Estimating subnational preferences across the European Union

Journal Article

Author(s):
Lipps, Jana; Schraff, Dominik

Publication date:
2021-01

Permanent link:
https://doi.org/10.3929/ethz-b-000381443

Originally published in:
Political Science Research and Methods 9(1), https://doi.org/10.1017/psrm.2019.38

Funding acknowledgement:
186002 - Regional Inequality and the Political Geography of EU Support (SNF)

The European Political Science Association, 2018

Estimating subnational preferences across the European Union *

JANA LIPPS AND DOMINIK SCHRAFF

Subnational analyses of political preferences are substantively relevant and offer advantages for causal inference. Yet, our knowledge on regional political preferences across Europe is limited, not least because there is a lack of adequate data. The rich Eurobarometer data is a promising source for European-wide regional information. Yet, it is only representative for the national level. This paper compares state-of-the-art methods for estimating regional preferences from nationally representative Eurobarometer data, validating predictions with regionally representative surveys. Our analysis highlights a number of challenges for estimating regional preferences across Europe, such as data availability, variable selection, and over-fitting. We find that predictions are best using a Bayesian additive regression tree with synthetic post-stratification.

*Jana Lipps is doctoral student at the European politics research group at ETH Zurich, Switzerland. Dominik Schraff is postdoctoral researcher at the European politics research group at ETH Zurich, Switzerland. We thank Thomas König, Lucas Leemann, Matthew Loveless, Santiago Olivella, and Fabio Wasserfallen for helpful comments and advice.

Subnational variation in political preferences across Europe has attracted increasing attention in recent years. [1] Elections in Europe have repeatedly demonstrated the divergences of political preferences between regions (e.g., Jones, Johnston, and Pattie 1992; Schraff 2019). A prominent example is the British vote to leave the European Union, which was informed by strong geographical divides.[2] The substantial variation in living conditions across the territories of EU members states likely affects a wider range of political preferences (Beramendi 2012; Pittau, Zelli, and Gelman 2010). Therefore, the regional level is a substantively relevant unit of analysis for our understanding of European politics.

In addition, an investigation of subnational public opinion across Europe holds great promises for innovative research designs in comparative politics. A meso-level perspective provides a good balance between generality (macro-level) and specificity (micro-level) (Georgiadou, Rori, and Roumanias 2018). Here, regional-level data offers the chance to strengthen causal inference (cf, Rodden 2010). Usually, country-level data shows limited variability and suffers from many potential con-founders. Subnational data allows researchers to hold constant confounding factors from the national levels. Further, valid measures of regional preferences enable insightful data description, case selection, and mixed method designs. Noisy and biased regional opinions can provide a flawed picture of how regions scale on politically relevant dimensions. Consider the example of regional variation in Euroscepticsm. Valid estimates of regional preferences are required to adequately rank regions and distinguish between more or less Eurosceptic

[1]This paper uses the terms subnational and regional exchangeable. Both describe geographical areas within EU member states.

[2]https://www.economist.com/news/britain/21701257- results-paint-picture-angry-country-divided-class-age-and-region-country-divided

areas. However, we so far have limited knowledge on regional public opinions as large European comparative social surveys are only representative on the national level.This paper addresses this gap by evaluating methods to derive more valid, European-wide regional preferences.

The Eurobarometer (EB) is the European comparative social survey with the broadest temporal and geographical coverage.[3] However, the EB is not designed to provide representative samples on the subnational level.[4] In this paper, we use data from regionally representative surveys to evaluate how well we can estimate regional preferences from standard EB data. Our goal is to develop a model that can be applied as a default strategy to estimate various regional political preferences across Europe. Our evaluation study finds that a Bayesian additive regression tree (BART) with synthetic post-stratification provides the best predictions, while requiring less researcher intervention.

Predicting regional preferences

In recent years, political scientists have developed various tools to estimate regional public opinion from nationally representative surveys. Below, we shortly introduce the available strategies. Since, we evaluate a number of alternative strategies, we have limited space to provide detailed explanations of the methods. For more extensive discussions we want to refer readers to literature referenced below.

[3]It is important to note that the Eurobarometer comes with some other issues regarding methodological transparency, the quality of field work, and measurement validity (Saris and Kaase 1997; Harkness 1999). Other surveys, such as the European Social Survey, are stronger in this regard, but are more limited in their scope.

[4]In this paper, the subnational/regional level refers to the NUTS 2 (Nomenclature of territorial units for statistics) level, which describes the basic regions for the application of regional policies.

Disaggregation

Disaggregation is the simplest strategy to estimate regional preferences from nationally representative survey data. National survey data is disaggregated to the regions respondents reside in. For each region, the average public opinion is calculated, including survey weights if available. This approach is only valid if national surveys have prohibitively large samples or regionally targeted sampling procedures (Hanretty, Lauderdale, and Vivyan 2018).

Global smoothing in multi-level regression

Current approaches to deal with the noisiness of regional information in the standard EB data apply multi-level regression (Mr) (Pittau, Zelli, and Gelman 2010, e.g.,). The hierarchical structure in the standard EB gives rise to a two-level model, with random effects for countries and regions. Regional estimates should improve due to partial-pooling, as sparsely sampled regions more strongly rely on their respective country average. Additionally, good regional predictors should improve predictions by making regions more comparable (Hanretty, Lauderdale, and Vivyan 2018). Regional predictors are also an important extension to the more complicated multi-level approaches introduced below. However, a major short-coming of these methods refers to the selection of regional co-variates. Eurostat, for example, offers a rich set of variables on the regional level. Selecting a regional predictor frequently seems arbitrary and brings the risk of over-fitting the data.

Multi-level regression with classical post-stratification

An extension of simple Mr is multi-level regression with post-stratification (MrP). This method gained increasing attention in the past decade and is currently the standard approach for estimating subnational preferences from nationally representative data (Buttice and Highton 2013; Hanretty, Lauderdale, and Vivyan 2018; Lax and Phillips 2009; Leemann and Wasserfallen 2017; Toshkov 2015; Warshaw and Rodden 2012). MrP comes with the advantages of simple Mr, but additionally allows to correct for systematic differences between the estimation sample and the true population (Hanretty, Lauderdale, and Vivyan 2018; Lax and Phillips 2009). Multi-level models are estimated with a set of demographic characteristics (e.g., gender, education, age). Predictions from the multi-level model are then weighted with census data on the true population structure of the demographics.

Data requirements for classical MrP are quite demanding. Researchers need regional census data on all individual level predictors included in the multi-level model. This comes with a number of challenges. First, census data needs to be congruent with items in the survey data. Education, for example, is measured in the Eurobarometer as the age when respondents finished full-time education. This is not comparable to regional census data from Eurostat, which reports the highest level of educational attainment. Second, regional census data is only available for a limited set of variables, which are not necessarily good predictors of certain political preferences. Yet, the gains of post-stratification depend on the extent to which individual level predictors explain regional differences (Leemann and Wasserfallen 2017).

Multi-level regression with synthetic post-stratification (MrsP)

A recent advancement in the methodological literature is multi-level regression with synthetic post-stratification (MrsP). MrsP addresses some of the weaknesses of classical MrP. More specifically, MrsP allows researchers to use individual level variables that are powerful predictors, but are not available from census datasets. The inclusion of such powerful predictors can greatly enhance predictions of regional preferences.

Consider an example where one specifies a model with gender, education, and occupational status as individual level predictors. Classical MrP requires regional census data on all three predictors for post-stratification: meaning, true joint probabilities. MrsP does not need such detailed data. For MrsP, one collects regional data on the relative gender, education and occupation shares (regional margins). Joints are created by multiplying margins for each logical combination of the categories of individual level predictors. These are called synthetic joints because multiplication assumes that individual level predictors are independent from each other. However, recent research has found that synthetic joints work very well, even if the zero-correlation assumption is violated (Leemann and Wasserfallen 2017).

A major restriction to MrsP is - again - model building. For example, a good individual-level predictor of EU trust is not necessarily a good predictor of redistributive preferences. To build general models, it therefore is important to select variables for synthetic joints that are potentially informative for many political preferences. Otherwise, one risks building idiosyncratic models that do not travel well to other settings.

Post-stratified BART

A recently proposed alternative to the regression-based methods introduced above is a Bayesian additive regression tree (BART) (Chipman, George, and McCulloch 2012; Montgomery and Olivella 2018). BART is an ensemble method that estimates many decision trees to subsequently learn the structure of the data. More specifically, BART uses a Bayesian back-fitting algorithm in which each tree tries to accommodate the residuals of the previous tree. The overall fit is derived by regularizing across all trees, which makes each tree a weak learner in the ensemble. There are two major advantages of this tree-based method. First, it is not fitting a global model, but rather estimates highly interactive effects across partitions of the data (a classical tree feature). BART therefore implements deep interactions to predict regional preferences as envisioned by Ghitza and Gelman (2013). Second, it selects the best predictors and their interactions in an iterative process. This allows researchers to provide a set of predictors and let the algorithm select and interact them in a non-parametric way, just as the data calls for them. As this is done across an ensemble of trees (e.g., 200 trees), we avoid over-fitting the data. BART predictions can also be post-stratified with classical or synthetic joints. This provides a further safeguard against bad predictions.

Of course, there are also shortcomings to BART. First, it is a black box method which is more difficult to interpret compared to multi-level regression models. Second, the Bayesian back-fitting algorithm is computationally more costly. Finally, a few hyper-parameters have to be set by the researcher before running the BART model. Researchers have to define the number of trees the BART model should estimate. Here, Chipman, George, and McCulloch (2012) have shown that 200 trees provide reliable predictions. Further, researchers can set parameters for the priors to control the influence of individual trees

on the overall fit. We have chosen parameters to ensure a conservative fit by limiting the influence of individual trees.[5] However, we have varied all hyper-parameters and the resulting predictions remain very similar.

Data

We are using the 2015 Standard Eurobarometer 83.3 to predict regional EU trust and compare the results with the regionally representative data from the 2015 Flash Eurobarometer. Trust in the European Union is the only political preference that is congruently surveyed across both datasets. The 2015 Flash EB provides regionally representative public opinions as the sampling procedure is based on randomly drawn telephone numbers from local registers. Moreover, the Flash EB samples 300 respondents per region, which reduces the statistical error substantially compared to Standard EB surveys.[6] Table A1 in the Appendix provides a more detailed comparison of the sampling procedures. After cleaning and aligning the two datasets, we are able to analyze EU trust across 184 EU regions and 22 member states. Note that we have to omit the very small countries, which do not show subnational variation (e.g., the Baltic states, Malta, Cyprus, and Luxembourg). Depending on the co-variates used for post-stratification, the number of regions decreases to 154. The intra-regional correlation (ICC) for EU trust in the EB 83.3 is 0.04 on the regional and 0.06 on the country level.[7] This shows that subnational variation is around two-thirds of the size of the between-country variation.

[5]For this, we chose a high value of ten for the "k" parameter (Chipman, George, and McCulloch 2012).

[6]Previous studies have also used representative population estimates to validate the performance of MrP (Toshkov 2015).

[7]See Table A2 in the Appendix

We add regional-level predictors from Eurostat to improve the model fit. We retrieve complete regional data on GDP per capita, area size, population density, tertiary education shares, and median household income. We use all regional predictors in the BART model, because the Bayesian back-fitting algorithm does select on the most relevant predictors. However, we have to select manually some variables for the multi-level regressions to avoid over-fitting. Therefore, we choose a parsimonious set of predictors for the regression-based methods. Here, we account for the effect of economic conditions on EU trust (Foster and Frieden 2017), adding the regional GDP. Moreover, we add a variable measuring whether a region lies in eastern, western, northern or southern Europe ("grand region").[8]

For post-stratification, we implement two strategies. First, we apply classical post-stratification using regional demographics from the 2011 census conducted by Eurostat (classical MrP). The 2011 EU census provides a limited set of demographics for all EU regions. The variables that are congruent with the standard EB are gender, age, and marital status. Multi-level regression results are therefore post-stratified by all logical combinations of two gender, six age (ten year groups from 15 years onward, with the last category being "65 years and older"), and four marital status (single, widow, (re-)married, divorced or separated) categories.

Second, we implement synthetic post-stratification by adding a promising individual-level predictor. We were looking for a predictor that potentially speaks to many political preferences. This is to ensure a certain generality of the model and avoid over-fitting. We

[8]Note that we compare models that do not have the same set of predictors. One might suggest that a "fair" comparison would require identical inputs. However, the different estimation strategies vary in their ability to capitalize on the available data. As our goal is to develop general models that provide best predictions, we provide each strategy with the adequate data. This also highlights the advantages and disadvantages of the different strategies.

chose the general left-right measure, which is surveyed in the EB regularly. We group respondents into the three categories "left", "center", and "right", and assume that political preferences are informed by the divides across these ideological camps. Margins on left-right self-placement are taken from the regionally representative 2013 Quality of Government survey [9]. Combining this with regional margins on gender, age, and marital status from the census data allows us to create a table of synthetic joints. Table A3 in the Appendix provides an overview of the data sources and coding schemes for the level-1 predictors.

Table 2 presents the different strategies and predictors used in the analysis. Note that the multi-level models are estimated with a reduced set of regional predictors. Contrarily, the BART models include the full set of regional predictors. Here, all regional predictors are country-mean centered to capture the within-country variation. Further, to account for variation between countries, we add the country mean of each regional predictor. This results in a set of 10 level-2 predictors for the BART models.

TABLE 1 *Overview of predictors and strategies*

Strategy	Level-1 Predictors	Level-2 Predictors
MrP	gender, age, marital status	region, grand region, country, gdp
MrsP	gender, age, marital status, left-right	region, grand region, country, gdp
BART CJ	gender, age, marital status	gdp, population density, area size, tertiary education rate, median household income
BART SJ	gender, age, marital status, left-right	gdp, population density, area size, tertiary education rate, median household income

Note: Regional predictors for the BART models enter the algorithm group-mean centered and as country mean. BART CJ = BART with classical joints; BART SJ = BART with synthetic joints

[9]https://qog.pol.gu.se/

Evaluation results

Figure 1 presents the results of the different estimation strategies with post-stratification (results without post-stratification can be found in Figure A1 of the Appendix). Note that the results are based on the 154 regions for which we have complete data on all variables. For each estimation strategy, we plot the predicted regional mean against the regional representative mean. The 45 degree reference line indicates perfect correspondence. Disaggregation performs worse. The average deviation between predicted and representative values is at 12 percent. Only a third of the predicted values fall into the confidence intervals of the representative values. Table A4 in the Appendix demonstrates that this large degree of noise is present across different datasets and dependent variables. Overall, we find that every method improves upon simple disaggregation.

A multi-level model with synthetic post-stratification leads to moderate improvements. On average, the predicted mean is off by 9 percentage points from the representative mean. RMSE suggests that estimations are even 12 percentage points off on average. Yet, 44 percent of predicted means now fall within the confidence interval of the representative mean, which is a large improvement. Using only census data for post-stratification - as in classical MrP - improves predictions, though only marginally. Error terms indicate between 8 and 10 percentage points difference and coverage rises to 47 percent. It seems that the left-right orientation does not help the predictive performance of the regression model. Here, classical joints from the census more accurately predict EU trust. Of course, this finding can change for different dependent variables.

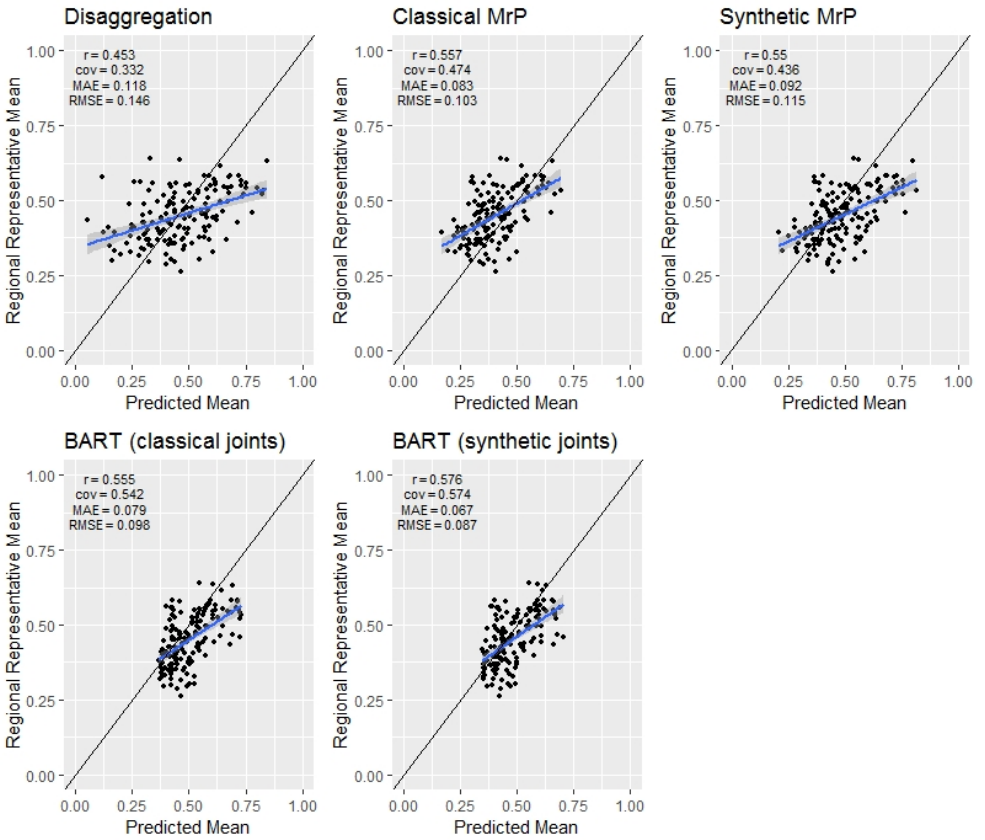
A BART model with synthetic post-stratification and regional predictors provides the best predictions. We get a RMSE of around 0.09 and an MAE of 0.07. This is a substantial gain in precision. RMSE, for instance, is reduced by 40 percent compared to

disaggregation. Coverage increases by 25 points to 58 percent. Now, a clear majority of predictions fall within the confidence interval of the representative means. A BART model with classical joints is somewhat less accurate. Here, coverage is at 54 percent and errors are larger. Even though BART does select on the most relevant predictors, we do not see large gains by adding regional covariates. Most likely, the available regional covariates are not very informative of the outcome. We also see that post-stratification does not make a difference. The raw predictions perform similar to the post-stratified predictions (compare Figure A1 in the Appendix).[10]

A number of insights emerge from our evaluation study. Prediction gains will always depend on the specific modeling strategy. It could well be that there are much better models than the ones we estimated, for example using individual-level predictors that are more closely related to the outcome variable. These, however, run the risk of being idiosyncratically tuned towards the specific outcome and data under investigation. For instance, one could use trust in national political institutions as predictor of EU trust, since these variables have been shown to be substantially related (Muñoz, Torcal, and Bonet 2011). Indeed, we can tune a MrsP model including martial status, age, left-right, and national parliament trust to approach the gains reported under the synthetic BART model. Figure A2 in the Appendix presents the results for this extended MrsP model. The strategy also reaches a MAE of 0.07. This, however, comes at the cost of building a more idiosyncratic model, which could perform worse in new settings.

[10]There are many potential explanations on why post-stratification does not improve predictions. Most likely, the representativeness of the EB is fairly good with regard to the level-1 predictors used. This result is not unique to our analysis. The post-stratified BART application presented in Montgomery and Olivella (2018) does also not gain a lot from re-weighting. Still, post-stratification is a good safeguard against bad predictions.

Figure 1. The Performance of Prediction Strategies



Note: r = Pearson's Rho; cov = Share of predictions that fall within the 95 percent confidence interval of the representative mean; MAE = Mean Absolute Error; RMSE = Root Mean Squared Error

There is less of a concern regarding over-fitting with classical joints, since there is not much of a choice in the selection of individual-level predictors for post-stratification. Currently, only age, gender, and marital status are compatible across the EB and the Eurostat census. Yet, there still is a risk of over-fitting coming from the selection of regional predictors. Here, researchers have two options. Either they use MrP, but then they

should keep the set of regional predictors parsimonious. Alternatively, researchers can use BART with classical joints to automatize the selection of predictors and capitalize on the other advantages of BART (e.g., deep interactions). In our application, this improved predictions, especially with regard to coverage. Also note that a major advantage of classical joints is that they work for the complete case universe. We therefore recommend using BART with classical joints for applications that require a complete case universe.

For applications that can accept a smaller European-wide sample, synthetic BART is most promising. And, of course, synthetic BART will also be most promising if synthetic joints for all EU regions will become available in the future (using our or other sensible individual-level predictors). In comparison to MrP, BART provides better predictions and reduces the risk of over-fitting. Still, a few tuning parameters of the BART model require researcher judgment. However, we have not seen major differences in our predictions when varying them. The Appendix present the performance of synthetic BART in a new setting. Figure A3 in the Appendix compares disaggregation and synthetic BART for estimating regional social trust in Spain. Here, we use regionally representative data from the QoG survey to validate our results. The predictive performance is equally good.

Discussion and outlook

This paper implements a number of strategies to improve regional estimates of public opinion from standard EB surveys. Our evaluation study indicates that BART with synthetic post-stratification performs best, substantially increasing the correspondence between regional estimates and the representative values. MrP and BART with classical joints from the EU census also works well and allow researchers to estimate regional preferences for the complete case universe. We also show that MrsP has great potential,

but increases the risk of over-fitting (see Figure A2 in the Appendix). Here, the selection of predictors remains a challenge, asking for careful judgment by researchers.

The paper demonstrates that a lot of choices in the modeling process are dictated by data availability. There is a limited set of variables in the EU census, the regionally representative surveys, and the EBs. And, of course, there are only a few variables that can be measured consistently across the different data sources. A lot of the limitations in our application come from the fact that we attempt to estimate regional preferences across the whole EU. This is a highly relevant endeavor for European-wide research. However, it should be noted that there is more potential in estimating regional preferences for single countries, using more detailed national surveys.

Considering our results, one might ask whether the presented gains are 'good enough'? Our analysis makes clear that the raw EB data is sub-optimal and that we can improve data quality substantially using estimation procedures. The gains of our analysis are within the range of gains reported in previous evaluation studies (Hanretty, Lauderdale, and Vivyan 2018). Even though the 57 percent coverage reached seems moderate, it substantially exceeds previous results reported for Eurobarometer data (Toshkov 2015). Moreover, Table A5 in the Appendix demonstrates that the reduction in noise affects the substantive interpretations we draw from the data. Table A5 uses regional EU trust as dependent variable and shows that coefficient estimates deviate substantially for the more noisy strategies. However, even our best predictions come with substantial noise. Here, one should consider that our reference values from the Flash EB do have a sampling error.

Even though the substantial noise in EB data is a challenge, we are optimistic that scholars will be able to use the EB's rich regional information. We hope that our analysis lays the ground for the estimation of European-wide regional preferences, enabling more disaggregated comparative research.

References

Beramendi, Pablo. 2012. *The Political Geography of Inequality. Regions and Redistribution.* New York: Cambridge University Press.

Buttice, Matthew K., and Benjamin Highton. 2013. "How does multilevel regression and poststratification perform with conventional national surveys?" *Political Analysis* 21 (4): 449–467. ISSN: 14764989. doi:10.1093/pan/mpt017.

Chipman, Hugh A., Edward I. George, and Robert E. McCulloch. 2012. "BART: Bayesian additive regression trees." *Annals of Applied Statistics* 6 (1): 266–298. ISSN: 19326157. doi:10.1214/09-AOAS285. arXiv: 0806.3286.

Foster, Chase, and Jeffry Frieden. 2017. "Crisis of trust: Socio-economic determinants of Europeans' confidence in government." *European Union Politics* 18 (4): 511–535. ISSN: 17412757. doi:10.1177/1465116517723499.

Georgiadou, Vasiliki, Lamprini Rori, and Costas Roumanias. 2018. "Mapping the European far right in the 21st century: A meso-level analysis." *Electoral Studies* 54, no. 1 (August): 103–115. ISSN: 02613794. doi:10.1016/j.electstud.2018.05.004. https://linkinghub.elsevier.com/retrieve/pii/S026137941830026X.

Ghitza, Yair, and Andrew Gelman. 2013. "Deep interactions with MRP: Election turnout and voting patterns among small electoral subgroups." *American Journal of Political Science* 57 (3): 762–776. ISSN: 15405907. doi:10.1111/ajps.12004.

Hanretty, Chris, Benjamin E. Lauderdale, and Nick Vivyan. 2018. "Comparing Strategies for Estimating Constituency Opinion from National Survey Samples." *Political Science Research and Methods* 6 (3): 571–591. ISSN: 2049-8470. doi:10.1017/psrm.2015.79. http://www.journals.cambridge.org/abstract%7B%5C_%7DS2049847015000795.

Harkness, Janet. 1999. "In pursuit of quality: Issues for cross-national survey research." *International Journal of Social Research Methodology* 2 (2): 125–140. ISSN: 14645300. doi:10.1080/136455799295096.

Jones, K., R. J. Johnston, and C. J. Pattie. 1992. "People, Places and Regions: Exploring the Use of Multi-Level Modelling in the Analysis of Electoral Data." *British Journal of Political Science* 22 (3): 343–380. ISSN: 14692112. doi:10.1017/S0007123400006426.

Lax, Jeffrey R., and Justin H. Phillips. 2009. "How Should We Estimate Public Opinion." *American Journal of Political Science* 53 (1): 107–121. ISSN: 00925853. doi:10.1111/j.1540-5907.2008.00360.x.

Leemann, Lucas, and Fabio Wasserfallen. 2017. "Extending the Use and Prediction Precision of Subnational Public Opinion Estimation." *American Journal of Political Science* 4 (61): 1003–1023. ISSN: 00925853. doi:10.1111/ajps.12319.

Montgomery, Jacob M, and Santiago Olivella. 2018. "Tree-Based Models for Political Science Data." *American Journal of Political Science* 62 (3): 729–744. ISSN: 00925853. doi:10.1111/ajps.12361. http://dx.doi.org/10.7910/.

Muñoz, Jordi, Mariano Torcal, and Eduard Bonet. 2011. "Institutional trust and multilevel government in the European Union: Congruence or compensation?" *European Union Politics* 12 (4): 551–574. ISSN: 14651165. doi:10.1177/1465116511419250.

Pittau, M. Grazia, Roberto Zelli, and Andrew Gelman. 2010. "Economic disparities and life satisfaction in European regions." *Social Indicators Research* 96 (2): 339–361. ISSN: 03038300. doi:10.1007/s11205-009-9481-2.

Rodden, Jonathan. 2010. "The Geographic Distribution of Political Preferences." *Annual Review of Political Science* 13 (1): 321–340. ISSN: 1094-2939. doi:10.1146/annurev.polisci.12.031607.092945. http://www.annualreviews.org/doi/10.1146/annurev.polisci.12.031607.092945.

Saris, W. E., and M. Kaase. 1997. *Eurobarometer: measurement instruments for opinions in Europe*. Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. https://nbn-resolving.org/%7B%5C%%7D0Aurn:nbn:de:0168-ssoar-49742-1.

Schraff, Dominik. 2019. "Regional redistribution and Eurosceptic voting." *Journal of European Public Policy* 26 (1): 83–105. ISSN: 1350-1763. doi:10.1080/13501763.2017.1394901. https://doi.org/10.1080/13501763.2017.1394901.

Toshkov, Dimiter. 2015. "Exploring the performance of multilevel modeling and post-stratification with eurobarometer data." *Political Analysis* 23 (3): 455–460. ISSN: 14764989. doi:10.1093/pan/mpv009.

Warshaw, Christopher, and Jonathan Rodden. 2012. "How should We measure district-level public opinion on individual issues?" *Journal of Politics* 74 (1): 203–219. ISSN: 00223816. doi:10.1017/S0022381611001204.