Does oil volatility scale with increasing price?

James Hamski (james.hamski@spsmail.cuny.edu)

5/19/2017

1 Introduction

In a commodity trading market the price level is expected to be tied to the system dynamics, namely supply, demand, and delivery of the commodity being traded. Volatility, the variation in price over time, reflects uncertainty in the balance of these system factors. This research paper aims to answer a simply posed question: "does oil price volatility scale with price?"; i.e., can we expect to observe larger price swings when the price is near $100 per barrel vs $20 per barrel?

If oil price volatility reflects uncertainty about supply and demand dynamics, it isn't immediately clear whether we should expect volatility to depend on price level. Higher oil prices are associated with "tightness" in the supply market, meaning there is little excess capacity to increase production, therefore we may expect swings higher but some base price support that results in lower measured volatility. Likewise, low prices may suggest excess capacity that can buffer shocks to the oil delivery system, dampening volatility. Despite these "just so" arguments, a multitude of factors such as storage dynamics, supply chain distruptions, the ability of producers to increase production to bring more oil to market or shut in production capacity in response to prices ("rebalancing"), and market speculation complicate this picture and suggest it must be studied empirically.

If it is found that oil price volatility is dependent on price level, the relationship may follow a scaling formula. For instance, if we can expect volatility of $1/barrel when oil is at $20/barrel, can we expect volatility of $5/barrel at $100/barrel price levels via a simple linear scaling rule? Three methods are presented in this research paper to answer this question: (1) regression modeling of price and volatility, (2) viewing volatility within oil price regimes, and (3) using multivariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) modeling.

Note that in this research paper *oil price* is used to specifically mean spot-traded crude oil. This represents only one component of the oil markets, and most of the actual oil price is determined by futures and long term delivery contracts (need cite). This research paper is concerned with understanding the energy system using pricing information. In this way, it differs from much of the published research in that it is not concerned with forecasting prices or volatility. Nor is it addressing exogeneous system elements such as equity markets or interest rates, though the literature shows that the crude oil market and larger economic indicators are intertwined (add cite). Instead, it contributes to our understanding of the system dynamics of an essential energy commodity.

2 Exploratory Data Analysis

2.1 Data Source

The data source is the West Texas Intermediate (WTI) nominal (i.e. not inflation adjusted) daily spot price record from the U.S. Energy Information Administration. The WTI series was filtered to the date range January 2, 1986 through December 30, 2016.

Table 1: Oil price series date and price ranges.

Date Range	Price Range
Min. :1986-01-03	Min. : 10.25
1st Qu.:1993-09-01	1st Qu.: 19.38
Median :2001-06-11	Median : 28.01
Mean :2001-06-19	Mean : 42.87
3rd Qu.:2009-03-31	3rd Qu.: 63.47
Max. :2016-12-30	Max. :145.31

2.2 Returns and Volatility

In this research paper, volatility is characterized two ways: (1) 5-day historic volatility and (2) 30-day historic volatility. In addition, the relationship between the returns themselves and price level is investigated. Single-period returns were calculated as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

In some sections, their absolute values are used.

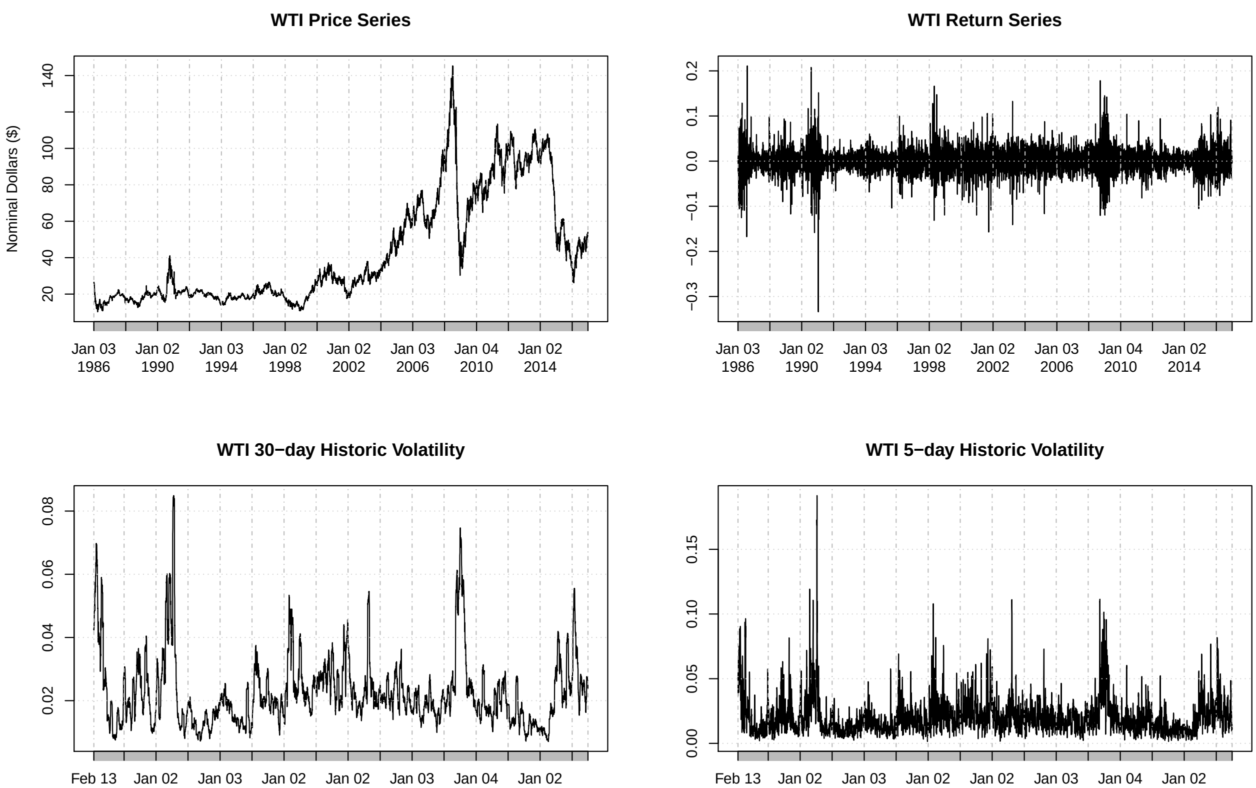
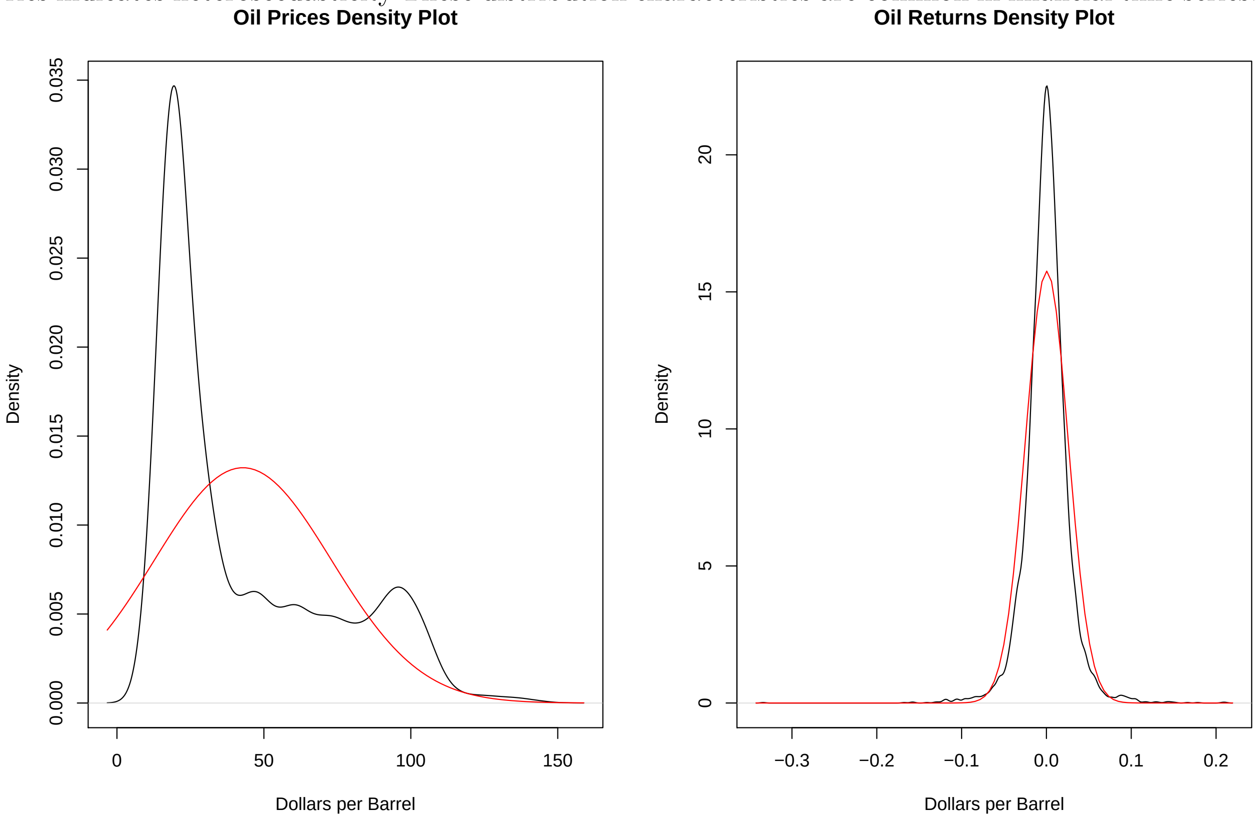


Figure 1: Price level, return, and volatility series plots of WTI spot oil prices.

A seen in Figure 1, most of the series from 1986 through 2004 contains prices between \$10/barrel and \$40/barrel. This results in a price series with left skew and a long right tail (Figure 2). The return series indicates heteroscedasticity These distribution characteristics are common in financial time series.



3 Times series exploration

A challenge in analyzing financial time series in general, and spot oil market prices specifically, is that the variance structure may be independent, but not identically distributed. Oil prices exhibit periods of low volatility (i.e. relatively constant prices) and periods of high volatility (i.e. changing prices). This is referred to as volatility clustering. This violates the assumption in the most frequently used time series model, the autoregressive integrated moving average (ARIMA) model.

4 Price-Volatility Regression Analysis

Relating price level and the measures of volatility at each time in the series is a simple exploration of the research problem. The covariance and correlation measures of vectors representing price versus returns, 30-day, and 5-day historic volatility indicate a weak, negative relationship.

Table 2: Covariance and correlation of price level and returns, 30-day, and 5-day historic volatility.

	Covariance	Correlation
return	-0.0396337	-0.0719290
30-day	-0.0640183	-0.1864200

	Covariance	Correlation
5-day	-0.0529310	-0.1212286

```
## numeric(0)
```

```
## numeric(0)
```

```
## numeric(0)
```

The case of 30-day historic volatility indicates a negative relationship between price level and volatility. However, this result appears to be due to a cluster high volatility around $20 per barrel, creating a leverage point. Residual analysis indicates that this is not a good relationship to model with linear regression. Residual analysis for the other volatility measures also display

As seen in Figure 1, nominal oil prices have spent time as high as $145 per barrel. However, the majority of the time series is far lower, with a median price of $28 per barrel. This means that the dataset is unbalanced and higher price levels represent a smaller portion of the dataset. In addition, oil price (and financial time series in general) exhibits volatility clustering. Therefore, it is anticipated that this simple regression model based on price and volatility is not the best possible solution to the question of characterizing the dependency of volatility on price.

5 Comparing Volatility across Price Regimes

Changepoint detection aims to detect the point or points where the statistical properties of a sequence of observations change (Killick and Eckley, 2014). Changepoint detection was used to detect change in the oil price mean throughout the period of record using the Pruned Exact Linear Time (PELT) algorithm (Killick et al. 2012). The time series between these changepoints represent "price regimes" (i.e. time series between changepoints) which have generally similar mean oil price compared to the entire record.

Changepoint detection proceeds by minimizing a cost function over possible locations and number of changepoints. The cost function:

```
## numeric(0)
```

```
##
## Call:
## lm(formula = sd ~ med, data = price.regime.descriptive)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.107 -1.135 -0.713  1.310  5.576
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.96662    1.51254   1.300  0.21611
## med          0.08285    0.02529   3.276  0.00602 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

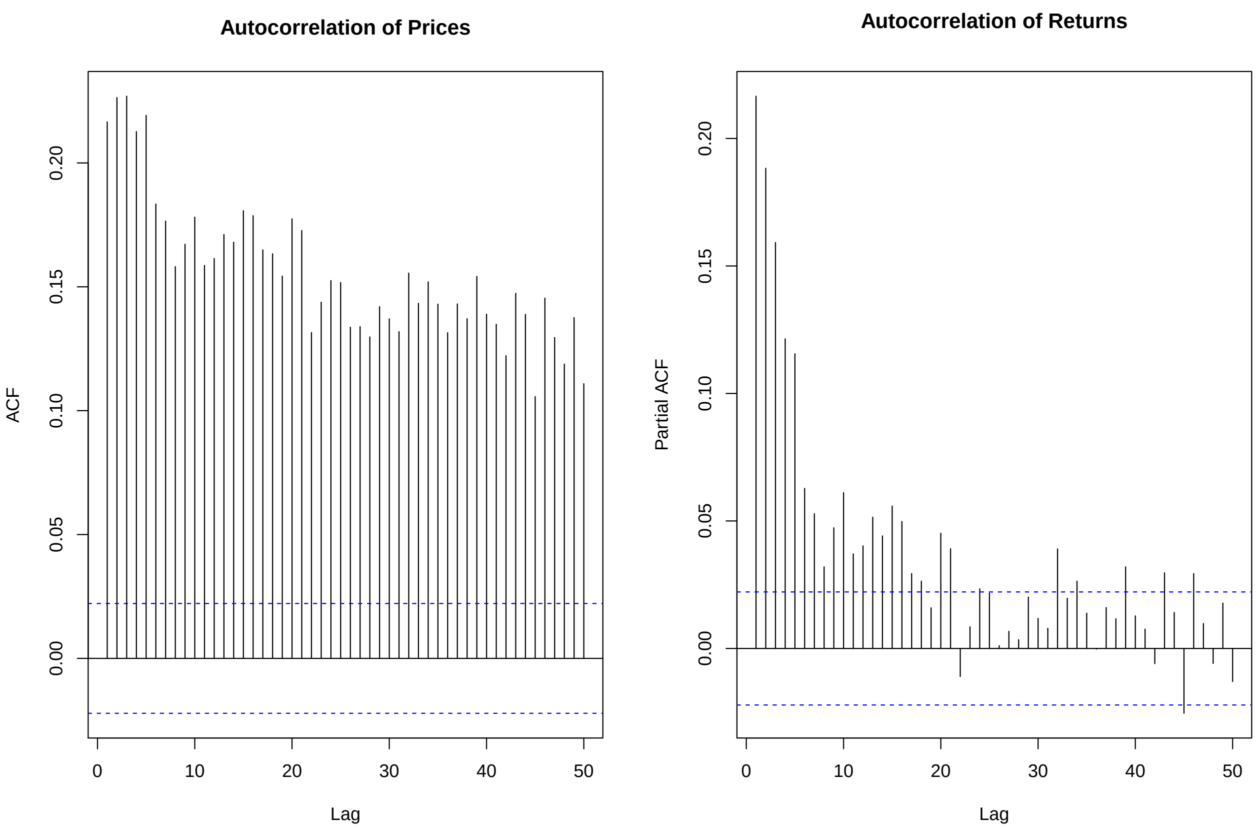


Figure 2: Autocorrelation plots for prices and returns.

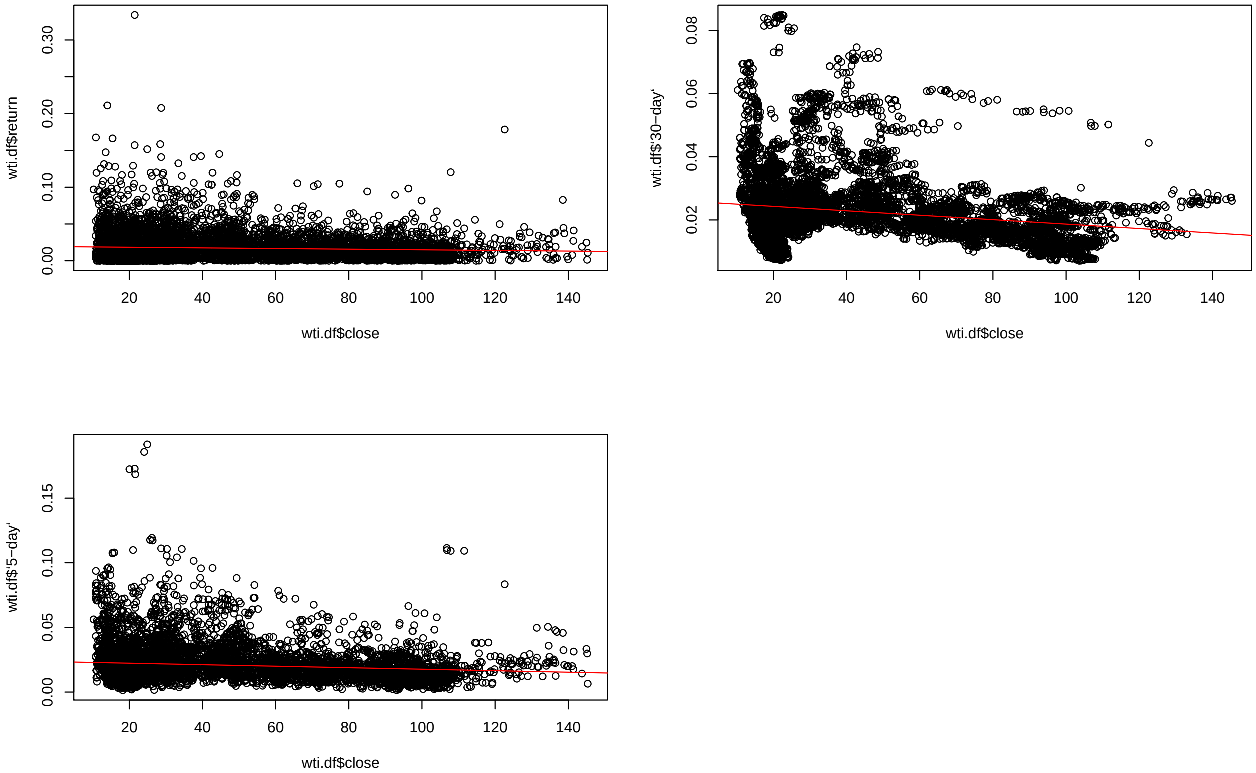


Figure 3: Scatterplots with a linear model relating spot oil price with returns, 30-day, and 5-day historic volatility.

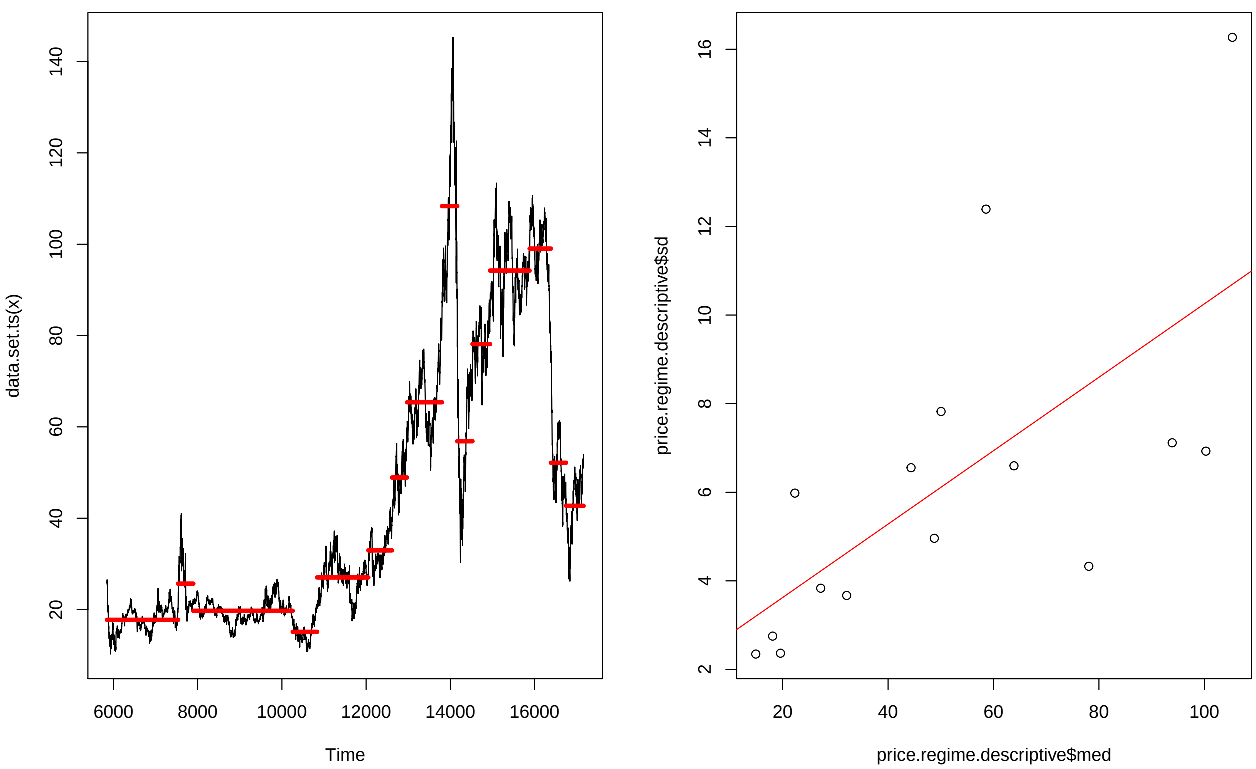


Figure 4: Some caption

```
##
## Residual standard error: 2.921 on 13 degrees of freedom
## Multiple R-squared:  0.4522, Adjusted R-squared:   0.41
## F-statistic: 10.73 on 1 and 13 DF,  p-value: 0.006024
```

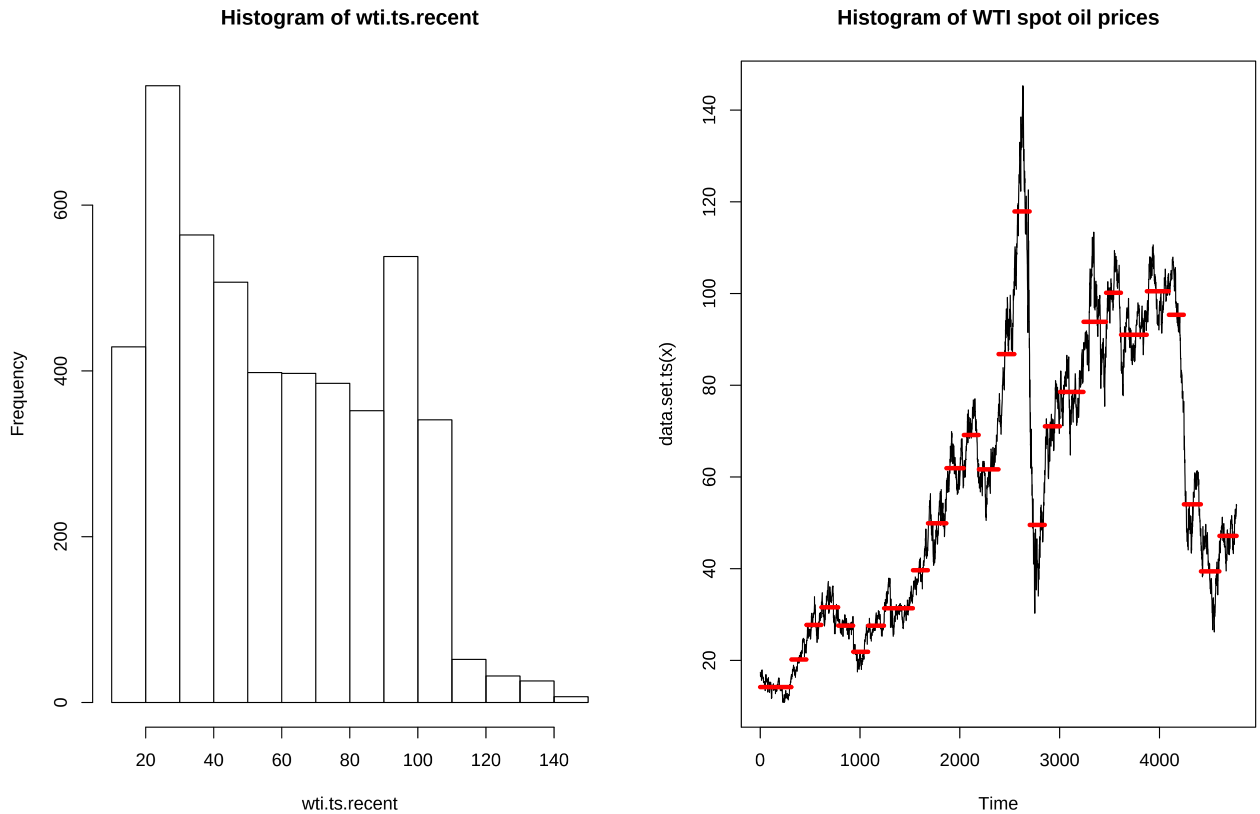


Figure 5: Histogram showing price distribution after limiting to 1998 through 2016.

```
## numeric(0)

##
## Call:
## lm(formula = sd ~ med, data = price.regime.descriptive.2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2593 -1.0648 -0.3374  0.9833  6.1756
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.64302    0.96084   0.669     0.51
## med          0.07285    0.01478   4.930 4.96e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.215 on 24 degrees of freedom
```

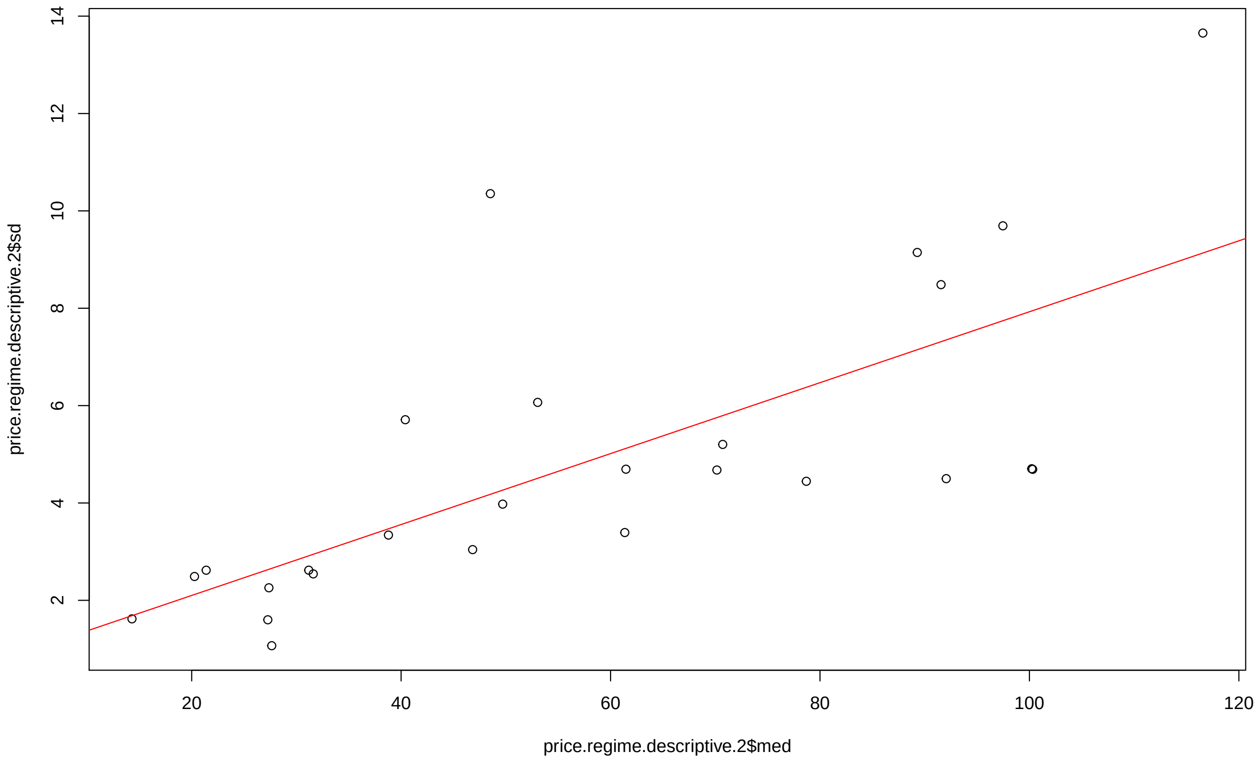


Figure 6: Price regimes and linear model after limiting to 1998 through 2016.

```
## Multiple R-squared:  0.5032, Adjusted R-squared:  0.4825
## F-statistic: 24.31 on 1 and 24 DF,  p-value: 4.961e-05
```

This section's analysis using changepoint analysis to break the oil price series into price regimes as defined by changes in the series mean is highly dependent upon model parameters. The PELT changepoint detection method identifies different changpoints depending on the minimum segment length and penalty parameters, and based on the section of the time series used. An interactive Shiny App was created and may be access in order to try different parameterizations of the changepoint model and view the effects on the relationship between medial regime price and standard deviation within that regime.

6 Multivariate GARCH Model

7 Acknowledgements

I would like to thank John Kemp, Thompson Reuters energy journalist, for posing the question investigated in this research paper. In addition, I thank Tancred Lidderdale and Mason Hamilton from the U.S. Energy Information Administration for additional information pertaining to the question.