

Explaining Instability in the Stability and Growth Pact: The Contribution of Member State Power and Euroskepticism to the Euro Crisis

Nicole Rae Baerg
University of Mannheim
nicole.baerg@uni-mannheim.de*

&

Mark Hallerberg
Hertie School of Governance
hallerberg@hertie-school.org

Abstract

The Stability and Growth Pact clearly failed to prevent the euro crisis. We contend that the failure was due largely to the ability of the Member States to undermine the Pact's operation. The European Commission served as a "watchdog" to monitor fiscal performance. The Member States themselves, however, collectively had the ability to change the content of the reports for individual states. We confirm the expectation that powerful Member States had the most success in undermining the role of the Commission. Perhaps more surprisingly, we find supporting evidence for our argument that governments with euroskeptic populations behind them were also more successful in weakening the Commission's warnings. These results have broader theoretical implications concerning which mechanisms explain country-specific outcomes under a shared rule. Another contribution is the creation a new dataset of European Commission assessments of Member State economic programmes and Council of Minister revisions.

*Previous versions of this paper were presented at the "The Political Economy of the Euro Crisis" workshop in Zurich, March 2014 and the "Crisis, Institutions, and Banking Union" Workshop at the Federal Ministry of Finance in Berlin, June 2014. The authors would like to thank William Roberts Clark, Mark Copelovitch, Jeffrey Frieden, Mareike Kleine, Will Lowe, Robert Thomson, Stefanie Walter, and three anonymous reviewers for helpful comments. Research assistance was generously provided by Andrei Ilas, Kyle Ott, Alyssa Taylor, and Grzegorz Wolszczak. Replication materials are available at our [Dataverse Website](#).

Introduction

The Maastricht Treaty set guidelines for what macro-economic goals European Union Member States would need to reach to qualify for participation in the coming common currency, with deficit and debt targets the most visible. Initially, the expectation was that only a few countries would be able to participate. When it became clear that there would be an expanded Eurozone, which would include not only core countries like Germany and France but more peripheral countries like Spain and even Greece that had traditionally run large budget deficits, there was growing concern that European-level rules were needed. Otherwise some governments would be tempted to run fiscal policies that could become unsustainable and leave the other Member States with the unpalatable choice of either bailing out the troubled government or letting it suffer from the consequences of a default. The Member State governments therefore agreed to a Stability and Growth Pact meant to avoid this outcome. Clearly, the need for financial packages for multiple Member States during the euro crisis means the Pact failed. This paper explores why the Pact failed, and it focuses especially on its institutional design.

Ongoing research examines how institutional design is related to compliance and the enforcement of rules and we build on this literature. Scholars have shown that there are significant benefits to flexibility in rules. The argument is that, under conditions of uncertainty, countries have a hard time committing to rules. A lack of flexibility constrains countries even further and, as a result, countries only agree to shallow commitments. Alternatively, rule flexibility allows countries to make deeper commitments, which addresses their commitment problem (Johns, 2013; Kucik and Reinhardt, 2008; Mansfield, Milner and Rosendorff, 2002).¹ In the European context, scholars have also indicated that domestic institutions can both help and hinder compliance with EU level rules (Heipertz and Verdun, 2010; Koehler and König, 2015; König and Mäder, 2013; Richard and Caraway, 2014). While rule flexibility may allow states to make deeper commitments, Member

¹A similar logic follows with escape clauses which allows for temporary deviations thus ensuring greater flexibility (see Rosendorff and Milner (2001)).

States also vary in the extent to which they can exploit flexibility. In other words, the degree of rule flexibility may not apply equally across all Member States. This article reports the first empirical evidence that both Member State voting power and domestic level euroskepticism is systematically associated with weakening (rather than breaking) EU level rules.

When are some governments successful at weakening the application of the rules they have agreed to at the international level? This paper documents the selective weakening of the assessments for Member States under the Stability and Growth Pact (SGP) both prior to, and during, the euro crisis. While “growth” is part of the title of the SGP, its primary goal is to keep deficits low to minimize the likelihood that a fiscal problem in one Member State affects another. The weakening of the assessments under the Pact meant that the main tool of the European level to prevent larger deficits and the buildup of debt levels in Member States was defanged. The ultimate causes to the crises are of course many and complex, and they include a number of factors highlighted by others in this volume such as international and external costs, policy announcements, and concerns about migration (see [Bernhard and Leblang](#); [Genovese, Wassmann and Schneider](#) and [Walter](#) in this volume).

Yet, there are three ways that this weakening contributed directly to the crisis. First, it meant that several large and/or euroskeptic Member States probably entered the crisis with weaker public finances than had they followed the original Commission recommendations. Second, given that we find that even small, euro-friendly states had the Commission’s text weakened when the big and/or euroskeptic states were receiving milder EU-level surveillance too, countries like Greece and Ireland that later received aid packages, also received weaker surveillance while other countries weakened and edited the Commission’s recommendations. Finally, one can also speculate about the indirect effects on Member State behavior. Those Member States that did not receive aid package had to pay for them. The fact that large states and especially those with euroskeptic populations entered the crisis with weaker finances meant that they were less likely and, possibly, less able to agree to debt restructurings because of the consequences for their own public finances.

In order to measure and explain the weakening of the assessment criteria, we examine the “preventive” arm of the SGP, which requires every European Union Member State to submit annual macro-economic programmes to a supranational body, the European Commission, for evaluation. Those programmes include information on deficits, debts, and the economic growth trajectories of Member States. The Commission, in turn, writes a report on each Member State that constitutes draft text for the intergovernmental Council of Economic and Finance Ministers (ECOFIN, or Council) for a formal statement to the Member States. The key benchmark of the SGP is that Member States should have a budget deficit of three percent or less. In this paper, we refer to the Commission’s power to comment on national programmes and make concrete recommendations as its “watchdog” function. The Commission may also recommend “corrective” measures. The Council, based on an initial recommendation from the Commission, may deem a Member State with deficits above this figure an “excessive deficit” country. If a Member State ignores the Council’s ruling, the Council, again with initial Commission recommendation, can decide to fine a non-complying country. In practice, in the lifetime of the Pact to date (1998 through 2015) the Commission has not recommended fines and the Council has not imposed them. An important reason why there have been no fines, we argue, is that Member States succeed in blocking critical text against them from the Commission early in the process. The “preventive” arm sets up the “corrective” arm; no action against a Member State can begin unless the Council identifies clear issues first.²

As we show below, some Member States systematically weaken the Commission’s assessment whereas others do not. We document large variation in the Council’s editing of Commission evaluations of Member States across countries and over time. The question that this paper addresses

²In terms of actual performance, the Stability and Growth Pact did not succeed in keeping budgets “close to balance or in surplus” since its introduction in 1999. Two large founding members of the European Community, France and Germany, were especially visible in their violation of the Pact in 2002 and 2003. The Council repeatedly ignored the Commission’s advice, and the Commission took the Council to the European Court of Justice. The outcome of the court case did not materially affect the overall institutional setup—it still left the Member States themselves, voting through the Council by qualified majority, to decide whether to impose the Commission’s recommendations.

is therefore, what types of countries are successful in weakening the assessments of them? Borrowing insights from the literature on judicial politics and literature on partisanship and elections, we examine whether or not Member State power and euroskepticism helps to explain why some countries weaken the assessments while other countries do not. We further examine whether or not the partisan hue of the government and the election cycle helps to explain weakening and editing. In this paper, we find that Member States with certain characteristics, by weakening the original Commission assessments, eroded the power of the European Commission in its role as a “watchdog.” Thus, in the case of the SGP, weakening the Commission assessments lowered Member States’ commitments to the EU and made it easier for powerful and euroskeptic Member States to pursue uncooperative strategies (see also [Chapman et al. \(2015\)](#); [Copelovitch \(2010\)](#) who find similar results in other domains).

To measure the extent of changes to the Commission’s assessment of Member States and to explain the variation in those changes, this paper presents a new dataset constructed from human coding and computer textual analysis of existing Commission reports on Member State Stability and Convergence Programmes and the changes the Council of Ministers made to them. Using this new data, we examine whether Member State power, euroskepticism, and elections help explain whether the Council changed the Commission’s recommendations. As a preview of our results, we find that countries with the largest voting power on the Council (10 votes prior to the Nice Treaty, 29 votes after) are more likely to receive favorable edits to the Commission’s original text. The same is true for countries with relatively euroskeptic domestic populations—they are more likely to have the Council change the Commission’s original assessment of it. Finally, countries are more likely to receive weakened recommendations and there is more editing of them when other Member States are receiving the same treatment. We conclude from our analysis that powerful Member States and Member States with euroskeptic domestic populations were able to receive the necessary number of votes in the Council to undermine the watchdog function of the European Commission, which undermined institutional oversight mechanisms that may have lowered the

debt levels of several Member States especially during “good times” and left them less exposed once the crisis was underway. This aspect of our theory leads to an interesting implication we explore further in the conclusion—Member State governments may have incentives to encourage more euroskepticism among their populations so that they have better bargaining outcomes at the European level.³

The Stability and Growth Pact

Before explaining Member State behavior under the Stability and Growth Pact, it is necessary to explain how it worked and which actors were involved. The Maastricht Treaty provided the road map to create a common currency and an independent central bank to regulate it. While fiscal policies were to remain the domain of Member States, the Treaty anticipated the need for some sort of macro-economic policy coordination.

The European Monetary System crisis in 1992-1993, combined with a general worsening of budget deficits across the Union, heightened concerns about how Member States would perform once the Economic and Monetary Union began. Two of the so-called Maastricht criteria focused on fiscal policy and set reference values of a general government deficit no larger than 3% of GDP and general government debt burden no larger than 60% of GDP. The Commission used these values, in turn, to judge whether countries met the requirements to join the Euro Area. In practice, the emphasis in the run-up to the introduction of the euro was on the deficit figure, and Member States had to get their deficits at or below 3% if they wanted to adopt the euro.⁴ However, there were only vague provisions for what to do once the euro was created. A state with a fiscal crisis could hurt all members of the Euro Area, and the remaining countries could subsequently find it in their best interest to assist this state. Consequently, there seemed little to prevent states from

³Thank you to Ben Ansell and David Samuels for encouraging us to develop this point.

⁴For more details, see [Hallerberg \(2004\)](#), Chapter 2.

relaxing their fiscal stances once they joined.⁵

These concerns led to the creation of the Stability and Growth Pact, to which the Member States agreed at the Dublin Summit in December 1996 (see [Heipertz and Verdun \(2010\)](#) for more detail about the Pact's creation). All states were expected to have budgets that were "close to balance" or in surplus over the medium-term. This means that they were to adjust their balances according to the economic cycle – in economic good times they were to run surpluses, while in bad times they could run deficits. The Commission then had the responsibility to review each programme, and to make recommendations to the ECOFIN Council on whether the programme met European goals and whether the goals set in the programmes were realistic, given domestic conditions. ECOFIN would have to approve the Commission's recommendation by qualified majority vote of all Member States, which meant that "countries that 'sinned' retained the right to vote and needed only a few additional countries prospective sinners among them to block such steps" ([Schuknecht et al., 2011, pg. 9](#)).⁶

Perhaps the biggest shock to the Pact in the initial years was the behavior of two of the founding members of the European Union, France and Germany. They had problematic deficits beginning in 2002. The Commission recommended an "early warning" to Germany in 2002, but the ECOFIN Council did not pass it, and an informal agreement reached with Germany meant that Chancellor Schröder avoided getting such a warning in a national election year. In autumn 2003, the Commission judged that the deficits in France and Germany were excessive and that they would continue, and it attempted to begin the sanctioning process against each Member State. The ECOFIN Coun-

⁵Article 103 of the Maastricht Treaty was judged as a "no bailout clause" because it stated that neither the Union level nor the other Member States were liable for the commitments of other governments. There was concern, however, that this was not a credible position, and indeed, the financial packages for Greece, Ireland, and Portugal in 2010 and 2011 confirmed this fear.

⁶The corrective arm followed from the Commission's assessment of a given state's domestic policy. If it judged both that an excessive deficit existed and that a state had not made any progress in eliminating it, the Commission could recommend to the Council that the state make a non-interest-bearing deposit with it of up to 0.5 percent of GDP, depending on the size of the deficit. If the Member State continued to neglect necessary reforms, the deposit could become a fine.

cil, however, balked at following the Commission's recommendation. The Commission's response in January 2004 was to take ECOFIN to the European Court of Justice. In its judgment, the Court ruled that the Council could not suspend a procedure without the Commission's recommendation (Morris, Ongena and Schuknecht, 2006, pg. 17-8), but this was a hollow victory—no additional action was taken against France and Germany.

There was a clear sense both in the academic and policy communities that something was wrong with the Pact, resulting in a slew of reform proposals (see Begg et al. (2004); Schelkle (2005); Schuknecht (2004)).⁷ The Member States themselves agreed to a concrete reform in March 2005. Under the revised Excessive Deficit Procedure (EDP), more factors were explicitly considered when determining whether a country had an excessive deficit, such as if spending fostered international solidarity or if it promoted the unification of Europe.⁸ Member States also set their own medium-term objectives, which could include the future fiscal effects of major structural reforms in their calculations. Once a state was found to have an excessive deficit, it could receive extended deadlines so that it had a longer period to correct the deficit than before the revision. Countries experiencing negative growth were exempt, whereas the previous requirement for automatic exemption was a decline in economic output of 2 percent of GDP.⁹ In the end, the Pact was weakened.¹⁰

In summary, the process under the Stability and Growth Pact is more flexible than a simple 3 percent deficit rule, and yet, the core procedure has remained the same. The Commission evaluates every Stability and Convergence Programme and issues a report where it can point out where it

⁷The best summary of the proposals appears in Fischer, Jonung and Larch (2006) appropriately-titled paper “101 Proposals to Reform the Stability and Growth Pact. Why so Many?”

⁸The clause on “international solidarity” was widely seen as a way to mollify France, which wanted defense spending to be exempt, while “European unification” would allow Germany to include the costs of German unification.

⁹An excellent review of the changes in the Pact appears in Morris, Ongena and Schuknecht (2006).

¹⁰A further reform came during the euro crisis. The so-called “Six Pack” routinized the “preventive” part of the SGP through a new “European Semester” (see Fabbri (2013)). It also put more emphasis on a debt level of 60 percent or below, not just on the 3 percent deficit target. Perhaps most importantly for this article, it changed the decision-making rules – see the Appendix for further details.

finds that the given Member State has an issue. These texts, however, are not the final decision – the ECOFIN Council has to approve each one by qualified majority (through 2011), and the Council could (and did) change the Commission’s text. Conveniently for our purposes, the various texts at the different stages of the process are publicly available. In fact, the debate between the Council and the Commission are embedded into textual changes and edits, which makes systematic analysis of the texts useful for extracting data.

Political Determinants of Weakening and Editing

The European Commission evaluates the macro-economic policies and macro-economic performance of every government. Once the evaluation is complete, the Commission reports the evaluation in the form of a textual document and gives it to the Council. For this paper, if the Council accepts the Commission’s evaluation unchanged, then we say that the Council (on behalf of the Member States) is abiding by the rules. If the Member State changes the text, then we examine whether the Member State succeeds in weakening the textual substance and/or changes the textual phrasing. Weakening and editing the Commission’s recommendations is important because the Council, by augmenting the Commission’s evaluation, is intervening in the oversight mechanism that the Commission is tasked to perform.

Why do governments care about what the European Union says about them in the first place? Can all Member States weaken and edit the Commission’s recommendations or can only a subsample of countries amend? Why does weakening and editing vary across countries and over time? It is important to note that the final document is a “Council recommendation” to a given government, which supersedes whatever the Commission has written before. A negative recommendation from the Council leaves the government open to criticism from both the press and the opposition. But, as we show in this paper, governments vary in their their ability to change the Commission’s textual document and also vary in their sensitivity to what the “watchdog” writes. This paper offers

the first look at the political determinants of changes to the recommendations.

We focus on the “preventive” arm of the Pact and we measure the variation between the Commission assessments and final European Council opinions of the Stability and Convergence Programmes. The empirical section goes into further detail, but we consider two dimensions: whether the Council weakens the substantive message of the Commission’s assessment and whether the Council edits the language of the recommendations. Weakening, therefore, is when the Council (as the representative of the Member States being evaluated) changes the evaluation of the Commission such that the evaluations are *substantially* less stringent. As a second measure, when the Council edits the language of the texts such that the textual phrasing of the assessment is different from that initially proposed, we call this editing.

Our explanation for why we see variation in the weakening and editing of the Commission’s recommendations borrows logic from the judicial politics literature on separation of powers and the comparativist literature on elections. Like in theories of courts where judicial functions are constrained by other branches of the government, we assume that the European Commission operates with political constraints imposed upon it from the Council of Ministers. Thus, the Commission’s ability to make recommendations is dependent on the power of the Member State and the public support it is afforded (Staton, 2010; Vanberg, 2005). In particular, we expect that as Member State power increases so too can the ability for the Member State to change what the Commission writes about it in the first place. We also note, however, that the Commission also has the means to increase its power especially when public support for the EU is high. Thus, an increase in euroskepticism in the domestic population of a given Member State should weaken the relative position of the Commission and strengthen the relative position of the Member State. Alternatively, a decrease in euroskepticism in a given Member State should increase the relative position of the Commission. Finally, Member States may have different sensitivities in or near an election and are expected to be particularly sensitive to what the “watchdog” writes in an election year than during other times. Therefore, another political reason we may expect to see variation in

weakening and editing has to do with elections. Thus, according to our framework Member State power, country-specific euroskepticism, and elections are all possible political determinants that might help explain the extent and with what frequency a Member State will weaken and edit the Commission's recommendations against it.

In addition to these factors, we also account for other explanations prominent in the literature. In particular, we examine whether the partisan hue of the government matters, whether economic conditions matter, and also whether shared common shocks prompt Member States to behave in ways similar to one another. In each case, we discuss below the logic of the argument, highlighting the empirical predictions for how these factors relate to amending the Commission's evaluation. The following section then offers an empirical test of these determinants. As a highlight of our findings, we find that large Member States and euroskeptic Member States are positively associated with weakening the Commission's recommendations and editing the text that the Commission writes about them. We also find that during the euro crisis, the propensity for weakening and textual editing went down, but that the profile of the Member States that engaged in weakening and editing holds throughout the period.

3.1 Member State Power

When examining multilateral lending and conditionality, scholars have found that powerful countries exert significant influence ([Thacker, 1999](#)) and that state power shapes the size and the conditionality of loan disbursements ([Copelovitch, 2010](#)). We see two ways to conceptualize the concept of "power" in the European context. One focuses on the main Member State players, while the second focuses on the formal voting rules that provide the relative weights for every Member State. The first approach is consistent with Stone's observation in his discussion of France and Germany's behavior under the Stability and Growth Pact that "the rules of the Pact are difficult

to enforce against powerful countries” (Stone, 2011, pg. 118).¹¹ Applying this logic to our set of dependent variables, powerful states might be more likely to weaken and edit the Commission’s recommendations because the potential punishment they face is less than punishments faced by small states. That is, what can the European Commission do if Germany wants to amend the Commission’s recommendations? In contrast, if Portugal receives specific recommendations for adjustments to its budget policy, it may be more vulnerable to pressure from the European Commission. It is also worth considering, however, that large Member States like France and Germany have a more difficult time complying with the Pact’s three percent deficit rule, which arguably has an impact on the behavior outlined here. Thus, it is not necessarily the case that large Member States (always) get (exactly) what they want (see for example Buti and Pench (2004); Fatas and Mihov (2003)).

This logic then begs the question about whether a strategic Commission would choose not to highlight large Member States in the first place. If the Commission were perfectly constrained by the wishes of large Member States, the Commission would make recommendations only for small states or draft recommendations that Member States accept. Empirically, however, we observe that the Council frequently changes the Commission’s texts for large states, suggesting that even large Member States do not get what they want from the Commission. While this brings up the question of whether large states would get even stricter recommendations if they were not so powerful, it suggests that, if the European Commission were perfectly constrained, we would get a null or negative sign on the coefficient for political power in our analysis – large states do not need to ask for changed Commission texts for them because the Commission would have written texts the large states wanted from the beginning. The other extreme is one where the European Commission is perfectly unconstrained by the Member States. In this case, the Commission would write whatever the Commission wanted and the Member States would attempt to change the Commission’s

¹¹Similarly, Schure and Verdun (2008) break Member States into “small,” “medium,” and “large” groups and argue that the size of Member States affects the amount of discretion they want to give to the European Commission.

assessment to the best of their ability. This would not explain, however, why the Commission gave less stringent recommendations leading up to the euro crisis. Thus, it seems that Member States, even large states, are at least partially constrained by what the Commission does and equally, the Commission is at least partially constrained by the Member States.¹² ¹³ Our expectation is that large states manage to weaken, and to change, the Commission's text on them more frequently than small states.

Rather than expect "big" states to get their way while "small" states do not, a second approach is to measure relative power according to the formal voting rules for approving a recommendation to a given Member State. There are four separate voting regimes over our time period. In the first, 1998-2002, it is the Treaty of Amsterdam rules. Under this regime, the largest state has 10 votes and the smallest state has 2 votes. The second voting regime is from 2003-2008, or under the Treaty of Nice. This Treaty introduced a so-called "triple majority" as the definition of QMV. The next voting regime is 2009-2010, or Lisbon. Under these rules, the Commission only needs a qualified majority of the relevant set of Member States' votes. Finally, another regime was in place as of 2011. According to these rules a Member State needed a qualified majority to block what the Commission wants rather than QMV to pass it. It takes a few large states, or a combination of a large and a few small ones to block changes to the Commission's texts. In this paper, we initially follow Stone with a dichotomous categorization of state power. In the robustness section we also check for whether or not each voting regime matters (see also [Schure and Verdun \(2008\)](#))

¹²There is evidence that the Commission engages in such strategic behavior in cases it brings to the European Court of Justice. That is, it presents the cases it expects to win, (see [Carrubba \(2005\)](#); [Fjelstul and Carrubba \(2015\)](#)) however, whereas in writing the assessment criteria, the Commission has to write something on every country and the guidelines for what types of report the Commission writes is somewhat standardized.

¹³This is the baseline model, and if it explains everything other possible explanations are not relevant. This goes in particular for the argument that interaction in the Eurogroup among ministers constitutes peer pressure, and if there is enough pressure on offenders in this off-the-record arena then the offenders will reform their ways. We would like to emphasize that we know of no academic piece that makes this argument, but discussion in Brussels and elsewhere on the utility of the Eurogroup focuses on peer pressure to resolve disputes. It may very well be, of course, that peer pressure is useful for smaller issues, but that the more visible compliance or non-compliance under the Stability and Growth Pact would not be subject to peer pressure.

and details in the Appendix on differences in the formal voting rules over time).

3.2 Public Support and Euroskepticism

Rather than focus only on the dynamics of voting within the Council, there are relevant variables outside the Council that affect whether the Commission gets its original text. We expect that public support for the EU helps to explain to what extent the Commission can prevail against the wishes of the Member States. Like the Member States, we assume that the Commission would like to write recommendations according to its own preferences, which, we assume, occurs when it can perform its role as a “watchdog.” We also assume that receiving support from the Member State’s domestic population can help bolster the Commission’s strength relative to the Member State. In particular, we associate EU level public support for the Commission, which is how much or how little the domestic population supports EU level institutions.

Arguments usually assume that citizens want their governments to honor their treaties. [Chaudoin \(2014b\)](#), however, finds in a study using a survey experiment that voters worry about keeping to a treaty only when they do not have a strong preference over a given policy: a voter who opposes free trade would prefer that her government not comply under a trade treaty. In another paper, Chaudoin also finds that domestic populations are more likely to be sensitive closer to elections and that this behavior induces governments to behave strategically with voters ([Chaudoin, 2014a](#)).

While most of the previous empirical literature has examined this in the context of trade, such a theory should also apply to the case of the European Union. We know that voters have different opinions across Member States on the “treaty” itself. Some, such as Ireland, have populations that were, on average, euro-optimists (at least before the crisis), while others like the United Kingdom had more euroskeptic populations. Because the treaty here relates to the project of European integration, we expect that Member States with citizens who are generally euro-optimists are more likely to want leaders to follow what the European Union says. In fights with the European Com-

mission, a population may be more likely to side with the Commission where there is enthusiasm for European integration. In contrast, in euroskeptic Member States, we expect the reverse – governments are more likely to fight to change negative Commission assessments knowing that in any showdown with the Commission the body is unpopular.

3.3 Elections

In addition to voters having different preferences, governments may also have different sensitivities to their domestic populations over the electoral cycle. Governments are presumably more sensitive to what populations think of their performance and may not want Commission criticisms of them during elections. Moreover, there may be more “bad” behavior for the Commission to complain about in an election year, as governments may try to boost the domestic economy by spending more and/or cutting taxes. Heipertz and Verdun, in particular, argue that electoral considerations were especially prominent in the original design of the Stability and Growth Pact and so they may also matter in explaining textual changes to the Commission’s recommendations (Heipertz and Verdun, 2010). This is the domestic story we explore in our main set of results, with the expectation that there is more editing of Commission texts in election years.^{14 15}

A summary of our predictions is in order. As a reminder of the order of play, Member States

¹⁴There are additional domestic factors we consider in the robustness section. There is a developed literature on Member State compliance under European law. The predictor for behavior in many of these studies is the position the Member State took prior to passage of the given law. Thomson, Torenvlied and Arregio (2007) consider the impact of the negotiation stage on the implementation stage, with the focus on directives and subsequent infringements and delays. Such directives must be transposed into national legislation. They find no evidence of country-specific effects, but they do find that countries that opposed the directive during its passage are more likely to face an infringement case. Similarly, König and Luetgert (2009), in a large-*n* quantitative study, find that countries that fail to transpose their directives on time were more likely to have been opposed during passage.

¹⁵Given that France and Germany pushed through the Stability and Growth Pact, the previous literature would suggest that these Member States would be more likely to comply. Yet these countries were widely seen as the first to break the Pact. This behavior suggests that country positions in the mid-1990s when the Stability and Growth Pact was originally passed are not a domestic reason why a government would be more likely to push for changes in Commission assessments. At the same time, the prediction would be the opposite of the one based on the size of the state, that is, France and Germany should be more likely to comply, and as a robustness test later in the paper with explore this argument in more detail.

submit either convergence or stability programmes yearly to the European Commission. The Commission, in turn, evaluates the programme and submits an assessment to the ECOFIN Council. The Council then decides whether to accept the Commission's evaluation as-is, strengthen the evaluation, or weaken it. In addition to the direction of changes, the Council also decides with what intensity to edit the Commission's text. Finally, once handed the new text by the ECOFIN Council, the government then decides whether to comply with an explicit recommendations. Given the theoretical expectations outlined above, in the next section we evaluate the following expected relationships.

When Member States have oversized power on the Council of Ministers, they can more easily act against the preferences expressed in the European Commission's recommendations. As a result, we argue that large Member State are more likely to change the Commission's recommendations in a direction more favorable to the Member State (i.e. weakening). Alternatively, small Member States cannot do this as easily. Our first testable hypothesis is therefore:

Hypothesis 1: *Big Member States will have more edited, and weakened, texts than small Member States.*

It is not only the Member States that have a stake in the recommendations, however. We argue that when public support for the Commission is high the European Commission is better able to perform its watchdog function than when public support for the Commission is low. We therefore expect the following:

Hypothesis 2: *Countries with domestic populations that are more euroskeptic will have more edited, and weakened, texts when the Commission has lower domestic support than when the Commission is afforded greater public support.*

Finally, we also expect that Member States are more sensitive to what the Commission writes in an election year. Similarly, governments may engage in stimulating the economy so there may

be more reasons for the Commission to write more stringent recommendations. This suggests our final hypothesis:

Hypothesis 3: *Because of election concerns, Member States in an election year will be more sensitive to what the Commission writes about them and they will have more edited, and weakened, texts.*

Research Design and Analysis

4.1 Dependent Variable

We examine variation between the European Commission assessments and European Council opinions of the Stability and Convergence Programmes between 1998/1999 and 2012/2013. The countries in the study consist of the EU-15.¹⁶ One might wonder why we include the three countries that chose not to join EMU, namely Denmark, Sweden, and the United Kingdom. We see these states as interesting sources of variation and as possible checks against sample selection bias. Euro out countries must submit convergence programmes yearly, and those programmes undergo the same review as the stability programmes that the Eurozone members face. As “outs,” they are not subject to the same possible sanctions that Eurozone members face, and in particular, they are not subject to possible fines, so it may be that they do not face the same incentives. By including them in the statistical analyses, we can empirically assess whether or not the behavior of the “in” countries is different from the “out” countries.

While our focus is on the EU-15 because these are the original Member States at the beginning of the Stage III of Economic and Monetary Union, which introduced this monitoring procedure for

¹⁶Countries are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, and United Kingdom.

all Member States, to be clear that our findings are common after other countries joined we report results from an expanded EU-27 sample in the robustness section that follows.

As mentioned, we develop two new measures that account for disagreements between the Commission's assessments and the wishes and opinions of the Member States, as reflected by the Council. The first measure examines the direction of Member State amendments and reports, specifically whether editing by the Member State made the Commission's amendments weaker or left the text the same. The second measure tracks the frequency that Member States change the Commission's text. Together, these new measures capture both the direction and also the intensity of amendments.

Weakening: Using human coding of Member States' edited Commission documents, we measure the direction of the amendments between the Commission's assessments and Council's opinions and put them into two categories: 1. "No change"– there is no noteworthy or substantial variation (defined more precisely below)¹⁷ 2. "Weakening"– the Council Opinion has placed less stringent demands or in general is less critical of a country programme than the Commission assessment.

Before one accepts the validity of the measure, a few words about coding are in order. We do not consider statements that are substantively similar yet vary in the level of detail.¹⁸ Omissions in the statement are considered substantively significant if the Commission holds a country to a specific outcome but the Council does not. Treatment of omissions is especially relevant because there is variation in the length and detail in various reports. Finally, we do not record a distinction in the magnitude of weakening; if there is a perceived change in one direction or the other in

¹⁷We also coded cases where there was a "strengthening." There were 28 such cases across the EU-15, which were comparatively rare (12 percent of cases), and we combined them with the null cases.

¹⁸For example, for Greece 1998/1999, the Commission indicated that "structural reform is needed in order to improve the efficiency of the Greek economy, particularly in its large public sector." The Council's response was as follows: "The Council welcomes the structural reforms included in the programme which are geared towards the labour market, the social security system and the wider public sector. The Council urges the Greek Government to implement them as scheduled and pursue further the reform effort in order to enhance the potential and efficiency of the Greek economy." This case was coded as no significant change.

the statement, we simply record it as such. Commission and Council documents are classified as having a weakening, a strengthening, both a weakening and a strengthening, or no change in a given year. Table 1 provides examples from each class.

Given these provisos, our first, hand coded dataset includes 224 annual Commission and Council reports on Member State programmes.¹⁹ Figure 1 “Weakenings by Country” shows the distribution of weakening by Member States over the sample period. There are 67 instances of the Council opinion weakening the Commission assessment. This amounts to approximately 30 percent of the observations, suggesting that weakening is common.²⁰ Especially striking is the clustering of weakened Council reports. While the Commission’s fight with France and Germany certainly got the most press attention in autumn 2003, most states received lighter Council comments that year than what the Commission first proposed in their original report. Also noteworthy is that at least two states received a weaker report in every year but 2006. The lack of cases of weakening in the most recent years is also noteworthy, especially after the 2011 reforms and the crisis period.

There is also clear variation by Member State. The total number of weakening varies from a high of 11 (France) to a low of 1 (Denmark, Luxembourg, Ireland and Spain). Collectively, France and Germany had nearly 26 percent of all cases of weakening, despite constituting around 14 percent of the dataset. While this relationship is perhaps unsurprising, the surprise is that many states received lighter comments from the Council. The “out” states of Denmark, Sweden, and the United Kingdom have proportional numbers; they constitute 22 percent of the weakening and they also represent 20 percent of the cases.

Editing: While the direction of changes depicts how Member States oppose the Commission’s evaluations, another equally important measure is the frequency or intensity of editing to

¹⁹Our total dataset has 225 rows with one missing value of weakening for Denmark 2002 and 36 missing values for our second dependent variable, edit distance. We check to make sure that missing data is not skewing our estimates by using multiple imputation in Amelia.

²⁰We also considered whether the Council on balance strengthened, rather than weakened, the Commission text. There are such 28 instances, or about 12 percent of observations. A majority of states have one such case. For the purposes of the analysis that follows, we treat strengthening cases as “no change.”

the Commission’s evaluations. In order to measure the frequency that a Member State opposes the Commission’s text, our second dataset constructs a metric of the Member State’s editing.²¹ To do this, we use automated textual analysis. First, we find the instances of textual differences between the Commission and the Council’s text using textual comparison software. Second, examining only the changed text, we calculate the Levenshtein distance metric for each edit made. The Levenshtein distance metric is a measure that computes the minimum number of single word deletions, insertions, or substitutions that are required to match the new text to the old text (Sørensen, 2007). For example, if the Council edits the initial recommendation from “2010” to “2010 in view of the uncertainty connected to the economic downturn,” then the Levenshtein distance metric is 10 because there are 10 word insertions necessary to transform the old text to the new text. The Levenshtein distance is commonly used in linguistics, including spelling algorithms, and in statistical models to evaluate language similarity. The textual distance measure is particularly useful here because it picks up word changes in the text, even if small, between the Commission and the Council reports.²²

Second, after calculating the distance metric for each edit per country document (there are for example over 5000 edits in any given year), we sum the total number of Levenshtein distance measures per country-year. As seen in Figures 1 and 2 and like the measure for weakening, editing of Commission documents varies substantially not only across time but also across countries. Countries such Germany, Italy, and France tend to have, on average, higher edit intensity than other, less powerful countries. This pattern is analogous to the observed pattern for weakening.

In addition to the raw numbers, reported by the dash lines, Figure 2 illustrates the trend in textual editing and the trend in weakening together on the same plot reported by the solid lines so

²¹The sample includes the same time period as before with all EU-15 Countries from 1998/1999 to 2012/2013.

²²Another possibility is to calculate micro changes, such as changes to characters, or alternatively, macro changes such as changes to the number of sentences in paragraphs. Measuring word changes is most effective because words can account for small and potentially meaningful additions to text, while simultaneously, not being too sensitive to cosmetic edits such as changes from British to American spelling and punctuation.

that we can observe the similarity and the differences between the measures. One important source of variation between the measures is the differences in the timing of their peaks; the weakening measure peaks earlier than the edit intensity measure. The benefit of looking at the direction and the intensity over time and across countries is that it allows us to capture changes in the use of this particular form of bargaining, and highlights the changing relationship between the Member States as represented by the Council, and the Commission.

One potential problem with the Levenshtein distance measure is that the distance measure does not measure lexical similarity between sentences. For instance, consider two sentences with the same number of words but slightly different meanings such as: “we should expect a 2 percent increase” and “it’s likely that a 2 percent increase.” This example has a calculated distance measure of zero because it has the same number of words and no substitutions have been made. Despite this, however, additional statistical analysis and the correlation of the Levenshtein distance measure with weakening suggests that both measures capture the phenomenon that we examine throughout the paper. We also think that the distance measure contributes something important to the study of political institutions, as much of what political institutions do is provide a forum for negotiations over time and this process often produces textual information, which, as we show here, is useful for comparative research.

4.2 Explanatory Variables

Member State Power: We examine whether or not a country is large in terms of vote share they have on the Council. If a country had 10 votes on the Council prior to the Nice Treaty, which is equivalent to having 29 votes after the Nice Treaty, we code this country as “large.” If a country has fewer than 10 votes on the Council, the country is coded as “small.” Furthermore, as discussed previously, we also account for whether or not the country is an “in group” or “out group” member, with Denmark, Sweden, and the United Kingdom coded as “outs” and the other countries coded as “ins.” In the robustness section, we also examine different voting rules.

Public Support: As a measure of public support for the European Commission, we measure a population's expected favorability towards the European Union as a proxy for domestic support and thus an indirect measure of the political strength of the European Commission in its ability to perform its watchdog function. The data that we use is from Euro-barometer. We code the values for each country in response to the question:

Taking everything into consideration, would you say that (your country) has on balance benefited or not from being a member of the European Community (Common Market)?

The question captures the perceived benefits and costs of EU membership.^{23 24}

Elections: We examine two variables concerning the government's perspective on the Stability and Growth Pact. To measure the impact of electoral considerations on a government's fiscal policies, we code for the percentage of a given year that precedes an election (Franzese, 2002). That is, an election held on July 1 would be coded as 0.5 in the same year and 0.5 in the previous year. As a different specification, we also use the forward version of the variable, which is both an additional control for endogeneity and well as a test for whether Member State governments anticipate the effects of a Commission document on later elections.

4.3 Controls

Partisanship: One potentially important consideration that might help explain the impact of domestic politics on Council behavior is partisanship. Grieco, Gelpi and Warren find that shifts leftward make it less likely that governments comply with legal commitments to capital market

²³In a recent working paper Fjelstul and Carrubba (2015) also use Euro-barometer as a proxy for domestic level public support for the European Commission

²⁴Unfortunately, Eurostat chose not to ask this question in 2012. More on the implications of this for our empirical model appears below. Also note that audience costs in this way assume that costs are born by forward looking Member States who are less concerned with reputation and more concerned with winning a larger share the next election. This seems most appropriate for coalition politics in the EU context.

integration (Grieco, Gelpi and Warren, 2009). Similarly, the budget balance requirements in the Stability and Growth Pact could be popular among right-wing governments but not among left-wing governments. In order to measure this, we employ Benoit and Laver's left-right partisanship of the Prime Minister and the Finance Minister for all 15 Member States (Benoit and Laver, 2006) which appears in Hallerberg and Wehner (2015).

There are a few theoretical concerns with partisanship, however. First, given the presence of coalition governments in most EU-15 countries that often include parties from both the left and the right (for example Belgium and Finland), it is hard to envision a clear qualified majority from one partisan block in place in the Council. Second, the argument does not seem consistent with the evidence – Germany and the United Kingdom were both left-leaning when the European Commission suggested concrete recommendations while France and Italy were both right-leaning. Instead of the partisanship of the government as a whole, however, the partisanship of the Finance Minister may matter. Because the Finance Ministers are responsible for economic policy (and are held accountable for their economic policymaking), it is perhaps more likely that partisan concerns are stronger for the Finance Ministry than the head of the coalition government. In the empirical analysis, we test this directly by using the partisanship of both the Prime Minister and the Finance Minister. To offer a preview of the results, we find that neither the partisanship of the Prime Minister nor the Finance Minister matters; because of this and our theoretical concerns about the relevance of this variable, we exclude it from the final specifications presented in Tables 2 and 3 (but see the results in the robustness analysis reported in Table A.3).

Economic Performance: In order to test whether or not economic performance matters for weakening and editing, we collect data on Member State output gaps. In order to be consistent with the data used by the Commission itself, we use the estimate of the output gap from the European Commission. The output gap measures whether countries are performing better (in terms of changes in GDP) than potential output (also measured as change in GDP). Countries that report a positive output gap are growing above potential whereas countries with a negative output gap are

growing below potential. Note that this is different than actual growth. For example, Ireland and Germany can have the same growth rate but different output gaps. If Ireland is growing slower than potential, Ireland will have a negative output gap, whereas if Germany is growing above potential, Germany will report a positive output gap even if the countries exhibit the same economic growth rate. The potential growth rates are then subtracted from the actual growth rates. We use this measure of economic performance because the output gap is a key concept in the SGP criteria applied by the European Commission.

Shared Behavior/Common Shocks: Finally, in order to account for shared common shocks that may impact Member States in important similar ways, we construct a spline smoother of the yearly trend for both measures of the dependent variable. A positive coefficient on our shared Common Shock variable suggests a positive correlation between the likelihood of weakening and editing when other countries are weakening and editing whereas a negative coefficient estimate suggests the that countries are not behaving similarly. Again in the case of multilateral institutions, scholars indicate that lax enforcement of rules can lead to bad behavior whereas stricter enforcement of rules can lead to good behavior (Copelovitch, 2010; Dreher and Vaubel, 2004).²⁵

4.4 Results

Our first set of results examines the weakening the Commission's recommendations. We want to test whether Member State power, public support, and elections matter for whether Member States weaken the Commission's recommendations.

Because the first measure of our dependent variable, substantive weakening of the Commission's assessment criteria, is coded 0 or 1, we use a logit model. As suggested before, we think of the dependent variable as a latent variable that estimates the probability making the Commission's recommendations substantively less stringent for the Member State.

²⁵As a further robustness test, we also include a quadratic time trend to account for the u-shape. We find no appreciable differences in the results.

The estimates and the standard errors from the logit model are shown in Table 2 and Figure 3. The model is a mixed effect glm model that accounts for both time and country random effects. When we estimate the effects of Member State characteristics on the likelihood of weakening, the results show that large Member States are also more likely to have the Commission texts written about them weakened. In particular, holding the other variables at their observed values, the probability of a weakening increases from .27 to .41 if one moves from a small to a large state. This provides us with some indication that politically powerful Member States (at least as measured by votes on the Council) are more likely (or more able) to weaken the Commission's assessments against them.

We also find some supportive evidence for Hypothesis 2. The more that the Member State's domestic population is euroskeptic, the higher the average number of substantive weakenings to the Commission's assessments. A move from the least euroskeptic Member State, Ireland in 1998, with just 3 percent of the population euroskeptic towards the EU, to the most euroskeptic Member State, which is Sweden in 2003, where an estimated 59 percent of the population feels skeptical towards the EU, increases the probability of weakening the European Commission's recommendation by almost 50 percent (.48).

We also check whether or not these estimated effects vary by different country characteristics, as estimates of variables in logit models depend on values of other input variables. In Figure 4, we show the estimated positive relationship between euroskepticism of the EU and the probability of weakening the Commission's recommendations for large and small countries. The x-axis, which measures euroskepticism, is restricted to the interval where we actually observe data. Large Member States or those states with outsized voting power (solid-line) have an estimated higher probability of weakening than those Member States that do not have such outsized power (dotted-line). Furthermore, as shown above, countries with euroskeptic domestic populations (or alternatively, countries where the Commission is likely to have lower support) are more likely to weaken what the Commission writes about than countries with less euroskeptic domestic populations.

In terms of our third hypothesis, we find no significant results. We cannot reject the null that being in an election year has no effect on the probability of weakening the Commission's assessments against the Member State. Similarly, we examined different lag structures of when an election takes place. None of these alternative specifications was statistically significant or affected the substantive results. In fact, one of our most robust findings is that the election variable is always indistinguishable from zero and in every model specification when predicting weakening.

There is little evidence that euro "out" countries are less likely to weaken the Commission's suggestions—the variable is not statistically significant, although Figure 3 indicates that the estimate is negative. Where a Member State is in its economic cycle also does not have a substantive impact. We considered several alternatives here to get at the underlining concepts. For economic performance, the timing could be relevant, and we checked the performance in the previous and well as current year. We also considered changes instead of levels; if a Member States is experiencing a marked economic slowdown, it may expect a weakened Commission text because of economic factors. Finally, we find evidence of a positive relationship between higher average values of weakening and shared common shocks; this provides some indication that Member States may engage in common behaviors on the Council, where a higher number of changes in texts is associated with higher editing among the Members in general.

In addition to Member State macro-characteristics, we are also interested in the effects of the partisanship of the government. Coalition politics in particular make the partisanship explanation unlikely in the EU context, however, we test for this explicitly. We find a slight negative relationship between partisanship of the Prime Minister and the probability of weakening, though the relationship is not significant by traditional cutoff points. Upon further inspection, we also find no substantive relationship between partisanship and weakening. We also run the same model again only this time using the partisanship of the Finance Minister. Here we find a slightly stronger negative result; more right-leaning finance ministers are less likely to weaken the Commission's report, however, the results are also insignificant and weak. Therefore, these findings, coupled with our

null findings for elections, provide no evidence that domestic institutional variables matter. Furthermore, we think that the limited effects of elections and partisanship on predicting whether or not a Member State will weaken the Commission’s assessment criteria is one of the important empirical contributions that we uncover using this new dataset.

Our next step is to examine whether the results are similar when we examine editing. We run the same analysis as before, but, because the edit distance measure is continuous, we use a linear probability model. As shown in Figure 5, once again, we find that Member States with outsized voting power are more likely to have a greater number of edits than smaller Member States. A one-unit increase in Member State size, from Spain to Italy, for example, is associated with approximately 300 additional edits to the European Commission’s recommendations. Furthermore, as before, being an “out” Member State is negatively associated with editing, though again, the sample size makes the uncertainty intervals around the estimate large here as well –“out” countries are estimated to make on average 100 or so fewer edits than “in” countries, even when we account for other things that we think are important such as euroskepticism and the country’s economic conditions. To put this into perspective, this means that, for every year, Euro Member States make at least 100 more changes to what the “watchdog” writes than countries that are subjected to the same oversight process but are not subjected to EU level fines, although the uncertainty around the estimate includes the null hypothesis estimate.

Unlike our previous findings, however, we do not find evidence that euroskeptic countries are more likely to edit intensively. This finding suggests that euroskeptic countries pursue a strategy of making substantive changes to the assessment criteria rather than engage in significant editing of the texts. Perhaps even more importantly, like the above findings, we find no systematic evidence that domestic institutional variables matter for weakening or editing. For example, we find no evidence that an election year matters for the number of edits made to the recommendations.²⁶

²⁶Surprisingly, there is some evidence that future elections are negatively associated with editing today but not with weakening. Future research should try to tease out whether or not this is because the Commission is easier on Member States.

Finally, the estimate for common shocks is also positive and significant. As above, these findings imply that a one unit increase in the average intensity of editing generates additional editing by other Member States.

In sum, the statistical analysis suggests that large Member States make on average a higher number of edits and are more likely to weaken the European Commission's texts about them. In addition, "out" countries are less likely to change the Commission's text and make fewer edits. Member States with relatively more euroskeptic populations are also more likely to weaken text but interestingly, we do not find any evidence that they systematically edit more than their euro-positive counterparts. This finding conforms with the theory that euroskeptic Member States have lower public support for the European Commission, which allows them to more easily weaken the power of the Commission's "watchdog function." It does not, however, explain why euroskeptic Member States do not try to change the textual information through editing more frequently. Also important for having confidence in our results is that the substantive interpretation of both the logit and the linear probability models produce similar findings across model specifications.

4.5 The Euro Crisis

The above analysis suggests that Member States are more likely to weaken when others are weakening and editing more intensively. Furthermore, we might expect that Member State behavior might be different before and after the euro crisis. During the euro crisis, Member States were encouraged by the European Commission to run larger deficits throughout the crisis than before the crisis. In fact, deficit spending was encouraged by the Commission and, in 2009, the Commission suggested that Member States should generally spend money to combat the crisis. Therefore, we expect less oversight during the crisis and therefore less editing and weakening at the end of the sample period.

We check for differences in the estimated effects before and after the euro crisis by including a dummy variable for the crisis period (2008-2011). In these tests, we find that Member States are on

average less likely to make amendments, either weakening or textual changes, though the effects are insignificant at traditional cutoff levels. The fact that the coefficient is negative conforms to the decline in the measures of the dependent variables shown in Figure 2.

Another issue may have to do with varying coefficients for the euro crisis and pre-euro crisis period.²⁷ As our dataset indicates, the number of edits first increases over the period and then declines. One way to think about this is to consider a simple interaction model. In this case there are two regimes, where the first set of estimated coefficients reports those periods where Member States actively try to weaken the Commission's evaluations, and the second estimated coefficients reports when Member States do not actively try to edit and weaken the Commission's evaluations. We examine this by interacting the crisis variable with the common shock variable. This gives us a separate effect of the average amounts of weakening and editing conditional on whether or not we are in the pre-crisis or crisis period. Unfortunately, due to data limitations, we cannot evaluate the estimated coefficients after the voting reforms in 2011. When we account for possible variation in the coefficients in this simple way, the variables for political power and euroskepticism remain stable and significant predictors for weakening. In addition, the crisis year estimate has a negative, though insignificant, estimated relationship with weakening as before. The common shock estimate continues to have a positive and significant effect, suggesting that when the average number of edits increases, weakening increases. This is consistent with our previous findings. The estimate for the interaction term of crisis and the common shock variable is positive and insignificant; this suggests that the results hold across the sample broadly.

We also run the same model using the alternative dependent variable, edit intensity. Our findings are similar. Political power remains a positive and significant predictor of edit intensity. The common shock variable, measuring the trend number of edits, is also positively associated with

²⁷ Johnston, Hancké and Pant (2014) argue that the build up to the crisis is about variation in competitiveness rather than government borrowing or levels of debt. Irrespective of the incentives for why Member States behave this way, if the mechanisms that we think matters are at work, we would still observe heterogeneity in the ability of Member States to change the texts.

edit intensity, again, suggesting that common behaviors may be at work. The crisis variable is negative though again insignificant by traditional cutoff values. Finally, the interaction between the crisis years and the trend number of edits is insignificant. These results are shown in Table A.1

4.6 Robustness Checks

This section evaluates whether the estimates reported above are stable in different model specifications.

Our first set of tests concerns our measure for state power. Instead of operationalizing power according to whether a given state had 10 votes on the Council under the Pre-Nice Treaty rules, we create an indicator variable for only France and Germany and rerun the results. We find similar results (reported in Table A.1 in the Appendix): political power, in this case measured only by France and Germany, is a strong and stable predictor of textual weakening. We also drop France and Germany from the sample entirely, and re-run the results using the 13 remaining euro countries. In this case we find an attenuated relationship between Member State power and weakening, however, the euroskeptic variable remains stable and significant, as do the other variables. In the case of predicting editing, even without France and Germany, we find a positive and significant relationship between Member State power and editing.

We also examine a more nuanced measure of the ability of a given state to change the Commission's proposed text, which we label "vote blocking." The four separate voting regimes are described above and in the Appendix. The variable represents the blocking percentage a state needs to attract beside its own vote to prevent the Council from blocking the Commission's text. France, for example, needed 17 percent of outstanding votes in 2000 while Finland needed 25 percent to block a Commission text. Our expectation is that the sign on the variable should be negative. Prior to 2010, for example, states voted on their own cases, which meant that large states needed fewer votes than small states. After 2010, states could no longer vote on their case, but only eurozone countries could vote on other eurozone members. Using the percentage required to block, this

measure for state power is not significant but the direction of the relationship is as expected. The euroskeptical variable remains positive and significant for the weakening specification.

In two other specifications, we change the timing for the euroskepticism variable to include the current year's level of expressed euroskepticism. As we mentioned above, there is some worry that the current year's level may be endogenously related to weakening, however, we use the current year as a robustness check to confirm that the results are stable. Like the other models, we find that euroskepticism is a strong predictor of weakening whether we use the current or forward specification. Furthermore, this remains true whether we specify large country power using only France and Germany, or the larger sample of countries with voting power above 10.

When we run the results using the dependent variable for editing, however, using only the France and Germany specification for power, the political power variable is positive but no longer significant at traditional cutoff points. Thus, there is some evidence that it is not just the behavior of France and Germany, but the larger sample of countries with voting power in the Council, that matters. Furthermore, we also test the effects of euroskepticism in the current and forward versions of the variables on text editing. While the current period is positive, though insignificant using traditional cutoffs, the future variable is negative. These findings are reported in [Table A.1](#) and [Table A.2](#) in the Appendix.

Instead of being limited to only the EU-15, we also ran the results using an expanded sample of all EU-27 countries. The results, presented in [Table A.3](#), are consistent with the results from the restricted sample above. Here we find that countries that are powerful and countries that are more euroskeptical are also more likely to weaken and edit more intensively. While we think that the results from the expanded sample are helpful in bolstering our claims, we still believe that the analysis is most appropriate for the smaller sample of countries, but it is reassuring to find the same results with an expanded set of countries.

In summary, the results from our robustness analyses are consistent with the evidence that we provide in the previous section with one exception, namely the relative ability of a state to block a

Commission change instead of the dummy variable for “large” and “small” Member States is not significant. This finding is consistent with (Stone, 2011), who notes that the big states get their way in the European Union regardless of the particularities of the actual voting rules in a given policy area.

In the other cases, political power is positively associated with weakening and editing. The level of euroskepticism in the domestic population matters, especially in predicting whether or not Member States will weaken the Commission’s text. What is perhaps most important is what we do not find: in none of the specifications do the domestic institutional variables matter. Therefore, our main empirical finding is that domestic variables, such as elections and the partisans hue of the government, are not systematically related to observed textual weakening of the Commission’s evaluations, nor does being in an election year or partisanship predict textual editing. Thus, there is little evidence that domestic explanations help to explain when and which Member States can enjoy flexibility in the rules, at least in this context.

Conclusion

This paper explores one mechanism at the European level that was meant to prevent a debt-fueled crisis in the Eurozone, namely the Stability and Growth Pact (SGP). We find that Member States were able to undermine the “watchdog” function of the European Commission in the Council of Ministers. We test for the determinants of Council changes in Commission assessments using human coding and automated textual analysis of Commission texts and Council edits. The ability of large or euroskeptic Member States to amend recommendations generated incentives for Member States to rewrite Commission criticism of their economic plans. The consequence was that Member State behavior systematically eroded the power of the oversight mechanism. This behavior contributed to the eurocrisis. It eroded the main institutional bulwark at the European Union level against higher deficits and debts. Of course, it is possible that the Member States would have sim-

ply ignored tougher language under the “corrective” arm of the Pact, but the general weakening of the Commission assessments left little scope for testing out the procedures under this part of the Pact.

Yet, it is not the case that every Member State had equal success in eroding the Pact. Large states were able to push their weight around. While journalists noted this concerning France and Germany in 2003, our study indicates that this was a common occurrence over several years. Moreover, we add an important variable to the debate, namely the opinions of voters concerning the international institution meant to evaluate the given government’s performance. It suggests that Member States that want to improve their negotiating outcomes at the European Union level should encourage more euroskeptic populations at home. This may help to explain, for example, why we see heads of government making speeches against EU regulations. The ability of the British government to negotiate a “British Rebate” would arguably fit a model where domestic euroskeptic appeals are made for Member State gain at the EU level. This finding is not without qualifications, however. The situation facing the Cameron government before a referendum on EU membership suggests that this type of strategy has the potential to backfire. A Member State may be able to get more from an international organization by having a skeptical domestic audience, yet, being too skeptical may mean foregoing potential benefits if that skepticism ultimately pushes the given state out of membership in the organization.

Looking forward, one question is whether changes in the framework at the end of 2011, be it through the “Six Pack,” “Two Pack,” or “Fiscal Compact” will change the patterns we observe in our study. Take the two main findings in this paper. In terms of euroskepticism, one goal of the European Semester is to increase the involvement of national parliaments as “watchdogs.” Moreover, as European election results in May 2014 suggested, European populations are even more euroskeptic than before. This may embolden Member State governments to fight what the Commission says about them and to work harder to change judgments they find unacceptable. Furthermore, there is additional evidence that the level of euroskepticism of a domestic population affects the inter-

action of domestic institutions on the European stage. One recent study, [Hallerberg, Marzinotto and Wolff \(2015\)](#), finds that more euroskeptic parliaments are also the ones that are engaged in checking the Commission recommendations. While the actors in this study are parliaments and not the governments themselves, such evidence suggests that greater euroskepticism may continue to affect the way that national institutions interact with the institutions of the European Union.

When it comes to large Member States, the relationship that we find in this paper might also change. We find evidence that a large state is more likely to change the Commission's document, and we also do not find an effect of blocking power in our robustness tests, but we have observations only from 2011 and 2012, and the the rules in place from 2011 require a qualified majority to block a Commission text rather than a qualified majority to pass it. If the formal rules are all that matter, this change in the rules should return some power back to the Commission. As [\(Stone, 2011\)](#) documents, however, large states have outsized influence in international organizations especially during extraordinary times. Germany has clearly played the type of role Stone describes since the beginning of the euro crisis, and one would anticipate that it would continue to do so now. If we combine more euroskeptic populations with German insistence that the rules now be respected, we would expect increasing tensions on the application of this framework, as we indeed saw in Summer 2015 with the discussions over the continued fiscal difficulties in Greece. The results of our paper suggest that effective governance from European Union level institutions will remain difficult.

An additional surprise is a non-finding – while domestic population preferences seem to influence the ability of a given government to weaken the Commission's assessment, upcoming elections and the partisan hue of the government do not. These results, of course, may be dependent on the time frame we examine, where there seemed to be little partisan disagreement about the desirability of small budget deficits. The increasingly contentious debates about the limits of “austerity” packages suggest that partisan differences may become more relevant in the future. Nevertheless, this non-finding concerning elections and the partisan hue of the government is of interest. Our

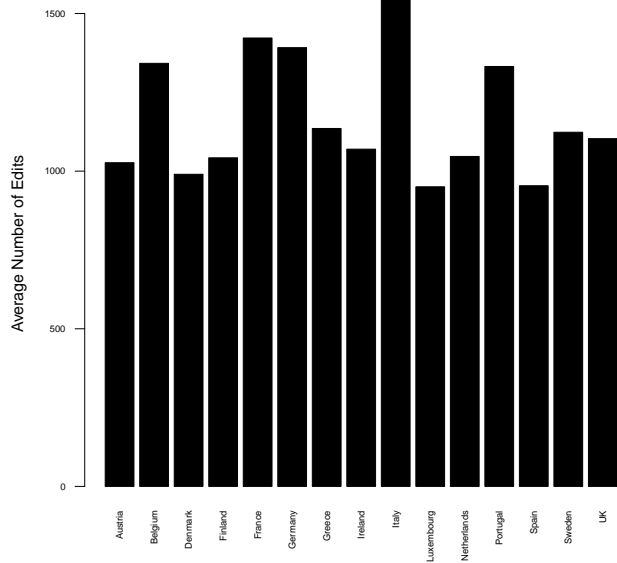
findings suggest that the underlying populations preferences, rather than the government of the day, is more important in explaining government action at the European level.

Finally, we also show how the use of human coding and computer textual analysis can be fruitfully employed to measure and analyze important aspects of institutional behavior. This is important as scholars often pay attention to whether countries comply with the rules without first uncovering whether, which, and how much countries can use the institutional framework to negotiate the terms even before they make the decision of whether or not to comply. Equipped with tools that allow for the computation of textual changes, we show how these tools can be used to build new and domain specific measures. Models similar to the one explored here could also be extended to account for deliberation and bargaining in committees, such as in legislative or central bank committees, or as important information that tracks variation in actors' positions across space and over time.

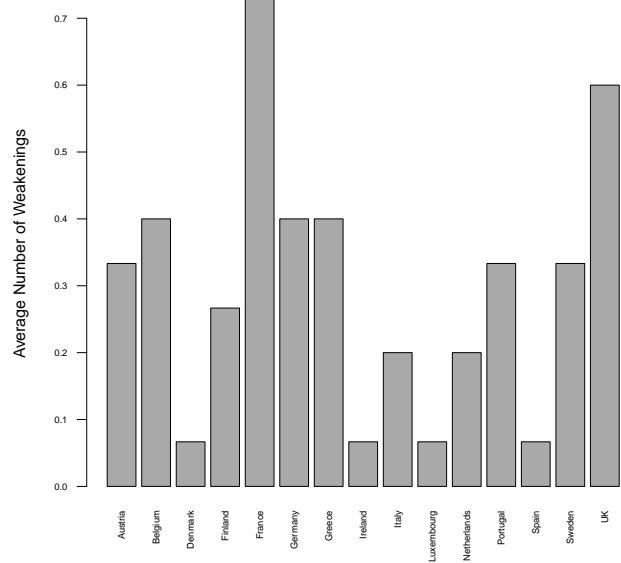
Table 1: Example Coding for Weakening Measure

Example of Commission and Council Assessments of Stability and Convergence Programs

Examples	Country and Year	Commission Assessment	Council Assessment
No Change	Italy (2002-2003)	8/1/03: "Given Italy's high debt, large primary surpluses will be required for many years."	21/1/03: A) "Given Italy's high debt, primary surpluses on the order of 5% of GDP will have to be maintained for many years."
Weakening	France (2010-2011)	6/7/11: Based on the assessment of the updated stability programme...the Council is of the opinion that the macroeconomic scenario underpinning the budgetary projections is too favourableThe aim of the Renewal of Social Dialogue Act and of the 2008 Labour Market Modernisation Act was to modernise social dialogue and to address the issue of labour market dualism. However, the reform did not tackle contractual segmentation and generally fell short of the measures needed to resolve the dualism of the French labour market.	7/12/11: "Based on the assessment of the updated stability programme...the Council is of the opinion that the macroeconomic scenario underpinning the budgetary projections is C 213/8 Official Journal of the European Union 20.7.2011 EN....The aim of the Renewal of Social Dialogue Act and of the 2008 Labour Market Modernisation Act was to modernise social dialogue and to address the issue of labour market dualism. The latter Act notably introduced a new procedure for terminating permanent contracts by mutual agreement (rupture conventionnelle) which is now increasingly used.
Weakening	United Kingdom (2004-2005)	16/2/05: A) "Deficit projections to 2006/2007 have been revised upward relative to the previous update despite an essentially unchanged macroeconomic outlook , while..." B) Paragraph concerning possibility of crossing 3% deficit rule (see Council cell) C) "For 2004/05, evidence of significant progress remains unconfirmed in outturn data and there is a high degree of uncertainty over both expenditures and revenues, leaving the risk noted above of a deficit higher than 3% of GDP."	A) "Relative to the 2003 update, deficit projections for both this financial year and the next have been revised upwards; beyond..." B) In a paragraph warning of UK's potential to break 3% rule, Council adds the following sentence: " However, given the decreased volatility of the UK economy in recent years and the cautious approach on the growth assumptions underlying the public finance projections, the risk appears small in the outer years. " C) "There is a clear risk that the budgetary outcome could be worse than projected in the programme in the short term."



(a) Editing by Country



(b) Weakening by Country

Figure 1: Across Country Average Variation in Editing and Weakening

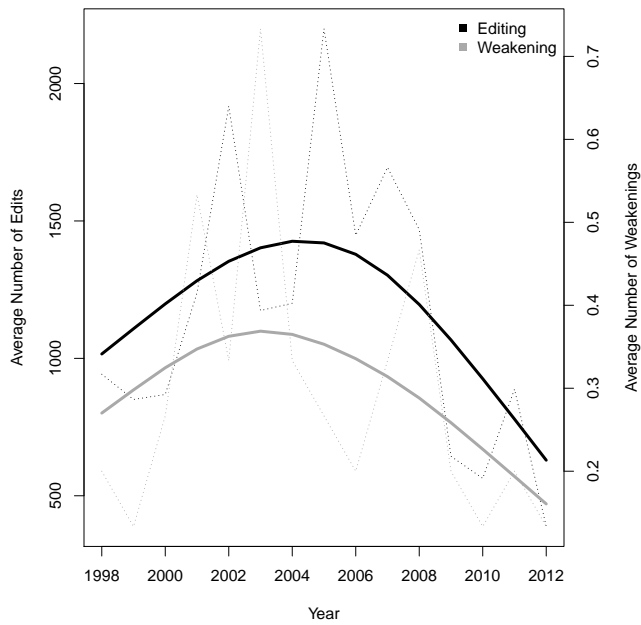


Figure 2: Over Time Variation in Average Editing and Weakening by Year

Table 2: Political Determinants of Weakening and Editing

	<i>Dependent variable:</i>	
	Weakening	Editing
	<i>generalized linear mixed-effects</i>	<i>linear mixed-effects</i>
	(1)	(2)
<i>LargeState_t</i>	0.84* (0.48)	307*** (117)
<i>OutputGap_t</i>	0.03 (0.09)	-21 (19)
<i>EuroOut_t</i>	-0.42 (0.57)	-113 (123)
<i>Euroskeptic_{t+1}</i>	5.37*** (2.00)	-248 (434)
<i>ElectionYear_t</i>	-0.22 (0.56)	-9 (119)
<i>CommonShock_t</i>	5.74*** (1.19)	1*** (0.1)
Intercept	-4.48*** (0.84)	-75 (169)
Observations	209	177

Note: Models include country and year effects

*p<0.1; **p<0.05; ***p<0.01

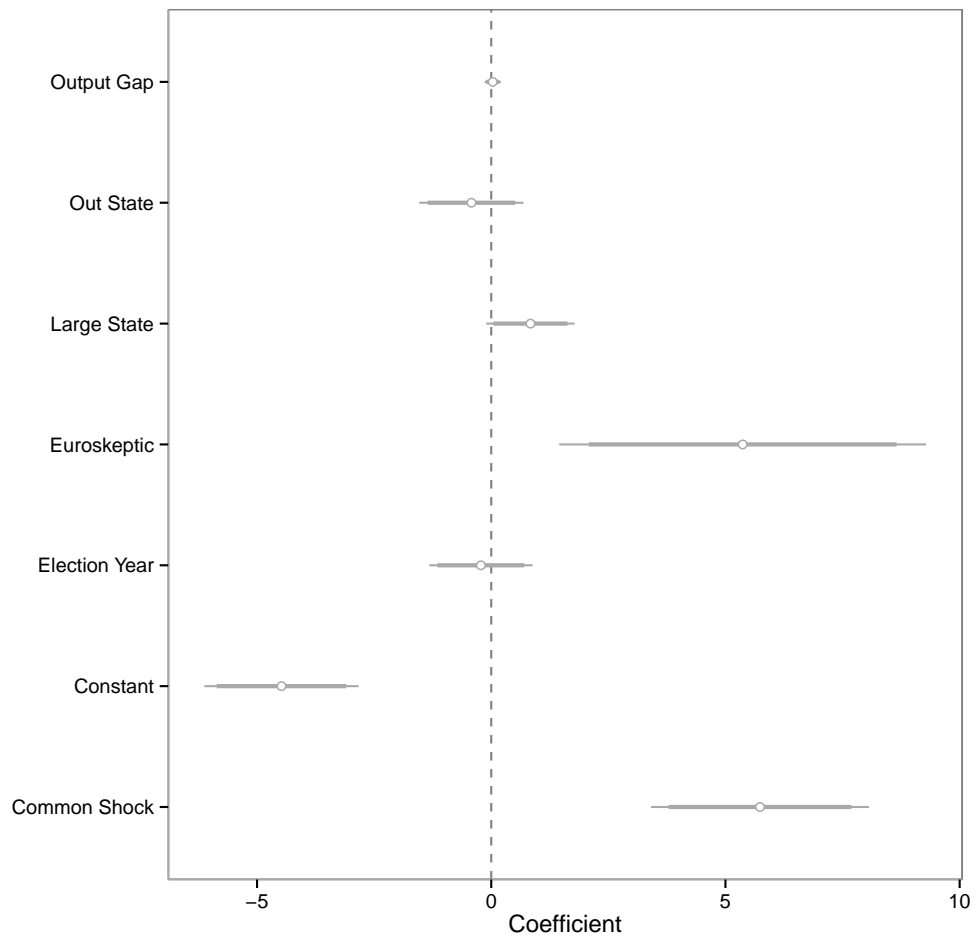


Figure 3: Political determinants of Weakening from mixed GLM with country and year random effects

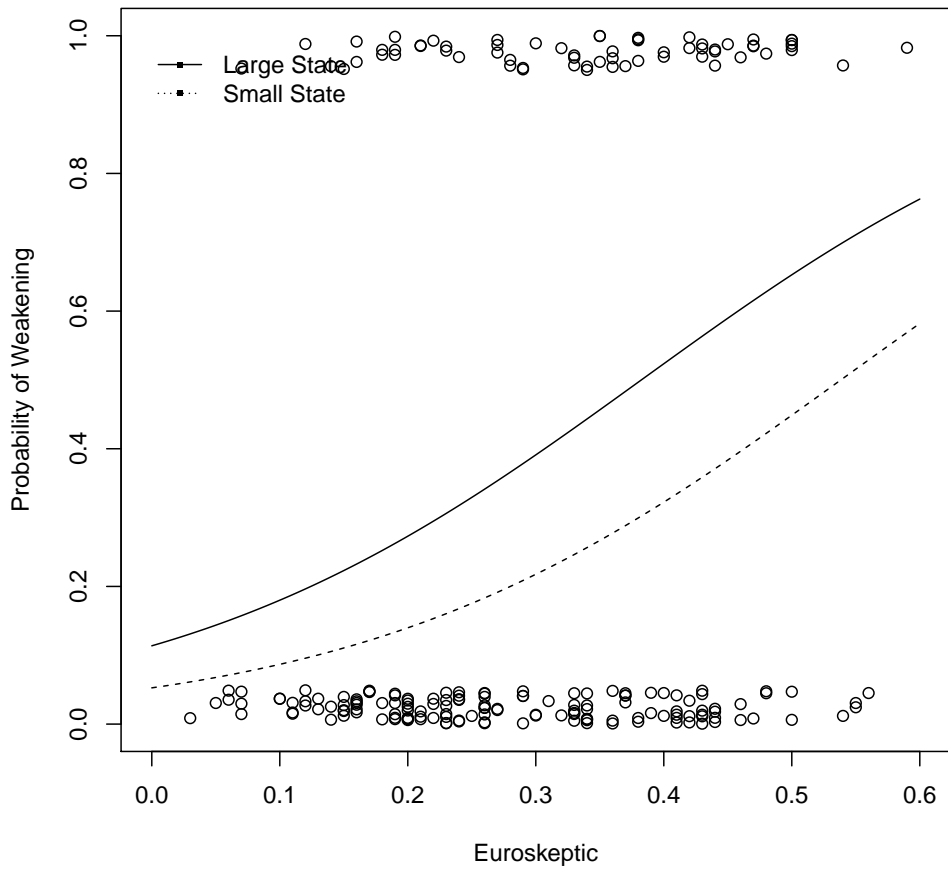


Figure 4: Effects of Euroskepticism on the probability of Weakening by Country Size

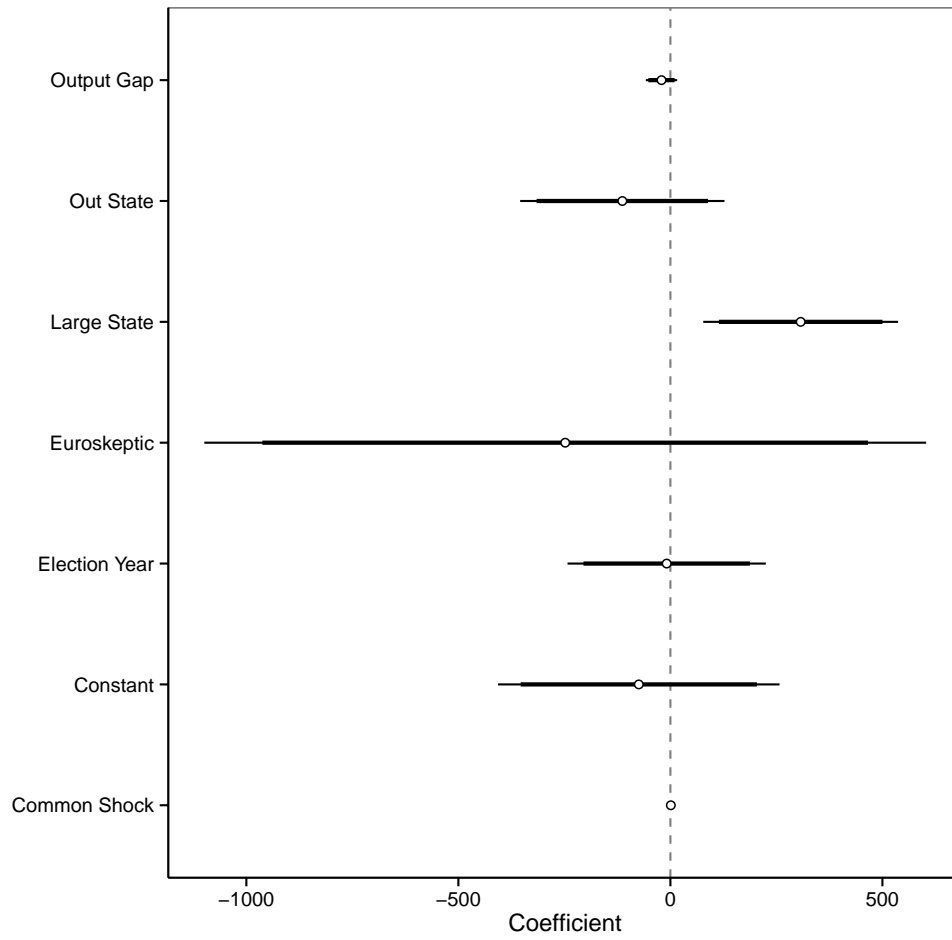


Figure 5: Political determinants of Editing from mixed LM with country and year random effects

Table 3: Effects of Financial Crisis on Weakening and Editing

	<i>Dependent variable:</i>	
	Weakening <i>generalized linear mixed-effects</i>	Editing <i>linear mixed-effects</i>
	(1)	(2)
<i>LargeState_t</i>	0.84* (0.48)	308*** (117)
<i>OutputGap_t</i>	0.01 (0.10)	-24 (21)
<i>OutState_t</i>	-0.44 (0.57)	-115 (123)
<i>Euroskeptic_t</i>	5.37*** (2.00)	-245 (435)
<i>ElectionYear_t</i>	-0.25 (0.57)	-10 (120)
<i>CommonShock_t</i>	5.68*** (1.20)	1*** (0.1)
<i>CrisisYear_t</i>	-0.23 (0.45)	-36 (104)
Intercept	-4.38*** (0.86)	-45 (191)
Observations	209	177

Note: Models include country and year effects

*p<0.1; **p<0.05; ***p<0.01

Appendix A

Coding of the Dependent Variables

We construct two dimensions using both human coding and textual analysis techniques. The human coding follows procedures discussed above. Documents are compared for substantive changes to the textual draft changes between the Commission's suggestions and the Council's amendments. This approach measures the direction of the recommendation, but does not measure the intensity or the magnitude of the changes.

In order to examine the intensity, we turn to techniques from computational social sciences. The machine learning follows the procedures summarized above. We first compile a list of textual changes using textual comparison software. We then use an add on package to extract the reported changes into a metadata file. Using two user made python scripts, we then extract the old text and new text and, running a levenshtein distance measure using the `nlk.metrics` package from NLTK, we then compute the metric for each edit. Finally, we then sum the total number of edits by country year and use this as a measure of country-year edit intensity.

Voting Rules in the Council of Ministers

The voting rules in the Council of Economic and Finance Ministers (ECOFIN) are remarkably complicated for the evaluation of stability and convergence programs. After 2009, they are also not the same rules as for other types of legislation. Below we summarize the rules and trace what the voting rules mean for a typical large state like France. In the first, 1998-2002, it is the Treaty of Amsterdam rules. Under this regime, the largest state has 10 votes and the smallest state has 2 votes. In order for the Commission text to pass, 62 of 87 votes are needed. States can also vote on their own cases. The second voting regime is from 2003-2008, or under the Treaty of Nice. This Treaty introduced a so-called "triple majority" as the definition of QMV. As Tsebelis and Yataganas notes, however, this had the effect of strengthening the Council vis-a-vis the Commission because it

increased the threshold needed for the Council to get to a qualified majority. In practice, this meant that it became somewhat easier for a large state to block a given Commission proposal. The next voting regime is 2009-2010, or Lisbon. Under these rules, the Commission only needs a qualified majority of the relevant set of Member States' votes. One major change was that a Member State did not get to vote on its own case. While this may seem to privilege the Commission, this was not the case for Eurozone states – only the euro states could vote on their cases, while everyone voted on the non-euro cases. This meant that the relevant number of votes to see a Commission proposal blocked for a large Eurozone country like France actually decreased rather than increased, so France actually had more power than before Lisbon. Finally, another regime was in place as of 2011. According to these rules a Member State needed a qualified majority to block what the Commission wants rather than QMV to pass it. It takes a few large states, or a combination of a large and a few small ones to block changes to the Commission's texts. Below are details on the coding:

I. 1998-May 2004

- i. A qualified majority required 62 of 87 votes to pass legislation. Member states had between 2 (Luxembourg) and 10 (France, Germany, Italy, United Kingdom) votes.
- ii. Note that member states potentially vote on their own cases. This means that the blocking minority is smaller for large states than for small. France, for example, needs only 15 more votes. Together with another large state and a medium state (e.g., Greece), it can block proposed text from the European Commission.

II. May-October 2004

- i. Transitional period. Member states continue to have between 2 and 10 votes, 88 out of 124 needed to pass legislation.

III. Nice Treaty: November 2004-November 2009

- i. Have 25 member states until December 2006, then 27 member states.
- ii. The vote must come from member states representing 62% of the population.
- iii. A triple majority is needed:
 - Qualified majority threshold of 72.3% of votes through December 2006, then 73.91%.
 - 232 of 321 votes from November 2004 until December 2006.
 - 255 of 345 votes from January 2007 until November 2009.
- iv. The vote must come from a majority of states.

For the purposes of this paper, we make calculations based on the first threshold only.

Note that member states potentially vote on their own cases. This means that the blocking minority is smaller for large states than for small. France, for example, needs 62 more votes. Together with two large states and a small state (e.g., Greece), it can block proposed text from the European Commission.

IV. Treaty of Lisbon: December 2009-December 2010

- i. The size of qualified majority threshold, population and number of state requirements technically remain the same, but there are important changes to voting for the purposes of this paper.
- ii. First, member states can no longer vote on their cases. The actual QMV is approximately 74% of the remaining member states. Second, everyone votes on the Commission documents for non-eurozone member states. Only eurozone states may vote on a eurozone member state.

iii. The practical consequence is as follows. A blocking minority for a euro out like the United Kingdom is 83 of 345 votes. A blocking minority for France is 47 of 180 votes. France, therefore, only needs to find one large state (e.g, Italy), one medium state (e.g., Greece) and one small state (e.g., Slovakia). Compared to the rules under Nice, therefore, while France can no longer vote on its case it needs fewer allies to block a Commission text.

V. Reverse Qualified Majority: from 2011

i. The same rules on voting are in place, but now it take a qualified majority to block a Commission text. This significantly increases the threshold. France now needs 137 votes. In 2011 Estonia joined the eurozone, so the required votes increased from 134 to 137.

Table A.1: Robustness Check: EU-15

	France and Germany		No France and Germany		Coming Election		Excessive Debts		Partisanship of PM or FM	
	Weakening	Editing	Weakening	Editing	Weakening	Editing	Weakening	Editing	Weakening	Editing
<i>FranceGermany_t</i>	1.20** (0.59)	257 (172)								
<i>LargeState_t</i>			0.36 (0.61)	307* (166)	0.85* (0.48)	300** (120)	0.92* (0.50)	224* (117)	1.09** (0.53)	263* (137)
<i>OutputGap_t</i>	0.03 (0.09)	-18 (19)	0.05 (0.10)	-23 (20)	0.04 (0.09)	-22 (19)			0.01 (0.10)	-30 (20)
<i>ExcessiveDeficit_t</i>							-0.09 (0.45)	208** (90)		
<i>PartyPM_t</i>									-1.59 (1.82)	-164 (400)
<i>PartyFM_t</i>									-0.42 (1.69)	152 (379)
<i>OutState_t</i>	-0.08 (0.56)	-57 (146)	-0.19 (0.57)	-119 (141)	-0.42 (0.57)	-117 (126)	-0.48 (0.60)	-49 (126)	-0.48 (0.61)	-101 (139)
<i>Euroskeptict_{t+1}</i>	5.37*** (1.91)	-74 (457)	5.59*** (1.99)	-200 (466)	5.40*** (1.99)	-246 (437)	5.24*** (1.95)	-282 (411)	5.36** (2.19)	-371 (483)
<i>ElectionYear_t</i>	-0.24 (0.56)	-12 (120)	0.23 (0.61)	19 (129)						
<i>ElectionYear_{t+1}</i>					0.20 (0.56)	-223* (116)	0.24 (0.57)	-208* (115)	0.39 (0.61)	-244** (123)
<i>CommonShock_t</i>	5.70*** (1.18)	1*** (0.1)	6.13*** (1.27)	1*** (0.1)	5.70*** (1.19)	1*** (0.1)	5.79*** (1.21)	1*** (0.1)	5.93*** (1.28)	1*** (0.1)
Intercept	-4.45*** (0.82)	-88 (180)	-4.80*** (0.89)	-74 (186)	-4.60*** (0.82)	-15 (169)	-4.57*** (0.81)	-84 (169)	-3.63*** (1.15)	17 (270)
Observations	209	177	181	154	209	177	207	175	193	161

Notes: *** p < .01; ** p < .05; * p < .1

Table A.2: Voting Proportion to Block Changes

	Weakening <i>generalized linear</i> <i>mixed-effects</i> Voting Block Proportion (1)	Editing <i>linear</i> <i>mixed-effects</i> (2)
<i>VoteBlocking_t</i>	-0.37 (1.69)	-87 (315)
<i>OutGap_t</i>	0.04 (0.10)	-17 (20)
<i>OutState_t</i>	-0.44 (0.62)	-115 (146)
<i>Euroskeptic_{t+1}</i>	6.37*** (2.02)	96 (452)
<i>ElectionYear_t</i>	-0.24 (0.56)	-11 (120)
<i>CommonShock_t</i>	5.69*** (1.21)	1*** (0.1)
Intercept	-4.44*** (1.01)	-61 (213)
Observations	209	177

Note: Models include country and year effects

*p<0.1; **p<0.05; ***p<0.01

Table A.3: Robustness Checks for the EU 27

	<i>Dependent variable:</i>	
	<i>Weakening generalized linear mixed-effects</i>	<i>Editing linear mixed-effects</i>
	(1)	(2)
<i>LargeState_t</i>	1.01** (0.48)	174* (101)
<i>OutputGap_t</i>	0.13** (0.06)	-7 (10)
<i>OutState_t</i>	0.16 (0.43)	-112 (78)
<i>Euroskeptic_{t+1}</i>	5.03*** (1.68)	376 (320)
<i>ElectionYear_t</i>	-0.12 (0.50)	-142 (98)
<i>CommonShock_{t+1}</i>	6.16*** (1.13)	1*** (0.1)
Intercept	-4.83*** (0.78)	-162 (142)
Observations	279	217

Note: Models include country and year effects *p<0.1; **p<0.05; ***p<0.01

References

- Begg, Iain, M Buti, Hendrik Enderlein, L Pench, Waltraud Schelkle and M. Weale. 2004. "Reforming Fiscal Policy Co-ordination under EMU: What Should Become of the Stability and Growth Pact?" *Journal of Common Market Studies* 42(5):1023–1059.
- Benoit, Kenneth and Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.
- Bernhard, William and David Leblang. 2016. "Sovereign Debt, Migration Pressure, and Government Survival." *Comparative Political Studies*. Advance online publication.
- Buti, Marco and Lucio. Pench. 2004. "Why Do Large Countries Flout the Stability Pact? And What Can Be Done About It?" *Journal of Common Market Studies* 42(5):1025–1032.
- Carrubba, Clifford. 2005. "Courts and Compliance in International Regulatory Regimes." *The Journal of Politics* 67(3):669–689.
- Chapman, Terrence, Songying Fang, Xin Li and Randall W. Stone. 2015. "Mixed Signals: IMF Lending and Capital Markets." *British Journal of Political Science* FirstView:1–21.
- Chaudoin, Stephen. 2014a. "Audience Features and the Strategic Timing of Trade Disputes." *International Organization* 68:877–911.
- Chaudoin, Stephen. 2014b. "Promises or Policies? An Experimental Analysis of International Agreements and Audience Reactions." *International Organization* 68:1–22.
- Copelovitch, Mark. 2010. "Master or Servant? Common Agency, Preference Heterogeneity, and the Political Economy of IMF Lending." *International Studies Quarterly* 54(1):49–77.
- Dreher, Axel and Roland Vaubel. 2004. "The Causes and Consequences of IMF Conditionality." *Emerging Markets Finance & Trade* 40(3):26–54.

- Fabbrini, Sergio. 2013. "Intergovernmentalism and Its Limits Assessing the European Unions Answer to the Euro Crisis." *Comparative Political Studies* 46(9):1003–1029.
- Fatas, Antonio and Ilian Mihov. 2003. "The Case for Restricting Fiscal Policy Discretion." *The Quarterly Journal of Economics* 118(4):1419–1447.
- Fischer, Jonas, Lars Jonung and Martin Larch. 2006. "101 Proposals to reform the Stability and Growth Pact. Why so many? A Survey." *Economic Papers* 267:3–48.
- Fjelstul, Joshua C and Clifford J Carrubba. 2015. "The Politics of International Oversight: Strategic Monitoring and Legal Compliance in the European Union." Working Paper.
- Franzese, Robert J. 2002. "Electoral and Partisan Cycles in Economic Policies and Outcomes." *Annual Review of Political Science* 5:369–421.
- Genovese, Federica, Pia Wassmann and Gerald Schneider. 2016. "The Eurotower Strikes Back: Crises, Adjustments and Europe's Austerity Protests." *Comparative Political Studies*. Advance online publication.
- Grieco, Joseph M., Christopher F. Gelpi and T. Camber Warren. 2009. "When Preferences and Commitments Collide: The Effect of Relative Partisan Shifts on International Treaty Compliance." *International Organization* 63(2):341–355.
- Hallerberg, Mark. 2004. *Domestic Budgets in a United Europe: Fiscal Governance from the End of Bretton Woods to EMU*. Ithaca, NY: Cornell University Press.
- Hallerberg, Mark, Benedicita Marzinotto and Guntram Wolff. 2015. "The Europeanization of Eurosceptic Parliaments? The Changing Role of National Parliaments Under the European Semester."
- Hallerberg, Mark and Joachim Wehner. 2015. "Crisis, Partisanship, and the Selection of Economic Policy-Makers." Manuscript, Hertie School of Governance and London School of Economics.

- Heipertz, Martin and Amy Verdun. 2010. *Ruling Europe: The Politics of the Stability and Growth Pact*. Cambridge University Press.
- Johns, Leslie. 2013. "Depth versus Rigidity in the Design of International Agreements." *Journal of Theoretical Politics* 26:468-495.
- Johnston, Alison, Bob Hancké and Suman Pant. 2014. "Comparative Institutional Advantage in the European Sovereign Debt Crisis." *Comparative Political Studies* 47:1771–1800.
- Koehler, Sebastian and Thomas König. 2015. "Fiscal Governance in the Eurozone: How Effectively Does the Stability and Growth Pact Limit Governmental Debt in the Euro Countries?" *Political Science Research and Methods* 3(02):329–351.
- König, Thomas and Brooke Luetgert. 2009. "Troubles with Transposition? Explaining Trends in Member-State Notification and the Delayed Transposition of EU Directives." *British Journal of Political Science* 39:163–194.
- König, Thomas and Lars Mäder. 2013. "Non-conformable, partial and conformable transposition: A competing risk analysis of the transposition process of directives in the EU15." *European Union Politics* 14(1):46–69.
- Kucik, Jeffrey and Eric Reinhardt. 2008. "Does Flexibility Promote Cooperation? An Application to the Global Trade Regime." *International Organization* 62:477–505.
- Mansfield, Edward D., Helen V. Milner and B. Peter Rosendorff. 2002. "Why Democracies Cooperate More: Electoral Control and International Trade Agreements." *International Organizations* 56(3):477–513.
- Morris, Richard, Hedwig Ongena and Ludger Schuknecht. 2006. "The Reform and Implementation of the Stability and Growth Pact." European Central Bank Occasional Paper No. 47, June.

- Richard, Stephanie J. and Teri L. Caraway. 2014. "International Negotiations in the Shadow of Elections." *International Organization* 68(3):701–720.
- Rosendorff, B. Peter and Helen V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4):pp. 829–857.
- Schelkle, Waltraud. 2005. "The Political Economy of Fiscal Policy Coordination in EMU: From Disciplinary Device to Insurance Arrangement." *Journal of Common Market Studies* 43(2):371–391.
- Schuknecht, Ludger. 2004. "EU Fiscal Rules: Issues and Lessons from Political Economy." European Central Bank Working Paper Number 421.
- Schuknecht, Ludger, Philippe Moutot, Philipp Rother and Jürgen Stark. 2011. "The Stability and Growth Pact: Crisis and Reform." European Central Bank Working Paper Number 129.
- Schure, Paul and Amy Verdun. 2008. "Legislative Bargaining in the European Union The Divide between Large and Small Member States." *European Union Politics* 9(4):459–486.
- Sörensen, Kenneth. 2007. "Distance Measures Based on the Edit Distance for Permutation-type representations." *Journal of Heuristics* 13(1):35–47.
- Staton, Jeffrey K. 2010. *Judicial Power and Strategic Communication in Mexico*. Cambridge University Press.
- Stone, Randall. 2011. *Controlling Institutions: International Organizations and the Global Economy*. Cambridge: Cambridge University Press.
- Thacker, Strom. 1999. "The High Politics of IMF Lending." *World Politics* 52(1):38–75.
- Thomson, Robert, Rene Torenvlied and Javier Arregio. 2007. "The Paradox of Compliance: Infringements and Delays in Transposing European Union Directives." *British Journal of Political Science* 37:685–709.

Vanberg, Georg. 2005. *The Politics of Constitutional Review in Germany*. Cambridge University Press.

Walter, Stefanie. 2016. "Crisis politics in Europe: Why austerity is easier to implement in some countries than in others." *Comparative Political Studies*. Advance online publication.