

Crowdsourced Replication Initiative – Study Extension*

Danilo Freire[†]

Robert McDonnell[‡]

23 October 2018

1 Summary

We suggest five extensions to Brady and Finnigan (2014):

- Add the 2016 ISSP survey wave to the original sample
- Use multiple imputation to address the issue of missing values
- Employ extreme bounds analysis to assess predictor robustness
- Add a hierarchy to the model to explore regional effects
- Use machine learning models to test whether the results are driven by parametric assumptions or not

2 Sample

We propose to extend the original study by adding the latest ISSP wave to the sample. This wave was not available when Brady and Finnigan (2014) published their article. Including the 2016 wave to the dataset has two main benefits to our analysis: 1) it increases the sample size, which allows us to observe smaller treatment effects; 2) it provides us with new data to test the explanatory power of the original models.

*All information necessary to replicate our contribution to the project is available at <https://github.com/danilofreire/crowdsourced-replication-iniciative>. Please contact the authors should you require any assistance to reproduce the results.

[†]Postdoctoral Research Associate, The Political Theory Project, Brown University, 8 Fones Alley, Providence, RI, 02912, danilo_freire@brown.edu, <http://danilofreire.com>. Corresponding author.

[‡]Data Scientist, First Data Corporation. Views expressed are author's own.

Our sample will be restricted to the developed countries which are included in all three ISSP waves. They are, in alphabetic order: Australia, France, Germany, Great Britain, Japan, New Zealand, Norway, Spain, Sweden, Switzerland, and the United States of America.

3 Multiple Imputation

We will also replicate the analysis using multiple imputation. Multiple imputation consists of adding values to missing observations by using a model that incorporates the randomness in data. The process is usually repeated a few times to get a more realistic sense of the uncertainty of the estimates (Little and Rubin 1989; Rubin 1987). Under reasonable assumptions, multiple imputation is approximately unbiased, increases the power of statistical tests, and is easy to compute using modern statistical software (Lall 2016). We will use the *Amelia II* package (Honaker et al. 2011) for the R statistical language (R Core Team 2018) and replicate all the analyses we include in this proposal.

4 Regional Hierarchies

It is well known that Western European countries have a tendency to maintain larger welfare states (Ebbinghaus and Manow 2001), however, this dynamic was not modelled in the original work. Our sample contains seven Western European countries out of the total of eleven. We propose to add an extra level of hierarchy to the statistical models to explore this further. This higher level will model regional differences by grouping the countries into two groups – ‘Europe’ and ‘Not Europe’. In practical terms, this can be easily done by re-writing the original models in a probabilistic programming language, such as Stan (Carpenter et al. 2017). In this way, observations which may be correlated with one another (because of shared Western European attitudes towards social expenditure, for example) can be grouped together and this correlation can be explored. We may also decide to explore different groupings: ‘Anglo-America’ (the United Kingdom and the U.S.A) and the rest, in order to exploit ‘welfare retrenchment’ in the first group (Ebbinghaus and Manow 2001).

5 Extreme Bounds Analysis

We also intend to assess the robustness of our findings with extreme bounds analysis (EBA). The estimation of EBA consists in running every possible permutation of the explanatory variables including in a given model. The coefficients are then analysed as a distribution. In particular, we are interested in the share of the coefficients that lies above or below zero, so that we can evaluate whether the explanatory variables are sensitive to changes in the model specification (Leamer 1985; Sala-i Martin 1997).

It is nonetheless tricky to estimate hierarchical models using extreme bounds analysis. We may therefore include all variables as a single level model, either removing country fixed effects, or including it as dummies.

6 Machine Learning Models

Lastly, we intend to employ machine learning tools to assess whether the results obtained are due to parametric parameter assumptions. Machine learning methods such as random forests (Breiman 2001) or gradient boosting (Friedman 2001) have many desirable properties, such as ‘highly accurate predictions, robustness to noise and outliers, internally unbiased estimate of the generalisation error, efficient computation, and the ability to handle large dimensions and many predictors’ (Muchlinski et al. 2015, 7). Thus, random forest allows the researcher to estimate very flexible models with minimal assumptions. Unlike parametric methods such as ordinary least squares or logistic regressions, the analyst does not have to impose any distributional form to the data-generating process. As a result, random forest is able to effectively uncover complex, nonlinear interaction effects in the data without prespecification (Jones and Linder 2015).

Whereas machine learning methods perform well without multilevel specifications, recent R packages such as MixRF are able to estimate mixed effects models¹. In the case the package does not work well with our data, another sensible option would be to include the country-level predictors as features in the model.

¹More information is available at <https://github.com/randel/MixRF>.

References

- Brady, D. and Finnigan, R. (2014). Does Immigration Undermine Public Support for Social Policy? *American Sociological Review*, 79(1):17–42.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1).
- Ebbinghaus, B. and Manow, P. (2001). *Comparing welfare capitalism*. Taylor & Francis.
- Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, pages 1189–1232.
- Honaker, J., King, G., and Blackwell, M. (2011). Amelia II: A Program for Missing Data. *Journal of Statistical Software*, 45(7):1–47.
- Jones, Z. and Linder, F. (2015). Exploratory data analysis using random forests. pages 1–31.
- Lall, R. (2016). How Multiple Imputation Makes a Difference. *Political Analysis*, 24(4):414–433.
- Leamer, E. E. (1985). Sensitivity Analyses Would Help. *The American Economic Review*, 75(3):308–313.
- Little, R. J. and Rubin, D. B. (1989). The Analysis of Social Science Data with Missing Values. *Sociological Methods & Research*, 18(2-3):292–326.
- Muchlinski, D., Siroky, D., He, J., and Kocher, M. (2015). Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. *Political Analysis*, 24(1):87–103.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons.
- Sala-i Martin, X. X. (1997). I Just Ran Two Million Regressions. *The American Economic Review*, 87(2):178–183.