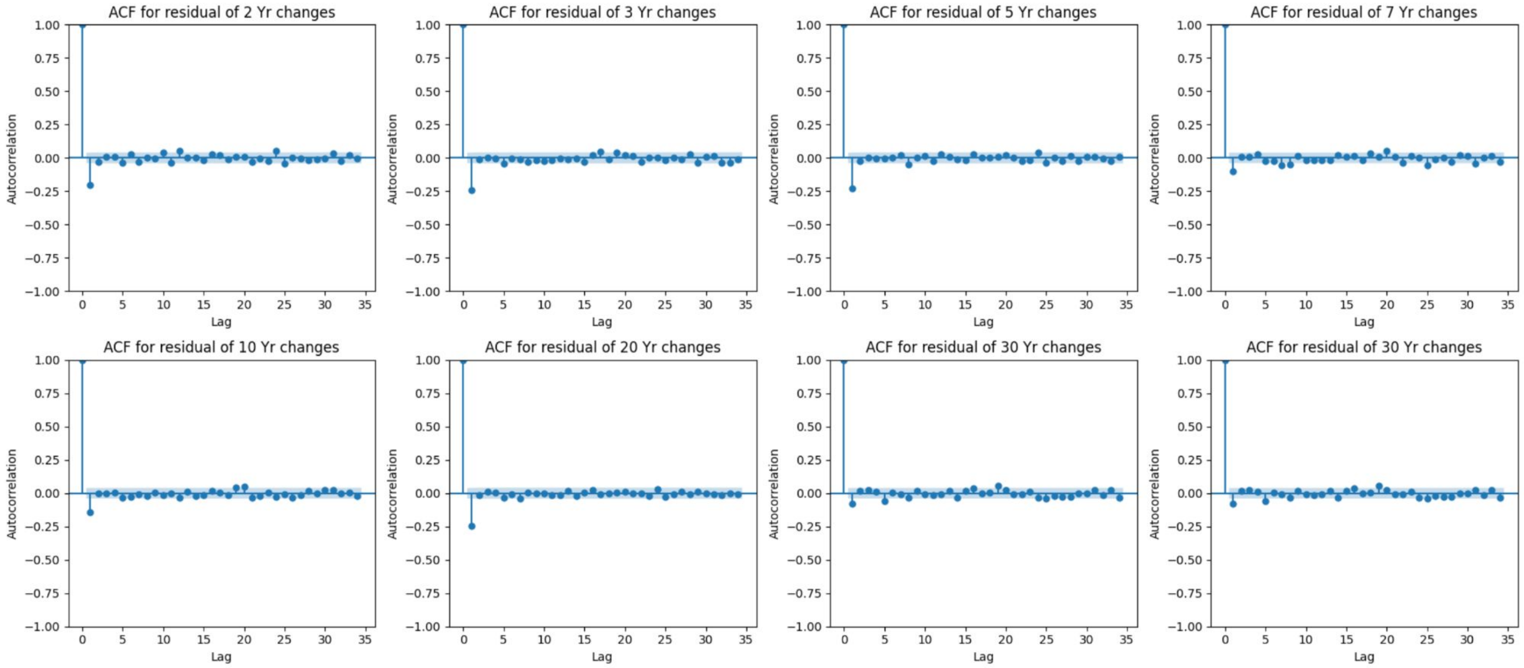
5.2 Search for good feature f (Ctd.)

Residual:

Short-term: mean reversion

Long-term: no significant momentum

Examine the residuals' ACF:



5.2 Search for good feature f (Ctd.)

$residual(r_{t+1}) \sim residual(r_t)$ can be great.

The only problem: feature (signal) $residual(r_t)$ is too 'quick'!

Can we use some technique to slow it down? **EWMA!**

$$residual(r_{t+1}) \sim EWMA(residual(r_t), halflife = h)$$

Next day's residual

Our feature_t:
EWMA of residual with different half-lifes!

5.2 Search for good feature f (Ctd.)

Examine t-statistics:

Recall that our goal is to 'maximize' $RHS = \Sigma_{j=1}^{d} residual(r_j^{(t+1)}) \cdot f_j^{(t)}$

Here for each j, we calculate $residual(r_j^{(t+1)}) \cdot f_j^{(t)}$ and do t-test.

```
Half-life = 0:
2 Yr:
corr=0.2059
t-stats = 8.32, p-value = 0.0
------------------------------
3 Yr:
corr=0.241
t-stats = 10.12, p-value = 0.0
------------------------------
5 Yr:
corr=0.2283
t-stats = 10.39, p-value = 0.0
------------------------------
7 Yr:
corr=0.0994
t-stats = 3.5, p-value = 0.0
------------------------------
10 Yr:
corr=0.1428
t-stats = 4.76, p-value = 0.0
------------------------------
20 Yr:
corr=0.2453
t-stats = 7.8, p-value = 0.0
------------------------------
30 Yr:
corr=0.0778
t-stats = 2.43, p-value = 0.02
------------------------------
```

```
Half-life = 10:
2 Yr:
corr=0.1058
t-stats = 4.18, p-value = 0.0
------------------------------
3 Yr:
corr=0.2036
t-stats = 8.24, p-value = 0.0
------------------------------
5 Yr:
corr=0.1268
t-stats = 5.34, p-value = 0.0
------------------------------
7 Yr:
corr=0.0793
t-stats = 2.48, p-value = 0.01
------------------------------
10 Yr:
corr=0.1155
t-stats = 3.84, p-value = 0.0
------------------------------
20 Yr:
corr=0.1613
t-stats = 6.16, p-value = 0.0
------------------------------
30 Yr:
corr=0.0423
t-stats = 1.34, p-value = 0.18
------------------------------
```

```
Half-life = 30:
2 Yr:
corr=0.0544
t-stats = 2.22, p-value = 0.03
------------------------------
3 Yr:
corr=0.171
t-stats = 7.07, p-value = 0.0
------------------------------
5 Yr:
corr=0.0684
t-stats = 3.0, p-value = 0.0
------------------------------
7 Yr:
corr=0.0683
t-stats = 2.23, p-value = 0.03
------------------------------
10 Yr:
corr=0.1153
t-stats = 4.01, p-value = 0.0
------------------------------
20 Yr:
corr=0.0942
t-stats = 4.23, p-value = 0.0
------------------------------
30 Yr:
corr=0.0427
t-stats = 1.43, p-value = 0.15
------------------------------
```

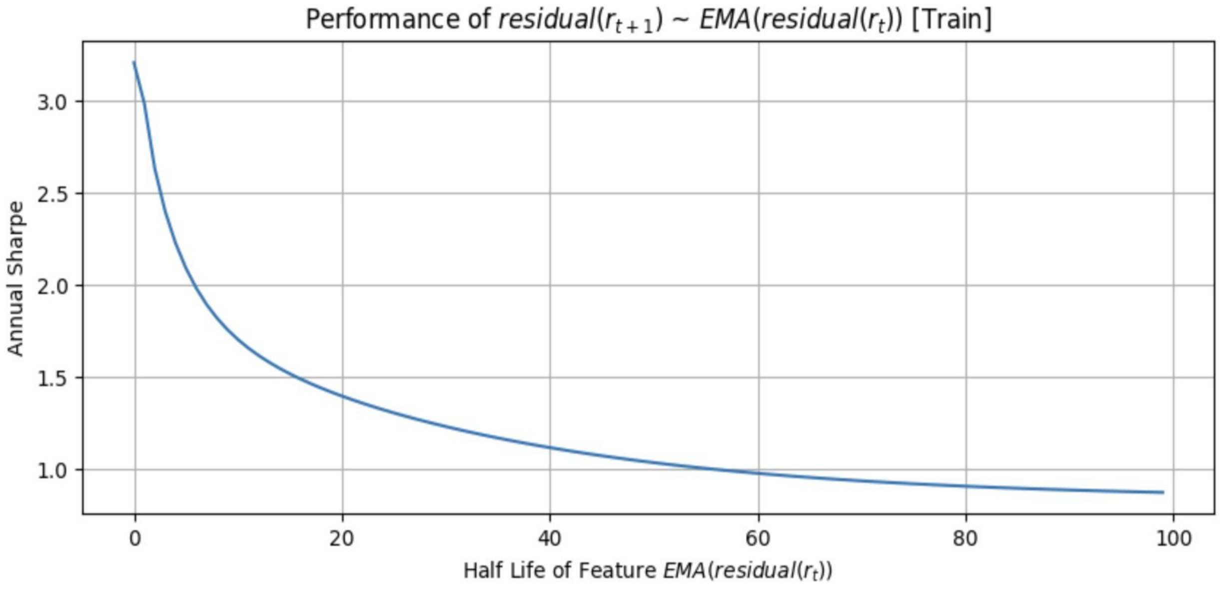
```
Half-life = 50:
2 Yr:
corr=0.0377
t-stats = 1.56, p-value = 0.12
------------------------------
3 Yr:
corr=0.1517
t-stats = 6.46, p-value = 0.0
------------------------------
5 Yr:
corr=0.0482
t-stats = 2.17, p-value = 0.03
------------------------------
7 Yr:
corr=0.0601
t-stats = 2.04, p-value = 0.04
------------------------------
10 Yr:
corr=0.1137
t-stats = 3.88, p-value = 0.0
------------------------------
20 Yr:
corr=0.0634
t-stats = 3.06, p-value = 0.0
------------------------------
30 Yr:
corr=0.0404
t-stats = 1.35, p-value = 0.18
------------------------------
```

```
Half-life = 100:
2 Yr:
corr=0.027
t-stats = 1.11, p-value = 0.27
------------------------------
3 Yr:
corr=0.1279
t-stats = 5.61, p-value = 0.0
------------------------------
5 Yr:
corr=0.0347
t-stats = 1.59, p-value = 0.11
------------------------------
7 Yr:
corr=0.0563
t-stats = 1.93, p-value = 0.05
------------------------------
10 Yr:
corr=0.109
t-stats = 3.45, p-value = 0.0
------------------------------
20 Yr:
corr=0.033
t-stats = 1.67, p-value = 0.1
------------------------------
30 Yr:
corr=0.0396
t-stats = 1.23, p-value = 0.22
------------------------------
```

5.2 Search for good feature f (Ctd.)

In Residual space:

$$residual(r_{t+1}) \sim EWMA(residual(r_t), halflife = h)$$



5.2 Search for good feature f (Ctd.)

In Portfolio space:

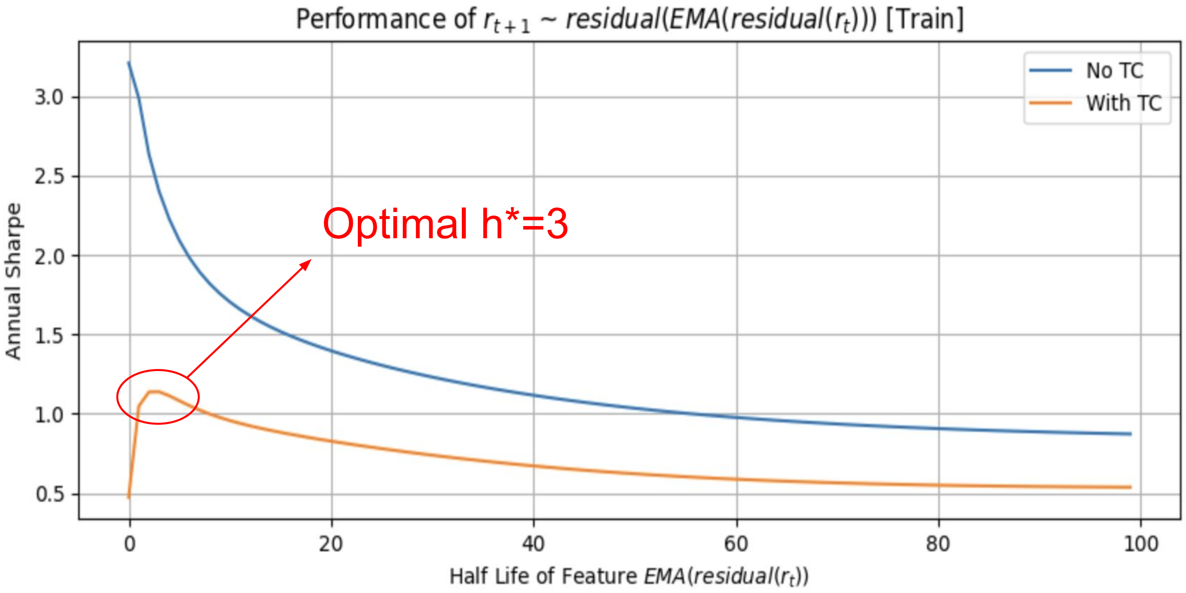
$$r_{t+1} \sim residual(EWMA(residual(r_t), halflife = h))$$

transaction cost = 0.1*ΔDV01

Assumptions

·we trade only **ATM IRS** (so that the cost of opening new position is zero) to realize the strategy

·everyday we clear the position and open new position to realize daily P&L



5.3 Test Optimal Strategy

In portfolio space, the optimal strategy is

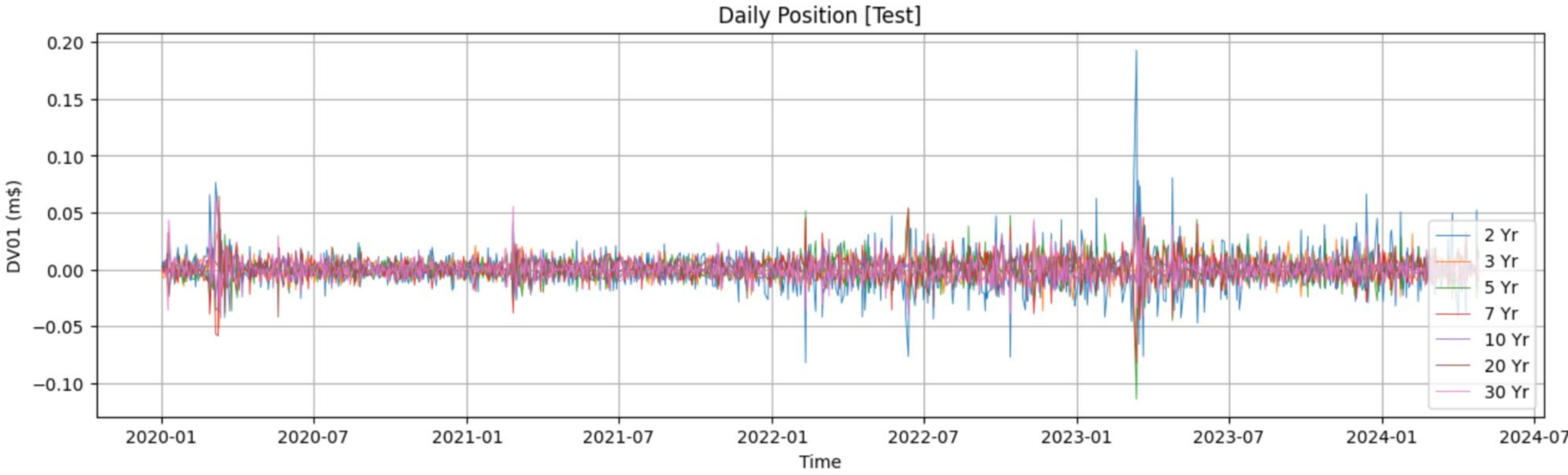
Our best feature_t

$$r_{t+1} \sim residual(EWMA(residual(r_t), halflife = 3))$$

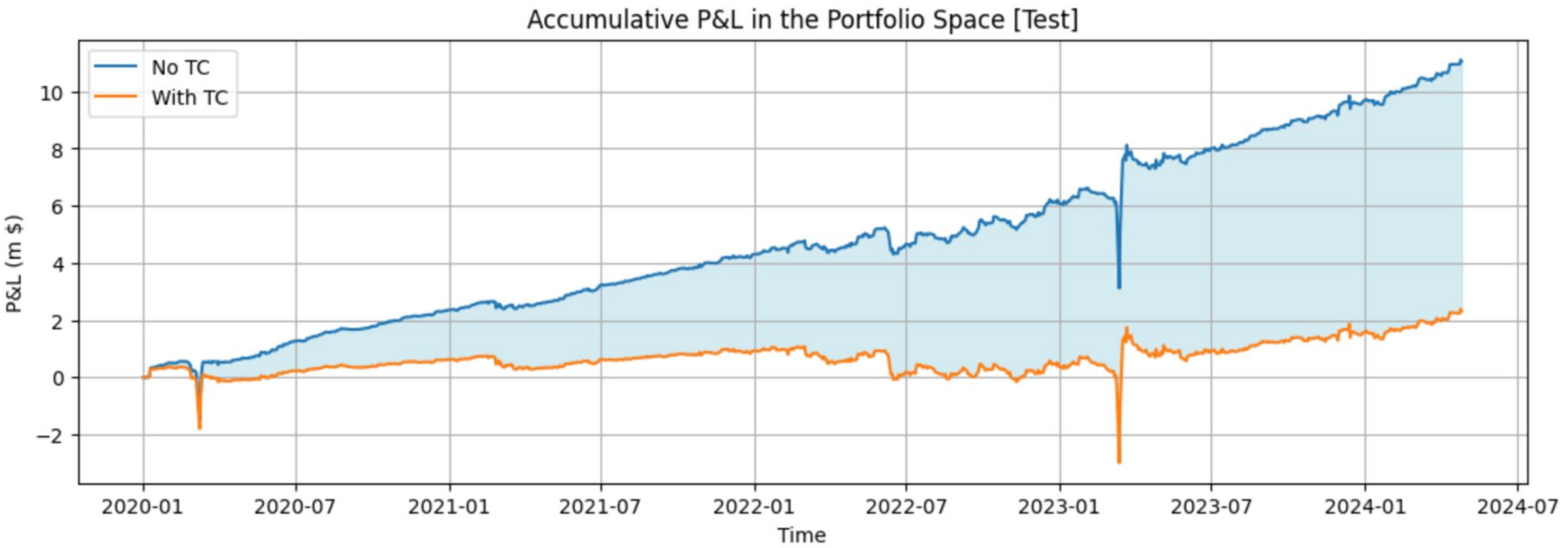
Recall day t+1 P&L= $\sum_{j=1}^{d} r_j^{(t+1)} \cdot residual(f_j^{(t)})$

Let's visualize the strategy's performance in **test period (2020-2024.4)**!

5.3 Test Optimal Strategy (Ctd.)



5.3 Test Optimal Strategy (Ctd.)




```
With transaction cost:
                         Annual Value
Probability of Success        0.634
Mean P&L                      2.577
Volatility of P&L             2.163
Sharpe Ratio                  1.191
----------------------------------------
No transaction cost:
                         Annual Value
Probability of Success        0.532
Mean P&L                      0.540
Volatility of P&L             2.147
Sharpe Ratio                  0.252
```

High winning rate!

Transaction cost consumes most of the profit, which has always been a problem in Fixed Income!

Thank you!