



Council Checks of the Commission under the European Semester: Does Member State Power and Euroscepticism Still Matter?*

NICOLE BAERG¹  and MARK HALLERBERG² 

¹University of Essex, Colchester ²Hertie School, Berlin

Abstract

The European Commission recommends evaluations of the economic plans of member states. The Council then provides final text. Previous research suggests that the Council selectively edits the Commission's recommendations. Both large member states and those with eurosceptic populations are likely to have the Council weaken what the Commission writes about them. We examine the determinants of the Council's editing of these texts after the introduction of the European Semester in the period 2011–18. We also account for different varieties of euroscepticism as the literature has become more differentiated. Using metrics of textual changes based on automated and hand-coding, we find little evidence of a systematic relationship between euroscepticism and textual editing. We do, however, find evidence that member state's voting power and euro status matters. Our findings suggest that eurosceptic no longer affects textual editing of such documents under the European Semester.

Keywords: European semester; economic and monetary union; European Commission

Introduction

After September 2008, or when Lehmann Brothers collapsed, several European Union (EU) member states introduced fiscal expansions as a reaction to the Global Financial Crisis (GFC). By 2010, fiscal difficulties in Greece and elsewhere led to a reversal in the direction of fiscal policy towards more contractions and concerns about how large deficits and corresponding debts would impact other member states. The EU established at this time 'The European Semester', which was meant to address perceived weaknesses in European economic policy.

There were several reasons why actors from governments to European parliamentarians judged the existing system of economic governance at the European level as inadequate and wanted to change it. The response to the European Commission's call for joint fiscal expansions at the beginning of the GFC was uneven; some member states expanded while others did not. One early lesson was that economic coordination alone was insufficient. Later, once it became clear that the budget deficit of Greece was much higher than initially reported to European authorities, there were also concerns that member states did not know

*This article and Special Issue were both supported in part by a Jean Monnet Network entitled 'The Politics of the European Semester: EU Coordination and Domestic Political Institutions (EUROSEM)' Agreement no. 600110-EPP-1-2018-1-CA-EPPJMO-NETWORK (Grant agreement no. 2018-1359), with the support of the Erasmus+ programme of the European Union. Research for Mark Hallerberg's contribution is part of the Cluster of Excellence "Contestations of the Liberal Script" (EXC 2055, Project-ID: 390715649), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy.

enough about what other member states were doing. Many saw a general lesson that not enough was done to prevent member states from running large deficits in the first place.

Previous literature finds that there were political as well as economic reasons for how, in the pre-European Semester system, *some* (though not all) member states could avoid stringent reforms (for example Baerg and Hallerberg, 2016; Mariotto, 2019). The European Commission (henceforth ‘Commission’) wrote annual reports on the economic plans (either Convergence or Stability Programmes) that member states submitted to it. It was expected to play a ‘watchdog’ function and point out member states whose economic policies were not in line with European rules. These rules, in turn, were meant to promote a common economic policy under the ‘Economic and Monetary Union’ and to minimize negative externalities from policies in other states. Importantly, the Commission’s document was not the last word. The Council, which meets under the Economic and Financial Affairs Council configuration (Ecofin) of the member states, had an opportunity to rewrite the formal document that evaluated member state economic policy both contemporaneously and as planned. Previous research showed that the Council was more likely to weaken the content of the document provided to member states (and to the press) when member states were large (such as France and Germany) or if the domestic populations of member states were generally eurosceptic – that is, less supportive of the EU (Baerg and Hallerberg, 2016). This meant that there was a political logic embedded in the guidance that the Council provided to member states as well as an economic one. To use an analogy: previous research found that the watchdog seemed to bark at smaller, friendly dogs that liked the EU while leaving bigger and/or unfriendly dogs alone.

This process was part of the ‘preventive’ part of the Stability and Growth Pact, and, as we document below, the European Semester was intended to strengthen this part. There were changes to European economic governance more generally, such as reverse qualified majority voting on Commission proposals on some issues, that strengthened the position of this actor (Franchino and Mariotto, 2020). As we will demonstrate below, however, Council edits to Commission texts continued post reforms. In this paper, we explore whether a similar political logic continued for member state editing in the period after the introduction of the European Semester.

We make two key departures from previous work. First, and consistent with Baerg and Hallerberg (2016), to measure textual editing we use both quantitative and qualitative techniques. Applying both automatic and human coding techniques to measure the amount, type, and depth of editing, we contribute an understanding of the patterns of textual editing made to the Commission’s reports across countries and over time in the post-reform era. We are especially interested in substantive edits – or changes to the meaning of the texts. As before, we explore whether substantive edits are correlated with politically powerful member states and those with more eurosceptic populations. Second, and different from Baerg and Hallerberg (2016), we consider different classes or types of textual editing. Baerg and Hallerberg (2016) use the ‘Levenshtein distance’ measure, which counts all changes across the two documents. In addition to total changes, we also count the frequency of edit types including deletions, insertions, and substitutions and include these as separate measures of editing rather than just a composite. Additionally, we consider different types of euroscepticism. We hypothesize that different types are likely to yield contradictory political pressures for textual editing. A Commission that is popular in a given member state may be less likely to have its recommendations edited. At the

same time, editing may be more common where public support for ‘a better life outside the EU’ is high. The other member states, who vote in the Council, may be more willing to support changes in such cases where there is concern that there is support for the member state exiting the Union.

To preview our results, we find that the total number of edits depend on country size, with large member states editing more than small, which we argue likely stems from their voting power in the council. Counter to previous findings, however, we also find that edits are more common in member states where trust in the European Commission is higher. Yet these are the results based on total textual changes only – when we examine whether political variables help predict the substance of textual changes, which we measure using human coding of edits, we do not find a relationship. In sum, while member states, especially large and EU out member states, continue to edit what the Commission writes about them, we find little evidence suggesting that euroscepticism is the political logic behind the editing.

These findings provide an interesting twist on the debate about the loci of politics in European Union politics. While Zürn (2019) notes that European policy is studied mostly at the national level even though the European level is also of interest (for example Schmidt, 2019), our paper examines how politicization of the recommendation process affected the relationship between the Commission and Council. In the period in the run-up to the GFC, the loci of politics played out in the Council as politically influential member states weakened criticism that the Commission wrote about them. In the new period, we find that this is not the case. But it would be incorrect to argue from our findings that the policy arena itself became less political. Rather the “game” likely shifted to other arenas. That the European Semester created a more technical exercise may have been a contributing factor to greater, and more open, politics over economic policy more generally. It may have contributed to a strategic Commission (Nugent and Rhinard, 2019) and/or to more public fights in the Council (Hobolt and Wratil, 2020). Understanding the consequences of this shift is needed future research.

I. Theory and Mechanisms

Do member states care about what the Commission writes about them? Do edits to the drafts of the Commission’s texts merely represent the fixing of minor errors and fine-tuning, or do they contain evidence of substantively more interesting dynamics such as domestic political pressures and power politics? Previous research found that, in the lead-up to the GFC, the Council weakened evaluations the European Commission had originally prepared for member states with more eurosceptic populations as well as the evaluations for politically influential states (for example Baerg and Hallerberg, 2016; Mariotto, 2019). Given the reforms made under the European Semester, the question is whether these two factors predict when the Council weakens the Commission’s evaluations of member state economic programmes in the post-reform era.

While we are interested in whether the system worked the same in the more recent period, we also seek to build on more recent scholarship on ‘euroscepticism’. The concept of euroscepticism has multiple meanings that are empirically interesting. Foster and Frieden (2017) argue that euroscepticism and one’s national identity are intertwined and reinforcing. The implication is that if an individual has a stronger national identity this may also imply that they are more likely to be eurosceptic. Alternatively, De Vries (2018, p.

6) argues that euroscepticism is fundamentally about attributing success and blame to different levels of government relative to one's own expectations. 'When national conditions are good [...] euroscepticism is mostly likely to develop. When national conditions are bad, however, EU support is the most likely outcome when no viable alternative to membership is present.' Relatedly, scholars have found that the Euro crisis triggered shifting national versus European loyalties, which may have consequences for the changing nature of European identity (Bauer, 2020; Matthijs and Merler, 2020).

In addition to relative support for the EU as a whole, citizens may hold opinions about key EU institutions. Indeed, the Directorate-General for Communication (DG COMM) is the Commission department responsible for explaining EU policies to outside audiences. DG COMM records and reports the share of the population with confidence in EU institutions including the European Parliament, European Central Bank and the European Commission.

A key institution for our purposes is the Commission, which operates in an increasingly politicized environment and responds to political pressure.¹ Williams and Bevan (2019) find that the Commission takes more unilateral action when member states are more eurosceptic. Similarly, Williams (2016) shows that greater aggregate public euroscepticism is associated with the issuance of more reasoned Commission opinions. Van Der Veer and Haverland (2018) find a positive relationship between a member state's level of politicization and the scope of recommendations towards that country by the Commission. These authors also argue that the Commission is responsive to polarization when it drafts Country-Specific Recommendations (CSRs) and that the Commission increases the scope of its recommendations when faced with more polarized public opinions of the EU. In a similar vein, Rauh (2019) indicates that, with increases in politicization, the Commission is interested in making benefits more disperse and that this is particularly true for more salient issues.

II. Causal Mechanisms

In this section, we explain the three mechanisms that we think may act as possible channels linking the relationship between public support for the EU and textual editing of the European Commission's recommendations. We start with a simple assumption. Any member state government would prefer to avoid external criticism of its policies. The press may pick this up and note any perceived problems that the Council's report identifies. Similarly, the political opposition will jump on external criticism of the government.

First imagine the case where there is no political calculus. In this case, voters' demands do not permeate the Council nor do power politics between small and large states (more below) matter. If this is true then the number of textual editing across the draft should reflect minor or idiosyncratic changes. Furthermore, we would not expect edits to vary systematically across member states by political or economic characteristics, nor would we expect minor changes to vary that much over time. Alternatively, if editing varies as a consequence of political and economic variables, then the next question is how are they associated?

¹For an excellent paper on the determinants of trust and the European Commission see Schafheitle *et al.* (2020).

The directional relationship is likely to be complex. On the one hand, greater domestic support for the Commission may lead member state governments to have a greater appetite for participating in EU level decision-making. If ministers indeed embody the symbolic interests of their member states, council edits might reflect engagement (rather than criticism or concern) about what the EU writes about them. On the other hand, member state governments may be hesitant to go against a popular Commission. If the Commission is unpopular, then a member state government may be emboldened to promote changes in the Council that weaken the criticism of that government. Note that Baerg and Hallerberg (2016) find a positive relationship between euroscepticism and edits, which suggests that this latter mechanism may be at work. We therefore test this argument.

H1: The recommendations for countries with less public trust toward the European Commission will be edited more extensively.

Second, the perceived costs (benefits) of EU *membership*, in addition to institutional trust, might matter for member state editing. Notice that this channel is distinct from the perceived penalties of institutional trust discussed above. When considering institutional trust, the cost of belonging to the economic governance system of the EU is paid by the country who is the recipient of the ‘imposition of discretionary and severe sacrifices’. Alternatively, when we move to EU membership and benchmarking, the perceived cost of belonging to the economic governance system of the EU is paid by those individuals who perceive that they would be better outside of the EU (namely the expected ‘dividend’ that they forego by not being able to go it alone). As a result, this cost is likely to be negatively associated with trust. For example, Portuguese paying the sanctions might also think that there is no viable alternative to EU membership and so also feel, from the perspective of benchmarking, that the EU is good. Alternatively, Germans might feel that their national conditions are being dragged down compared to their own expectations of life outside of the EU, and so being inside the EU is a worse outcome despite having low instrumental costs. As a second hypothesis, therefore, we also test the following hypothesis.

H2: The recommendations for countries with less support for being inside the EU are also those that have a higher number of textual edits.

As a final potential mechanism, we consider the relative political power that a member state has in the Council of Ministers, which was tested in earlier research. As argued by Mariotto (2019), despite the agenda-setting role played by the Commission, the Council holds veto power and can amend the Commission’s proposed recommendations. This set-up, in combination with the supermajority rule of the Council, may favour changes to the Commission’s proposal. Member states’ voting rules were reformed in 2011. According to these rules, a member state needed a qualified majority to block what the Commission wanted rather than qualified majority vote to pass it. In practice, it takes a few large states, or a combination of a large and a few small states, to block changes to the Commission’s texts. Large member states are therefore able to make a greater number of edits, on average, as they are both stronger and their partnership is needed by smaller states, which might encourage logrolling. In other words, it is more difficult for small states to get changes to their recommendations passed and, as a result, we might see more editing by larger voting states. Applying this argument, powerful states are more likely to water down the Commission’s

recommendations because of the influence they exert in the Council.² Therefore, and following the same logic as in Baerg and Hallerberg (2016) and Mariotto (2019), we expect that larger countries with greater voting weights are more successful in weakening the Commission's recommendations. A third hypothesis we test as follows.

H3: The recommendations for countries that are larger member states will have a higher number of textual edits than smaller states.

III. Data and Methods

Sample and Time Period

We are interested in whether political factors are related to textual edits during the European Semester period. The particularities of the European Semester are described elsewhere in this special issue (see for example D'Erman *et al.* (this issue)). Empirically, we cover the time period 2011–18. We focus on the 'original' 15 member states of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom, which are the same countries considered by Baerg and Hallerberg (2016). There are only 13 countries in our sample, however, as we are missing Greece because it is a 'programme' member state the entire period and does not submit the same documents. We also drop Luxembourg as it approximates a city state. Another concern is that a few country years have missing data because the EU reports were not available. This is the case for Finland in 2012, and Ireland and Portugal in 2013.

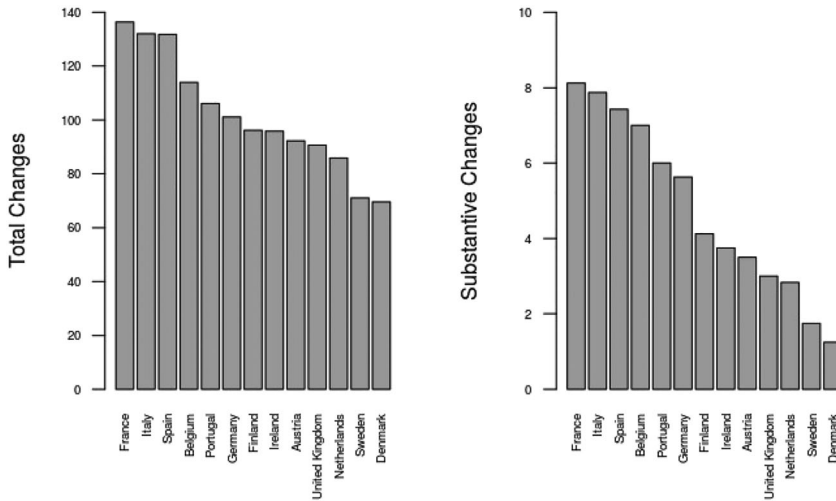
IV. Dependent Variable: Textual Editing

Because we are interested in evaluating changes to the European Commission's recommendations by the Council for each member state, our main dependent variable is the number of *textual edits*. We measure this in five ways: four measures are based on quantitative coding and one measure based on hand coding. First, we calculate the raw number of textual edits using the summary statistics produced by the document compare tool in Adobe Acrobat. We purposely use this methodology, which we think is the most easily reproducible by the largest number of researchers, including researchers without significant experience doing quantitative text analysis. We downloaded and used a free trial version of Adobe Acrobat Pro. We used Adobe Acrobat's compare documents tool to compare the two versions of PDF files available from the European Council website. This tool compares the country specific recommendations reports of the European Commission with the final modified version of the text by ECOFIN. After running the software, the software creates additional summary quantities based on the two input pdf documents. These summaries are our measures and include *total changes*, *replacement* of words, *insertions* of new words, and *deletions* of words. We record these numbers into a dataset and reproduce these steps for all country-years in our sample data.

Second, given that the quantitative measures of textual editing do not take into account context, we create a measure of whether there are substantive changes across the drafts

²Another interpretation for this mechanism are exchange models where asymmetric salience and vote trading play key roles – see for example Arregui *et al.* (2006) and König and Proksch (2006) – and we thank an anonymous reviewer for pointing this out.

Figure 1: Textual Editing by Country: Average Number of Total Changes and Average Number of Substantive Changes



based on human coding of textual differences using a human coder. In order to make this task easier, we first split the documents into paragraphs using the reshape function in the R package *quanteda* (Benoit *et al.*, 2018). Then, using the *Diffir* package, we highlight differences between the Commission and Council documents in each paragraph. The *Diffir* output highlights texts of comparable sections (in this case paragraphs) across the documents. The coder then evaluates each country-year-paragraph and codes whether there were substantive changes observable across the texts. The individual codes the paragraph ‘0’ if there was no substantive changes across the paragraphs and ‘1’ if there was a substantive change across the paragraphs (see Figure 1).

Our method to distinguish substantive changes was to consider non-substantial or superficial changes. Superficial changes are coded for the following traits: changes in Articles mentioned by the Commission and the Council; EU decisions mentioned by the Council and not by the Commission; text added by the Council that does not affect any recommendation or ‘context’ analysis; the same words but grammatical change or verbal tense (that does not change the meaning or urgency of the recommendation); use of acronyms; and finally repetitive and same changes made by the Council to the Commission text in the majority of the Member States.³ Any other changes to the texts are coded as a substantive change.

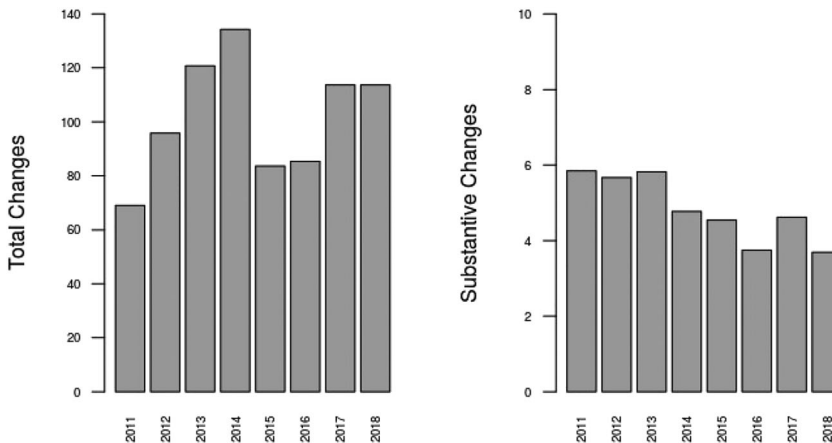
The same person did this coding for all countries and for all years in our dataset. For validation, we took a random sample of substantive changes and both authors assessed whether they agreed with the classification. To make our final annual measure, we then sum the total number of substantive changes from paragraphs into an aggregate country-year total, which reports the total number of substantive changes, *total substantive*, for a given country-year. Table 1 shows examples of coded substantive changes

³For more information, please see the Online Appendix for the complete coding protocol.

Table 1: Examples of Substantive Textual Changes

France 2017	Both texts mention the objective of making a fiscal effort. However, the Commission text mentions explicitly the need to pursue fiscal policy. The Council text only says to ‘pursue substantial fiscal effort’.
Germany 2011	Commission text says ‘remove unjustified restrictions on the craft sector’. The Council text says ‘[...] on certain crafts’.
Germany 2013	The Commission text mentions that ‘The programme plans gross debt to fall to 80.5% of GDP in 2013 and to remain on a downward path thereafter’. The Council text mentions 80 per cent.
Ireland 2014	The Commission text emphasizes that the unemployment rate among young people peaked above 30 per cent in late 2012 and early 2013. the Council says only in mid-2012.
Ireland 2017	The Council specifies to use windfall gains ‘arising from the strong economic and financial conditions’. Not mentioned in the Commission text, which only mentioned ‘proceeds from asset sales’.
Italy 2013	The Council text deleted the highlighted part: ‘Achieve the planned structural primary surpluses in order to put the very high debt-to-GDP ratio (forecast to be 132.2% of GDP in 2014) on a steadily declining path’.

Figure 2: Textual Editing by Year: Average Number of Total Changes and Average Number of Substantive Changes



whereas Figure 2 illustrates the average and substantive metrics of changes across countries in the sample. Importantly, we see similar patterns for both average total changes measured using the automated approach and substantive changes using the human coded approach. France edits the text most while Denmark edits the texts least on average for both metrics.⁴

⁴Figure A1 in the Online Appendix shows the across-year variation in editing over the sample period.

Rather than look only at across-country variation, we also examine variation in textual editing that occurs over time. One question is whether there is a sustained trend, either upwards or downwards, across the sample. In looking across time, we find that our quantitative measure of total textual changes varies significantly more than our human coded measure, which stays relatively flat between four and six edits. The quantitative measure follows an inverted U-shape, going from a low of just over 60 edits in 2011 to a high of almost 140 edits in 2014 and then declining again. As shown in Figure A1 in the Online Appendix, when we decompose the types of changes, we see that most of these textual changes reflect an increase in substitutions in the texts rather than other forms of editing.

V. Independent Variables

Euroscepticism

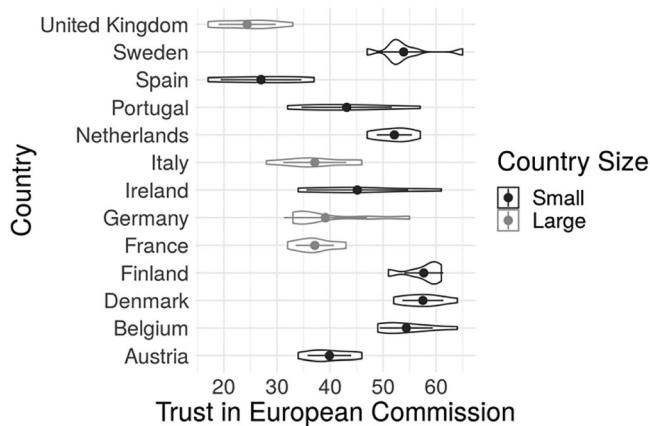
We examine two ways that euroscepticism may matter for textual editing. Our first measure of euroscepticism is trust in the European Commission, *Trust Commission*. This indicator measures confidence among EU citizens in the specific institution responsible for economic governance in the EU. The indicator is based on the Eurobarometer Survey, which is a survey conducted twice a year since 1973, and this specific variable is taken from the Spring version of the survey as it is in line with the time period of the Commission. We also use the Autumn version as a robustness check. The metric that we use is the share of positive opinions (people who declare that they tend to trust) about the Commission. Citizens are asked to express their confidence levels by choosing from the following alternatives: ‘tend to trust’, ‘tend not to trust’ and ‘do not know’ or ‘no answer’.

This variable is an important measure of euroscepticism because it specifically asks about trust that citizens have in the institution responsible for producing the first textual document. Another benefit is that the measures also have good coverage, covering the entire sample and time period. One possible problem with the measure is that because the response is elicited by an interviewer there may be *response biases*, either social desirability bias or acquiescence. A second problem is that some respondents may conflate their opinions about different EU institutions (that is, European Central Bank, European Parliament) despite the fact that they are being asked specifically about the European Commission.⁵

As a second measure of euroscepticism, we also want to examine benchmarking, which we measure with the indicator, *Better Out*. This indicator measures whether respondents agree that they would be better off outside of the EU. This question is also from the Eurobarometer Survey and asks ‘(English version) Please tell me to what extent you agree or disagree with each of the following statements. (OUR COUNTRY) could better face the future outside the EU.’ As above, the measure is aggregated to the national level and we use the average proportion of respondents that say they would be better-off outside of the EU in a given country-year. This

⁵Eurobarometer.

Figure 3: Euroscepticism: Institutional Trust



measure also has good sample and time coverage, with the exception of the year 2011.^{6,7}

As foreshadowed in the theoretical section, euroscepticism towards the Commission and opinions about being better off outside of the EU are not only conceptually different but are also inversely related. The correlation of the measures is $\rho = (-0.61)$, which suggests that the measures are moderately correlated. The fact that there are possibly different channels of euroscepticism that may affect textual editing is exactly what we want to test. In order to account for both pathways, we include both measures of euroscepticism into the statistical models simultaneously.

Large Country

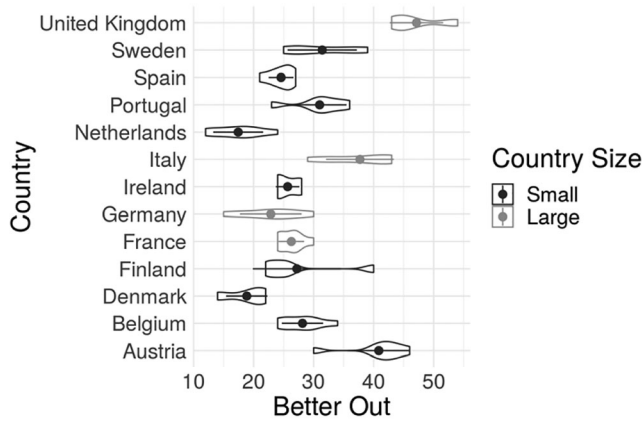
Finally, as per our third hypothesis, we account for whether a country has ‘outsized’ voting power in the Council. We measure voting power based on the vote share that a member state has on the Council before the Nice Treaty in 2011. If a country had ten votes on the Council prior to the Nice Treaty (which is equivalent to having 29 votes after the Nice Treaty) we classify this country as ‘large’ and code it as ‘1’. Alternatively, if a country has fewer than ten votes on the Council, we classify the country as ‘small’ and code it as ‘0’. The countries that are large according to this measure are France, Germany, Italy, and the United Kingdom (UK).

Figures 3 and 4 illustrates the distributions of the levels of euroscepticism by measure and by country size. As we can see, levels of trust in the European Commission ranges from just over 20 per cent to over 60 per cent across the sample. Large countries (those

⁶We use the descriptive data that the Commission reports. On weighting countries, the report notes that the sampling points were drawn systematically from each of the administrative regional units, after stratification by individual unit and type of area.

⁷As a robustness check, we complete the series using random imputation using the R package *imputeTS* (Moritz and Bartz-Beielstein, 2017). We use univariate rather than panel data estimation for the missing data. The only data missing in our panel is for 2011. While imputation in general is a well-known problem and widely covered by routines such as *Amelia* (Honaker and King, 2010), in our case, the missingness is actually a univariate time series rather than panel missing data.

Figure 4: Euroscepticism: Benchmarking



with large voting power in the Council) are also those countries with lower levels of trust, on average. We observe much larger sample variation in the *Better Out* measure than we do for the measure of *Trust Commission*. The UK is unsurprisingly the most eurosceptic member state according to both measures; Spain is simultaneously not so trusting yet also perceives life outside of the EU as being worse. This points to the above argument that the perceptions of the imposition of costs may not translate into perceptions that life outside of the EU is perceived as better by citizens. In contrast to Spain, Italy has an average amount of trust in the Commission but also scores relatively high on perceptions that life would be better outside of the EU.

VI. Control Variables

Euro out

Countries outside of the Euro Area have a different relationship to EU reporting standards. To account for this, we consider whether the country is an ‘in group’ or ‘out group’ member. We code Denmark, Sweden and the United Kingdom as ‘outs’ and the other countries coded as ‘ins’. We observe significant variation in public opinions towards the EU in these countries. As shown in Figures 3 and 4, Denmark and Sweden score relatively high on both metrics for their support for the EU. Unsurprisingly again, the UK is the least supportive of the Commission and also the most likely to feel that their interests are best served outside of the EU.

Economic Variables

In addition to the political calculus that governments make when it comes to editing their country reports, we test whether a country’s economic fundamentals also matter. To account for the economic conditions of member states, we include variables for unemployment, per capita GDP, and Gross Government Debts (as a percentage of GDP). All of the economic data downloaded is from EUROSTAT and we use the EUROSTAT data as it is

Figure 5: Trust: Total Changes

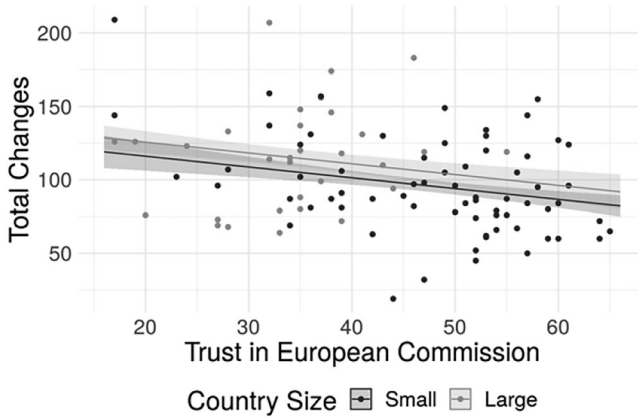
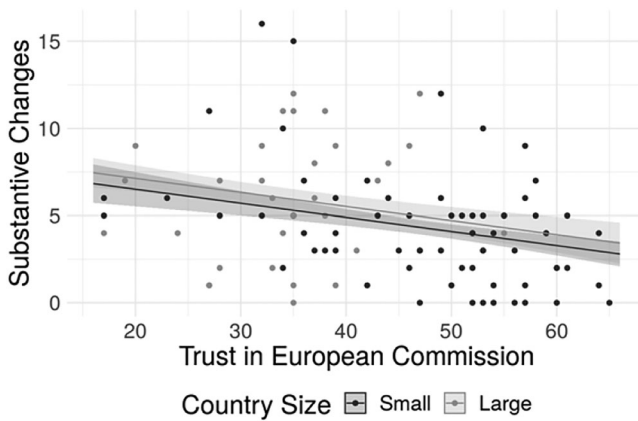


Figure 6: Trust: Substantive Changes



also the data used by the European Commission in making their assessments of the member states.⁸

VII. Analysis and Results

Before we present the statistical analysis, we first want to show graphically the relationships among textual editing, euroscepticism, and country size, which are our main political variables of interest. Figure 5 examines the relationship between trust in the Commission and the total number of textual changes made to the Commission’s recommendations. Here we see a positive relationship for both large and small member states. What is also interesting is that we can see once again that the large member states are also those that have, on average, the lowest levels of trust in the Commission. Alternatively,

⁸European Commission.

when we examine the relationship between trust and substantive edits as we do in Figure 6, we see a more nuanced picture – as trust in the Commission increases, substantial edits to the Commissions’ reports go down. With the variable *Better Out?* we find no obvious relationship between public opinions about being outside of the EU and the amount of changes. Looking at the data, Figure 7 again shows some clustering such that small countries are less likely to report being better outside of the EU when compared to large countries. Figure 8 shows little to no relationship between substantial edits and public perceptions of being better outside of the EU, though one does see clear indication that larger countries make a greater number of edits than smaller countries. These graphical relationships are based on predictions that do not account for other variables that might matter, such as the economic conditions of the member states. Therefore, in the next sec-

Figure 7: Better out of EU: Total Changes

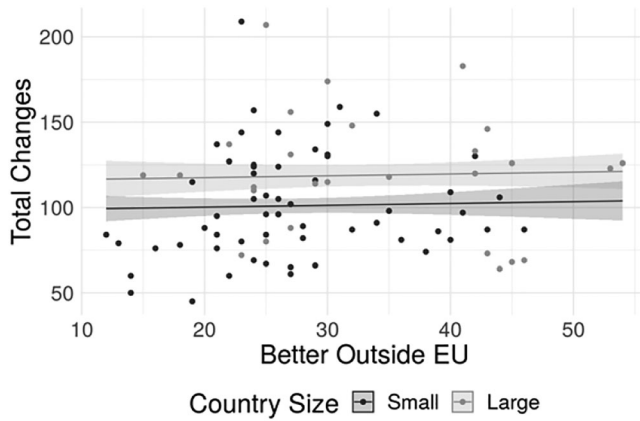


Figure 8: Better out of EU: Substantive Changes

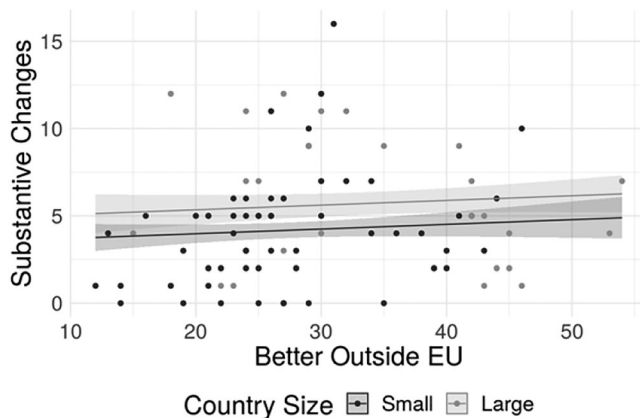


Table 2: Estimating Predictors of Textual Editing

	Dependent variable				
	Total Edits	Substantive Edits	Replacements	Insertions	Deletions
	(1)	(2)	(3)	(4)	(5)
Trust	0.756 (0.517)	0.051 (0.057)	0.164 (0.415)	0.337*** (0.089)	0.258*** (0.085)
Better Out EU	0.113 (0.449)	0.002 (0.051)	-0.023 (0.360)	0.083 (0.077)	0.047 (0.074)
Growth	-0.100 (1.090)	-0.082 (0.113)	-0.358 (0.883)	-0.226 (0.191)	0.429*** (0.182)
Income (logged)	4.671 (21.603)	-1.883 (2.658)	9.884 (17.134)	2.033 (3.649)	-6.300* (3.478)
Debt	0.349 (0.223)	0.039 (0.027)	0.241 (0.177)	0.111*** (0.038)	0.004 (0.036)
Unemployment	4.222** - (1.892)	0.155 (0.221)	3.037 - (1.509)	0.860*** (0.323)	0.431 (0.307)
Large	20.366** (10.122)	0.941 (1.246)	12.264 (8.020)	4.426*** (1.706)	3.795*** (1.626)
Out	-25.437** (12.649)	-1.903 (1.567)	-18.253* (10.020)	-2.443 (2.132)	-4.597** (2.032)
Constant	-33.534 (241.360)	17,997 (29.790)	-71.945 (191.284)	-37.725 (40.709)	65.177* (38.800)
Observations	87	87	87	87	87
R ²	0.352	0.255	0.333	0.405	0.298
Adjusted R ²	0.285	0.178	0.265	0.344	0.226
F-Statistic	42.282***	26.679***	38.987***	53.032***	33.077***

Note: Random effects linear panel model. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

tion, we move to more formal statistical tests in order to evaluate our hypotheses while controlling for member states' economic conditions.

The main statistical model that we present is a linear, random effects model. The unit of analysis is country-year and we have 13 out of the original 15 member states (no Greece and no Luxembourg) for the years 2011 to 2018 for a total of 100 observations. We use random rather than fixed effects for a number of reasons including small sample size, slow moving variables, and time invariant predictors (both Large and Out variable are time-invariant) (Plümper and Troeger, 2011; Clark and Linzer, 2015) though we include models with fixed effects (thus dropping the large and euro-out variables) in the robustness section and in models in the Online Appendix. The results do not differ extensively.

Table 2 presents the model results. Model (1) shows that there is a positive but statistically insignificant association between trust in the European Commission and total number of edits. This finding is the opposite to what is expected by our first hypothesis (H1). Next, we examine the decomposition of the total editing into their various parts (Models 3–5). Here we see that the positive relationship is observed no matter the type of textual changes across the texts. An increase in trust in the Commission is most strongly associated with textual replacements, followed by insertions, and finally deletions, with both insertions and deletions statistically significant at traditional cut-off points.

When we move to measuring substantive editing, however, (Model 2), we see neither a large nor statistically significant relationship between trust in the Commission and textual editing. While the coefficient is negative (as expected by H1), the coefficient is almost zero. So as to check that this result is not a function of the type of statistical model, we also run a Poisson model to account for the count data, however the results are similar with the exception that Debt and Out countries reach standard levels of significance.

So as to check that this result is not a function of the type of statistical model, we also run a Poisson model to account for the count data, however the results are similar with the exception that Debt and Out countries reach standard levels of significance.

Taken together, these results provide little evidence for our hypotheses H1 or H2: we do not observe that domestic trust in the European Commission is associated with a lesser number of quantitative or substantive textual edits. Indeed, if anything, we find some (albeit weak) evidence that member states more trusting of the Commission may be more likely to make quantitative edits across drafts.⁹

Concerning the other possible channel of euroscepticism, *Better Out*, this variable is also neither substantively nor statistically associated with total textual editing nor any of the decomposed subcomponents. While there is a slight negative relationship reported across some of the models (and which can be seen in Table 2), again we find very little support for our hypothesis (H2) using this measure. Euroscepticism vis-à-vis benchmarking is not statistically associated with the number or type of textual edits of country recommendations across the drafts of texts.

Next, we examine our final hypothesis (H3), and test whether being a *Large Country* matters. Here we do find a strong positive relationship between those countries that are

⁹We also check whether the results differ when we use a human coded variable of substantial changes that measures de-emphasis of the Commission's recommendations. We find no significant differences with the exception that the trust variable is positive and significant at traditional levels. See the Online Appendix and Tables A2 and A3 for details.

large (in terms of voting power) and the number of textual edits (Models 1, 3–5). Large countries edit the Commission's recommendations significantly more than small countries with large countries making an estimated 20 more total edits. When it comes to the decomposition of the edits, we find that large countries make a greater number of textual replacements. Textual insertions and textual deletions, while positive and statistically significant, are associated more weakly than these other forms of editing. Interestingly, when we shift to examining qualitative rather than quantitative measure of edits (Model 2), we again find again null results.

We next examine the effects of our economic control variables on textual editing. Here we find that editing is somewhat sensitive to member states' economic conditions. In particular, we find that debt levels and unemployment seem to be positively associated with textual editing but again find little evidence that it is associated with substantive editing.¹⁰ Also as before, once we decompose the total edits into their different component parts, we see that textual replacements are again the main driver of the results.

Finally, the *Out* variable, which represents an indicator for those countries outside of the Euro area, edit much less than 'in' countries on average. Indeed, 'out' countries make approximately 25 fewer edits than countries inside of the Euro, which is substantively almost as strong as the effects of the Large country indicator. Also as before, most of these edits are replacements rather than other types of edits. Out countries are also predicted to reduce their substantial edits by two, which is substantially very large, though this variable is not significant at standard thresholds. In summary, we find little evidence that euroscepticism predicts the level of textual changes between drafts of the Commission's text and the Council's changes. We especially find no evidence that euroscepticism matters for predicting substantive changes to the texts despite expectations that different types of euroscepticism may work on editing through different channels. Given that previous research did find that euroscepticism mattered for editing, our results are suggestive that the European Semester reforms weakened the impact of a previous political logic (through euroscepticism) on member states' recommendations. As mentioned in the theoretical section, scholars have found that euroscepticism is increasing alongside nationalism. One question, therefore, relating to the COVID-19 crisis is whether the crisis will increase nationalism and in turn, euroscepticism, and whether (or to what extent) this channel will stay closed.

On the other hand, inequalities remain as a consequence of member states' status with the EU: Countries with larger voting power continue to edit more and Euro area 'out' countries continue to edit less and substantially so. A new finding is that textual replacements are the kind of textual editing most strongly related to country status, which we think is an area of important future research.

VIII. Alternative Mechanisms and Robustness Checks

In this section, we indicate that our results are not sensitive to a variety of alternative model specifications and omitted variables. As we mention above, we take a pluralistic view of politics and do not account for inter-elite behaviour. To check whether elites

¹⁰In running a Poisson count model, we find the income and debt variables reach traditional levels of statistical significance as reported above.

may matter, we examine the role of the finance minister and code whether the finance minister (FM) is from a eurosceptic party. The logic here is that a eurosceptic FM would be more likely to push for edits than a europhile FM. To code whether or not the FM is eurosceptic, we examine the European party family to which that the FM belongs. We code FMs that belong to the European Conservatives and Reformists in the EU Parliament as Eurosceptic or '1' and code FMs that belong to: the European People's Party, Progressive Alliance of Socialists and Democrats, Renew Europe Group, Alliance of Liberals and Democrats for Europe, or Independents in the EU Parliament as being Pro-Europe or '0'. Including this variable into the model, we find a negative but insignificant relationship between eurosceptic FMs and textual editing.

A second concern relates to our variable 'large country', which we argue proxies political power on the Council. Of course, voting power in the EU is also associated with population size and it may be the case that it is the size of the eurosceptic domestic population rather than the member state's political power that matters. For example, it may be that larger countries have more complex economies and governance systems and therefore receive more comprehensive recommendations, which makes them edit more. Alternatively, they may even just have a greater capacity for editing. To account for this, we include population directly into the model and interact this variable with 'Trust in the Commission'. This effectively up-weights citizens living in large countries and down-weights citizens living in smaller countries. Interestingly, we find that population size exerts no independent effects on textual editing nor do we find that the interaction between trust and population size is particularly important.

Finally, a concern might be the use of random effects (see above). To run a fixed effects model, we remove the country-invariant 'large country' and 'out country' variables and run the models again this time using a fixed effects specification. As in our previous results, we see that 'Trust in the Commission' is positively associated with quantitative changes to the text but not associated with qualitative changes. The variable 'Better Out of EU' is also not associated with textual changes. When we look at the economic variables that were sensitive previously, we see that unemployment and debt are also no longer strong predictors of textual changes.

We also run additional robustness tests including using the data on Trust from the October wave of the survey, imputing missing values, and using a count model for the substantive edits. These and other checks such as tests for multicollinearity are available in the Online Appendix in Tables A1–A4.

In summary, we find little evidence of a systematic relationship between euroscepticism and textual editing. There is perhaps a small and statistically significant relationship between Trusting the Commission and inserting new textual content into the Commission's report, but we find little evidence that euroscepticism matters during our sample time-period. Including population also does little to change our results. Indeed, it has little substantive effect, which suggests that the 'Large Country' variable is likely picking up power in the council rather than country size.

As these country specific features of the data (Voting Size, Euro out) seem to have the most explanatory power, and following Bokhorst's (this issue) focus on specific cases under the European Semester, in the next section, we examine more carefully variation in the context of textual changes (see Table 3) by looking at Austria and Italy more closely. In addition to sharing a geographical border, these countries score similarly on

Table 3: Robustness Checks: Predictors of Textual Editing

	<i>Dependent variable:</i>				
	<i>Total Edits</i>	<i>Substantive Edits</i>	<i>Replacements</i>	<i>Insertions</i>	<i>Deletions</i>
	(1)	(2)	(3)	(4)**	(5)
Trust	0.272 (1.202)	0.013 (0.117)	-0.215 (0.992)	0.412** (0.198)	0.075 (0.196)
Better Out EU	0.344 (0.782)	0.004 (0.076)	0.070 (0.645)	0.121 (0.129)	0.153 (0.127)
Growth	0.115 (1.334)	-0.011 (0.130)	0.169 (1.101)	-0.363 (0.220)	0.308 (0.217)
Income (logged)	27.986 (169.655)	-27.400 (16.555)	52.563 (140.048)	-39.966 (27.921)	15.390 (27.635)
Debt	0.034 (0.850)	-0.180** (0.083)	0.069 (0.702)	-0.158 (0.140)	0.123 (0.138)
Unemployment	-0.729 (5.089)	-0.148 (0.497)	1.765 (4.201)	-1.373 (0.837)	-1.121 (0.829)
Euroscaptic FM	-7.845 (21.218)	-1.236 (2.070)	-4.719 (17.515)	-2.556 (3.492)	-0.570 (3.456)
Population	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Trust* Population	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Observations	87	87	87	87	87
R^2	0.062	0.107	0.090	0.284	0.288
Adjusted R^2	-0.241	-0.182	-0.204	0.053	0.059
F Statistic (df = 9; 65)	0.480	0.862	0.716	2.867***	2.928***

Note: Fixed effects linear panel model. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

eurocepticism. Despite this, Italy is a large voting member state whereas Austria is not. Italy's economic situation was significantly weaker over the time period compared to Austria as well. Capitalizing on this variation, we examine what types of substantive coding the countries engage in and see whether this gives us a clue into patterns of politicization (or not) under the European Semester.

Austria and Italy

We focus on the types of substantive edits that are made by examining two countries: Italy and Austria. They tend to vary on the amount of editing they do, with Italy editing more than Austria, despite being similar in terms of their level of eurocepticism. Another interesting similarity is that political parties in these member states are shown to combine elements of both soft and hard eurocepticism at the same time (Heinisch *et al.*, 2020). Furthermore, while their eurocepticism does not vary widely, their voting weights on the Council does vary, with the Italy being "large" and Austria "small" in our dataset. We try to exploit this variation by examining more closely any obvious places of difference in their substantive edits.

Returning to our variable, substantive edits, we go deeper to consider whether a substantive change involves textual 'de-emphasis', 'emphasis', or 're-wording'. De-emphasis includes those cases where the text of the Commission is more explicit than the resulting text after Council editing. For example, in Austria 2014, the Commission says, 'Social Security contributions and payroll taxes amount to almost 50% of gross wages'. In the Council text, the document refers to 'The tax wedge amounts to almost 50% of labour costs', which de-emphasizes expected labour market concerns. Similarly in Italy, 2014 the Commission text mentions explicitly that there is a need for more direct 'taxes on diesel and petrol'. On the contrary, in the Council text, the Council is more general and states that the country should 'ensure more effective environmental taxation'.

In contrast to de-emphasis, there is also emphasis. Emphasis includes those cases where the text of the Commission is *less* explicit than the resulting text after Council editing. For example, in Austria, the Council emphasizes that the educational outcomes have to improve 'disadvantaged young people including those with a migrant background' whereas the Commission is not that specific (does not mention the word disadvantaged) (Austria, 2014). If we look to Italy, we also sometimes observing additional emphasis. The Council actually even increases the rate of unemployment reported by the Commission, from 10.9 per cent to 11.2 per cent, which in effect, makes the Commission's recommendations for Italy even more stringent (Italy, 2018).

The final category that we examine is 're-wording'. Re-wording refers to observed changes to language such that the meaning changes slightly. For example, in Austria, the Council uses the word 'agreed' whereas the Commission uses the word 'decided' (Austria, 2012). Similarly, the Commission says 'Austria's public expenditure on pensions is relatively high compared to the rest of Europe', while the Council refers to 'other Member States'. (Austria, 2017) These examples are more than superficial changes as they change slightly the meaning of the text; however, they are ambiguous in terms of whether they make the recommendations harder or easier.

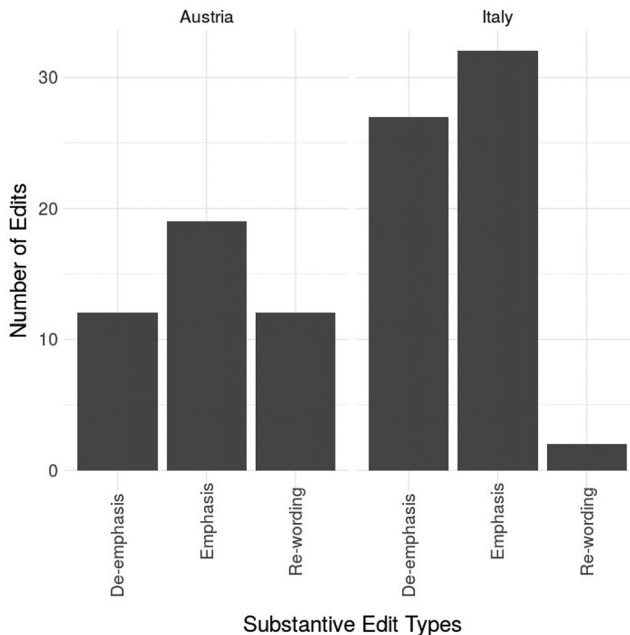
Figure 9 shows that while both Austria and Italy both engage in de-emphasis and emphasis, a much larger share of Austria's substantive editing is re-wording when compared to Italy.

Also interesting is that both countries engage in both emphasizing and de-emphasizing the Commission's text and they do so in relatively similar proportions, with Italy having a slightly larger share of de-emphasizing edits. In aggregate, Italy edits more than Austria under all three categories. As Italy is both a large country but also has a weaker economy, it is hard to say for certain which variable is driving these results. However, examining the documents using 'key word in context' analysis of the Council and Commission's text, we find that Italy emphasises more about debt, labour market reforms, unemployment, and growth risks than Austria. While we are cautious in terms of making too general claims, we find that at least within these two countries, economic rather than political concerns are the dominating explanation for textual editing.

Conclusion

This article explores both political and economic explanations for the editing of the European Commission's evaluation of member state's economic policies in the Council of Ministers. Previous research finds that member states were able to undermine the 'watch-dog' function of the European Commission in the Council of Ministers (Baerg and Hallerberg, 2016). We re-examine the role of political variables in Council editing during the period of the European Semester. New to this paper, we also decompose

Figure 9: Type of Substantive Editing Austria and Italy



euro-scepticism into different possible ways that member states can engage in textual editing. While we find that Large and Euro Out member states still exhibit different behaviours when it comes to the total number of edits of the Commission's recommendations, we do not find any associations with substantive changes.

This is a notable development from previous work that examined the functioning of the system in the pre-European Semester period. While much of the literature notes an increase in politicization over time, which would imply continued substantive edits according to state size and how euro-sceptic a population is, we find little evidence of this. As Koop *et al.* (2020) persuasively argue, member states choose what arenas to politicize. Our paper suggests that member states were successful in depoliticizing this particular pathway through the reforms of the European Semester.

This does not mean, however, that the economic governance as a whole is depoliticized. Other papers in this special issue examine how politicization affects the implementation and enforcement of supranational economic policy (for example Van der Veer, this issue). We focus on one particular nexus of the relationship between the Commission and Council that played an important role in the pre-European Semester period. Other authors argue that the European Commission has become more political over this time period (Nugent and Rhinard, 2019). The European Commission itself may have become more strategic in the types of recommendations that it made to member states in this period.

Nevertheless, this article suggests a change in practice in how one part of the process operated. To continue with our analogy, the 'watchdog' barks as before. But there is no evidence that it barks only at small, friendly dogs. In the time period 2011–18, and for 'dogs' from the EU-15, the 'watchdog' barked somewhat indiscriminately.

Correspondence:

Mark Hallerberg
Hertie School
Friedrichstr. 180, 10117 Berlin, Germany
email: hallerberg@hertie-school.org

References

- Arregui, J., Stokman, F.N. and Thomson, R. (2006) 'Compromise, Exchange and Challenge in the European Union'. In Thomson, R., Stokman, F.N., Achten, C.H. and König, T. (eds) *The European Union Decides* (Cambridge: Cambridge University Press), pp. 124–52.
- Baerg, N. and Hallerberg, M. (2016) 'Explaining Instability in the Stability and Growth Pact: The Contribution of Member State Power and Euro-scepticism to the Euro Crisis'. *Comparative Political Studies*, Vol. 49, No. 7, pp. 968–1009.
- Bauer, S. (2020) *Citizens' Support for the European Union* (New York: Springer).
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S. and Matsuo, A. (2018) 'quanteda: An r Package for the Quantitative Analysis of Textual Data'. *Journal of Open Source Software*, Vol. 3, No. 30, p. 774.
- Bokhorst, D. (this issue) 'The Influence of the European Semester: Case Study Analysis and Lessons for its Post-Pandemic Transformation'.
- Clark, T. and Linzer, D. (2015) 'Should I use Fixed or Random Effects'. *Political Science Research and Methods*, Vol. 3, No. 2, pp. 399–408.

- De Vries, C.B. (2018) *Euro-scepticism and the Future of European Integration* (Oxford: Oxford University Press).
- D'Erman, V., Schulz, D. F., Verdun, A. and Zagermann, D. (this issue) 'The European Semester in the North and in the South: Domestic Politics and the Salience of the EU-Induced Wage Reform in Different Growth Models'.
- Foster, C. and Frieden, J. (2017) 'Crisis of Trust: Socio-economic Determinants of Europeans' Confidence in Government'. *European Union Politics*, Vol. 18, No. 4, pp. 511–35.
- Franchino, F. and Mariotto, C. (2020) 'Politicisation and Economic Governance Design'. *Journal of European Public Policy*, Vol. 27, No. 3, pp. 460–80.
- Heinisch, R., McDonnell, D. and Werner, A. (2020) 'Equivocal Euro-scepticism: How Populist Radical Right Parties Can Have Their EU Cake and Eat It'. *Journal of Common Market Studies*, Vol. 59, No. 2, pp. 189–205.
- Hobolt, S.B. and Wratil, C. (2020) 'Contestation and Responsiveness in EU Council Deliberations'. *Journal of European Public Policy*, Vol. 27, No. 3, pp. 362–81.
- Honaker, J. and King, G. (2010) 'What to do about Missing Values in Time-Series Cross-section Data'. *American Journal of Political Science*, Vol. 54, No. 2, pp. 561–81.
- König, T. and Proksch, S.-O. (2006) 'A Procedural Exchange Model of EU Legislative Politics'. In Thomson, R., Stokman, F.N., Achten, C.H. and König, T. (eds) *The European Union Decides* (Cambridge: Cambridge University Press), pp. 211–38.
- Koop, C., Bressanelli, E. and Reh, C. (2020) 'Actors under Pressure: Politicisation and Depoliticisation as Strategic Responses'. *Journal of European Public Policy*, Vol. 27, No. 3, pp. 329–41.
- Mariotto, C. (2019) 'Negotiating Implementation of EU Fiscal Governance'. *Journal of European Integration*, Vol. 4, No. 4, pp. 465–86.
- Matthijs, M. and Merler, S. (2020) 'Mind the Gap: Southern Exit, Northern Voice and Changing Loyalties since the Euro Crisis'. *Journal of Common Market Studies*, Vol. 58, No. 1, pp. 96–115.
- Moritz, S. and Bartz-Beielstein, T. (2017) 'imputeTS: Time Series Missing Value Imputation in R'. *R Journal*, Vol. 9, No. 1, pp. 207–18.
- Nugent, N. and Rhinard, M. (2019) 'The "Political" Roles of the European Commission'. *Journal of European Integration*, Vol. 41, No. 2, pp. 203–20.
- Plümpert, T. and Troeger, V.E. (2011) 'Fixed-Effects Vector Decomposition: Properties, Reliability, and Instruments'. *Political Analysis*, Vol. 19, No. 2, pp. 147–64.
- Rauh, C. (2019) 'EU Politicization and Policy Initiatives of the European Commission: The Case of Consumer Policy'. *Journal of European Public Policy*, Vol. 26, No. 3, pp. 344–65.
- Schmidt, V.A. (2019) 'Politicization in the EU: Between National Politics and EU Political Dynamics'. *Journal of European Public Policy*, Vol. 26, No. 7, pp. 1018–36.
- Schafheitle, S., Weibel, A., Meidert, M. and Leuffen, D. (2020) 'The Road to Trust. A Vignette Study on the Determinants of Citizens' Trust in the European Commission'. *Journal of Common Market Studies*, Vol. 58, No. 2, pp. 256–75.
- Van Der Veer, R.A. and Haverland, M. (2018) 'Bread and Butter or Bread and Circuses? Politicisation and the European Commission in the European Semester'. *European Union Politics*, Vol. 19, No. 3, pp. 524–45.
- Williams, C.J. (2016) 'Issuing reasoned opinions: The Effect of Public Attitudes towards the European Union on the Usage of the "Early Warning System"'. *European Union Politics*, Vol. 17, No. 3, pp. 504–21.
- Williams, C.J. and Bevan, S. (2019) 'The Effect of Public Attitudes Toward the European Union on European Commission Policy Activity'. *European Union Politics*, Vol. 20, No. 4, pp. 608–628.

Zürn, M. (2019) 'Politicization Compared: At Global, European, and national Levels'. *Journal of European Public Policy*, Vol. 26, No. 7, pp. 977–95.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1. Supporting information.