Math 23C Term Project

Robi Rahman and Sharon Tai

24 March - 14 May 2021

Question

Can the price changes in certain commodities reveal the kind of recession the U.S. is in? How do qualitatively different recessions affect commodities' prices? We are specifically interested in the dotcom crash of the early 2000s, the Great Recession, and the COVID-19 pandemic.

Motivation

"The stock market is not the economy." This refrain frequently sounds. We have decided to assess whether commodity prices that could signal the health of the general public correlated with unemployment rates across two decades. The three significant recessions in this time period affected different populations differently. Could the differences in commodity prices signal which populations were significantly affected, and how affected they were? Are there differences in how each type of recession affects them?

In general, the risk-adjusted long-run expected returns to all publicly traded assets are the same throughout the economy. If one commodity had a higher expected return than another, traders would sell the less profitable asset and buy the more profitable one, until the costs of each are proportionate to their future returns. The best null hypothesis for changes in asset prices is that all of them (specifically all the ones that are equally risky) will rise or fall by the same amount, all other factors equal.

However, different goods change in value differently in relation to different world events, and some commodities are correlated with other ones. We have selected a list of commodities that we hypothesize will behave differently during qualitatively different recessions. Did some of the selected commodities respond differently to these qualitatively different recessions?

Auxiliary questions we'll be considering are 1. What the price signals about the good and its consumers 2. Whether there were supply and demand shocks that affected the prices

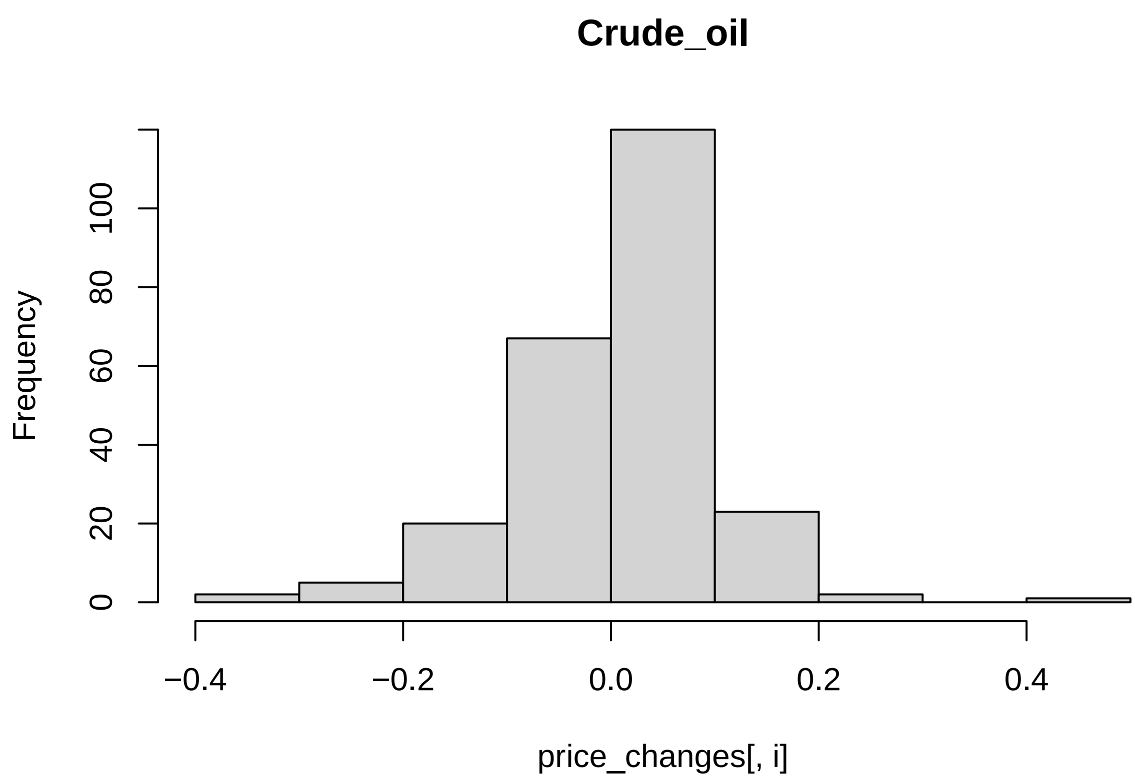
Hypothesis

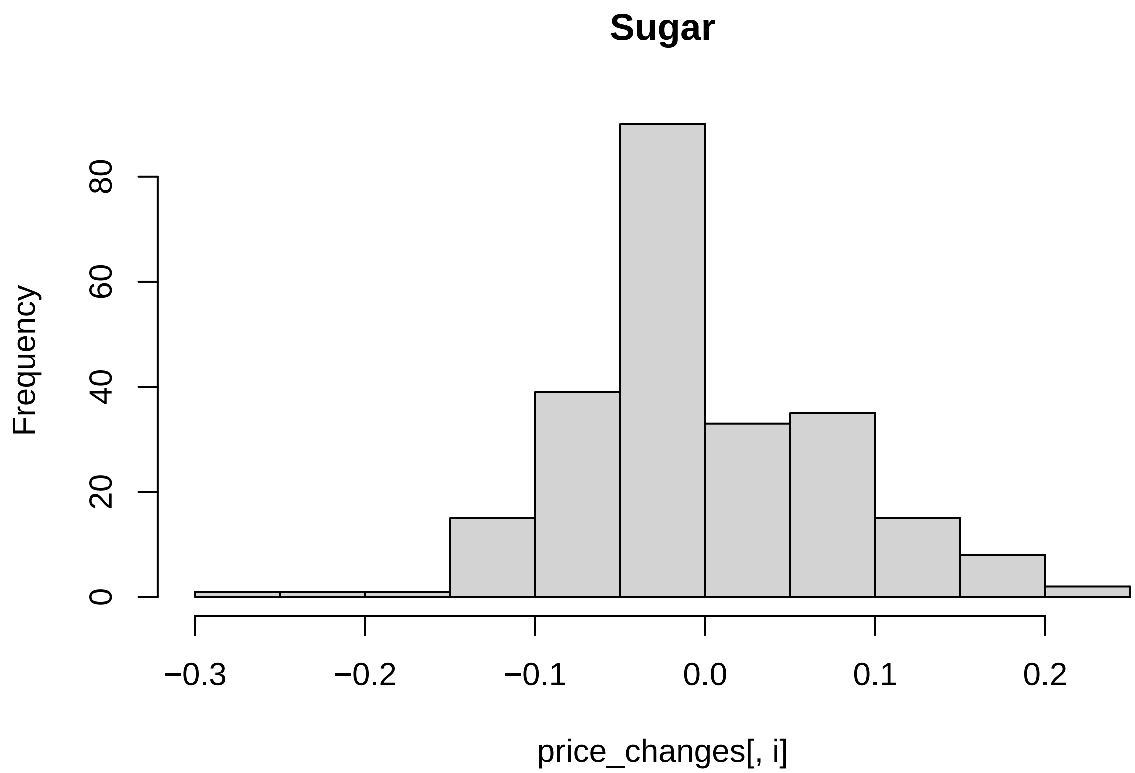
Pre-register our hypotheses here!

Analysis

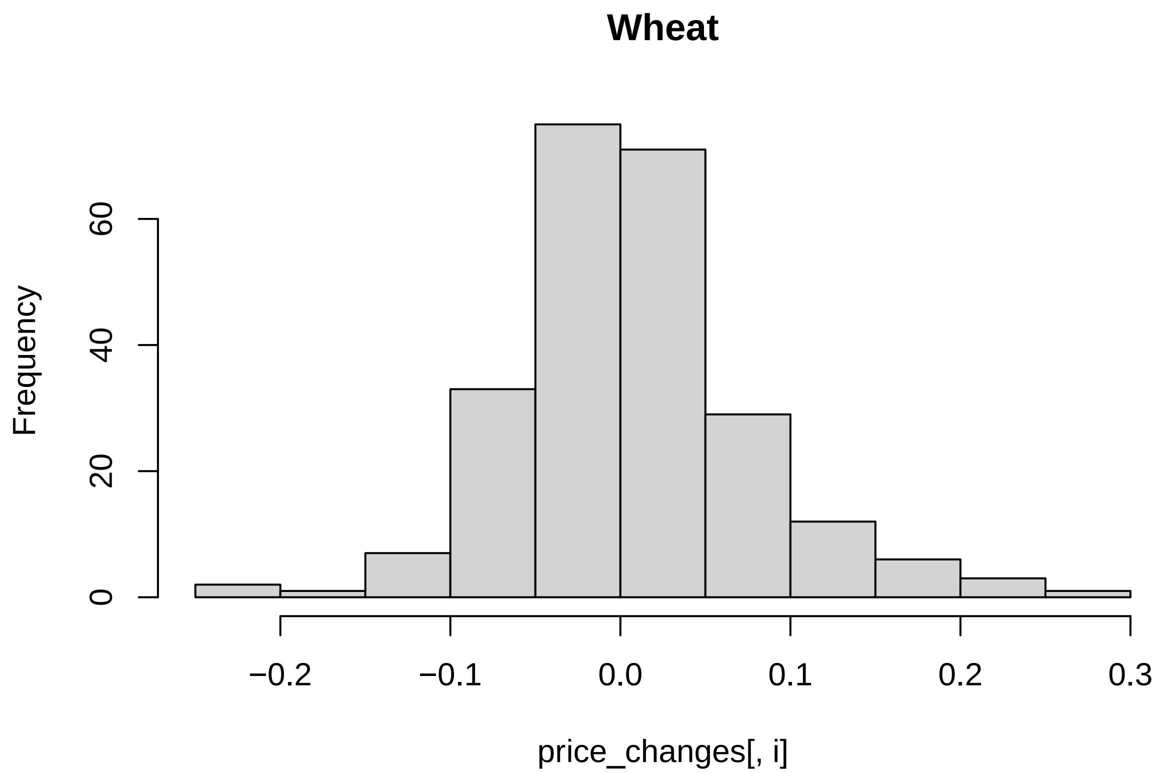
```
commodity_prices <- read.csv("source_data/commodities data.csv")
recession_dates <- read.csv("source_data/monthly recession indicator.csv")

recession_dates <- recession_dates[c(-242),] # Removed the entry for March 2021 so that the datasets ha
commodity_prices[,9] <- decomma(commodity_prices[,9]) # Remove commas from the price of gold column.

goods <- c("Month","Crude_oil", "Sugar", "Soybeans", "Wheat", "Beef", "Rubber", "Cocoa_beans", "Gold",
names(commodity_prices) <- goods
```

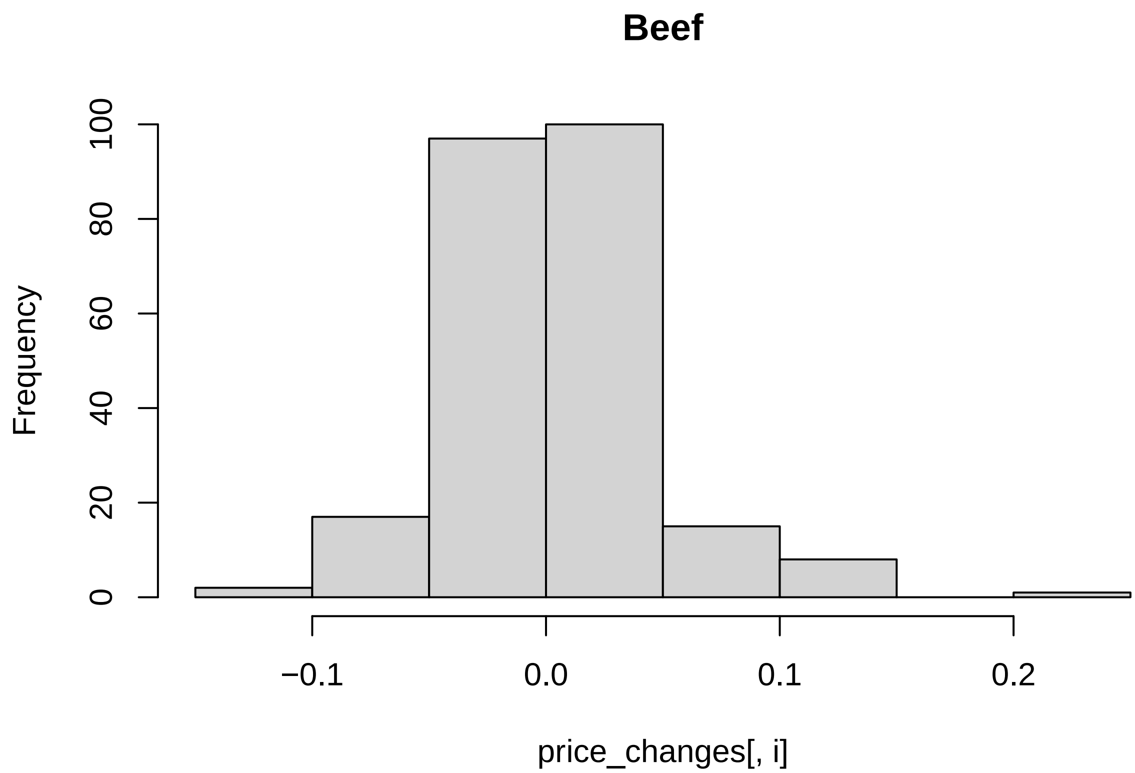
```r
# Adding a column to distinguish the three different recessions in our date range.
recession_dates$which_recession <- recession_dates$USREC
recession_dates[84:101,3] <- 2*recession_dates[84:101,3]
recession_dates[230:241,3] <- 3*recession_dates[230:241,3]

# Find month-over-month price changes for each commodity
price_changes <- commodity_prices
for (c in 2:11) {
  for (r in 2:241) {
    price_changes[r,c] <- commodity_prices[r,c] / commodity_prices[r-1,c] - 1
  }
}
price_changes$recession_bool <- recession_dates$USREC
price_changes$which_recession <- recession_dates$which_recession
price_changes <- price_changes[c(-1),]

# Display a histogram of the price changes
for (i in 2:11) {
  name <- goods[i]
  hist(price_changes[,i], main=name)
}
```

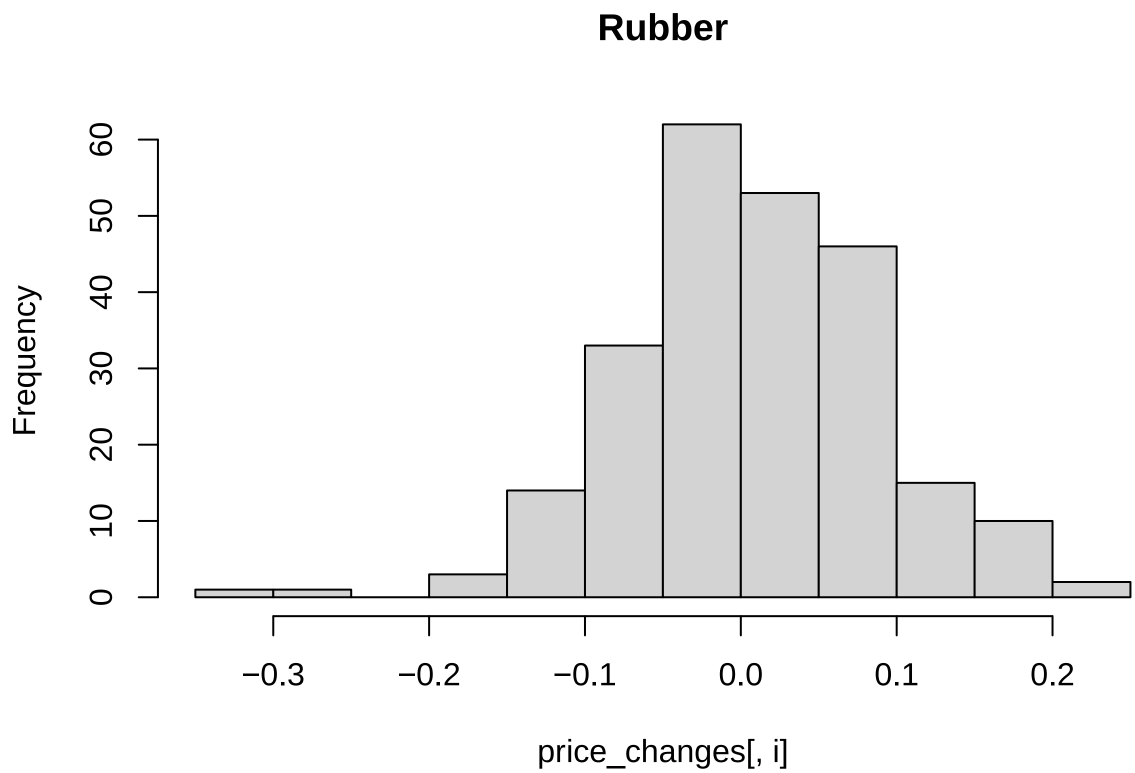


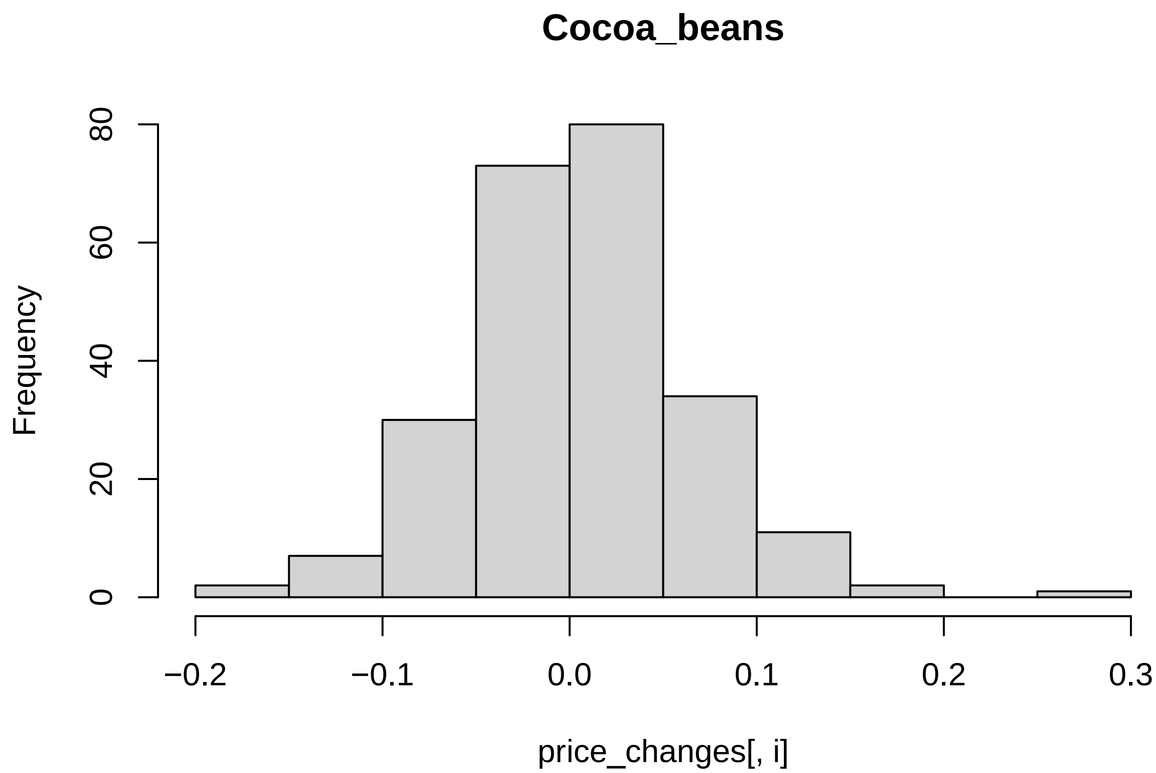

Sugar
Frequency
0
20
40
60
80
−0.3
−0.2
−0.1
0.0
0.1
0.2
price_changes[, i]


Soybeans
Frequency
0
20
40
60
80
−0.2
−0.1
0.0
0.1
0.2
price_changes[, i]

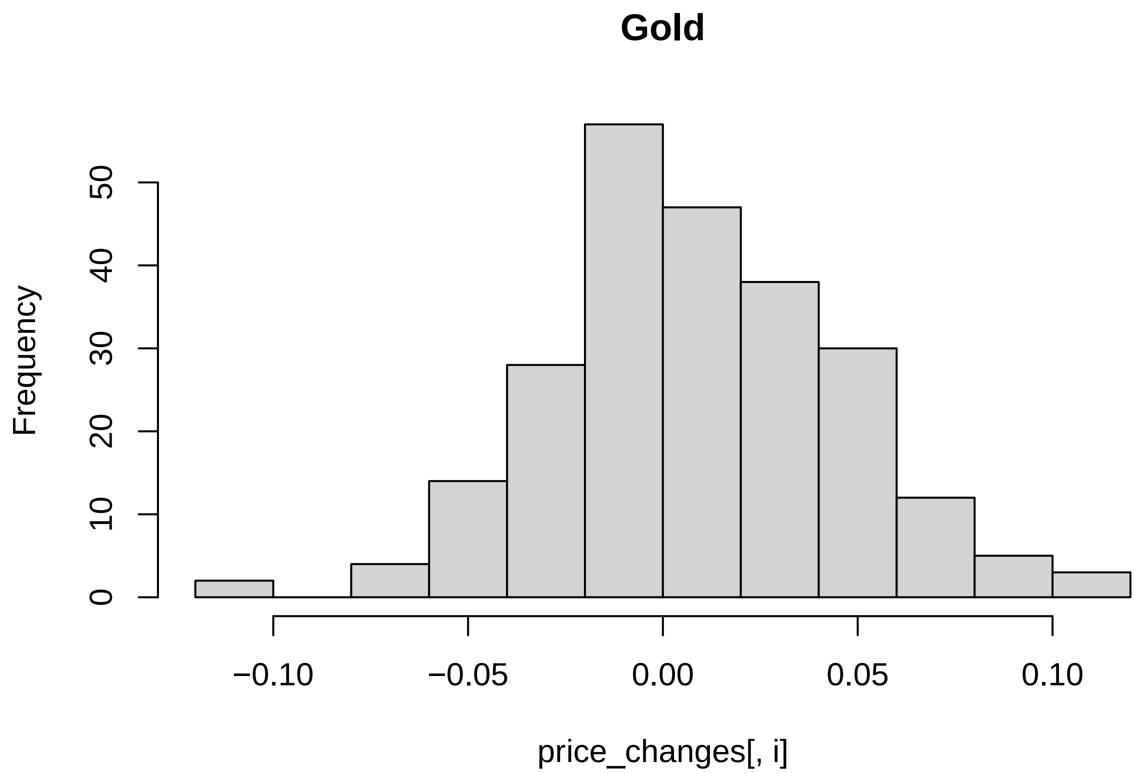

Wheat
Frequency
0
20
40
60
−0.2
−0.1
0.0
0.1
0.2
0.3
price_changes[, i]

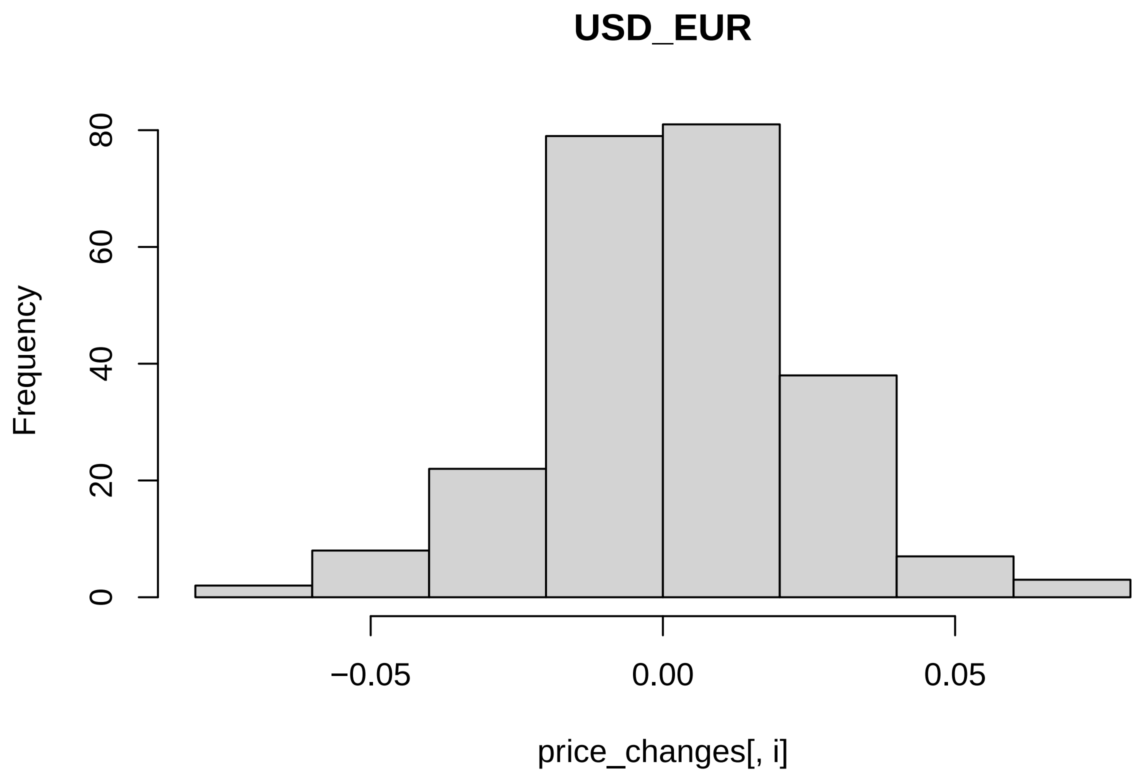

Beef
Frequency
0
20
40
60
80
100
−0.1
0.0
0.1
0.2
price_changes[, i]

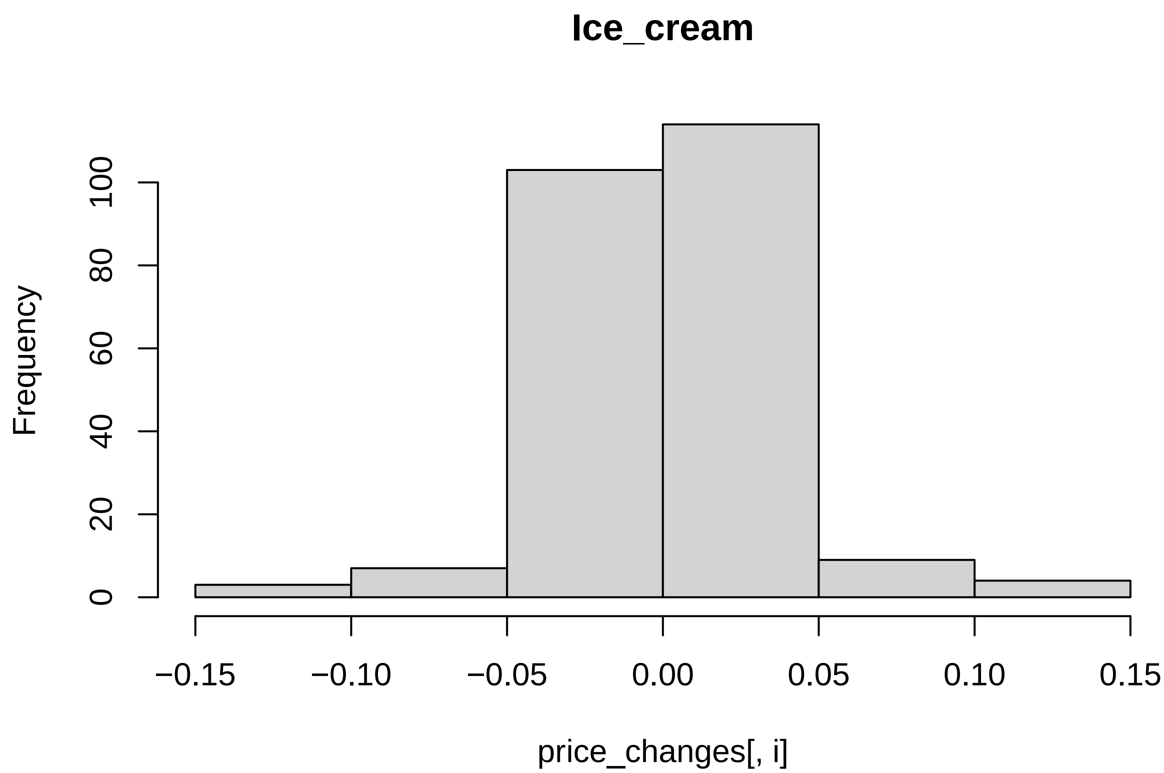

Rubber
Frequency
0
10
20
30
40
50
60
−0.3
−0.2
−0.1
0.0
0.1
0.2
price_changes[, i]



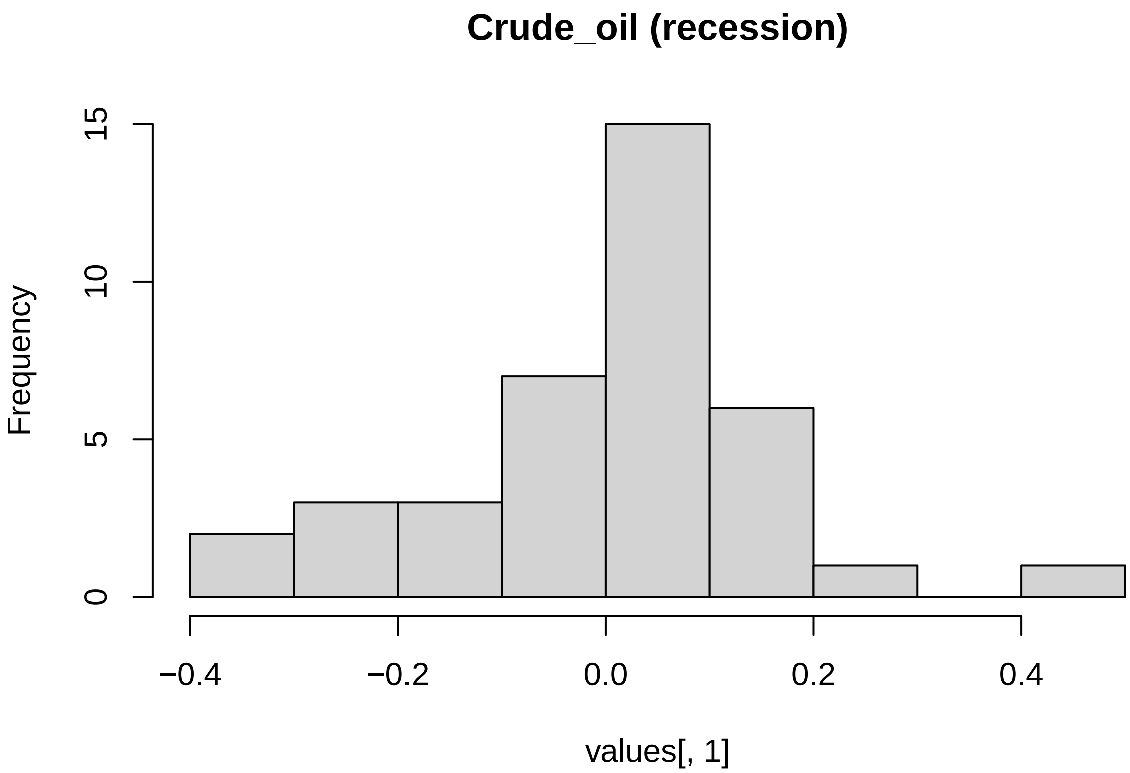
Cocoa_beans
Frequency
0
20
40
60
80
−0.2
−0.1
0.0
0.1
0.2
0.3
price_changes[, i]

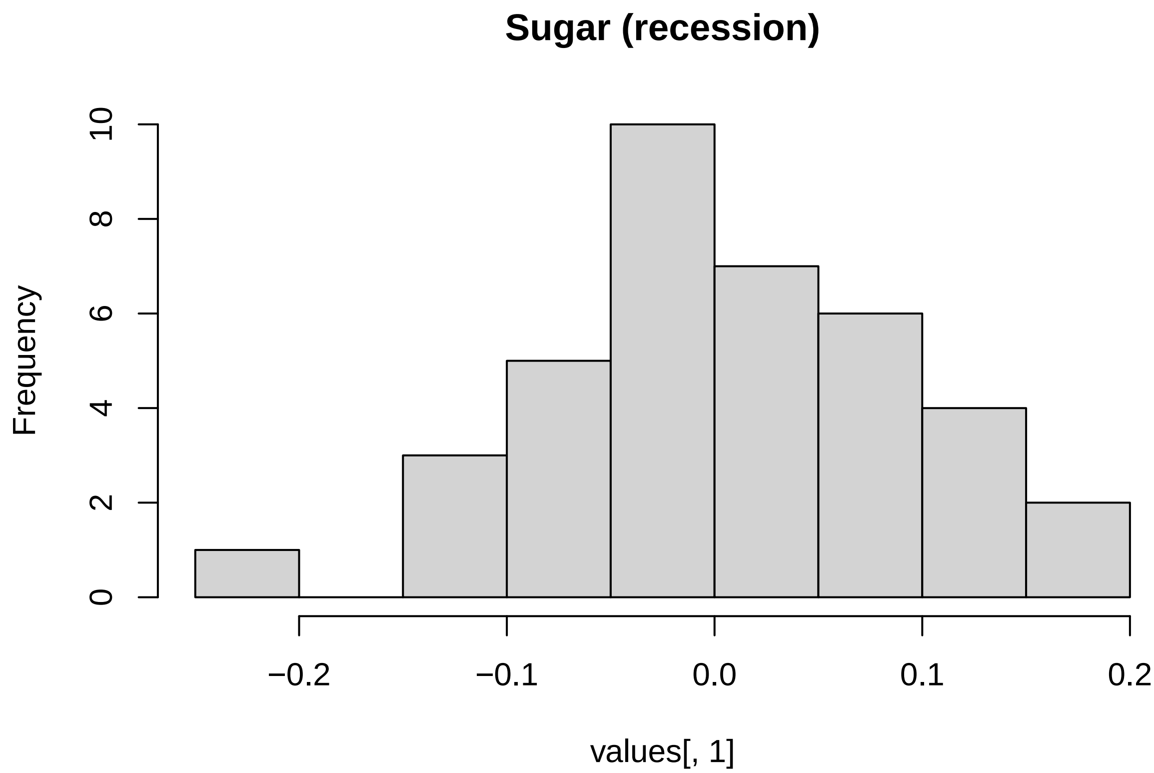

Gold
Frequency
0
10
20
30
40
50
−0.10
−0.05
0.00
0.05
0.10
price_changes[, i]

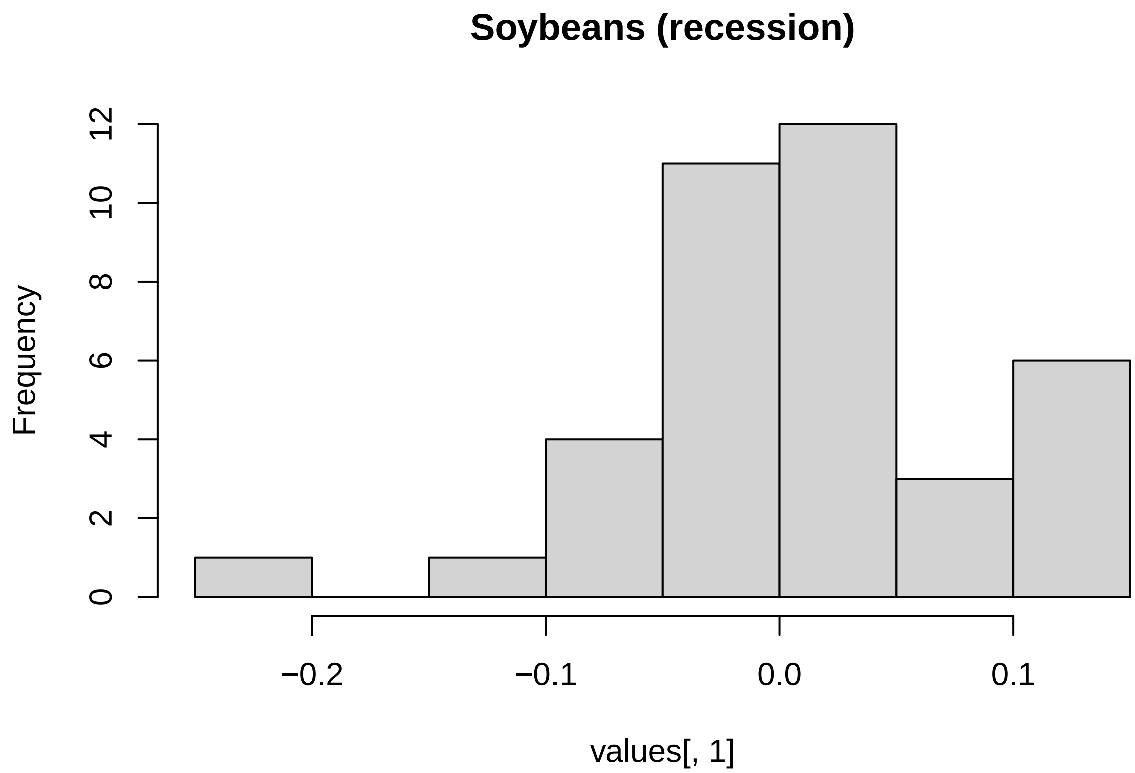

USD_EUR
Frequency
0
20
40
60
80
−0.05
0.00
0.05
price_changes[, i]



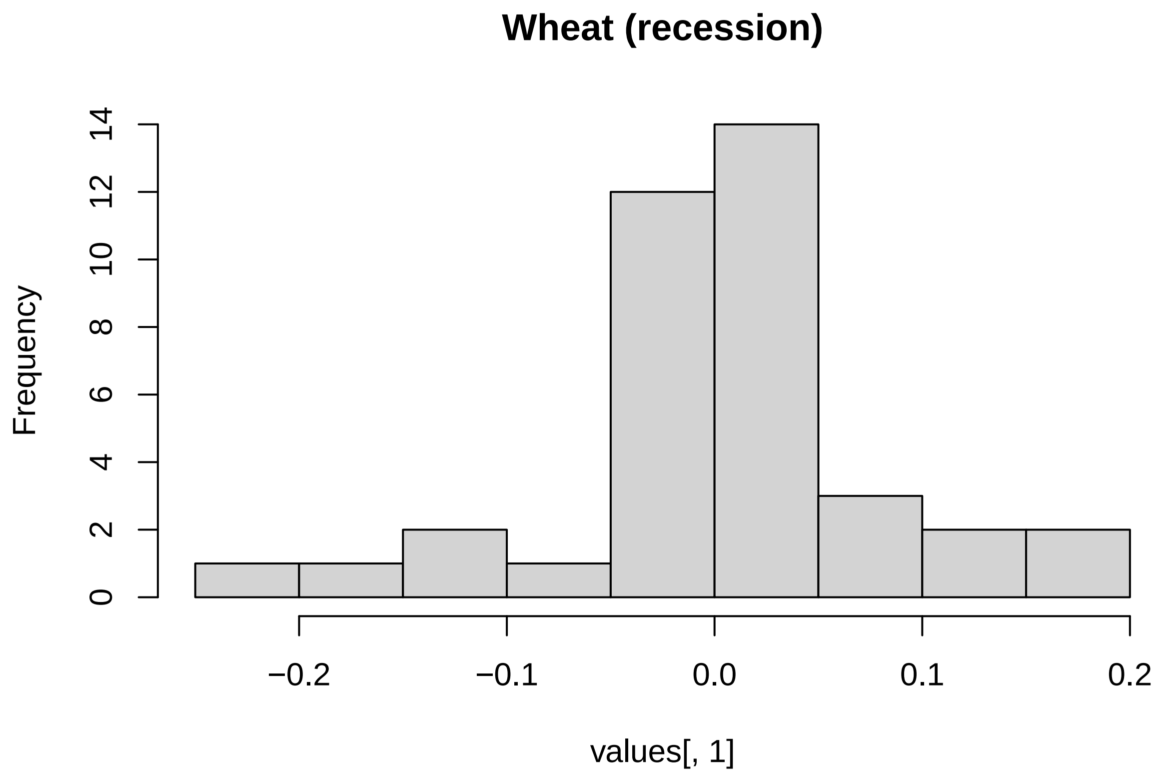

```r
# Repeating that, but during recessions only
for (i in 2:11) {
  name <- paste(goods[i],"(recession)")
  values <- price_changes[,c(i,13)]
  values <- values[values[,2] == 1,]
  hist(values[,1], main=name)
}
```

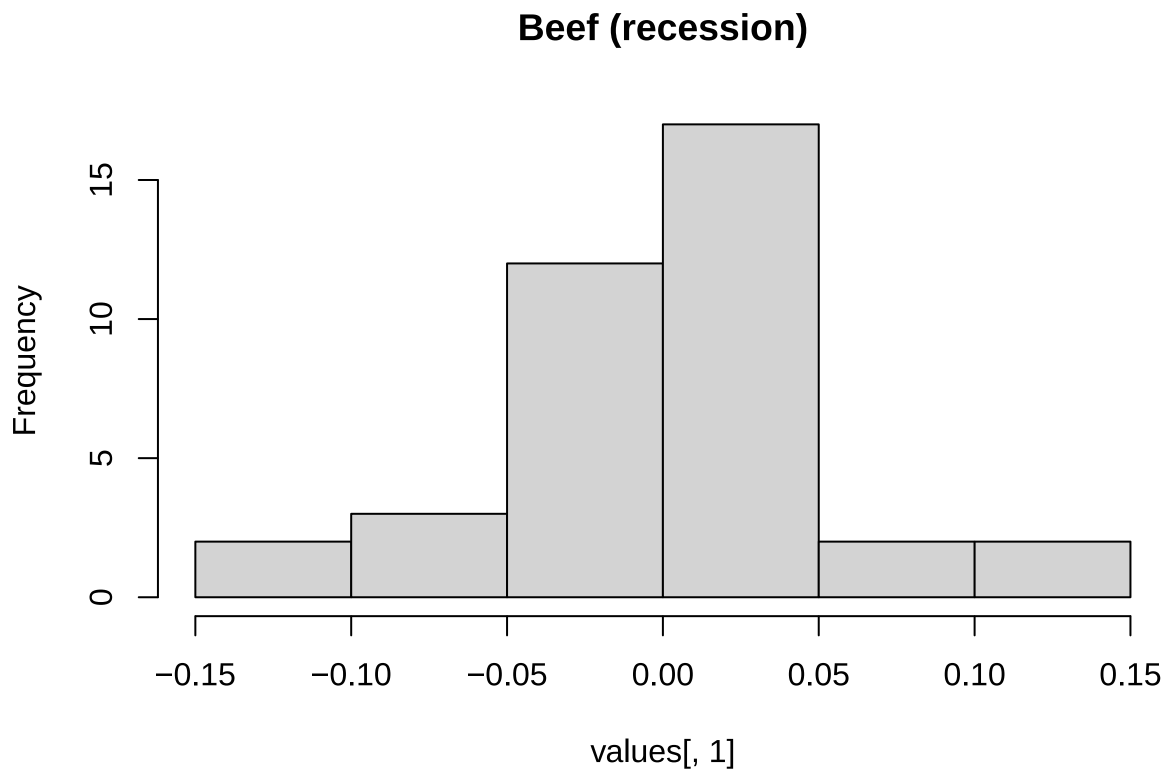

Crude_oil (recession)
Frequency
0
5
10
15
−0.4
−0.2
0.0
0.2
0.4
values[, 1]

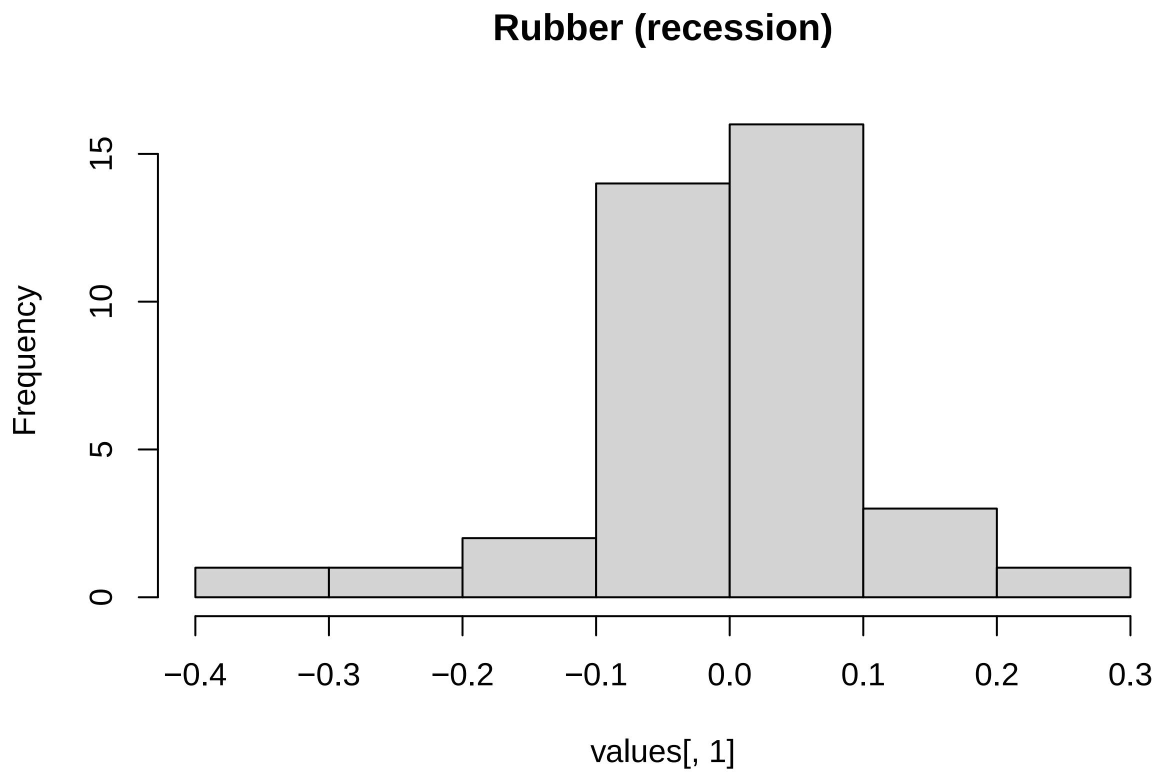

Sugar (recession)
Frequency
0
2
4
6
8
10
−0.2
−0.1
0.0
0.1
0.2
values[, 1]



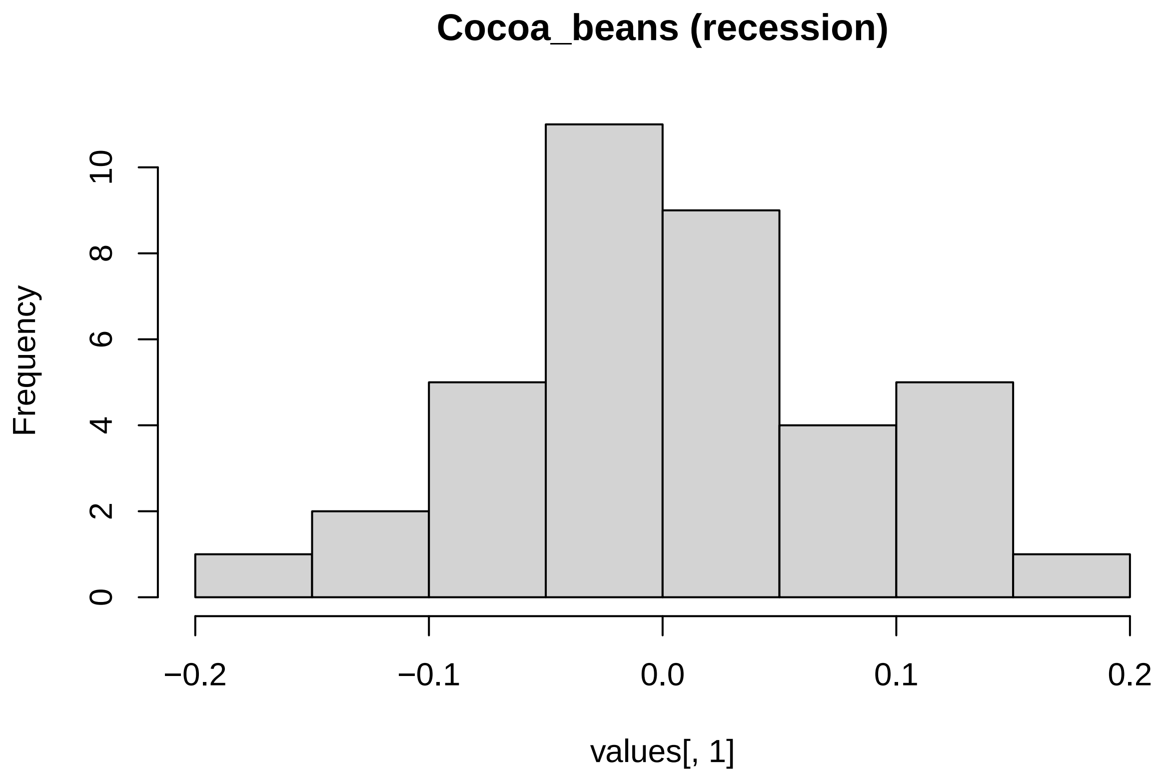
Soybeans (recession)
Frequency
0
2
4
6
8
10
12
−0.2
−0.1
0.0
0.1
values[, 1]

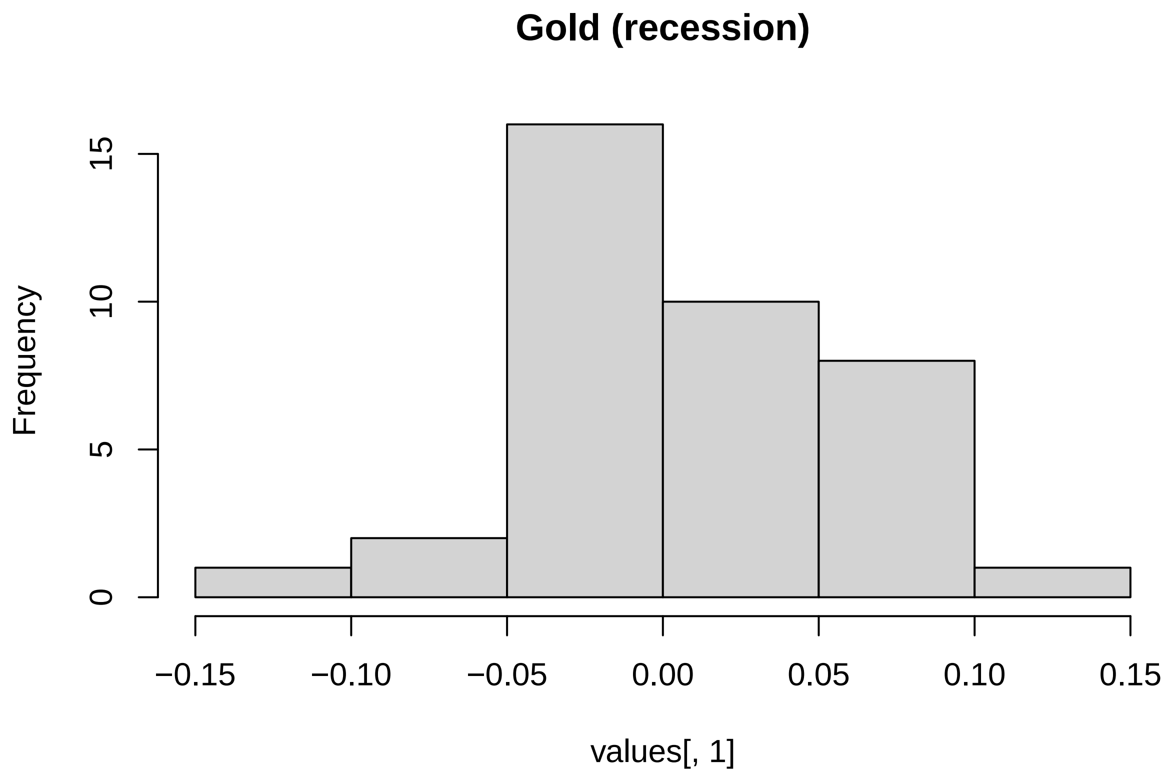

Wheat (recession)
Frequency
0
2
4
6
8
10
12
14
−0.2
−0.1
0.0
0.1
0.2
values[, 1]

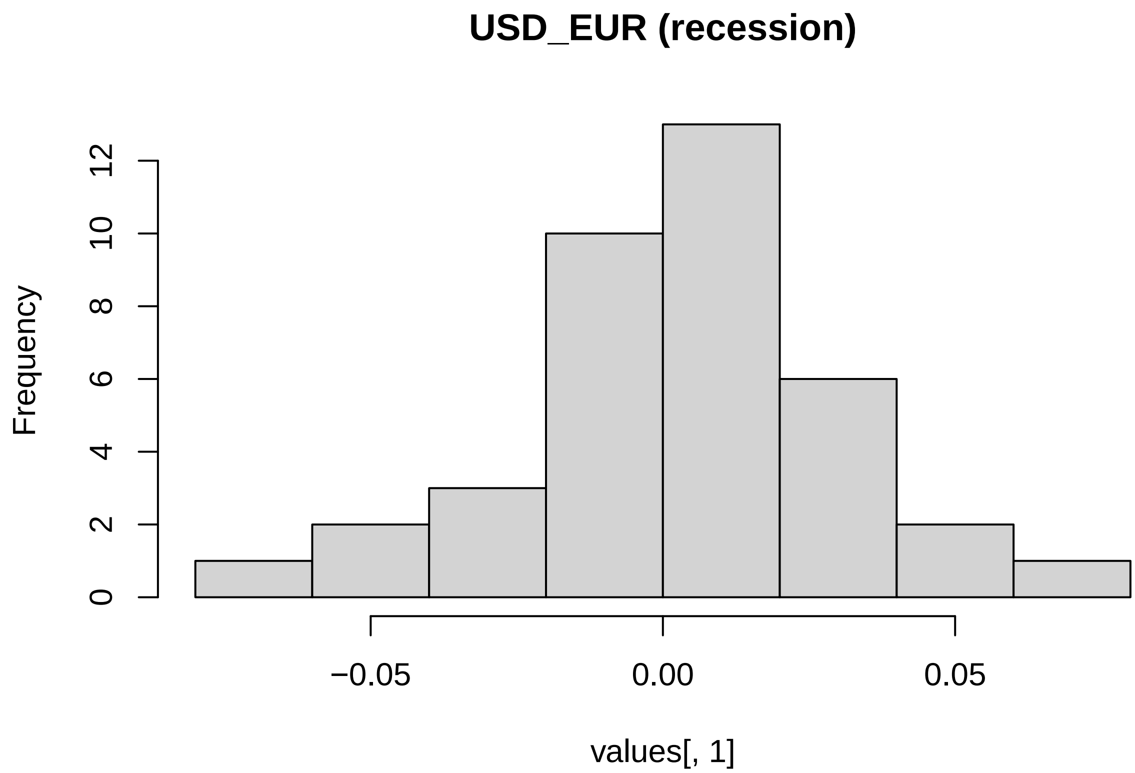

Beef (recession)
Frequency
0
5
10
15
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
values[, 1]

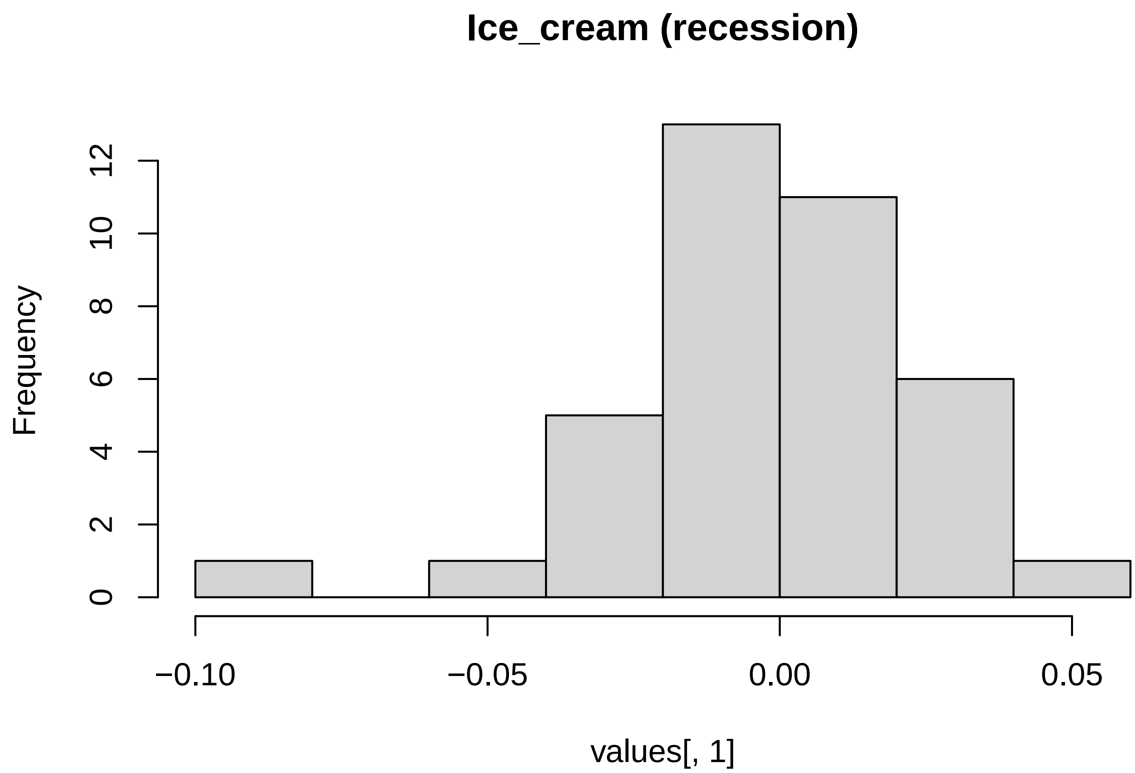


Rubber (recession)
Frequency
0
5
10
15
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
values[, 1]

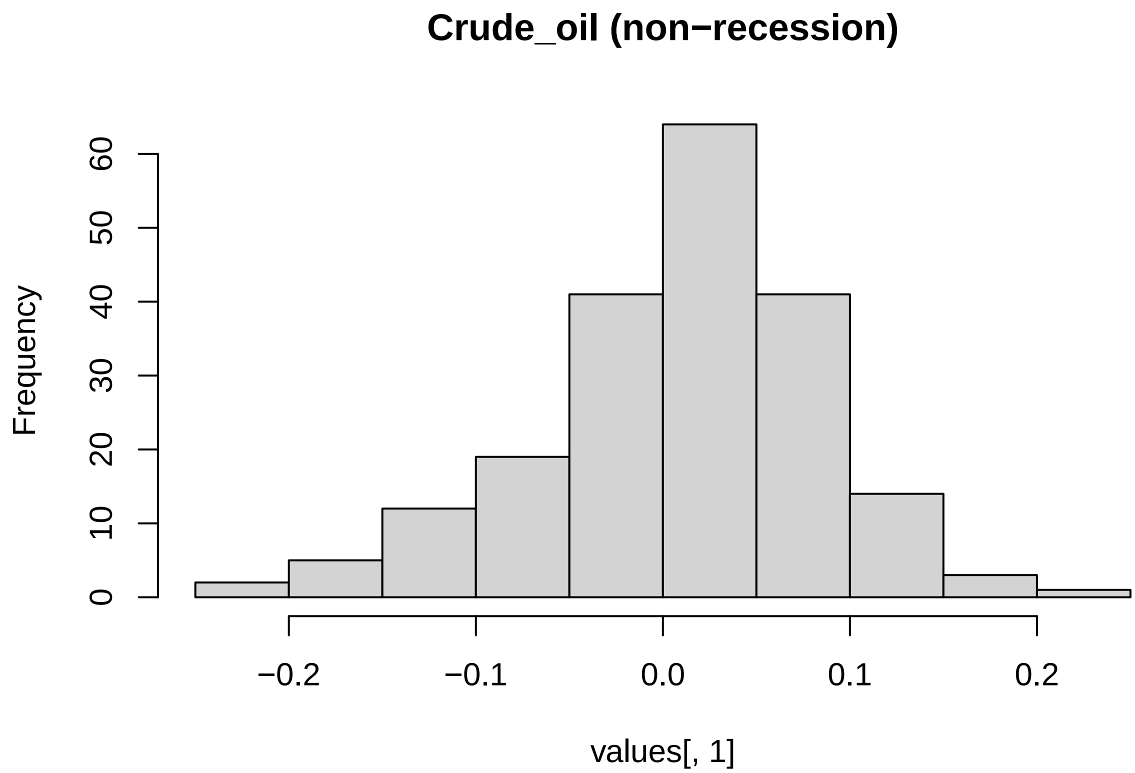

Cocoa_beans (recession)
Frequency
0
2
4
6
8
10
−0.2
−0.1
0.0
0.1
0.2
values[, 1]

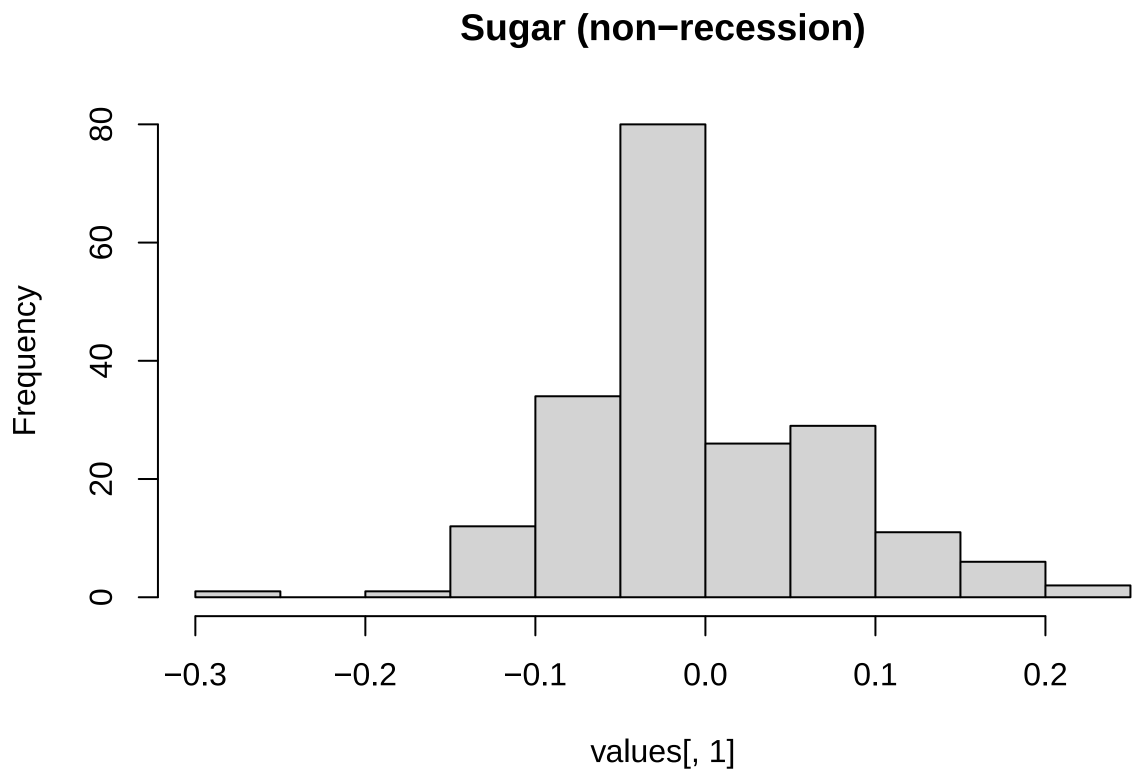

Gold (recession)
Frequency
0
5
10
15
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
values[, 1]

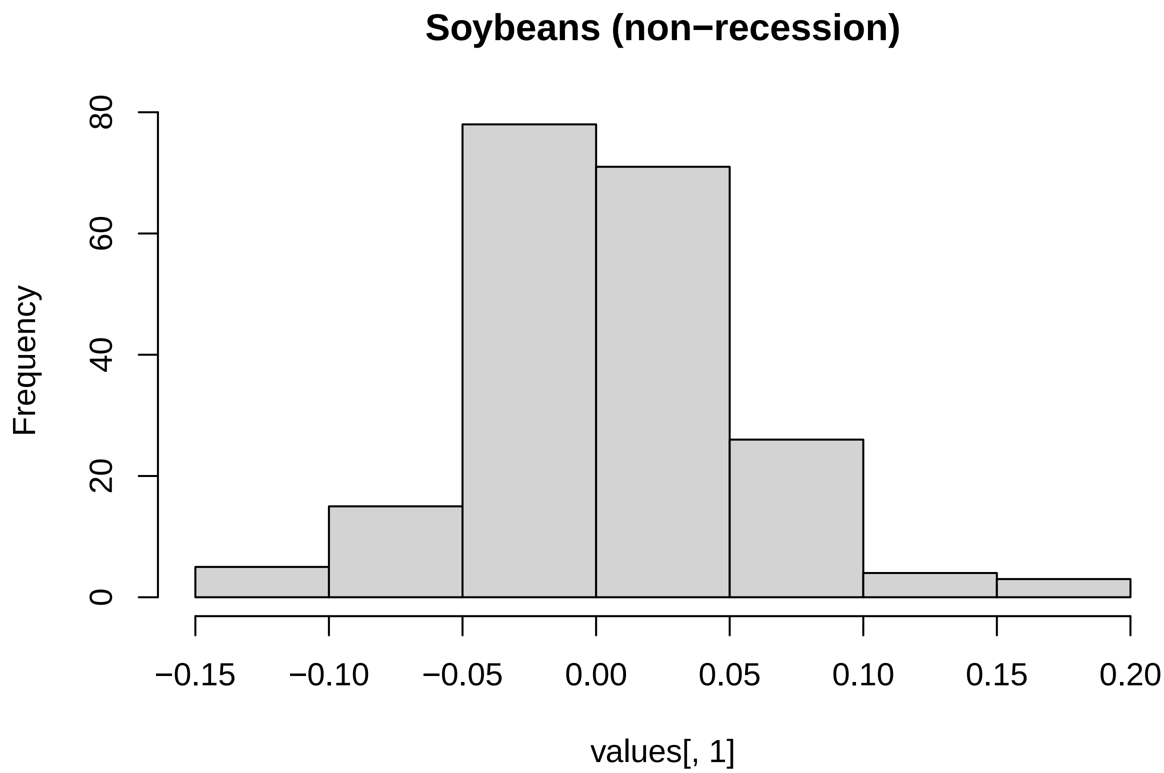

USD_EUR (recession)
Frequency
0
2
4
6
8
10
12
−0.05
0.00
0.05
values[, 1]



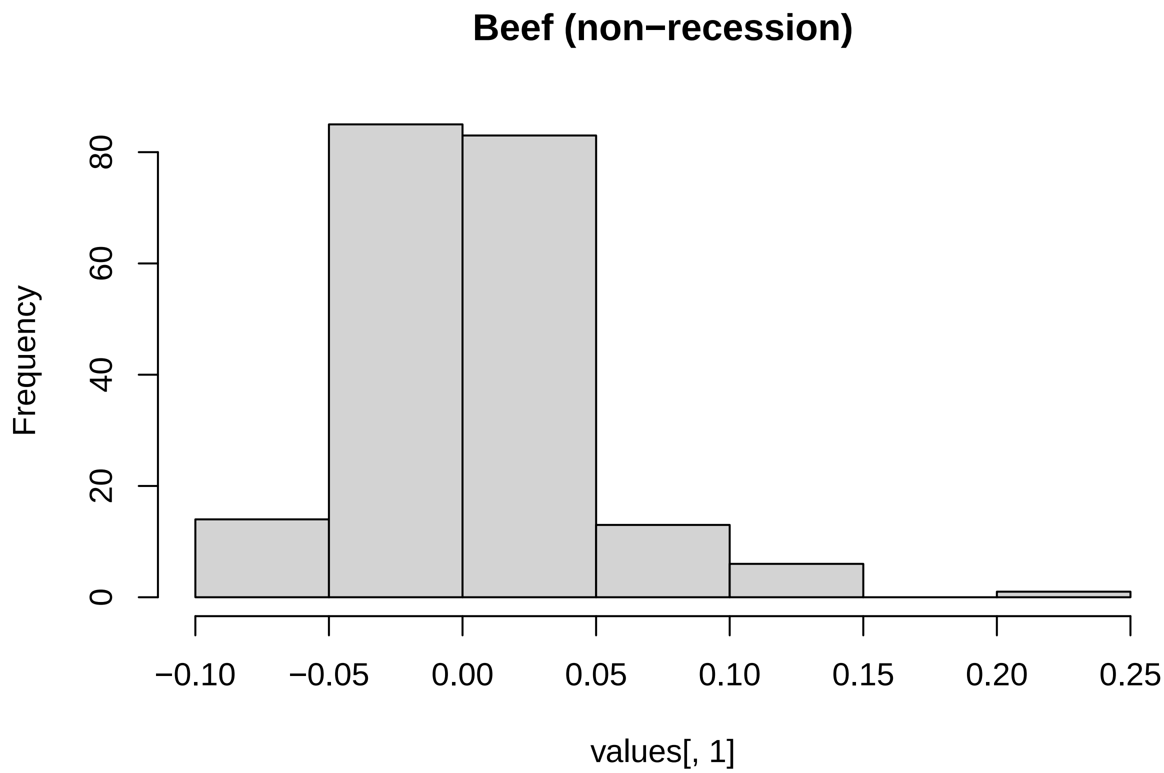

```r
# Repeat for months not in recession
for (i in 2:11) {
  name <- paste(goods[i],"(non-recession)")
  values <- price_changes[,c(i,13)]
  values <- values[values[,2] == 0,]
  hist(values[,1], main=name)
}
```

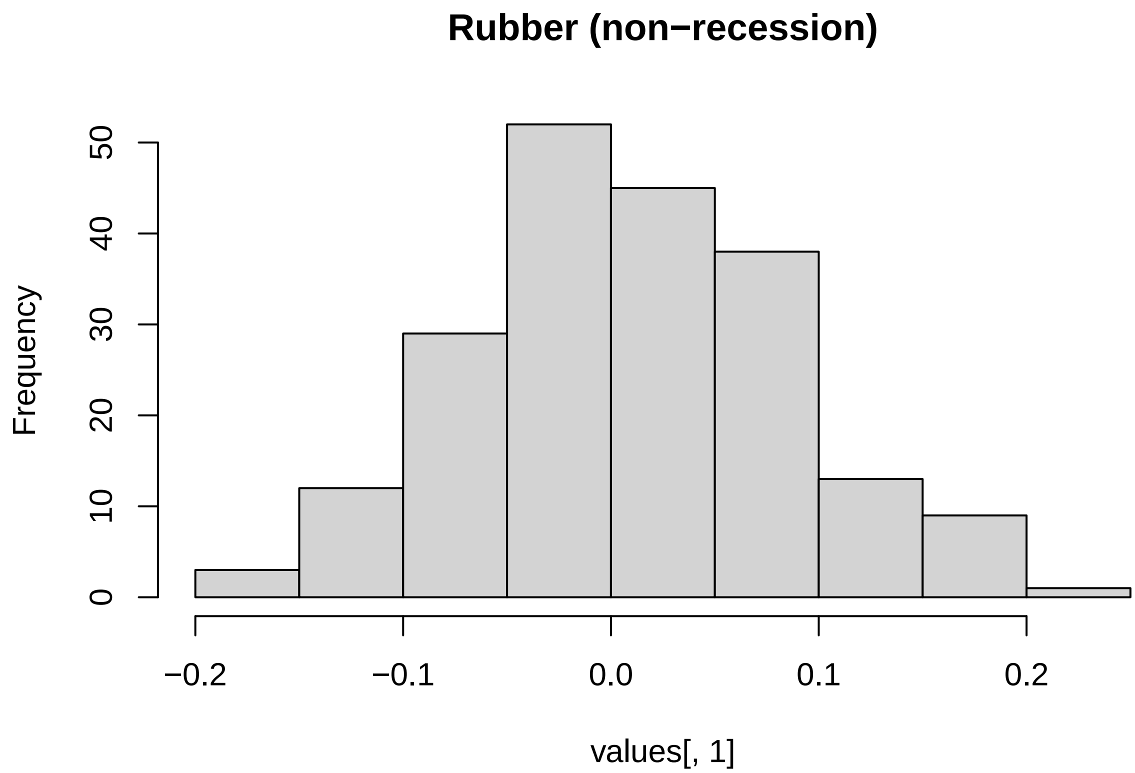

Crude_oil (non−recession)
Frequency
0
10
20
30
40
50
60
−0.2
−0.1
0.0
0.1
0.2
values[, 1]


Sugar (non−recession)
Frequency
0
20
40
60
80
−0.3
−0.2
−0.1
0.0
0.1
0.2
values[, 1]

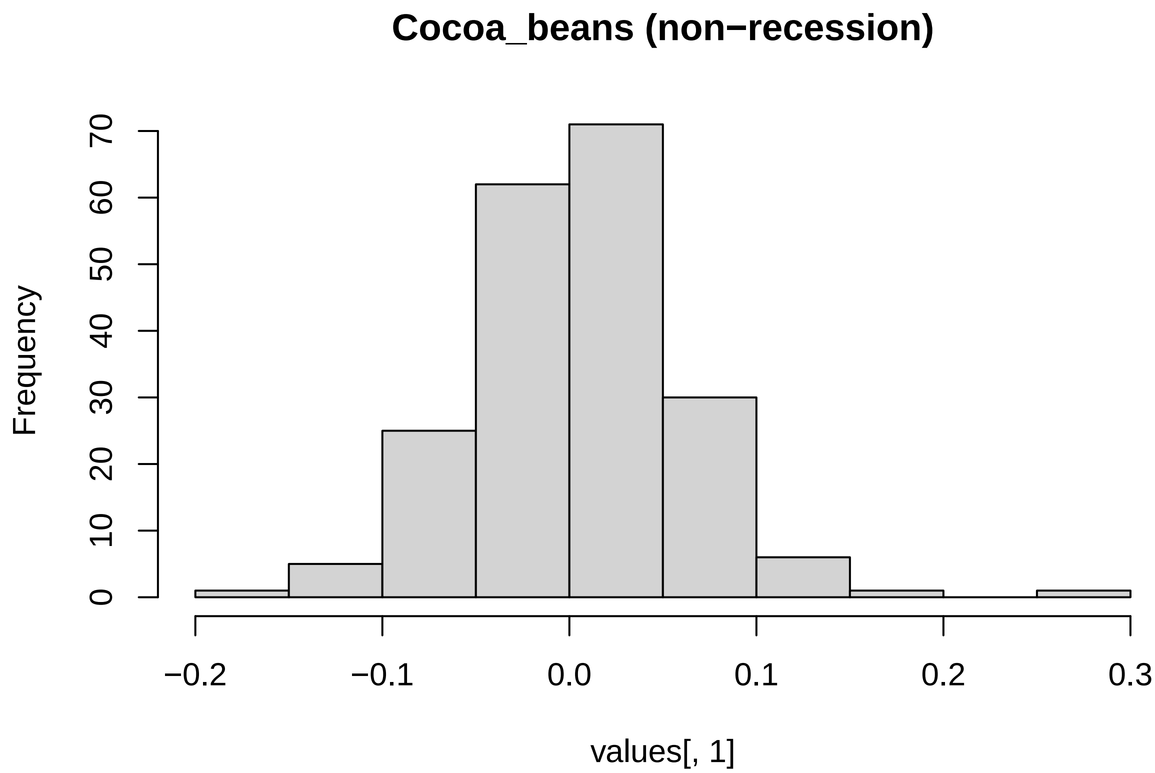

Soybeans (non−recession)
Frequency
0
20
40
60
80
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20
values[, 1]

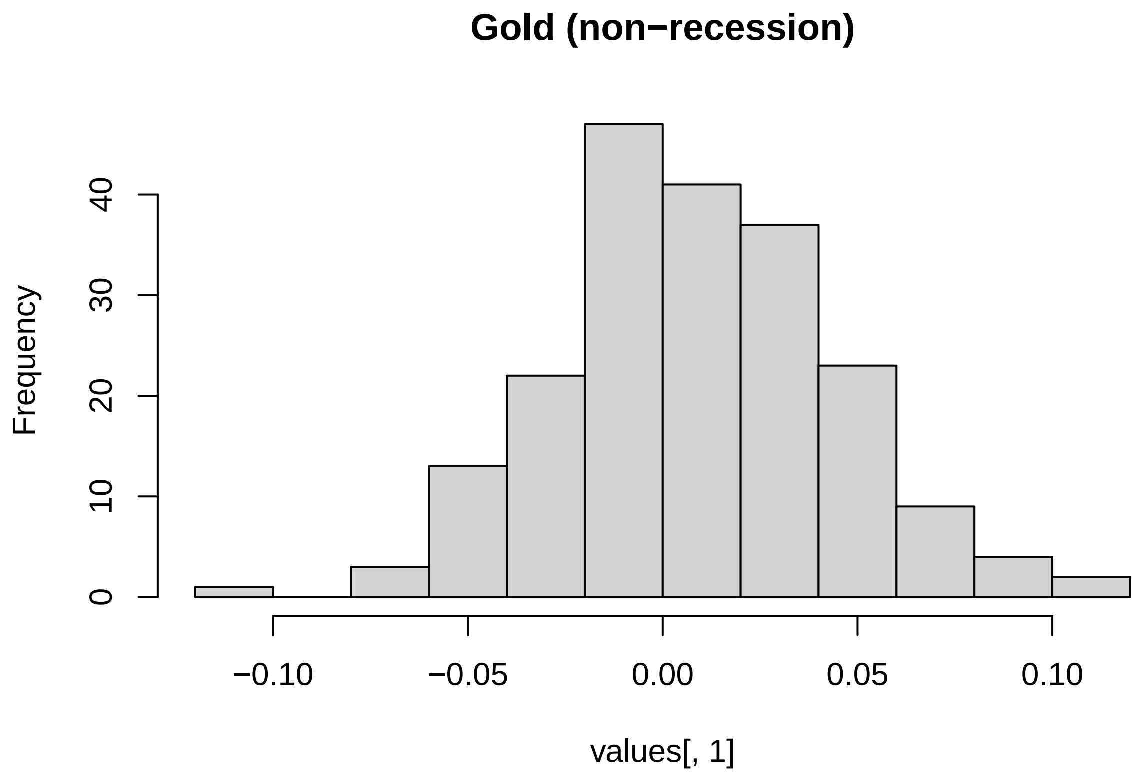

Wheat (non−recession)
Frequency
0
10
20
30
40
50
60
−0.2
−0.1
0.0
0.1
0.2
0.3
values[, 1]

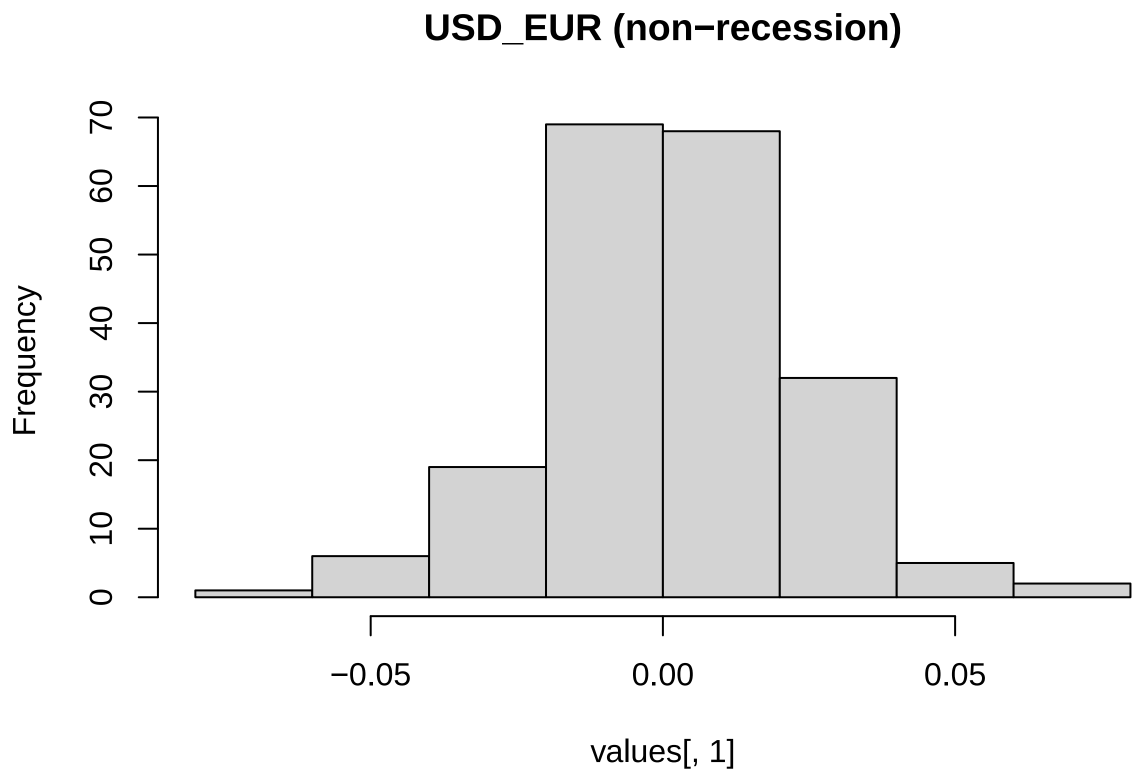

Beef (non−recession)
Frequency
0
20
40
60
80
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20
0.25
values[, 1]

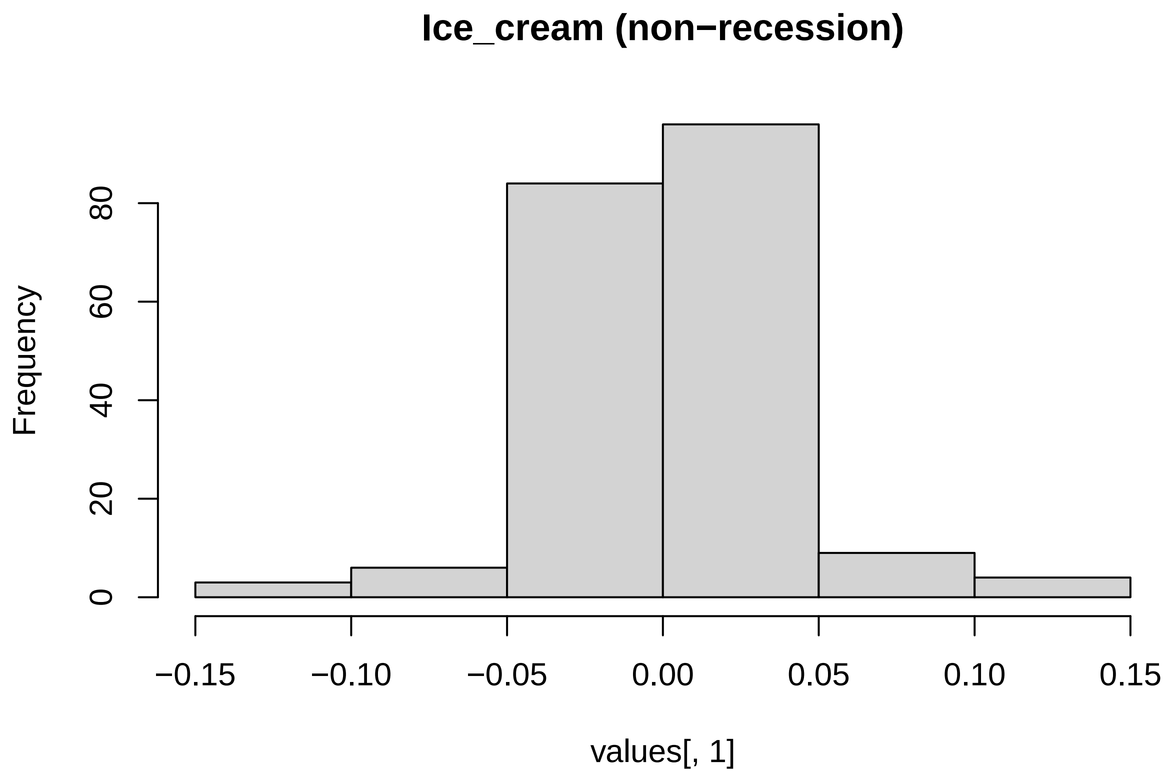


Rubber (non−recession)
Frequency
0
10
20
30
40
50
−0.2
−0.1
0.0
0.1
0.2
values[, 1]


Cocoa_beans (non−recession)
Frequency
0
10
20
30
40
50
60
70
−0.2
−0.1
0.0
0.1
0.2
0.3
values[, 1]


Gold (non−recession)
Frequency
0
10
20
30
40
−0.10
−0.05
0.00
0.05
0.10
values[, 1]

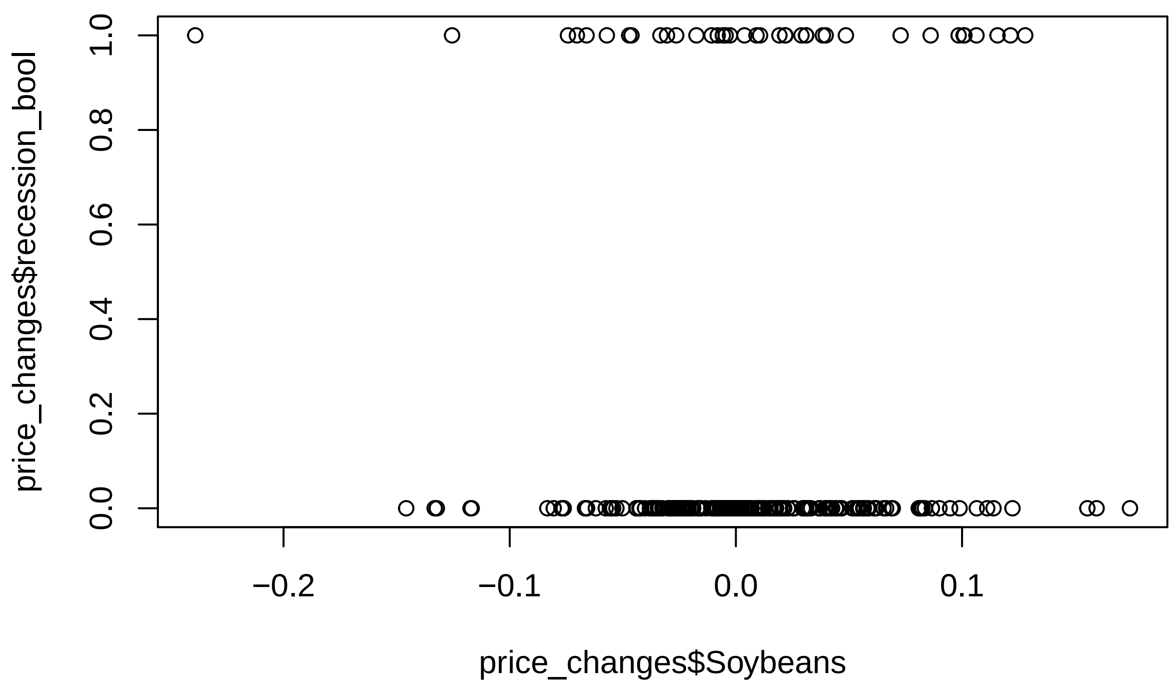

USD_EUR (non−recession)
Frequency
0
10
20
30
40
50
60
70
−0.05
0.00
0.05
values[, 1]




```
soybeans_regression <- glm(recession_bool ~ Soybeans, data = price_changes, family = 'binomial')
summary(soybeans_regression)
```

```
##
## Call:
## glm(formula = recession_bool ~ Soybeans, family = "binomial",
##     data = price_changes)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6622  -0.5949  -0.5817  -0.5656   2.0888
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6825     0.1794  -9.379   <2e-16 ***
## Soybeans      1.5869     3.2948   0.482     0.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 209.71  on 239  degrees of freedom
## Residual deviance: 209.48  on 238  degrees of freedom
## AIC: 213.48
##
## Number of Fisher Scoring iterations: 4
```

```r
plot(price_changes$Soybeans, price_changes$recession_bool)
```

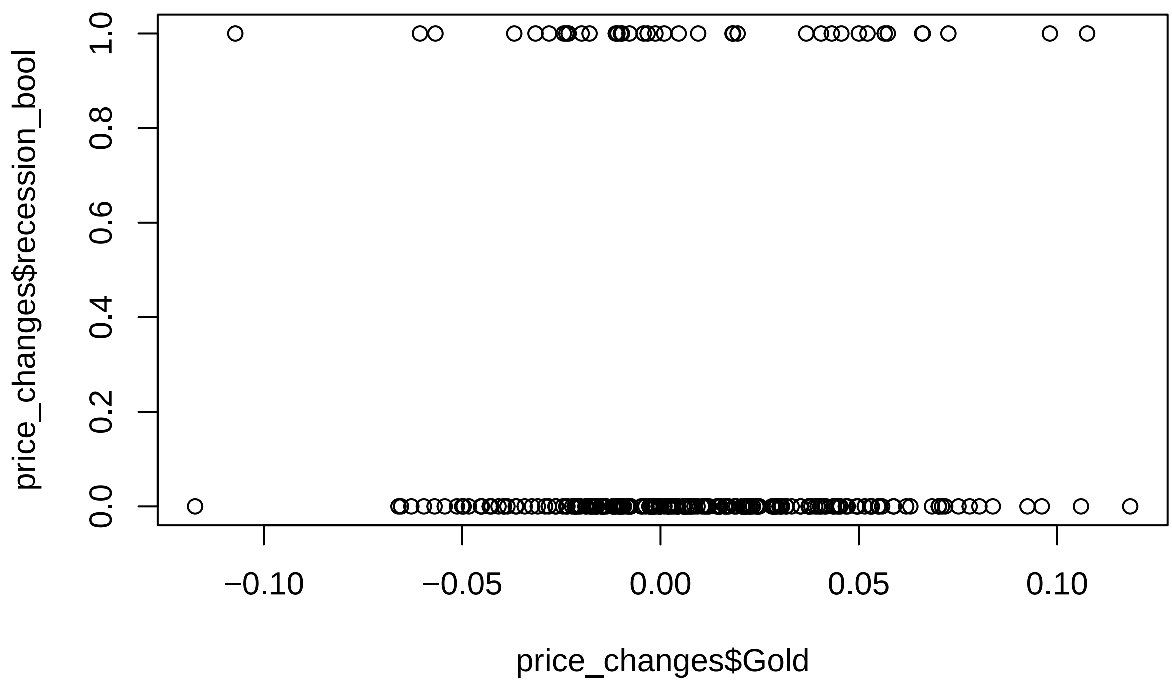



```r
gold_regression <- glm(recession_bool ~ Gold, data = price_changes, family = 'binomial')
summary(gold_regression)
```

```
##
## Call:
## glm(formula = recession_bool ~ Gold, family = "binomial", data = price_changes)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6153  -0.5911  -0.5845  -0.5761   1.9676
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6793     0.1827  -9.194   <2e-16 ***
## Gold          0.9375     4.7325   0.198    0.843
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 209.71  on 239  degrees of freedom
## Residual deviance: 209.67  on 238  degrees of freedom
## AIC: 213.67
##
```

```
## Number of Fisher Scoring iterations: 3
```

```r
plot(price_changes$Gold, price_changes$recession_bool)
```




```r
goods_vs_recession_logistic_model <- glm(
  recession_bool ~ Crude_oil + Sugar + Soybeans + Wheat + Beef + Rubber + Cocoa_beans + Gold + USD_EUR +
  data = price_changes, family = 'binomial')
summary(goods_vs_recession_logistic_model)
```

```
##
## Call:
## glm(formula = recession_bool ~ Crude_oil + Sugar + Soybeans +
##     Wheat + Beef + Rubber + Cocoa_beans + Gold + USD_EUR + Ice_cream,
##     family = "binomial", data = price_changes)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.8139  -0.6123  -0.5662  -0.4928   2.0971
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.7006     0.1882  -9.036   <2e-16 ***
## Crude_oil    -1.2750     2.2732  -0.561    0.575
## Sugar         2.3158     2.6110   0.887    0.375
## Soybeans      3.3223     3.9135   0.849    0.396
## Wheat        -2.3030     2.8125  -0.819    0.413
## Beef          1.2275     4.4501   0.276    0.783
```

```
## Rubber       -1.6263     2.5382  -0.641    0.522
## Cocoa_beans   0.6074     3.1873   0.191    0.849
## Gold          0.6411     5.3050   0.121    0.904
## USD_EUR       0.5395     9.3769   0.058    0.954
## Ice_cream    -1.0465     5.1222  -0.204    0.838
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 209.71  on 239  degrees of freedom
## Residual deviance: 207.41  on 229  degrees of freedom
## AIC: 229.41
##
## Number of Fisher Scoring iterations: 4
```

```r
negative_goods_regression <- glm(recession_bool ~ Crude_oil + Wheat + Rubber + Ice_cream, data = price_c
summary(negative_goods_regression)
```

```
##
## Call:
## glm(formula = recession_bool ~ Crude_oil + Wheat + Rubber + Ice_cream,
##     family = "binomial", data = price_changes)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6870  -0.6013  -0.5820  -0.5460   2.0757
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6582     0.1774  -9.349   <2e-16 ***
## Crude_oil    -0.2640     2.0091  -0.131    0.895
## Wheat        -0.7643     2.5456  -0.300    0.764
## Rubber       -0.8586     2.4213  -0.355    0.723
## Ice_cream    -1.5616     5.0401  -0.310    0.757
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 209.71  on 239  degrees of freedom
## Residual deviance: 209.22  on 235  degrees of freedom
## AIC: 219.22
##
## Number of Fisher Scoring iterations: 4
```

```r
# Probability of a recession if all of these goods lost 100% of their value this month:
inv.logit(-1.6582+0.2640+0.7643+0.8586+1.5616)
```

```
## [1] 0.8569641
```

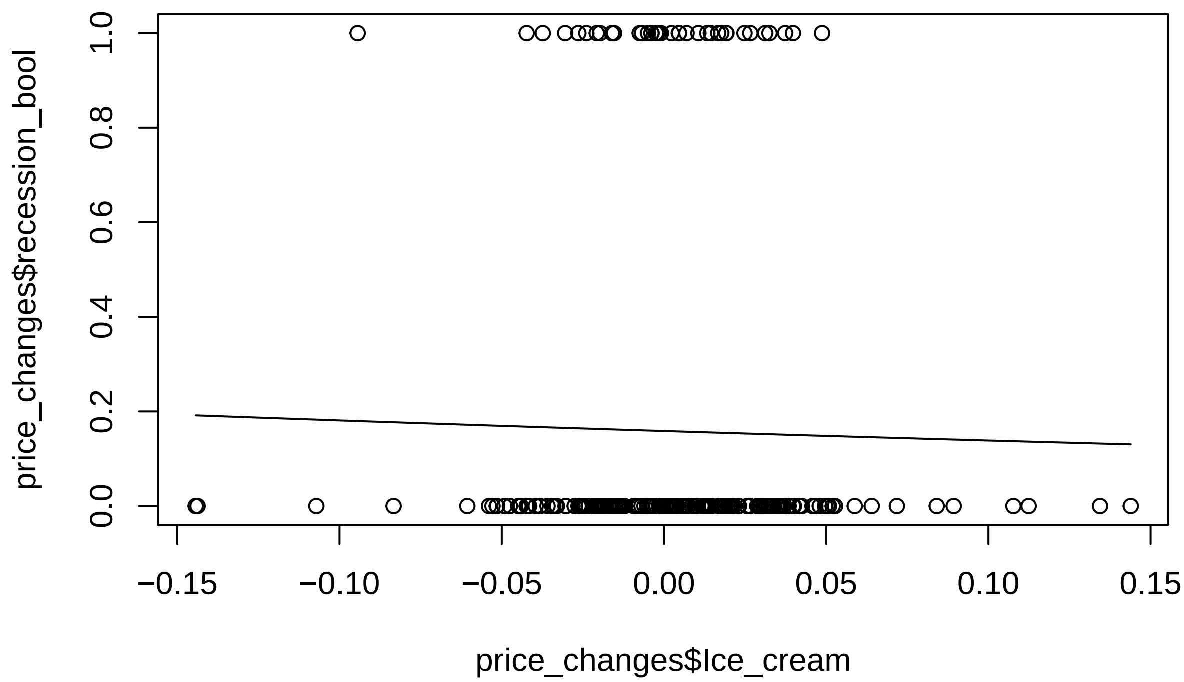
```r
# Probability of a recession if all of these goods doubled in price last month:
inv.logit(-1.6582-0.2640-0.7643-0.8586-1.5616)
```

```
## [1] 0.00601958
```

```r
icecream_regression <- glm(recession_bool ~ Ice_cream, data = price_changes, family = 'binomial')
summary(icecream_regression)
```

```
##
## Call:
## glm(formula = recession_bool ~ Ice_cream, family = "binomial",
##     data = price_changes)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6524  -0.5946  -0.5845  -0.5708   1.9529
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6687     0.1769  -9.434   <2e-16 ***
## Ice_cream    -1.5900     4.9778  -0.319    0.749
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 209.71  on 239  degrees of freedom
## Residual deviance: 209.61  on 238  degrees of freedom
## AIC: 213.61
##
## Number of Fisher Scoring iterations: 3
```

```r
plot(price_changes$Ice_cream, price_changes$recession_bool)
curve(inv.logit(-1.5900*x-1.6687), add=TRUE)
```




```
icecream_predictions <- as.factor(predict(icecream_regression, newdata=price_changes, type='response')
confusionMatrix(icecream_predictions, reference = as.factor(price_changes$recession_bool==1))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction FALSE TRUE
##      FALSE   109   18
##      TRUE     93   20
##
##                Accuracy : 0.5375
##                  95% CI : (0.4722, 0.6019)
##     No Information Rate : 0.8417
##     P-Value [Acc > NIR] : 1
##
##                   Kappa : 0.0366
##
##  Mcnemar's Test P-Value : 2.16e-12
##
##             Sensitivity : 0.5396
##             Specificity : 0.5263
##          Pos Pred Value : 0.8583
##          Neg Pred Value : 0.1770
##              Prevalence : 0.8417
##          Detection Rate : 0.4542
##    Detection Prevalence : 0.5292
```

```
##       Balanced Accuracy : 0.5330
##
##        'Positive' Class : FALSE
##
```

```
recession_predictions <- as.factor(predict(goods_vs_recession_logistic_model, newdata=price_changes, typ
confusionMatrix(recession_predictions, reference = as.factor(price_changes$recession_bool==1))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction FALSE TRUE
##      FALSE   104   19
##      TRUE     98   19
##
##                Accuracy : 0.5125
##                  95% CI : (0.4474, 0.5773)
##     No Information Rate : 0.8417
##     P-Value [Acc > NIR] : 1
##
##                   Kappa : 0.0081
##
##  Mcnemar's Test P-Value : 5.55e-13
##
##             Sensitivity : 0.5149
##             Specificity : 0.5000
##          Pos Pred Value : 0.8455
##          Neg Pred Value : 0.1624
##              Prevalence : 0.8417
##          Detection Rate : 0.4333
##    Detection Prevalence : 0.5125
##       Balanced Accuracy : 0.5074
##
##        'Positive' Class : FALSE
##
```