Project title: Generating representative counterfactuals in election campaigns via the use of discrete choice experiments and post-stratification

Project leaders: Dr. Roberto Cerina, Assistant Professor of Statistics
r.cerina@maastrichtuniversity.nl
Roberto Cerina - Maastricht University
Dr. Roselinde Kessels, Assistant Professor of Statistics
r.kessels@maastrichtuniversity.nl
Roselinde Kessels - Maastricht University

Proposal: Forecasting the outcome of elections has been a major research topic in the political science literature. Understanding voter preferences and predicting voter choice in elections is of critical importance to politicians, as they can use the insights gained from the exercise to fine-tune their campaign strategies.

Most researchers base their forecasts on opinion polls, mainly because this source of information is abundant and readily available before elections. However, opinion polls are not without limitations as measurements of political perceptions and preferences: firstly, traditional random-digit-dial (RDD) polls suffer from low-response rates and over-sampling of politically enthusiastic voters, whilst online polls are also biased by digital literacy rates and access to the web, amongst factors; secondly, polls typically don't have a systematic way to estimate interesting counterfactuals that can be of interest to political campaigns - e.g. *`In which constituencies would the Conservative party improve their chances of winning if they changed their immigration/ abortion /covid policy, ceteris paribus ?'*

To solve these limitations, we propose to enhance the non-representative data obtained from opinion polls with data from discrete choice experiments (DCEs), and use post-stratification and regularisation techniques to generate representative counterfactual maps of the votes under different party-policy and candidate-attribute regimes.

Economic voting theory underlying DCEs assumes that voters maximize their utility. The theory attempts to explain the political choices made by voters based on their assessment of the national economy and their own personal economic situation. *Imagine for example that you must choose between two different alternatives, which will govern the country in the future. It is up to you to weigh the pros and cons. Which alternative do you prefer?*

Alternative A	**Alternative B**
• You will receive 350£ more in salary each month • Increasing unemployment • Rishi Sunak as prime minister	• You will receive 150£ more in salary each month • Decreasing unemployment • Liz Truss as prime minister
O	**O**

This is an example of a choice task of a DCE that can be presented to respondents. The levels of three attributes are varying, and several choice tasks are shown. The total of all choice sets assigned to the respondents is called the experimental design. We can

maximize the information content of the DCE data by computing a carefully selected design according to experimental design principles such as Bayesian D-optimality.

After performing the DCEs, there is still the question of how to generalise the results to each of the relevant political constituencies (e.g. 650 parliamentary constituencies in the UK). In a best-case scenario, the number of participants to a DCE will be in the range of 1,000 to 5,000, and hence simple aggregation of the participants' choices at the constituency-level would not confer enough statistical power to make representative constituency-level inference. We therefore propose to use Multilevel Regression and Post-Stratification (MrP) to generalise the experiments' results to the constituency-level, by accounting for potential selection effects and sample biases through Hierarchical Bayesian modeling.

The objective of this proposal is therefore threefold:

1. To design and perform DCEs that can examine voter choices under different policy- and candidate-regimes, and tease out heterogeneity and complexity in these preferences;
2. To analyse choice data from DCEs using MrP to account for selection effects in the sample and generalise findings to the smaller lower-level areas, whilst preserving uncertainty in preferences via a Hierarchical Bayesian approach;
3. To develop an R package to design DCEs for electoral campaigns and generate easily interpretable counterfactual vote maps.

Keywords: Voter choice modelling, discrete choice experiments, multilevel regression and post-stratification, opinion polls, forecasting, counterfactual vote maps

Requirements candidate: **Background in computational statistics and statistical programming (absolute must!)**, Proficient in English, Feel for political science

Top 5 selected publications:

Cerina R, Duch, R (2021), "Polling India via regression and post-stratification of non-probability online samples", *PLOS ONE*, *16*(11), [e0260092]. https://doi.org/10.1371/journal.pone.0260092

Cerina, R, Duch, R (2020), "Measuring public opinion via digital footprints", *International Journal of Forecasting*, *36*(3), 987-1002. https://doi.org/10.1016/j.ijforecast.2019.10.004

Kessels R, Jones B, Goos P (2015), "An improved two-stage variance balance approach for constructing partial profile designs for discrete choice experiments", *Applied Stochastic Models in Business and Industry*, *31*(5), 626-648. https://doi.org/10.1002/asmb.2065

Luyten J, Tubeuf S, Kessels R (2022), "Rationing of a scarce life-saving resource: Public preferences for prioritizing COVID-19 vaccination", *Health Economics*, *31*(2), 342-362. https://doi.org/10.1002/hec.4450

Mouter N, Boxebeld S, Kessels R, van Wijhe M, de Wit GA, Lambooij MS, van Exel J (2022), "Public preferences for policies to promote COVID-19 vaccination uptake: A discrete choice experiment in the Netherlands", *Value in Health*, *25*(8), 1290-1297. https://doi.org/10.1016/j.jval.2022.03.013

Maastricht University

Approved by the academic department, Prof. Dr. Alexander Grigoriev

AACSB ACCREDITED

EFMD EQUIS ACCREDITED

ASSOCIATION OF MBAS ACCREDITED