

Do Past Returns Predict Future Results?

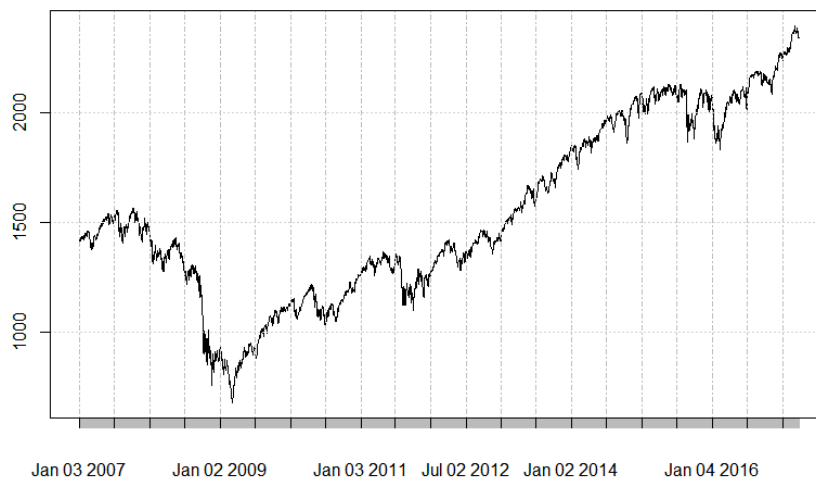
Do Past Returns Predict Future Results?

Overview

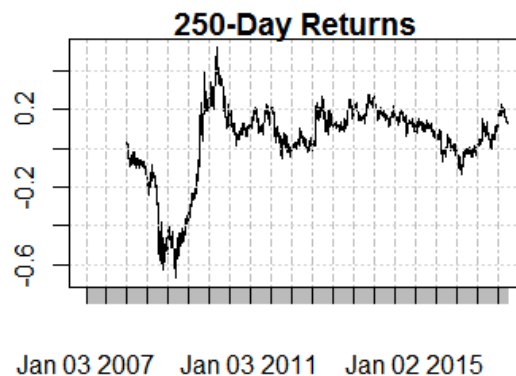
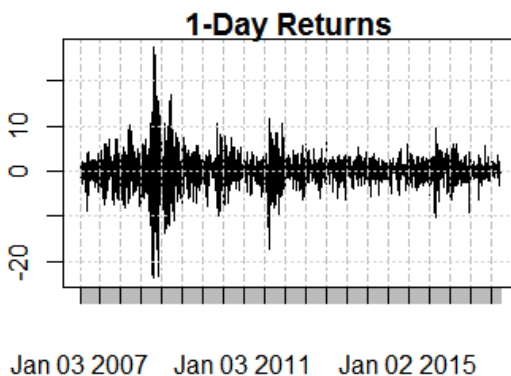
A fundamental question in stock analysis is how to estimate future returns. Modern Portfolio Theory assumes stock returns are normally distributed with a given mean and variance. The variance can be estimated using historical volatility. This has been shown to be largely persistent over time. The expected return is commonly estimated using historical returns, but is this appropriate?

Historical Returns for the S&P 500

The chart below shows the adjusted closing price of the S&P 500 for the last 10 years:



Returns on this investment look dramatically different based on the time the investment is made and the duration for which it is held. 1-day returns are near-zero and have extremely high variance. If a stock is held for a year (250 trading days) the returns range from 0-20% with much smaller variance except for an abnormal period during the crash of 2008.

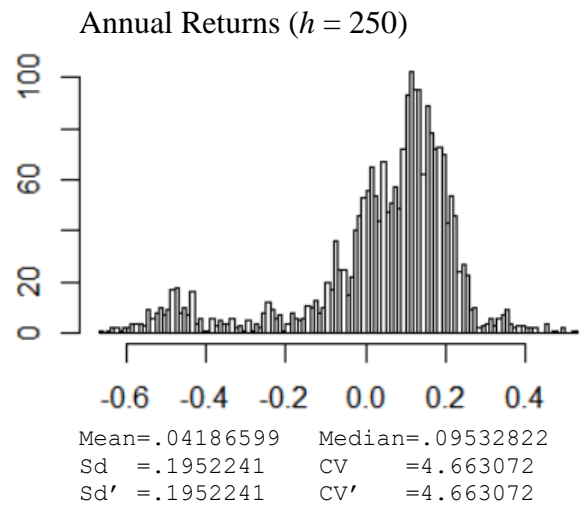
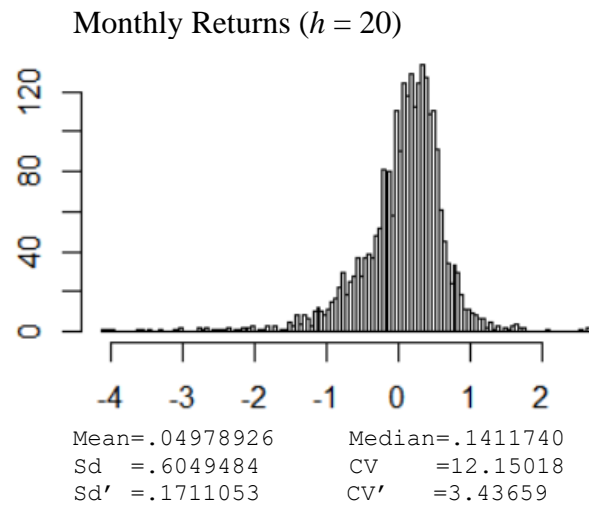
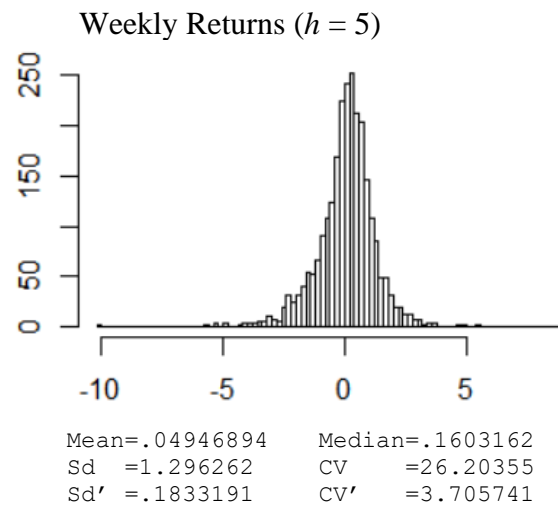
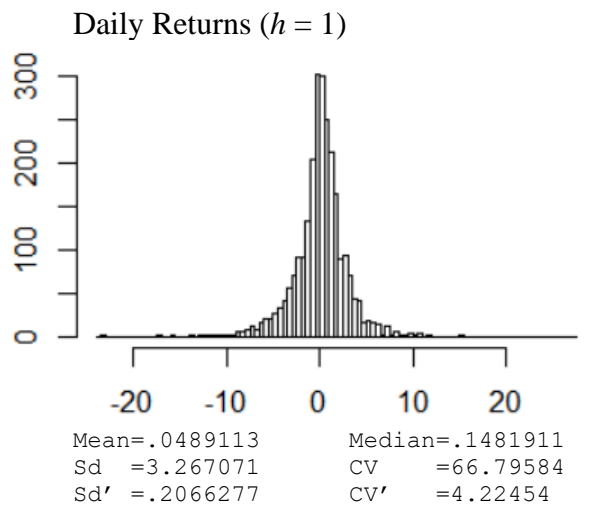


Do Past Returns Predict Future Results?

We can calculate the mean and standard deviation for different holding periods. The comparison will make more sense if we annualize the results so that they are on the same scale.

$$\begin{aligned} \text{Annualized (Log) Return over holding period } h & r = \left(\frac{250}{h}\right) \cdot \ln\left(\frac{S_{t+h}}{S_t}\right) \\ \text{Normalized Standard Deviation} & \sigma' = \sigma\sqrt{h/250} \end{aligned}$$

Regardless of the holding period, each of the charts shows a mean return ranging between 4.19% and 4.98% and a median that's 5-10% higher than the mean. The standard deviation of the daily returns is incredibly high. If we just pick a single day at random and make a 1-day investment, the standard deviation of returns is 3,267%. However, if we do this for each of the 250 days in a year, we would expect this to decrease to 20.66%. Normalized in this way, each of the holding periods has similar volatility ranging from 17.11 to 20.66%.

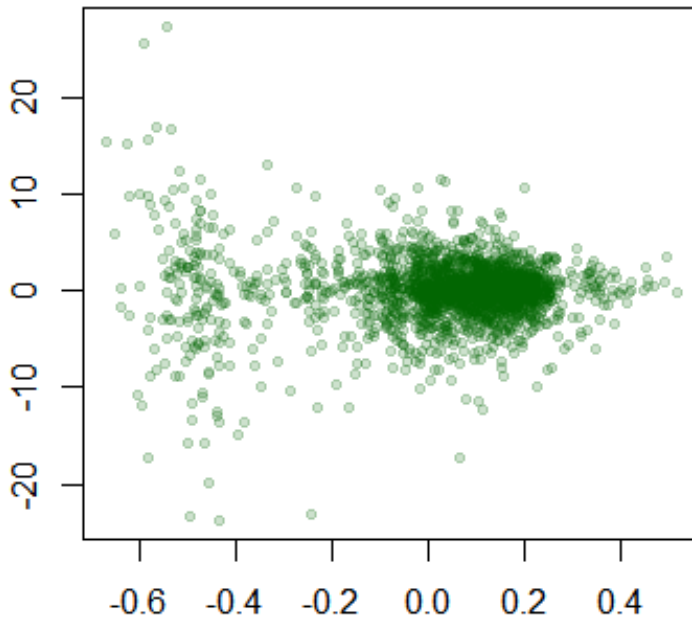


\ Indicates annualized values

Do Past Returns Predict Future Results?

Estimating Future Returns

Now we must ask: how do estimate future returns? Can we simply take the historical averages and use these as the expected values? Let's put that theory to the test. Let's try to estimate daily returns by average the last years' worth of daily returns. We do this for the last 5 years of data and have a hard time seeing any relationship in the data:



A linear regression produces the following coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.1492	0.0938	1.591	0.112
slope	-0.4285	0.6974	-0.614	0.539

Neither coefficient is statistically significant, and the R-squared is 0.0003027, showing just what we see in the plot: that the regression is not a good predictor. Also note that the model gave us a coefficient implying a negative relationship between past returns and future results. Really, the most significant part of this model is the intercept which tells us the average return is 14.92%. The intercept is calculated across all of our test data, meaning that we could not have estimated it without using out-of-sample data. We need to remove the intercept if we really want to understand how our historical average return works as a predictor of future returns. This results in the following model:

	Estimate	Std. Error	t value	Pr(> t)
slope	0.4527	0.4239	1.068	0.286

The estimate is still not significant, but at least it is a positive coefficient. This model implies that we should take the average daily return over the last 250 days and multiply it by .4527 to get an estimate for tomorrow's 1-day return. This is not exactly an inspiring result.

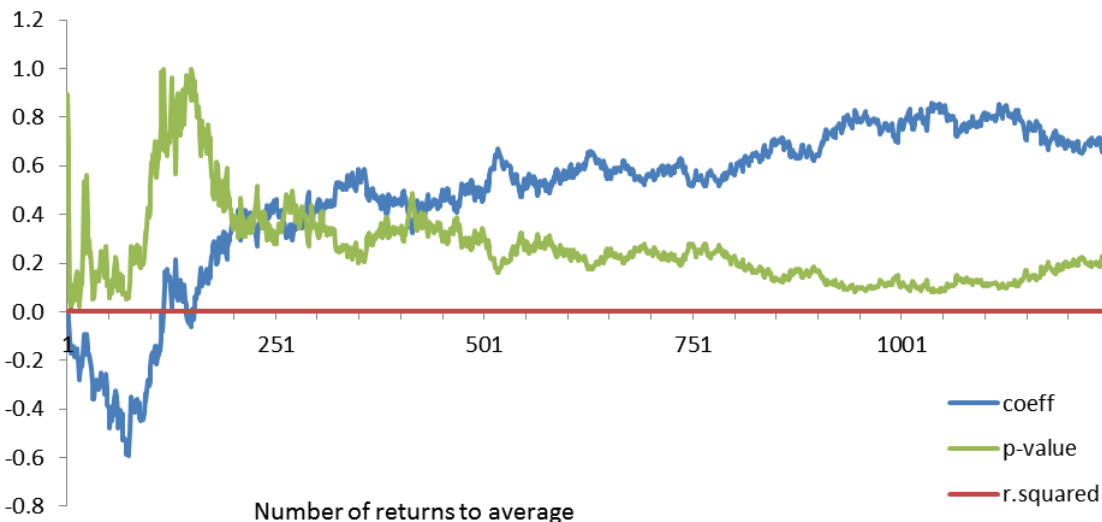
Do Past Returns Predict Future Results?

Is there a better way to average past returns, and does the model become more predictive for different holdings periods? Let's find out.

1-Day Returns

We performed the regression to predict 1-day returns using various moving averages of historical 1-day returns. The results are below. Notice that:

- The R squared is always near zero. None of these are good, predictive models.
- The p-value hovers most of the time around 10-30%. This would fail to be statistically different from zero under a 5% or 1% significance requirement.
- The coefficient actually begins negative. It reaches its most negative point at 72-74, meaning that if we average daily returns over 3-4 months, we should actually expect future performance to do the opposite. This is an interesting result because it implies that stocks do swing to unwarranted extremes in the short-term.
- After 1 year of data the coefficient is at 0.5 and it levels off around 0.8. This means we can take several years of daily stock returns and use them as a long-term estimate of future returns. However, the relationship is not 1-to-1. The model actually says we should assume future returns to be less than what we've seen in the past.

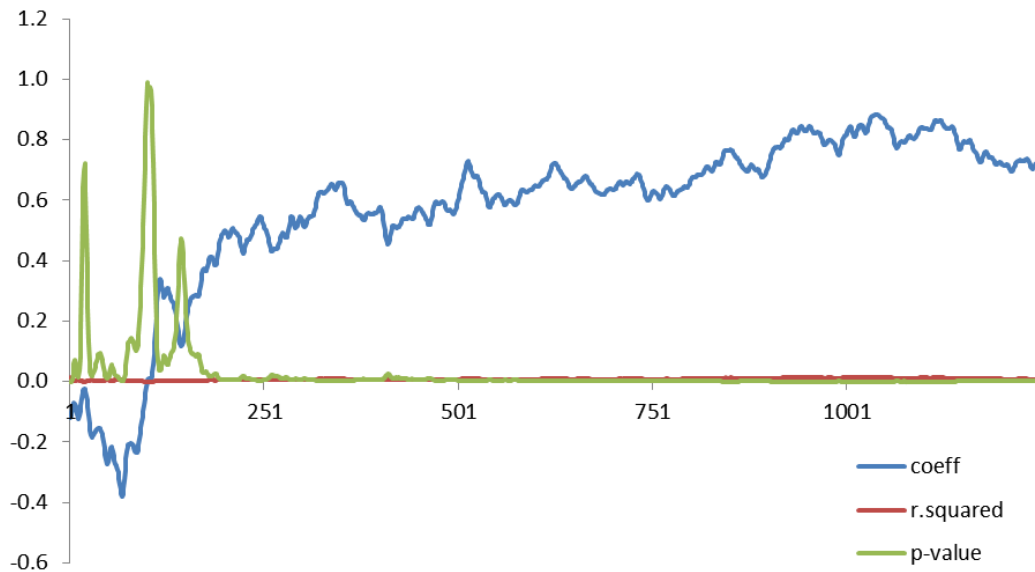


The results imply that we can use a long-term (2-5 years) average of daily returns to estimate future ones. We may also consider a contrarian strategy in the short-term, buying stocks that have fallen or selling stocks that have risen significantly in the past couple of months. None of these will be great predictors of success, and actual returns will still vary significantly on both the positive and negative side of our expectations.

Do Past Returns Predict Future Results?

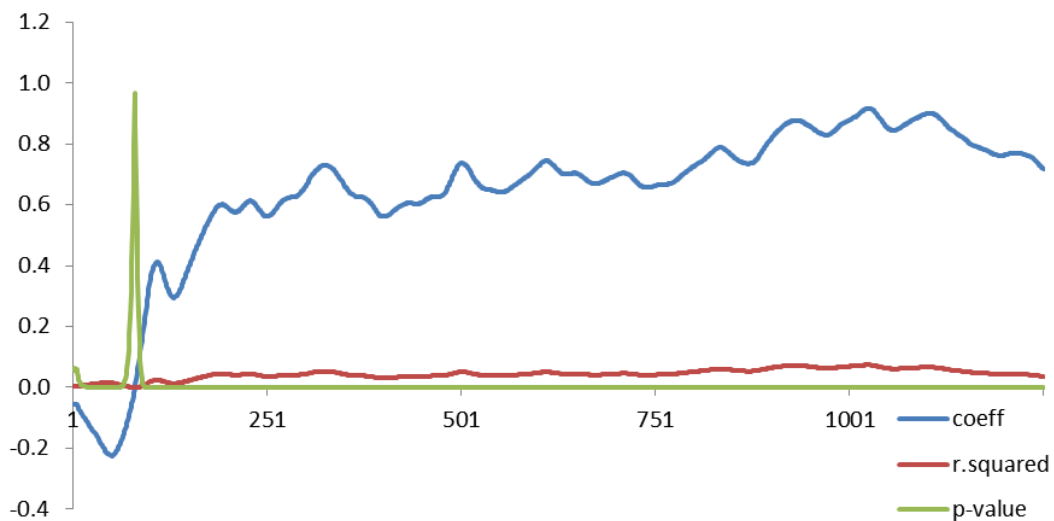
5-Day Returns

We do this again with 5-day returns, estimating how we would do if we left our money in the market for about 1 week. Results are similar in that we still see low R-squared, and a coefficient that switches from negative in the short term to positive in the long term. One big difference is that we now have an estimate with a low enough p-value to be statistically significant.



20-Day Returns

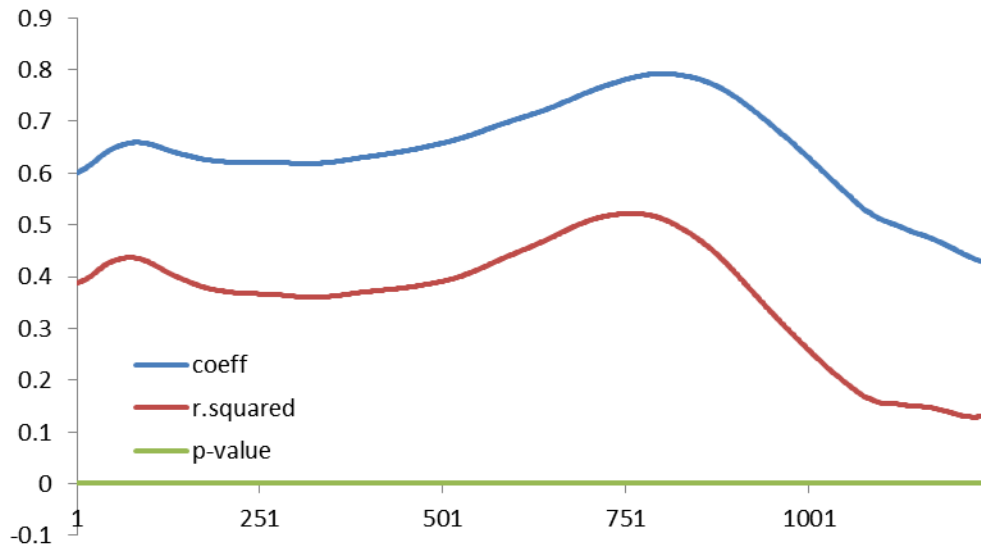
20-day returns show a similar pattern. However, the model is starting to become more predictive as the R-squared is beginning to differ from zero slightly.



Do Past Returns Predict Future Results?

Annual Returns (h=250)

Annual return data begins looking much different. We no longer see the negative short-term coefficient. We also see much higher levels of significance and R-squared values that are increasing to about 40%. Averaging 250-day returns over 3 years of data seems to be the best policy. Using more data actually decreases the coefficient in the estimate and the predictive power of the model.



Extrapolating Results

Extrapolating results from this analysis may be risky. Two general extrapolations do seem plausible:

- That there is a negative correlation between short-term historical and future-results
- That over the long-term there is a positive (albeit not 1-to-1) correlation between historical and future results

However, we should be careful about getting too detailed in our take-aways. As an example, the evidence that using 3-years of data as the best predictor for 250-day returns may be specific to this data set. There is a good chance that this was affected by our choice of time periods to analyze. The period from 2009 to 2015 was relatively stable in terms of market behavior. 2008 was a period of great loss. It may be that we are losing predictive power in the models by including data from those abnormal years in some of our models. It seems safer to conclude that the long-term average, even using more than 3 years of data, would be a good predictor, assuming that we remove abnormal periods of returns that do not seem pertinent to current market conditions. Of course, determining when the market moves from being “normal” to being “abnormal” can be even more important than making good short-term investments. Such “Black Swan” periods as 2008 or perhaps even 2017 can wipe away years of accumulated wealth in a matter of weeks or days.