Appendix to Chapter 2

Appendix to Chapter 4

Appendix to Chapter 5

Appendix to Chapter 6

APPENDIX TO CHAPTER 2

Studies were included in the review if they met the eligibility criteria presented in Table A2.1 (Liberati et al., 2009). Our literature search was executed using the electronic database of the Web of Science Core Collection. We chose this database as it is the most comprehensive social scientific database while adhering to high standards of academic quality: it includes only peer-reviewed work and its journals are thoroughly reviewed before inclusion. We do not include books, because research covered in books is often also published in journals and their inclusion requires the use of databases with lesser quality controls.

Table A2.1. Eligibility criteria

No.	Subject	Criterion
1.	Topic	Studies must focus on the different elements, such as the actors, institutions and policies, and processes, such as processes of inter-institutional bargaining and legitimisation, of the political system underpinning EU financial regulation or economic governance.
2.	Design	Studies must have a positive (non-normative) goal and orientation.
3.	Publication status	Studies must be published in peer-reviewed journals.
4.	Language	Studies must be published in English.
5.	Year of publication	Studies must be published between 1999 and 2016.

Two separate queries were applied to the Core Collection database using its Boolean operators; the query on financial regulation used the search terms [financial regulation] and [European Union], and the query on economic governance replaced the first field by [economic governance]. The use of Boolean operators ensures terms did not have to occur adjacent to each other for studies to end up in the search results. For example, studies of which the abstract only included the terms 'European Union' and 'economic' were also retrieved and screened (e.g. Borrás & Radaelli, 2015). In a similar vein, search terms regarding politics were not applied during any search strategies because these would have been too restrictive. Both queries were temporally restricted to 1999-2016 and were further refined to yield only academic articles. These queries were last executed in January 2017 and generated a total of 246 and 572 hits, respectively.

One review author then applied the inclusion criteria to the abstracts of these studies. The query on financial regulation yielded 63 studies eligible for review. The query on economic governance yielded 106 new eligible studies and 11 duplicates also covered by the first query. A total of 638 studies were excluded, because they did not involve financial regulation or economic

governance as defined above (520)[34], did not involve the underlying political system (69)[35] or the EU (36), or because they were not published in English (13).

The final step in the selection procedure involved screening the full texts to exclude studies with a non-empirical orientation, being articles with normative goals or literature reviews. While both types of publications draw on empirical data and provide valuable insights, neither were included: the former makes selective use of empirical data to develop a normative argument, whereas the latter draws studies into the review that may not be eligible or that are already covered in the review. This step excluded 32 normative articles and an additional seven literature reviews, resulting in a final population of 138 eligible studies across both policy areas. The studies were published in 44 different journals; the major outlets being the Journal of European Public Policy (27), Journal of Common Market Studies (22) and Journal of European Integration (10). The full selection process is displayed in Figure A2.1.

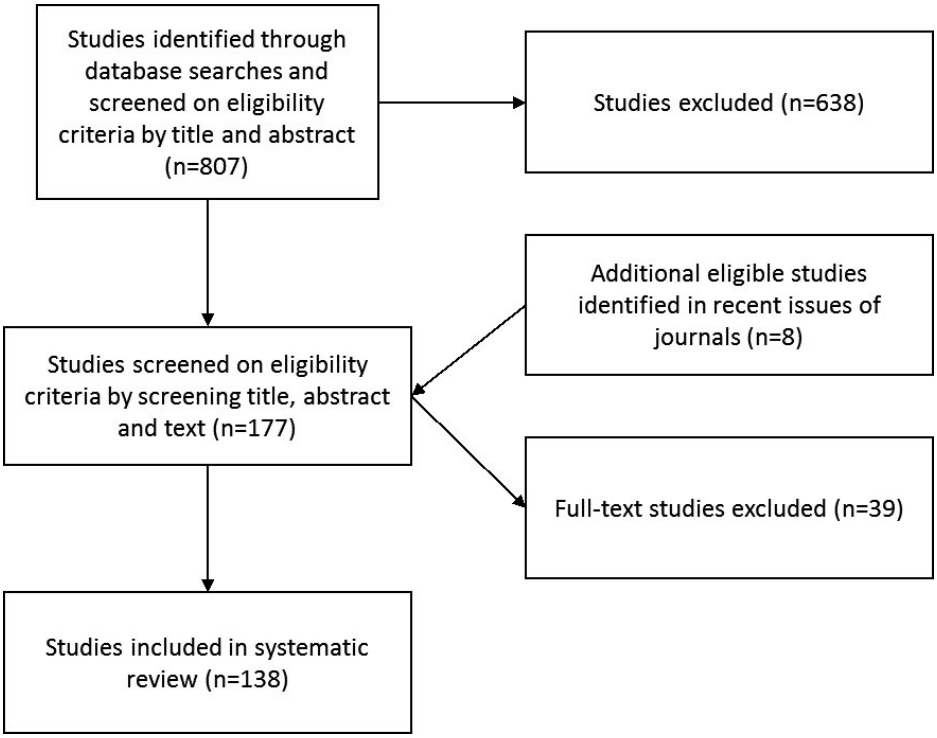


Figure A2.1. PRISMA flowchart: process for identifying and retaining studies

34 The topics of these excluded studies varied widely, but dominant issues in this dropped literature were environmental policy, employment policy or other policy related to specific sectors of the economy, regional development (funds) and international accounting standards.

35 Examples of such studies were studies that reviewed the effectiveness of policy or studies that presented a purely economic or financial analysis of the crisis.

Upon inclusion, the studies were coded. These codes provided metadata, as well as data on a study's design, use of theory and units of analysis. For each study, we also extracted its central statement of intent (its goal) and its corresponding findings. Coding inevitably is a subjective process and we took appropriate steps to limit the associated bias. One author first coded 80 studies while another author reviewer coded 20 articles that were deemed difficult to code. This pilot was then used to revise the codebook. In a second round, one reviewer coded all 138 studies while the other recoded a subsample of 30 randomly selected studies. Discrepancies in coding were resolved by discussing the coding of the study in question and revising the coding if necessary. The codebook, as well as a quantitative scoping exercise, can be found below.

QUANTITATIVE SCOPING

Academic affiliation and territorial scope of study

The majority of studies are conducted at research institutions located in the North-West of the EU (63,8%); 78,6% in financial regulation and 58,1% in economic governance. Universities in the United Kingdom are responsible for 40,5% of studies on financial regulation and 29,1% on economic governance, strongly outperforming the EU's other large member states: these figures are 11,9% and 14,0% for Germany (North-West) and 4,8% and 5,8% for France (South) respectively. Whereas the UK's interest in financial regulation is largely explained by the size of its financial sector, its dominance in economic governance is remarkable given the UK's opt-outs to the ERM II (the first stage of joining EMU) and to the corrective arms of the EU's macroeconomic and fiscal surveillance frameworks. This finding thus seems to reflect the general dominance of UK-based academics (who are not necessarily UK nationals) in English journals.

The contribution from research institutions in Southern member states is much lower: 19,8% for economic governance and 9,5% for financial regulation. Since 2011, the relative share of contributions from Southern Europe across both areas has risen from 7,7% to 17,9%. If one assumes that national context has some degree of influence on the type of research conducted, we find this increasing diversity to be a positive development; especially since the crisis reinforced the cleavage between publics and governments of Northern (creditor) and Southern (debtor) member states (Baglioni & Hurrelmann, 2016; Dehousse, 2016; Schimmelfennig, 2016; Seikel, 2016). The EU's east is rarely featured (2,2% for both areas combined). Three countries constitute the source of the 15,9% of studies conducted outside the EU: the United States (13,0%), Canada, and Switzerland (1,5% each).

In turn, 71,7% of the studies examining the politics of financial regulation or economic governance focus on the EU in its entirety or its institutions, as opposed to a focus on single member states (10,8%), or comparing member states (12,3%) or the EU and US (5,1%). Those studies that focus on specific regions within the EU focus predominantly on its North-West (10,9%). The EU's south, as well as non-EU-wide cross-region comparisons, both account for 5,1% of the literature.

STUDY DESIGN AND ROLE OF THEORY

With 82,6% share of total studies, the case study is the dominant research design used to study the politics of financial regulation or economic governance. This corroborates Franchino's (2005) findings on the prevalence of case studies in European integration research, and can be explained by the case study's relative advantages in studying macro-level phenomena (where there are usually insufficient observations for large-N designs) and in taking a process-oriented approach (Haverland & Van der Veer 2017). Medium- or large-N research (14,5%) are by no means less relevant for this area of research; these studies often examine the attitudes of the public or governing elites (e.g. Banducci et al., 2003; Wonka & Rittberger, 2011) or interest group mobilisation (e.g. Chalmers, 2015).

Given the variety of designs employed, we did not assess the methodological validity of these studies. Instead, we assessed whether studies explicitly reported on the way they collected and/or analysed empirical data: merely 7,9% of case studies did, against 80,0% of large-N studies. In total, 79,0% of studies did not explicitly report on their methods, making any attempts at judging the validity of their findings a near-impossible endeavour. In the period since 2011, this average share of studies lacking explicit methodological reporting increased by 12 percentage points compared to the period 1999-2010.

We distinguish between four different roles played by theory in these studies (Lijphart 1971; Eckstein 1975; Haverland & Van der Veer 2017). *Descriptive ideographic* studies constitute 18,8% of the reviewed studies. These studies offer rich empirical description but make no use of theories or theoretical concepts (e.g. Levitt, 2012; van Middelaar, 2016). Without such theoretical abstraction, their language does not travel beyond the case(s) investigated by the study. A key identifier for descriptive ideographic studies was the lack of a paragraph or section explicating the theories or concepts employed.

The largest share of studies is *theory-informed descriptive* (41,3%). These studies have an explicit theoretical grounding and are written up using more abstract theoretical concepts. While due to abstraction their findings travel beyond the case(s) examined in the study, they offer no critical reflection on the theory used. A typical example of theory-informed descriptive studies are those that use empirical data to illustrate the relevance of a theory (e.g. Heinisch, 2000; Orenstein, 2008).

Only 1,4% of studies are classified as *hypothesis generating*. These heuristic studies explicitly build on empirical data to develop theory (e.g. Borrás & Radaelli, 2011). *Theory confirming/infirming* studies make up the fourth category, which constitutes 38,4% of studies reviewed. Whereas in theory-informed descriptive studies empirics do not feed back onto the theory, in these studies this relationship works both ways: theory is used to explain empirics and empirics are used to reflect on the applicability of a theory (Quaglia, 2015a; Verdun, 2013). Given the strong reliance on hypotheses testing in large-N research, it is unsurprising that 85,0% of this subset of studies explicitly reflect on the applicability of theory in their conclusions. For case studies, this percentage was only 28,1%.

Since the crisis, the relative share of theory-informed descriptive studies has increased substantially, from 26,9% before 2011 to 44,6% since. The share of theory confirming/infirming studies rose from 29,0% between 1999-2013 to 46,1% between 2014-2016. Descriptive ideographic studies have remained constant in absolute numbers and have therefore decreased relative to other designs. This implies theory is increasingly important in research on the politics of financial regulation and economic governance, and the field is increasingly feeding its findings into the broader theoretical debates it draws on. We interpret this as a positive sign of a maturing of the literature–largely catalysed by the economic crisis.

The PRIMSA-checklist also requires the analyst to make an assessment of the methodological rigor of the reviewed studies. However, in many cases, we have found it impossible to assess the methodological quality of eligible studies, as they stem from vastly different research traditions that rely on different ontologies, methodologies and common practices of reporting on these. For example, a strict review based on methodological transparency would instil a bias in favour of positive comparative politics studies, and against studies stemming from a critical political economy tradition. For this reason, we have only looked at the share of studies which reports on the methodology used: 21% of eligible studies clearly reported on the methodology used, against 79% which did not.

PERSPECTIVES ON EUROPEAN INTEGRATION

Research on EU politics is often informed by theories of EU integration. We distinguish between five different perspectives[36], whereas the first three are most explicitly mid-range theories. *Neo-functionalism,* with its emphasis on functional, political and cultivated spill over, was a core perspective in only 1,4% of studies reviewed (e.g. Niemann & Ioannou, 2015). We encountered *multi-level governance*, which emphasises the interdependencies and shared authority between the subnational, national and supranational levels, in 6,5% of studies (e.g. Meyer, 2005). *Intergovernmentalism,* which privileges the role of EU member states and their national interests in explaining European integration, was used by 13,0% of studies (e.g. Bressanelli & Chelotti, 2016). Moving towards more generic theories, *Institutionalism,* which views institutions (and their different components: as rules of the political game constraining the realisation of actor preferences, as shared norms that shape decisions, and as processes of path dependency) at the EU level as the dominant factors shaping the integration process, was central to 34,0% of studies (e.g. Featherstone, 2005). We did not encounter studies from a *post-functionalist* perspective, which

36 The perspective on European integration we identify in a given study is not necessarily identical to the theory it uses. For example: Spendzharova (2014) develops a theory to predict member state regulatory preferences using their levels of foreign bank ownership and domestic bank internationalisation. Despite its firm nesting in the intergovernmentalist tradition, her theory is smaller in scope.

sees the ongoing politicisation of the EU and the rise of identity politics as a constraint on further integration. In 8,0% of the studies reviewed multiple theories were purposely pitted against each other to assess their relative explanatory capacity (Jones et al., 2016). For 37,0% of studies a clear presence of one of the aforementioned perspectives was not identifiable (Torres, 2013).

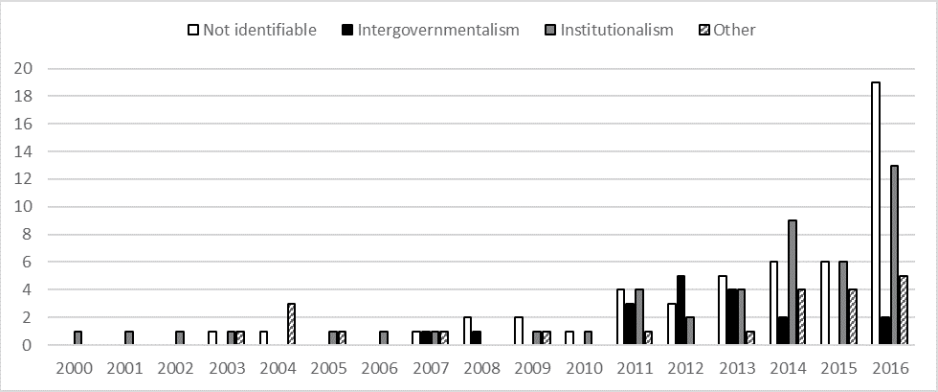


Figure A2.2. Perspectives on European integration over time

Figure A2.2 plots the use of these perspectives over time. There are no visible patterns in the use of perspectives before the crisis period. Then, we witness a short surge in intergovernmentalist studies between 2011 and 2014. This pattern is linked to the large battery of legislative reforms negotiated by EU member states in the early years of the crisis–such as the European Financial Stability Facility (Begg, 2012), Capital Requirements Directive (Howarth & Quaglia, 2013) and Banking Union (Spendzharova, 2014) – and the intergovernmental nature of the Euro area bailouts (e.g. Zahariadis, 2012).

Once dust begins to settle on overhauled architectures of EU economic governance and financial regulation, scholars begin explaining the institutional change and its consequences; the decrease in intergovernmentalist studies after 2014 coincides with a steady rise in the number of institutionalist studies since 2012. These studies are concerned with issues such as the impact of crisis reform on Commission-Council relations (Bocquillon & Dobbels, 2014), the legitimacy of the EU's increased authority in the post-crisis period (White, 2015) and the development of Germany's institutional power throughout the crisis (Steinberg & Vermeiren, 2016).

However, the most notable increases are in studies which do not clearly advocate dominant perspectives on European integration. A partial explanation for this strong increase is that the bulk of this work covers the more routine workings of the system, and thus does not cover integration per se.

APPENDIX TO CHAPTER 4

Table A4.1. Descriptives: MIP scoreboard indicators - Breach variables

Indicator	Indicative threshold	Min.	Max.	Mean	SD	NA
3-year backward moving average of the current account balance as percent of GDP	-4% and +6%	0	1	0.218	-	-
net international investment position as percent of GDP	-35%	0	1	0.5	-	-
5-year percentage change of export market shares measured in values	-6%	0	1	0.483	-	-
3-year percentage change in nominal unit labour cost	+9% (EA) and +12% (non-EA)	0	1	0.151	-	2
3-year percentage change of the real effective exchange rates based on HICP/CPI deflators	-/+5% (EA) and -/+11% (non-EA)	0	1	0.121	-	-
private sector debt (consolidated) in % of GDP	133%	0	1	0.474	-	18
private sector credit flow in % of GDP	14%	0	1	0.038	-	18
year-on-year changes in house prices relative to a Eurostat consumption deflator	6%	0	1	0.138	-	-
general government sector debt in % of GDP	60%	0	1	0.517	-	-
3-year backward moving average of unemployment rate	10%	0	1	0.333	-	-
year-on-year changes in total financial sector liabilities	16.5%	0	1	0.026	-	18
3-year change in p.p. of the activity rate	-0.2%	0	1	0.138	-	-
3-year change in p.p. of the long-term unemployment rate	+0.5%	0	1	0.517	-	-
3-year change in p.p. of the youth unemployment rate	+2%	0	1	0.454	-	-

Notes: N = 174, J = 27, K = 7.

Table A4.2. Descriptives: Other variables

	Min.	Max.	Mean	SD	NA
CSRs: Number of words	9	1139	288.724	200.992	-
CSRs: Share of social investment in average recommendation	0.003	0.474	0.216	0.105	-
Political power	0.007	0.136	0.038	0.035	4
EMU	0	1	0.632	-	-
Kurtosis (Polarisation)	-0.871	0.907	-0.12	0.398	-
Euroscepticism Nat. Parliament	0	1	0.787	-	-
Election	0	59	23.69	15.211	-
Govt. pos. anti-pro EU	0	7	5.509	1.314	-
Govt. pos. left-right	1.05	8.68	5.611	1.629	-

Notes: N = 174, J = 27, K = 7.

DISTRIBUTION AND TRANSFORMATION OF DEPENDENT VARIABLE (NUMBER OF WORDS)

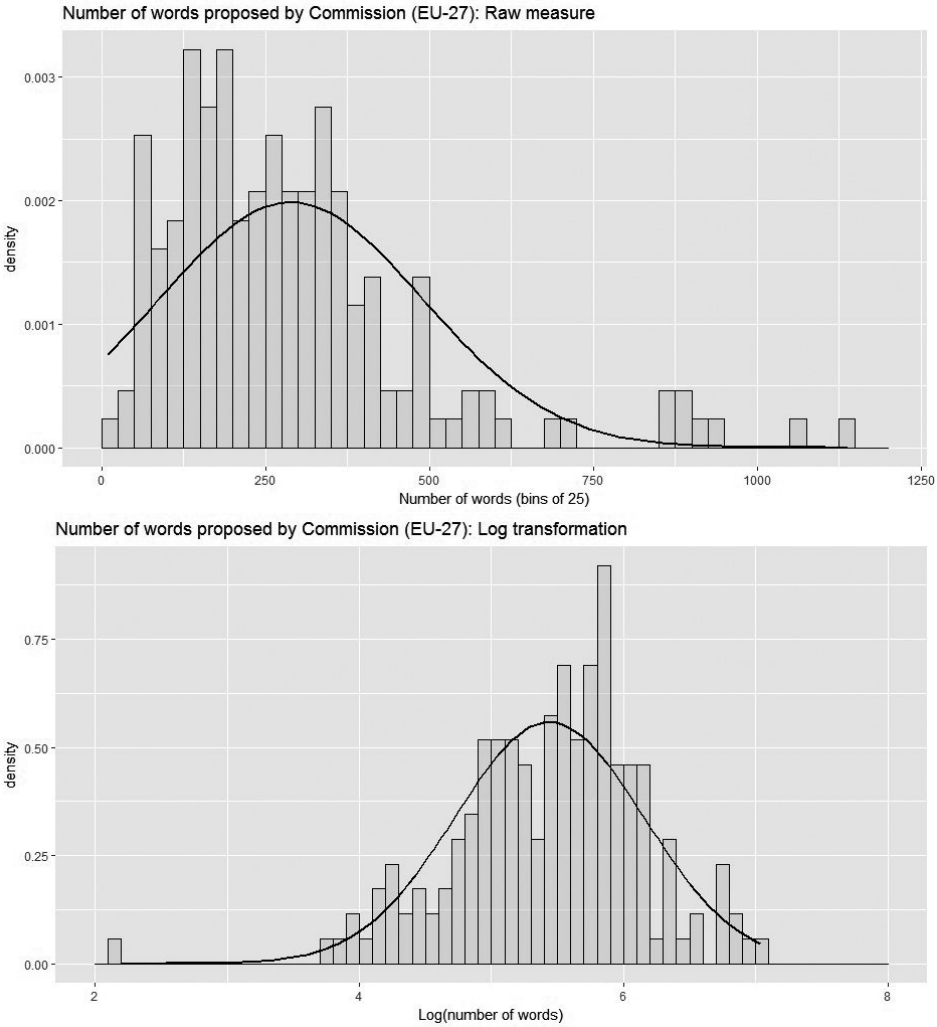


Figure A4.1. Density plots of outcome variables before and after log transformation

LATENT DIRICHLET ALLOCATION (LDA)

The outcome variable for models B reflects the proportion of topics explicitly related to social investment in the average recommendation per country-year. This measure is bounded between 0 and 1, with 0 indicating the average share of recommendations in a given country-year recommending social investment being 0%, and 1 indicating this share to be 100%. Thus, a country

receiving three recommendations in a given year, in which the shares of covered topics related to social investment are 20%, 60% and 40%, receives a value of 0.4 on this measure.

We use the statistical software R and its text mining features (most notably R's tm, Quanteda, topicmodels and ldatuning packages) to run the LDA model used to estimate the topics covered in the recommendations. First, we extract the recommendations from the Commission documents, being all text in the enumeration between "HEREBY RECOMMENDS […] to:" and "Done at Brussels, […]". We split these texts by recommendation (N = 824) and clean the texts using conventional pre-processing steps, including the removal of frequently (> 60%) and infrequently (< 1%) occurring terms. LDA models are fully unsupervised, but require the researcher to ex-ante set the number of topics the model should converge upon. To estimate the optimal number of topics, we use the ldatuning package (https://cran.r-project.org/web/packages/ldatuning/ldatuning.pdf) to estimate the statistically optimal number of topics present in the total corpus of recommendations. The ldatuning package calculates four different metrics. Griffiths2004 optimises the log-likelihood of replicating the terms as found in the corpus generatively under varying values for topic number *k*. This metric uses Markov Chain Monte Carlo simulations to iterate over a range of possible values for *k*, simulating a log-likelihood of finding the terms per document given *k* and optimizing this log-likelihood over a range of possible values for *k*. CaoJuan2009 minimises topic density (cosine distance), which is a measure of the multiway correlation between topics under varying values of *k*. Lower levels of density indicate less overlap between topics and a more stable topic structure. Arun2010 minimises Kullback-Leibler divergence, which is a measure of how much information is lost by changing the parameterisation of the topic model (altering *k*). This metric thus aims to find *k* such that in moving from the data to a summary of the data (i.e. a parameterised topic model based on *k* topics), minimal information is lost. The Deveaud2014 is an adapted measure based on Kullback-Leibler divergence, which also seeks to optimise divergence between topics over a range of values of *k*. Based on these metrics, we choose 30 topics. We then fit the LDA model.

We use three steps to qualitatively validate the fitted model. First, we look at the highest scoring terms per category and assess whether grouped terms are consistent (convergent validity) and whether topics sufficiently differ from each other (discriminant validity) (see Table A4.3). Second, we group the full-text recommendations by the topics they score highest on and compare these recommendations to assess whether their content is sufficiently congruent (see Table A4.4). We then name these topics and identify a topic as related to social investment if all full-text recommendations scoring highest on this topic advocate social investment (see Table A4.3). For example, topics 25 and 27 cover recommendations advocating improvements to the accessibility for disadvantaged groups to a member state's healthcare and pension systems, respectively. While these are clearly oriented towards social investment, these topics also cover recommendations advocating measures to make these systems more financially sustainable. As such, neither topic is solely related to social investment.

In the final step, we cross-validate our measure with the Commission's own classification of topics covered in the CSRs by member state, which are available for 2016 and 2017 (2016: http://ec.europa.eu/europe2020/pdf/csr2016/csr2016-overview-table_en.pdf and 2017: https://ec.europa.eu/info/files/2017-european-semester-policy-areas-covered-csrs_en). There is a substantial association between our measure and the measure based on the Commission's own topic classifications (r = 0.65), which given our conservative definition of 'social investment' strongly cross-validates our measure.

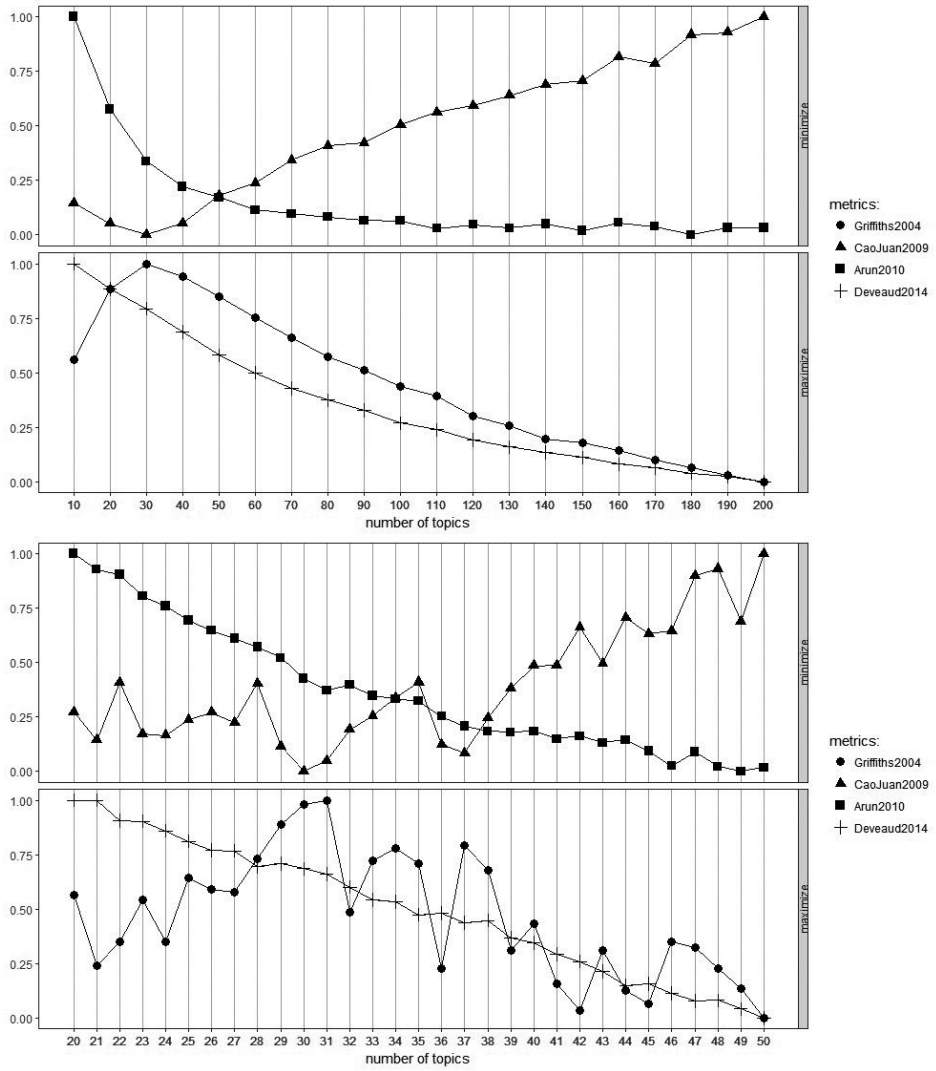


Figure A4.2. LDA-tuning results – Converging from 200 to 30 topics

CROSS-CLASSIFIED MULTILEVEL REGRESSION DIAGNOSTICS

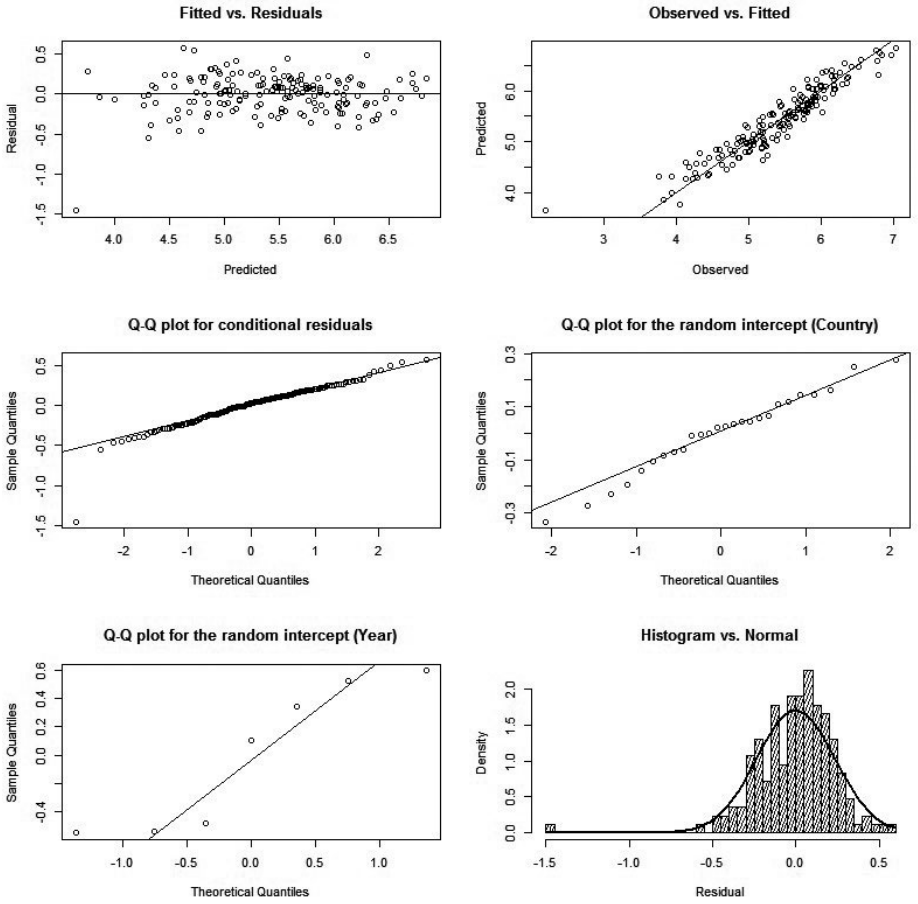


Figure A4.3. Regression diagnostics (Model A3)

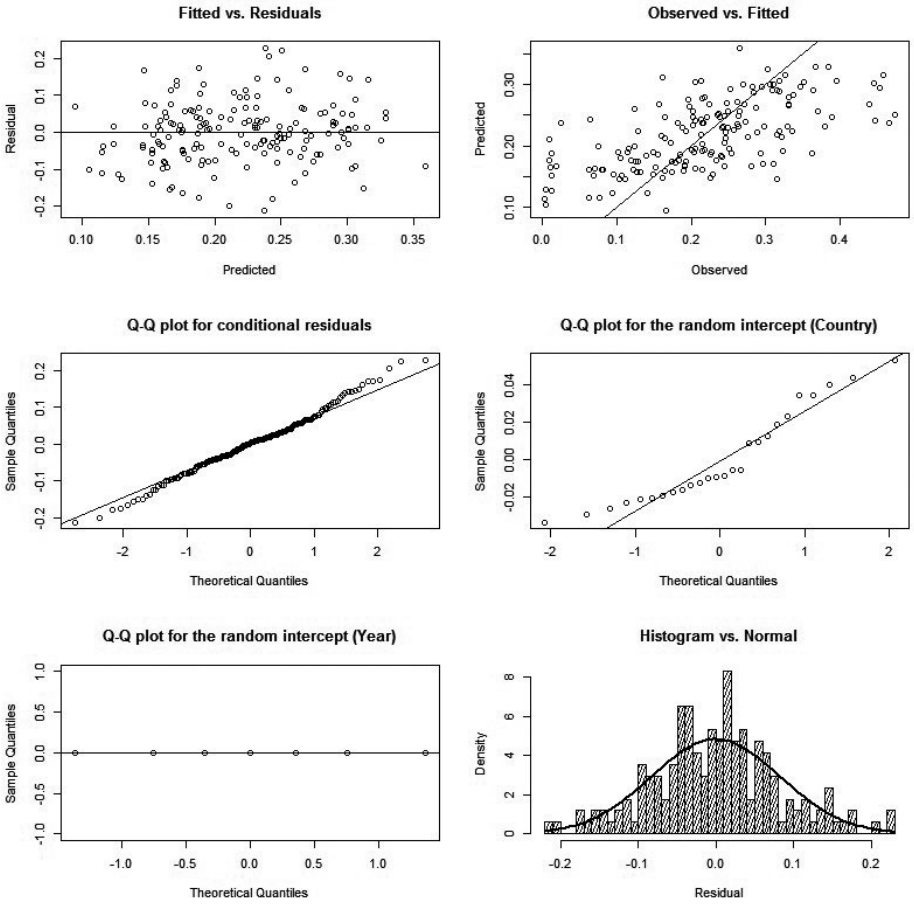


Figure A4.4. Regression diagnostics (Model B3)

Table A4.3. LDA – Topics and first 20 terms per topic by weight

	Topic	N	Social investment	First 20 terms
1	Public sector reform - Judiciary	17	No	insolv, case, restructur, legal, proceed, effici, procedur, includ, court, length, improv, system, framework, reduc, judici, corpor, process, particular, commerci, administr
2	Public sector reform - General	10	No	reform, implement, adopt, ensur, plan, speed, acceler, financ, step, complet, access, particular, act, legisl, promot, institut, rule, enhanc, assess, effect
3	Employability - Disadvantaged groups	20	Yes	particip, increas, children, strengthen, improv, childcar, qualiti, labour, market, educ, lifelong, afford, group, particular, care, avail, facil, women, mainstream, disadvantag
4	Fiscal policy - Ensure fiscal adjustment	22	No	budgetari, fiscal, term, medium, strengthen, rule, framework, ensur, growth, bind, requir, reinforc, strategi, govern, object, expenditur, signific, debt, plan, thereaft
5	Employability - Young and low-skilled workers	44	Yes	unemploy, labour, improv, activ, market, skill, work, social, employ, polici, benefit, address, youth, coverag, increas, system, long, term, peopl, low
6	Fiscal policy – Limit or avoid deviation MTO	20	No	term, medium, object, pension, budgetari, long, sustain, system, deviat, allow, ensur, fiscal, reform, financ, link, account, limit, take, achiev, financi
7	Structural policy - Energy sector reform	37	No	energi, transport, electr, network, improv, gas, sector, market, effici, regul, competit, capac, interconnect, ensur, strengthen, border, cross, particular, independ, cost
8	Innovation and R&D	23	No	invest, innov, research, public, privat, improv, busi, increas, infrastructur, foster, develop, cooper, polici, ensur, support, prioritis, financ, sector, regul, includ
9	Fiscal policy – Limit and/or avoid deviation MTO II	17	No	fiscal, polici, growth, line, stabil, requir, pursu, prevent, ongo, pact, arm, ensur, public, effort, sustain, substanti, translat, strengthen, take, achiev
10	Public sector reform – Quality and cost-efficiency	44	No	public, improv, administr, procur, effici, servic, implement, strengthen, corrupt, manag, effect, qualiti, step, transpar, busi, reduc, capac, fight, increas, procedur
11	Labour market - Linking supply to demand	16	No	servic, employ, public, effect, train, ensur, job, market, labour, polici, support, increas, capac, strengthen, activ, unemploy, implement, includ, qualiti, assist
12	Fiscal policy – Pursue growth-friendly strategy	23	No	fiscal, implement, term, budgetari, consolid, ensur, continu, posit, sound, strategi, object, growth, preserv, expenditur, medium, friend, public, complianc, revenu, envisag
13	Structural policy - Local government	15	No	level, govern, local, region, across, spend, administr, coordin, public, ensur, includ, central, effici, consist, increas, transpar, system, account, adopt, plan
14	Labour market - Social security	20	Yes	social, employ, servic, nation, effect, minimum, strengthen, market, labour, activ, scheme, integr, establish, particular, improv, polici, link, incom, youth, coordin
15	Labour market - Supply side reform	36	No	wage, social, partner, consult, develop, ensur, competit, product, accord, practic, nation, labour, reform, system, set, index, minimum, employ, creation, condit

Table A4.3. LDA – Topics and first 20 terms per topic by weight (continued)

	Topic	N	Social investment	First 20 terms
16	Fiscal policy – Ensure fiscal adjustment II	33	No	gdp, deficit, fiscal, adjust, toward, debt, correct, achiev, excess, structur, medium, object, term, budgetari, expenditur, use, ensur, annual, govern, ratio
17	Financial sector reform - Housing market	23	No	hous, market, household, tax, mortgag, increas, includ, properti, reform, rent, suppli, debt, system, price, deduct, interest, reduc, rental, plan, taxat
18	Market regulation - Effectiveness and efficiency	10	No	law, end, adopt, implement, ensur, provid, nation, market, framework, support, enterpris, incent, establish, uniti, author, includ, review, legisl, regulatori, action
19	Financial sector reform - General	45	No	bank, sector, non, financi, loan, improv, perform, financ, restructur, capit, asset, ensur, credit, busi, access, address, includ, supervis, smes, particular
20	Public sector reform - State-owned enterprises	15	No	state, enterpris, own, manag, asset, complet, govern, implement, ensur, corpor, strategi, review, compani, ownership, privatis, clear, commerci, function, divest, plan
21	Fiscal policy – Ensure fiscal adjustment III (excessive deficit)	37	No	deficit, excess, correct, structur, ensur, specifi, adjust, effort, implement, budgetari, expenditur, progress, suffici, year, strategi, toward, term, medium, growth, benchmark
22	Employability - Education reform	55	Yes	educ, train, vocat, improv, school, market, labour, qualiti, earli, skill, higher, leav, increas, outcom, system, reduc, relev, includ, peopl, youth
23	Labour market - Demand-side reform	39	Yes	labour, market, reduc, improv, peopl, earner, work, employ, low, tax, migrant, high, background, disincent, facilit, servic, enhanc, young, social, unemploy
24	Public sector reform - Tax system reform	34	No	tax, reduc, taxat, improv, complianc, shift, system, environment, base, vat, incom, collect, properti, growth, burden, particular, corpor, consumpt, labour, broaden
25	Healthcare reform	11	No	care, cost, increas, healthcar, health, system, effect, improv, reform, sector, public, qualiti, strengthen, hospit, social, spend, particular, effici, reduc, includ
26	Public sector reform - General II	7	No	strengthen, framework, plan, improv, order, nation, regul, fund, provid, institut, strateg, capac, establish, qualiti, implement, link, action, resourc, particular, made
27	Financial sector reform - Pension system	57	No	age, retir, pension, statutori, life, expect, long, increas, earli, sustain, term, link, ensur, older, system, worker, employ, improv, effect, scheme
28	Labour market - Social protection II	14	Yes	labour, market, employ, perman, increas, system, step, pension, particip, rate, protect, effort, reduc, general, promot, contract, without, ensur, reform, inter
29	Market regulation - Deregulation	54	No	servic, competit, sector, remov, barrier, restrict, regul, retail, regulatori, reduc, market, profess, particular, includ, profession, entri, open, busi, improv, burden
30	Structural policy - Energy-efficiency and renewables	26	No	energi, promot, build, effici, develop, reduc, particular, incent, continu, effort, improv, product, transport, includ, step, target, activ, toward, high, increas

Table A4.4. Top three recommendations per topic

Topic	ID	Score
1	SLV2014_6	0,990069
	CRO2014_7	0,984537
	SLV2013_9	0,939091
2	LAT2013_4	0,65786
	POR2014_8	0,643543
	UK2017_2	0,494671
3	CZE2012_3	0,923322
	CZE2016_3	0,863336
	SLK2016_2	0,792135
4	EST2014_1	0,982599
	HUN2014_1	0,952932
	ITA2014_1	0,842167
5	LIT2012_3	0,977406
	LIT2013_3	0,961849
	FIN2011_3	0,923322
6	LAT2015_1	0,929787
	EST2015_1	0,878622
	AUS2016_1	0,874432
7	POL2013_6	0,985353
	POL2014_5	0,98184
	BUL2012_7	0,975405
8	GER2016_1	0,973861
	POL2012_5	0,906469
	ITA2011_5	0,842231
9	POL2017_1	0,979618
	HUN2017_1	0,973861
	UK2017_1	0,973014
10	SLK2014_6	0,985353
	BUL2014_5	0,954675
	CZE2014_7	0,944876
11	SPA2014_3	0,991981
	CZE2011_4	0,974656
	HUN2011_4	0,973861
12	SWE2013_1	0,953372
	FIN2013_1	0,907198
	SWE2014_1	0,872686
13	FRA2014_2	0,799471
	FRA2015_2	0,717453
	EST2013_5	0,675587
14	ROM2014_4	0,832916
	BUL2012_3	0,815712
	BUL2013_3	0,791151
15	BEL2013_3	0,980105
	FRA2014_3	0,967622
	MAL2012_4	0,966501
16	MAL2015_1	0,935246
	HUN2015_1	0,915547
	POL2015_1	0,889031
17	SWE2016_1	0,984245
	NL2014_2	0,98184
	SWE2017_1	0,980105
18	SPA2013_7	0,938813
	POR2014_7	0,63958
	SPA2012_1	0,527062
19	IRE2014_6	0,987544
	CRO2014_8	0,973861
	IRE2014_5	0,967094
20	SLV2013_8	0,990934
	SLV2014_5	0,928706
	LIT2011_4	0,821265
21	DEN2012_1	0,978569
	POL2012_1	0,958206
	AUS2012_1	0,938886
22	HUN2013_6	0,935304
	ITA2014_6	0,929517
	SLV2012_5	0,885315
23	GER2014_2	0,966949
	NL2012_3	0,966501
	NL2011_3	0,960078
24	ITA2013_5	0,98184
	BEL2014_2	0,978569
	SPA2013_2	0,978003
25	ROM2014_3	0,966501
	IRE2015_2	0,96357
	IRE2014_2	0,934194
26	UK2014_6	0,96357
	ITA2011_6	0,625676
	CZE2011_6	0,611255
27	AUS2013_2	0,97611
	MAL2011_2	0,973861
	NL2011_2	0,971144
28	POL2014_3	0,95689
	POL2013_4	0,953833
	NL2016_2	0,943186
29	DEN2015_2	0,968995
	CYP2012_6	0,960078
	CRO2016_4	0,953372
30	LUX2013_6	0,94396
	EST2012_4	0,902937
	LUX2014_5	0,897003

APPENDIX TO CHAPTER 5

OPERATIONALISATION OF CONTROL VARIABLES

Gross Government Debt and Cyclically-adjusted Government Deficit

These measures capture the levels of government debt and deficit as percentages of GDP, i.e. the EDP's two main fiscal enforcement criteria, at a given time point. The data for these variables was taken directly from the European Commission's AMECO database, which is itself sourced from Eurostat (the Commission's Statistical Office) (European Commission, 2019a). Since 2005, the Commission uses the *Cyclically-adjusted Government Deficit* because it captures a member state's structural fiscal stance in a way that is less dependent on the state of the economy (for details, see Mourre, Isbasoiu, Paternoster, & Salto, 2013).

Political power

I capture *political power* as the voting power of a member state in the Council, using the Shapley-Shubik index of the power distribution under the rules of the Treaty of Nice for 2005-2013 and under the Treaty of Lisbon for 2014-2018. In many policy areas, voting in the Council occurs by qualified majority voting, which requires 55% of member states who together represent 65% of the EU population to pass a vote. As such, the voting power of EU member states is strongly dependent on their populations. The Shapley-Shubik index takes this into account and is based on the probability that a member state is pivotal in turning a losing coalition into a winning one (Hix & Høyland, 2011). It is calculated by determining all possible constellations of coalitions in the Council, and then determining for how many of those potential situations the member state in question would be able to cast the decisive vote. A score of 0 indicates a member state has zero influence over the outcome of decisions in the Council, whereas a score of 100 would indicate the hypothetical situation in the vote of the member state by itself determines the outcome of every vote. For more details, see Napel and Widgrén (2011).

Government positions on the EU and economic matters

A *government's position on the EU* is calculated using the same data as the measure of the mobilisation of Eurosceptic parties in a member state's parliament. For each party in government, I used the CHES scores to determine the position of the party on the anti-pro EU scale. If the party ruled alone, the party's score is also the final score for this observation. If the party was part of a coalition government, I used the average position of all coalition parties, as weighed by their seats in parliament relative to the total seats held by the coalition. A *government's left-right position* on economic matters is calculated in the same way by using a ten-point left-right scale from Döring and Manow's (2016) ParlGov dataset. On this scale, scores below and above 5 indicate

economically left- and right-wing parties, respectively. This measure again uses seat-weighted scores for coalition governments.

Electoral cycle

Finally, I capture the member state's position in its *electoral cycle* as the number of years until the next parliamentary election in this a member state. A score of 0 indicates that an election occurs within the current half-year (t), and a score of 1 indicates an election is one year away (t_{+2}).

MODEL SPECIFICATION

This section presents the mathematical specification of the OMM presented in this contribution (for a more extensive description of this model, see de Haan-Rietdijk et al., 2017). For the probability of switching state (as given in Table 5.1), $\pi_{ijnt} = p(s_{nt} = j | s_{n(t-1)} = i)$, where s_{nt} is the state of country n in year t. The likelihood for the data (s) from time t=2 onwards is given by

$$f(s|\pi) = \prod_{t=2}^{T} \prod_{n=1}^{N} \prod_{i=1}^{S} \prod_{j=1}^{S} (\pi_{ijnt})^{[s_{tn}=j]} \cdot {}^{[s_{(t-1)}=i]}$$

The individual-level transition probabilities (π_{ijnt}) are derived from their logits (a_{isnt}), as given by

$$\pi_{ijnt} = \frac{\exp(\alpha_{ijnt})}{\sum_{s=1}^{s} \exp(\alpha_{isnt})}$$

where

$$\alpha_{ijnt} = \mu_{ij} + \beta_{ij}^{1} \chi^{1}{}_{nt} + \beta_{ij}^{2} \chi^{2}{}_{nt} + \ldots \beta_{ij}^{12} \chi^{12}{}_{nt} + u_{on} + \epsilon_{ijnt}$$

In this last equation, all predictors are grand mean centred ($\chi_{nt} - \bar{\chi}$) and most are lagged χ_{t}–1 (see the Data and Methods section for details on the lagging of predictors). μ_{ij} is the average logit for transitioning from state i to state j, where ϵ_{ijnt} is country n's deviation from μ_{ij} at t, and u_{on} is country n's average deviation from μ_{ij}. Importantly, the intercepts (μ_{ij}) and their error terms (u_{on} u_{on} and ϵ_{ijnt}) are set to 0 whenever $i = j$, which means stability (the member state remains in the same state from t to t_{+1}) is the reference for the transition logits. Similarly, coefficient estimates (β_{ij}) are set to 0 whenever $i = j$.

Regarding prior specification, I follow the diffuse, regularizing prior specifications recommended by de Haan-Rietdijk et al. (2017), who in turn base their recommendations on the work of Andrew Gelman and others. For the intercepts (μ_{12} and μ_{21}) and regression coefficients (β_{ij}^{1} to β_{ij}^{12}), these are independent Cauchy prior distributions with scale parameter 10 for the intercepts and 2.5 for the coefficients, and location parameter 0 in all cases. The multivariate prior for the covariance matrices of the random logit variances is the Inverse-Wishart (IW) distribution with an identity matrix and degrees of freedom equal to the number of random effects (n = 28).

Estimation procedure

The OMM was estimated using R 3.6.1 and JAGS 4.3.0. (Plummer, 2017). I ran four MCMC chains, using 10,000 burnin iterations and 50,000 sampling iterations each, with a thinning parameter of 10 to reduce autocorrelation in the samples. The trace and density plots respectively showed no trend and looked unimodal, indicating adequate convergence. I also ran two separate chains with randomly generated starting values for the sampler. These arrived at the same estimates, again indicating adequate convergence. Finally, pair plots showed no signs of collinearity between the estimates for any of the included predictors. Due to the volume of this output, these plots are available upon reasonable request.

POSTERIOR PREDICTIVE CHECKS

Posterior predictive checks are the only method to assess the fit of a Bayesian Markov model, and I rely on checks presented by de Haan-Rietdijk et al. (2017) and Shirley et al. (2010). The underlying logic of such predictive checks is that the model parameters and their uncertainty can be used to simulate predicted outcomes, which can then be compared to the empirically observed outcomes. If these simulated outcomes are not extreme in relation to the empirical outcomes, the model is a good fit of the empirical data. Here (as well as for the check presented in the main text), I used the model parameters and the empirical data for the predictor variables to simulate 20,000 new predicted outcome variables. For each simulation, I calculated the mean proportion of years countries spent in EDPs and the standard deviation of this proportion over countries. Lastly, I also calculate the distribution of the proportion of years over countries. I then compared these with the actual empirical outcomes (Figure A5.1).

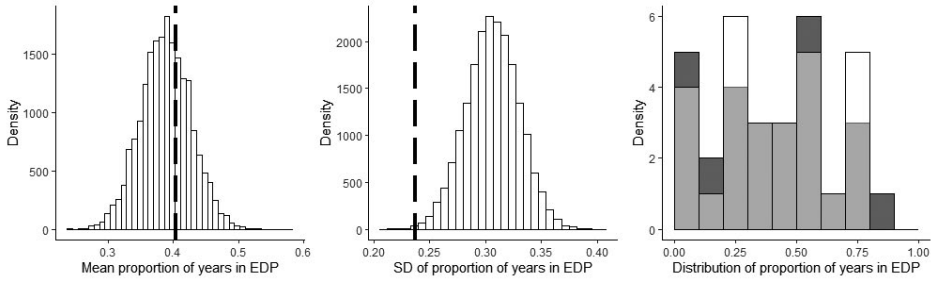


Figure A5.1. Posterior predictive check 2

Note: Posterior predictive checks for the proportion of years countries spend in the EDP in 2005-2017. The dashed line in the first two panels is the empirical value for the statistic. In the last panel, the distribution of model predictions is displayed in dark grey and the empirical distribution in white.

The model accurately predicts the average proportion of years countries spend in the EDP, but overestimates the variance of this proportion over countries. This is evident in the last panel

in Figure A5.1, which shows the model predicts a slightly flatter distribution than the empirical distribution. Figure A5.2 presents a similar posterior predictive check, which compares the simulations to the data regarding the number of state switches countries experienced. The model's predictions regarding the mean proportion of switches across countries are again accurate. The model again overestimates the variance in the proportion of switches between countries, although the overestimation is less extreme here. The model overestimates the number of countries that switch twice and thrice, and slightly underestimates other frequencies.

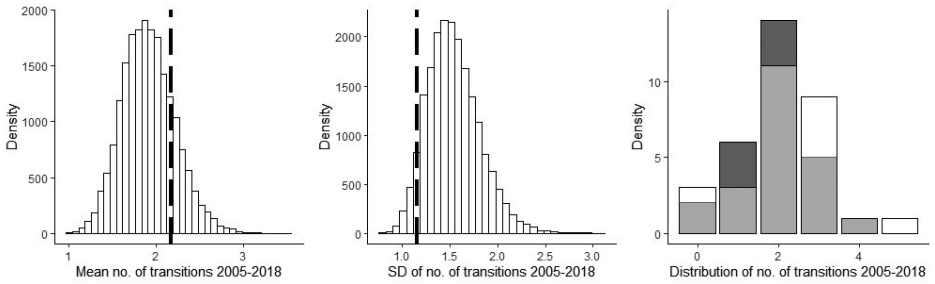


Figure A5.2. Posterior predictive check 3
Note: Identical to Figure A5.1, but for the number of between-state transitions.

Lastly, I checked the model's predictive accuracy against the predictive accuracy of a mixed-effects OMM that only includes predictors for member state's nominal government deficits and gross government debts. This check is presented in Figures A5.3-5: these figures are essentially the same predictive check as presented in Figure 5.3, but here an additional (dotted) line presents the predictions made by the OMM which included only the deficit and debt predictors.

These plots reveal that, as expected, the substantive effects of the political predictors are small compared to the fiscal predictors. The full model does a better job at predicting state switches in the cases of Austria, Belgium, Bulgaria, Cyprus, Denmark, Estonia, Finland, Germany, Greece, Hungary, Latvia, Italy, Poland, Portugal, Spain and the United Kingdom (16 countries). It performs worse in the cases of Malta, the Netherlands and Romania (3 countries). Differences are negligible in the cases of Croatia, the Czech Republic, France, Ireland, Lithuania, Luxembourg, Slovakia, Slovenia and Sweden (8 countries). The overall small differences between both models again suggest that Euroscepticism may push the Commission to signal resolve only in the margins of its discretionary space. However, differences in predictions are likely to be suppressed in these plots, as the random effects included in the simpler OMM absorb a large portion of the structural differences in levels of Euroscepticism between member states.

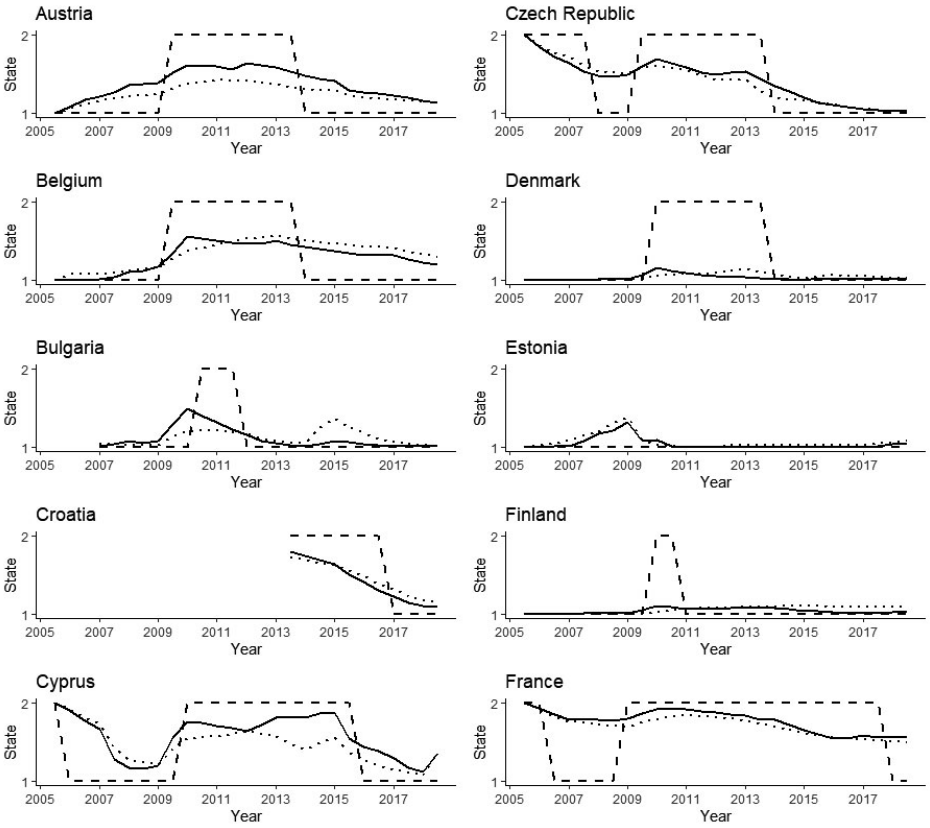


Figure A5.3. Posterior predictive check 4

Note: Posterior predictive checks for the country-specific EDP-trajectories in 2005-2017. Solid lines display the model's predicted trajectories and dashed lines display the empirical trajectories. Dotted lines display the trajectories predicted by a model including only the deficit and debt predictors. Predicted trajectories were calculated by taking the mean of state predictions for a country-year.

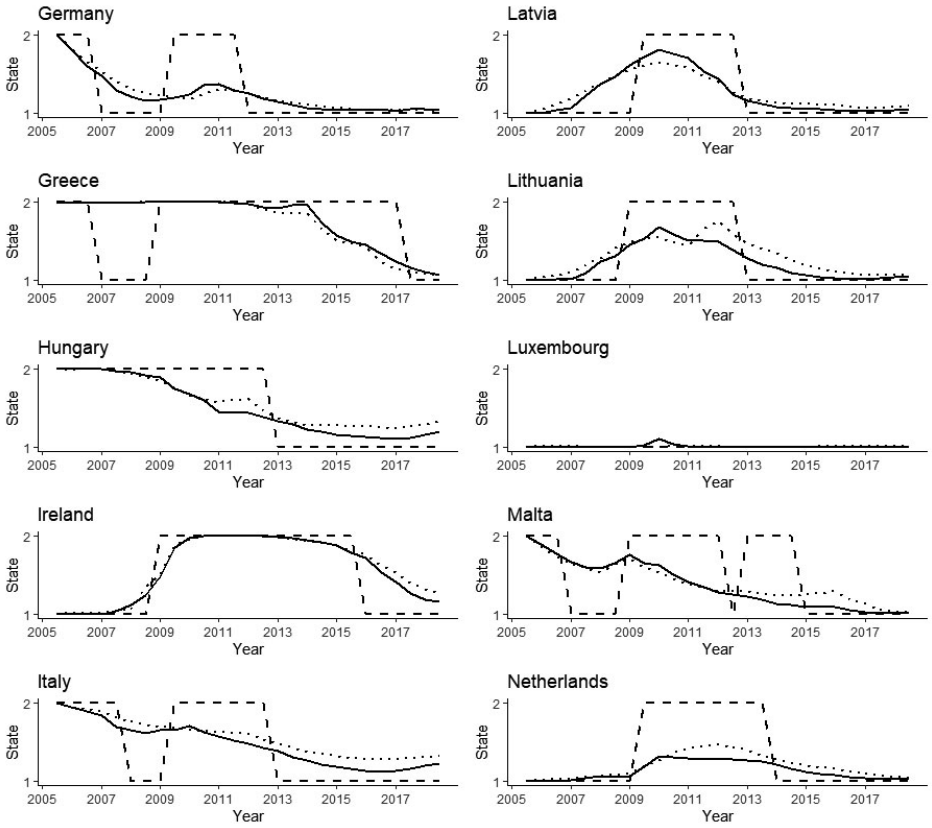


Figure A5.4. Posterior predictive check 4 - continued

Note: Posterior predictive checks for the country-specific EDP-trajectories in 2005-2017. Solid lines display the model's predicted trajectories and dashed lines display the empirical trajectories. Dotted lines display the trajectories predicted by a model including only the deficit and debt predictors. Predicted trajectories were calculated by taking the mean of state predictions for a country-year.

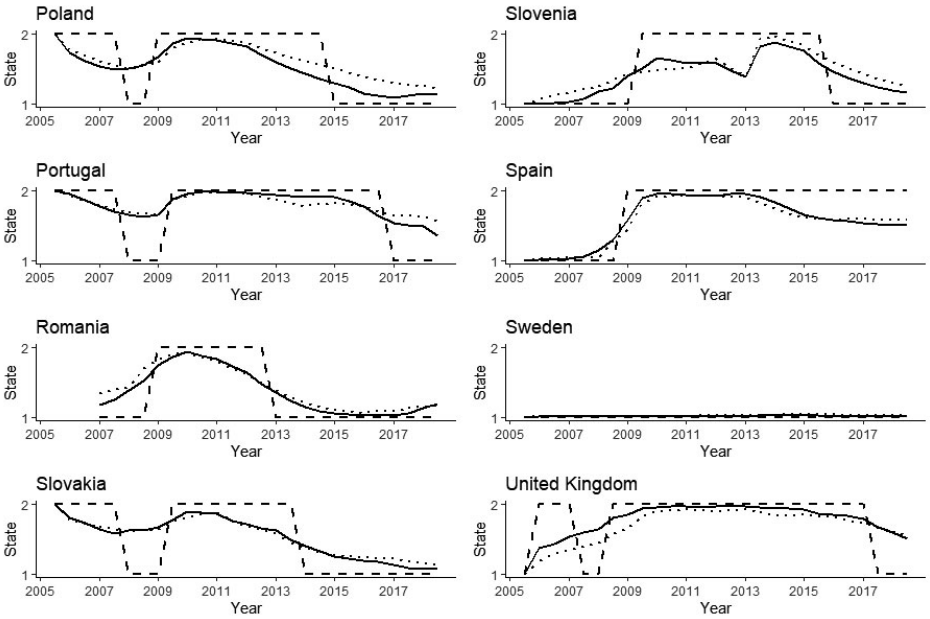
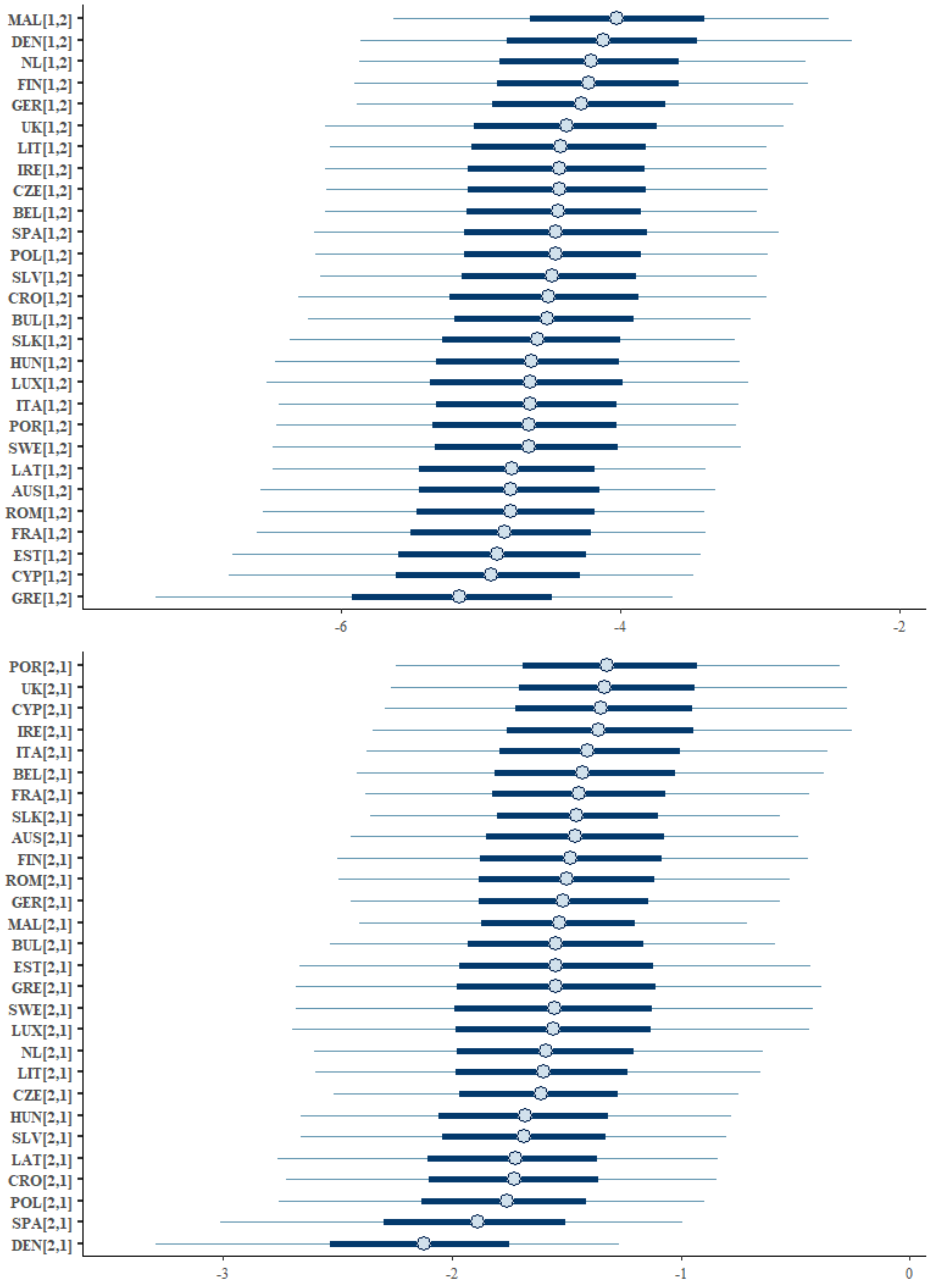


Figure A5.5. Posterior predictive check 4 - continued

Note: Posterior predictive checks for the country-specific EDP-trajectories in 2005-2017. Solid lines display the model's predicted trajectories and dashed lines display the empirical trajectories. Dotted lines display the trajectories predicted by a model including only the deficit and debt predictors. Predicted trajectories were calculated by taking the mean of state predictions for a country-half year.



Figures A5.6. Random effects of the full OMM

ADDITIONAL ROBUSTNESS CHECKS USING ALTERNATIVE MODELLING STRATEGIES

In order to further assess the robustness of the OMM's results, I conducted a number of additional analyses using more simplistic models that predict when countries move into an EDP. This exercise shows both the robustness of the OMM's results, as well as its superiority over alternative specifications.

A first issue is that less complex model specifications do not properly account for the state-dependency of the data. A simple logit model is capable of predicting whether a country is subject to an EDP or not, but treats all time-points as independent and cannot properly account for the longitudinal dependency of the data (where s_{nt} depends on s_{nt}–1). In other words, it does not understand the substantive difference between a country switching into or remaining subject to an EDP. For non-hierarchical data structures, the inclusion of lagged dependent variables or autoregressive structures could circumvent this. However, such solutions are highly problematic in combination with random (or fixed) effects that account for clustering as they break the assumption that error terms are independent from the model's predictors. This leads to erroneous estimates and inferences (Xu, DeShon, & Dishop, 2019). Hence, the first step towards estimating any alternative model is to remove all observations of countries being subject to an EDP, when they were in the same state at t_{-1}. This way, the only observations retained are those of countries not subject to EDPs at t, or those for which an EDP is launched at t. On such a sample, a logit would predict the opening of EDPs, albeit on a skewed subset of the empirical data. This subsample of the data consists of 454 of the 756 observations.

I subsequently estimated five different models using this subsample (M1-5), which are presented in Table A5.1. Models 1-4 are frequentist logistic multilevel models that predict the opening of EDPs. Model 1 includes only the random intercept for member states, as well as the two fiscal predictors. Model 2 adds to this the four political control variables. Model 3 subsequently adds the predictors for the reputational threats faced by the Commission, as well as the interactions included in the model. Models 1-3 do not account for temporal clustering or dependence in the data in any way. Model 4 is identical to Model 3, except that it also adds a random intercept for t. This does not provide a solution to the time-dependency within clusters (t is dependent on t_{-1}), but provides an important robustness check for the temporal dependence across clusters (observations for different countries at the same point in time are more likely to be similar). This type of temporal dependence is not accounted for in the OMM, which only accounts for dependency within clusters; including the random intercept for t there would again break the assumption of independent errors. Finally, Model 5 presents the results of a Bayesian Cox Proportional Hazard model with diffuse regularizing priors. As the OMM, this survival model accounts for temporal dependence within clusters, but does not include a random intercept for t.

First and foremost, Models 1-3 and Model 5 show the strong robustness of the OMM's estimates using these alternative model specifications, at least regarding the results for the opening

Table A5.1. Results using alternative models

	M1			M2		
Predictors	Log-Odds	Conf. Int	p	Log-Odds	Conf. Int	p
(Intercept)	-2.96	-3.48 – -2.43	<0.001	-3.07	-3.64 – -2.50	<0.001
Gross government debt12	-0.26	-1.25 – 0.74	0.614	-0.61	-1.70 – 0.48	0.276
Cycl.-adjusted government deficit12	3.43	2.29 – 4.57	<0.001	3.66	2.44 – 4.87	<0.001
Voting power12				0.75	-0.21 – 1.71	0.124
Public Euroscepticism MS12						
Euroscepticism MS parliament12						
Euroscepticism creditor countries12						
Government position EU12				-0.28	-1.21 – 0.65	0.558
Government position left-right12				-0.34	-1.30 – 0.62	0.484
Electoral cycle12				0.24	-0.63 – 1.12	0.586
Pub. Eurosc. MS·Deficit12						
Pub. Eurosc. MS·Eurosc. MS par.12						
Pub. Eurosc. MS·Eurosc. cred.12						
PSD Random Int. MS12						
σ2	3.29			3.29		
τ00	0.00 MS			0.00 MS		
Groups	28 MS			28 MS		
N	454			454		
Events						

of EDPs. However, Models 1-3 fail to detect any between-group variance (τ_{00}), which strongly suggests these models have difficulties in returning reliable estimates. The addition of a random intercept for *t* in Model 4 absorbs all variance that is explained by the predictors in Model 3, to such an extent that even the estimate for the cyclically-adjusted deficit is no longer statistically significant. Model 4 suddenly also detects substantial unexplained variance across countries (τ_{00}), while Models 1-3 did not.

Table A5.2. Fit improvement across Models 1-4

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
M0	2	197,5584	205,7946	-96,7792	193,5584	NA	NA	NA
M1	4	156,4294	172,9018	-74,2147	148,4294	45,12899	2	1,59E-10
M2	8	160,6292	193,574	-72,3146	144,6292	3,800241	4	0,433715
M3	14	144,9915	202,6449	-58,4958	116,9915	27,63767	6	0,00011
M4	15	98,81339	160,5848	-34,4067	68,81339	48,17814	1	3,89E-12

Note: M0 is a baseline model including only the (random) intercept.

M3			M4			M5		
Log-Odds	Conf. Int	p	Log-Odds	Conf. Int	p	Log-Odds	Cred. Int.	Haz. Rate
-4.94	-6.56 – -3.33	<0.001	-44.74	-93.30 – 3.82	0.071	-3.052	-5.224 – -0.926	NA
-0.50	-1.80 – 0.79	0.446	0.06	-50.92 – 51.04	0.998	-1.251	-3.257 – 0.726	0,286
3.78	1.99 – 5.57	<0.001	50.92	-17.13 – 118.98	0.142	5.743	3.461 – 8.667	312,116
0.70	-0.44 – 1.84	0.227	17.38	-32.30 – 67.07	0.493	1.693	0.233 – 3.145	5,433
3.81	1.36 – 6.26	0.002	-5.11	-75.65 – 65.42	0.887	4.119	1.055 – 7.243	61,511
-0.55	-1.95 – 0.86	0.446	4.20	-45.71 – 54.11	0.869	-0.735	-2.801 – 0.966	0,48
-4.87	-7.62 – -2.12	0.001	0.67	-79.00 – 80.34	0.987	-4.201	-7.304 – -1.448	0,015
-0.22	-1.31 – 0.87	0.689	8.47	-21.10 – 38.05	0.574	-0.48	-1.847 – 0.931	0,619
-0.29	-1.32 – 0.73	0.575	2.30	-31.04 – 35.64	0.893	-0.035	-1.256 – 1.192	0,965
-0.07	-1.05 – 0.90	0.881	-3.74	-42.37 – 34.89	0.850	0.188	-0.935 – 1.317	1,207
3.17	-0.51 – 6.85	0.091	38.35	-121.46 – 198.17	0.638	4.343	-0.231 – 9.422	76,934
-1.55	-4.83 – 1.74	0.356	-8.47	-170.65 – 153.70	0.918	-3.336	-8.415 – 1.077	0,036
6.76	2.58 – 10.94	0.002	7.77	-95.70 – 111.23	0.883	7.08	2.396 – 11.931	1187,645
						1.2	0.14 – 2.47	NA
3.29			3.29					
0.00 MS			514.09 MS					
			2436.60t					
28 MS			28 MS			28 MS		
			27t					
454			454			454		
						28		

THE NEGATIVE BASE EFFECT OF PUBLIC EUROSCEPTICISM IN CREDITOR STATES

Why does the counterintuitive, credibly negative coefficient for the effect of Euroscepticism in creditor states occur in all models except Model 4? In short, it is a consequence of inevitable misspecification. As explained above, in an ideal world the modelling strategy given the data in this contribution accounts for both within-country temporal dependence (where s_{nt} depends on s_{nt}–1), and the between-cluster temporal dependence (where observations for Bulgaria and Spain in 2014 are more likely to be similar than those for the same countries in different years). But the inclusion of random or fixed effects in a model that also controls for within-cluster dependence leads to severe estimation errors (Xu et al., 2019). However, in the OMM and Models 3 and 5, the omission of a random effect for *t* yields a biased estimate of the coefficient for Euroscepticism in creditor states, because the latter is the only variable in the model with the same scores across countries (i.e., it is invariant across groups).

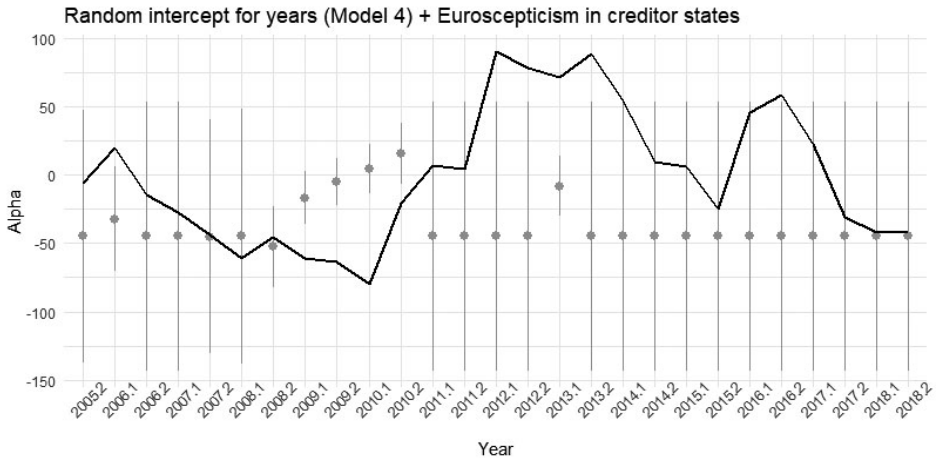


Figure A5.7. Creditor Euroscepticism and unexplained variance over time

Figure A5.7 visualises the issue clearly by plotting the observed values of this predictor against the distribution of the random intercept for *t* included in Model 4, which *only* controls for between-cluster temporal dependence. A large number of time points have random effects with large (and consistent) confidence intervals: these are time points in which no EDPs were opened, and as such there was insufficient information to estimate the intercept (recall also that this model uses only a subsample of the data). For the periods where EDPs were opened, the estimates of the random intercept are more accurate. Hence, Figure A5.7 indicates that, on average, levels of Euroscepticism in creditor states were relatively low in the period when many EDPs were opened. When the random effect for *t* is excluded, the predictor for Euroscepticism in creditor states therefore erroneously absorbs some of the unexplained variance from the period 2008-2010. This period coincides with the onset of the crisis, when a large number of EDPs were opened as most EU member states ran into severe fiscal problems.

Whereas the exclusion of a control for between-cluster temporal variation in the OMM biases the estimate for this predictor, it does not impact the substantive implications of the findings presented in this contribution. First, Model 4 indicates that the inclusion of a control for between-cluster dependence at the expense of a control for within-cluster dependence leads to far more problematic inference errors. Second, this contribution does not argue for a direct effect of Euroscepticism in creditor states on the Commission's enforcement of EDPs. Instead, it argues that this variable conditions its response to another reputational threat, being Euroscepticism in the target state. The overall conclusion that the positive effect of Euroscepticism in the target state is stronger when Euroscepticism in creditor states is higher is therefore not dependent on the base effect of Euroscepticism in creditor states. Instead, Figure A5.7 suggest that reputational considerations regarding Euroscepticism likely played less of a role in late-2008 and early-2009, which is not unsurprising given the sudden upward surge of member state debts and deficits at the onset of the crisis.

APPENDIX TO CHAPTER 6

QUANTITATIVE EXPLORATION OF POLITICISATION DIMENSIONS

In order to get an idea about levels of politicisation in the three member states covered in the case studies, as well as politicisation in creditor countries, these levels were visualised over time using quantitative indicators for each dimension. Importantly, the figures below were not explicitly part of the process tracing exercise, but were used to get an understanding of the politicisation conditions relevant for these cases prior to case selection. The three dimensions of societal EU politicisation (salience, polarisation and mobilisation), as well as levels of issue salience, were visualised for the three target countries (the UK, Finland and Italy), as well as for two creditor states (Germany and the Netherlands). These two creditor states were selected because both countries are hard-line proponents of fiscal discipline in the EU and are influential in shaping EMU politics in that regard (Germany is the EU's economic powerhouse and the Netherlands has emerged as the lead country in the Hanseatic League).

Salience EU/EC is measured using monthly aggregates of newspaper articles mentioning the terms 'European Commission' or 'European Union' in major business newspapers in the respective member states in the Lexis Nexis database.[37] The selected newspapers were Kauppalehti (Finland), Handelsblatt (Germany), NRC Handelsblad (The Netherlands) and the Financial Times (the UK). In case of Italy, no major business newspaper was available and Corriere della Sera, a major daily newspaper, was used instead.

The *polarisation* of public opinion on the EU is captured by calculating the kurtosis of the distribution of citizen opinions on the EU in each member state, using the weighted individual-level responses to the Eurobarometer item: "In general, does the EU conjure up for you a very positive, fairly positive, neutral, fairly negative or very negative image?" Kurtosis is a measure of the shape of the distribution of opinions, where *lower* scores indicate that more observations are farther from the mean and therefore indicate stronger public contestation of EU legitimacy. The weighted overall variance in responses to this item is highly correlated to the kurtosis measure (r = -0.686). However, variance is equally affected by a few extreme outliers as it is by more frequent deviations from the mean opinion. This makes kurtosis a better measure of polarisation (Rauh, 2016).

Mobilisation can be conceptualised in different ways, and the preferred measure in this study would be SGP-specific, such as public protests addressing EU fiscal policy (Rauh, 2016). However, this data is not readily available. Thus, for this quantitative exploration, mobilisation is captured in more general terms through the strength of Eurosceptic challenger parties in a member state's

37 For Finland, Germany, Italy, and the Netherlands, these terms were translated to Finnish, German, Italian and Dutch, respectively.

parliament. Voting behaviour has been a primary focus of studies on political mobilisation in political science for decades (e.g. Bond et al., 2012; Holbrook & Mcclurg, 2005; Kriesi, 2008). Eurosceptic parties were identified using the Chapel Hill Expert Survey seven-point anti-pro party position scale on European integration (Polk et al., 2017). This data was used to create a continuous measure indicating the share of seats occupied by parties in a national parliament that score below four on the scale; these can be classified as 'soft' or 'hard' Eurosceptics (Ray, 2007).

Issue salience is captured through the share of respondents answering the economic situation, rising prices/inflation, unemployment and/or pensions to the following Eurobarometer question: *What do you think are the two most important issues facing (OUR COUNTRY) at the moment?* These four categories were chosen because they highlight different elements of a more general public concern with economic insecurity, and higher scores on this variable indicate a greater likelihood that citizens will take an interest in specific developments regarding EU fiscal surveillance in their own countries and across the EU more generally.

Polarisation, mobilisation and issue salience are visualised separately for each member state under surveillance. For the two creditor states, the average of the two countries was visualised. In all cases, polarisation, mobilisation and issue salience are plotted as deviations from the EU average (with averages given as dashed lines and member state values as grey dots). Lastly, the two fiscal indicators key to surveillance under the SGP are were also plotted for the three countries under surveillance, as a percentage of GDP and relative to the SGP thresholds. The data for these indicators was taken from Eurostat.

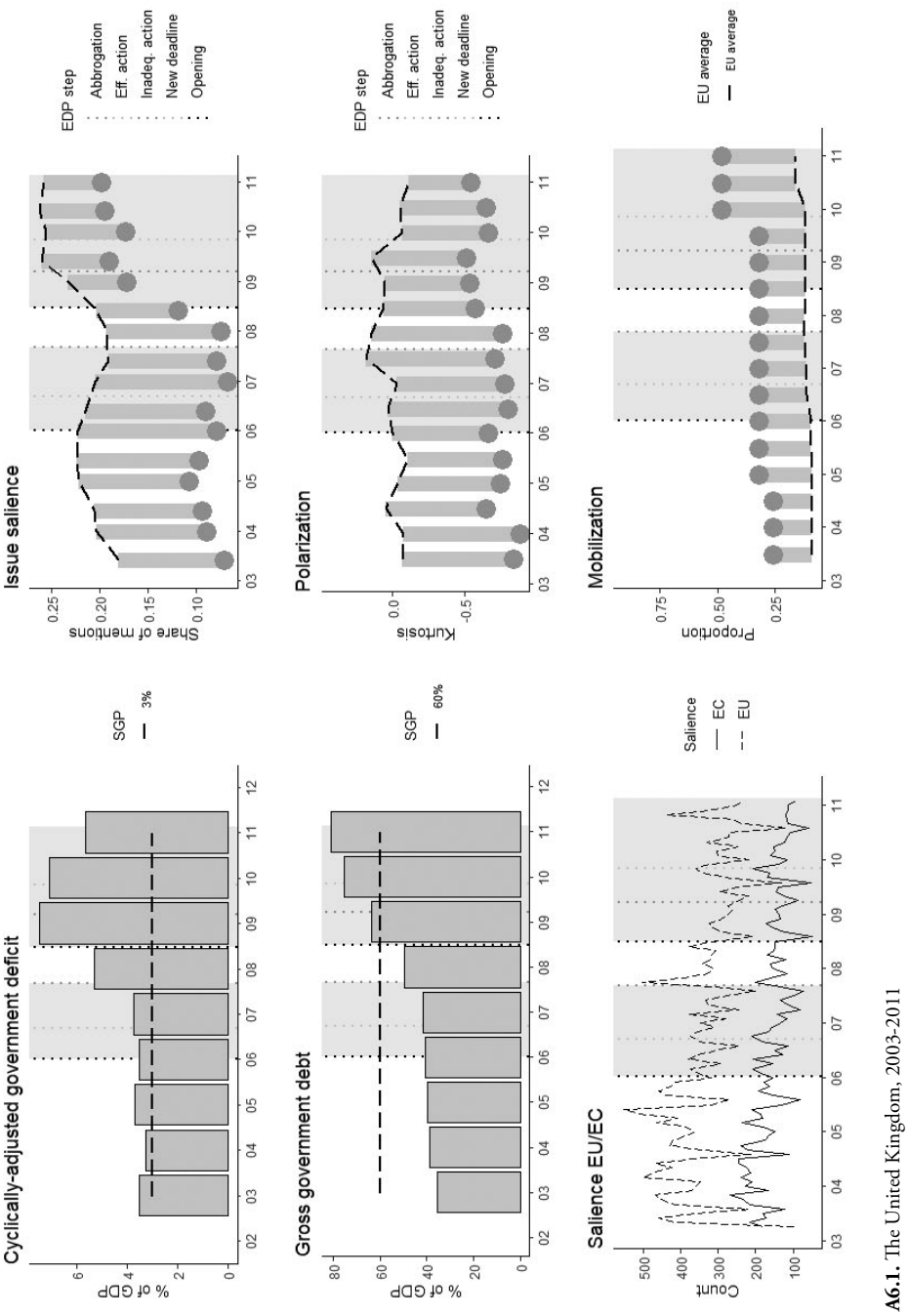


Figure A6.1. The United Kingdom, 2003-2011

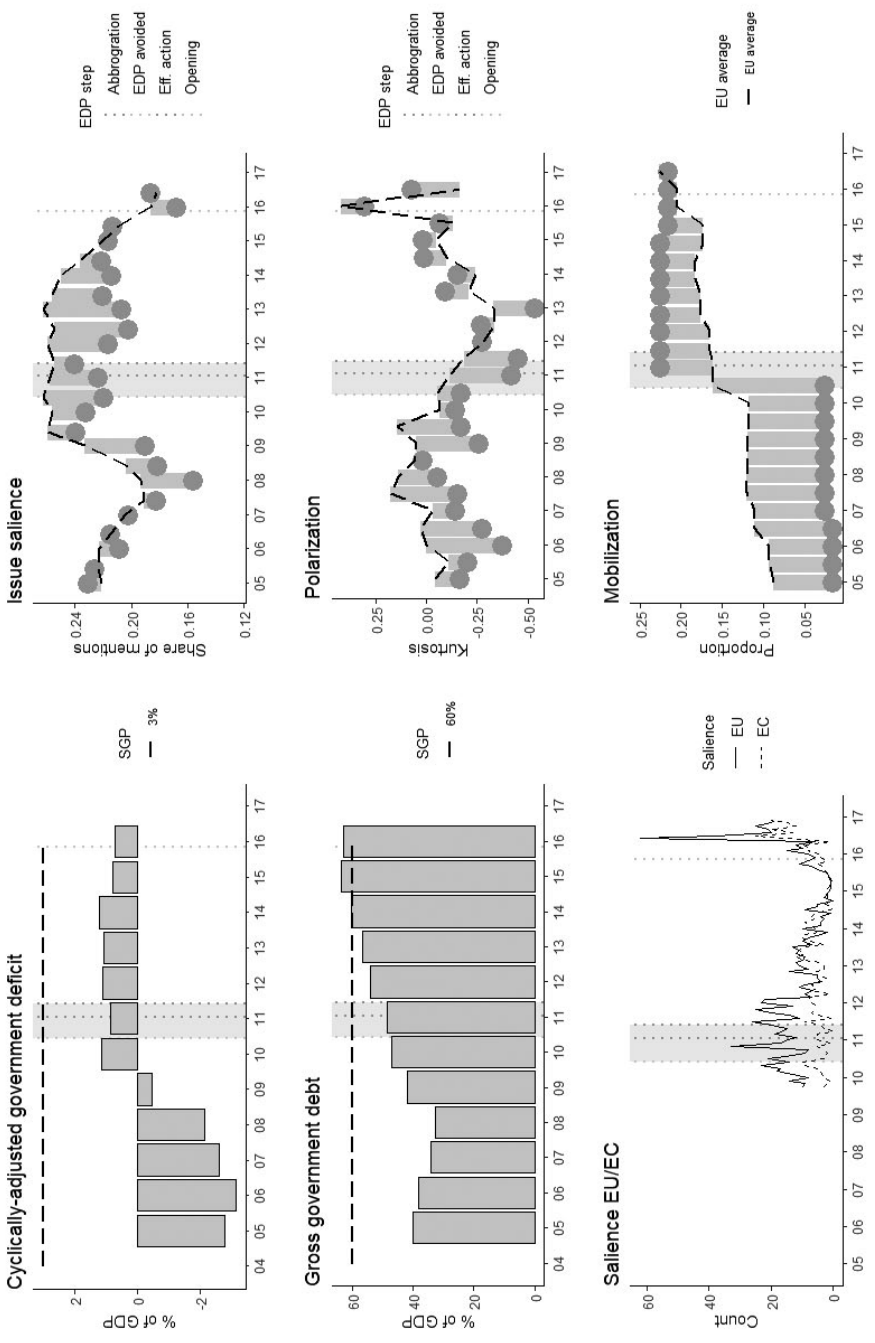


Figure A6.2. Finland, 2005-2017

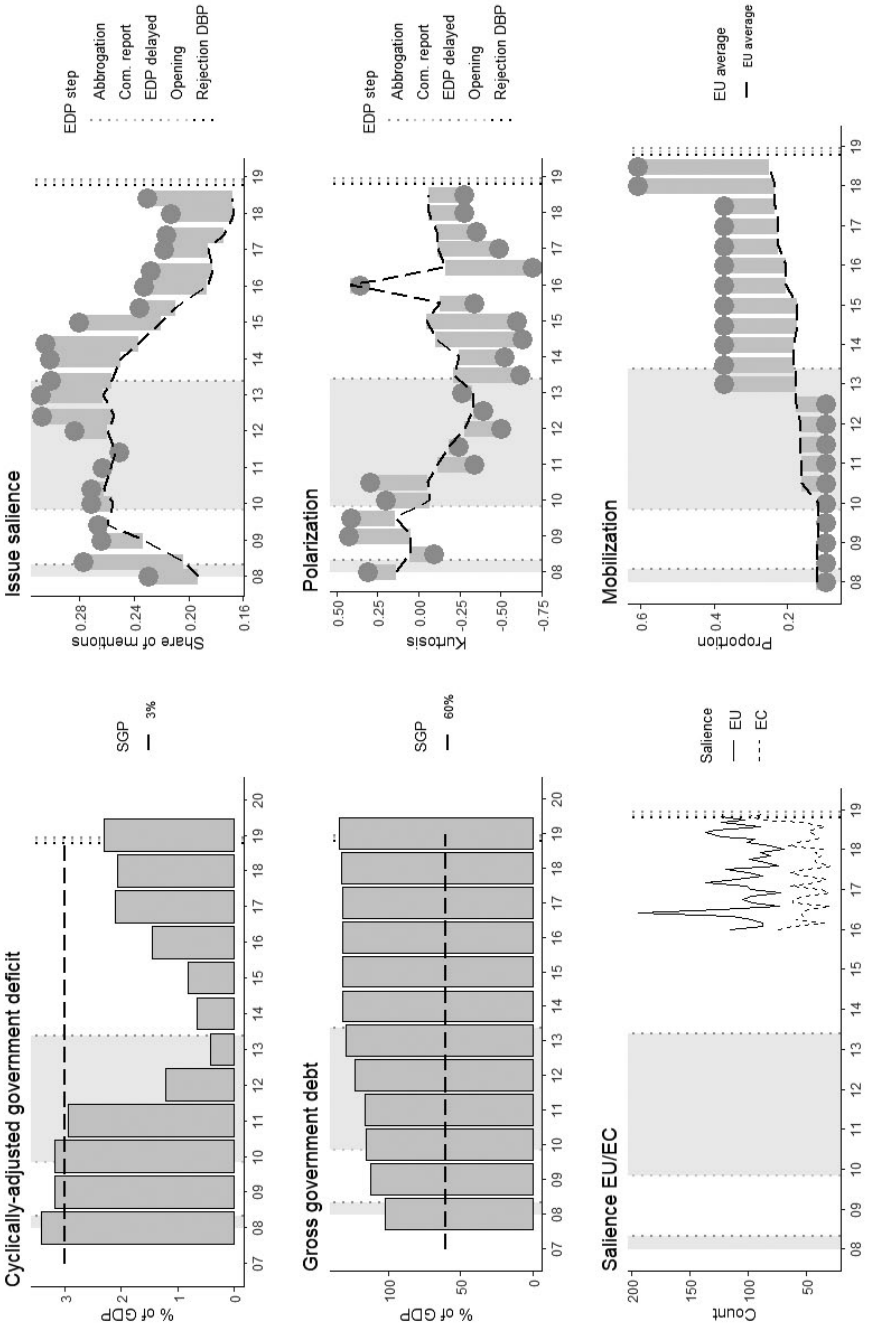


Figure A6.3. Italy, 2008-2018

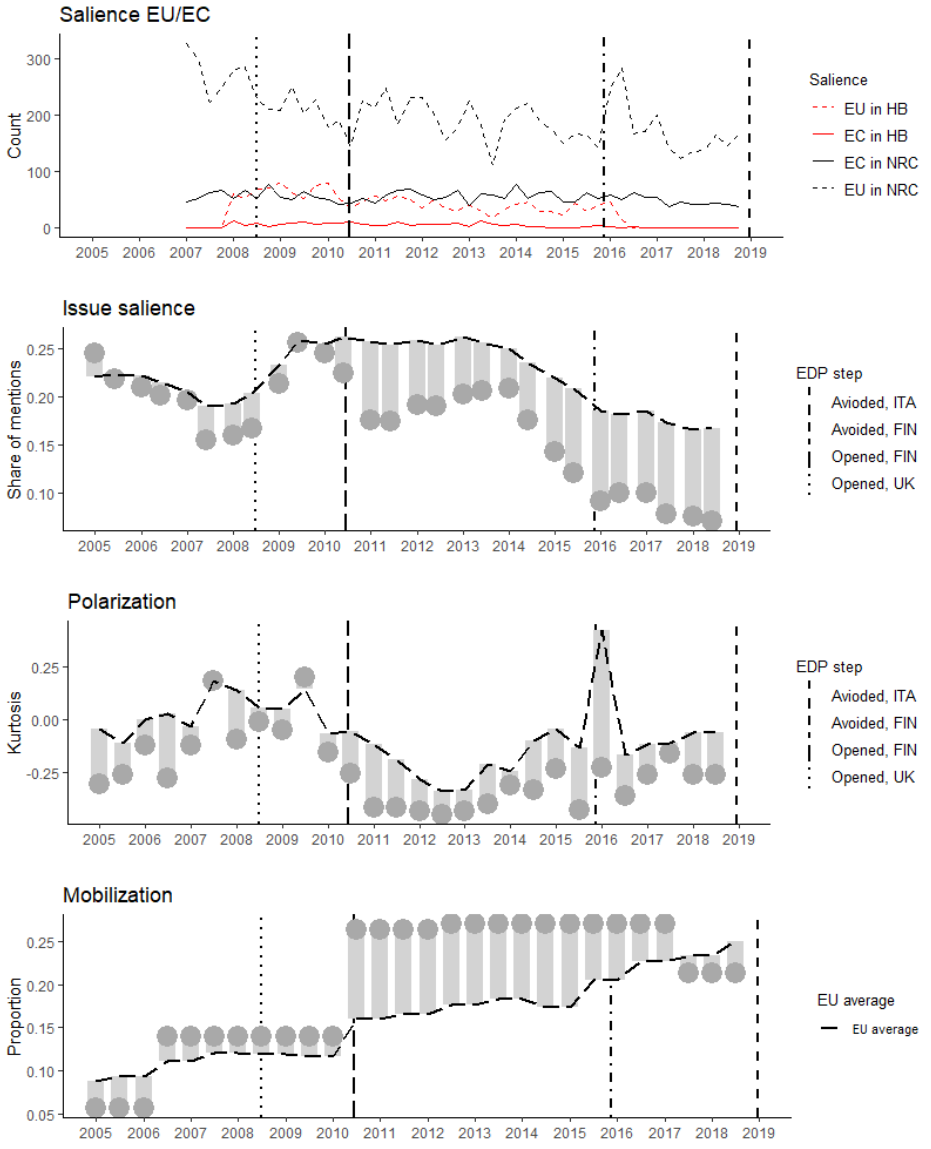


Figure A6.4. Germany and the Netherlands, 2005-2018

LIST OF INTERVIEWS (I)

The interviews were conducted between November 2018 and May 2019, with lengths ranging between 40-120 minutes. The list below gives the positions of interviewees at the time of the interview. However, given the small network of experts working on EU fiscal surveillance, many interviewees previously held other positions in which they were involved in SGP enforcement across EU institutions, and some of the observations shared by them in interviews stem from these periods. Given the high risk of identifiability (e.g. many member states often only have a single expert working on EU fiscal surveillance, and the Commission has a single official working on a country-desk), it is not possible to disclose nationalities or previously held positions.

I1. Commission official, 18-11-2018
I2. Member state official, 09-01-2019
I3. Council secretariat, 04-02-2019
I4. Council secretariat, 04-02-2019
I5. Commission official, 13-02-2019
I6. Commission official, 19-03-2019
I7. Former EFC official, 08-04-2019
I8. Commission official, 12-04-2019
I9. Former EFC president, 02-05-2019
I10. Commission official, 02-05-2019
I11. Former EFC official, 07-05-2019
I12. Member state official, 08-05-2019

LIST OF DOCUMENTS (D)

D1. (European Commission, 2008c)
D2. (European Commission, 2008a)
D3. (ECOFIN Council, 2008)
D4. (European Commission, 2008b)
D5. (European Commission, 2010a)
D6. (European Commission, 2015a)
D7. (European Commission, 2015b)
D8. (European Commission, 2015a)
D9. (European Commission, 2015c)
D10. (European Commission, 2018a)
D11. (European Commission, 2018b)

LIST OF NEWSPAPER ARTICLES (N)

N1. (The Guardian, 2008)
N2. (The Daily Telegraph, 2008b)
N3. (The Daily Telegraph, 2008a)
N4. (Mail on Sunday, 2008)
N5. (Daily Mail, 2008)
N6. (Financial Times, 2011)
N7. (Kauppalehti, 2010)
N8. (Kauppalehti, 2015)
N9. (Politico, 2018b)
N10. (Financial Times, 2018c)
N11. (Politico, 2018c)
N12. (Il Giornale, 2018)
N13. (La Stampa, 2018)
N14. (Financial Times, 2018d)
N15. (Financial Times, 2018a)
N16. (e.g. New York Times, 2018)
N17. (Politico, 2018a)
N18. (Reuters, 2018)
N19. (Financial Times, 2018b)
N20. (Handelsblatt, 2008)
N21. (NRC Handelsblad, 2008)
N22. (De Volkskrant, 2010)
N23. (NRC Handelsblad, 2010)

INTERVIEW QUESTIONNAIRE

The following questionnaire was used to structure the semi-structured interviews. It was adapted on a case-by-case basis to suit the interviewee's position and the cases that were discussed.

Introduction

1. A general introduction on the project and the purpose of the interview was provided. Interviewees were informed of their rights under the GDPR and were asked to consent to the interview. Interviewees were asked whether I could record the interview and how I should refer to the interview.
2. What are your activities at [organisation] with regard to the SGP?
3. [If not Commission staff:] In what ways are you in contact with the European Commission?

Credibility of enforcement

4. How would you describe the Commission's reputation in relation to the enforcement of the SGP?
5. What does the Commission do to try to maintain the credibility of enforcement when applying the Pact?
6. To what extent do you believe all member states are treated equally?
7. How would you describe the style of enforcement under Barroso?
8. How would you describe the style of enforcement of the Juncker Commission?
9. Commissioner Moscovici has repeatedly talked about 'politically intelligent enforcement' of the SGP. What do you think this entails?

Methodology and discretion

10. How do you view the methodology of the Commission for assessing the fiscal stance of a member state?
11. In your view, how much discretion does the Commission have in shaping the enforcement of the SGP?
12. How standardised are the economic analyses underpinning Commission assessments?
13. [If not Commission staff:] Does [your organisation] check the Commission's analyses?

Reputation and public opinion

14. In your view, how concerned is the Commission with upholding a reputation as good fiscal watchdog?
15. To what extent have public perceptions of the EU in the target country been a concern in the enforcement of the corrective arm of the Pact?
16. And perceptions outside the target country?

17. How important is the Commission's reputation when applying the corrective arm of the SGP?

Examples of case-specific questions

18. An EDP was (not) opened for [Member state] in [year]. Can you tell me what happened?
19. What options were on the table before this decision was taken?
20. Why did the Commission go for this option?
21. Was there much debate about the application of exception clauses? Between whom?
22. Why did [member state] (not) present relevant factors?
23. How contested was this decision in the college?
24. To what extent was the EDP considered necessary at the time?
25. How did the [member state] government respond to the (possibility of an) EDP? Why?
26. How contested was the launch of the EDP in the [member state] media?
27. Which other actors took an interest in the (potential) EDP?

Fiscal surveillance and [member state] politics

28. Can you tell me how public and political perceptions of EU fiscal surveillance in [member state] have developed over time?
29. To what extent have perceptions on the Commission's implementation of the Pact led to public criticism?
30. How was the [other member state's] case discussed in [member state]?
31. By whom?

Closing questions

32. To what extent do you foresee a scenario in which the Commission does impose a sanction for fiscally disobedient member states?
33. How successful has the Commission been in enforcing the SGP?
34. In your view, what are the major (political) difficulties for the Commission with respect to the enforcement of the SGP's corrective arm?
35. What do you believe should be done to improve the corrective arm of the SGP?
36. Do you have any further questions or is there anything you want to add?