**Penalized Regressions and Predicting Silver Prices**

**Susan Wiggins**

**Cases in Business Analytics BAN 525**

**Professor: Dr. Cetin Ciner**



**Table of Content:**

Introduction………………………………………………………………………………………………………………………………………3-4

Analysis and Model Comparison……………………………………………………………………………………………………….4-7

Interpretation……………………………………………………………………………………………………………………………………7-9

References………………………………………………………………………………………………………………………………………..10

Appendix A **(Ordinary Linear Regression)**…………………………………………………………………………………………11-13

Appendix B **(General Regression Lasso)**……………………………………………………………………………………………13-14

Appendix C **(General Regression Adaptive Lasso)**…………………………………………………………………………….14

Appendix D **(General Regression Elastic Net)**……………………………………………………………………………………15

Appendix E **(General Regression Adaptive Elastic Net)**…………………………………………………………………….16

Appendix F **(General Regression Adaptive Lasso with Cauchy)**………………………………………………………..17

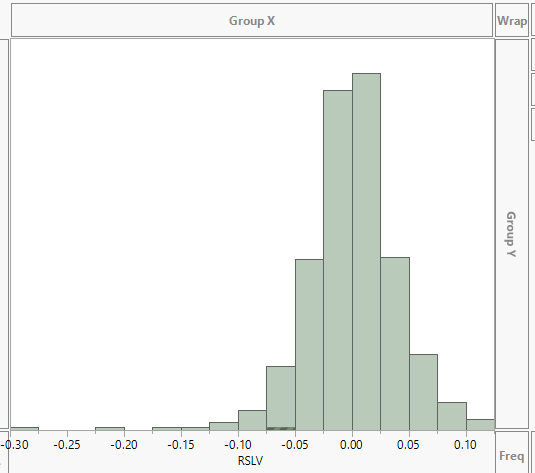
Appendix G **(Comparison Model)**……………………………………………………………………………………………………..18

Appendix H **(Elastic Net Final Profiler Results)**…………………………………………………………………………………18-20

Appendix I **(General Summary Variable Importance Final)**………………………………………………………………21-23

**Introduction**

Silver a precise metal is considered the work horse of the industrial market. Silver has made the following impacts: conductive properties making it useful in the computer industry, anti-bacterial properties making silver highly used in health care market, again conductive properties making useful in solar panels effecting the energy market, and of course the jewelry market. The question at hand is should you make an investment in silver based on using penalized regression for the prediction of silver prices. This case study of prediction of silver prices will cover a large number of financial market variables based on a pre-generated dataset and address the explanation of silver prices in a financial market as well as any lagged effects. The dataset consists of variables from currency, bond, and stock markets. The data is weekly and consisted in the time series of March 18, 2009 through March 6, 2019. The dataset was prepared based on the following criterion: raw prices were converted into percentages, a difference of a natural log was established, continuous compound returns, returns are noted with R labeling, and L denotes lagged instances. The dependent variable being measured in this study is silver prices (RSLV) and the predicator candidates are exchange rates, interest rates, stock market, market volatility, and inflation. The data for silver (RSLV) displays a normal distribution based in graph below.



Therefore, a log transformation is not applied within this analysis. Six methods will be addressed in this

Case: ordinary linear regression, Lasso, Adaptive Lasso, Elastic Net, Adaptive Elastic Net, and Adaptive

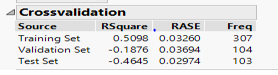
Lasso with Crauchy distribution. Ordinary linear regression will always be used as the benchmark in this

process; however, due to a large number of variables in the dataset random error can occur whereby causing problems such as high variance and poor forecast. Therefore, a focus on penalized regression has been studied in this case for predicting silver prices. Penalized regression ideally shrinks the variables to zero that have little significance in the analysis and we only see the most important variables in the model. Furthermore, variable selection is performed as well as estimation at the same time. Penalized Regression also know as general regression will be related to all the following modeling methods Lasso, Adaptive Lasso, Elastic Net, Adaptive Elastic Net, and Adaptive Lasso with Crauchy. In addition, lambda is a tuning parameter which considers the shrinkage. Cross Validation is used in our efforts of towards finalizing lambda for penalization.

**Analysis and Model Comparison**

The dataset employed in this case study is large and sometimes modeling can be affected when Big data is taken into account. Ordinary Least Squares (OLS) a modeling technique most highly used in prediction methods can cause problems when Big data is being used. For instance, OLS modeling can cause large variances in the data, models random noise, and poor forecasting can exists. Therefore, other machine learning techniques need to be implemented such as Penalized Regression. As noted previously penalized regression has a shrinkage effect on the variables whereby removing unimportant variables in our analysis. Lasso and Elastic Net are two modeling effects with same characteristics they both involve variable selection and perform estimation at the same time. On the other hand, adaptive lasso and adaptive elastic net are used more when there is an already noted implication of what variables might already have significance in the question being analyzed. These select variables would be penalized less based on this assumption. Dr. Cetin Ciner points out that Adaptive methods, “use initial estimates of OLS to weight the variables to give you an idea about which variables are more important” (27). All these methods introduced as penalized regression consider normal distribution as a default. However, the final modeling method Adaptive Lasso with Cauchy considers extreme values and might give better predications. Next, Cross validation is deployed in the predication technique for silver prices. Estimates are then built on the data after the completion of each modeling affect occurs. In the case predication of silver prices, the cross validation is established on 60/20/20 split of the data. Therefore, sixty percent of training data is used in estimation to row 309, twenty percent used for validation in order to stop the modeling process between row 310 and 413, and twenty percent used for testing for unbiased analysis in predication ability of the model from row 414 to end. The final interpretation of the model will be based on the results of the test data due to the basis of new observations.

Since the silver case dealt with time series data, the predictions are included at the end of the analysis making validation and test dead last in the column. JMP allows the user to set up the training, validation, and testing based on the labeling technique of 0, 1, 2. Zero is associated with training, ones are associated with validation, and twos are indicated by testing data. Next, modeling techniques of OLS, Lasso, Adaptive Lasso, Elastic Net, Adaptive Elastic Net, and Adaptive Lasso with Cauchy are initiated in the process to discover silver prices. First, ordinary linear regression was performed on variables listed from RFXB to LRIYR while removing the RSLV dependent variable from the list. The results for the OLS are represented in the following chart.

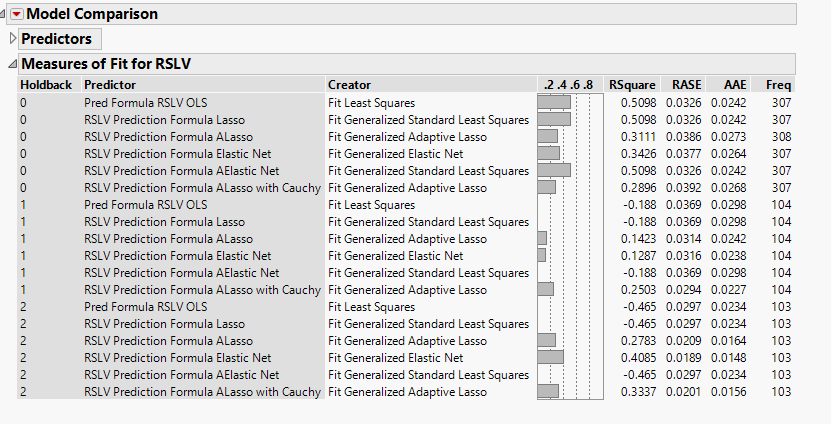


The testing data reveals that -46.45% of the data can be interpreted from the model based on the RSquare. The root average square error (RASE) score is also low at .02974 for OLS. However, the OLS model performs negatively for RSquare which indicates this is not a good modeling technique for predication of silver prices. RSquare value needs to be closer to positive 1.0 for a good interpretation of model.

Next, the following penalized modeling methods were performed Lasso, Adaptive Lasso, Elastic Net, Adaptive Elastic Net, and Adaptive Lasso with Cauchy. I found that all these modeling methods where not identical to the OLS due to the Lambda training value not being zero. All Lambda values where positive in these modeling methods indicating something was affected using these models. The following table indicates the lambda values reported in this case study for predication of silver prices.

|  |  |
| --- | --- |
| **Name Model Method** | **Training Lambda** |
| General Regression Lasso | 0.0663111 |
| General Regression Adaptive Lasso | 0.011254 |
| General Regression Elastic Net | 0.0113677 |
| General Regression Adaptive Elastic Net | 0.0113677 |
| General Regression Adaptive Lasso with Cauchy | 6.129274 |

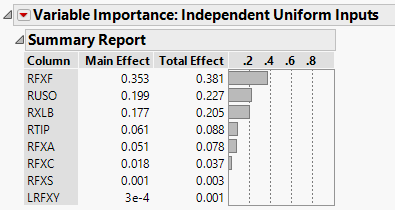
In addition, the results of each one of these models are listed in the appendix. A model comparison was run for analysis of best results of modeling predications and chart is listed below.



Due to the factor that each one of the new modeling techniques brought new information earlier based on the lambda values being a positive value, I was able to compare the best modeling for selection of Elastic Net. The reason behind this selection is based on the comparison values of RSquare being the highest at .4085. The testing data in the holdback cross validation indicated by (2) reveals that 40.85% of the data can be interpreted from the model. Further evidence of Elastic Net selection is the lower score for root average square error (RASE) of .0189 and average absolute error (AAE) of .0148.

**Interpretation**

The parameter estimates for the Elastic Net model are listed as RFXF (Swiss Franc), RUSO (Oil commodity), RXLB (material sector index), RTIP (inflation measure), RFXA (Austrian Dollar), RFXC (Canadian dollar), RFXS (Swedish krona), and LRFXY (Japanese Yen). The exchange rate group are listed in the following in the results RFXF (Swiss Franc), RFXA (Austrian Dollar), RFXC (Canadian dollar), RFXS (Swedish krona), and LRFXY (Japanese Yen). The RXLB (material sector index) is a stock market group. Parameter estimates are listed in the appendix section. The Canadian dollar is a general importance towards commodities. Canada is a big commodity exporter so then it would show developments in commodities. The stock market groups of material sector are of definite interest due to these two categories having possible increases or decreases in the amount of oil demand. The Japanse Yen is of general importance in silver prices due to the huge impact of using this commodity in production of products for such industries as computer market and energy market. The marker for Chi-square in the Elastic Net indicates estimates were best for RXLB, RFXF, and RUSO indicated in red. Next, the variable importance is included in the predication of silver prices and the following table gives us the results below.



The total effect considers the variations in the data and the RFXF Swiss franc explains 38 % of the variations in the silver price. The most important variable to look at is the Swiss franc when looking at explanation of silver prices. The second variable to look at pertaining to silver prices is the oil commodity at 23% and the third variable of interest is the material sector with total effect at 21%. RTIP, RFXA, RFXC, RFXS, and LRFXY are all below 10% for variable importance. The model prediction profiler reveals that the all variables which were indicated important in the model have a positive relationship. As you increase the average of the RFXF (Swiss Franc), RUSO (Oil commodity), RXLB (material sector index), RTIP (inflation measure), RFXA (Austrian Dollar), RFXC (Canadian dollar), RFXS (Swedish krona), and LRFXY (Japanese Yen) the price of silver also increases.

In conclusion, the selected model to predict silver prices was Elastic Net. The highest variable for fluctuations in silver price is related to the Swiss franc which is part of the exchange rate group at 38 percent and has a positive relationship. Swiss franc shines in crisis because is safe currency. The second most important variable associated with silver price is the commodity oil and has a positive relationship on silver prices. The third variable studied with higher percentage ranking in the Elastic Net is the material sector taking into account 21% and has a positive relationship when the material sector rises so does silver price. Supply and demand for silver at a material standpoint as well as industrial can significantly change the price of silver for consumers. Silver stated earlier is used in a lot of different markets from computers, energy, healthcare, and manufacturing of jewelry. The testing data reveals that 36% of the data can be interpreted from the Elastic Net model based on the RSquare. In addition, if the variables for RFXF Swiss franc, RUSO oil commodity, and RXLB material sector rise then silver prices will also increase as well as the other variables listed in the summary report earlier.

**Reference**

Ciner, Cetin. August 2019. Module 1. “Cross Validation and Predicting Gold Prices.” Slides 1-53. Retrieved from: <https://uncw.instructure.com/courses/24252/pages/module-1-lecture-materials-cross-validation-stepwise-regression-and-risk-factors-for-gold-prices?module_item_id=524495>

Ciner, Cetin. September 2019. Module 2. “Penalized Regression and Bank Revenues.” Slides 1-54. Retrieved from: <https://uncw.instructure.com/courses/24252/pages/module-2-lecture-materials-penalized-regressions-cross-validation-and-bank-revenues?module_item_id=524501>

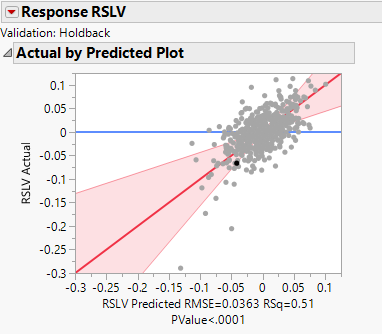
Lear Capital Editor. “4 Fascinating Industrial Uses for Silver.” Lear Capital the Precious Metals Leader. 17 Oct. 2016. Web. 7 Sept. 2019. <https://www.learcapital.com/news-blog/4-industrial-uses-for-silver/>

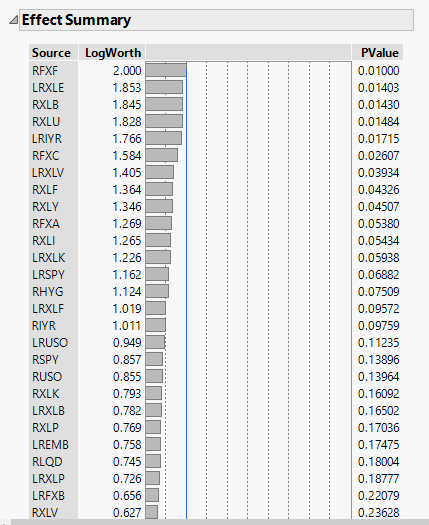
Pines, Lawrence. “4 Reasons you Might Invest in Silver.” Commodity.com. Web. 7 Sept. 2019. <https://commodity.com/precious-metals/silver/trading/>

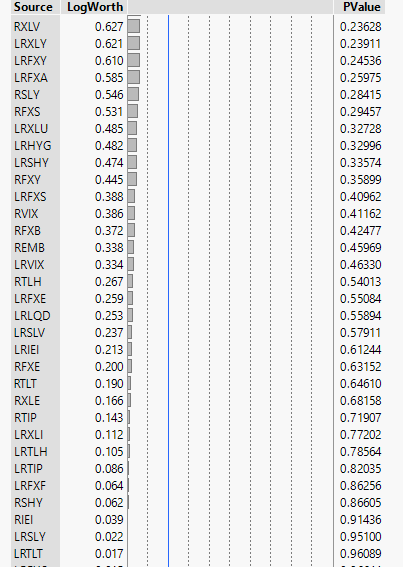
Web. Picture.<https://www.google.com/search?biw=1366&bih=657&tbm=isch&sa=1&ei=TGlxXfbBMM-d5wKHw4mAAQ&q=silver+coins+&oq=silver+coins+&gs_l=img.12..0i67j0l9.18949.18949..20290...0.0..0.88.88.1......0....1..gws-wiz-img.FBBXrndEvRY&ved=0ahUKEwj2l4CovLrkAhXPzlkKHYdhAhAQ4dUDCAY#imgrc=-U6EohF-t8_aXM:>

**Appendix A**

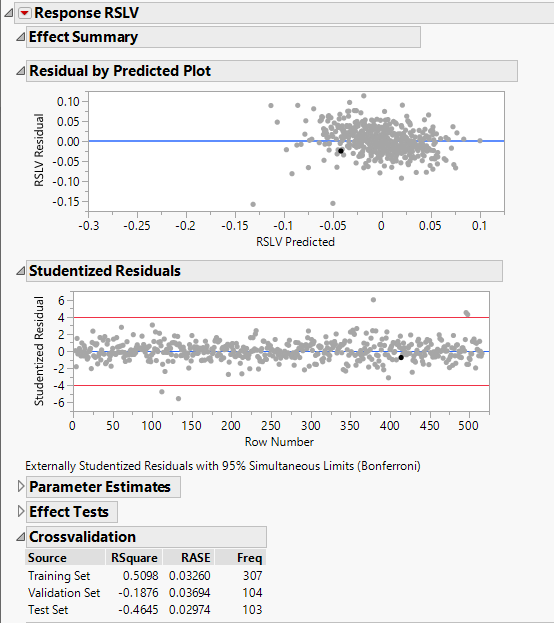
**Ordinary Linear Regression**





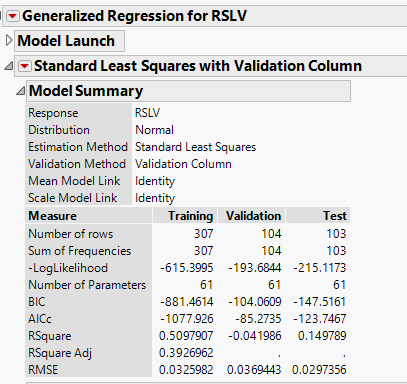
 

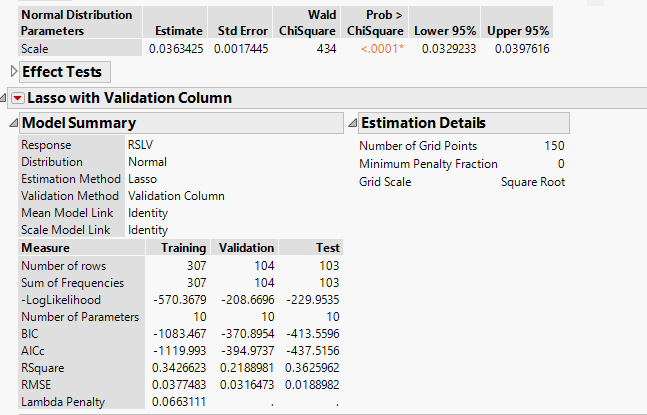




**Appendix B**

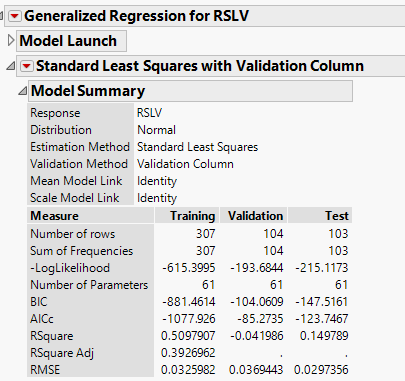
**General Regression Lasso**

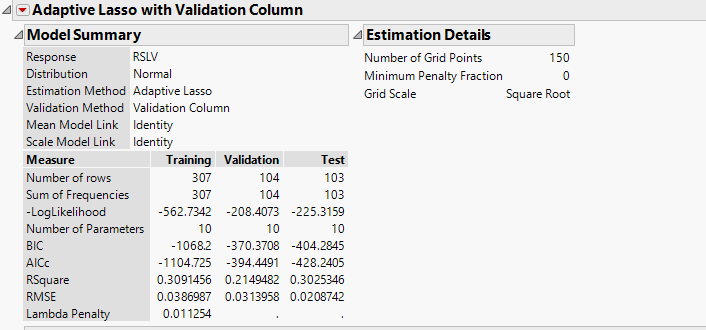




**Appendix C**

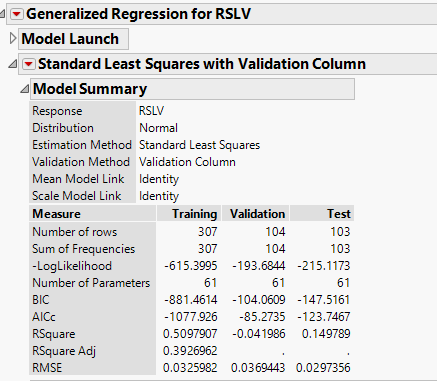
**General Regression Adaptive Lasso**

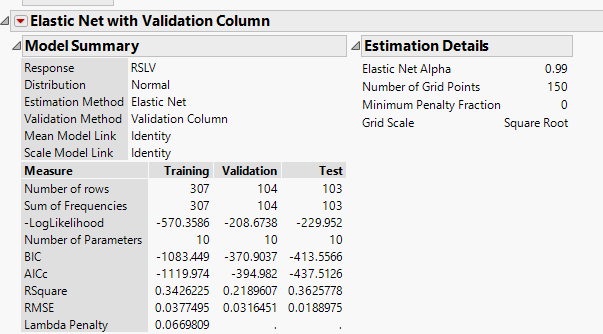




**Appendix D**

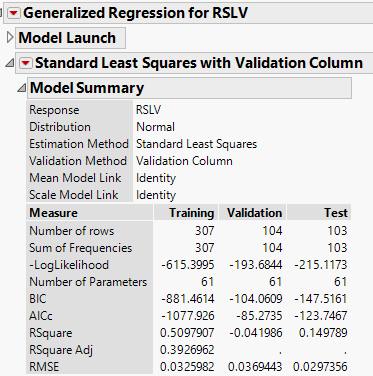
**General Regression Elastic Net**

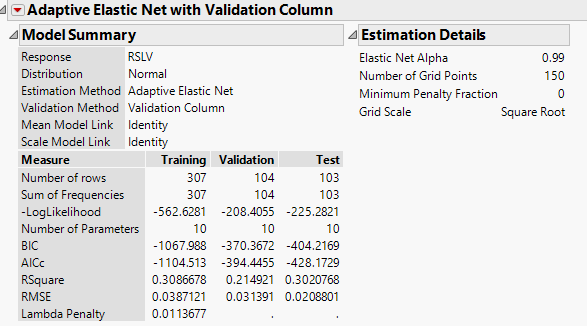




**Appendix E**

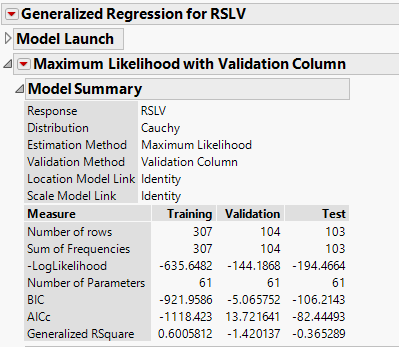
**General Regression Adaptive Elastic Net**

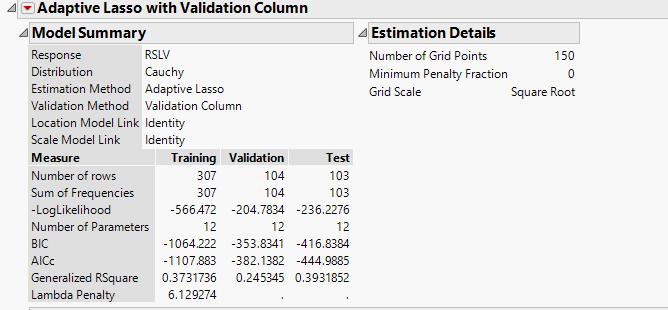




**Appendix F**

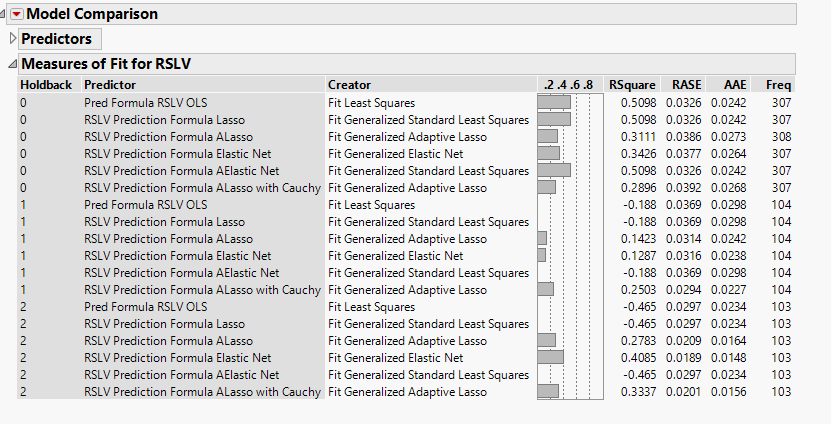
**General Regression Adaptive Lasso with Cauchy**





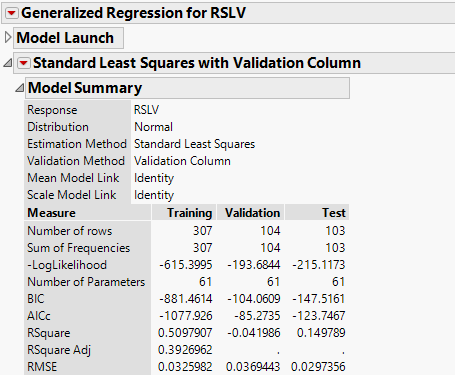
**Appendix G**

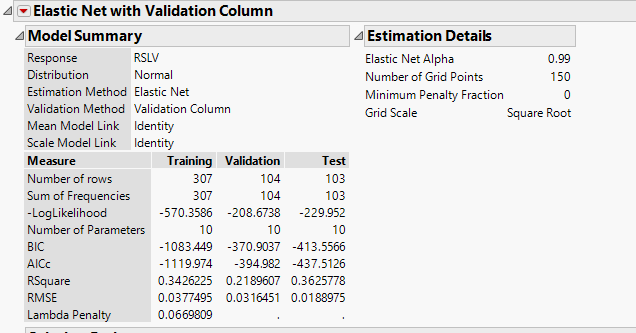
**Model Comparison**

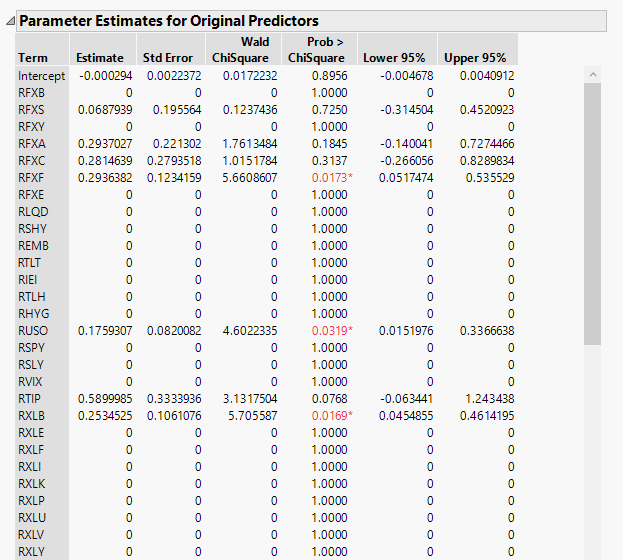


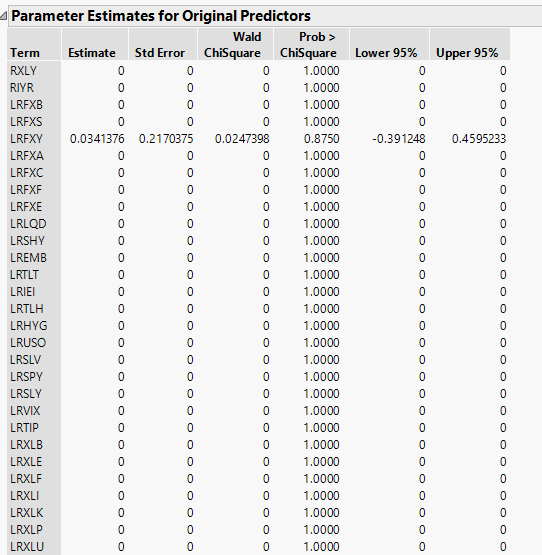
**Appendix H**

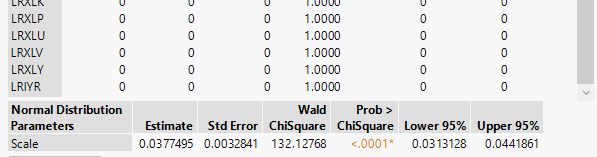
**Elastic Net Final Profiler Results**

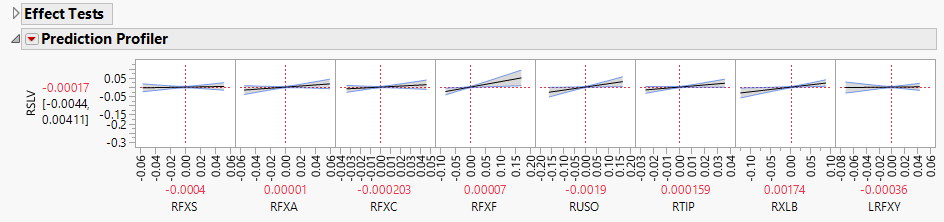












**Appendix I**

**General Summary Variable Importance Final**

