

Analysis Part, Q2

Zhenyu(Daniel) Bai

Initial Data Interpretation:

- S&P: Standard & Poor's 500 Index, showing the general performance of prominent American equities overall. It might show how Tesla stock aligns with the general market trend.
- DX(Dollar Index): a measure of the value of the U.S. dollar relative to a basket of foreign currencies. It theoretically would have an impact on the trade performance of Tesla Stock.
- Gold: theoretically correlated to the movement of DX
- VIX(Volatility Index): represents the market's expectations for the relative strength of near-term price changes of the S&P 500 Index (SPX). It can indicate the market sentiment and the expectation of market participants.
- IRX(13-week U.S. Treasury Yield): key indicator of short-term interest rates in the US, providing insights into the current state of the economy, investor sentiment, and expectations about future monetary policy
- TYX(30-year U.S. Treasury Yield): a critical benchmark for long-term interest rates in the US and

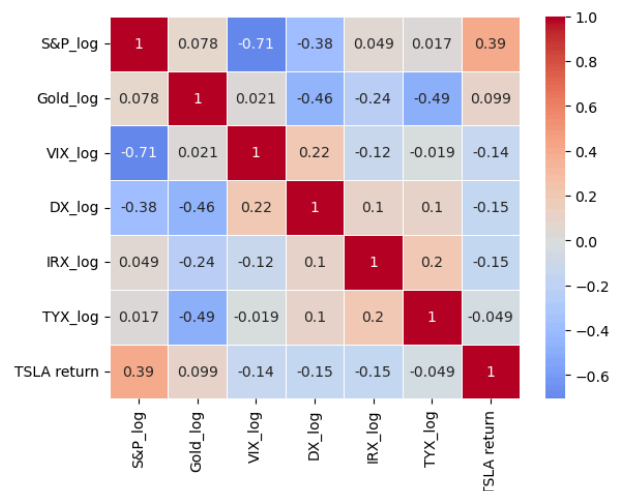
Baseline Model:

Firstly, the baseline linear regression is set up to measure the impact of each variable on the TSLA return. Since the TSLA return is the log difference of Tesla stock price, it captures the changing direction and magnitude of the stock price. In order to align with the measurement, all variables in the base model are also applied with the log difference transformation. On top of that, MinMaxScaler on the range of -1 to 1 is also applied on the data to reduce variance differences in the explanatory variables.

The log transformations applied to these variables typically help stabilize variance and make relationships more linear. The correlations shown are based on historical data and represent how these variables have moved together or in opposition during the time frame of the data.

The image shows the correlation between variables. There are several significant correlations.

1. S&P_log and VIX_log (-0.71): The strong negative correlation is expected as the VIX measures market volatility and is often inversely related to stock market



performance. When the stock market (S&P 500) is performing well and investor confidence is high, volatility tends to be low, and vice versa.

2. S&P_log and DX_log (-0.38): A moderate negative correlation suggests that a stronger dollar is sometimes associated with lower returns in the stock market, potentially due to the impact on international earnings of U.S. companies and the comparative attractiveness of U.S. assets.
3. Gold_log and DX_log (-0.46): Gold is negatively correlated with the dollar because their values are binded. As the dollar strengthens, gold becomes more expensive in other currencies, which can dampen demand and lower prices.
4. Gold_log and TYX_log (-0.49): Long-term yields reflecting an expectation of economic growth and inflation can lead to lower gold prices, as gold is a non-yielding asset and is less attractive when yields rise.

From the discussion above, it can be shown that strong multicollinearity exists in the training data, causing bias in the estimations of regression coefficients or inflating their variances.

Before reducing the collinearity, we can first analyze the regression result from the base model. Only S&P_log has a significant impact on the stock return on the significance level of 0.05. From the R-squared, we can also see that 0.06 of the changes in stock price can be explained by the model for the evaluation dataset.

Feature Engineering

There can be several measures to reduce the multicollinearity.

We can choose between Gold_log or DX_log because of the inverse relationship that the dollar value is directly linked to the gold value. After some experimenting, I decided to not use both of them because they also have impacts on the S&P_log as well.

Also, I chose to use TYX-IRX to reflect the difference between people's long-run and short-run economic expectations, aiming to capture the spread between long-term and short-term interest rates as a measure of the yield curve. The absolute log transformation can help in dealing with negative values and stabilizing variance.

```

OLS Regression Results
=====
Dep. Variable:      TSLA return      R-squared (uncentered):    0.230
Model:              OLS              Adj. R-squared (uncentered): 0.178
Method:             Least Squares     F-statistic:               4.427
Date:               Fri, 16 Feb 2024   Prob (F-statistic):        0.000580
Time:               21:23:34          Log-Likelihood:            -15.675
No. Observations:   95                AIC:                       43.35
Df Residuals:       89                BIC:                       58.67
Df Model:            6
Covariance Type:    nonrobust
=====
                    coef      std err      t      P>|t|      [0.025      0.975]
-----
SGP_log      0.4390      0.100      4.317      0.000      0.232      0.620
Gold_log     0.0384      0.090      0.427      0.670     -0.140      0.217
VIX_log      0.1165      0.079      1.480      0.142     -0.040      0.273
DX_log       0.0477      0.097      0.494      0.622     -0.144      0.240
IRX_log      -0.1653      0.118     -1.404      0.164     -0.399      0.069
TYX_log      0.0158      0.090     -0.175      0.862     -0.164      0.195
=====
Omnibus:                 3.436      Durbin-Watson:           1.848
Prob(Omnibus):            0.179      Jarque-Bera (JB):         2.826
Skew:                     0.403      Prob(JB):                 0.243
Kurtosis:                 3.254      Cond. No.                 2.58
=====

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

```

Final Results

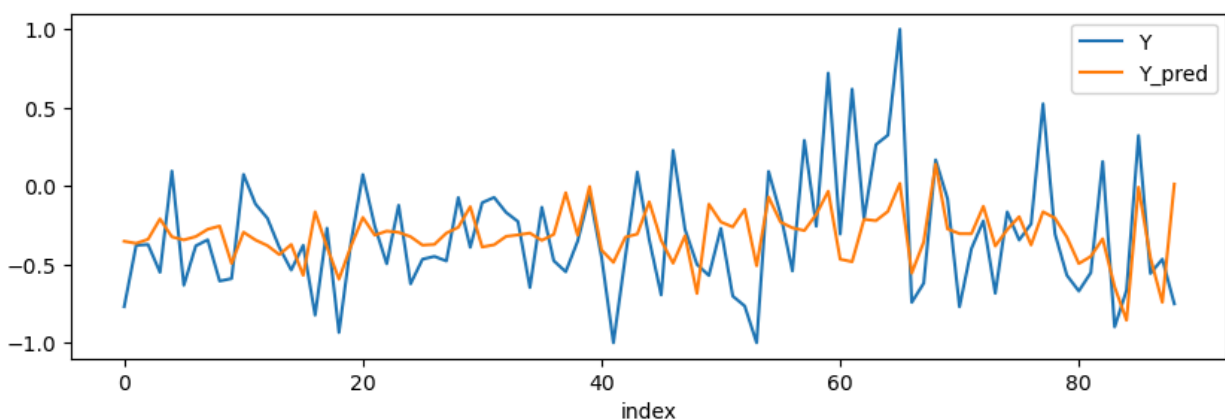
Firstly, there is a strong and significant positive relationship between Tesla return and the changes in S&P index, indicating that Tesla has a strong market share and influences in the US market. At the same time, we can also infer that Tesla stock return is sensitive to the market movement.

The positive coefficient of 0.3198 for 'TYX-IRX' actually indicates that, holding other factors constant, a one-unit increase in the absolute log difference between TYX (30-year Treasury yield) and IRX (13-week Treasury bill rate) is associated with a 0.3198 increase in Tesla's return. This suggests that as the spread between long-term and short-term interest rates increases, Tesla's return tends to increase as well.

This positive relationship suggests that Tesla's stock returns may benefit from a steepening yield curve. A steepening yield curve is often interpreted as an expectation of future economic growth, which could be favorable for Tesla's business outlook. For investors, the positive coefficient for 'TYX-IRX' could indicate that Tesla's stock is a more attractive investment during periods when the economy is expected to grow.

If we analyze the model performance, we can also see an improvement in the estimation. Contrary to the original R^2 of 0.06 in the evaluation dataset, the R^2 improves to 0.40, which means that the base model has been overfitting.

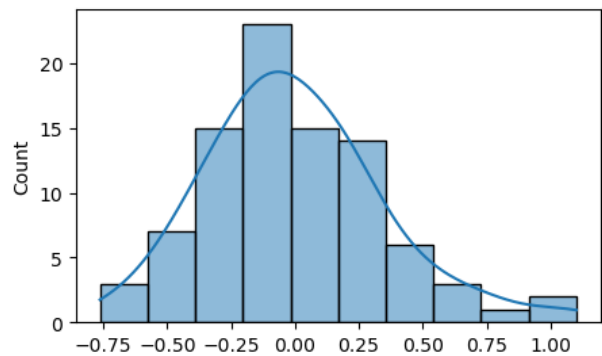
```
independent variables: ['S&P_log' 'TYX-IRX']
Coefficients: [ 0.32873018 -0.28339501]
Intercept: -0.4354714685988167
=====Evaluation=====
Mean squared error (MSE): 0.08
Coefficient of determination ( $R^2$ ): 0.40
=====OLS Regression Results=====
Dep. Variable:      TSLA return      R-squared (uncentered):      0.328
Model:              OLS              Adj. R-squared (uncentered):  0.308
Method:              Least Squares   F-statistic:                  16.83
Date:                Sun, 25 Feb 2024 Prob (F-statistic):           1.11e-06
Time:                15:17:22        Log-Likelihood:               -35.509
No. Observations:    71              AIC:                          75.02
Df Residuals:        69              BIC:                          79.54
Df Model:             2
Covariance Type:     nonrobust
=====
              coef      std err      t      P>|t|      [0.025      0.975]
-----
S&P_log      0.4627      0.139      3.328      0.001      0.185      0.740
TYX-IRX      0.3198      0.077      4.172      0.000      0.167      0.473
=====
Omnibus:              11.544      Durbin-Watson:           1.645
Prob(Omnibus):        0.003      Jarque-Bera (JB):        12.196
---
Notes:
[1]  $R^2$  is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```



I think this is caused by the limitation in the training dataset. The stock price is generally related to the demand and supply relationships of investors in the stock market. For an international company like Tesla, its stock performance will also be directly linked to international trade, which will also be affected by the strength of the dollar and interest rates, since they will influence the

export of the products and the supply chain cost, which will be more complex than the given data.

As it has been shown in the residual diagram, which is nearly a normal distribution, the model could capture the general trend of the stock return movement but fails to explain the stock in more details. It might require better models and larger datasets to achieve.



Conclusion

Overall, the regression analysis provides valuable insights into the relationship between Tesla's stock returns, market movements, and interest rate return spreads. It underscores the importance of broader market trends and monetary policy expectations in influencing the stock's performance. However, the complexity of stock price movements necessitates a broader analytical framework incorporating more variables and possibly different modeling techniques to capture the multifaceted influences on Tesla's returns.