**University of Amsterdam, AY 2022/2023**

**Mattia Guarnerio, Student Number: 14350920, e-mail:** [**mattia.guarnerio@student.uva.nl**](mailto:mattia.guarnerio@student.uva.nl)

**Final Paper for *Fixed & Random Effects Modelling*; Research Master’s Social Sciences, Year 1, Semester 2; Prof. Dr. T.W.G. van der Meer**

**Euro(di)vision**

*Euroscepticism and Political Ideology: Why Time and Context Matter*

**Table of Contents**

Introduction and Theory, p. 2

Research Design, p. 3

Results: The Empty Model, p. 7

Results: Random-Intercept Within-Between Framework, p. 10

Results: Random-Slopes Within-Between Framework, p. 13

Results: Robustness Checks, p. 16

Conclusions and Limitations of the Research, p. 20

Appendix, p. 21

References, p. 26

**Introduction and Theory [357 words]**

Since the aftermath of Brexit, European decision-makers, opinion leaders, and academics alike have raised their concerns regarding the survival of European Union’s political project of supra-national integration. Drawing upon the so-called cue-taking and benchmarking approaches, I contend that European constituents ground their attitudes towards such effort on national proxies – i.e., cues, or benchmarks – to overcome their information shortfalls, caused by the issue’s inherent complexity (Hobolt & De Vries, 2016). Thus, in this paper my substantive interest is studying the influence of the nation-based, contextual proxies of the political left and right on the European citizens’ support for further unification at the individual level. More specifically,

**RQ:** How does the national-level left-right political leaning of anEU member country autonomously impact individual support for further integration in the European Union?

Althoughthe existing literature shows that Eurosceptic tendencies pertain to both fringes of the left-right spectrum (Van Elsas et al., 2016) in the last two decades extreme right-wing Euroscepticism has become ever more salient (Vasilopoulou, 2018), with empirical evidence suggesting that time is a key component in shaping the relationship between political ideology and Euroscepticism (Van Elsas & Van Der Brug, 2015). Moreover, as radically conservative claims are grounded on the opposition to any European involvementthat transcendsthe national interest, I expect right-leaning national cues to be more impactful on individual-level beliefs concerning the European Union’s integration and more influential among citizens witha pre-existing rightward political orientation.Consequently,

**H1:** On average and ceteris paribus, citizens living in countries that are structurally right-leaning show less support for further integration in the European Union, even when controlling for their individual-level left-right political positioning.

**H2:** On average and ceteris paribus, citizens living in countries that are increasingly right-leaning show less support for further integration in the European Union, even when controlling for their individual-level left-right political positioning.

**H3:** On average and ceteris paribus, the country-level effect of structurally right-leaning national political contexts is stronger on those citizens with a pre-existing rightward political orientation.

**H4:** On average and ceteris paribus, the country-level effect of increasingly right-leaning national political contexts is stronger on those citizens who express a rightward political orientation.

**Research Design and Methods [592 words]**

I employ a subset (N = 98664) of the Rounds 4 (2008), 6 (2012), and 9 (2018) of the European Social Survey (ESS ERIC, 2018a, 2018b, 2021b) a source that provides open-access, high quality dataconcerning individual-level political beliefs and socio-demographic attributes, characterised by random probability sampling and high response rates. It contains information on persons aged 15 and over and resident within private households in 23 EU member countries**1**: Belgium, Bulgaria, Croatia**2**, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy**3**, Latvia**4**, Lithuania, The Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom**5**. I integrate my subset with national-level indicators on economic performance and perception of corruption provided by the World Bank (2021) and Transparency International (2008, 2012, 2018).

I operationalise personal support for further integration in the European Union as the response to the question: “some say European unification**6** should go further. Others say it has already gone too far. […] what number on the scale best describes your position?” (ESS ERIC, 2021a). Thus, the dependent variable is expressed on a Likert scale ranging from 0 – “Unification has already gone too far" – “Unification should go further” (see Table 1). I analysesuch individual-level outcome in a three-level hierarchical data structure, with respondents at Level-1 (N = 98664), surveys nested within countries at Level-2 (time-variant, J = 66) and countries at Level-3 (time-invariant, C = 23).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 1 | Level-1 (individual, N = 98664) Variables (ESS ERIC, 2018a, 2018b, 2021b) | | | | | |
| Concept | Analytical Role | Label | Name | Type | Range |
| Respondent | Level-1 identifier | Respondent's identification number | idno | Numeric integer | / |
| Respondent's support for European integration | Dependent variable | European Union: European unification go further or gone too far | euftf | Categorical ordinal | 0 | Unification already gone too far to 10 | Unification go further |
| Interaction effect of living in increasingly right-wing countries | Independent variable | Interaction of individual-level political positioning with within-country, time-variant component | within\_lr\_int | Continuous | -2.241 to 2.340 |
| Interaction effect of living in structurally right-wing countries | Independent variable | Interaction of individual-level political positioning with between-country, time-invariant component | between\_lr\_int | Continuous | -3.390 to 3.267 |
| Respondent's political positioning | Control variable | Placement on left right scale, grand mean-centred | lrscale\_m | Categorical ordinal | -5.108 | Leftmost to 4.913 | Rightmost |
| Respondent's gender | Control variable | Respondent is female | female | Binary | '0' | Male '1' | Female |
| Respondent's age | Control variable | Age of respondent, grand mean-centred | agea\_m | Continuous | -36.277 to 53.839 (years) |
| Respondent's education | Control variable | Years of full-time education completed, grand mean-centred | eduyrs\_m | Continuous | -13.072 to 46.928 (years) |

Since I am substantively interested in both structural and over-time aspects of living in an EU member country, the fixed effects approach, widely regarded as the “gold standard” default (Schurer & Yong, 2012) in the political science field, would be inappropriate for my research because it can only derive within-effects. Instead, I employ the within-between random effects specification suggested by Bell, Fairbrother, and Jones (2019)**7**. This formulation allows to explicitly modeltime-variant within-country effects, and time-invariant between-country effects, without assuming they are equal, as in the mainstream random effects framework, which yields an uninterpretable weighted average of the two estimates (Bell & Jones, 2015).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 2 | Level-2 (country, time-variant, J = 66) Variables  (ESS ERIC, 2018a, 2018b, 2021b; The World Bank, 2021; Transparency International, 2008, 2012, 2018) | | | | | |
| Concept | Analytical Role | Label | Name | Type | Range |
| Surveys nested within countries | Level-2 identifier**8** | Respondent's country \* wave (Level-2) ID number | cntry\_wave | Categorical nominal | 401 | Belgium, 2008 to 923 | United Kingdom, 2018 |
| Time-variant within-country effect of political positioning | Independent variable | Time-variant divergence from time-invariant mean of  placement on left right scale at the country level | cntry\_lrscale\_diff | Continuous | -0.458 | leftmost, Italy, 2012 to 0.250 | rightmost, Croatia, 2008 |
| Time-variant within-country effect of nation's economic performance | Control variable | Time-variant divergence from time-invariant mean of GDP per capita, PPP (current international $, standardized) | gdp\_z\_diff | Continuous | -0.568 sd |  Ireland, 2012  to 1.043 sd |  Ireland, 2018 |
| Time-variant within-country effect of perception of corruption in country | Control variable | Time-variant divergence from time-invariant mean of Corruption Perceptions Index in country | cpi\_diff | Continuous | -8.312 | Poland, 2008 to 6.333 | Lithuania, 2018 |

Accordingly, the two main independent variables are the time-variant divergence from the time-invariant mean of placement on a left-to-right Likert scale (see Table 2), and the time-invariant average placement on a left-to-right Likert scale (see Table 3, p. 5), at the country level. Moreover, I evaluate **H3** and **H4** by including cross-level interactions (see Table 1, p. 3) among these two predictors and the Level-1 measurement of placement on a left-to-right Likert scale, which is additionally employed as a control variable for assessing **H1** and **H2**. Other Level-1 confounders are the respondent’s age, and years of full-time education completed (see Table 1, p. 3). In line with the within-between random effects framework, at Levels 2 and 3 I control for time-variant divergences from time-invariant country means, and time-invariant country averages, of standardized GDP per capita, adjusted for PPP**9**, and the Corruption Perceptions Index**10** (see Tables 2 and 3, pp. 4-5).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 3 | Level-3 (country, time-invariant, C = 23) Variables  (ESS ERIC, 2018a, 2018b, 2021b; The World Bank, 2021; Transparency International, 2008, 2012, 2018) | | | | | |
| Concept | Analytical Role | Label | Name | Type | Range |
| Country | Level-3 identifier**11** | Respondent's country ID number | cntry\_num | Categorical ordinal | 1 | Belgium, to 23 | United Kingdom |
| Time-invariant between-country effect of political positioning | Independent variable | Average placement on left right scale at the country level  (time-invariant) | cntry\_lrscale | Continuous | -0.612 | leftmost, Germany, to 0.665 | rightmost, Latvia |
| Time-invariant between-country effect of nation's economic performance | Control variable | GDP per capita, PPP (current international $,  standardized, time-invariant) | gdp\_z | Continuous | -1.755 sd | Bulgaria to 1.899 sd | Ireland |
| Time-invariant between-country effect of perception of corruption in country | Control variable | Corruption Perceptions Index (time-invariant) | cpi | Continuous | -26.397 | Bulgaria to 24.287 | Denmark |

To address the anti-conservativeness of the confidence intervals computed by standard maximum likelihood (ML) methods that rely on the assumption of normally distributed test statistics (Stegmueller, 2013) on models with less than thirty upper-level clusters – i.e., countries, time-invariant – I ground all population inferences on Patterson and Thompson’s (1971) restricted maximum likelihood (REML) estimation, implementing the so-called *m − l − 1* denominator-degrees-of-freedom (DDF) adjustment heuristic by hand**12**, as recommended by Elff et al. (2021)**13**. Sinceit is an intuitive and computationally very fast technique and given the marked disparity in sample sizes between Level-2 units and Level-1 observations, I apply it to the time-variant effects as well.

**1.** I exclude Austria, Greece, Luxembourg, Malta, and Romania from my analytic sample. For these countries, there is either no information in the European Social Survey regarding the years of substantive interest – i.e., Luxembourg, and Malta – or there is too much missing data at Level-2, with only one survey available for 2008, 2012, or 2018 – i.e., Austria, Greece, and Romania.

**2.** Croatia has missing data at Level-2 for 2012. However, since in multi-level-oriented, political science research it is key to include the highest possible number of upper-level clusters – i.e., in this instance, countries, at Level-3 – to aid the practitioner in carrying out appropriate statistical inference, I deem the absence of only one survey concerning the years of substantive interest to be an acceptable information loss. I check whether my findings are robust to the more restrictive assumption of no missing data at Level-2 in the *Results: Robustness Checks* section (see pp. 16-19).

**3.** Italy has missing data at Level-2 for 2008. However, since in multi-level-oriented, political science research it is key to include the highest possible number of upper-level clusters – i.e., in this instance, countries, at Level-3 – to aid the practitioner in carrying out appropriate statistical inference, I deem the absence of only one survey concerning the years of substantive interest to be an acceptable information loss. I check whether my findings are robust to the more restrictive assumption of no missing data at Level-2 in the *Results: Robustness Checks* section (see pp. 16-19).

**4.** Latvia has missing data at Level-2 for 2012. However, since in multi-level-oriented, political science research it is key to include the highest possible number of upper-level clusters – i.e., in this instance, countries, at Level-3 – to aid the practitioner in carrying out appropriate statistical inference, I deem the absence of only one survey concerning the years of substantive interest to be an acceptable information loss. I check whether my findings are robust to the more restrictive assumption of no missing data at Level-2 in the *Results: Robustness Checks* section (see pp. 16-19).

**5.** The United Kingdom is not a EU member country since 2020. However, since in multi-level-oriented, political science research it is key to include the highest possible number of upper-level clusters – i.e., in this instance, countries, at Level-3 – to aid the practitioner in carrying out appropriate statistical inference, and this study concerns years from 2008 to 2018, when Great Britain was still part of the European Union, I deem its incorporationin the analytic sample to be justifiable. I check whether my findings are robust to influential Level-2 and Level-3 clusters in the *Results: Robustness Checks* section (see pp. 16-19).

**6.** In Note 17 of the ESS Wave 9’s main questionnaire, the authors specify that “unification refers to further integration rather than further enlargement” (ESS ERIC, 2021a, p. 13). Therefore, I can safely assume there is a substantive equivalence between the two terms, and that the measured social phenomenon is indeed support towards further EU integration.

**7.** This random-effects formulation is obtained by subtracting the time-invariant grand mean – i.e., Level-3 – from the time-variant values at Level-2, generating a variable that measures time-variant divergences from time-invariant means. I employ this de-meaned version of the within-between specification, and not the Mundlak model (1978), because I am not interested in controlling for the raw value of the time-varying predictor. Furthermore, it would be substantively meaningless to estimate the effect of a survey – i.e., Level-2 – moving from a country – i.e., Level-3 – to another, by holding Level-2 characteristics constant. Thus, I wish to compute the between effect, and not the contextual effect (Bell et al., 2019).

**8.** In the analysis, I also employ the corresponding character labels, saved in the *cntry\_wave\_char* variable, when needing a more interpretable data visualisation.

**9.** Adjusting for purchasing power parity (PPP) accounts for the disparities in price levels across different countries and their respective currencies.

**10.** The Corruption Perceptions Index (CPI) is an indicator that aims to measure the perceived levels of corruption in different countries around the world, on a numeric scale ranging from 0 to 100, with higher scores representing greater transparency. It is grounded on evaluations gathered from experts and several reputable sources – i.e., international organizations, research institutions, and business entities. Recent literature shows that perception of corruption within the national context impacts the subject-specific development Eurosceptic beliefs and tendencies (Baute, 2023).

**11.** In the analysis, I also employ the corresponding character labels, saved in the *cntry* variable, when needing a more interpretable data visualisation.

**12.** This heuristic provided by Elff et al. (2021) is an approximation of the Satterthwaite (1946) and Kenward-Roger (Kenward & Roger, 1997) Denominator-Degrees-of-Freedom (DDF) adjustment techniques. When dealing with simple multi-level data structures, the practitioner can ground the p-value calculations on the Student t’s two-tailed distribution with *m* – *l – 1* degrees of freedom, with *m* as the number of upper-level clusters, *l* equalling the number of contextual variables. In the statistical software Stata, this correction must be carried out by hand, as Elff et al. (2021) illustrate in the Online Appendix to their article.

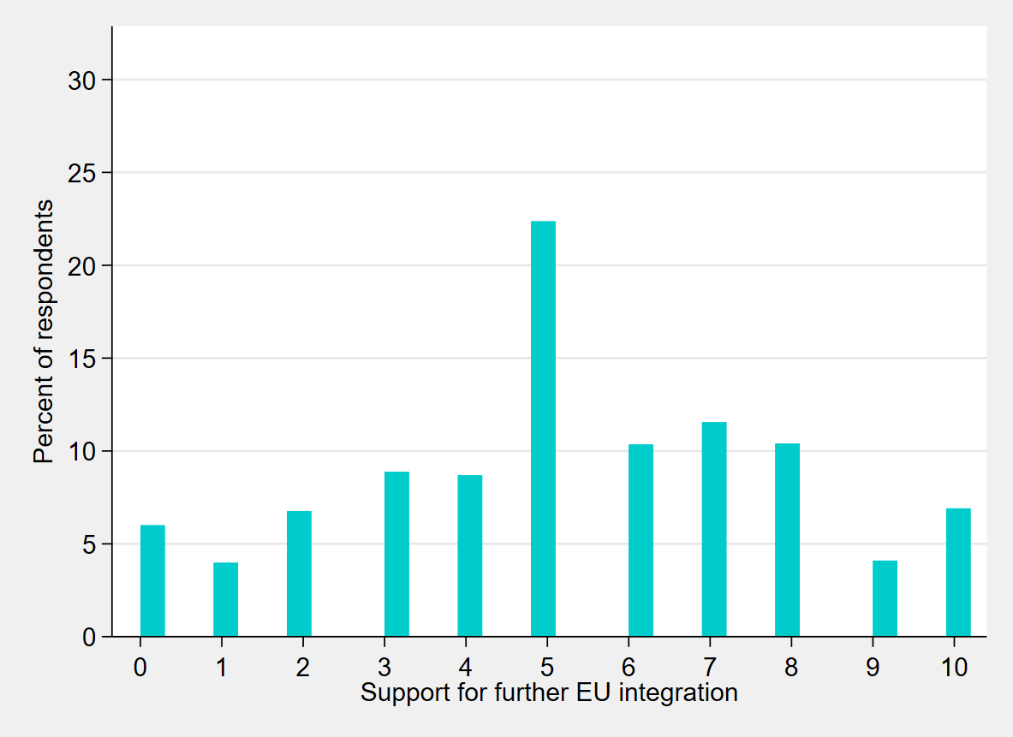
**13.** The more accurate Satterthwaite (1946) and Kenward-Roger (Kenward & Roger, 1997) DDF adjustment techniques can be utilised within the framework of the *mixed* Stata command, but they are far too computationally demanding for the hardware at my current disposal.

**Results: The Empty Model [252 words]**

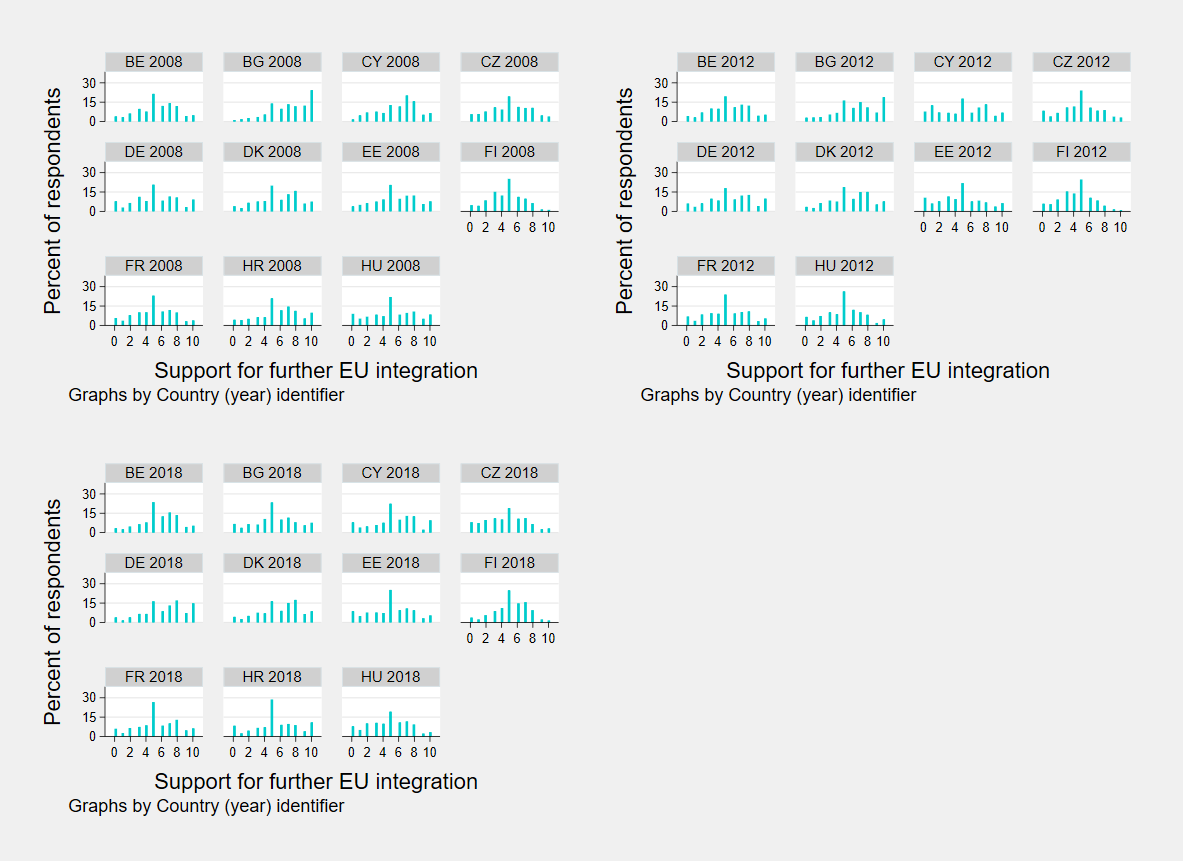


I fit a variance-components model – i.e., an empty model (see Table 4) – to quantify the amount of variance observable at each hierarchical level of the data structure, by not including any independent or control variables in the specification (Rabe-Hesketh & Skrondal, 2008, pp. 77–78). The degree of support for more integration in the European Union is at an overall mean of 5.304, on a Likert scale ranging from 0 to 10, denoting that, on average, respondents from the 23 selected EU member countries slightly lean towards approving furtherunification, rather than opposing it. The Intra-Class Correlation (ICC), which measures the similarity of individuals within the same higher-level unit – i.e., surveys, or countries – equals0.029 at Level-3, and 0.060 at Level-2. Most of the diversity in the outcome variable isexplained by individual-level differences, but approximately 2.9% of the total is due to time-invariant, between-country attributes. Moreover, when accounting for Level-3 clustering, around 6.0% of the remaining variability is situated at thewithin-surveylevel. These numbers are small, yet non-negligible**14**. Accordingly, the likelihood ratio test (see Table 5) yields a naïve p-value that is extremely close to zero, indicating that the hypothesis of equal fit between a single-level and a multi-level model must be rejected**15**, thus suggesting that takingnestinginto consideration provides better fit to the data. The small fluctuations can be visualised in Figures 1, 2a, 2b, and 3 (see pp. 8-9), which respectively show the Level-1, Level-2, and Level-3 distribution of the degree of support towards further integration in the European Union.

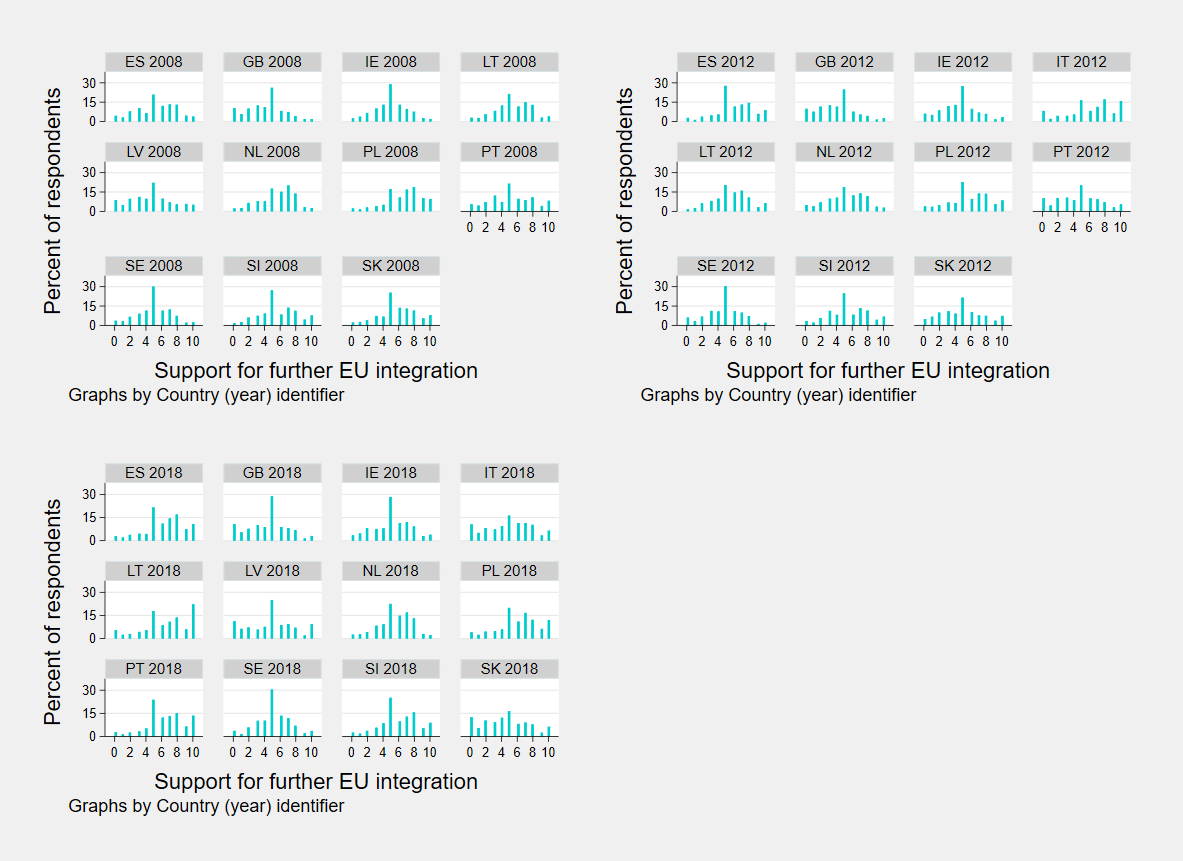


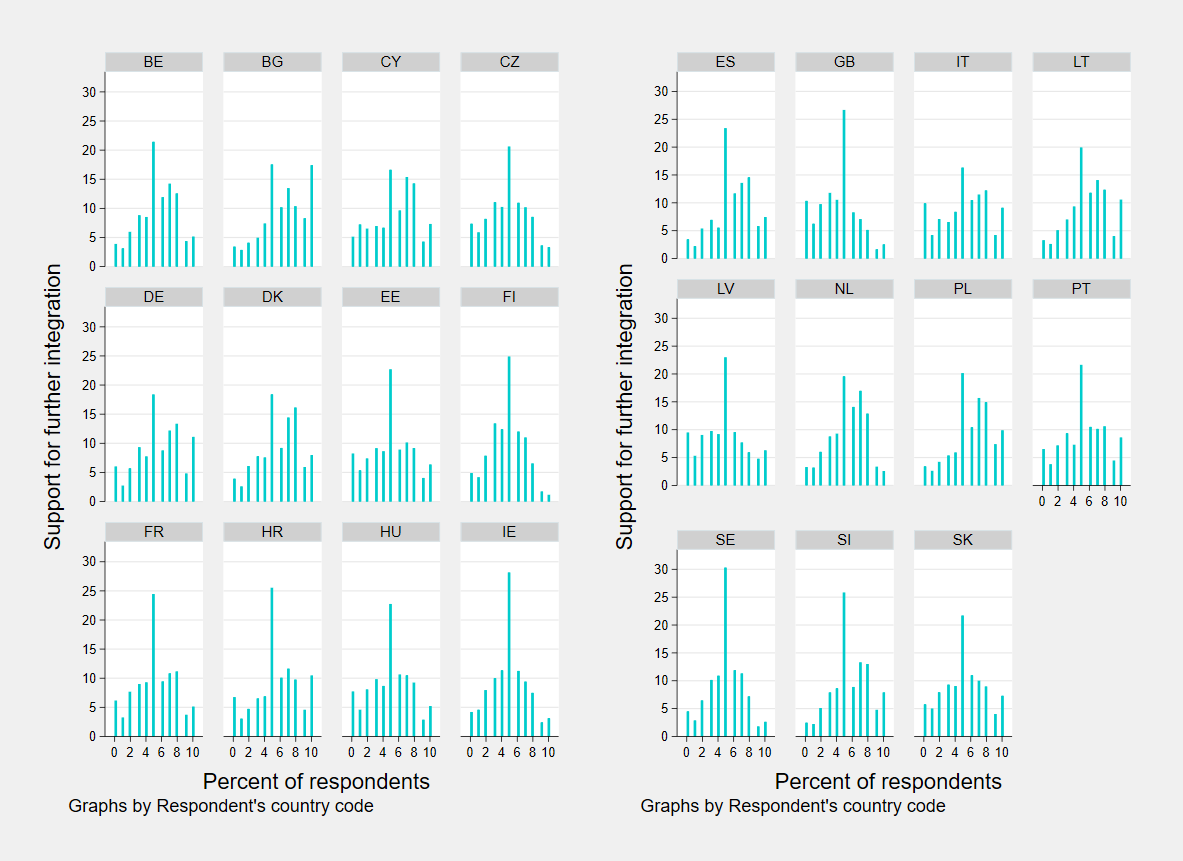
****

**Figure 1 –** *Level-1 (individual) distribution of degree of support towards further EU integration*



**Figure 2a –** *Level-2 (country, time-variant) distribution of degree of support towards further EU integration, showing EU member countries from Belgium to Hungary*

****

**Figure 2b –** *Level-2 (country, time-variant) distribution of degree of support towards further EU unification, showing EU member countries from Spain to Slovakia*

**Figure 3 –** *Level-3 (country, time-invariant) distribution of degree of support towards further EU integration*

**14.** 95% confidence intervals which show that both Intra-Class Correlations are non-null in the population are available on request.

**15.** Since the null hypothesis is on the boundary of the parameter space, the reported p-value is conservative. However, if such naïve p-value for the conservative test already leads to the rejection of the null hypothesis at the 5% level of significance, then the appropriately adjusted p-value, being invariably lower, will not substantially change an interpretation concerning model fit.

**Results: Random-Intercept Within-Between Framework [458 words]**

I fit a linear random-intercept model in line with the within-between formulation to addressthe unrealistic assumption that individual-level degrees of support towards further EU integration expressed by participants who took the same survey, or live in the same country, are uncorrelated given the observed covariates. In other words, I split the total residual, ξijc, into three error components: ζjc, shared among respondentsto the same ESS Wave, ζc, common to all co-nationals, and εijc, unique for everyone. Consequently, each observation has a subject-specific intercept β1 + ζjc + ζc (Rabe-Hesketh & Skrondal, 2008, pp. 127–128). Together with Level-1, Level-2, and Level-3 exogeneity**16**, it is assumed that the variance of the Level-1 residual is homoscedastic given the covariates and the random intercepts, the variance of the Level-2 random intercept is homoscedastic given the covariates and the Level-3 random intercept, and the variance of the Level-3 random intercept is homoscedastic given the covariates. Moreover, it is supposed that the Level-1 residuals are uncorrelated for any two Level-1 units – i.e., two respondents – given the covariates and the random intercepts, Level-2 random intercepts are uncorrelated for any two Level-2 clusters – i.e., two surveys – given the covariates and the Level-3 random intercept, and Level-3 random intercepts are uncorrelated for any two Level-3 clusters – i.e., countries, time-invariant – given the covariates (Rabe-Hesketh & Skrondal, 2008, pp. 128–129). I do not include cross-level interactions to test for **H3** and **H4** because multi-level models that do not allow for random slopes on the lower-level components of the latter often lead toseverely anti-conservative statistical inferences (Heisig & Schaeffer, 2019).



Substantively meaningful results for the random-intercept within-between framework are reported in Table 6**18** (see p. 11). As I expected, on average and ceteris paribus, citizens living in countries that are increasingly right-leaning show less support for further unification in the European Union, even when controlling for their individual-level left-right political positioning. The effect is significant at the conventional 5% level – i.e., p-value < 0.05**19** – thus providing corroboration to **H2**. Contrariwise, although the time-invariant, between-country effect of residing in a structurally right-leaning country on the subject-level endorsement for EU integration is negative, it is not significant in the population**20**. Therefore, **H1** is not substantiated by empirical evidence. It is important to note that within and between effect sizes cannot be straightforwardly confronted, as it is still under debate whether the specification allows for such comparison**21**. Turning to Level-1 and Level-2 variances, these are all reduced relative to the empty model, while Level-3 variance remains unchanged. Accordingly, the Level-3 Intra-Class Correlation is approximatively equivalent, while the Level-3 Intra-Class Correlation is unaltered (see Table 4, p. 7). This reflects the finding that including covariates in the formulation helpsto account for a non-negligible amount of residual variation in the outcome variable at Level-1 and Level-2 – i.e., respondents, and surveys. On the other hand, Level-3 variances and ICCs suggest that structural, country-specific factors continue to contribute a consistent amount of variation to the outcome variable even after accounting for the covariates at lower levels, implying that the time-invariant, between-country effects have a relatively stable impact on support towards further EU integration, unaffected by the specification of lower-level predictors.

**16.** Level-1 exogeneity is the assumption that the Level-1 residual εijc has zero expectation or mean, given the covariates and the random intercepts. Analogously, Level-2 exogeneity is the supposition that the survey-specific random intercept ζjc has zero expectation or mean given the covariates and the Level-3 random intercept. Level-3 exogeneity adheres to the same underlying principle, whereby the Level-3 random intercept ζc is postulated to have zero expectation or mean given the covariates.

**17.** This heuristic provided by Elff et al. (2021) is an approximation of the Satterthwaite (1946) and Kenward-Roger (Kenward & Roger, 1997) Denominator-Degrees-of-Freedom (DDF) adjustment techniques. When dealing with simple multi-level data structures, the practitioner can ground the p-value calculations on the Student t’s two-tailed distribution with *m* – *l – 1* degrees of freedom, with *m* as the number of upper-level clusters, *l* equalling the number of contextual variables. In the statistical software Stata, this correction must be carried out by hand, as Elff et al. (2021) illustrate in the Online Appendix to their article.

**18.** I only report results for the intercept, and Level-2 and Level-3 independent variables (see Tables 2 and 3, pp. 4-5). Fixed-effects coefficients and inferences for confounders are available upon request.

**19.** Throughout this paper, significance is invariably established at the 5% level.

**20.** The utilization of robust p-value estimations does not alter the interpretation that would have been derived from employing standard inferential statistics.

**21.** Furthermore, in this research the within- and between-country effect sizes are quite uninformative even in absolute terms. Since the range of values taken by the main Level-2 and Level-3 independent variables in the analytic subsample is very narrow, the estimations of the fixed-effects parameters are inflated, while their actual effect on the 0 to 10 Likert scale is smaller than it might appear.

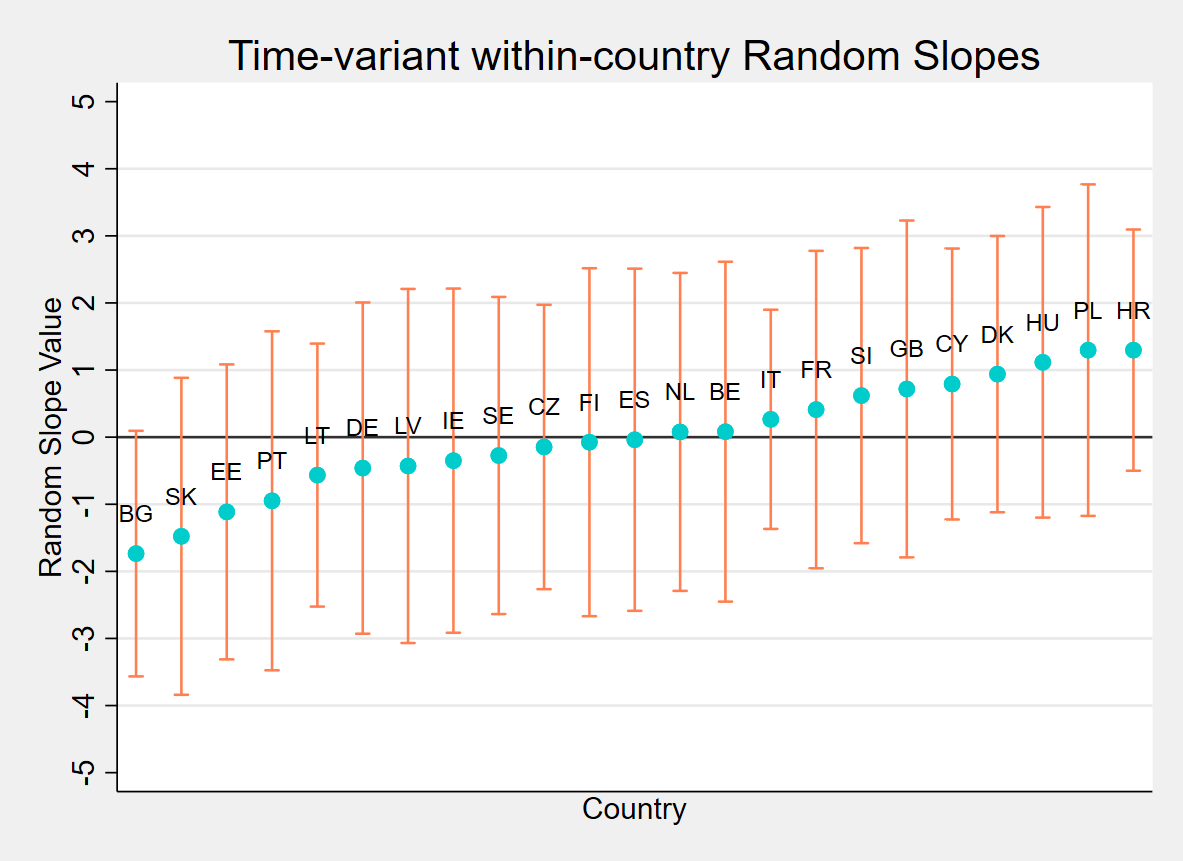
**22.** I do not explicitly carry out a test for model fit comparing the random intercept model and the empty model with no covariates, for two reasons. First, within the Stata environment likelihood-ratio tests cannot be applied to two specifications with different fixed-effects parts. Second, the choice to fit the random-intercept model is purely theory-driven, so taking decisions based on a likelihood-ratio test would not make any substantive sense.

**Results: Random Slopes Within-Between Framework [701 words]**

I fit a linear random-slopes model in line with the within-between formulation to addressthe unrealistic assumption that all individuals in the analytic subsample share the same coefficients for the independent variables predicting their degree of support towards further EU integration. In other words, two new terms are added to the specification for each random-effect predictor, to which one applies assumptions analogous to the random intercept’s: ζjc, the survey-specific random slope, and ζc, the country-specific random slope (Rabe-Hesketh & Skrondal, 2008, pp. 188–191). Since inferences for regression parameters depend on the established covariance structure, it is recommended to select the latter a priori, before turning to the mean structure (Rabe-Hesketh & Skrondal, 2008, p. 325). Thus, I decide to set an unstructured covariance matrix, as the cue-taking and benchmarking frameworks lead me to anticipate that each pair of observations within any cluster – be it Level-2, or Level-3 – exhibits a distinct correlation. National proxies change as the political context mutates over time – especially in the span of four, or even six years – and are country-dependent by definition (Hobolt & De Vries, 2016).



In my random-coefficients model specification, I incorporate random slopes at Level-3 for the main Level-2 independent variable – i.e., the time-variant, within-country effect of living in a left- or right-wing nation – and at Levels 2 and 3 for the participant’s (Level-1) endorsement forEuropeanunification. The former choice is theory-driven, as the existing literature shows how personal conceptions of left and right, on which the ESS measurements are grounded (ESS ERIC, 2021a), reflect the EU citizens’national and cultural backgrounds (Bauer et al., 2017). The latter is predominantly methodological, because multi-level models that do not allow for random slopes on the lower-level components of cross-level interactions often lead toseverely anti-conservative statistical inferences (Heisig & Schaeffer, 2019). Accordingly, I integrate the subject-level interaction effects of living in structurally, or increasingly right-wing countries, to assess **H3** and **H4**, respectively. On the other hand, I have no strongtheoretical expectation in favour or against allowing for supplementary random slopes for Level-1 and Level-2 control variables. Nevertheless, I separately test for their addition, following Hox et al.’s (2010) recommendations. In all instances but one – i.e., letting the parameter for the respondent’s gender vary across surveys– I experience serious convergence problems, guiding me towards the decision of not including any extra random coefficient**24**. I prioritise model stability and interpretability, focusing on the fixed effects and key random effects that align with the research questions and theoretical framework of my study.

Substantively meaningful results for the random-intercept within-between framework are reported in Table 7**25** (see p. 13). As I hypothesised, on average and ceteris paribus, citizens living in countries that are increasingly right-leaning show less support for further unification in the European Union, even when controlling for their individual-level left-right political positioning. The effect is significant, thus providing corroboration to **H2**. Contrariwise, although the time-invariant, between-country effect of residing in a structurally right-leaning country on the subject-level endorsement for EU integration is negative, it is not significant in the population**26**. Therefore, **H1** is not substantiated by empirical findings. Lastly, fixed-effects coefficients for both interactions imply that, in the analytic subsample, on average and ceteris paribus, the country-level effect of increasingly, or structurally, right-leaning national political contexts is stronger on those citizens with a pre-existing rightward political orientation. However, as neither parameter reaches statistical significance, there is limited evidence in support of **H3** and **H4**.

**Figure 4 –** *Caterpillar plot of time-variant, within-country random slopes*

Turning to random-effects variances, the variance of the Level-3 random slope for the time-invariant, within-country effect is remarkably high, pointing to the existence of substantial variation between countries. I decide against computing the coverage interval containing about 95% of the slopes in the population of countries to visualise its magnitude, because it would be untenable to assume that the 23 states comprised in the data set are a random sample of all the EU member nations**27**. Instead, I provide a country-by-country plot of the predictedrandom slopes (see Figure 4, p. 14). Although the significant Level-2 time-variant, within-country parameter exhibits an overall negative pattern across the 23 nations, it appears that such effect is markedly country-dependent, in line with the cue-taking and benchmarking frameworks (Hobolt & De Vries, 2016). However, given the large confidence intervals shown in Figure 4, it is prudent to exercise utmost caution when interpreting this finding. Further data collection at Level-3 is warranted to obtain more precise estimates**28**.

**23.** This heuristic provided by Elff et al. (2021) is an approximation of the Satterthwaite (1946) and Kenward-Roger (Kenward & Roger, 1997) Denominator-Degrees-of-Freedom (DDF) adjustment techniques. When dealing with simple multi-level data structures, the practitioner can ground the p-value calculations on the Student t’s two-tailed distribution with *m* – *l – 1* degrees of freedom, with *m* as the number of upper-level clusters, *l* equalling the number of contextual variables. In the statistical software Stata, this correction must be carried out by hand, as Elff et al. (2021) illustrate in the Online Appendix to their article.

**24.** Furthermore, the inclusion supplementary random effects at all levels leads to convergence issues when computing DFBETAs for detecting influential surveys and countries.

**25.** I only report the results for the intercept, interaction effects, Level-2, and Level-3 independent variables (see Tables 1, 2, and 3, pp. 3-5), and the respondent’s political positioning. Random-effects covariances, and fixed-effects coefficients and inferences for control variables, are available upon request. I do not compute ICCs for Levels 2 and 3. Their estimates are conditional on the random slopes’ values, thus uninformative and hard to interpret due to the mathematical properties of the independent variables. I do not explicitly evaluate model fit in the article because my theoretically substantiated random-slope and random-intercept formulations exhibit different fixed-effects parts (see Note 22, p. 12). However, to justify my procedural choices in the *Results: Robustness Checks* section (see Note 29, p. 18), I still compare the random-intercept and the random-slopes specifications, without including predictors for interaction effects (see Table A, Appendix, p. 21).

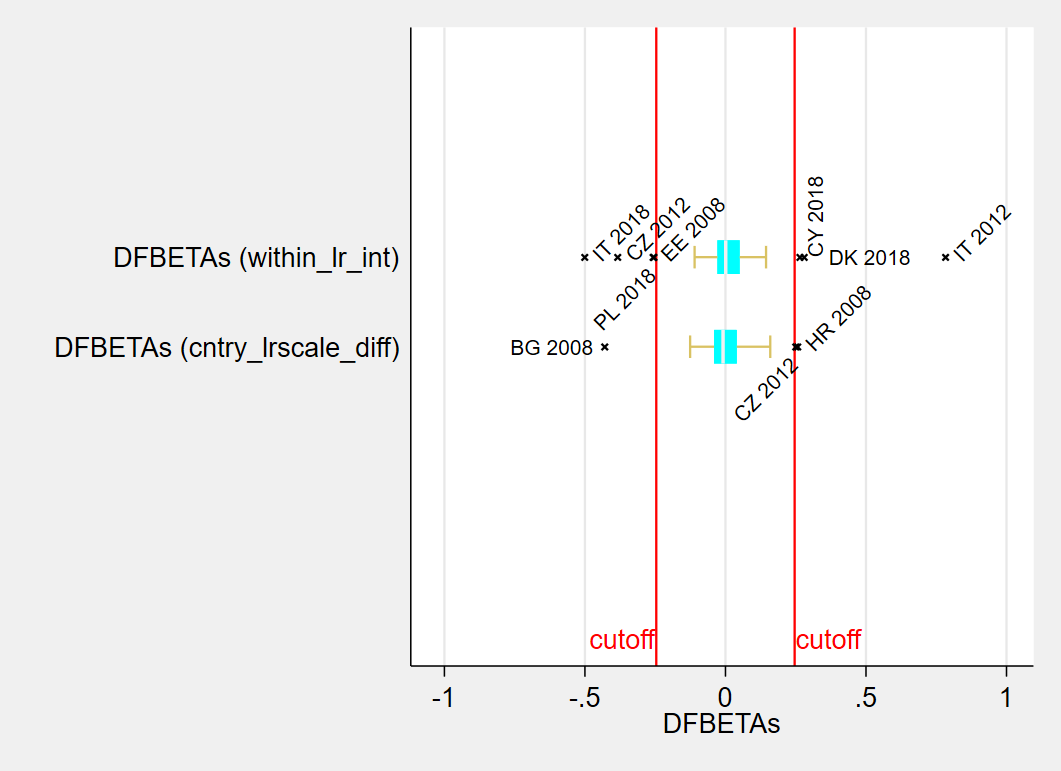
**26.** The utilization of robust p-value estimations does not alter the interpretation that would have been derived from employing standard inferential statistics.

**27.** Austria, Greece, Luxembourg, Malta, and Romania are not missing at random because I exclude them basing on data availability in the European Social Survey (see Note 1, p. 5).

**28.** It is important to note that, even in the best possible post-Brexit scenario, with all the EU member countries being surveyed, the Level-3 clusters would be 28.

**Results: Robustness Checks [496 words]**

I now turn to the execution of several robustness checks, to verify the results found within the random slopes framework**29**. I first validatewhether upholding the more restrictive sampling assumption of no missing data at Level-2 – i.e., surveys – leads to markedly different outcomes. The countries I exclude in this assessment are Croatia, Italy, and Latvia**30**, by setting a fixed-effect dummy variable at Level-3 to prevent loss of statistical power (Langford & Lewis, 1998; Van Der Meer et al., 2010). Results for this procedure are reported in Table B (see Appendix, p. 21). The findings obtained within the random-slopes formulation consistently withstand the more stringent assumption of no missing data at Level-2.



**Figure 5 –** *Level-2 DFBETAs of time-variant, within-country effects (see Tables 1 and 2, pp. 3-4)*

I subsequently search for clusters of influential upper-level units, by evaluating the DFBETAs diagnostics for the main Level-2 and Level-3 variables, and their interaction effects with Level-1 political positioning (see Tables 1, 2, and 3, pp. 3-5). A key issue is that Stata does not support this robustness check for three-level data structures. Furthermore, even when separating the hierarchy, it is incapable of calculating DFBETAs for country-specific, time-invariant observations – i.e., Level-3 – within the random-slopes framework, because their estimation is too computationally complex. I am thus forced to limit my assessment of Level-3 outliers to the random-intercept model**31**. Results for survey-level influential cases are shown in Figures 5 and 6 (see pp. 16-17). I identify a grand total of 15 influential surveys**32**, out of 66 in total, nine of which impact time-variant, within-country effects, and six of which regard time-invariant, between-country effects. This is a remarkably high number of outliers, which leads me to suspect that my models suffer from misspecification**33** (Van Der Meer et al., 2010). Nevertheless, I carry out separate and joint tests by including different fixed-effect dummy variables at Level-2 to prevent loss of statistical power (Langford & Lewis, 1998), reported in Tables C, D, and E (see Appendix, pp. 22-23). Most importantly, when factoring out the influence of all the Level-2 anomalies, the time-invariant, between-country effect becomes significant at the 5% level, putting the rejection of **H1** into serious question.

Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente

**Figure 6 –** Level-2 *DFBETAs of time-invariant, between-country effects (see Tables 1 and 3, pp. 3-5)*

Immagine che contiene testo, diagramma, linea, schermata

Descrizione generata automaticamente

**Figure 7 –** Level-3 *DFBETAs of time-variant, within-country effects (see Tables 1 and 2, pp. 3-4)*

DFBETAs of Level-3 outliers are reported in Figures 7 and 8 (see pp. 17-18). I identify a grand total of 5 influential countries**32**, out of 23 in total, two of which impact time-variant, within-country effects, and three of which regard time-invariant, between-country effects. This is a remarkably high number of outliers, which corroborates the suspect that my models suffer from misspecification**33** (Van Der Meer et al., 2010). Nevertheless, I carry out separate and joint tests by including different fixed-effect dummy variables at Level-3 to prevent loss of statistical power (Langford & Lewis, 1998), displayed in Tables F, G, and H (see Appendix, pp. 24-25). The findings obtained within the random-slopes formulation consistently withstand all robustness checks concerning Level-3 anomalies. However, considering the outcomes of the tests factoring out influential surveys, and the signals ofmodel misspecification, this brings me to the conclusion that **H1** cannot be fully disconfirmed within this research’s framework**34**.

Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente

**Figure 8 –** Level-3 *DFBETAs of time-invariant, between-country effects (see Tables 1 and 3, pp. 3-5)*

**29.** All robustness checks are run on the random-slopes, within-between formulation, the best model specification from a theory- and data-driven perspective (see Table A, Appendix, p. 18).

**30.** Italy is missing data at Level-2 for 2008, while Croatia, and Latvia are missing data at Level-2 for 2012.

**31.** Simple bivariate scatterplots would be remarkably complex to interpret because the main outcome variable – i.e., support for further EU integration – is not measured on a continuous scale. Consequently, I employ the *mltcooksd* package, which provides estimates of Cook’s distance and DFBETAs for the second-level units in hierarchical mixed models. Since it is built as a post-estimation command within the *xtmixed* framework, I cannot compute the Student’s t-test statistics. However, this does not constitute a concern for robustness checks, as I do not need to carry out statistical inferences to identify Level-2 and Level-3 outliers.

**32.** I consider an upper-level unit as influential when its corresponding DFBETAs statistic exceeds the cut-off value of 2√nx (Belsley, 2004).

**33.** This is particularly worrying because the random-effects within-between formulation is still prone to misspecification of unobserved confounders. In other words, omitting Level-1 confounders in the within-between formulation can still bias estimates of higher-level effects (Bell & Jones, 2015).

**34.** I choose to not carry out robustness checks for heteroscedasticity, for two reasons. Firstly, in the within-between framework, assuming that random intercepts are normally distributed, when that is not the case, introduces only modest biases (Bell et al., 2019). Secondly, and most importantly, Stata does not allow utilisation of the sandwich estimator when employing the REML method, which I deem to be more pertinent for conducting appropriate statistical inference (Elff et al., 2021).

**Conclusions and Limitations of the Research [451 words]**

In this paper, I focus on analysing the associations between the country-based, contextual proxies of the political left and right, and the European citizens’ support for further EU unification at the micro level. More specifically, I evaluate the autonomous impact of the national-level left-right leaning of anEU member state on the support for more integration in the European Union, testing the so-called cue-taking and benchmarking theories (Hobolt & De Vries, 2016). By employing a random-effects framework in its within-between formulation (Bell et al., 2019), I showrobust empirical evidence that corroborates **H2**. Contrariwise, there is no indication of the existence of a significant interaction between this time-variant, within-country effect, and the subject-specific political positioning, leading me to confidently reject **H4**. In sum, I can conclude that, on average and ceteris paribus, citizens living in nations that are increasingly right-leaning exhibit less endorsement towards further integration in the European Union, independently of their political beliefs. Turning to the hypotheses regarding structural effects, the overall picture is insufficiently clear to reach a conclusiveverdict. My findings suggesta negative, but non-significant time-invariant, between-country effect, and a positive, and analogously non-significant interaction effect. Yet, results for the main between-effect of interest are not robust when factoring out all influential Level-2 units – i.e., surveys. Moreover, the number of Level-2 and Level-3 outliers is remarkably high, an indicator of potential model misspecification (Van Der Meer et al., 2010). Thus, I can unambiguouslydismiss **H3**, ascertaining that, on average and ceteris paribus, the country-level effect of structurally right-leaning national political contexts is not stronger on those citizens with a pre-existing rightward political orientation. However, I cannot decisively disconfirm **H1** – i.e., the presence of a significant structural effect – within this research’s framework.

This conclusion may seem rather unsurprising, as this study suffersfrom several shortcomings. First, the literature suggests that measuring an individual’s political ideology with a left-to-right scale may be inappropriate, as there is limited interpersonal comparability in how questionnaire respondents interpret the abstract concepts of left and right (Bauer et al., 2017). Second, lack of computational power and up-to-date tools within the statistical software Stata greatly restrictsthe scope and reliability of my robustness checks, meaning that I cannot take random-slopes into account when searching for Level-3 outliers. Third, there are serious reasons to suspect that my models are mis-specified, which is particularly concerning since omitting Level-1 confounders in the within-between formulation can still bias estimates of higher-level effects (Bell & Jones, 2015). However, introducing supplementary control variables would likely yield additional convergence issues. Future researchin this substantive field should consider a wider breadth of confounders, collect data for a larger number of time points and countries, and employ state-of-the-art hardware for statistical modelling.

**Appendix**















**References**

Bauer, P. C., Barberá, P., Ackermann, K., & Venetz, A. (2017). Is the Left-Right Scale a Valid Measure of Ideology?: Individual-Level Variation in Associations with “Left” and “Right” and Left-Right Self-Placement. *Political Behavior*, *39*(3), 553–583. https://doi.org/10.1007/s11109-016-9368-2

Baute, S. (2023). Mass Euroscepticism revisited: The role of distributive justice. *European Union Politics*, *0*(0). https://doi.org/10.1177/14651165231170789

Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, *53*(2), 1051–1074. https://doi.org/10.1007/s11135-018-0802-x

Bell, A., & Jones, K. (2015). Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, *3*(1), 133–153. https://doi.org/10.1017/psrm.2014.7

Belsley, D. A. (2004). *Regression diagnostics: Identifying influential data and sources of collinearity*. Wiley.

Elff, M., Heisig, J. P., Schaeffer, M., & Shikano, S. (2021). Multilevel Analysis with Few Clusters: Improving Likelihood-Based Methods to Provide Unbiased Estimates and Accurate Inference. *British Journal of Political Science*, *51*(1), 412–426. https://doi.org/10.1017/S0007123419000097

ESS ERIC. (2018a). *ESS4—Integrated file, edition 4.5 [Data set]*. Sikt - Norwegian Agency for Shared Services in Education and Research. https://doi.org/10.21338/ESS4E04\_5

ESS ERIC. (2018b). *ESS6—Integrated file, edition 2.4 [Data set]*. Sikt - Norwegian Agency for Shared Services in Education and Research. https://doi.org/10.21338/ESS6E02\_4

ESS ERIC. (2021a). *ESS9 Data Documentation*. Sikt - Norwegian Agency for Shared Services in Education and Research. https://doi.org/10.21338/NSD-ESS9-2018

ESS ERIC. (2021b). *ESS9—Integrated file, edition 3.1 [Data set]*. Sikt - Norwegian Agency for Shared Services in Education and Research. https://doi.org/10.21338/ESS9E03\_1

Heisig, J. P., & Schaeffer, M. (2019). Why You Should *Always* Include a Random Slope for the Lower-Level Variable Involved in a Cross-Level Interaction. *European Sociological Review*, *35*(2), 258–279. https://doi.org/10.1093/esr/jcy053

Hobolt, S. B., & De Vries, C. E. (2016). Public Support for European Integration. *Annual Review of Political Science*, *19*(1), 413–432. https://doi.org/10.1146/annurev-polisci-042214-044157

Hox, J. J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, *64*(2), 157–170. https://doi.org/10.1111/j.1467-9574.2009.00445.x

Kenward, M. G., & Roger, J. H. (1997). Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*, *53*(3), 983. https://doi.org/10.2307/2533558

Langford, I. H., & Lewis, T. (1998). Outliers in Multilevel Data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *161*(2), 121–160. https://doi.org/10.1111/1467-985X.00094

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, *46*(1), 69. https://doi.org/10.2307/1913646

Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, *58*(3), 545–554. https://doi.org/10.1093/biomet/58.3.545

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata*. Stata Press Publication.

Satterthwaite, F. E. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, *2*(6), 110. https://doi.org/10.2307/3002019

Schurer, S., & Yong, J. (2012). *Personality, Well-being and the Marginal Utility of Income: What Can We Learn from Random Coefficient Models?* [Working Paper]. Victoria University of Wellington. http://hdl.handle.net/10063/2040

Stegmueller, D. (2013). How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches: How Many Countries? *American Journal of Political Science*, *57*(3), 748–761. https://doi.org/10.1111/ajps.12001

The World Bank. (2021). *World Bank Open Data*. https://data.worldbank.org/

Transparency International. (2008). *Corruption Perceptions Index 2008*. Transparency.org. https://www.transparency.org/en/cpi/2008

Transparency International. (2012). *Corruption Perceptions Index 2012*. Transparency.org. https://www.transparency.org/en/cpi/2012

Transparency International. (2018). *Corruption Perceptions Index 2018*. Transparency.org. https://www.transparency.org/en/cpi/2018

Van Der Meer, T., Te Grotenhuis, M., & Pelzer, B. (2010). Influential Cases in Multilevel Modeling: A Methodological Comment. *American Sociological Review*, *75*(1), 173–178. https://doi.org/10.1177/0003122409359166

Van Elsas, E. J., Hakhverdian, A., & Van Der Brug, W. (2016). United against a common foe? The nature and origins of Euroscepticism among left-wing and right-wing citizens. *West European Politics*, *39*(6), 1181–1204. https://doi.org/10.1080/01402382.2016.1175244

Van Elsas, E. J., & Van Der Brug, W. (2015). The changing relationship between left–right ideology and euroscepticism, 1973–2010. *European Union Politics*, *16*(2), 194–215. https://doi.org/10.1177/1465116514562918

Vasilopoulou, S. (2018). The Radical Right and Euroskepticism. In J. Rydgren (Ed.), *The Oxford handbook of the radical right* (pp. 122–140). Oxford University Press.