

Have UK's electricity prices become to dependent on Natural Gas?

Energy Industry Analytics

Candidate number: 28

02/06/2022

Abstract

As the world is transitioning to renewable energy sources, it is important to understand the implications. This report aims to highlight the United Kingdom's increased dependence on Natural Gas as Coal is reduced, and Renewable Energy grows. Specifically, I investigate the effect Natural Gas price has on electricity prices through a Multivariate Regression model and a Granger Causality test. I also investigate whether Electricity production from Gas and Wind influences electricity price volatility. To end the report I use an ARIMA forecast with three different scenarios to show the development in electricity price for a given development in the Gas price. The main results is that gas prices significantly affect electricity prices, both the multivariate regression and the Granger Causally test gives this result. The analysis on the price volatility did not give any conclusive results. The implications of the dependence on natural gas for the UK is that if a large part of the natural gas becomes unavailable because of embargos on Russia then the UK have to pay more for natural gas or turn back to coal for a stable energy source to partner renewable energy.

Contents

1 Introduction	2
1.1 Motivation	2
1.2 Short Review of Literature	3
1.3 Data	3
1.4 Background	4
2. Energy in the United Kingdom	7
2.1 United Kingdom's Energy Mix	7
2.2 Electricity and Natural Gas in the UK	8
2.3 UK's dependency on Norwegian Natural Gas	9
3. Cleaning of Data	11
4. Correlation	13

5. Regression	15
5.1 Variables	15
5.2 Regression in levels	17
5.3 Stationarity	18
5.4 Regression in differences	20
5.5 Autocorrelation	21
5.6 Cochrane-Orcutt	22
5.7 Homoskedasticity	22
5.8 Overfitting - Regularisation	24
5.9 Granger Causality	25
6. Garch Model of Natural Gas and Electricity price	26
6.1 Electricity Return	26
6.2 Natural Gas return	29
6.3 Gas, Wind, Nuclear and Coal production's effect on electricity price volatility	31
7. ARIMA-model	33
8. Conclusion	37
8.1 Summary of results	37
8.2 Weaknesses of study	37
References	37
Sources	37
Data	39

1 Introduction

1.1 Motivation

My motivation behind this project is two folded. I wanted to learn more about the electricity market as it is gaining importance today and facing several challenges. Second, I want to learn more about Natural gas because it have an increasingly important role in Energy markets and Norway is a major exporter. Norway covers more than 20% of European gas consumption and is Europe's second-largest supplier. Awareness of natural gas has increased because of increased prices the last year and the Russian invasion of Ukraine. In the upcoming summer, I will have an Internship at the Norwegian Energy company Equinor. Equinor mainly have Oil and Gas in their portfolio, but also have an ambition of becoming an Offshore Wind major, which is another reason for wanting to focus on natural gas and the implications of more renewable energy in this report.

As the world are transitioning to clean energy combined with growing energy demand, the energy markets are going through changes, where renewable energy are increasingly replacing fossil fuels such as coal, oil and natural gas. Prices for electricity and gas have soared in Europe the last year and the increased energy-prices are one of the main reasons the world are battling inflation. The fast rebound in overall energy demand strained supply chains for coal and natural gas, pushing up electricity prices.

1.2 Short Review of Literature

Recent literature has highlighted the growing relationship between natural gas and electricity (Alexopoulos, T.A. (2017)). This study finds that natural gas is a critical determinant of electricity prices. Other literature has investigated the effects the changing energy mix have on wholesale power prices in the United States. (Mills et al., 2021) This study finds that wind and solar have contributed to lower average annual energy prices, however also finds natural gas prices as the dominant driver.

These studies have similar goals as my report, but the focus is on the US market, while mine is on the UK market.

1.3 Data

For this report I will make use of several datasets, including electricity, gas, oil, coal and carbon prices as well as export and import data of Gas for Norway and the United Kingdom. I will also use datasets for the daily Energy generation by source in the UK and monthly weather data.

```
sources2021 = read.csv("Source2021.csv")
sources2020 = read.csv("Source2020.csv")
sources2019 = read.csv("Source2019.csv")
sources2018 = read.csv("Source2018.csv")
sources2017 = read.csv("Source2017.csv")
sources2016 = read.csv("Source2016.csv")
sources2015 = read.csv("Source2015.csv")
sources = rbind(sources2015,sources2016,sources2017,sources2018,sources2019,sources2020,sources2021)

overview_sources = read.xlsx("overview.xlsx")

GasExpNo = read.xlsx("~/Studier/NHH/V2022/ENE434/Project/Norway_GasExp.xlsx")

GasImports = read.xlsx("GasImportUK.xlsx")
GasImportsY = read.xlsx("GasImportUKYear.xlsx")

Cons2021 = read.csv("Cons2021.csv")
Cons2020 = read.csv("Cons2020.csv")
Cons2019 = read.csv("Cons2019.csv")
Cons2018 = read.csv("Cons2018.csv")
Cons2017 = read.csv("Cons2017.csv")
Cons2016 = read.csv("Cons2016.csv")
Cons2015 = read.csv("Cons2015.csv")

coalProd = read.xlsx("CoalProduction.xlsx")
RenewProd = read.xlsx("Renewable.xlsx")

price2022 = read.csv("EL_Price_2022.csv")
price2021 = read.csv("EL_Price_2021.csv")
price2020 = read.csv("EL_Price_2020.csv")
price2019 = read.csv("EL_Price_2019.csv")
price2018 = read.csv("EL_Price_2018.csv")
price2017 = read.csv("EL_Price_2017.csv")
price2016 = read.csv("EL_Price_2016.csv")
price2015 = read.csv("EL_Price_2015.csv")
```

```

GasPriceD = read.csv("GasPriceD.csv")

GasPriceM = read.csv("GasPriceM.csv")

Carbon = read.csv("CarbonPriceD.csv")

oilprice= read.xlsx("OilPrice.xlsx")
Coal = read.csv("CoalPrice.csv")

Weather = read.xlsx("WeatherUK.xlsx")

```

1.4 Background

The United Kingdom has an ambitious energy transition plan and is leading the way for the rest of Europe (Rystadenergy, 2022). The United Kingdom is the country in Europe with the second largest consumption of natural gas, only behind Germany. Most of the natural gas is either produced in the North Sea (majority Norway) or imported as LNG from countries such as the US and Qatar. UK uses natural gas to generate electricity and for heating. The Russian invasion of Ukraine have showed Europe's dependence of Natural gas, with several countries refusing to sanction Russian gas since it will have too great of an impact on their economy. The United Kingdom is not as dependent on Russia as the rest of Europe for its natural Gas, with less than 4% of their Natural Gas imported from Russia (GOV.UK,2022). This does not mean that the UK are not affected by the Russian Invasion, the natural gas prices have risen dramatically and as a result of this Energy prices have become more volatile.

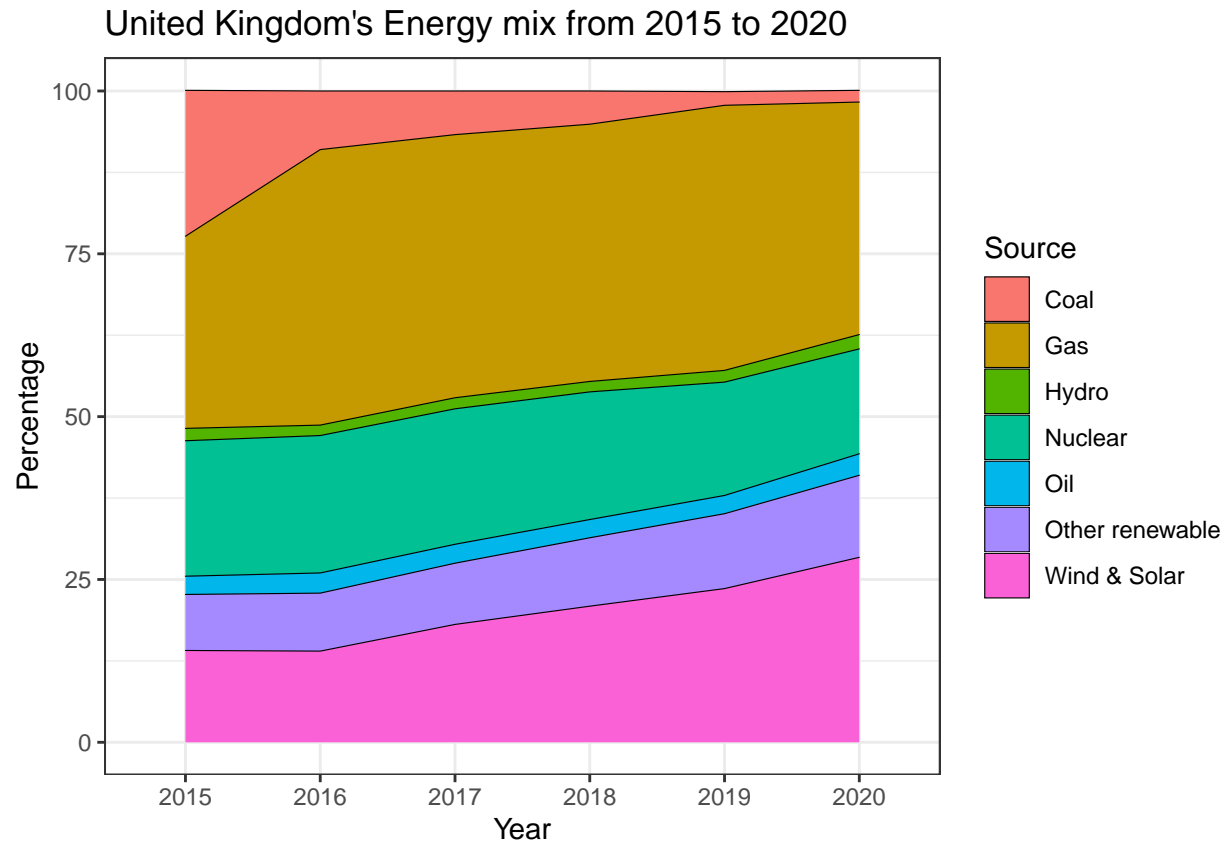
We have seen a shift in the Energy mix, characterized by the fact that Natural Gas is now seen as the most environmentally friendly fossil energy source and Coal has largely been replaced by gas for Power generation. I want to research the UK because as seen in the graph below their two main energy sources are Natural Gas and Renewable Energy.

```

colnames(overview_sources) = c("Year", "Gas", "Wind & Solar", "Nuclear", "Hydro", "Coal",
                                "Oil", "Other renewable")
overview_sources_long = overview_sources %>%
  pivot_longer(cols="Gas":"Other renewable", names_to="Source",
               values_to = "Percentage")

ggplot(overview_sources_long, aes(x=Year, y=Percentage, group=Source, fill=Source))+
  geom_area(alpha=1, size=0.2, colour="black")+
  ggtitle("United Kingdom's Energy mix from 2015 to 2020")+theme_bw()

```

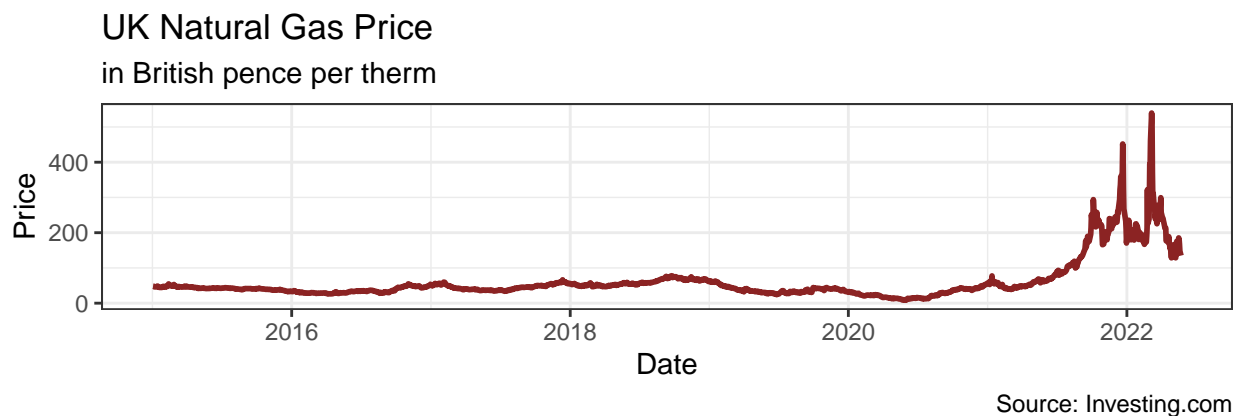
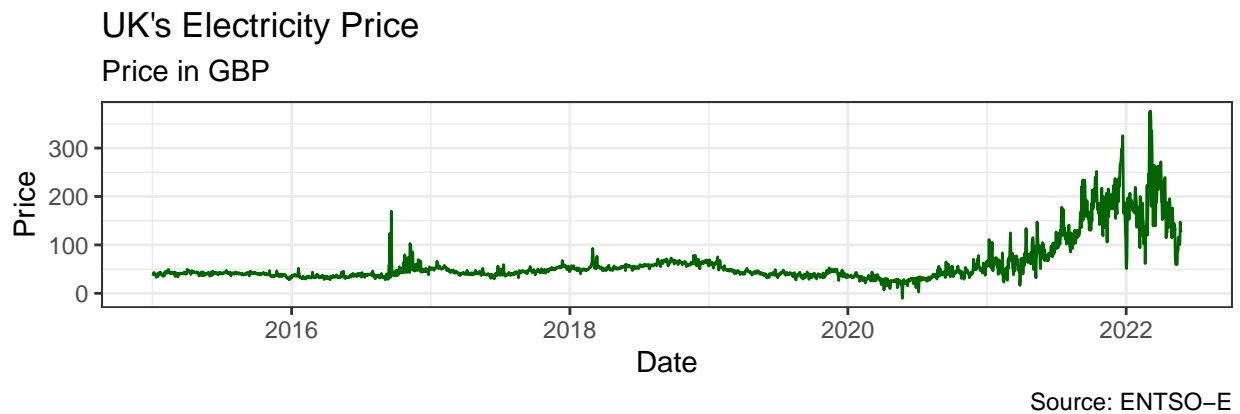


Below we see the surge in UK Natural Gas price and the UK electricity price. The natural gas prices are shown in British Pence per therm and the electricity price in Megawatt.

```
EL_price = rbind(price2015,price2016,price2017,price2018,price2019,price2020,price2021,price2022)
colnames(EL_price) = c("Time","Price","Currency","BZN.GB")
EL_price = EL_price %>% select(Time,Price)
EL_price[c("Date","x")] = str_split_fixed(EL_price$Time," ",2)
EL_price[c("Hour","y")] = str_split_fixed(EL_price$x,"-",2)
EL_price$Date = as.Date(EL_price$Date,format = "%d.%m.%Y")
EL_price = EL_price %>% select(Price,Date,Hour)
# See that there are a few 0's in the data,
#so I will replace them with the previous value.
EL_price[EL_price == 0] = NA
EL_price = fill(EL_price,Price,.direction = "down")
EL_price = EL_price %>% group_by(Date) %>%
  summarize(
    Price = mean(Price)
  )
EL_price = EL_price[2:2702,]
EPlot = ggplot(EL_price,aes(x=Date,y=Price)) + geom_line(color="darkgreen")+
  labs(
    title="UK's Electricity Price",
    subtitle = "Price in GBP",
    caption = "Source: ENTSO-E"
  )+ theme_bw()
```

```
#Daily
GasPriceD = read.csv("~/Studier/NHH/V2022/ENE434/Project/GasPriceD.csv")
GasPriceD$Date = as.Date(GasPriceD$Date,format = "%m/%d/%Y")
GasPriceD = GasPriceD %>% select(Date,Value)
colnames(GasPriceD) = c("Date","Price")
GasPriceRecent = GasPriceD[4528:6393, ]
GPlot = ggplot(GasPriceRecent,aes(x=Date,y=Price))+
  geom_line(color="brown4",size=1)+
  labs(
    title="UK Natural Gas Price",
    subtitle = "in British pence per therm ",
    caption = "Source: Investing.com"
  )+ theme_bw()
```

```
grid.arrange(EPlot,GPlot,ncol=1)
```



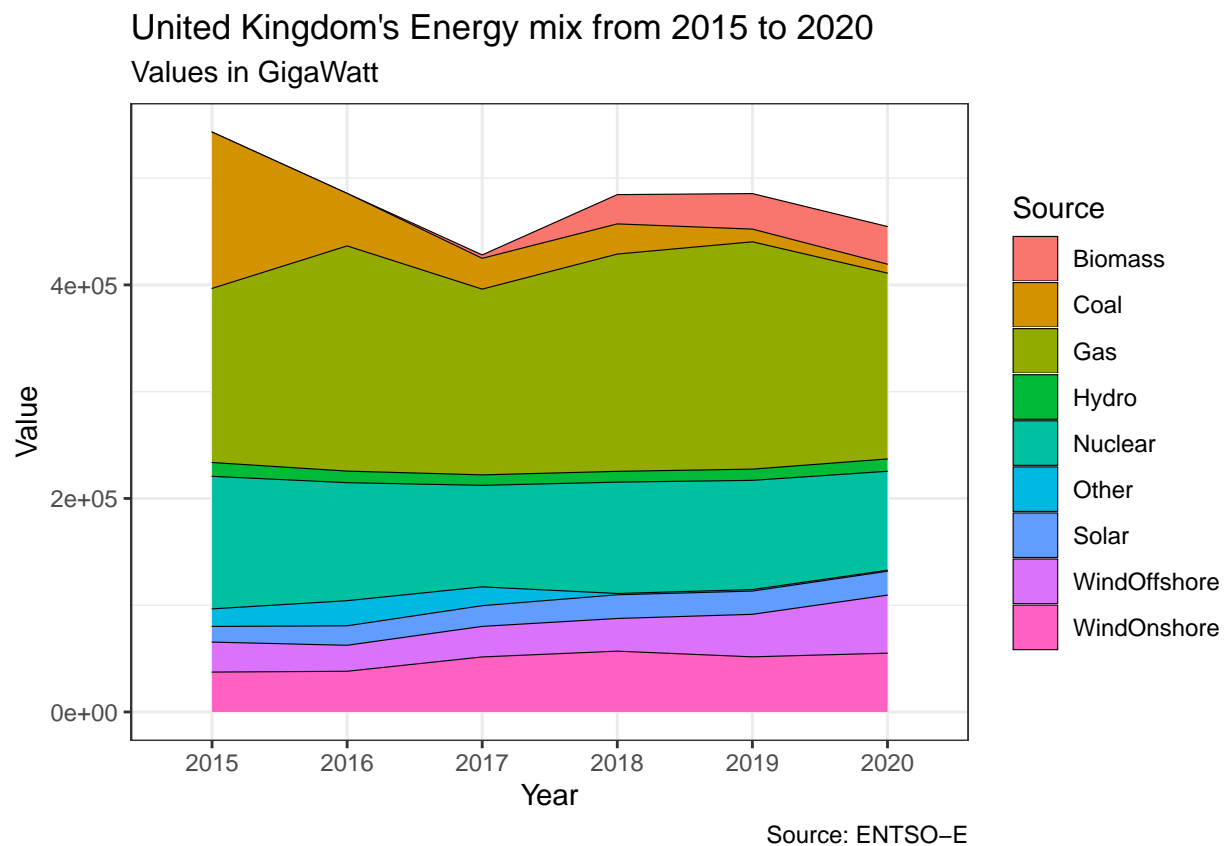
The charts show why it is interesting to focus this report on electricity and gas prices.

To answer if the UK have become to dependent on natural gas this report will try to address two research questions; Can changes in electricity prices mainly be explained by changes in the Natural Gas prices? Can Electricity production from Wind and Gas explain volatility in the electricity prices?

2. Energy in the United Kingdom

2.1 United Kingdom's Energy Mix

```
ggplot(sources_year,aes(x=Year,y=Value,group=Source,fill=Source)) +  
  geom_area(alpha=1,size=0.2,colour="black")+  
  labs(  
    title="United Kingdom's Energy mix from 2015 to 2020",  
    subtitle = "Values in GigaWatt",  
    caption = "Source: ENTSO-E"  
  )+theme_bw()
```



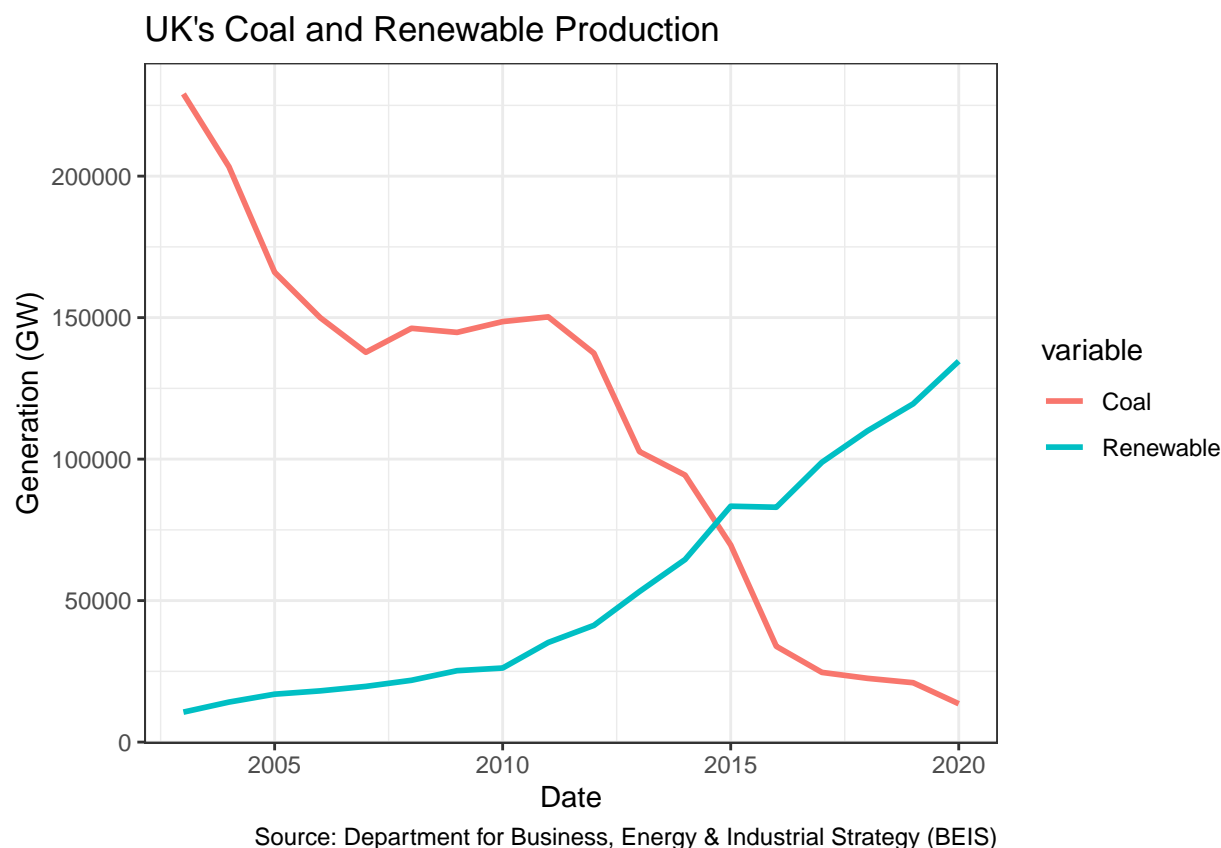
We notice three things about this interesting chart. First, we see that Natural Gas plays a big part in UK's energy mix. Second, we notice the phasing out of coal and third, there is a somewhat downward trend in total energy produced.

The UK's progress on lowering emissions has revealed vulnerabilities in their Energy system as the country transitions to net zero.

1. The UK has become very reliant on Natural Gas. This is correlated with the fall in coal generation in the UK over recent years. The coal reduction is a key reason why the UK has come a good way to net zero. Gas stands for over 35% of the energy generation in the UK.
2. The second is the increased reliance on renewables which varies in their output. The largest is Wind, which varies with wind speed. They also have solar which varies with sun hours. From the graph with UK's Energy mix, we see that Wind & Solar is over 25% of the UK's energy generation.

UK electricity price volatility is unlikely to go away as the energy transition continues towards more dependence on wind and solar. From the graph below we see that the UK have done well in removing Coal which is critical in reducing Emissions. The reduction in Coal has been replaced mainly by renewables and increase in natural Gas and the effect for the UK is larger dependence on Natural Gas and possibly more volatile electricity prices through intermittent renewable energy. Coal is a stable and secure energy source and replacing it with unstable energy sources are likely to change the price dynamics in the energy markets.

NULL



2.2 Electricity and Natural Gas in the UK

Electricity prices are mainly influenced by weather (changes demand) and the cost of the commodities generating the electricity. This can be Coal, Natural Gas prices and costs of producing Nuclear and Renewable energy.

Electricity prices are impacted heavily by rising gas prices, because of the importance of gas in electricity generation and gas-use in households for heating etc. Natural Gas is the most significant driver of the increase in electricity prices in the United Kingdom (Clean Energy Wire, 2022).

Electricity prices are also affected by the cost of carbon allowances. The carbon prices have increased significantly over the winter 2021. (Ofgem, 2022). Another factor that influences the electricity prices is renewable energy generation, which as seen in the figure above is mostly wind generation. More wind is usually associated with lower electricity prices, *ceteris paribus*. The varying generation of renewables can make electricity prices more volatile compared to fossil fuels like Coal and Natural Gas. The increase in renewables combined with decrease in Coal, makes the UK's reliance on Natural Gas larger.

Natural gas prices are affected by three main supply-side factors:

- gas production
- gas storage
- the level of imports and exports.

And three main demand-side factors:

- Variations in weather (winter and summer)
- Economic growth
- Availability and prices of other fuels. Large consumers can switch between gas, coal and petroleum, depending on the cost.

In short, the reason for the increase in Natural Gas prices the last year are: Low gas storage levels in Europe, lower imports from Russia to Europe and this leads to higher prices to be able to meet demand.

2.3 UK's dependency on Norwegian Natural Gas

Norway is one of the largest Natural Gas exporters in the world and this is often in the shadow of the Country's oil exports. While Norway is a small player in the world's oil markets, they are the third largest exporter of natural gas, only behind Russia and Qatar (Norwegianpetroleum.no, 2022).

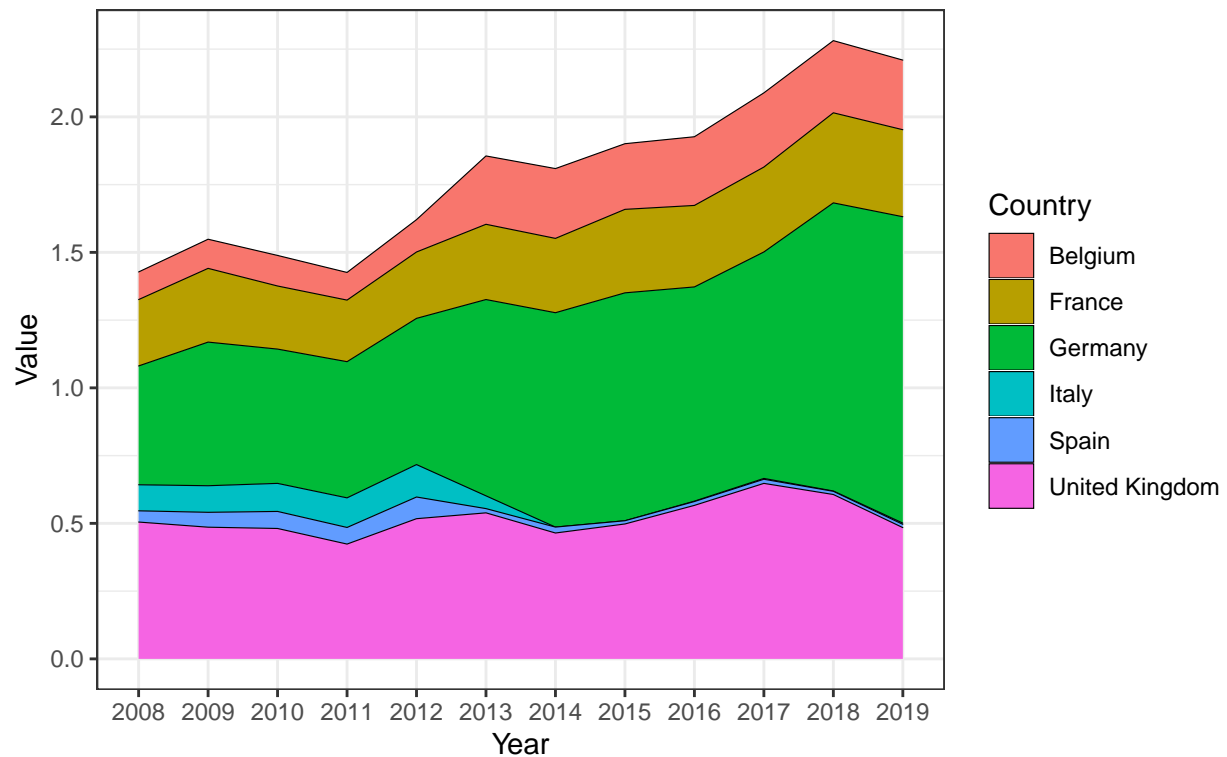
The figure below shows that Norway provides Europe with a lot of natural Gas. The two largest exports are Germany and United Kingdom.

```
GasExpNo = read.xlsx("~/Studier/NHH/V2022/ENE434/Project/Norway_GasExp.xlsx")
GasExpNo[c("Year", "Month")] = str_split_fixed(GasExpNo$Date, "M", 2)
#from TeraJoules to Mbo/e
GasExpNoYOilEq = GasExpNo %>% group_by(Year, partner) %>% summarise(
  Value = sum(Value, na.rm = TRUE)/6117.86/365
)
```

```
## 'summarise()' has grouped output by 'Year'. You can override using the
## '.groups' argument.
```

```
colnames(GasExpNoYOilEq) = c("Year", "Country", "Value")
GasExpNoYOilEq = GasExpNoYOilEq[!(GasExpNoYOilEq$Country=="Total"),]
ggplot(GasExpNoYOilEq, aes(x=Year, y=Value, group=Country, fill=Country))+
  geom_area(alpha=1, size=0.2, colour="black")+
  labs(
    title="Norway's Natural Gas Exports (Million Barrels of Oil Equivalent per day)",
    caption = "Source: Eurostat")+theme_bw()
```

Norway's Natural Gas Exports (Million Barrels of Oil Equivalent per day)



Source: Eurostat

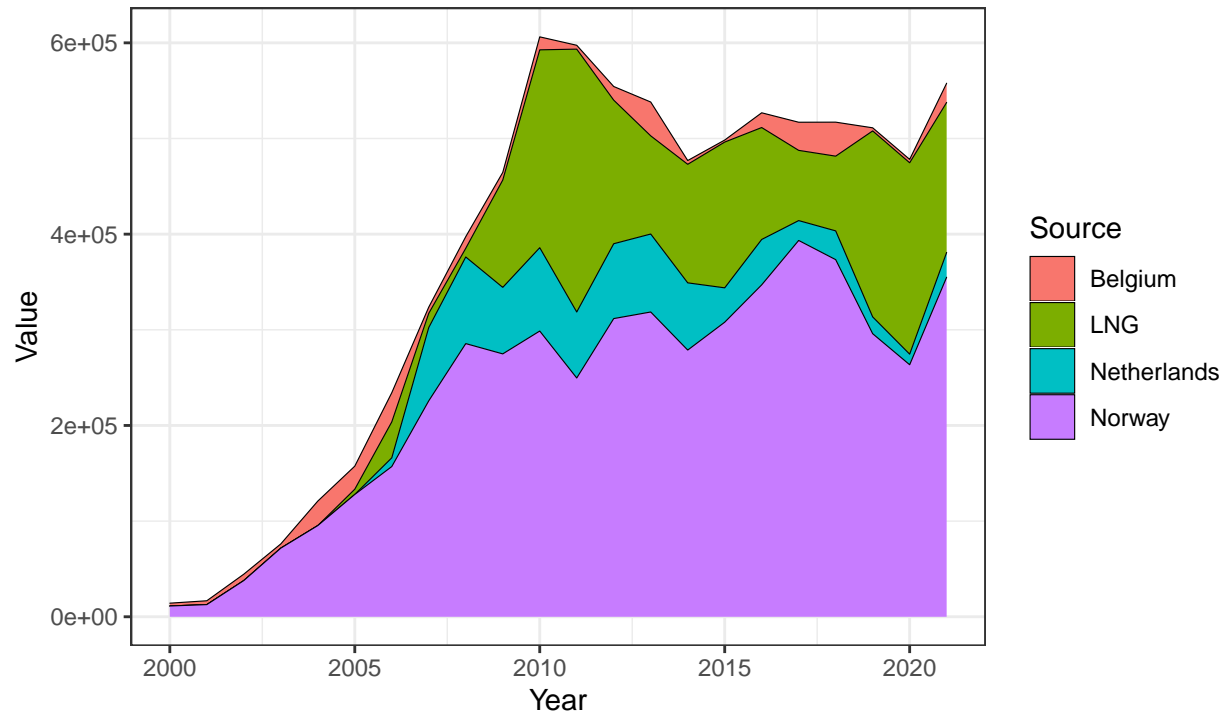
As we can see, Norway are a major Natural Gas exporter and the second largest export destination is the United Kingdom. The demand for Norwegian Gas will likely soar after Russia's invasion of Ukraine.

```
GasImportsY = read.xlsx("GasImportUKYear.xlsx")
UKimportDivided = GasImportsY %>% select(Year,Norway,Belgium,Netherlands,LNG)
UKimportDivided = UKimportDivided %>% pivot_longer(cols="Norway":"LNG",names_to = "Source",values_to = "Value")

ggplot(UKimportDivided,aes(x=Year,y=Value,group=Source,fill=Source))+
  geom_area(alpha=1,size=0.2,colour="black")+
  labs(
    title="UK's Natural Gas Imports",
    subtitle = "Values in GigaWatt",
    caption = "Source: Department for Business, Energy and Industrial Strategy (BEIS)"
  )+theme_bw()
```

UK's Natural Gas Imports

Values in GigaWatt



Source: Department for Business, Energy and Industrial Strategy (BEIS)

This figure shows where the United Kingdom imports its oil from, and we see that Norway is the largest part of the imports. When it comes to Natural Gas excluding LNG, almost all UK's gas import comes from Norway.

3. Cleaning of Data

In the next section I clean the datasets so it is easier to use later.

```
sources1 = sources
sources1$Wind = sources1$WindOffshore+sources1$WindOnshore
sources1 = sources1%>% select(Time,Gas,Coal,Nuclear,Wind)

sources1[c("Date","x")] = str_split_fixed(sources1$Time," ",2)
sources1[c("Hour","y")] = str_split_fixed(sources1$x,"-",2)
sources1$Date = as.Date(sources1$Date,format = "%d.%m.%Y")
sources1 = sources1%>%select(Gas,Coal,Nuclear,Wind,Date,Hour)
#values in gigawatt
sources1$Gas = sources1$Gas/1000
sources1$Nuclear = sources1$Nuclear/1000
sources1$Coal = sources1$Coal/1000
sources1$Wind = sources1$Wind/1000
sources1 = sources1 %>% group_by(Date) %>% summarise(
  Gas = sum(Gas,na.rm = TRUE),
  Coal = sum(Coal,na.rm=TRUE),
  Nuclear = sum(Nuclear,na.rm=TRUE),
```

```

    Wind = sum(Wind,na.rm=TRUE)
  )
sources1$Month = as.yearmon(sources1$Date,"%Y-%m")
sources1 = sources1 %>% group_by(Month) %>% summarise(
  Gas = sum(Gas,na.rm = TRUE),
  Coal = sum(Coal,na.rm = TRUE),
  Nuclear = sum(Nuclear,na.rm = TRUE),
  Wind = sum(Wind,na.rm = TRUE)
)
sources1 = sources1[1:72,]

```

```

Consumption = rbind(Cons2015,Cons2016,Cons2017,Cons2018,Cons2019,Cons2020,Cons2021)
colnames(Consumption) = c("Time","Forecast","Value")
Consumption[c("Date","x")] = str_split_fixed(Consumption$Time," ",2)
Consumption[c("Hour","y")] = str_split_fixed(Consumption$x,"-",2)
Consumption$Date = as.Date(Consumption$Date,format = "%d.%m.%Y")
Consumption = Consumption%>% select(Value,Date,Hour)
Consumption$Value = as.numeric(Consumption$Value)
#converting from Megawatt to Gigawatt
Consumption$Value = Consumption$Value/1000
# See that there are a few NA's in the data, so I will replace them with the previous value
Consumption = fill(Consumption,Value,.direction = "down")
Consumption = Consumption[0:113100,]
Cons_day = Consumption %>% group_by(Date) %>% summarize(
  Value = sum(Value)
)
#Cons_month
Cons_month = Cons_day %>% select(Date,Value)
Cons_month$Month = as.yearmon(Cons_month$Date,"%Y-%m")
Cons_month = Cons_month %>% group_by(Month) %>% summarise(
  Value = sum(Value)
)
Cons_month = Cons_month[0:72,]

```

```

EL_priceM = EL_price
EL_priceM$Date = as.yearmon(EL_priceM$Date,"%Y-%m")
EL_priceM = EL_priceM %>% group_by(Date) %>% summarise(
  ElPrice = mean(Price,na.rm=TRUE)
)
EL_priceM$Date = as.Date(EL_priceM$Date,format="%Y-%m")

```

```

GasPriceM = read.csv("GasPriceM.csv")
GasPriceM = GasPriceM %>% select(Date,Value)
GasPriceM = GasPriceM[216:304,]

```

```

oilprice = oilprice[49:120,]
oilprice$Price = as.numeric(oilprice$Price)

```

```

Weather$Date = as.Date(Weather$Date,format = "%d.%m.%Y")
WeatherM = Weather[169:240,]

```

```
Carbon = Carbon %>% select(Date,Value)
Carbon$Date = as.Date(Carbon$Date,format = "%m/%d/%Y")
Carbon_Month = Carbon
Carbon_Month$Month = as.yearmon(Carbon_Month$Date,"%Y-%m")
Carbon_Month = Carbon_Month %>% group_by(Month) %>% summarise(
  Value = sum(Value)
)
Carbon_Month = Carbon_Month[75:146,]
```

```
Coal = Coal[661:732,]
```

4. Correlation

The graphs with Electricity price and gas price gave an impression of the correlation between the two, but plotting them in the same chart will make it easier to see.

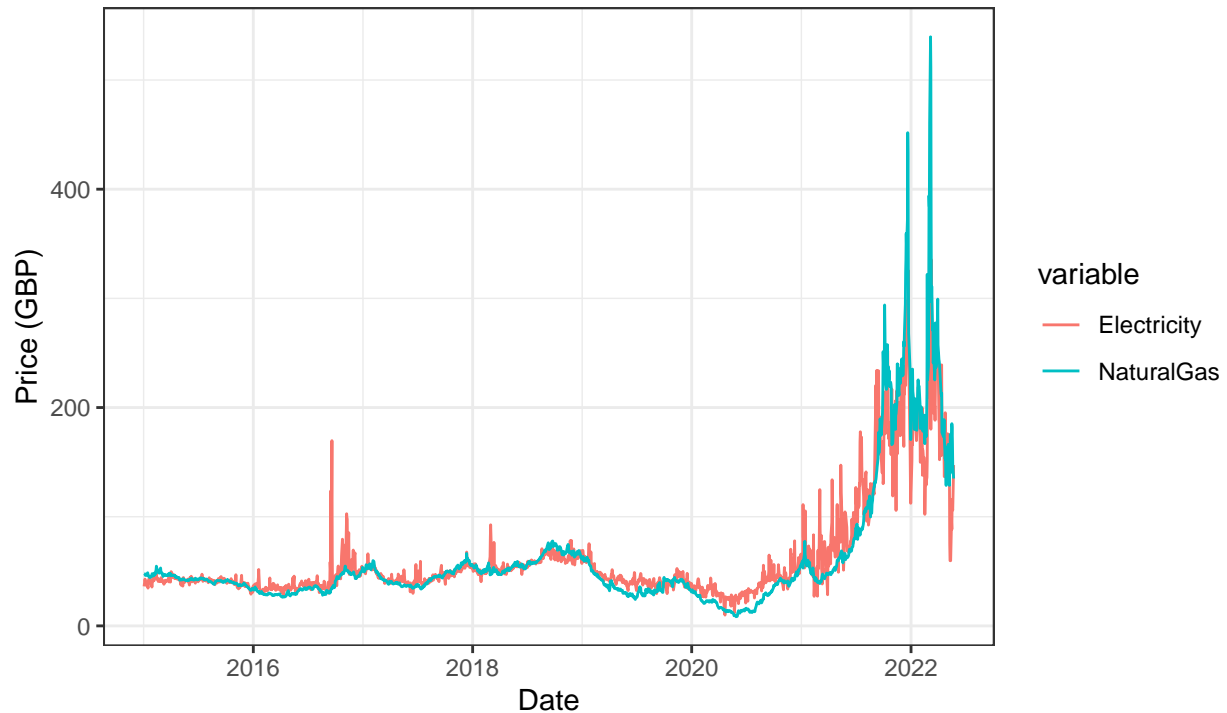
```
Prices = merge(EL_price,GasPriceD,by="Date")
colnames(Prices)=c("Date","Electricity","NaturalGas")

PricesMelted = reshape2::melt(Prices,id.var="Date")

ggplot(PricesMelted,aes(Date,value,col=variable))+
  geom_line()+xlab("Date")+ylab("Price (GBP)")+
  labs(title="UK Natural Gas Price & UK Electricity Price",
       subtitle = "Natural Gas in British pence per therm and Electricity in MW/h",
       caption = "Source: Investing.com & ENTSOE-E")+theme_bw()
```

UK Natural Gas Price & UK Electricity Price

Natural Gas in British pence per therm and Electricity in MW/h



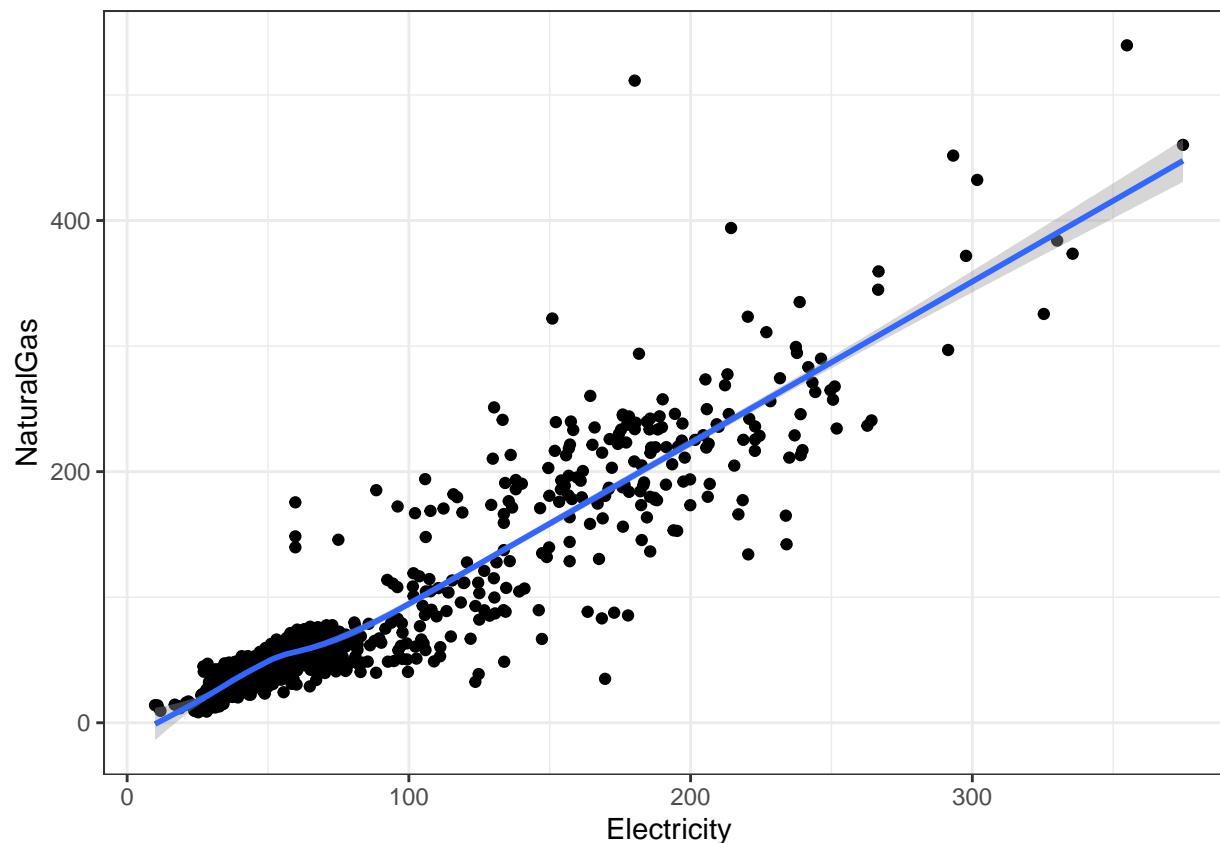
Source: Investing.com & ENTSOE-E

The graph shows that the prices of Electricity and Natural Gas follow each other very closely. As mentioned earlier, gas is a very important part of the energy mix of the UK. The visualization of Electricity prices and Natural Gas shows a clear correlation, but the electricity price seems to vary a lot more than the gas price. An important note is that the two prices are in different measures, the chart implies that the prices are almost the same but this is far from true. Natural gas prices are on average three to four times lower than electricity prices in the UK. I should have plotted the prices in the same measure, but since I did not find a conversion rate I trusted, I chose to keep the original measure.

Let's see how Electricity and Natural Gas prices are correlated in the UK.

```
ggplot(Prices,aes(Electricity,NaturalGas))+geom_point()+geom_smooth()+theme_bw()
```

```
## 'geom_smooth()' using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



```
cor(Prices$Electricity,Prices$NaturalGas)
```

```
## [1] 0.9344652
```

The chart and the correlation coefficient of 0.93 shows a strong correlation between Natural Gas prices and Electricity prices. There is a linear positive relationship between the variables. In the next section I will investigate if it only is correlation or an actual causal relationship between gas prices and electricity prices.

5. Regression

I will now apply a more formal analysis, where the goal is to create an estimate of the variables that explain the electricity price.

5.1 Variables

```
Gdf = EL_priceM
Gdf$GasPrice = GasPriceM$Value
Gdf = Gdf[0:72,]
Gdf$OilPrice = oilprice$Price
Gdf$CoalPrice = Coal$Value
Gdf$Carbon = Carbon_Month$Value
```

```
Gdf$WindSpeed = WeatherM$WindSpeed
Gdf$SunHours = WeatherM$SunHours
Gdf$Temp = WeatherM$Temperature
Gdf$Consumption = Cons_month$Value
Gdf$GasGen = sources1$Gas
Gdf$CoalGen = sources1$Coal
Gdf$NuclearGen = sources1$Nuclear
Gdf$WindGen = sources1$Wind
```

In the model of the electricity price the following variables are included:

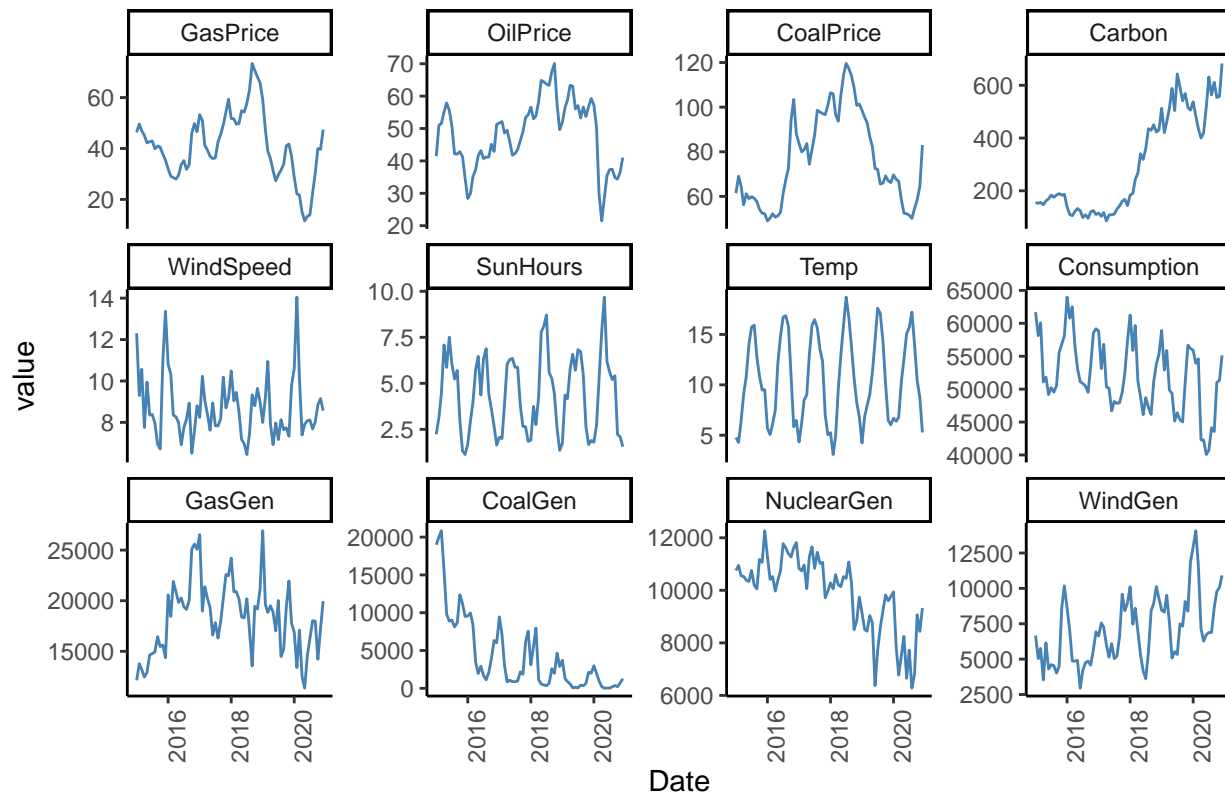
- IELPrice : A lagged variable for electricity price. I will control for the effect of past values on the current price.
- GasPrice : We are interested in the Gas Price effect on electricity Price.
- OilPrice: Included since it is a proxy for the world economy. When the world economy is good, prices are generally higher.
- CoalPrice: The price of coal is included because it is a substitute for Natural Gas in electricity generation and heating. Even if the use of coal is minimal in the UK today, it is useful to include it in the regression.
- Consumption: The consumption of electricity
- WindSpeed: The monthly average Wind Speed in the United Kingdom. Included since it impacts the electricity generated from wind.
- SunHours: The monthly average sun hours in the United Kingdom. Included since it impacts the electricity generated from solar.
- Temp: Monthly average temperatures in the United Kingdom. Included because it affects the demand of electricity in households.
- Carbon: Carbon Price is a possible determinant of electricity prices. Higher carbon prices is often associated with higher power prices.
- Energy sources: Included because the amount of energy produced is likely to affect prices

The amount of variables restrict both the frequency and the period of the data. The data are from 2015 through 2020 in monthly data.

```
GdfMelt = reshape2::melt(Gdf,id.var="Date")
GdfMelt = GdfMelt[73:936,]

GdfMelt %>% ggplot(aes(x=Date,y=value))+geom_line(color="steelblue")+
  facet_wrap(~variable,ncol=4,scales = "free_y")+
  theme_classic()+theme(axis.text.x = element_text(angle=90))+
  ggtitle("Independent variables in the Regression")
```


Independent variables in the Regression



Here is a visualization of the independent variables.

5.2 Regression in levels

Right away I know from economic reasoning that there will be problems with multicollinearity in the model, because I might have included too many variables with the same effect on the outcome. Multicollinearity is when multiple independent variables are highly correlated and results in less reliable estimates. I first run a regression with the variables in levels and then I will use a VIF-test to check for multicollinearity.

```
# lag function not working properly for some reason
# I will compute lags using a for-loop
Gdf$LElPrice = 0
for(i in 2:nrow(Gdf)) {
  Gdf$LElPrice[i] = Gdf$ElPrice[i-1]
}
Gdf[Gdf == 0] = NA

lm_mod = lm(ElPrice ~
  LElPrice+
  GasPrice+
  OilPrice+
  CoalPrice+
  Consumption+
  WindSpeed+
  SunHours+
```

```

Temp+
Carbon+
GasGen+
CoalGen+
NuclearGen+
WindGen,
data=Gdf)
vif(lm_mod)

```

```

##      lElPrice      GasPrice      OilPrice      CoalPrice      Consumption      WindSpeed
##      7.949644     11.412318      2.760225      6.263350      9.286894      7.114513
##      SunHours        Temp        Carbon        GasGen        CoalGen      NuclearGen
##      6.330171      6.329259      6.721883      5.911329      7.307811      4.972510
##      WindGen
##      13.522292

```

There is a big problem with Multicollinearity in this regression. A VIF-value over 5 indicates multicollinearity. I therefore remove the variables: lElPrice, GasGen, CoalGen, NuclearGen and WindGen because these have high correlation with other variables.

```

lm_mod1 = lm(ElPrice ~
      GasPrice+
      OilPrice+
      CoalPrice+
      Consumption+
      WindSpeed+
      SunHours+
      Temp+
      Carbon,
      data=Gdf)
vif(lm_mod1)

```

```

##      GasPrice      OilPrice      CoalPrice      Consumption      WindSpeed      SunHours
##      4.049797      2.465196      2.974957      5.080570      1.586318      3.143326
##      Temp        Carbon
##      3.615450      1.858957

```

A VIF-value between 1 and 5 is good and we have corrected for multicollinearity.

5.3 Stationarity

To be able to create valid regressions and models, the time series of all our variables must be stationary. If our model is not stationary the estimated coefficients in the models will not be consistent. Using non-stationary series can lead to spurious results, and especially in our case where most Energy time series consist of seasonality it is important that we use stationary time series. Seasonality can make it seem like two variables are correlated, but they can be correlated because of a similar seasonality.

We know from visualizing the variables, and our knowledge about energy markets that most of the time series are non-stationary. Therefore, I take the first difference of the variables. To test if the first-differenced variables are stationary, I will use the Augmented Dickey-Fuller test. The null hypothesis is that the variables time series are non-stationary, and if the p-value is lower than 0.05, we reject the null hypothesis.

```
Gdf = Gdf %>% mutate(
  dElPrice = c(NA,diff(ElPrice)),
  dGasPrice = c(NA,diff(GasPrice)),
  dOilPrice = c(NA,diff(OilPrice)),
  dCoalPrice = c(NA,diff(CoalPrice)),
  dConsumption = c(NA,diff(Consumption)),
  dWindSpeed = c(NA,diff(WindSpeed)),
  dSunHours = c(NA,diff(SunHours)),
  dTemp = c(NA,diff(Temp)),
  dCarbon = c(NA,diff(Carbon)),
  #dGasGen = c(NA,diff(GasGen)),
  #dCoalGen = c(NA,diff(CoalGen)),
  #dNuclearGen = c(NA,diff(NuclearGen)),
  #dWindGen = c(NA,diff(WindGen))
)
```

All the variables are now in first difference, and a ADF test can be performed to test for stationarity.

```
dGdf = Gdf[2:72,]
adfELPrice= adf.test(dGdf$dElPrice)
adfGasPrice = adf.test(dGdf$dGasPrice)
adfOilPrice = adf.test(dGdf$dOilPrice)
adfCoalPrice = adf.test(dGdf$dCoalPrice)
```

```
adfCarbon = adf.test(dGdf$dCarbon)
adfCons = adf.test(dGdf$dConsumption)
adfWindSpeed = adf.test(dGdf$dWindSpeed)
adfSun = adf.test(dGdf$dSunHours)
adfTemp = adf.test(dGdf$dTemp)
```

```
AdfResults = matrix(c(adfELPrice$p.value,adfGasPrice$p.value,adfCarbon$p.value,
                      adfOilPrice$p.value,adfCoalPrice$p.value,adfCons$p.value,adfWindSpeed$p.value,adfSunHours$p.value,adfTemp$p.value),
                    nrow=9,
                    colnames= c("P-Value"))
row.names(AdfResults) = c("dElPrice","dGasPrice","dCarbon","dOilPrice","dCoalPrice","dConsumption","dWindSpeed","dSunHours","dTemp")
```

```
stargazer(AdfResults,float=FALSE)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: fre., jun 03, 2022 - 12:27:39

	P-Value
dElPrice	0.038
dGasPrice	0.010
dCarbon	0.015
dOilPrice	0.012
dCoalPrice	0.024
dConsumption	0.010
dWindSpeed	0.010
dSunHours	0.010
dTemp	0.010

A P-value of 0.010 means that the P-value is less than 0.01. The P-values from the Augmented Dickey-Fuller tests show that our variables is now stationary. The differenced El Price and Coal price has a higher P-value than the other series, but since it is under the 5% level, I accept it.

5.4 Regression in differences

The non-stationarity in the regression with levels makes it clear that it is necessary to estimate a regression in differences.

```
lm_mod2 = lm(dElPrice ~
             GasPrice+
             OilPrice+
             CoalPrice+
             Consumption+
             WindSpeed+
             SunHours+
             Temp+
             Carbon,
             #GasGen+
             #CoalGen+
             #NuclearGen+
             #WindGen,
             data=dGdf)
```

```
stargazer(lm_mod1,lm_mod2,float = FALSE)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: fre., jun 03, 2022 - 12:27:39

	<i>Dependent variable:</i>	
	ElPrice (1)	dElPrice (2)
GasPrice	0.550*** (0.049)	0.535*** (0.099)
OilPrice	-0.073 (0.048)	-0.065 (0.084)
CoalPrice	0.130*** (0.027)	0.047 (0.070)
Consumption	0.0002 (0.0001)	0.0001 (0.0001)
WindSpeed	-0.841*** (0.272)	-0.482 (0.289)
SunHours	-0.306 (0.268)	-0.154 (0.335)
Temp	0.117 (0.141)	-0.159 (0.192)
Carbon	0.008*** (0.002)	0.017* (0.009)
Constant	10.883 (9.080)	0.035 (0.374)
Observations	72	70
R ²	0.930	0.585
Adjusted R ²	0.921	0.531
Residual Std. Error	2.687 (df = 63)	3.068 (df = 61)
F Statistic	104.470*** (df = 8; 63)	10.760*** (df = 8; 61)

Note:

*p<0.1; **p<0.05; ***p<0.01

There are only two variables in the differenced regression that has a statistically significant, the Gas price and the Carbon price. The gas price are significant at a 99% significance level and the Carbon price at 90%. But even though we have a stationary regression and the estimates should be accurate, we have not taken autocorrelation into account. Autocorrelation will lead to standard errors being inaccurate. We will use the Durbin-Watson test to test for autocorrelation.

5.5 Autocorrelation

After the stationarity in the time series have been corrected for, a second condition needs to be corrected. Autocorrelation does not affect the estimates in the model, but it causes efficiency problems. In this case it is therefore important to correct for possible autocorrelation because the significance of the variables is of special importance for the research question.

Autocorrelation is the correlation of a variable's current value and its past values. Autocorrelation can be

visualized using a ACF-plot, but for now I am interested in the autocorrelation of the regression estimate and not each single variable and I will therefore use a Durbin-Watson statistic.

```
lmtest::dwtest(lm_mod2)
```

```
##
## Durbin-Watson test
##
## data: lm_mod2
## DW = 2.8545, p-value = 0.9999
## alternative hypothesis: true autocorrelation is greater than 0
```

A DW-test statistic around 2 will say that we have little autocorrelation in our data. There is autocorrelation in this case and the estimation should therefore be transformed to a Cochrane-Orcutt estimation. This is a procedure which adjust a linear model for serial correlation in the error term.

5.6 Cochrane-Orcutt

```
coch_mod = cochrane.orcutt(lm_mod2)
```

```
coch_mod
```

```
## Cochrane-orcutt estimation for first order autocorrelation
##
## Call:
## lm(formula = dElPrice ~ GasPrice + OilPrice + CoalPrice + Consumption +
##      WindSpeed + SunHours + Temp + Carbon, data = dGdf)
##
## number of interaction: 6
## rho -0.484208
##
## Durbin-Watson statistic
## (original): 2.85448 , p-value: 9.999e-01
## (transformed): 2.05144 , p-value: 4.38e-01
##
## coefficients:
## (Intercept) GasPrice OilPrice CoalPrice Consumption WindSpeed
## -0.014616 0.513186 -0.045631 0.086977 0.000072 -0.455637
## SunHours Temp Carbon
## -0.195202 -0.071524 0.021663
```

The Cochrane-Orcutt estimation have a Durbin Watson statistic of 2.05, which means that the autocorrelation have been taken into account. The regression estimate is then close to being reliable and possible to interpret.

5.7 Homoskedasticity

The regression also needs to fulfill the homoskedasticity assumption. Homoskedasticity can be tested using a Breusch-Pagan test. The null hypothesis is that there is homoskedasticity, meaning that the error term has a constant variance conditional on the explanatory variables. Under this assumption, there should be no correlation between the squared residuals and the independent variables.

```
bptest(coch_mod)
```

```
##  
## studentized Breusch-Pagan test  
##  
## data:  coch_mod  
## BP = 5.1582, df = 8, p-value = 0.7405
```

The results of the Breusch-Pagan test indicate that we cannot reject the null hypothesis of homoskedasticity at standard significance levels. Our estimation fulfills the homoskedasticity assumption.

```
stargazer(lm_mod1,lm_mod2,coch_mod,float = FALSE)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: fre., jun 03, 2022 - 12:27:39

	<i>Dependent variable:</i>		
	ElPrice	dElPrice	
	(1)	(2)	(3)
GasPrice	0.550*** (0.049)	0.535*** (0.099)	0.513*** (0.075)
OilPrice	-0.073 (0.048)	-0.065 (0.084)	-0.046 (0.064)
CoalPrice	0.130*** (0.027)	0.047 (0.070)	0.087* (0.051)
Consumption	0.0002 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
WindSpeed	-0.841*** (0.272)	-0.482 (0.289)	-0.456* (0.264)
SunHours	-0.306 (0.268)	-0.154 (0.335)	-0.195 (0.316)
Temp	0.117 (0.141)	-0.159 (0.192)	-0.072 (0.163)
Carbon	0.008*** (0.002)	0.017* (0.009)	0.022** (0.009)
Constant	10.883 (9.080)	0.035 (0.374)	-0.015 (0.232)
Observations	72	70	70
R ²	0.930	0.585	
Adjusted R ²	0.921	0.531	
Residual Std. Error	2.687 (df = 63)	3.068 (df = 61)	
F Statistic	104.470*** (df = 8; 63)	10.760*** (df = 8; 61)	

Note:

*p<0.1; **p<0.05; ***p<0.01

The Cochrane-Orcutt estimation indicates that differences in Gas price and Carbon Price are significant in explaining the difference in Electricity price. Wind speed and coal price is also significant. Notably the oil price is not significant, but the sign is negative which means that the oil price have failed to be a proxy for the world economy. The oil price is cyclical and tends to follow the world economy, and the economic intuition is that electricity prices is higher in an economic upturn.

5.8 Overfitting - Regularisation

Overfitting because of many explanatory variables in the regression can lead to increase variance in the estimation. Since multicollinearity have been checked, overfitting should not be a problem, but a Regularisation process is run to be sure. Lasso regression is a method that reduces the number of variables, but still gives more accurate prediction. Will use Lasso and not Ridge because there is a small number of significant parameters.


```

x = model.matrix(dElPrice ~
                GasPrice+
                OilPrice+
                CoalPrice+
                Consumption+
                WindSpeed+
                SunHours+
                Temp+
                Carbon,
                data=dGdf)[,-1]
y = dGdf$dElPrice

cv.out = cv.glmnet(x,y,alpha=1)
bestlam = cv.out$lambda.min
lasso_mod = glmnet(x,y,alpha=1,lambda=bestlam)
coef(lasso_mod)

```

```

## 9 x 1 sparse Matrix of class "dgCMatrix"
##              s0
## (Intercept)  1.073925e-01
## GasPrice     5.259040e-01
## OilPrice     .
## CoalPrice    1.719039e-02
## Consumption  4.499607e-05
## WindSpeed    -2.503720e-01
## SunHours     .
## Temp         -1.046504e-01
## Carbon       1.040111e-02

```

OilPrice and Sunhours are set to zero. The regularisation model indicates a similar result as the linear regression, where GasPrice, Carbon Price and CoalPrice is still important for explaining electricity prices.

5.9 Granger Causality

After we have a valid estimation which implies that changes Gas prices have a significant effect on changes in electricity prices we need to test the causality. I use the Granger-test to test for Granger causality from Gas prices to electricity prices. If the null hypothesis is rejected then we can conclude that Gas prices Granger-causes the corresponding electricity price.

```

grangertest(dElPrice~dGasPrice,data=Gdf)

## Granger causality test
##
## Model 1: dElPrice ~ Lags(dElPrice, 1:1) + Lags(dGasPrice, 1:1)
## Model 2: dElPrice ~ Lags(dElPrice, 1:1)
##   Res.Df Df       F    Pr(>F)
## 1      67
## 2      68 -1 14.356 0.0003265 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The low P-value clearly indicate rejection of the null hypothesis of no Granger Causality. The results implies that Gas prices Granger-cause electricity prices. In this case it is important to check that we do not have reverse-causality, because despite our null hypothesis was rejected, it is possible reverse causation is occurring. That means that changes in electricity prices causes change in the Gas price. Test this by performing the Granger-causality test in reverse.

```
grangertest(dGasPrice~dElPrice,data=Gdf)
```

```
## Granger causality test
##
## Model 1: dGasPrice ~ Lags(dGasPrice, 1:1) + Lags(dElPrice, 1:1)
## Model 2: dGasPrice ~ Lags(dGasPrice, 1:1)
##   Res.Df Df       F Pr(>F)
## 1      67
## 2      68 -1 0.0292 0.8649
```

For the reverse Granger-causality test we can not reject the null hypothesis which means that changes in electricity prices does not Granger-cause changes in Gas prices.

The prior analysis show that the Natural Gas price is a strong predictor for Electricity prices. There is evidence that the electricity price in the UK is highly dependent on the gas price. The regression shows that Gas price, carbon price and coal price has a significant effect on electricity prices, with the gas price having the largest magnitude and signficance.

Can now answer the first question: Based on a multivariate regression and a Granger Causality test we can say that changes in electricity prices is to a high degree explained by changes in Natural Gas price.

6. Garch Model of Natural Gas and Electricity price

Energy markets are getting an increased position in world economy. The volatility of Electricity and Gas prices has been increasing in recent years. With the UK moving towards net zero, they have to adapt to more price volatility in electricity prices. Factors like political instability(Middle East, Russia for example) and the increase of wind power generation will keep the energy markets volatile in the future.

GARCH models incorporate mean-reversion which is normal in energy markets. I will estimate univariate GARCH models for natural gas and electricity prices in the UK, using daily data from 2015 to 2022, that is as recent as May 2022. The data is obtained from respectively Investing.com and ENTSO-E.

I use Logarithmic first differences to get stationary data. This section presents a univariate GARCH(1,1) model for both the electricity price and the natural gas price.

6.1 Electricity Return

Electricity prices are more vulnerable than other commodities, because it is not possible to store. This results in periods with high volatility.

```
EL_price$return = (c(NA,diff(log(EL_price$Price))))
EL_price = EL_price[2:2701,]
EL_price[is.na(EL_price)] = 0

ElPricePlot = ggplot(EL_price,aes(x=Date))+
  geom_line(aes(y=Price),color="black")+
  theme_minimal()
```

```
geom_line(aes(y=return*100),color="red",alpha=0.4)+theme_bw()+
ggtitle("Electricity Price and Returns")
```

```
ELPrice_ts = as_tsibble(EL_price,index=Date)
adf.test(ELPrice_ts$return) #ADF test shows stationary series.
```

```
##
## Augmented Dickey-Fuller Test
##
## data: ELPrice_ts$return
## Dickey-Fuller = -19.175, Lag order = 13, p-value = 0.01
## alternative hypothesis: stationary
```

```
ELPrice_arima = ELPrice_ts %>% fill_gaps() %>%
  model(
    arimaMod = ARIMA(return)
  )
ELPrice_arima %>% report()
```

```
## Series: return
## Model: ARIMA(0,0,3)(0,0,2)[7]
##
## Coefficients:
##          ma1          ma2          ma3      sma1      sma2
##      -0.3680  -0.2215  -0.1362   0.1013   0.0787
## s.e.   0.0192   0.0202   0.0204   0.0196   0.0186
##
## sigma^2 estimated as 0.02276: log likelihood=1277.53
## AIC=-2543.06 AICc=-2543.02 BIC=-2507.65
```

I let R choose the optimal ARIMA model for our differenced data. R chooses a $ARIMA(0,0,3)(0,0,2)\{7\}$. This means a model with three MA-terms. In addition the model has two seasonal MA-terms which means that the model uses the error 7 days ago to predict the next observation.

I could have let the ARIMA model take the first-difference. But I do this myself so I am certain that I have stationary data.

```
garchModel = garchFit(~arma(0,3)+garch(1,1),data=ELPrice_ts$return, trace=F)
```

```
summary(garchModel)
```

```
##
## Title:
## GARCH Modelling
##
## Call:
## garchFit(formula = ~arma(0, 3) + garch(1, 1), data = ELPrice_ts$return,
## trace = F)
##
## Mean and Variance Equation:
## data ~ arma(0, 3) + garch(1, 1)
```

```

## <environment: 0x00000000328df5f8>
## [data = ELPrice_ts$return]
##
## Conditional Distribution:
## norm
##
## Coefficient(s):
##      mu      ma1      ma2      ma3      omega      alpha1
## 3.3561e-05 -3.8139e-01 -2.5622e-01 -5.1435e-02 9.0637e-05 1.2502e-01
##      beta1
## 8.8780e-01
##
## Std. Errors:
## based on Hessian
##
## Error Analysis:
##      Estimate Std. Error t value Pr(>|t|)
## mu      3.356e-05 4.487e-04 0.075 0.9404
## ma1     -3.814e-01 2.180e-02 -17.495 < 2e-16 ***
## ma2     -2.562e-01 2.340e-02 -10.948 < 2e-16 ***
## ma3     -5.144e-02 2.272e-02 -2.264 0.0236 *
## omega   9.064e-05 2.054e-05 4.413 1.02e-05 ***
## alpha1  1.250e-01 1.575e-02 7.940 2.00e-15 ***
## beta1   8.878e-01 1.214e-02 73.100 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 2340.442 normalized: 0.8668305
##
## Description:
## Fri Jun 03 12:27:41 2022 by user: 47930
##
##
## Standardised Residuals Tests:
##
##      Statistic p-Value
## Jarque-Bera Test R Chi^2 6225.968 0
## Shapiro-Wilk Test R W 0.9302793 0
## Ljung-Box Test R Q(10) 46.5271 1.150879e-06
## Ljung-Box Test R Q(15) 85.01517 8.358314e-12
## Ljung-Box Test R Q(20) 98.99356 1.906142e-12
## Ljung-Box Test R^2 Q(10) 19.01859 0.04002721
## Ljung-Box Test R^2 Q(15) 24.43087 0.05812862
## Ljung-Box Test R^2 Q(20) 30.22873 0.06623121
## LM Arch Test R TR^2 20.31787 0.0613069
##
## Information Criterion Statistics:
##      AIC      BIC      SIC      HQIC
## -1.728476 -1.713177 -1.728489 -1.722943

```

All MA-terms are significant which means that previous “shocks”, affects the pattern of volatility.

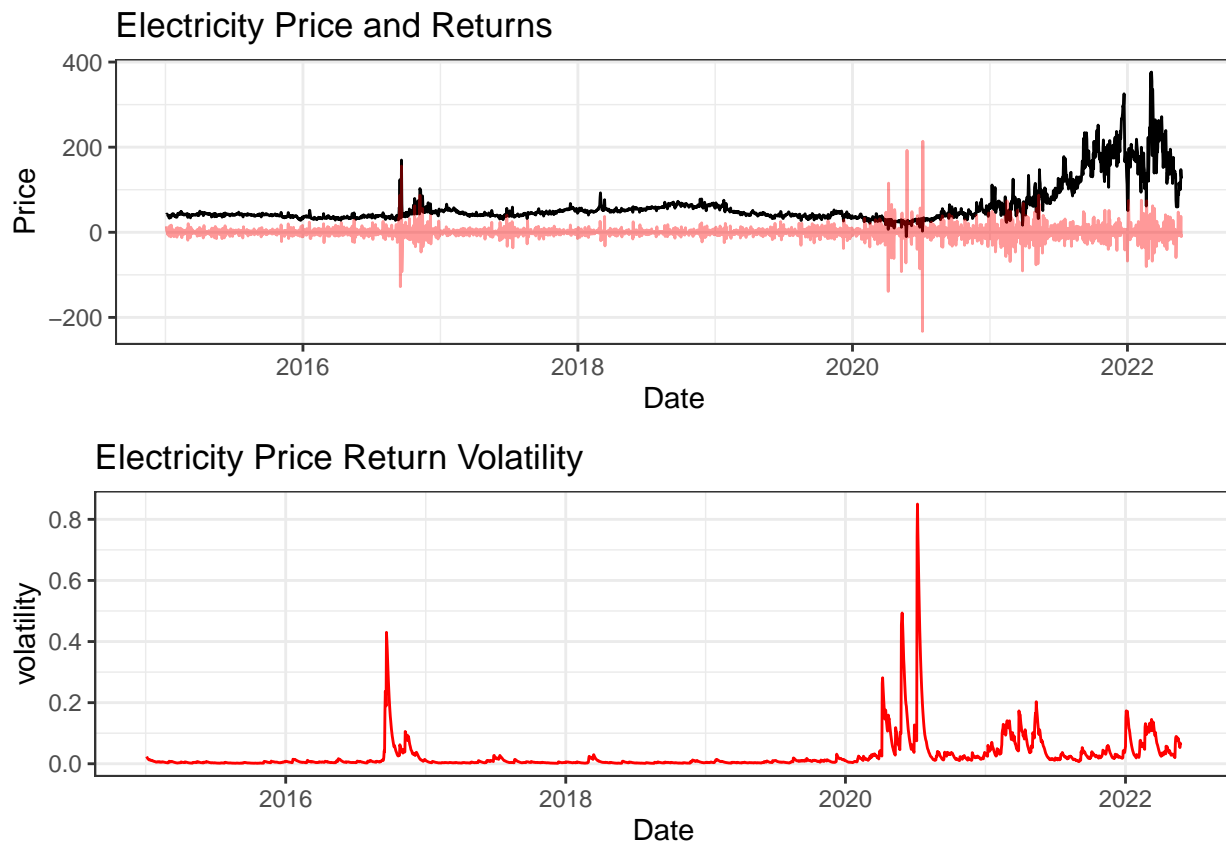
```

ELPrice_ts$volatility = garchModel@h.t

GarchEl = ggplot(ELPrice_ts,aes(Date,volatility))+
  geom_line(color="red")+theme_bw()+ggtitle("Electricity Price Return Volatility")

grid.arrange(ElPricePlot,GarchEl,ncol=1)

```



Surprisingly the most relative volatile period for price changes was in the start of Covid-19, but we see that the electricity price has had a volatile period the last years. The weather-dependent electricity supply from Wind and Solar can cause significant price volatility, which can lead to both higher and lower prices. Will investigate this in section 6.3.

6.2 Natural Gas return

I include the volatility of the changes in gas price return to give a point of reference for the electricity price volatility.

```

GasPriceGarch = GasPriceD[4528:6393, ] %>% select(Date,Price)
GasPrice_ts = as_tsibble(GasPriceGarch,index=Date)
GasPrice_ts = GasPrice_ts %>% mutate(
  returns = c(NA,diff(log(Price)))
)
GasPrice_ts = GasPrice_ts %>% dplyr::filter(!is.na(returns))
GasPricePlot = ggplot(GasPrice_ts,aes(x=Date))+
  geom_line(aes(y=Price),color="black")+

```

```

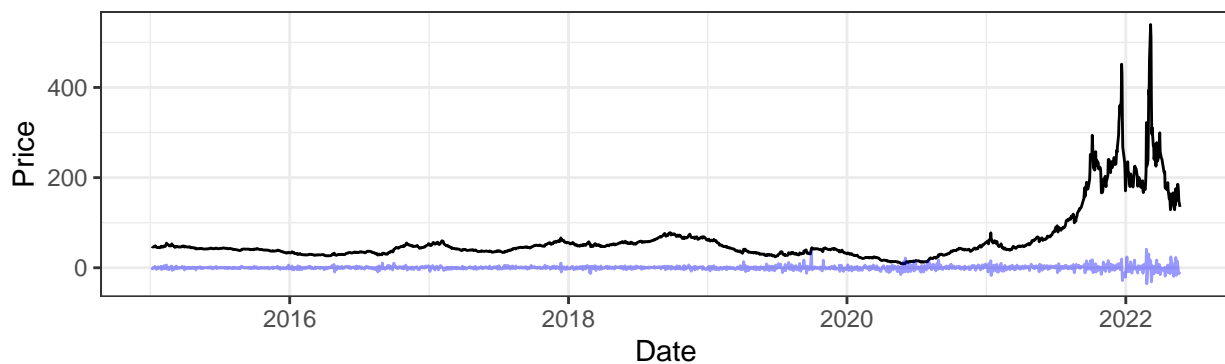
  geom_line(aes(y=returns*100),color="blue",alpha=0.4)+theme_bw()+
  ggtitle("Natural Gas Price and Returns")
adf.test(GasPrice_ts$returns) #stationary series

##
## Augmented Dickey-Fuller Test
##
## data: GasPrice_ts$returns
## Dickey-Fuller = -12.778, Lag order = 12, p-value = 0.01
## alternative hypothesis: stationary

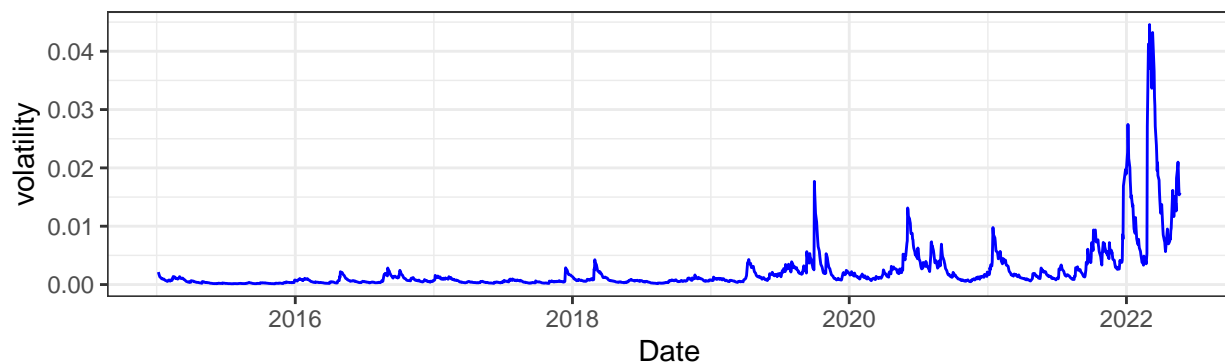
GasPrice_arima = GasPrice_ts %>% fill_gaps() %>%
  model(
    arimaMod = ARIMA(returns)
  )
garchModGas = garchFit(~arma(0,3)+garch(1,1),data=GasPrice_ts$returns, trace=F)
GasPrice_ts$volatility = garchModGas@h.t
GarchGas = ggplot(GasPrice_ts,aes(Date,volatility))+
  geom_line(color="blue")+theme_bw()+ggtitle("Conditional Gas Price Return Volatility")
grid.arrange(GasPricePlot,GarchGas,ncol=1)

```

Natural Gas Price and Returns



Conditional Gas Price Return Volatility



The changes in gas price returns are much less volatile than the electricity prices. This is in line with the impression we got from plotting the two prices. Electricity prices varies more than Gas prices. An economic explanation for this is that the demand and supply for electricity always have to be in sync, this can lead to volatile prices, whereas for Gas the demand and supply does not always have to meet because you can turn to other commodities.

6.3 Gas, Wind, Nuclear and Coal production's effect on electricity price volatility

We have seen that the three biggest energy sources in the UK are Gas, Nuclear and Wind. Historically it has been Coal. Nuclear seems to be relatively stable, while more Gas are used when there is not wind, and less when there are wind. To explore if the UK have become to dependent on Natural Gas it could be interesting to investigate how Gas, Wind, Nuclear and Coal generation effect electricity price volatility.

```
sourcesReg = sources1
sourcesReg$ElVolatility = ElPrice_ts2$volatility
sourcesReg$ElPrice = ElPrice_ts2$Price

GasCor = cor(sourcesReg$ElVolatility,sourcesReg$Gas)
WindCor = cor(sourcesReg$ElVolatility,sourcesReg$Wind)
CoalCor = cor(sourcesReg$ElVolatility,sourcesReg$Coal)
NuclearCor = cor(sourcesReg$ElVolatility,sourcesReg$Nuclear)

VolatilityCor = matrix(c(GasCor,WindCor,CoalCor,NuclearCor))

colnames(VolatilityCor) = "Price Volatility"
row.names(VolatilityCor) = c("Gas","Wind","Coal","Nuclear")
stargazer(VolatilityCor,title="Correlation Matrix",float = FALSE)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: fre., jun 03, 2022 - 12:27:44

Price Volatility	
Gas	-0.113
Wind	-0.010
Coal	-0.188
Nuclear	-0.313

The correlation between the different electricity sources and the price volatility shows that increased energy generation is correlated with decreased price volatility. The correlation between wind and volatility is the smallest and close to zero.

```
lm_vol = lm(ElVolatility~Gas+Wind+Coal+Nuclear,data=sourcesReg)

adfGas = adf.test(sourcesReg$Gas)
adfWind = adf.test(sourcesReg$Wind)
adfCoal = adf.test(sourcesReg$Coal)
adfNuclear = adf.test(sourcesReg$Nuclear)
adfVolatility = adf.test(sourcesReg$ElVolatility)
AdfResults3 = matrix(c(adfVolatility$p.value,adfGas$p.value,adfWind$p.value,adfCoal$p.value,adfNuclear$p.value),
colnames(AdfResults3) = c("P-Value")
row.names(AdfResults3) = c("Volatility","Gas","Wind","Coal","Nuclear")

stargazer(AdfResults3,float=FALSE)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: fre., jun 03, 2022 - 12:27:44

	P-Value
Volatility	0.010
Gas	0.010
Wind	0.010
Coal	0.013
Nuclear	0.010

All series are stationary.

```
dwtest(lm_vol)
```

```
##
## Durbin-Watson test
##
## data:  lm_vol
## DW = 0.13396, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
coch_lm_vol = cochrane.orcutt(lm_vol)
```

We see that there are a lot of autocorrelation.

```
bptest(coch_lm_vol)
```

```
##
## studentized Breusch-Pagan test
##
## data:  coch_lm_vol
## BP = 3.0723, df = 4, p-value = 0.5458
```

```
summary(coch_lm_vol)
```

```
## Call:
## lm(formula = ElVolatility ~ Gas + Wind + Coal + Nuclear, data = sourcesReg)
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.5957e-02  1.0604e-02   2.448  0.01444 *
## Gas          7.1475e-06  3.9792e-06   1.796  0.07259 .
## Wind        -2.3010e-06  4.9213e-06  -0.468  0.64015
## Coal        -2.0853e-05  1.1161e-05  -1.868  0.06184 .
## Nuclear     -8.7584e-06  2.3911e-05  -0.366  0.71418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0189 on 2347 degrees of freedom
## Multiple R-squared:  0.0032 , Adjusted R-squared:  0.0015
## F-statistic: 1.9 on 4 and 2347 DF, p-value: < 1.075e-01
##
## Durbin-Watson statistic
## (original):    0.13396 , p-value: 0e+00
## (transformed): 1.69145 , p-value: 3.813e-14
```


The Breusch-Pagan test indicates a homoskedastic regression and after transforming to a Cochrane-Orcutt there is less autocorrelation than at first but there is still substantial autocorrelation.

I will be careful to interpret this regression since it is likely to not be valid, but interestingly Gas has a positive sign, while Wind, Coal and nuclear have a negative sign. The effect of Gas and Coal are significant at a 90% significance level for explaining the electricity price volatility. The R-squared is also extremely low, which if the model was valid means that the different amount of types of energy does not explain the variance in the volatility.

7. ARIMA-model

Lastly I will create a forecast to understand how the electricity price would change from different Gas prices. The goal of the ARIMA-model is to create forecasts for the Electricity prices with three different scenarios for the Natural gas prices. Having analyzed the regression and a GARCH model, the question remains how Natural Gas prices will affect electricity prices in the future. Therefore in this section an ARIMA-model is used to forecast the electricity price based on different scenarios of the Natural gas price.

- The first scenario is no change in Natural Gas prices,
- The second scenario is increase in Natural Gas prices. 5 GBP increase in gas prices each month.
- The third scenario is decrease in Natural Gas prices. 5 GBP decrease in gas prices each month.

In the ARIMA-model monthly data from 2011 to may 2022 is used.

```
EL_priceM = EL_price
EL_priceM$Date = as.yearmon(EL_priceM$Date,"%Y-%m")
EL_priceM = EL_priceM %>% group_by(Date) %>% summarise(
  Price = mean(Price,na.rm=TRUE)
)
EL_priceM$Date = as.Date(EL_priceM$Date,format="%Y-%m")

El2011 = read.xlsx("ElPrices2011-2014.xlsx")
El2011$Date = as.yearmon(El2011$Date,"%Y-%m")
El2011$Date = as.Date(El2011$Date,format="%Y-%m")

EL_price = rbind(El2011,EL_priceM)

EL_price$returnEL = c(NA,diff(EL_price$Price))
EL_price = EL_price[2:137,]
```

```
adfEl = adf.test(EL_price$returnEL)
```

```
colnames(GasPrice) = c("Date","GasPrice")

GasPrice$returnGas = c(NA,diff(GasPrice$GasPrice))
GasPrice = GasPrice[169:304,]
```

```
adfGas = adf.test(GasPrice$returnGas)
```

```
AdfResults2 = matrix(c(adfEl$p.value,adfGas$p.value))
colnames(AdfResults2) =c("P-Value")

row.names(AdfResults2) = c("ElPrice","GasPrice")
stargazer(AdfResults2,float = FALSE)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: fre., jun 03, 2022 - 12:27:45

	P-Value
ELPrice	0.010
GasPrice	0.019

All variables are stationary, and we can move on to create a forecast.

```
prices = GasPrice
prices$ELPrice = EL_price$Price
prices$returnEL = EL_price$returnEL
prices$CarbonPrice = Carbon$CarbonPrice
prices$returnCarbon = Carbon$returnCarbon
prices$Date = as.Date(prices$Date)
prices$Date = yearmonth(prices$Date)
prices_ts = tsibble(prices, index=Date)
```

To make sure that a model with an exogenous variables, Gas price improves the predictive performance of the model compared to no exogenous variable we run a comparison and check if the exogenous variables actually improve the goodness of fit.

```
armax2 = prices_ts %>% fill_gaps() %>% model(
  mod1 = ARIMA(ELPrice~GasPrice),
  mod2 = ARIMA(ELPrice),
  mod3 = ARIMA(ELPrice~GasPrice+CarbonPrice)
)
glance(armax2) %>% arrange(AICc)
```

```
## # A tibble: 3 x 8
##   .model sigma2 log_lik   AIC  AICc   BIC ar_roots  ma_roots
##   <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl> <list>   <list>
## 1 mod3    31.9   -424.  858.  859.  873. <cpl [13]> <cpl [0]>
## 2 mod1    33.0   -427.  862.  862.  873. <cpl [13]> <cpl [0]>
## 3 mod2   123.   -516. 1046. 1047. 1066. <cpl [2]>  <cpl [4]>
```

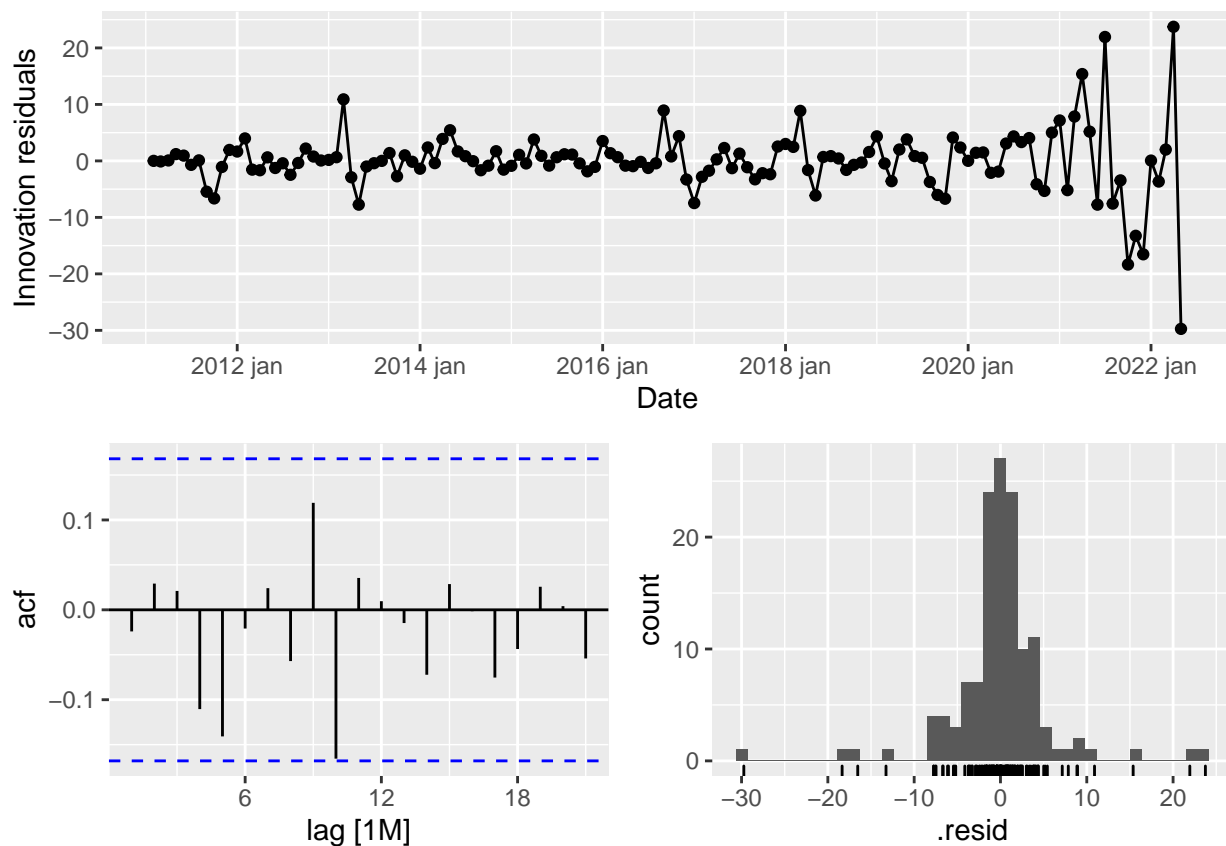
The AIC and BIC values for the model with Gas price are significantly lower than the model without an exogenous variable models, and this indicates a better fit to the data. I was also contemplating including carbon prices because as we saw from the Multivariate Regression, the carbon price have a significant effect on electricity price, but decided since I then also had to create a different scenarios for the carbon price that it would increase the noise in the forecast.

```
armax = prices_ts %>% fill_gaps() %>%
  model(ARIMA(ELPrice~GasPrice))
# Let R choose the optimal Arima model.
report(armax)
```

```
## Series: ELPrice
## Model: LM w/ ARIMA(1,1,0)(1,0,0)[12] errors
##
```

```
## Coefficients:
##          ar1      sar1 GasPrice
##        -0.6770  0.2919   0.7306
## s.e.    0.0786  0.1360   0.0237
##
## sigma^2 estimated as 32.99: log likelihood=-426.88
## AIC=861.75   AICc=862.06   BIC=873.38
```

```
#The model chooses a first difference model. We know that the first difference series of ElPrice and
#GasPrice are stationary.
armax %>% gg_tsresiduals()
```



The figure displays the residuals from the ARIMA model. It seems like the plotted residuals displays stationarity in our model. The ACF displays that we have little autocorrelation in the residuals. This implies that our ARIMA-model is valid. Seasonality is a big problem in power markets, but the ACF does not indicate any obvious seasonality. With a valid model, forecasts can be produced.

No Change:

```
scen1 = new_data(prices_ts,12) %>% mutate(
  GasPrice = rep(prices_ts$GasPrice[136],12) + cumsum(rep(0,12)),
)

armax1_forecast = forecast(armax,new_data=scen1)
NoChange = armax1_forecast%>% autoplot(prices_ts)+theme_bw()+ggtitle("No Change")
```

Increase:

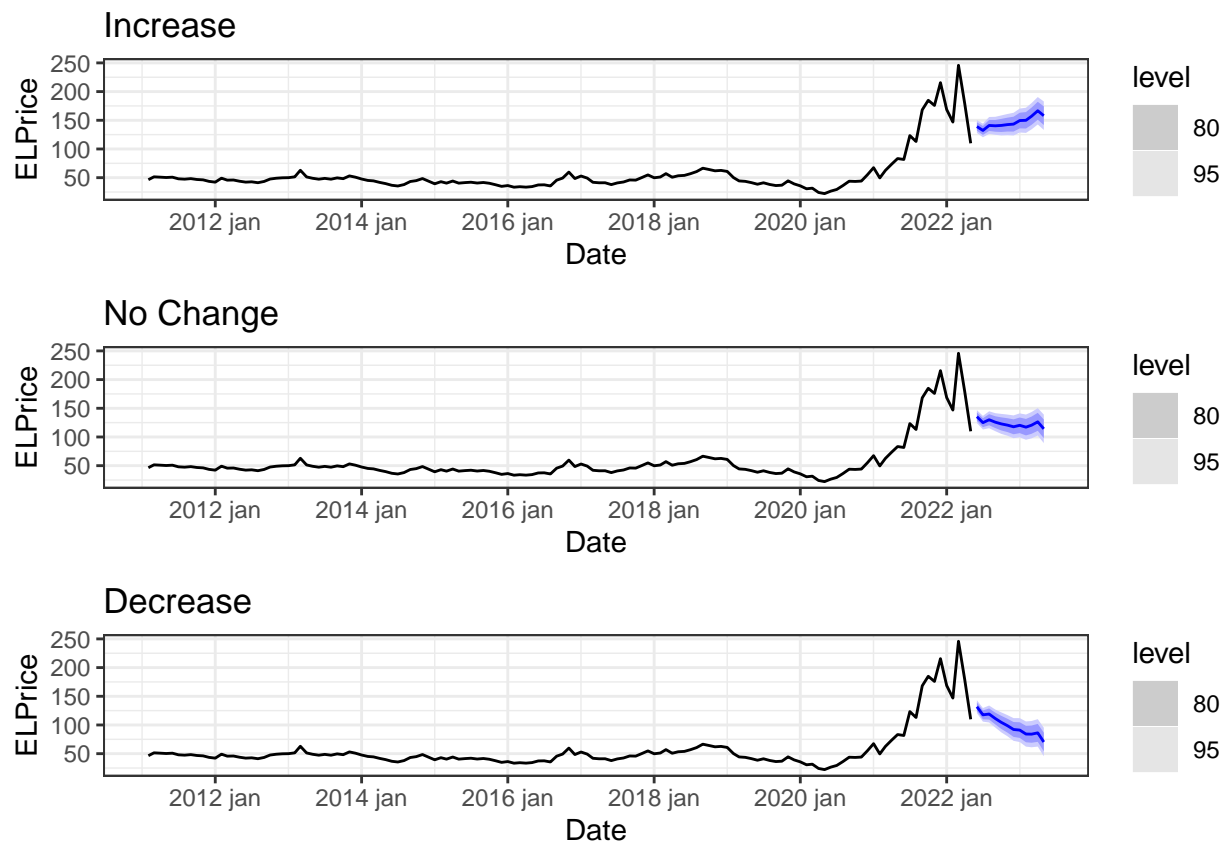
```
scen2 = new_data(prices_ts,12) %>% mutate(
  GasPrice = rep(prices_ts$GasPrice[136],12) + cumsum(rep(5,12)),
)

armax2_forecast = forecast(armax,new_data=scen2)
Increase = armax2_forecast%>% autoplot(prices_ts)+theme_bw()+ggtitle("Increase")
```

Decrease:

```
scen3 = new_data(prices_ts,12) %>% mutate(
  GasPrice = rep(prices_ts$GasPrice[136],12) + cumsum(rep(-5,12)),
)

armax3_forecast = forecast(armax,new_data=scen3)
Decrease = armax3_forecast%>% autoplot(prices_ts)+theme_bw()+ggtitle("Decrease")
```



As expected from the estimated significant relationship between electricity and gas prices, changes in gas prices makes a big impact on electricity prices. The uncertainty in these forecasts seem very small since the future has so many unknowns, but this show again how much of the electricity price that can be explained by the gas price.

We also see that the series with an AR process tends to see reversion to the series mean, because the decrease is greater than the upturn for the same magnitude of decrease/increase.

8. Conclusion

8.1 Summary of results

The aim of this report was to examine UK's dependence on Natural gas. The analysis show that gas price is the most important variable in explaining the electricity price. When the gas price increases, so will the electricity price. The short analysis on whether the price volatility is affected by the changing energy mix, gave results that are inconclusive and the model are likely to be invalid. Further analysis could be to analyse the impact the changing energy mix in the UK has on electricity price volatility. Reduced Coal use, and increased intermittent renewables, increases the dependence on Natural Gas and is likely to increase the price volatility.

8.2 Weaknesses of study

- Seasonality

I have not specifically dealt with seasonality in the analysis which can affect the coefficient estimates, but I have used first difference and the regression have stationary data and little autocorrelation which should deal with most of the seasonality.

- Monthly data in the Regression

I managed to find daily data for most of my datasets, but finding daily data for Temperature, Sun hours and wind speed was impossible with a time-constraint.

- Natural Gas prices and Electricity prices are affected by the same type of demand factors.

- Gas is used for electricity generation, but gas also works as kind of a substitute for electricity for heating. Therefore one should be careful to claim that gas price only affects electricity prices and not the other way around.

- European Energy crisis

The energy prices in europe have been extreme in the last half of 2021 and the start of 2022, because of different reasons. In the regression in section 5, the data period was restricted by the amount of variables and I only used data from 2015 through 2020, this means that the regression was not affected by the energy crisis.

The ARIMA forecast will be affected by this, because the forecast is based on former data, this could lead to inaccurate forecasts, but it is only a forecast over the next 12 months, and then the last year will be highly relevant.

References

Sources

1. Ofgem. (2022).Wholesale market indicators.

online

Available at: <https://www.ofgem.gov.uk/energy-data-and-research/data-portal/wholesale-market-indicators>

Accessed30May2022

2. Eia.gov. (2021). Factors affecting natural gas prices - U.S. Energy Information Administration (EIA).

online

Available at: <https://www.eia.gov/energyexplained/natural-gas/factors-affecting-natural-gas-prices.php>

Accessed 31 May 2022

3. for, D. (2022). Russia-Ukraine and UK energy: factsheet.

online

GOV.UK. Available at: <https://www.gov.uk/government/news/russia-ukraine-and-uk-energy-factsheet#:~:text=UK%20with%20Russia.-,Our%20single%20largest%20source%20of%20gas%20is%20from%20the%20>

Accessed 1 Jun. 2022

4. Norwegianpetroleum.no. (2022). *Exports of Norwegian oil and gas - Norwegianpetroleum.no.*

online

Available at: <https://www.norskipetroleum.no/en/production-and-exports/exports-of-oil-and-gas/>

Accessed 2 Jun. 2022

5. Rystadenergy.com. (2022). Offshore wind capacity additions in Europe to break records in 2022, driven by UK developments.

online

Available at: <https://www.rystadenergy.com/newsevents/news/press-releases/offshore-wind-capacity-additions-in-europe-to-break-records-in-2022-driven-by-uk-developments/>

Accessed 2 Jun. 2022

6. Economics Observatory. (2022). *Why are UK energy prices rising? - Economics Observatory.*

online

Available at: <https://www.economicsobservatory.com/why-are-uk-energy-prices-rising>

Accessed 2 Jun. 2022

7. Equinor.com. (2013). *Natural gas.*

online

Available at: <https://www.equinor.com/energy/natural-gas>

Accessed 2 Jun. 2022

Academic articles

8. Geissmann, T., & Obrist, A. (2018). Fundamental Price Drivers on Continental European Day-Ahead Power Markets
9. Alexopoulos, T.A. (2017). The growing importance of natural gas as a predictor for retail electricity prices in US. Energy
10. Mills, A., Wiser, R., Millstein, D., Carvallo, J.P., Gorman, W., Seel, J. and Jeong, S. (2021). The impact of wind, solar, and other factors on the decline in wholesale power prices in the United States.

Data

Energy Sources, Electricity Price and Energy Consumption in the UK:

11. Entsoe.eu. (2022). ENTSO-E Transparency Platform.

online

Available at: <https://transparency.entsoe.eu/>

Accessed 31 May 2022

Gas Prices, Coal Prices, Carbon Prices, and Import and Export data

12. Knoema. (2021). Home - Public Knoema Data Hub.

online

Available at: <https://public.knoema.com/>

Accessed 31 May 2022

Coal Production, Renewable Production and Weather data

13. GOV.UK. (2022). Research and statistics.

online

Available at: <https://www.gov.uk/search/research-and-statistics>

Accessed 31 May 2022