PORTFOLIO THEORY AND ASSET PRICING

Submission Number: 2

Group Number: 3 - A

Group Members:

Mantobaye Moundigbaye mantobaye@yahoo.fr
Louis Regnier-Vigouroux louisregnierv@gmail.com
Yonas Menghis Berhe Yonix500@gmail.com

8)
8.1)

According to our comparison table below, most of the sectors performed better on the LEI regressor.

Sector	Fund Name	Score	MSE
XLK	Technology Select Sector SPDR Fund	0.0948144	30.7937
SPY	SPDR S&P 500 ETF Trust	-0.0234417	177.41
XLY	Consumer Discretionary Select Sector SPDR Fund	-0.105958	40.7723
XLI	Industrial Select Sector SPDR Fund	-0.19269	16.4004
XLU	Utilities Select Sector SPDR Fund	-0.297895	20.8688
XLF	Financial Select Sector SPDR Fund	-0.587755	2.1849
XLE	Energy Select Sector SPDR Fund	-1.06844	65.3343
XOP	SPDR S&P Oil & Gas Exploration & Production ETF	-1.45582	1788.56
XLB	Materials Select Sector SPDR Fund	-2.55183	18.4086

These funds track the main sources of input and outputs of an economy. Collectively, they aggregate to most of the value generated in the economy. The high scores obtained from the regression results on the LEI are justified, since such indicators also measure the performance of the major means of production and consumption within the economy. And these are directly or indirectly influenced by these funds.

8.2)

The one sector that performed best on the CEI is the GDX fund. This fund tracks a market-capitalization-weighted index constituted of global gold-mining firms. This fund is mostly used by investors as a hedge against inflation, weakening dollar and financial or economic and geopolitical uncertainty. Therefore, if investors and market analysts assume that the current state of an economy is unstable then they are more likely to purchase GDX funds.

8.3)

Unfortunately, our results did not show any fund performing best on the lagging indicators. However, two funds (namely the XLV (health care) and XLP (consumer staples) funds) performed better when regressed on all combined indicators than separately for each category of the indicators. The lagging indicators mainly confirm long-term trends in the economy. This could explain why they do not correlate extensively with current economic trends.

9.1) Non-Technical Report

In this project, we try to find the best set of predicting indicators for each of the selected SPDRs. To this end, we have used various types of models to find out what are the most suitable indicators in terms of variance explained.

First, we extracted data for all types of indicators, namely LEI (Leading Economic Indicators), CEI (Coincidental Economic Indicators) and LAG (Lagging Indicators), and for the selected SPDRs and the equity fund. The data was gathered over a 4-year period from 2016 to 2019. Different sets of data showed different frequencies, varying from weekly to monthly to quarterly. To allow for comparison of all data, we therefore computed the monthly and quarterly returns for the selected ETF funds.

The first set of models used linear regression to try to explain the variation of each ETF. We performed four sets of 11 regressions (1 for each of the 11 ETFs). Each set of regressions used different explanatory variables and were applied to each of the 11 ETFs. The explanatory variables were, respectively, the LEI, the LAG, the CEI and all indicators combined. For each regression model, we collected the R-score and the Mean Squared Error (MSE) for comparison.

The collected results were sorted by order of decreasing score. The R-score or the coefficient of determination is used to select the best indicator for each of the ETFs. Therefore, for each ETF, the regression with the highest R-score corresponds to the regression using the best set of indicators (LEI, CEI, LAG, ALL).

For the second set of models, we used LASSO regression and Regression Trees methods. Lasso regressions are mostly used to reduce complexity and it employs models with fewer parameters. Regression tree is another type of model that predicts outcome using variables that results in the most gain of information. However, as we can see from the Fig 2 and 3, these models only showed improved performance for a few ETFs. More specifically the LASSO method did not show any improvement in the regression score. Results were however improved using regression trees for both the XLI and XLB funds.

In addition to these two models, we performed a cluster analysis. However, this method just shows how different points in time can be clustered based on some common characteristics such as similar market conditions. It does not tell us anything regarding the assignment of the best predictor indicator for each fund. Therefore, we cannot compare the output of this model to the models previously presented.

9.2)

The group first had a meeting to discuss ways on how to approach the second submission. The discussion focused on the type of program to use and how to divide the workload equally to all the group members.

After this meeting, the tasks were distributed as follows:

Louis Regnier-Vigouroux

- Data collection and building linear regression models

Mantobaye Moundigbaye

- Perform LASSO regression, Cluster analysis and Regression Trees

Yonas Menghis Berhe

- Model interpretation and Report writing

After each group member completed their assigned responsibilities, we decided to switch tasks and review each other's work. This review process allowed the group to solve outstanding issues and improve the overall technical solution through feedbacks. Moreover, this allowed to exchange ideas that could result in a better model interpretation.

As a result of the review process Louis and Mantobaye updated and made the necessary changes to the code as to reflect the discussions and according to the data changes suggest by the staff. Finally, Yonas collected the results and the overall full report, which was then reviewed by all other group members.

	level_0	Symbols	coefs	score	mse
16	CEI	GDX	[-0.00012632721134356605, -2.0605257635791854,...	0.310195	9.32997
5	LEI	XLK	[0.003889823168008847, 12.617256249695256, -12...	0.0948144	30.7937
42	ALL	XLV	[-3.0364915631986805e-05, -3.1120269706653044e...	0.0129777	0.00764618
11	LEI	SPY	[0.0025647436066845862, 32.82275711591052, 6.7...	-0.0234417	177.41
7	LEI	XLY	[0.005875543503329326, 15.817608096174878, -13...	-0.105958	40.7723
3	LEI	XLI	[-0.0016062668778153131, 26.692047383550317, 2...	-0.19269	16.4004
2	LEI	XLU	[-0.006053796504675176, 4.862813071743094, -0....	-0.297895	20.8688
44	ALL	XLP	[-9.192149198207807e-05, -8.67437452331971e-05...	-0.430988	0.00442911
1	LEI	XLF	[-0.0002365366503321703, 4.351429897603006, 25...	-0.587755	2.1849
0	LEI	XLE	[0.011181975409731123, -49.51658719476371, -48...	-1.06844	65.3343
10	LEI	XOP	[0.04778811373272929, -212.7529241459583, -169...	-1.45582	1788.56
9	LEI	XLB	[-0.0014544605679136901, 11.510225833081316, 1...	-2.55183	18.4086

Fig1. Linear Regression Table

	level_0	Symbols	coefs	score	mse	nb_predictors
16	CEI	GDX	[-0.00012632721134356605, -2.0605257635791854,...	0.310195	9.32997	NaN
5	LEI	XLK	[0.003889823168008847, 12.617256249695256, -12...	0.0948144	30.7937	NaN
42	ALL	XLV	[-3.0364915631986805e-05, -3.1120269706653044e...	0.0129777	0.00764618	NaN
11	LEI	SPY	[0.0025647436066845862, 32.82275711591052, 6.7...	-0.0234417	177.41	NaN
7	LEI	XLY	[0.005875543503329326, 15.817608096174878, -13...	-0.105958	40.7723	NaN
3	LEI	XLI	[-0.0016062668778153131, 26.692047383550317, 2...	-0.19269	16.4004	NaN
2	LEI	XLU	[-0.006053796504675176, 4.862813071743094, -0....	-0.297895	20.8688	NaN
44	ALL	XLP	[-9.192149198207807e-05, -8.67437452331971e-05...	-0.430988	0.00442911	NaN
1	LEI	XLF	[-0.0002365366503321703, 4.351429897603006, 25...	-0.587755	2.1849	NaN
0	LEI	XLE	[0.011181975409731123, -49.51658719476371, -48...	-1.06844	65.3343	NaN
10	LEI	XOP	[0.04778811373272929, -212.7529241459583, -169...	-1.45582	1788.56	NaN
9	LEI	XLB	[-0.0014544605679136901, 11.510225833081316, 1...	-2.55183	18.4086	NaN

Fig2. Lasso Regression Table

	level_0	Symbols	coefs	score	mse	nb_predictors	nmse
16	CEI	GDX	[-0.00012632721134356605, -2.0605257635791854,...	0.310195	9.32997	NaN	NaN
3	TREE	XLI	NaN	0.242737	NaN	NaN	-0.0189236
5	LEI	XLK	[0.003889823168008847, 12.617256249695256, -12...	0.0948144	30.7937	NaN	NaN
42	ALL	XLV	[-3.0364915631986805e-05, -3.1120269706653044e...	0.0129777	0.00764618	NaN	NaN
11	LEI	SPY	[0.0025647436066845862, 32.82275711591052, 6.7...	-0.0234417	177.41	NaN	NaN
7	LEI	XLY	[0.005875543503329326, 15.817608096174878, -13...	-0.105958	40.7723	NaN	NaN
2	LEI	XLU	[-0.006053796504675176, 4.862813071743094, -0....	-0.297895	20.8688	NaN	NaN
44	ALL	XLP	[-9.192149198207807e-05, -8.67437452331971e-05...	-0.430988	0.00442911	NaN	NaN
1	LEI	XLF	[-0.0002365366503321703, 4.351429897603006, 25...	-0.587755	2.1849	NaN	NaN
9	TREE	XLB	NaN	-0.759344	NaN	NaN	-0.0147625
0	LEI	XLE	[0.011181975409731123, -49.51658719476371, -48...	-1.06844	65.3343	NaN	NaN
10	LEI	XOP	[0.04778811373272929, -212.7529241459583, -169...	-1.45582	1788.56	NaN	NaN

Fig3. Regression Tree Table