

Reproducing a Study of Stochastic Volatility + Market Inefficiency

>>> Fun with Multi Linear Regression

Maks Pazuniak

Original Study:

FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

Examining the Sources of Excess Return Predictability: Stochastic Volatility or Market Inefficiency?

Kevin J. Lansing

Federal Reserve Bank of San Francisco

Stephen F. LeRoy

University of California, Santa Barbara

Jun Ma

Northeastern University

December 2018

Working Paper 2018-14

<https://www.frbsf.org/economic-research/publications/working-papers/2018/14/>

Stochastic Volatility

sto·chas·tic

/stə'kastik/

adjective

TECHNICAL

1. randomly determined; having a random probability distribution or pattern that may be analyzed statistically but may not be predicted precisely.

[the volatility of asset prices is not
constant]

The efficient-market hypothesis is a theory that asset prices fully reflect all available information.

A direct implication is that it is impossible to "beat the market" consistently on a risk-adjusted basis since market prices should only react to new information.

Findings:

>>> Not Investment
Advice

“

We show that the sentiment-momentum variable is positively correlated with fluctuations in Google searches for the term “stock market,” suggesting that the sentiment-momentum variable helps to predict excess returns because it captures shifts in investor attention, particularly during stock market declines.

”

Data Sources:

Variance Risk Premium: <https://sites.google.com/site/haozhouspersonalhomepage>

EOM Nominal S&P, Nominal Dividends / Nominal Risk Free Rate: <http://www.hec.unil.ch/agoyal/>

University of Michigan Consumer Sentiment: <http://www.sca.isr.umich.edu/>

Google Trends: <https://trends.google.com/trends/?geo=US>

Quandl API: <https://www.quandl.com/>

Yahoo Finance: <https://finance.yahoo.com/>

Study Covers 1990-03 through 2017-12

Key Terms & Independent Variables:

index: S&P 500

price-dividend ratio index closing value / cumulative nominal dividends >> " **pd** "

fed funds rate delta 12 month change in federal funds rate >> " **ff12_D** "

variance risk premium 3 month moving average in difference between implied volatility from option on the index and realized volatility of the index >> " **vrp3** "

fed funds rate delta 12 month change in federal funds rate >> " **ff12_D** "

consumer sentiment delta 12 month change in UM Consumer Sentiment >> " **sent12_D** "

excess stock return delta 1 month change in excess return (over the risk free rate) - a measure of return momentum >> " **ersf_D** "

interaction consumer sentiment delta X excess stock return delta >> " **sent_x_ersf_D** "

google search term momentum change in 1 month change in volume of google searches for the term "Stock Market" >> " **Google_D2** "

Target Variable:

excess return in month t+1 12 month change in federal funds rate >> " **ersf_t1** "

Results from Data Cleaning

Scorecard

	Study	Reproduction
ersf mean	0.57	0.65
pd mean	52.4	52.1
vrp3 mean	16.2	17.3
ff12 delta mean	-0.29	-0.29
sentiment delta mean	0.18	0.19
ersf delta mean	0.002	0.004
sentD X ersfD mean	-3.57	-3.63

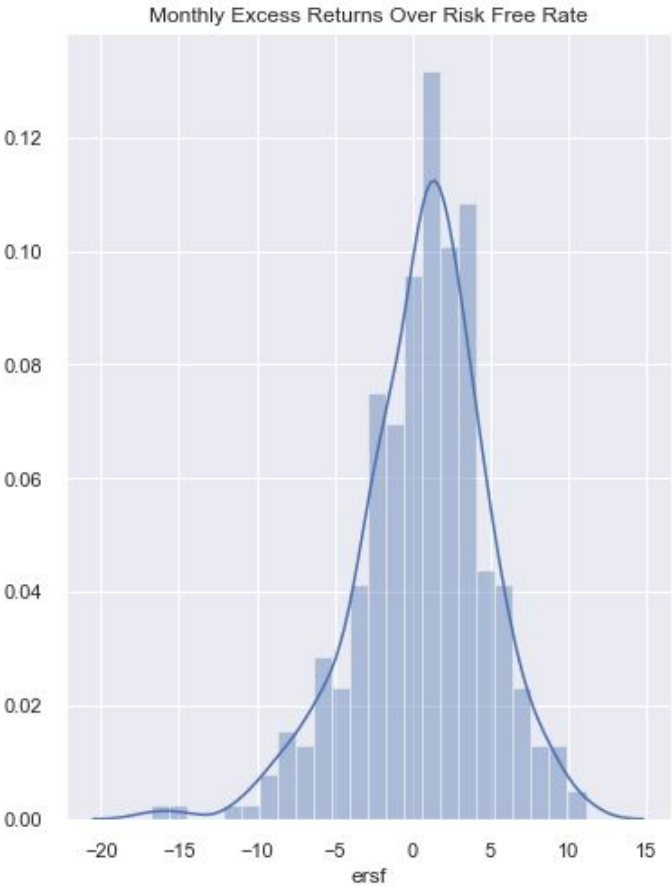
Results from initial data gathering and cleaning resulted in near-parity with the study, with the exception of variables related to the Excess Return;

Unclear where calculation differs at this time.

Mean: 0.65
Standard Deviation: 4.1

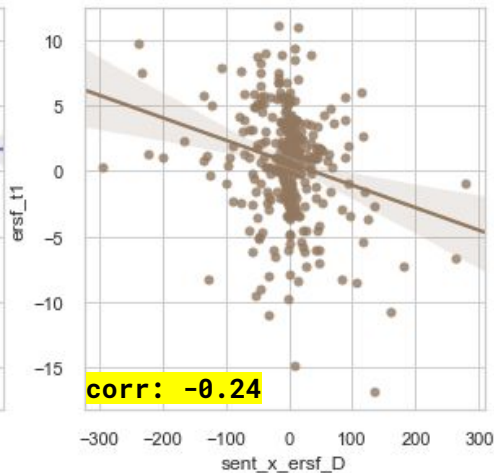
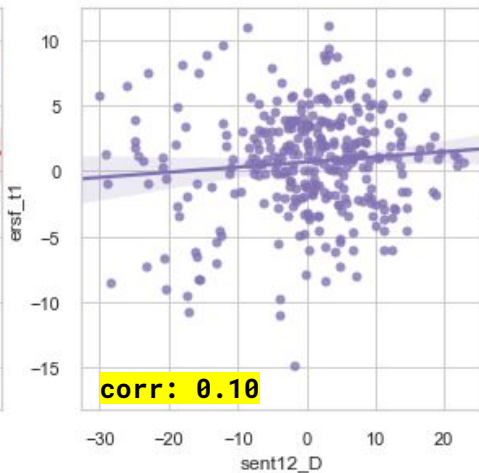
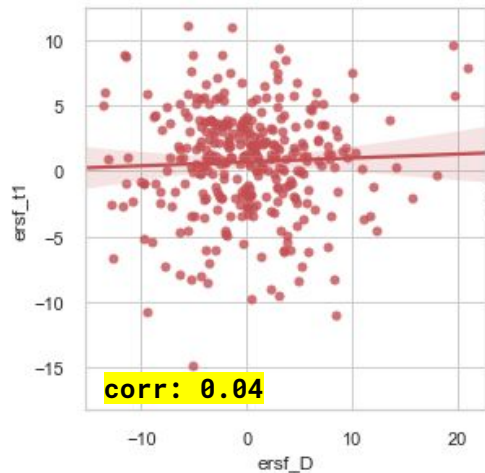
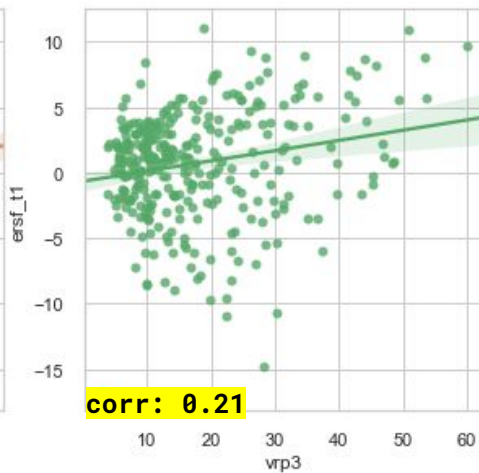
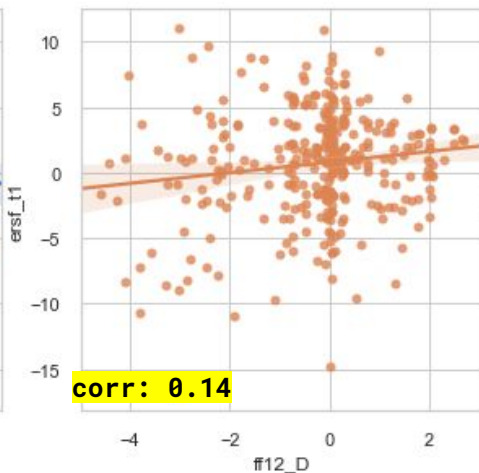
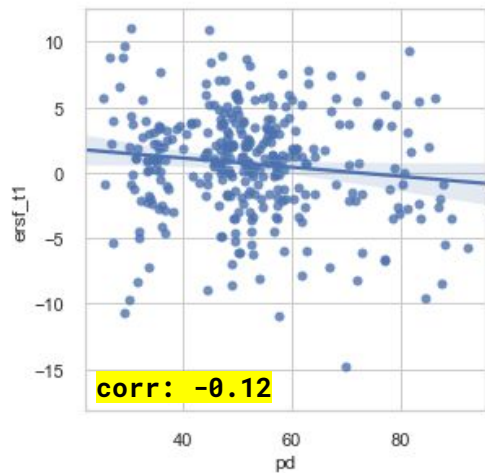
Minimum: -16.8%
Maximum: 10.9%

Distribution of Monthly Returns



Correlation Plots

target variable vs. independent variables



Baseline Model

price-dividend / federal funds rate / variance risk premium / sentiment change (12mo) X return momentum

OLS (Statsmodel)

Dep. Variable:	ersf_t1	R-squared:	0.177
Model:	OLS	Adj. R-squared:	0.167
Method:	Least Squares	F-statistic:	17.64
Date:	Tue, 23 Apr 2019	Prob (F-statistic):	3.96E-13
Time:	12:59:22	Log-Likelihood:	-910.22
No. Observations:	333	AIC:	1830
Df Residuals:	328	BIC:	1849
Df Model:	4		
Covariance Type:	nonrobust		

Original Study R-Squared: 17.3%

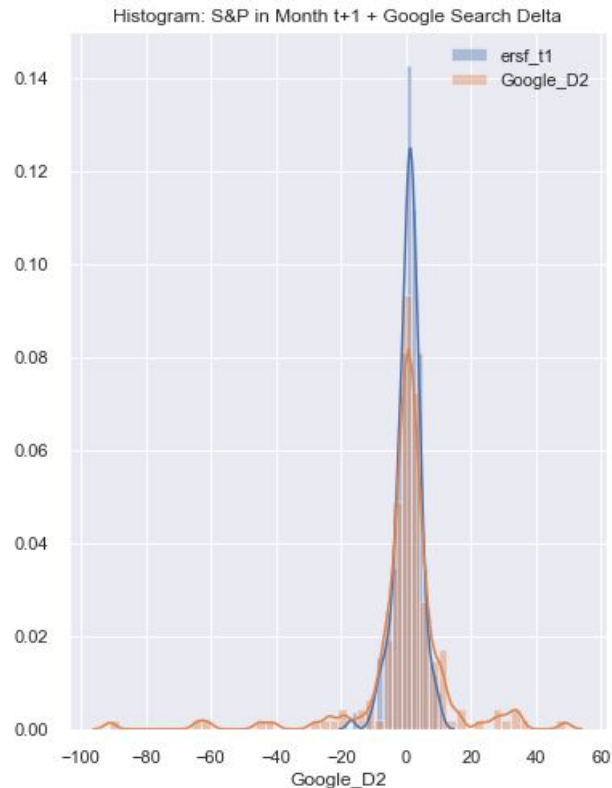
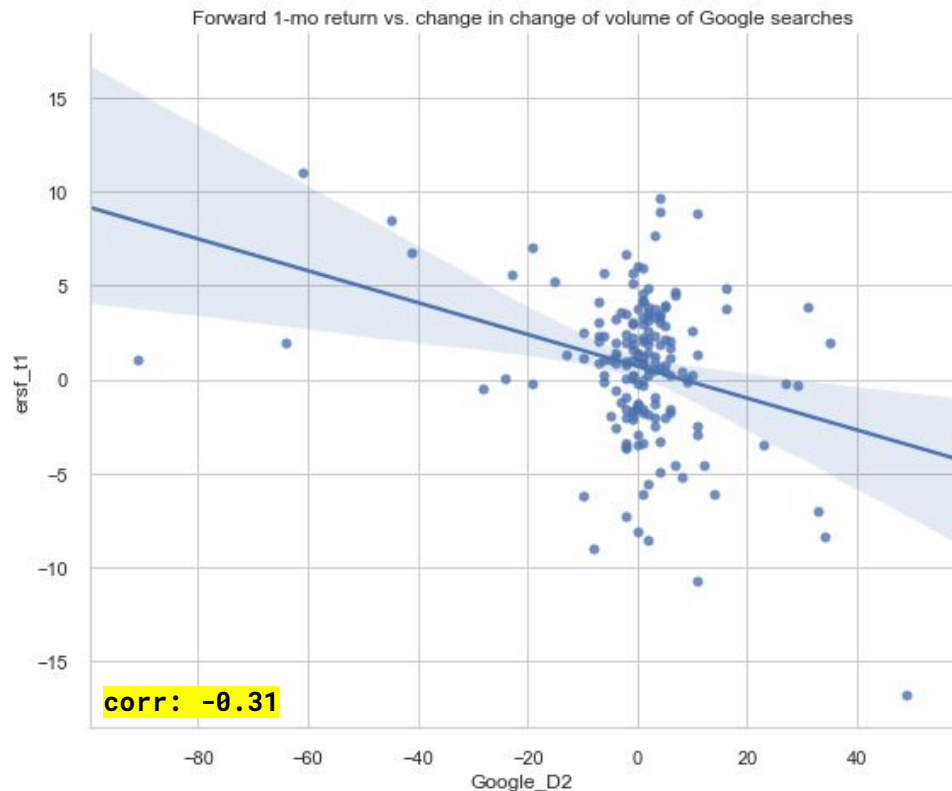
The original study included the ersf_D and sent12_D variables in their regression model;

Excluded here as p-values were fairly high, with minimal impact on the R-Squared value.

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1.995	0.831	2.402	0.017	0.361	3.630
pd	-0.059	0.015	-3.899	0.000	-0.089	-0.029
ff12_D	0.869	0.162	5.357	0.000	0.550	1.188
vrp3	0.112	0.020	5.494	0.000	0.072	0.153
sent_x_ersf_D	-0.014	0.004	-3.828	0.000	-0.021	-0.007
Omnibus:	30.347	Durbin-Watson:	2.003			
Prob(Omnibus):	0	Jarque-Bera (JB):	40.286			
Skew:	-0.663	Prob(JB):	1.79E-09			
Kurtosis:	4.071	Cond. No.				

Adding Change in Volume of Google Searches

google search term momentum change in 1 month change in volume of google searches for the term
"Stock Market" >> " Google_D2 "



Final Model

price-dividend / federal funds rate / variance risk premium / sentiment change (12mo) X return momentum / Google Delta

OLS (Statsmodel)

Dep. Variable:	ersf_t1	R-squared:	0.346	Original Study R-Squared: 34.2% The original study included the ersf_D and sent12_D variables in their regression model; Excluded here as p-values were fairly high, with minimal impact on the R-Squared value.
Model:	OLS	Adj. R-squared:	0.325	
Method:	Least Squares	F-statistic:	17.12	
Date:	Tue, 23 Apr 2019	Prob (F-statistic):	1.42E-13	
Time:	12:59:25	Log-Likelihood:	-431.11	
No. Observations:	168	AIC:	874.2	
Df Residuals:	162	BIC:	893	
Df Model:	5			
Covariance Type:	nonrobust			

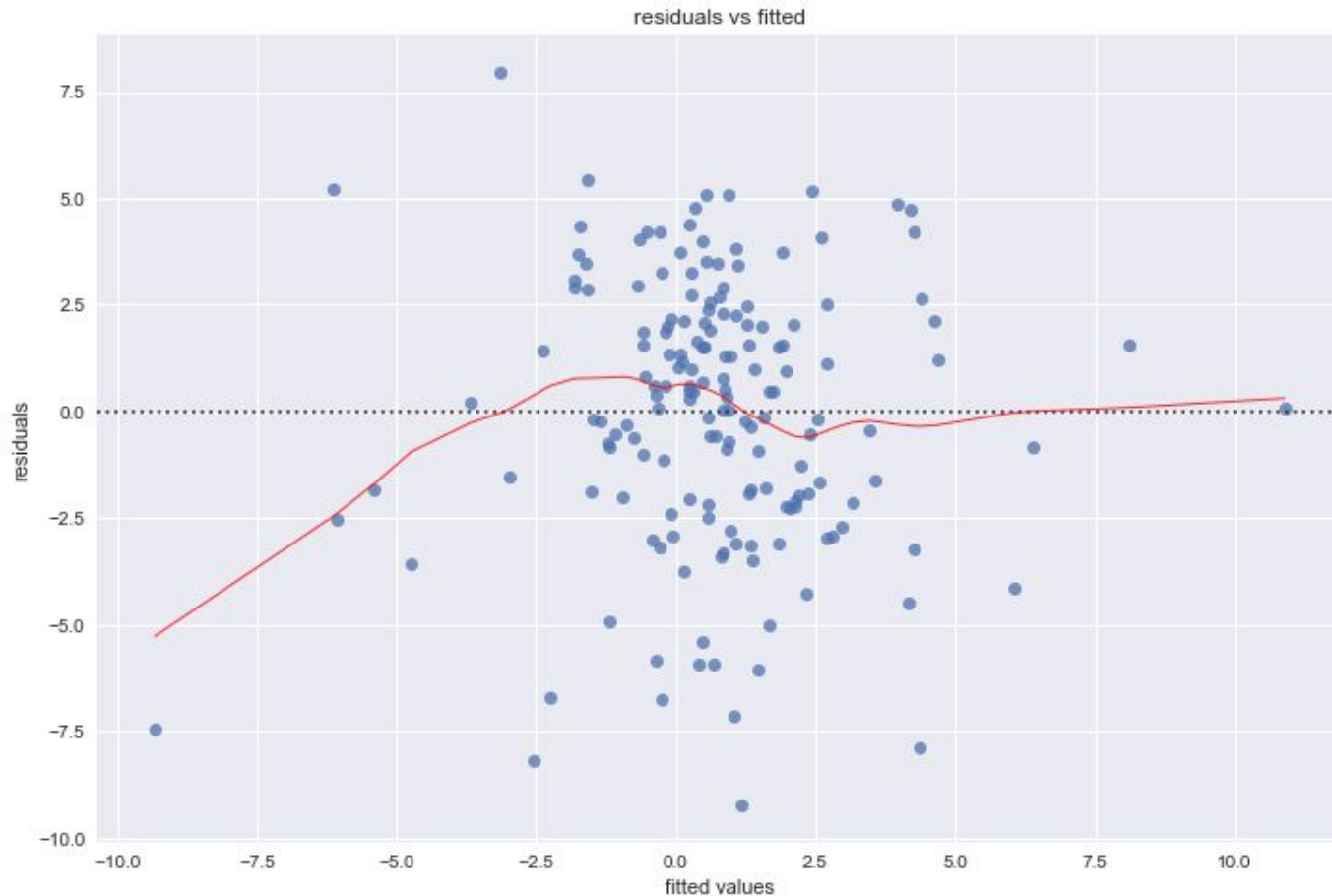
	coef	std err	t	P> t	[0.025	0.975]
Intercept	6.0135	3.044	1.976	0.0500	0.003	12.024
pd	-0.1419	0.057	-2.509	0.0130	-0.254	-0.03
ff12_D	1.6372	0.28	5.85	0.0000	1.085	2.19
vrp3	0.1383	0.029	4.76	0.0000	0.081	0.196
sent_x_ersf_D	-0.0094	0.004	-2.176	0.0310	-0.018	-0.001
Google_D2	-0.0742	0.017	-4.295	0.0000	-0.108	-0.04

Omnibus:	5.936	Durbin-Watson:	1.883
Prob(Omnibus):	0.051	Jarque-Bera (JB):	5.858
Skew:	-0.457	Prob(JB):	0.0534
Kurtosis:	3.025	Cond. No.	724

**Different
Timeframe:**

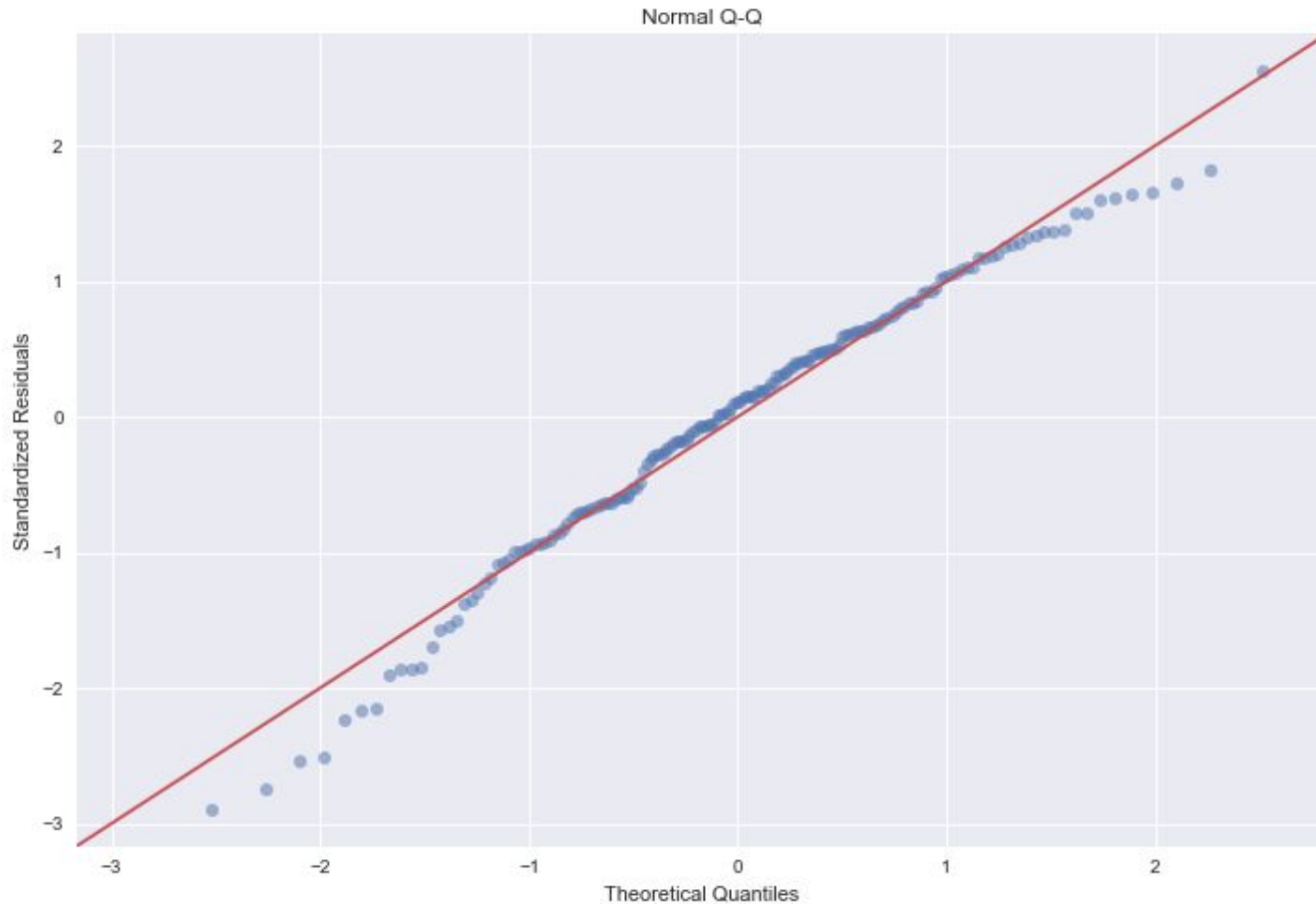
Google Trends
not available
until 2004.
Covers 2004-01
through 2017-12.

Final Model - Residuals vs. Fitted Values



The residual plots are fairly well distributed around the horizontal line, with no discernible pattern; a linear regression model appears to be appropriate.

Final Model - QQ Plot



The QQ Plot shows indicates that the residuals are fairly normally distributed, with some significant outliers towards the tails.

Predicting 2018 Forward Month Stock Returns

Running the Final Model on Untrained 2018 Data:

Applying a simple “Buy / Sell” Signal. If the Predicted Return is > 0 , Buy (or Hold), if Predicted Return is < 0 , Sell. Following this signal would have resulted in a -7.0% return on the year versus a -6.3% return with a pure “Buy and Hold” strategy. The amount show the results of investing \$10,000 on 12/31/2017, and following the monthly signal. While this is a useful heuristic for applying the model, the actual model Root Mean Squared Error was fairly high, at 5.6% (compared to 3.15% on the Trained Dataset).

month	pred_ret_next_mo	ersf	invested	correct	model	actual
2018-01	0.55	5.66	TRUE	TRUE	\$10,566	\$10,566
2018-02	-0.07	-3.88	FALSE	TRUE	\$10,566	\$10,155
2018-03	-2.23	-2.67	FALSE	TRUE	\$10,566	\$9,884
2018-04	6.68	0.28	TRUE	TRUE	\$10,595	\$9,912
2018-05	-2.66	2.16	FALSE	FALSE	\$10,595	\$10,126
2018-06	1.05	0.49	TRUE	TRUE	\$10,647	\$10,175
2018-07	-0.17	3.60	FALSE	FALSE	\$10,647	\$10,541
2018-08	0.92	3.01	TRUE	TRUE	\$10,967	\$10,858
2018-09	0.67	0.40	TRUE	TRUE	\$11,011	\$10,902
2018-10	0.96	-6.98	TRUE	FALSE	\$10,242	\$10,141
2018-11	-1.47	1.76	FALSE	FALSE	\$10,242	\$10,319
2018-12	4.24	-9.21	TRUE	FALSE	\$9,299	\$9,368
					-7.01%	-6.32%

Predicting 2018 Forward Month Stock Returns



The Model failed to predict some of the extreme market moves in 2018 and predicted some unusual, extreme moves as well.

Next Steps:

Incorporate Additional Variables:

- Volume Weighted Moving Average
- Twitter Sentiment
- Breadth - Advance / Decline
- Sentiment - Put / Call Ratios