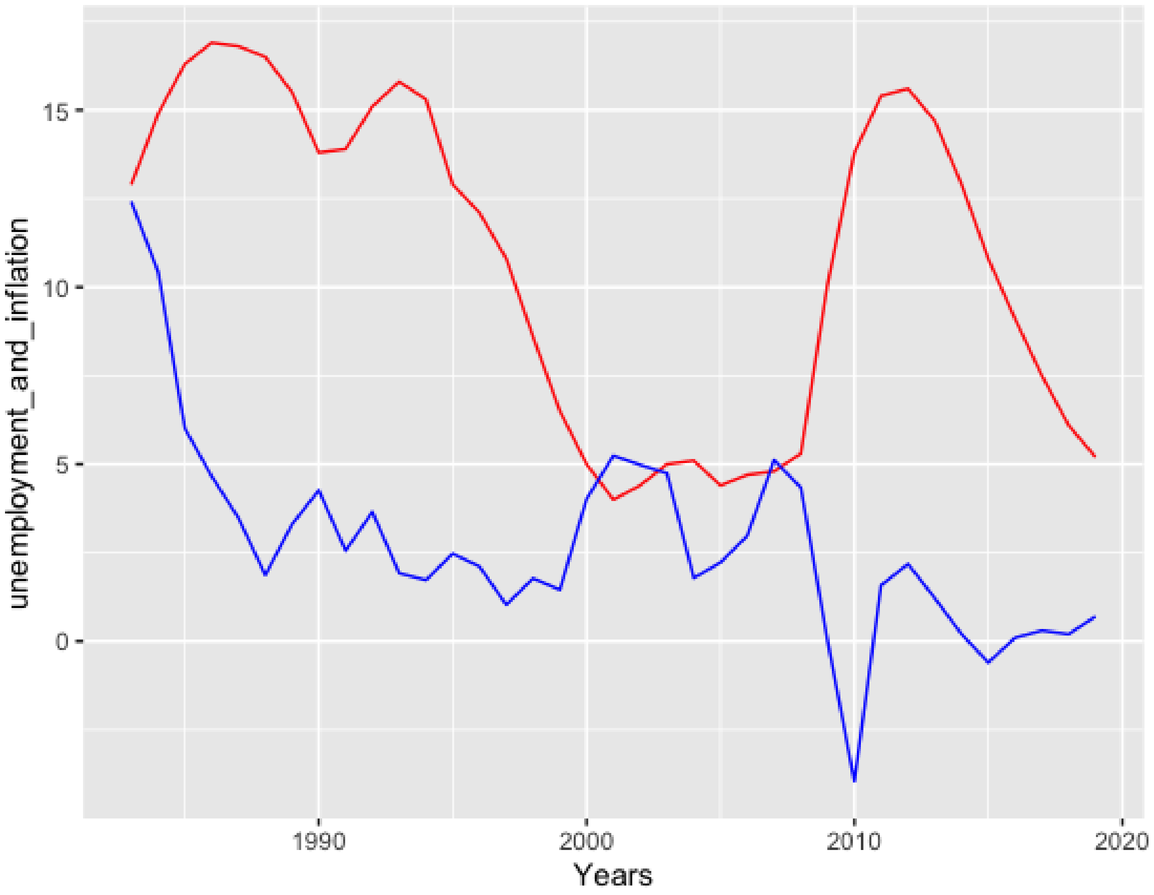
The Determinants of Inflation

Ireland

1983 - 2019

Piercesare Fagioli, Francesca Malfiore, Pietro Marini, Silvia Romanato

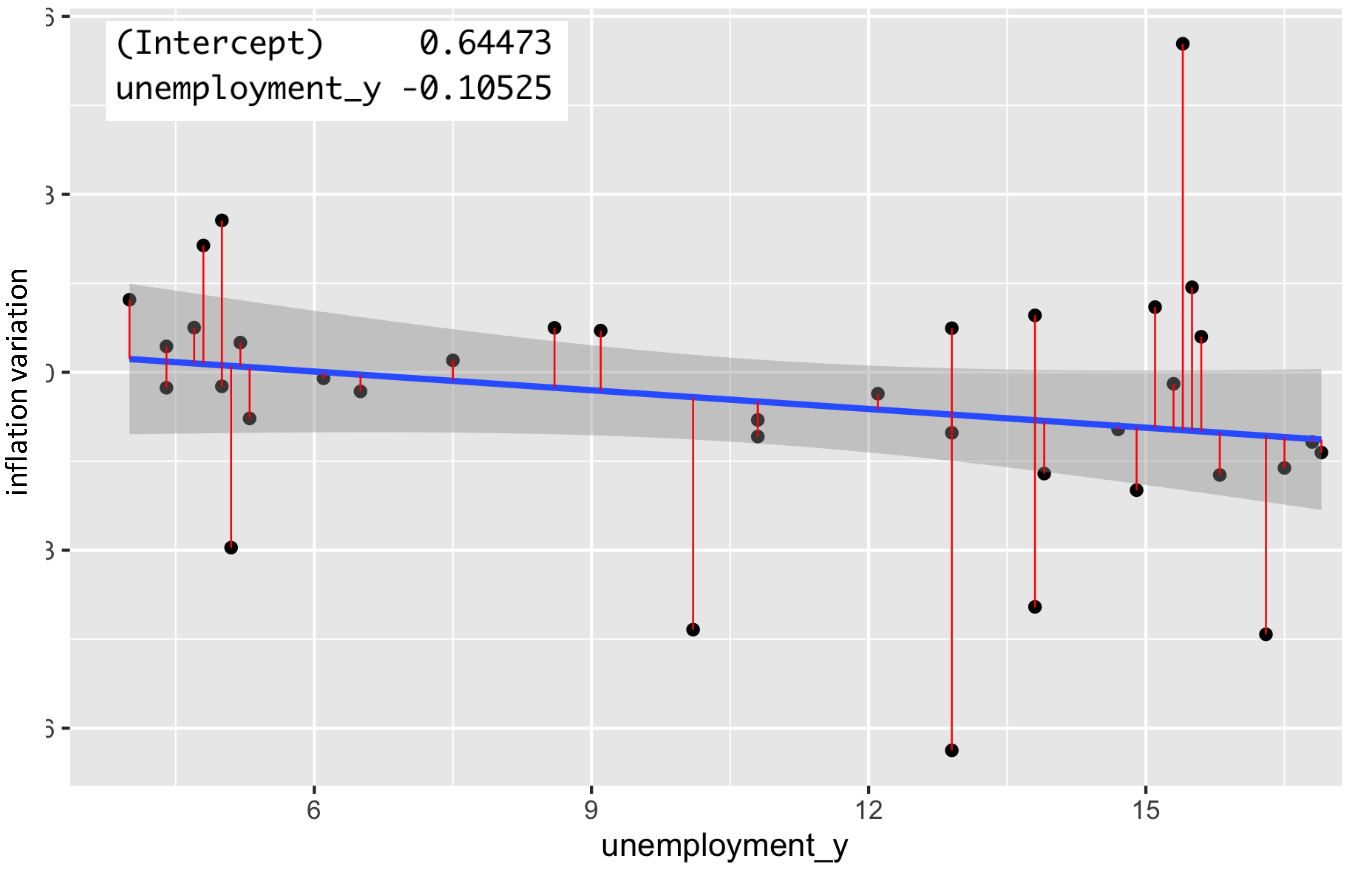
Variables Over Time



---: unemployment with yearly data between 1983 – 2019
---: inflation (plot 1) and change in inflation (plot 2) with yearly data

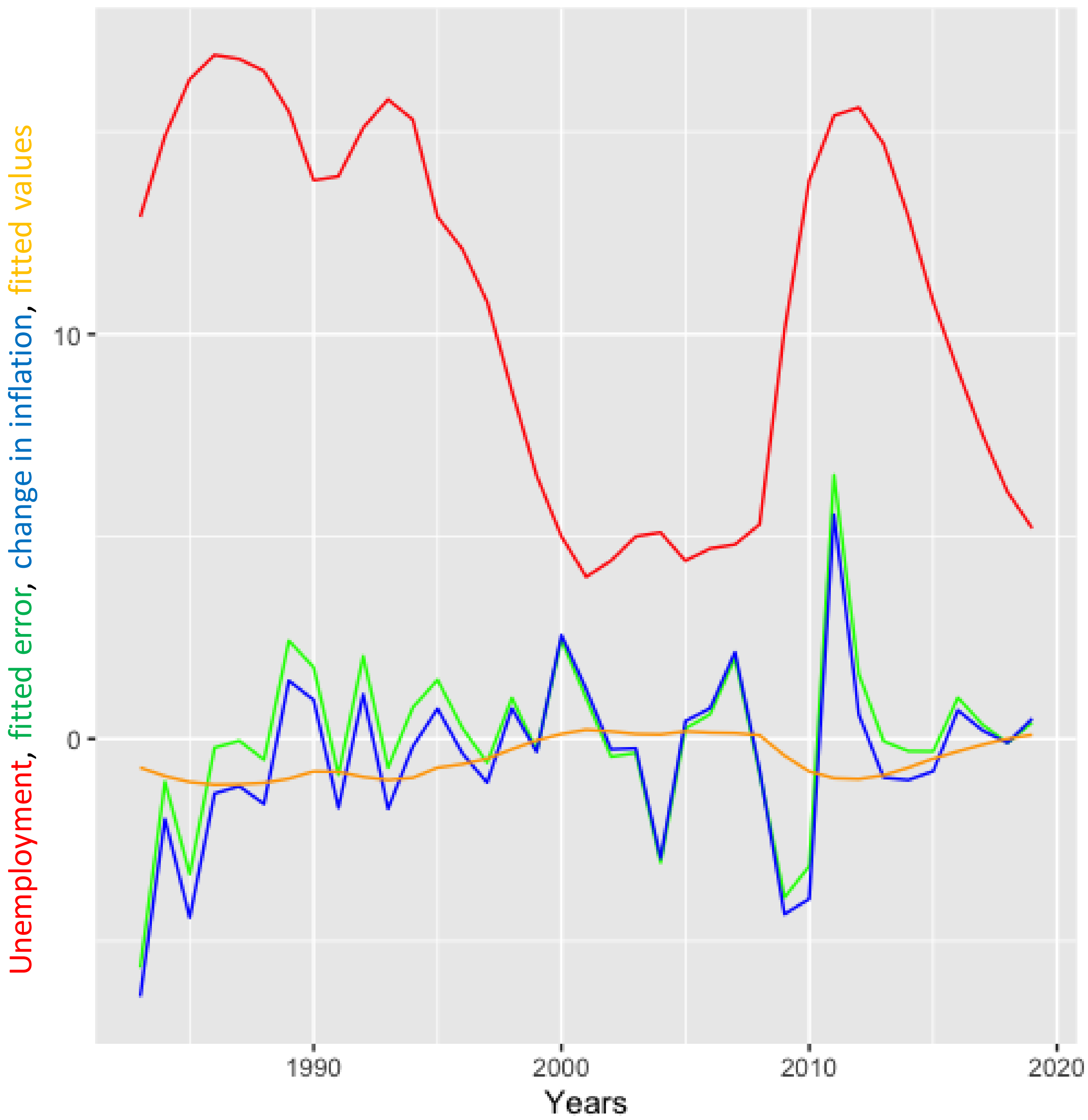
Linear Model: regressing Change in Inflation Year-on-Year (for each month) on Unemployment

- To begin with, we are going to analize yearly data (from 1983 to 2019).
- This is our linear regression: the blue line clearly represents the negative linear relationship between unemployment (on the x axis) and the variation in inflation (our proxy for expected inflation): our Phillips Curve.
- The dots are the realization of the random variable we are studying, so that the vertical distances (in red) from the fitted model (in blue) are the realizations of the errors: the residuals.
- The gray surface surrounding the main blue regression line represents the 95% confidence interval (α = 5%) in which our econometric model expects the realization of the random variable to fall in.
- As we will see, the 37 data points (one per year) provide a general overview of the main macro-trends the economy has faced in the last three decades. However, these are too scarce to give significance to the explanatory variable of our Phillips Curve.



Red: unemployment
Green: fitted error
Blue: inflation variation
Orange: fitted values
Plotted over time

Here, with the aid of a line graph, we can extrapolate the results about the fitted errors and the fitted values obtained with the preceding model, and plot them over time alongside with our variables.



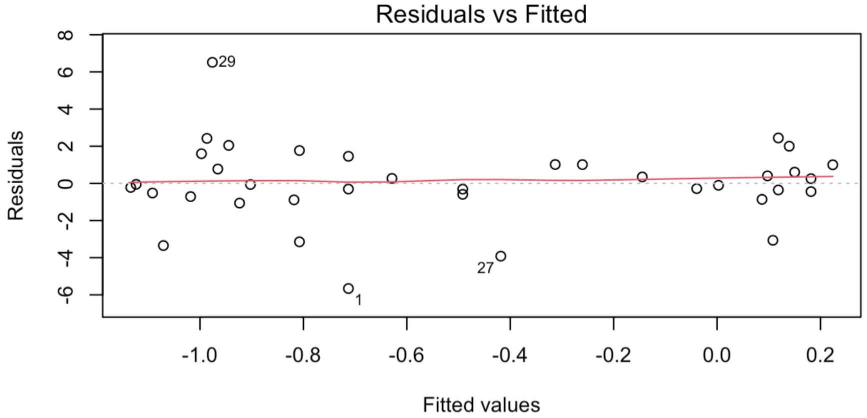
Yearly Inflation Model Summary

```
Residuals:
    Min      1Q  Median      3Q     Max 
-5.6613 -0.5916 -0.0598  1.0101  6.5154 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.64473    0.89640   0.719    0.477
unemployment_y -0.10525    0.07669  -1.372    0.179

Residual standard error: 2.118 on 35 degrees of freedom
Multiple R-squared:  0.05106,	Adjusted R-squared:  0.02395 
F-statistic: 1.883 on 1 and 35 DF,  p-value: 0.1787
```

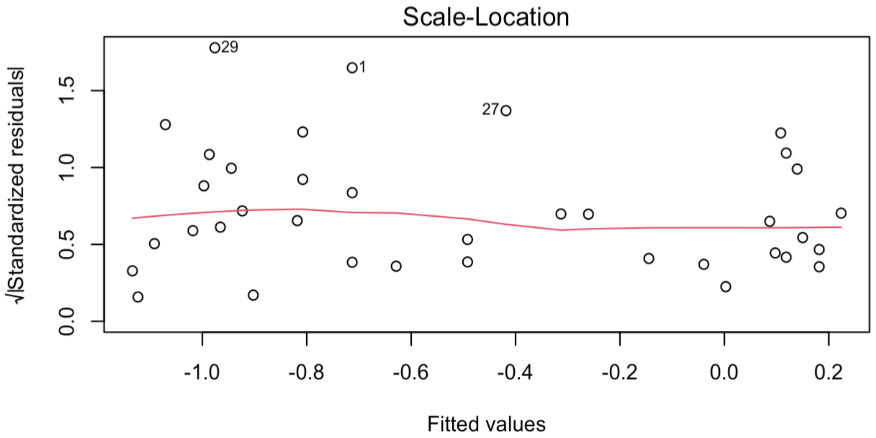
As we have already stated before, the Irish yearly data about inflation and unemployment seem not to have significance in the model we are considering.

In fact, the p-values (*Pr(>|t|)* *in the R-summary function*) are very high; remember we are performing Hypothesis Testing on the null hypotesis H0 that our regressors are not significant! Thus, with such high p-values, we do not reject H0, i.e. we assume that our regressors are not significant and so our model is *very imprecise and far from reality.*

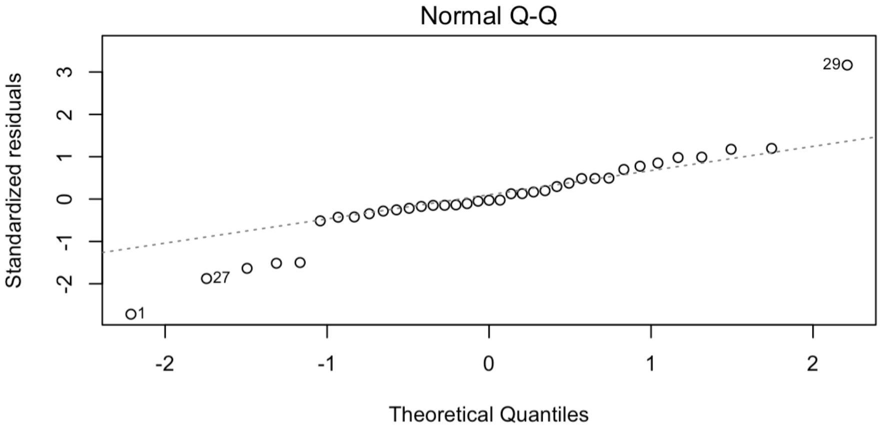
Graphical Tests



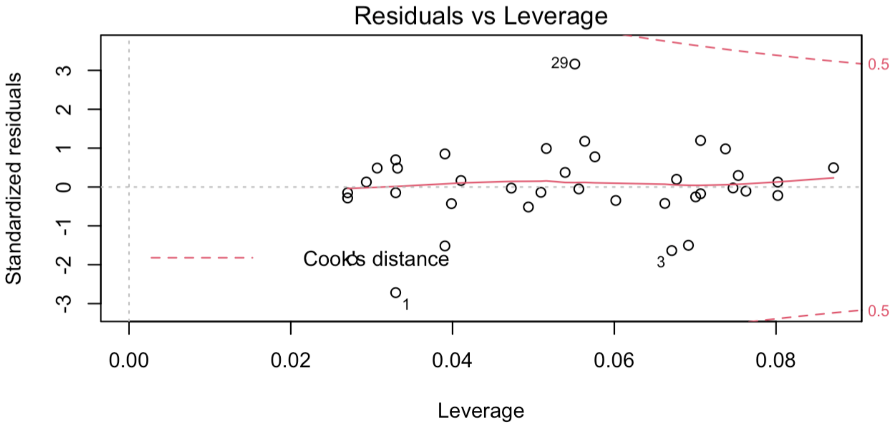
Ideally, the residual plot should show no fitted pattern. That is, the red line should be approximately horizontal at zero. The presence of a pattern may indicate a problem with some aspect of the linear model. This is OK!



This plot shows if residuals are spread equally along the ranges of predictors. It's good if you see a horizontal line with equally spread points. This is OK!

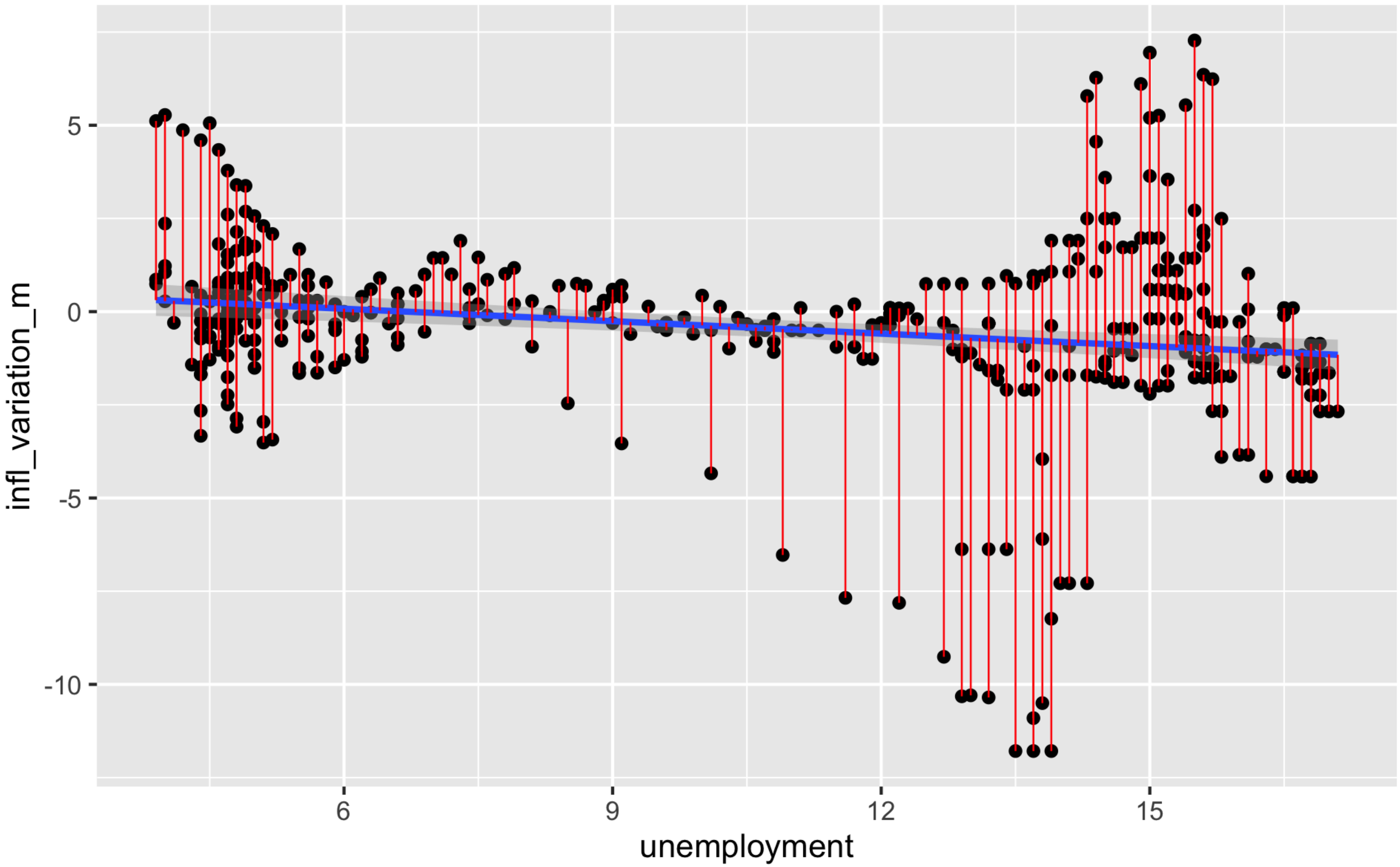


The QQ plot of residuals can be used to visually check the normality assumption. The normal probability plot of residuals should approximately follow a straight line. This is OK!



Used to identify influential cases, that is extreme values that might influence the regression results when included or excluded from the analysis. This test in particular is performing particularly badly.

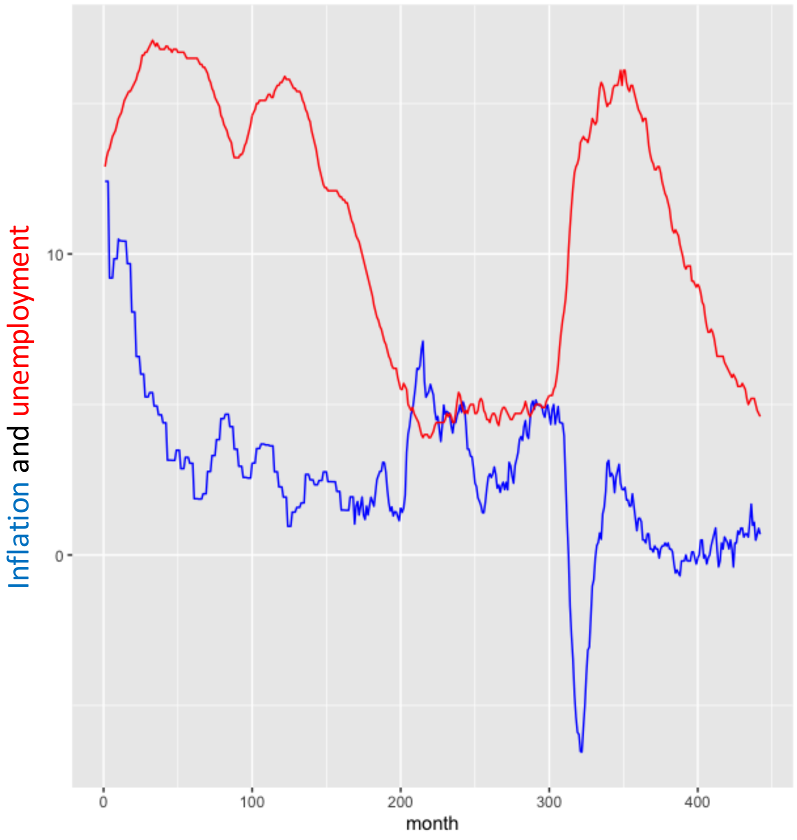
Linear Model: regressing Change in Inflation Year-on-Year (for each month) on Unemployment



In order to achieve proper significance for our model, we continue by considering monthly data.

This is our new linear regression. On the x axis we still find the unemployment and on the y axis the *delta*-inflation. The blue line (the regression) still represents the negative relationship between our two variables and the red lines are the residual errors. *We note how much the confidence interval has shrinked compared to the first linear regression since, with more data, uncertainty diminishes.* We will check this in the regression R-summary.

Perfectioning our visualization of inflation and unemployment using monthly data



- Continuing to analyze the <u>monthly</u> data:
- We now have much more data available and, by just even looking at the graphs in the first slides (*with particular with reference to slide 2*), we can note how much more precise we are with this new model!

Monthly Inflation Summary

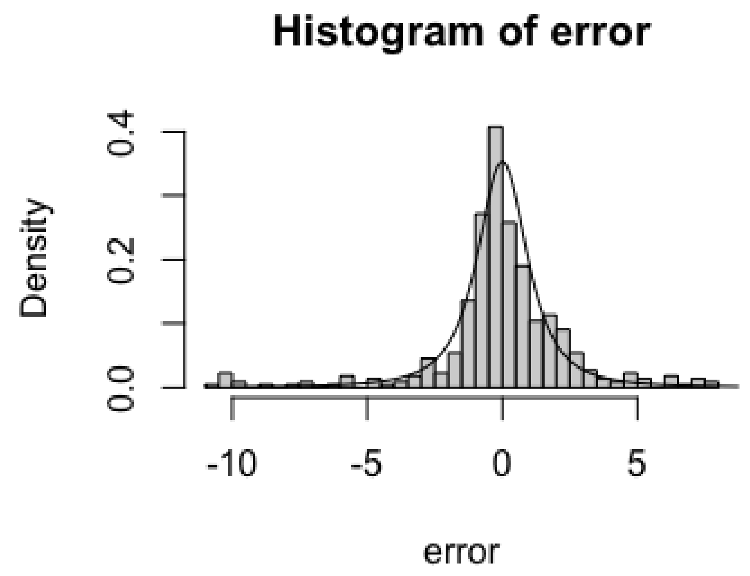
```
Residuals:
     Min       1Q   Median       3Q      Max 
-11.0309  -0.6999   0.0720   0.9228   8.2544 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)   0.74443    0.30610   2.432   0.0154 *  
unemployment -0.11125    0.02628  -4.233 2.81e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.531 on 440 degrees of freedom
Multiple R-squared:  0.03912,	Adjusted R-squared:  0.03694 
F-statistic: 17.91 on 1 and 440 DF,  p-value: 2.814e-05
```

Here, we may notice how a greater amount of data can dramatically improve the relevance of our statistic. In fact, the p-values are quite low for the intercept (*, basically a little more than 1%), and very low for unemployment (***, basically zero). Thus we can reject the null hypothesis H0 that our regressor is not significant in the model!

Adding Linear Regressors: 10y YIELD

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.75540    0.30368   2.487  0.01324 *
unemployment -0.03568    0.03723  -0.958  0.33836
yield_10      -0.12767    0.04490  -2.844  0.00467 **
```

The 10 years yield is a regressor with pretty great significance, even though it makes unemployment lose significance itself...

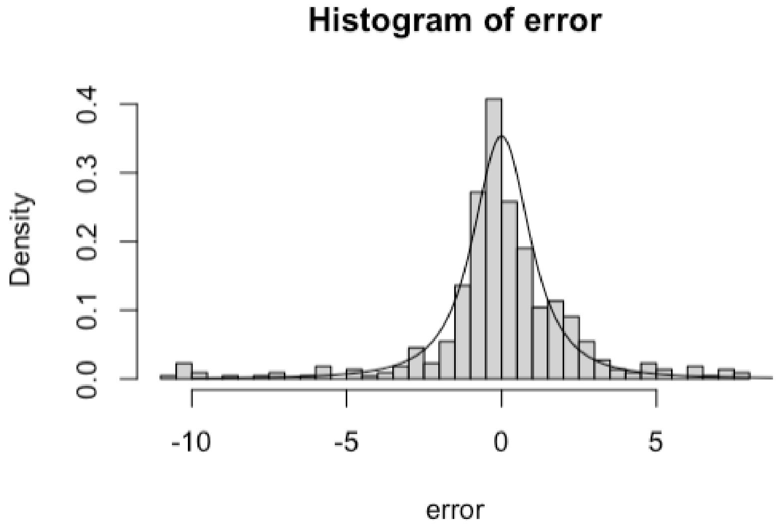


By *<u>forward selection</u>* we introduce the 10 years yield as additional explanatory variable.
This is because if the Treasury yields rises, so do the interest rates on consumer and business loans with similar lengths. Investors like the safety and fixed returns of bonds. Treasuries are the safest since they are guaranteed by the government. Other bonds are riskier. They must return higher yields in order to attract investors. To remain competitive, interest rates on other bonds and loans increase as Treasury yields rise.
When yields rise on the secondary market, the government must pay a higher interest rate to attract buyers in future auctions. Over time, these higher rates increase the demand for Treasuries. That's how higher yields can increase the value of the currency.

Adding an Alternative Linear Regressor: Share Prices Index (IRL)

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -3.984675   0.864128  -4.611 5.26e-06 ***
unemployment  0.132359   0.048918   2.706  0.00708 ** 
share_prices  0.035337   0.006068   5.823 1.12e-08 ***
```

Share Prices Index is a regressor with very high significance and it doesn't make the unemployment lose its own significance; for a merely *statistical* point of view, it seems a better regressor than the 10 years yield...

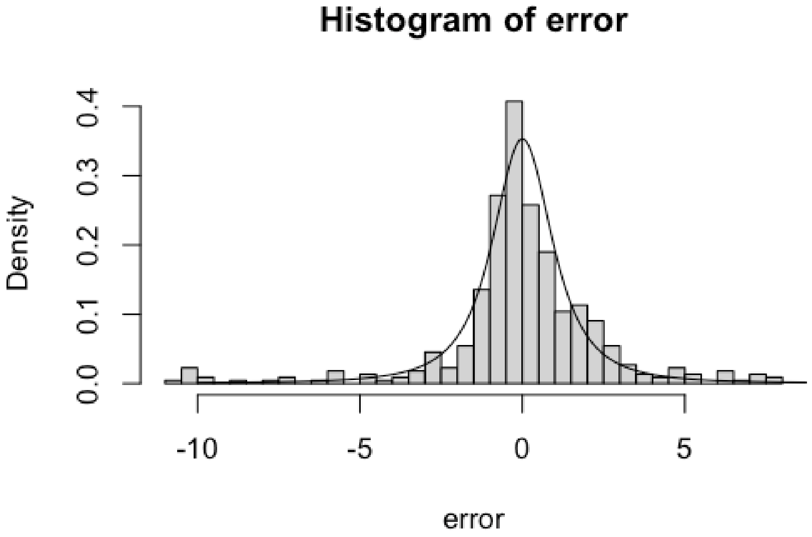


Here is the *economic* link and justification of why we introduced the Share Prices Index (i.e. a measure of changes in the market capitalisation of a basket of Irish shares): if the unemployment rate is higher, the general income (and therefore, cash to spend) will be limited. With less cash, people spend less money, and there's less demand for products. This can mean that stock prices can go down in many areas because there isn't as much of a demand for certain goods. Another reason is that investment in general in shares also diminishes.

Adding an Alternative Linear Regression: Changes in OIL PRICE

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.371567   0.292792   1.269   0.2051
unemployment -0.079378   0.025139  -3.158   0.0017 **
change_oil    0.051783   0.006908   7.496 3.67e-13 ***
```

Oil real price can be considered a *statistically* significant regressor.

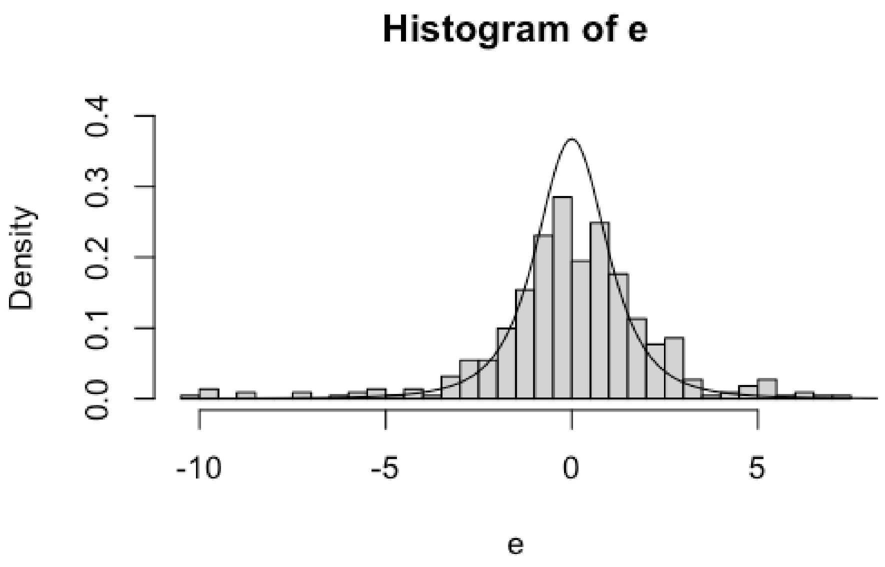


Oil and inflation are *economically* linked because oil is a major input in the economy—it is used in critical activities such as fueling transportation and heating homes—and if input costs rise, so should the cost of end products. For example, if the price of oil rises, then it will cost more to make plastic, and a plastics company will then pass on some or all of this cost to the consumer, which raises prices and thus creates inflation.

The full model with K-regressors

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -4.051303   1.055311  -3.839 0.000142 ***
unemployment  0.172529   0.046065   3.745 0.000204 ***
change_oil    0.054094   0.006829   7.921 1.97e-14 ***
share_prices  0.032949   0.007632   4.317 1.95e-05 ***
yield_10     -0.039421   0.056053  -0.703 0.482256    
```

In the joint model 3 out of 4 regressors keep statistical significance, a part from the 10 years yield that looses it.



Now, we consider all the regressors together in a model with k regressors (with k=4).

All together, 3 out of 4 regressors keep statistical significance, a part from the 10 years yield which loses significance. However, the model overall looks good with low p-values for the other regressors.

Testing the OLS assumptions (multi variate regression)

OLS TRIVIAL TEST:

- The covariance between the unemployment and the residuals is:

  COV (u, $\hat{\varepsilon}$) = -1.220603e-15 $\approx$ 0

- The expectation of the residual is approximately 0

  $\mathbb{E}\,[\hat{\varepsilon}] = -1.830499e - 16 \approx 0$

The covariance and the mean error should both be approximately 0.
Then this test is passed.

TESTS FOR SERIALLY CORRELATED ERRORS

- Durbin-Watson test:

  DW = 0.12103
  p-value < 2.2e-16

  BUT CAN'T USED IN DYNAMIC MODELS (we can't draw conclusions from that)

- Breusch-Godfrey test (for serial correlation of order up to 1):
  LM = 389.2
  p-value < 2.2e-16

Since the p-value is low, the H0 is rejected (uncorrelated errors)
and errors are serially correlated (LM > 0).
Thus. we will apply the HAC-robust standard errors.

TESTS FOR HOMOSCEDASTICITY:

- Goldfeld-Quandt test:
  GQ = 1.698, df1 = 216, df2 = 216, p-value = 5.576e-05

- Studentized Breusch-Pagan test:
  BP = 10.376, df = 4, p-value = 0.03455

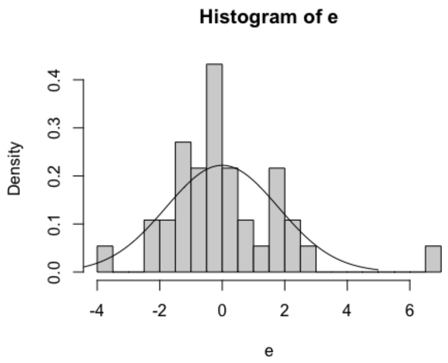
- White's Test:
  statistic 32.8
  p-value 0.0000662

All this tests have a low p-value then the homoscedasticity (null hypothesis) is rejected.
Then we have to apply heteroskedasticity-robust standard errors.

TEST FOR NORMALITY

- Jarque-Bera test:
  JB = 374.59, p-value < 2.2e-16

This test is not passed due to the value
of the JB statistic



Since in our multivariate regression model *heteroskedasticity* is an issue, we choose to adopt the *Heteroskedasticity-robust standard errors*.

In this way the coefficients of the regressors become:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-4.0513033	0.7212640	-5.6169	3.464e-08	***
unemployment	0.1725294	0.0475856	3.6257	0.0003222	***
change_oil	0.0540938	0.0101781	5.3147	1.706e-07	***
share_prices	0.0329493	0.0048485	6.7957	3.543e-11	***
yield_10	-0.0394213	0.0607704	-0.6487	0.5168779	

Comparing it to before, the model lost some significance but we don't have serially correlated errors.

Since in our multivariate regression model *serial correlation* is an issue, we choose to adopt the *HAC-robust standard errors*.

In this way the coefficients of the regressors become:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-4.051303	2.107267	-1.9225	0.0551876	.
unemployment	0.172529	0.104953	1.6439	0.1009232	
change_oil	0.054094	0.014867	3.6386	0.0003068	***
share_prices	0.032949	0.012277	2.6837	0.0075572	**
yield_10	-0.039421	0.167707	-0.2351	0.8142723	

Comparing it to before, the model lost some significance but we don't have serially correlated errors.

Thank you for your kind attention

The Ireland team:

Piercesare Fagioli
Francesca Malfiore
Pietro Marini
Silvia Romanato