AMSTERDAM
DATA
COLLECTIVE

GREEN VOTING EXPLAINED

The number of Green Members of the European Parliament (MEPs) has grown in recent years, thanks to research on the effects of climate change and to the increased focus of the European public on sustainability. Finding a pattern in Green voting behaviour would allow us to test the current theories in the literature and predict new trends. Many macroeconomic indicators such as GDP per capita, location and the support for traditional left-wing parties are known to be strong predictors of the vote share for Green political groups. In the following pages, we will show you how, using a popular classification method, it is possible to predict which countries will successfully elect at least one Green MEP.

Green Elections

Between 23 and 26 May 2019, the 28 member countries of the European Union were called to elect the ninth European Parliament since 1979. With rates varying throughout the Union, it is estimated that an average of 51% of eligible voters took part in the election, an all-time high and an unprecedented success for European democracy. Perhaps even more astounding is the fact that the Greens, typically a minority in the political landscape, scored their best result to date, with a total of 74 seats out of 751. The support for the Greens is not homogeneous across the European Union and certainly has not been through the years. As data scientists, we found it an interesting and challenging opportunity to develop my very own model to predict the performance of the Greens/EFA group in the European Parliament elections.

Greens/EFA
Founded in 1999, the group unites environmentalists and regionalists

Before 1999
They were preceded by the Greens and the Rainbow group (ARC)

Leadership
The current chairs are Philippe Lamberts and Franziska Keller

Priorities
Climate change, EU agricultural policies, ethical scrutiny for MEPs

Impressive results
Green MEP numbers increased by 42.3% during the last election!

Looking for answers

As a first step, we look at the available data to derive some initial insights.

The Data

We collected the results of all nine European Parliament elections from the European Parliament website, dividing the results by country. The share of Green MEPs will be our dependent variable, while the share of Left-wing MEPs will be part of the regressors. In addition, we collected some popular predictors of Green voting: GDP per capita, CO2 emissions, nuclear power usage and unemployment. We also assigned geographical indicators to each country, indicating whether the country is located in Northern, Southern, Western or Central Europe, to capture location fixed effects. At the end of the data wrangling process we have a dataset of 175 observations. The size of the dataset is limited because election data has a five-yearly frequency starting from 1979. As of 2019, the highest GDP per capita is found in Luxembourg, while the highest CO2 emissions are in Germany and in the United Kingdom. France has the most intensive nuclear power usage, and Greece has the highest unemployment rate.

The Literature

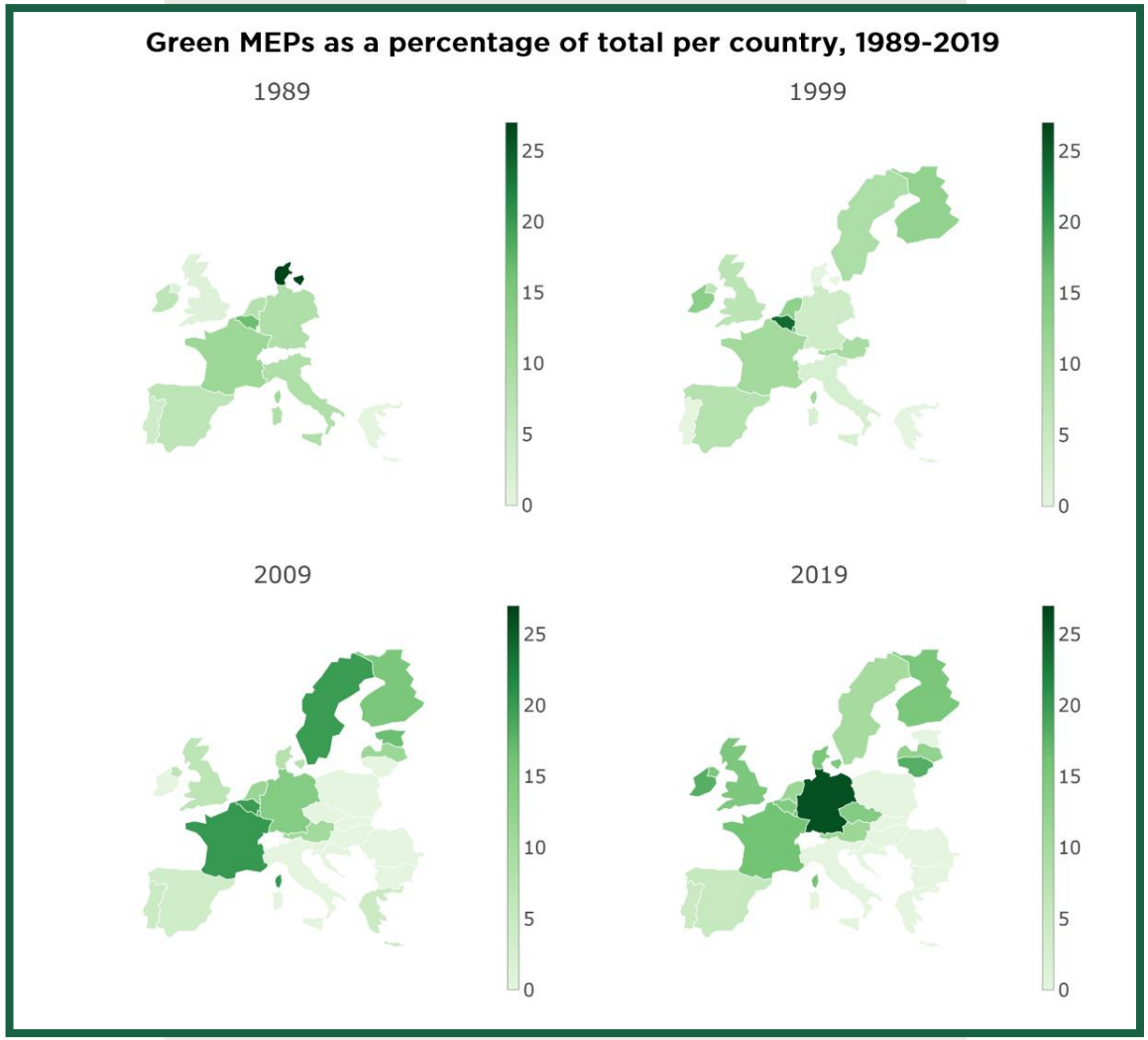
We compare two theories of Green voting: first, the "post-materialist" thesis of Inglehart (1971), who claimed that, as national income increases and a stable welfare state is in place, the focus of political debate shifts towards issues related to the quality of life, such as the environment; second, the "left-libertarian" thesis of Kitschelt (1988), who claimed that parties like the Greens combine an agenda similar to the traditional Left with libertarian subjects such as the reduction of bureaucracy and the decentralization of government to local autonomies,

"Popular predictors of Green voting: GDP per capita, CO2 emissions, nuclear power usage and unemployment."

and can only gain support in rich economies with a strong leftist government. We will test both theories by looking at the data.

Distribution of Green MEPs

On the right is a sequence of choropleths illustrating how the number of Green MEPs in all member countries of the European Union has changed over the years. In general, the consensus for Green parties has been growing. The share of seats in the European Parliament varies widely, typically between 0 and 25% but swinging sometimes higher, such as in Denmark in 1989 and in Germany in 2019. We already observe a pattern here: green votes are concentrated between Western and Northern Europe, in particular Denmark, Sweden, Belgium, France and Germany. These are all high-income economies with strong welfare states, but also a long tradition of socialist governments. Here the Greens reap a substantial proportion of the seats (between 20 and 30%), while their Southern European counterparts gain much lower numbers.



Explanatory Variables

On the right, a series of scatterplots that correlate our main explanatory variables. The y-axis represents the number of Green seats in proportion to the eligible seats for each country, the x-axis the log GDP per capita, each dot a member country of the EU and its size the proportion of left-wing seats to total seats. The different regions of Europe are highlighted in different colours. We observe a positive correlation between GDP and Green seats. We see that Northern Europe has almost always an average of between 10 and 20% Green seats while Southern Europe usually rests on the 0 line. Western and Central Europe are more homogeneous in their distribution. We also see that the countries were the Left is the strongest (the largest circles) are located at the bottom of each plot, which suggests that indeed there is a negative correlation between Left and Green seats.

Plotting correlations between Green MEPs, log GDP and Left MEPs, 1989-2019

In blue, Northern Europe; in green, Western Europe; in pink, Central Europe; in yellow, Southern Europe.

Modelling

In a perfect world, given the explanatory variables, we would be able to predict the exact amount of seats gained by the Greens. However, in the real world, voter behaviour is influenced by an endless list of factors, some of which can be measured, while others will never be captured in a quantitative model.

For this reason, we split our continuous range of seat shares into a binary set and define as 1 the outcomes in which the Greens obtain at least one seat, and 0 otherwise. To predict such outcomes, we use binary classification, and in particular Logistic Regression, which typically performs quite well when the number of regressors is limited. After a feature selection routine, we find that indeed log GDP, Left MEPs and Northern Europe are the strongest predictors in our setting. We then run a Logistic Regression on several combinations of countries and average the coefficients, thus improving robustness and external validity. The results are clear: log GDP and Northern Europe have a strong positive effect on the probability of electing at least one Green MEP, while Left MEPs have a negative effect.

Conclusion

High-income countries closer to Northern Europe and with lower support for the traditional left are more likely to elect Green MEPs. What does this tell us? Perhaps this corroborates Inglehart's "post-materialist" thesis: it is hard to worry about climate change when you struggle to pay your bills, but as soon as a certain threshold of wealth is reached, the focus of political debate in the left-wing shifts from the typical class struggles of the last century to issues related to the quality of life, such as the environment. Because Left and Greens compete for the same pool of voters, these priorities are determinant in the final electoral result. As the rest of the EU converges on the same growth path as the top economies, we should then expect the number of Green MEPs to soar even higher in the coming years.

"As soon as a certain threshold of wealth is reached, the focus of political debate shifts to issues related to the quality of life."

About Us

Partnering with us means relevant insights and usable models. We are passionate about advanced statistical techniques, machine learning and big data. It is our goal to help our clients develop the best data-driven solutions. Amsterdam Data collective is a team of quantitative specialists, combining the business acumen of a management consultant with the problem-solving abilities of a data scientist. Our consultants identify, quantify and provide advice on how to act upon opportunities and threats. If you trust us to be your data science partner, we trust you to take the right decisions.

Author

Raffaele Di Carlo
Consultant, Researcher