

# Financial Data Analysis Final Project

陳明鴻

National Chengchi University(NCCU)

Author Note

Ming-Hung Chen, Department of Business Administration, NCCU

Student ID: 107305103

Contact Info: 107305103@nccu.edu.tw

## **Introduction**

Back in the 2010s, while the EU announced the plan of ban on selling fuel automobiles in 2030, it seemed the end of traditional automobile enterprise is approaching. However, with the improvement of lithium batteries, most of the industry leaders started showcasing the premiere of their own solutions to the ban and at the same time numbers of new manufacturer have raise.

Comes with this worldwide EV trend, lots of newly founded manufacturers had been listed on the market and had caused commotions. For example, Tesla(TSLA) has reached 1,200 USD in 2021, Lucid(LCID) had grown 200% in 2021 and Nikola(NKLA) had a unreal 500% growth in 2020.

However, whether these ‘rookie’ companies worst a shot, and will perform like those golden geese in Silicon Valley, Google, Apple, Microsoft, etc., in the last two decades, or will they become flashes in the pan?

In this research, the answer will be found out by comparing various assets in the car industry and construct types of portfolio with these assets to test out the most profitable strategy.

## **Tools and Data Introductions**

The research will be conducted by:

1. Calculate individual assets’ return, cumulative return, and other statistics to examine whether newly founded car manufacturers outperform traditional manufacturers and have higher volatility.
2. Construct different types of portfolio by using in-sample data to calculate out-of-sample return and accurately test out which of them outperform the whole industry without being beaten by serious fluctuation of certain assets.
3. Conclude with ways to modify strategies in order to found the optimal strategy to invest in the car industry.
4. Note: individual assets’ data were retrieved from Yahoo Finance in daily format, and the final result will be display in percentage and annualized(if necessary). The date of the original data is between: 2020-09-21 to 2022-06-15, it is relatively short due to lots of New Energy(NE) stocks were listed in recent years.
5. Data in 2020-09-21~2021-09-21 will be used as sample data to construct portfolio weights. The out-of-sample period will be 2021-09-22 to 2022-06-15 with 185 total trading days
6. Risk Free Rate(RF) is set to 3% annually (mean of Sep. 2021 to Jun. 2022). Transaction Fee is set to 0.035%.

### **Assets choosing and brief introduction**

1. *Arrival(ARVL)*, a new Car manufacturer focus on electric buses.
2. *BMW(BMWYY)*, a traditional German Car manufacturer.
3. *BYD(BYDDY)*, a Chinese company makes car batteries and electric vehicles.
4. *Centro Electric Group(CENN)*, an Australian commercial vehicles manufacturer.
5. *Daimler AG(DMLRY)*, a prestigious car manufacturer with longest history.
6. *Ford(F)*, an American Car manufacturer with top market value in the industry
7. *General Motor(GM)*, a traditional American Car manufacturer with types of brands underneath: Jeep, Hummer, Chevrolet, etc.
8. *Honda Motor(HMC)*, a prestigious Japanese Car manufacturer.
9. *Ferrari(RACE)*, the most famous race car/super car manufacturer.
10. *Lucid(LCID)*, an American manufacturer focus on making luxury EVs.
11. *LI Auto(LI)*: a Chinese electric vehicle manufacturer which enter the market in 2015 and has top market share in Chinese market.
12. *Mullen Auto.(MULN)*: an American electric vehicle manufacturer which enter the market in 2015.
13. *NIO Inc.(NIO)*: a Chinese EV manufacturer which has top market share in Chinese market.
14. *Nikola(NKLA)*: an American manufacturer focus on making electric trucks.
15. *Porsche Auto Holdings(POAHY)*: a holding company which is the biggest shareholder of VW AG.
16. *Fisker(FSR)*: an American manufacturer founded in 2016.
17. *Stellantis(STLA)*: a traditional car manufacturer merged by Fiat Chrysler Auto and PSA Peugeot Citroën
18. *Toyota Motor(TM)*: a prestigious Japanese car manufacturer with great reputation on its reliability.
19. *Tesla(TSLA)*: an American manufacturer with the highest market value in car industry.
20. *Tata Motors(TTM)*: an Indian Car manufacturer.
21. *XPeng(XPEV)*: a Chinese electric vehicle manufacturer which has top market share in Chinese market.

### **Assets choosing Explanation:**

The main reason of choosing assets in car industry, including newly rise companies(NKLA, LCID, NIO, etc.) and prestigious car manufacturers with long history(Ford, GM, TM, etc.), is to examine how performance will be affected between portfolios with different strategies, especially those using variance as filters since newly rise companies often have higher volatility and traditional car manufacturers tend to have lower volatility.

## **Portfolio choosing and brief Introduction:**

1. *Fixed-weight(FW):*

Allocate equal weight( $1/21$ ) to all assets at beginning to check the whole industry's performance and set it as a benchmark to examine whether other portfolios outperform it.

2. *Price-weight(PW):*

Allocate weights according to assets' prices, assets with higher prices will have bigger weight in the portfolio. The feature of Price-weight portfolio is, serious fluctuation might appear while few assets with highest prices suddenly crash or soar. For example, Tesla which has the highest price would be the decisive one to the performance of Price-weight portfolio.

3. *Global Minimum Variance Portfolio(GMVP):*

Using assets that have lowest variance to pursue minimum variance.

4. *Mean Variance Frontier Portfolio(MVP):*

MVP is supposed to have better performance with lower variance comparing to GMVP due to the fact that MVP's return/Var. plot is a hyperbola but GMVP is only the left most point. In the research, comparison can be done by using multiple statistics tools.

5. *Mean Variance Portfolio with Risk Free asset(RFMVP):*

Combine MVP with risk-free asset can optimize MVP, and the comparison can also be done in the research.

6. *No Short-Sale Global Mean Variance Portfolio(NSGMVP):*

GMVP with a constraint that does not Short Sale asset, it might have lower return compare to GMVP but depends on market fluctuations.

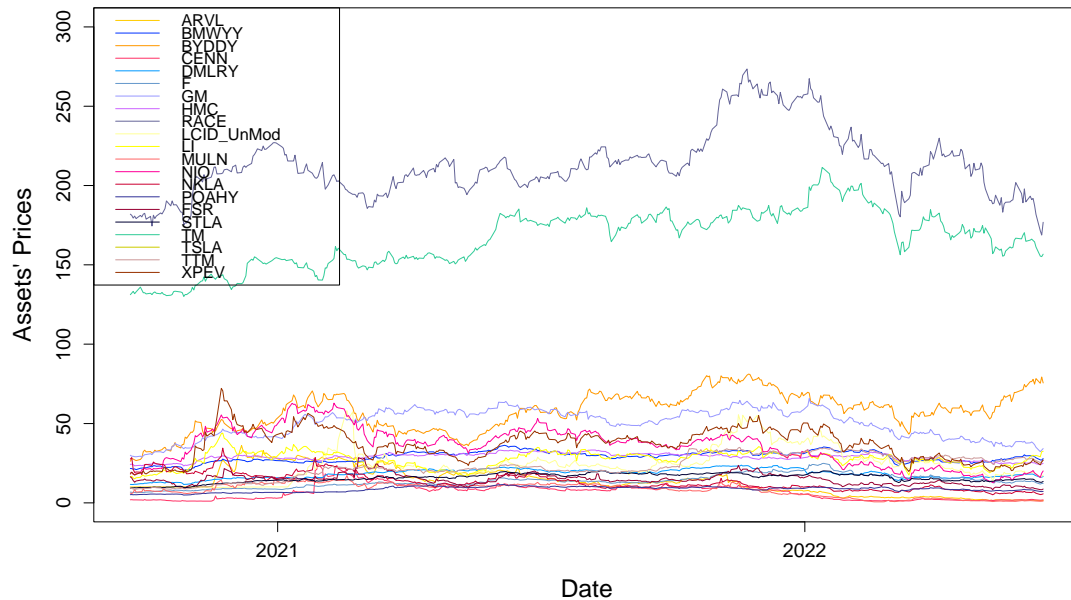
7. *Tangency Portfolio(TAN):*

Tangency Portfolio is the portfolio that maximizes the Sharpe ratio, which means it fully invests in risky assets. It might retrieve highest return during bull market but might crash more than other portfolio. The reason of choosing Tangency Portfolio is to examine how portfolio with high volatility will perform in industry with numbers of risky stocks.

## Individual Assets Comparison

To examine whether newly founded car manufacturers outperform traditional manufacturers, we first inspect the price fluctuation over the whole period:

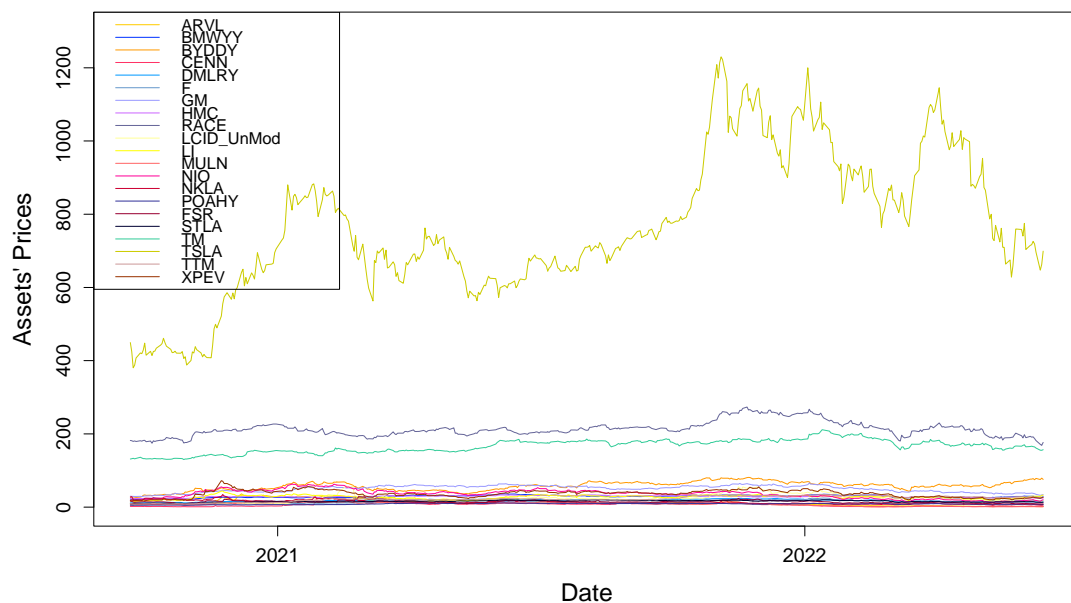
**Time Series Plot of Assets' Prices without TSLA**



In this plot(without TSLA), we can only noticed traditional car manufacturers might also experience serious fluctuations like Ferrari(RACE), and Toyota Motor(TM). As for BYD(BYDDY), it is soaring compare to other companies.

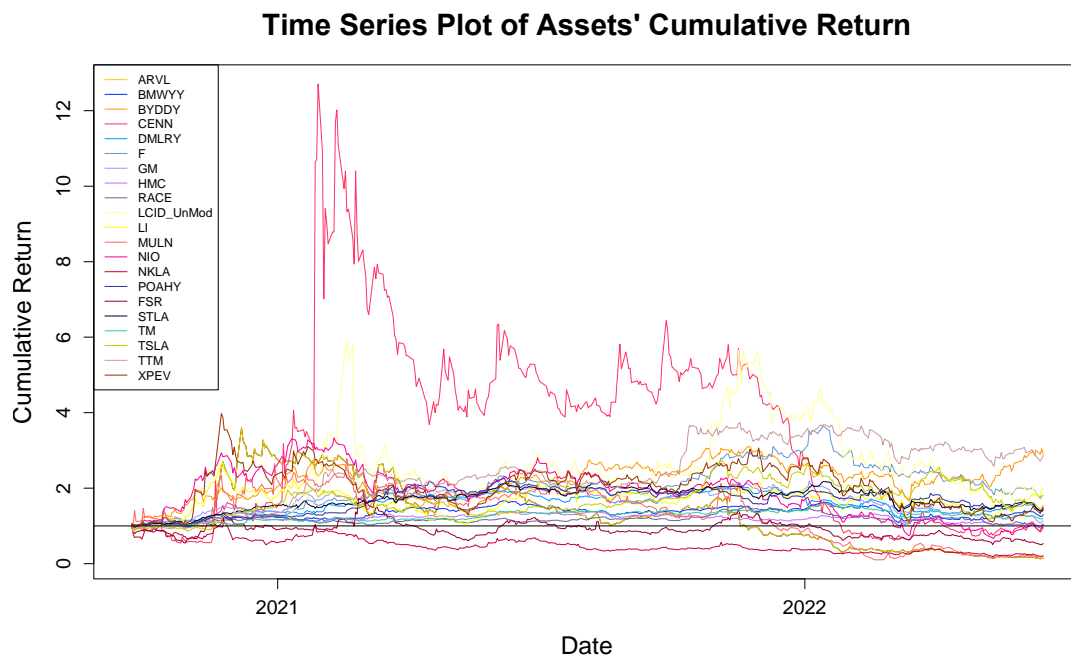
However, if we add TSLA into the plot:

**Time Series Plot of Assets' Prices**



By adding in TSLA, we can find out that TSLA's fluctuation is much more higher than RACE, GM. Therefore we cannot conclude that neither NE stocks nor non-NE related car manufacturers have higher volatility.

To further examine performance of target assets, calculation of cumulative return during the whole period is needed:

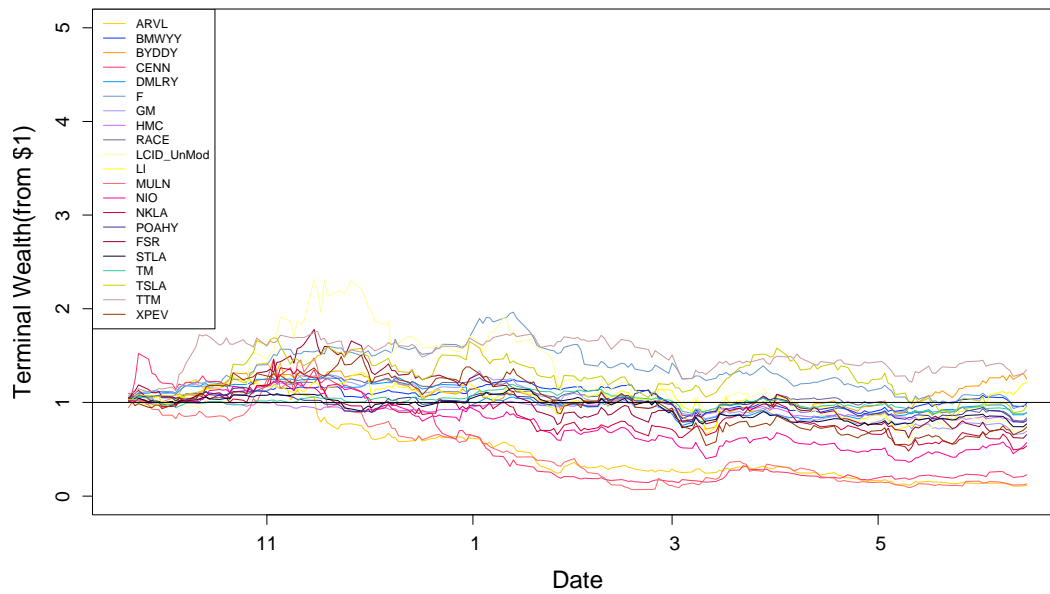


In this plot, CENN had the most cumulative return at the beginning of 2021 but drop to lower than 1 at around 2022, and LCID also had relatively good performance at 2021 but later crash in a few months. These two assets have shown the feature of New Energy related car manufacturers from our hypothesis.

In the out-of-sample period (will later be use in portfolio examinations)the cumulative return was a totally different case due to the high based-period it had. The whole industry seemed to have lower cumulative return with smaller fluctuation. The assets which possessed the top two highest cumulative return, TTM and BYDYY, had respectfully 176.0% and 133.5% cumulative return.

To shortly brief up, comparing prices, return, and cumulative return does not support the hypothesis:

## Time Series Plot of Assets' Cumulative Return During Out of Sample Period



## Statistics of Individual Assets Comparison

As a result, to verify the hypothesis, we will have to calculate these assets' daily returns' statistics (Min., Median, Max., mean, Std., skewness, kurtosis, ACF, annualized and shown in percentage):

	Nums.of	Orb:Min.	X1st.Qu.	Median	Mean	X3rd.Qu.	Max.	Var.	Std.	SKewness.	Kurtosis.	ACF1.
ARVL	437	-27.6	-3.8	-0.3	-0.225	2.5	49.7	11647.644	107.924	0.093	0.037	0.093
BMWYY	437	-6.7	-1	0.1	0.078	1.2	7.5	974.892	31.223	-0.007	0.005	0.056
BYDDY	437	-10.5	-2.5	0.1	0.324	2.7	22.5	4099.23	64.025	0.039	0.008	-0.059
CENN	437	-41.6	-4.6	-1.6	0.812	2.8	252.3	69480.017	263.591	0.591	0.511	-0.027
DMLRY	437	-17.5	-1.1	0.1	0.106	1.3	6.5	1235.738	35.153	-0.072	0.036	0.018
F	437	-9.7	-1.5	-0.1	0.173	1.8	11.7	1955.81	44.225	0.023	0.006	-0.013
GM	437	-8.9	-1.6	-0.1	0.062	1.6	9.7	1650.446	40.626	0.013	0.003	0.006
HMC	437	-4.7	-1	0.1	0.019	1	6	615.172	24.803	0.002	0.001	0.03
RACE	437	-7.6	-1	0.1	0.012	0.9	8.9	876.284	29.602	0.01	0.01	-0.033
LCID_UnMo	437	-38.6	-3	-0.1	0.341	2.5	31.8	10624.131	103.073	0.054	0.03	0.098
LI	437	-20.6	-2.5	-0.1	0.285	2.4	31.9	6601.753	81.251	0.059	0.023	-0.043
MULN	437	-39.1	-4.5	-0.9	0.145	3.1	144.9	33132.302	182.023	0.327	0.235	-0.031
NIO	437	-15.2	-3	-0.3	0.145	2.9	25.6	6864.713	82.854	0.047	0.009	-0.021
NKLA	437	-26.9	-3.7	-0.5	-0.167	3.3	23.6	9741.699	98.7	0.014	0.01	0.017
POAHY	437	-8.9	-1.2	0	0.109	1.3	11.7	1535.322	39.183	0.033	0.011	0.123
FSR	437	-17.2	-3.4	-0.6	0.021	2.5	38.6	8920.523	94.449	0.134	0.041	-0.019
STLA	437	-8.7	-1.2	0.1	0.116	1.4	11.5	1496.9	38.69	0.005	0.009	-0.015
TM	437	-6.2	-0.8	0.1	0.053	1	6.9	606.03	24.618	-0.005	0.008	0.046
TSLA	437	-12.2	-2.1	0.2	0.174	2.3	19.6	3713.441	60.938	0.012	0.008	-0.036
TTM	437	-12.6	-1.4	0.2	0.291	1.9	16.3	2523.937	50.239	0.044	0.017	0.002
XPEV	437	-15.1	-3.4	-0.2	0.265	2.8	33.9	9159.343	95.704	0.097	0.027	0.023

(Statistics of Individual Assets)<sub>NE Co. shown in lighter orange</sub>

First, most of NE related stocks have higher Max. and lower Min. compare to those others, for example, CENN has 252% Max. daily return and MULN has 144% Max. daily return, but LCID, LI, and MULN also have Min. return lower than 20%.

As for variances and standard deviations, NE stocks' are relatively high which verified the hypothesis of NE stocks having higher volatilities.

However, NE stocks and non-NE stocks do have relatively small skewness and kurtosis, which implied there will not be any fat tail. Also, the ACF implies that the assets' average correlation between data points are relatively small as well.

	Nums.of.Obs	Min.	X1st.Qu.	Median	Mean	X3rd.Qu.	Max.	Var.	Std.	SKewness.	Kurtosis.	ACFI.
ARVL	185	-27.6	-5.5	-1.1	-0.923	3.4	23.7	12176.181	110.346	0.003	0.006	-0.05
BMWYY	185	-6.7	-1.2	0.1	0.011	1.4	5.8	1190.616	34.505	-0.023	0.003	0.077
BYDDY	185	-8.7	-2.5	0.1	0.176	2.8	22.5	3645.372	60.377	0.059	0.021	-0.086
CENN	185	-16.4	-4.7	-1.6	-0.554	2.8	26.4	13383.09	115.685	0.06	0.007	0.099
DMLRY	185	-17.5	-1.1	0.2	-0.03	1.5	6.5	1633.681	40.419	-0.114	0.046	0.036
F	185	-9.7	-2	-0.2	0.035	1.8	11.7	2474.752	49.747	0.027	0.006	0.002
GM	185	-7.8	-1.9	-0.2	-0.162	1.6	7.5	1887.367	43.444	0.004	0	0.03
HMC	185	-4.7	-1.2	0	-0.096	1.1	4.1	687.462	26.22	-0.01	-0.001	-0.05
RACE	185	-7.6	-1.5	0	-0.076	1.1	8.9	1267.882	35.607	0.018	0.006	-0.045
LCID_UnMo	185	-18.3	-3.7	-0.5	-0.037	3.1	31.3	9823.357	99.113	0.063	0.016	-0.023
LI	185	-20.6	-2.4	-0.2	0.235	2.5	31.9	6975.714	83.521	0.057	0.032	0.025
MULN	185	-39.1	-7	-2.6	-0.21	3.1	144.9	57173.939	239.111	0.328	0.186	0.011
NIO	185	-15.2	-3.6	-0.6	-0.141	3.1	25.6	8146.404	90.257	0.05	0.01	0.032
NKLA	185	-14.9	-4	-0.7	-0.181	3.4	21.5	8369.262	91.484	0.036	0.005	-0.027
POAHY	185	-8.9	-1.4	-0.2	-0.094	1.2	11.7	1929.981	43.932	0.047	0.014	0.196
FSR	185	-12.8	-3.6	-0.3	-0.117	2.5	13	5770.395	75.963	0.025	0.001	0.003
STLA	185	-8.7	-1.6	-0.1	-0.097	1.6	11.5	2035.883	45.121	0	0.007	0.005
TM	185	-6.2	-1	0	-0.063	1.1	6.9	846.597	29.096	-0.004	0.005	0.039
TSLA	185	-12.2	-2.4	0.3	0.062	2.5	13.5	4554.541	67.487	-0.006	0.003	-0.025
TTM	185	-7.7	-1.6	-0.1	0.199	2.1	15.3	2679.636	51.765	0.058	0.012	-0.004
XPEV	185	-13.8	-3.6	-0.2	-0.01	3.2	29.6	8143.876	90.243	0.049	0.013	-0.048

(Statistics of Individual Assets, out-of-sample period)NE Co. shown in lighter orange

The statistics of individual assets in out-of-sample period also explain the same result that NE stocks possess higher Max. return and larger standard deviations

### Conclusion of Individual Assets Comparison

The Hypothesis of NE stocks have greater fluctuations is verified by examining returns with Max., Min., standard deviation. Without having greater cumulative returns, NE assets are more suitable for short-term trades to pursue high risk premium.

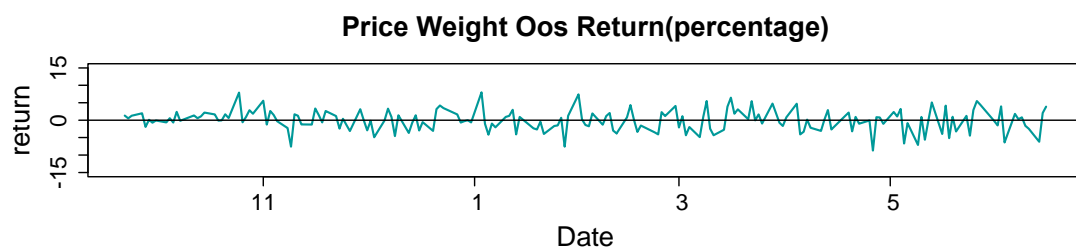
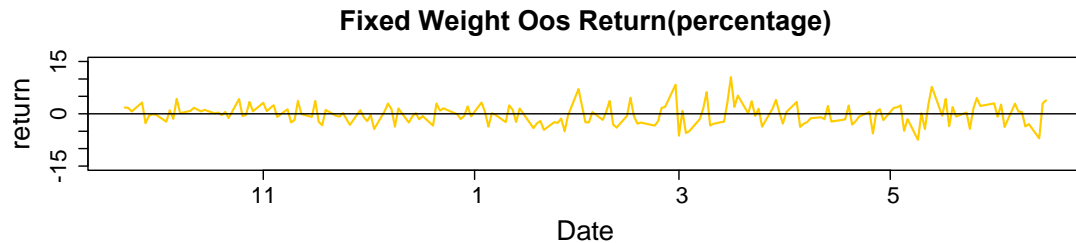
If one is tolerable with such high volatility, investing in NE stocks might performs incredibly well, if not, one should choose to invest in traditional car manufacturers' stocks and hold for long-term, which will supposedly retrieve nearly much cumulative return compare to NE stocks without enduring great risk.



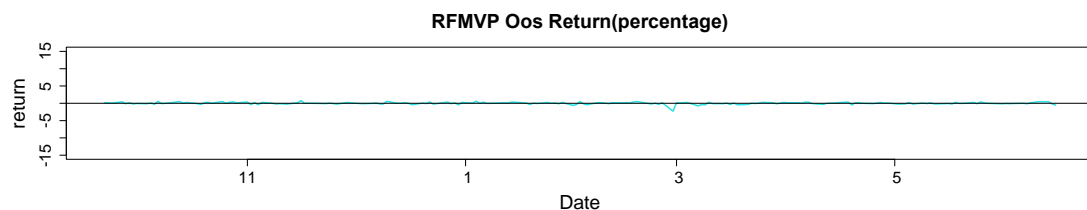
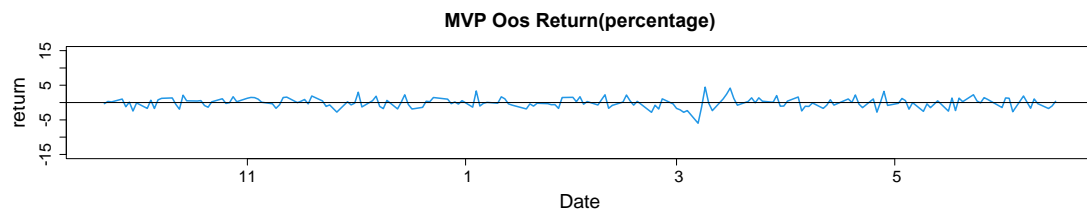
## Portfolio Comparison

### Out-of-sample Return Comparison

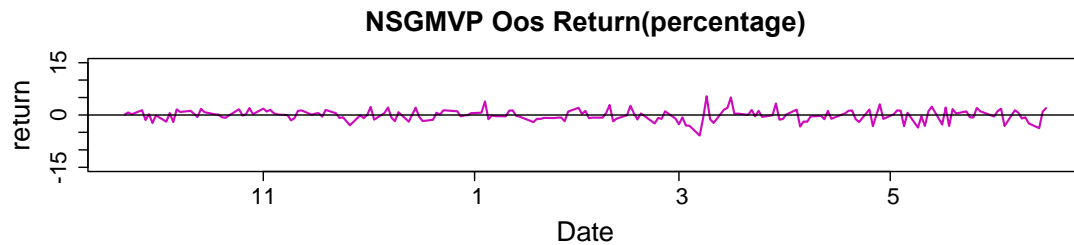
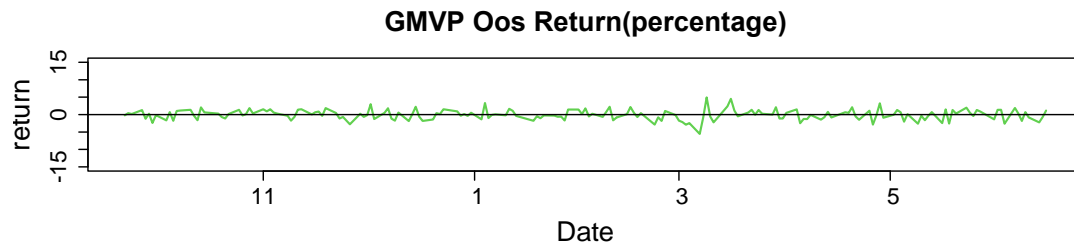
To examine which portfolio is the most suitable one in car industry, returns of each portfolio will be inspected:



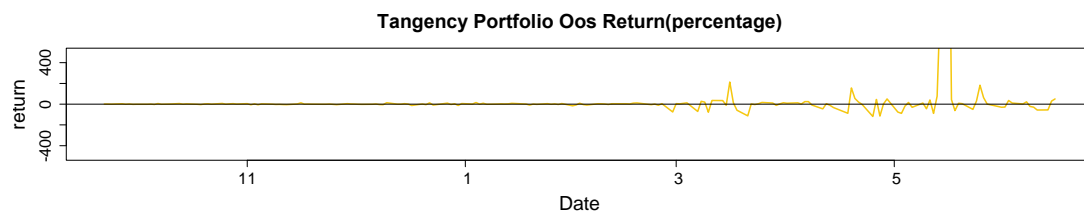
Assets relocations are needed in Fixed weight and Price weight portfolio, and price weight relocate according to asset's price. However, the overall returns of these two portfolio do not distinguish really well.



The goal of mean-variance optimization is to maximize an asset's return based on the same risk, and adding risk-free assets can further optimize it more. In this case, the RFMVP clearly choose to weight risk-free assets more than car industry stocks, which implies that risk-free assets have greater return on same scale of risk or the setting of Risk Free rate is slightly high. However, further inspection on cumulative return is needed to verify if risk-free assets have greater return compare to car industry stocks on same scale of risk.



Short-sale is an often-seen hedging tools in a portfolio to compromise the risk brought by over-grown assets. However, NSGMVP has the constraint that short-sale is not allowed, therefore greater fluctuations were caused in this case. Further inspections on cumulative return can verify which is the more suitable portfolio.



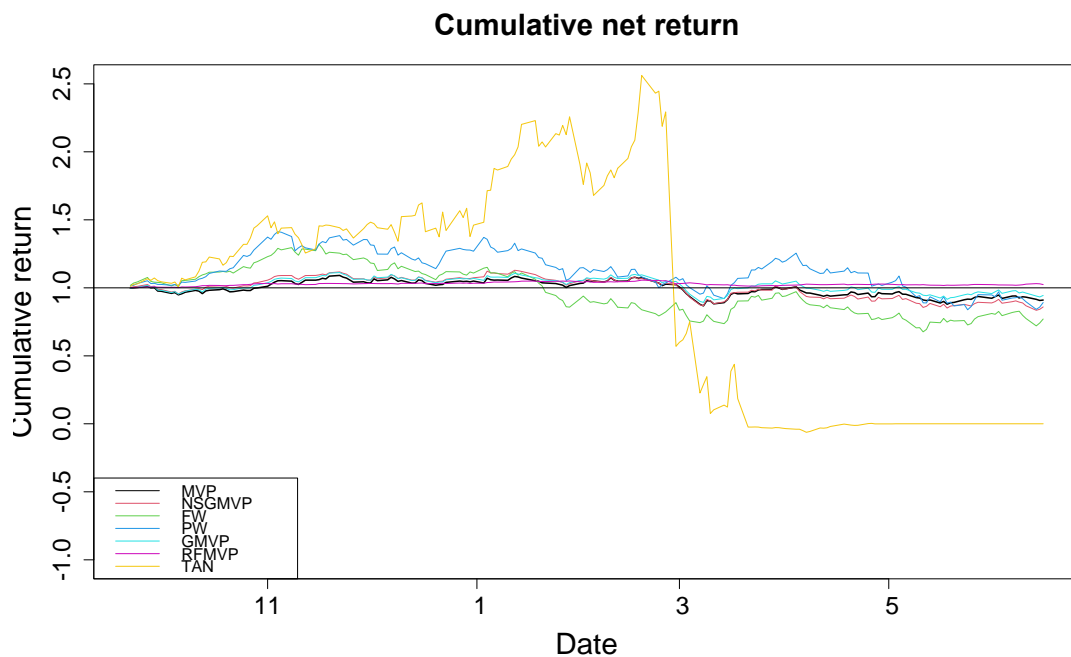
Noted: The return percentage's upper and lower limit in this plot is 400% ~ -400%

Tangency portfolio maximize the risk to retrieve higher return. In this case, assets with extremely high variance have been blended in, therefore serious fluctuations are inevitable.

### Conclusion of Comparisons of Portfolio Return

To sum up briefly, Tangency Portfolio, Fixed-weight and Price-weight seem to perform better and have greater fluctuations due to the fact that assets with high variance were in the list, which were the assets that GMVP, MVP kind of portfolio might avoid blending or might be given smaller weights. Nevertheless, comparisons in cumulative return is required to verify which portfolio performs the best.

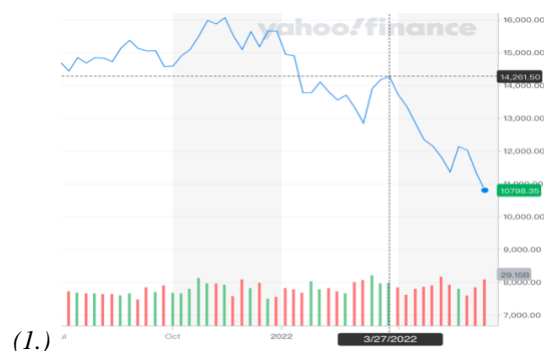
## Out-of-sample Cumulative Return Comparison



By inspecting cumulative return data, it has been denied that Tangency Portfolio, Fixed-weight and Price-weight Portfolio are more suitable portfolio. The reasons why they crashed and turned out to possess cumulative return lower than 1 is:

1. It is a relative high base-period for the whole car industry, and the whole market as well<sup>(1)</sup>.
2. Tangency Portfolio, Fixed-weight and Price-weight Portfolio contains high risk assets with greater weights compare to MVP, GMVP. Therefore while encountering market crash (mentioned in 1.) these portfolios were more likely to crash more seriously.

However, the final survivor: RFMVP which retain cumulative return that higher than 1 is mostly because of highly weighting in risk-free assets, not the strategy it applied on the car industry stocks.



*Nasdaq's crash from March 22*

## Statistics of Portfolios

To further inspect each portfolios' feature, examinations on each portfolios' statistic is required and will started from descriptive statistics:

	No..of.Obs.	Min.	X1st.Qu.	Median	Mean	X3rd.Qu.	Max.	Var.	Std.	SKewness.	Kurtosis.	ACF1.
FW	185	-7.47	-2.29	-0.162	-0.1	1.69	10.526	2091.94	45.738	0.022	0.003	0.049
PW	185	-8.724	-1.914	0.158	-0.016	1.695	7.995	2300.491	47.963	-0.009	0.001	-0.015
MVP	185	-5.98	-1.001	-0.016	-0.04	1.02	4.439	535.094	23.132	-0.007	0.005	0.055
RfMVP	185	-2.129	-0.11	0.029	0.013	0.156	0.719	19.35	4.399	-0.158	0.074	-0.065
GMVP	185	-5.532	-0.915	0.052	-0.02	0.929	4.944	545.557	23.357	0	0.004	0.046
NSGMVP	185	-5.889	-1.035	0.004	-0.069	1.082	5.362	674.855	25.978	-0.005	0.004	0.027
TAN	185	-118.234	-3.838	1.462	17.362	6.166	3167.935	14014896	3743.647	0.825	0.698	0.032
GM	185	-7.8	-1.9	-0.2	-0.162	1.6	7.5	1887.367	43.444	0.004	0	0.03
Ford	185	-9.7	-2	-0.2	0.035	1.8	11.7	2474.752	49.747	0.027	0.006	0.002
TSLA	185	-12.2	-2.4	0.3	0.062	2.5	13.5	4554.541	67.487	-0.006	0.003	-0.025
Lucid	185	-18.3	-3.7	-0.5	-0.037	3.1	31.3	9823.357	99.113	0.063	0.016	-0.023
NIO	185	-15.2	-3.6	-0.6	-0.141	3.1	25.6	8146.404	90.257	0.05	0.01	0.032

### (Descriptive Statistics of portfolio return, mixed with Individual Assets)

Mean of each portfolios' (except TAN Por.)return do not outperform example assets and their variances are not specifically lower than example assets'. Statements above might imply that constructing portfolios might not be better than buying single assets. However, it will have to be reminded that the out-of-sample period is a relative high base-period and the US market was undergoing a large adjustment.

Min. and Max. of TAN portfolio are extremely different from other portfolios because of certain risky assets were weighted relatively high.

	VaR(0.05)	ES(0.05)	LPSD
FW	-4.839	-6.043	2.029
PW	-4.986	-6.248	2.204
MVP	-2.437	-3.046	1.065
RfMVP	-0.442	-0.558	0.222
GMVP	-2.441	-3.055	1.059
NSGMVP	-2.76	-3.444	1.218
TAN	-370.54	-469.082	24.706
GM	-4.663	-5.807	2.012
Ford	-5.119	-6.429	2.08
TSLA	-6.931	-8.708	3.041
Lucid	-10.306	-12.915	3.977
NIO	-9.493	-11.869	3.74

### (Risk Statistics, mixed with Individual Assets)

Except TAN portfolio, other portfolios have relatively lower VaR(\$1 invested) and ES compare to both NE stocks and non-NE stocks. It can corroborate that by constructing portfolio instead of buy individual assets, risk can be decreased.

FW	PW	MVP	RfMVP	GMVP	NSGMVP	TAN
-0.614	-0.146	-0.568	0.086	-0.349	-0.781	1.168

### (Sharpe Ratio)

Only RFMVP and TAN Por. retrieve positive Sharpe ratio, which means that only there two portfolio outperform risk-free asset on the same scale of risk.

## **Conclusion**

It has been verified that most of the portfolios do not outperform certain single assets no matter NE stocks or non-NE stocks. The Sharpe ratio also implied that most of them cannot beat risk-free asset on the same scale of risk. However, it is important to notice that the out-of-sample period is in a relatively high base-period, and portfolios with risky assets might have encountered major lost. That is also the reason why RFMVP is the portfolio that outperform others in cumulative return. Plus, whether these portfolios perform poorly could be further verified by comparing with cumulative return of the whole market.

To improve the strategy, assets with high risk should still be added into the portfolio but would be recommend to invest in short-term and need to be aware of the base-period issue. During high base-period, if one insist in investing risky NE assets or portfolios, it is recommended to add more risk-free asset into the portfolio, like what RFGMVP does, as a hedging tool. The other way to optimize the portfolio is to weight more in iconic assets like TSLA, TM, BYDDY, etc., to prevent sudden crush brought by certain risky NE assets.

To sum up, the 7 types of portfolio do not perform well in the out-of-sample period, but it is still worthy of attention that certain portfolio. For example, Tangency portfolio retrieve 250% cumulative return in the first four month in the out-of-sample period. If at another time with a rather normal base-period, it can perform extremely well against other portfolios.

## Appendix

### Other Statistics of Portfolios

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.
FW	0.383	0.758	1.16	2.978	2.245	79.143	7.372
PW	0.192	0.456	0.666	1.981	1.119	61.614	6.384
MVP	0.385	1.226	2.2	15.569	4.71	1500.95	114.079
RfMVP	0.204	0.599	0.876	1.095	1.303	6.162	0.865
GMVP	0.414	1.124	1.97	10.284	4.009	1138.33	83.695
NSGMVP	0.521	1.49	2.429	15.737	4.66	1501.041	114.094
TAN	1.551	8.09	16.664	109.603	70.482	9511.013	700.988

### (Turn Over Rate)

TAN Por. has higher TOR, which means it sales and buys assets frequently.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.
FW	0	0	0	0	0	0	0
PW	0	0	0	0	0	0	0
MVP	0.178	0.214	0.271	0.279	0.338	0.414	0.066
RfMVP	0.51	0.773	0.834	0.845	0.981	1.112	0.142
GMVP	0.188	0.209	0.224	0.258	0.319	0.369	0.06
NSGMVP	0	0	0	0	0	0	0
TAN	0.51	0.773	0.834	0.837	0.966	1	0.132

### (Shortsale-long Ratio)

FW, PW, and NSGMVP have 0 SLR due to the fact that they are not allowed to short sale assets.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.
FW	0.048	0.048	0.048	0.048	0.048	0.048	0
PW	0.261	0.292	0.308	0.311	0.327	0.373	0.028
MVP	0.102	0.116	0.131	0.127	0.138	0.146	0.012
RfMVP	0.065	0.078	0.086	0.083	0.09	0.099	0.008
GMVP	0.12	0.131	0.135	0.136	0.139	0.153	0.007
NSGMVP	0.229	0.244	0.272	0.289	0.349	0.381	0.052
TAN	0.065	0.078	0.086	0.083	0.09	0.099	0.008

### (Herfindahl-Hirschman Index)

## Individual Assets' Return

