Incoherence in Regime Complexes:

A Sentiment Analysis of EU-IMF Surveillance.

Michael Breen

Dublin City University

Dermot Hodson

Birkbeck College, University of London

Manuela Moschella

Scuola Normale Superiore and the Centre for International Governance Innovation

Abstract: The proliferation of international institutions means that states can be subject to multiple, overlapping and potentially incoherent international obligations. The regime complexity literature draws attention to this problem but says little about its character and causes. This article investigates whether and why two key components of the international economic surveillance regime – the International Monetary Fund (IMF) and the European Union (EU) – impose conflicting obligations on the same states. Based on a comparative sentiment analysis of more than 400 surveillance documents and using differences in tone as a proxy for incoherence, our results show that the IMF was more pessimistic about member states' economic policies before the global financial crisis but less so thereafter. Our results suggest that differences in discretionary authority rather than the distribution of power drove such incoherence, with the EU's fiscal rules encouraging less pessimism before the global financial crisis and more pessimism thereafter.

Keywords: international institutions; EU; IMF; regime complexity; international economic surveillance; sentiment analysis

In February 2001, Ireland received a public rebuke over its economic policy from the European Union (EU). By pursuing expansionary budgetary policies in spite of 'an increasing extent of overheating', the EU's Council of Ministers (2001) concluded, the Irish government's macroeconomic policy mix was 'inappropriate'. Irish fiscal policy should have been 'neutral rather than expansionary', the International Monetary Fund (IMF) concluded in August 2001, while praising the government's tax policies (IMF, 2001). This difference of tone, though it should not be exaggerated, blunted the EU's efforts to hold Irish authorities to account. Ireland's finance minister Charlie McCreevy challenged the EU's recommendation and the IMF's analysis allowed him, as one commentator put it, 'to make the point to Brussels that it is the EU…not Ireland that is out of step on fiscal policy' (Clery, 2001).

This affair proved costly for Ireland, which allowed severe macroeconomic imbalances to accumulate, and for the EU, which was drawn into a sovereign debt crisis that the unwinding of imbalances in this member state and others eventually triggered. Nor was it an isolated case. Tensions between EU and IMF surveillance are common, a high-level study concludes, because of differences in views about the effectiveness of policy instruments and because EU rules may, on occasion, prevent EU member states from following IMF policy (Task Force on IMF Issues of the International Relations Committee of the European System of Central Banks, 2015).

Cases of this kind can be conceptualized as problems of incoherence in a regime complex. The concept of regime complexity was pioneered by Raustiala and Victor (2004: 279) to explore the 'array of partially overlapping and nonhierarchical institutions governing a particular issue-area'. Initially applied to the global governance of plant genetic resources, the last decade has seen empirical applications

of this regime complex approach to an array of issue areas, including trade, human rights and security (Faude and Gehring, 2017). Incoherence occurs when the elemental components of a regime complex – typically, but not exclusively, international institutions – produce 'multiple, possibly mutually contradictory, sets of regulations' (Gehring and Faude, 2014) that impose conflicting obligations on member states. Incoherence is problematic from both a systemic and individual state perspective. For instance, the existence of multiple international provisions provide state actors with forum shopping opportunities to escape regulation or obtain a favorable outcome (Alter and Meunier, 2009; Gomez-Mera and Molinari, 2014), thus threatening the credibility of international regimes (Drezner, 2009) and the effectiveness of international cooperation (Gehring and Faude, 2014). Furthermore, the existence of multiple international obligations may interfere in domestic political battles, by providing policy entrepreneurs new political resources for influencing and opening up domestic fights (Newmann and Posner, 2018).

In spite of these findings, considerable work remains to be done on how to measure incoherence in regime complexes and its determinants. In this article, we investigate the extent of – and factors driving – incoherence between two key players in the regime complex surrounding international economic surveillance: the IMF and the EU. The IMF is the premier global forum for international surveillance because of the universality of its membership and technical proficiency of its staff (Lombardi and Woods, 2008: 714). The EU goes well beyond the surveillance efforts of other regional organizations, in part, because of the governance challenges created by the euro (Savage and Verdun, 2016). This article provides the first, systematic test of two competing conjectures to explain incoherence in regime complexes. The first is that the elemental components of a regime impose different obligations on states where

these states wield different degrees of power within these institutions. The second expects incoherence to be associated with the differing degrees of discretionary authority exercised by the elemental components of the regime.

A major methodological innovation of this paper is to use sentiment analysis to measure incoherence in a regime complex. A method of quantitative text analysis with hitherto underexploited potential for students of international relations, sentiment analysis is used to derive a simple standardized measure of the tone of over 400 IMF and EU surveillance documents for the same group of EU member states over the period 1997-2014. Analyzing these documents as a whole rather than the recommendations within them, we treat differences in the tone of the language used within as a useful proxy for measuring policy coherence. Positive tone indicates the validation of member states' economic policies, whereas a negative tone brings with it an obligation to alter the status quo, albeit the kind of 'soft' obligation described by Simmons (2000) in her study of IMF surveillance.

Our results reveal the extent of incoherence within this regime complex. On average, the IMF is found to be more pessimistic than the EU in its assessment of EU member states' economic policies. The IMF was also significantly more pessimistic than the EU before the global financial crisis but it was less pessimistic thereafter. Using linear and panel regression analysis, we explain such incoherence not by differences in the distribution of power within the EU and IMF but by differences in the discretionary authority that the two organizations enjoy in performing surveillance. The influence of discretion is stronger for the EU, which appears to be more constrained than the IMF in its judgment of states' economic policies.

The first section of this article explores the problem of incoherence in regime complexes and sets out competing theoretical explanations for why incoherence

arises. The second discusses our decision to focus on IMF and EU economic surveillance and the methodology behind the sentiment analysis. The final section summarizes our results and discusses their significance for wider debates in international relations and EU studies.

The Problem of Incoherence in Regime Complexes

The problem of incoherence, as we define it, describes a situation in which the constitutive elements of a regime complex impose conflicting obligations on the same member states.[1] Obligation is not a binary variable though: it occurs on a sliding scale that ranges from binding legal rules to non-binding norms (Abbott et al., 2000: 404). We consequently conceive of incoherence as a spectrum that runs from a situation in which obligations imposed by institutions differ to a situation in which they are incompatible. In the extreme case of incompatibility, an international institution imposes obligations that require member states not to implement the obligations imposed by other institutions. In all other cases of incoherence, conflicting obligations give rise to a situation where states choose, prioritize or balance among multiple courses of action.

Why does incoherence arise? Blending insights from the literature on international regimes with the scholarship on international organizations, we advance two sets of explanations.

Power Distribution: Scholars of regime complexity assume that state power shapes the formation and persistence of regime complexes (Keohane and Victor, 2011: 8-9). We go a step further by conjecturing that differences in the distribution of power within a regime complex drive incoherence. Specifically, we expect states to be subject to differing obligations from the elemental components of a regime complex

in cases where such states wield different degrees of power within these elemental components.

In thinking about the impact of power on incoherence we start from two assumptions. First, all other things being equal, more powerful states carry greater weight within international organizations. We would expect the constituent elements of a regime to tread more softly with states that wield more decisional or economic power because the former is more politically and financially dependent on the latter or because more powerful states are less compliant. Second, we assume that a state's power is likely to vary across elemental components, *inter alia*, because the relative size of states varies according to their membership of international organizations and the governance structures of these bodies. On this basis, we expect a state that wields greater influence within one international institution than another will, all other things being equal, face more stringent obligations from the first than the second. This accounts, to some degree, for why US President Donald Trump faced greater pressure in the G20 than the G7 over his withdrawal from the Paris Agreement. Whereas the G7 summit in Charlevoix essentially took note of the difference between the United States and other members over the climate change accord (G7, 2018), all members except the United States signed up to a Climate and Energy Action Plan for Growth at the G20 summit in Taormina in a show of support for the Paris Agreement (G20, 2018). In view of these considerations, we formulate the following hypotheses:

H1: Decisional power: The greater the differences in decisional power between the elemental components of a regime complex, the greater the potential for incoherence.

H2: Economic power: The greater the differences in economic power between the elemental components of a regime complex, the greater the potential for incoherence.

Discretionary Authority: A second factor that may produce incoherence in a regime complex lies in the differing degrees of discretionary authority enjoyed by the elemental components of the regime. Viewed in principal-agent terms, state-principals can choose among varying rules and mechanisms to keep the international organization-agent accountable, identify agency slack and thus punish it including via a revision to the terms of delegation (Hawkins et al., 2006). Different control mechanisms are thus expected to shape the agent's performance, with more specific and intrusive rules designed to minimize the agent's discretion and thus the probability of deviant behavior.

Extending these insights to the case under investigation, it is thus plausible to expect that if two elemental components of the same regime complex differ in their level of discretion they are likely to differ in the obligations they place on states. This may happen because different levels of discretion allow one institution to act in a way that another might not be able to. For example, international institutions that act under rule-based delegation – the circumstance in which the state principal instructs the agent on how the agent has to pursue its mandate – are generally more constrained than institutions that operate under discretion-based delegation when it comes to imposing obligations on states (Hawkins et al., 2006: 27-8). While we see rules as determining discretion, the question of whether more or less discretion encourages international institutions to impose tighter or looser obligations on states is an empirical one. Dutiful agents will use their discretion to reinforce the interests of their principals, while overzealous agents may use discretion to impose more stringent

obligations (Tallberg, 2003). The more general theoretical claim that we wish to interrogate is that a regime complex in which the constituent elements are subject to different rules and hence different degrees of discretionary authority are more likely to impose incoherent obligations on states. Thus, for example, we might expect that the regime complex surrounding human rights in Europe will impose different obligations on states at times because the Court of Justice of the EU and the European Court of Human Rights operate under jurisdictions that differ in important respects (Douglas-Scott, 2006). This brings us to the following hypothesis:

H3. Design: The greater the differences in discretionary authority accorded to the elemental components of a regime complex, the greater the potential for incoherence.

Power and design certainly do not exhaust the list of potential drivers of incoherence. For instance, a burgeoning literature has drawn attention to recruitment patterns and professional networks to explain the (different) behavior of international institutions (Ban, 2015; Seabrooke and Tsingou, 2014). While it is plausible to foresee cases whereby incoherence arises because of variation in educational and professional backgrounds between officials in two institutions, we have not included in our empirical analysis a variable measuring the impact of professions for lack of comparable data between the IMF and the European Council and the Commission.

Data and measurement

Case selection

Today, most states have their economic policies monitored simultaneously by several intergovernmental organizations and other types of international institution. The G20, the G7, the Bank for International Settlements (BIS), the World Bank, the

Organization for Economic Cooperation and Development (OECD), the Financial Stability Board and the Association of Southeast Asian Nations are just some of the bodies that make up this regime complex. Since countries may be members of more than one of these international bodies, national governments are under the gaze of many eyes and subject to multiple recommendations that need not necessarily cohere. The interaction of other international institutions within this regime could be studied but the IMF and the EU provide an excellent starting point not only because their policy assessments are highly consequential for domestic political battles (Newmann and Posner, 2018) but also because the institutions are frequently at odds (Broome, Homolar and Kranke, 2017). It could be argued that comparing IMF and EU surveillance is problematic because the IMF's Articles of Agreement and EU Treaties impose differing degrees of obligation on states. Our response is that the obligations imposed by EU surveillance should not be overstated. To date, no financial penalties have been imposed on member states even though excessive deficits have been commonplace.[2] Likewise, following Simmons (2000), we see the IMF's Articles of Agreement as imposing low obligations on states rather than no obligation at all.

Sample

Our analysis equates differences in the tone of IMF and EU surveillance with differences in the degree of obligations imposed on member states and so policy incoherence. Where states are simultaneously subject to positive and negative evaluations of their economic policies by these institutions, this creates ambiguity about their obligations, a negative assessment encouraging policy change and a positive assessment favouring the status quo. A strongly negative assessment, likewise, brings more pressure than a weakly negative one. We focus on two key

surveillance documents: Public Information Notices (PINs)[3] and EU Council opinions on member states' stability programs. These documents are central to the conduct of surveillance. PINs summarize the Executive Boards' views on the Article IV reports prepared by Fund staff following country missions, which are usually conducted on an annual basis. Under the EU's Stability and Growth Pact, EU member states submit stability (or convergence) programs each year that set out medium-term budgetary plans in compliance with the EU's fiscal rules. It then falls to the Council, acting on a recommendation from the Commission to issue an opinion on these stability programs.

When the IMF and EU believe that governments are implementing 'good' policies, surveillance reports should contain less pessimistic language and, consequently, bring less pressure to change domestic economic policies. The corollary of this point is that pessimistic language places a higher degree of obligation on a state than positive language; whereas positive language validates the economic policies of the state under review, we contend, negative language implies that states must change their economic policies either because of bad policy choices or worsening economic conditions or both. Specifically, we compare PINs issued by the IMF Executive Board for the EU 15 between 1997-2014 with Council opinions on member states' stability programs during the same period. We focus on documents rather than recommendations contained within as this provides us with a richer view of surveillance documents. The combined corpus includes 442 cases (i.e. documents) and around 595,000 words.

To date, there has been no attempt that we know of to apply this methodology along these lines.[4] Hallerberg and Bridwell (2008) code differences between Commission recommendations concerning, and Council opinions on, stability and

convergence programs between 1998 and 2007. Baerg and Hallerberg (2016) use a similar research strategy to investigate the differences between European Commission assessments of member state economic programs and the assessment of the Council of Ministers. This analysis has the advantage of distinguishing between the nature and content of the obligation imposed on member states, but it deals with just two categories – differences between Commission and Council texts and the precision of Commission recommendations – and so loses much of the semantic nuance in EU surveillance efforts. Turning to the literature on the IMF, Fratzscher and Reynaud (2011) perform a content analysis of all PINs for a group of emerging markets over the period 1996-2007. Each document in this corpus is coded according to its 'favorableness' towards national economic policies. Manual coding of a concept like favorableness, or indeed incoherence, can provide a rich picture of the obligations placed on member states but the methodology raises serious concerns about reliability, as one coder's view of what constitutes favorable may differ considerably from another's, especially in documents that cover multiple policy areas (Krippendorff, 2012: 267-329). We redress this problem by combining an automatic content analysis of the words used by the IMF and EU and a sentiment analysis that distinguishes between words with different connotations.

Measuring Incoherence in Surveillance Reports

We measure incoherence by using differences in the tone of the policy assessments issued by the IMF and the EU. When the IMF and EU believe that governments are implementing 'good' policies, surveillance reports should contain less pessimistic language. Conversely, when they believe that governments are not doing so, they should contain more pessimistic language. There is virtually no informal language or

fundamental changes in the meaning of different words within our corpus. The range of words and expression is highly constrained, reflecting the formal nature of the process. It is very unusual to have comparable sets of documents with all of these attributes, which lends credibility to our analysis.

According to Loughran and McDonald (2011: 35), mainstream dictionaries fail miserably in financial contexts as the majority of words they identify as negative are not typically considered as such in the financial world. We use the sentiment dictionary that Loughran and McDonald developed to address this problem. When we apply their dictionary to our surveillance texts it generates word frequencies across the following categories: negative, positive, uncertain, strong, and weak words.[5] Figure 1 shows that the distribution is very similar across both types of surveillance, and that negative words are the most common type of word, followed by positive, uncertain, weak, and strong words.

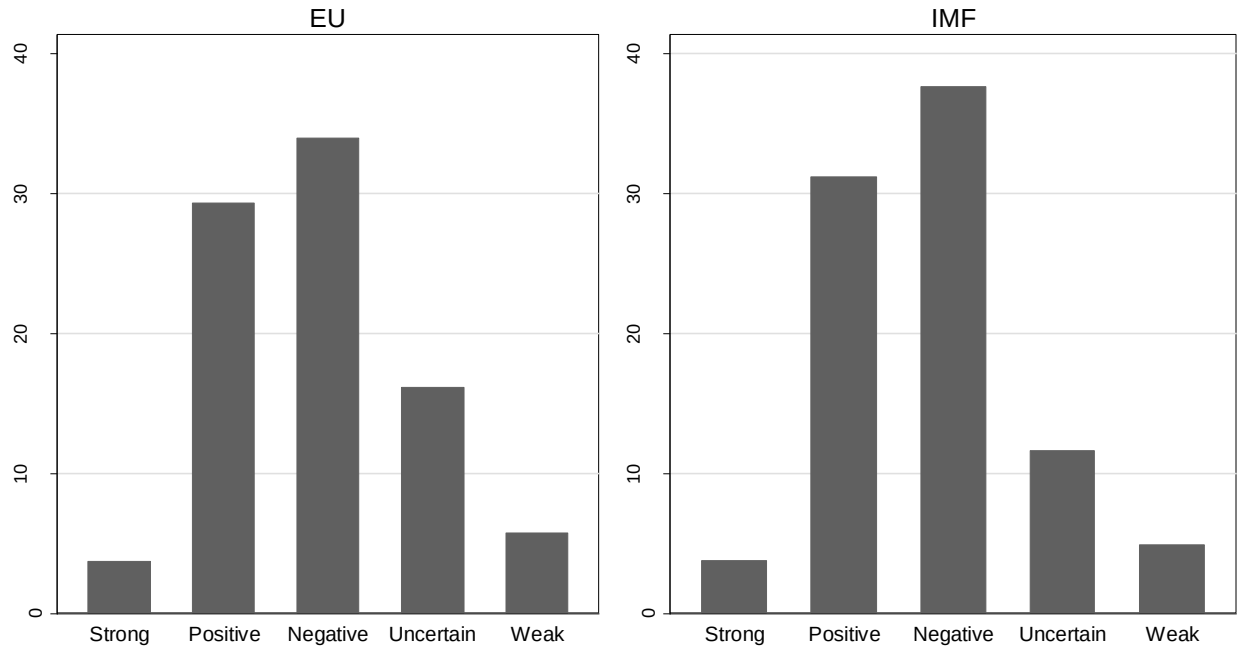

Figure 1. Categories of Language in International Economic Surveillance

Like Tetlock (2007), we use principal components factor analysis (PCA) to derive a single, replicable, measure of tone, which incorporates all word categories. The results, presented in Table 1, show that across both EU and IMF texts the first semantic component is heavily loaded with negative, weak, and uncertain types of language. By contrast, the second component is dominated by positive and strong language and contains much less weak and uncertain language. Therefore, it can be said that IMF and EU reports that score higher on the first factor contain more language associated with pessimistic evaluations.

Table 1. Principal Components of Surveillance Texts

EU sentiment					
	Comp1	Comp2	Comp3	Comp4	Comp5
Negative	0.50	0.21	-0.10	-0.77	-0.29
Positive	0.39	0.58	-0.53	0.41	0.22
Strong	0.41	0.30	0.83	0.20	0.07
Weak	0.46	-0.47	-0.10	0.41	-0.61
Uncertain	0.45	-0.54	-0.06	-0.11	0.69
IMF sentiment					
	Comp1	Comp2	Comp3	Comp4	Comp5
Negative	0.46	0.21	-0.78	0.34 -	0.02
Positive	0.35	0.51	0.00	0.77 -	0.01
Strong	0.19	0.64	0.52	0.51	0.10
Weak	0.54	-0.39	0.17	-0.03	0.71
Uncertain	0.56	-0.33	0.28	0.04 -	0.6946

Notes: 232 EU observations and 210 IMF observations from 1997-2014. Language Categories from Loughran and McDonald's(2011) Financial Sentiment Dictionary.

In Table 2, our summary statistics show that the average IMF assessment is slightly more pessimistic than its EU counterpart. However, the average masks substantial

variation over-time and across institutions.[6] Figure 2 below illustrates this variation, showing that the institutions are frequently at odds. Before the crisis, a typical IMF assessment contained 34 per cent more pessimistic language than an EU assessment. Since the crisis EU assessments have contained 53 per cent more pessimistic language.[7] Although the EU has moderated its language since 2010, it continues to be more pessimistic than the IMF. Overall, the lack of consensus in both good and bad economic times and sharply conflicting responses to the global financial crisis suggest that incoherence is a problem in international economic surveillance.

The type of language used in surveillance documents is best thought of as a controlled natural language, where grammar and vocabulary is restricted to remove ambiguity. This raises the possibility that the EU and IMF may use different procedures to control their use of sentimental words. However, an Independent-Samples t Test shows that the difference of means in the use of sentiment-charged words between the EU and the IMF is not different from zero, suggesting that we can compare the sentiment produced by both institutions.

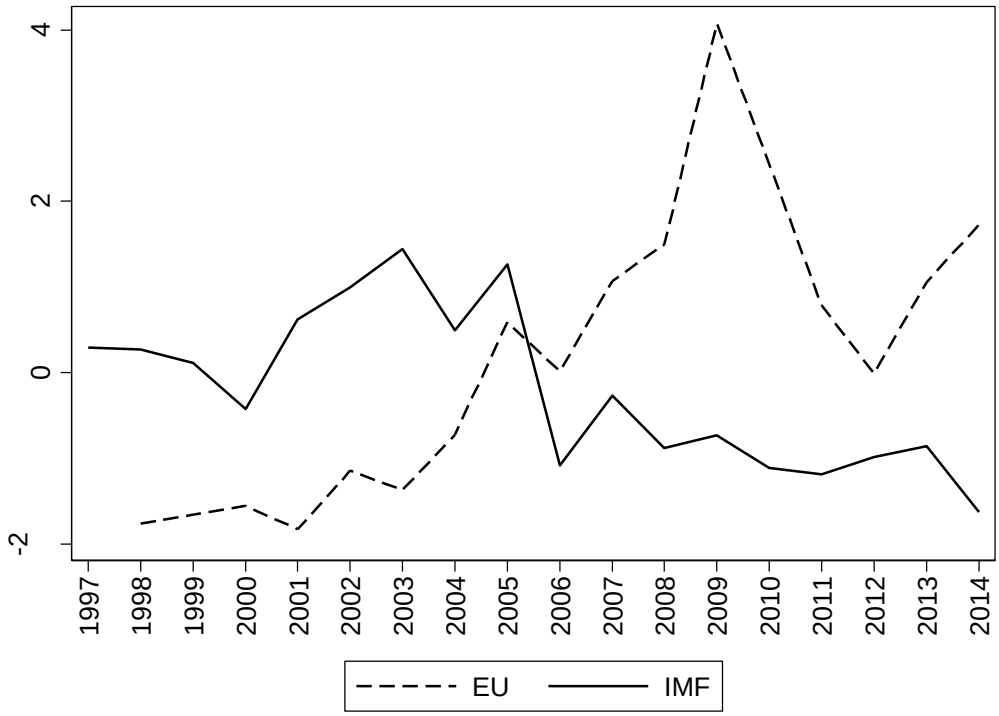

Figure 2. Median Pessimism in Surveillance Assessments

Explaining Incoherence

Having presented evidence on the degree of incoherence between IMF and EU surveillance, this section moves on to explore the factors that explain differences in sentiment. To measure differences in tone, we subtract the EU's pessimism score for a given country in a given year from the IMF's pessimism score pertaining to the same country in the same year. This exercise yields 156 observations of incoherence from 312 surveillance missions between 1997 and 2012. During this period, there were 397 episodes of EU-15 surveillance, including 206 EU Council reports and 191 IMF reports. We dropped 85 of these episodes because surveillance was delayed or deferred by at least one organization.[8] Table 2 provides summary statistics for this, and all of the variables in our empirical tests.

Table 2. Descriptive Statistics and Data Sources

	N	Mean	Min	Max	SD	Source

Incoherence	156	0.50	-6.54	6.44	2.61	(1)
EU tone	206	-0.16	-2.56	5.50	1.77	(1)
IMF tone	191	0.11	-2.41	4.23	1.47	(2)
Power Difference Index	156	-0.63	-2.49	0.64	1.03	(4)
EU Penrose Index	240	0.08	0.00	0.27	0.09	(4)
IMF Penrose Index	240	0.06	0.00	0.17	0.05	(4)
Ex. Deficit Pro.	240	0.25	0.00	1.00	0.43	(3)
WEO	215	0.03	-5.71	8.03	3.21	(3)
Fiscal space	240	0.92	-5.60	24.41	2.02	(1)
Δ Debt/GDP	240	0.01	-69.51	122.92	35.37	(3)
Δ Current account/GDP	240	-0.06	-18.44	19.01	6.92	(3)
Inflation (%)	240	0.03	-5.71	6.80	1.28	(3)
Growth (%)	240	2.28	-8.86	10.78	2.83	(3)
VIX Index	240	22.37	12.78	32.58	5.68	(7)
Spread (%)	233	0.62	-1.30	21.00	2.03	(6)
Austerity	224	0.35	0.00	1.00	0.48	(8)
GDP (log)	240	6.19	3.02	8.17	1.30	(5)
Deficit	230	0.03	-18.70	18.70	2.38	(3)
GDP forecast	157	2.58	-3.00	8.80	1.71	(5)

Notes: Variables collected for the period from 1997-2012. (1) EU Commission, (2) www.imf.org, (3) European Commission's AMECO Database, (4) Leech and Leech (2005) (5) World Economic Outlook, (6) OECD, (7) Chicago Board Options Exchange, (8) IMF database of action-based fiscal consolidation; Ağca and Igan (2013) for 2010 and 2011, authors' own calculations for 2012.

To test our conjectures concerning state power and discretionary authority, our independent variables are operationalized as follows:

Differences in the distribution of power: To measure decisional power we use Leech and Leech's (2005) Penrose index of the voting power of EU member states in the IMF Board of Governors. A similar index is constructed for qualified majority votes within the EU's Council of Ministers. We subtract the IMF and EU values of the Penrose indexes to create a single index which measures a state's ability to exert leverage over at least one organization within the regime complex. In further robustness tests we use the separate indexes. Economic power is measured using the natural log of gross domestic product (GDP). This is an appropriate measure of economic power in the EU-15, which are all considered wealthy advanced economies but differ considerably in economic size**.**

Differences in Discretionary Authority: To measure the degree of discretionary authority granted to the EU and IMF, we consider rules that inform the surveillance activities within each of the two institutions. The expectation here is that the degree of pessimism expressed by the EU or IMF changes as the rules underpinning their respective surveillance regimes begin to bite. An EU treaty protocol establishes two reference values in relation to which the European Commission and Council of Ministers is expected to judge excessive deficits: the ratio of the planned or actual government deficit to GDP of 3 per cent and the ratio of government debt to GDP of 60 per cent.[9] We code separate debt and deficit dummy variables to record years where a state has breached one of the critical thresholds. Temporary breaches of these thresholds are allowed in some cases. Therefore, we code an additional dummy variable which takes a value of '1' only when an EU member state is subject to an excessive deficit procedure.

To measure the degree of discretionary authority in the IMF, we use member countries' fiscal space, which is a key indicator that informs the Fund's surveillance efforts (Ostry et al., 2010). Simply put, fiscal space refers to the room that states have to undertake discretionary policy without raising concerns about the sustainability of public debt or market access. As a state's room for fiscal space narrows so too, we assume, will the IMF's discretionary authority. Fiscal space may not be as prescriptive as the EU's fiscal rules but the former puts the onus on the Fund to step up its interactions with the state in question. We operationalize this variable following Botev et al. (2016) by using two simple alternative measures of fiscal space. The first is the differential between interest rates and growth, which captures the extent to which the pace of economic growth can offset the impact of the interest rate on the debt ratio. The second is the ratio of government debt to tax revenue, which captures

the tax years that it would take to repay government debt. In both cases, an increase in the indicator can be read as a reduction in fiscal space.

Table 3 presents the findings from our linear regressions. The first column shows estimates that measure the impact of decisional and economic power. The second column includes our proxies for design features. The third column combines both power and institutional constraint variables in a single specification. The fourth column adds an interactive term to capture potential interactions between our rule indicators and the last column adds alternative rule indicators.

Table 3. The Determinants of Incoherence

	(1)	(2)	(3)	(4)	(5)
Power Index	-0.15		-0.35	-0.35**	-0.45
	(0.175)		(0.165)	(0.159)	(0.213)
GDP (log)	0.06		0.05	0.05	0.14
	(0.189)		(0.164)	(0.161)	(0.195)
Ex. Deficit Pro.		-2.52***	-2.72***	-2.86***	-2.36***
		(0.426)	(0.440)	(0.501)	(0.498)
Fiscal space *a*		0.37**	0.33**	0.21	0.33**
		(0.127)	(0.118)	(0.281)	(0.124)
EDP*Fiscal space *a*				0.16	
				(0.301)	
Deficit > 3					-1.11
					(0.662)
Debt > 60					0.16
					(0.351)
Fiscal space *b*					0.37
					(0.473)
Observations	156	156	156	156	156
R-squared	0.006	0.209	0.231	0.232	0.264

Robust standard errors clustered by country in parentheses
*** p<0.01, ** p<0.05

First, we find no support for the argument that differences in the distribution of power within the EU and IMF drive incoherence in the surveillance efforts of these organizations. The F-test associated with column 1 is low, indicating that our first model does not provide a better fit to the data than a model that contains no

independent variable. Thus, it would seem that state power within the surveillance regime complex has little to do with incoherence, as measured by differences in the tone of surveillance reports.

Our second major finding is that differences in discretionary authority drive incoherence between the IMF and EU (H3). Where the rules underpinning EU and IMF surveillance bite – and so reduce discretionary authority – we observe statistically significant changes in incoherence. When an excessive deficit procedure is in effect, thus requiring EU institutions to initiate disciplinary proceedings, we detect a large and statistically significant change. Specifically, it is associated with an 96.5 per cent decrease in standard deviations of our incoherence variable. We interpret the EDP being triggered as tending to push the EU towards a generally more pessimistic IMF. In practice, this suggests there is much potential for incoherence, as illustrated by the substantial divergence in sentiment in the years immediately following the global financial crisis when the EDP was triggered many times. In order to establish this we must examine the sentiment of each organisation separately, which we do in the next section.

Our fiscal space indicator is also statistically significant and correlated with an increase in our incoherence variable. A unit increase in this indicator corresponds with less fiscal space. Substantively, a one standard deviation increase in fiscal space is associated with 28 per cent of a standard deviation increase in our incoherence variable. We interpret a fall in fiscal space as pushing the IMF away from a more optimistic IMF. Like our previous finding regarding the EDP, we must test the sentiment of the IMF separately to establish that a decrease in fiscal space is met by more IMF pessimism.

Column 3, which combines our power and discretionary authority variables, largely supports these findings. In this column, the substantive and statistical significance of the excessive deficit procedure and fiscal space are broadly similar.Column 4 introduces an interaction term between our EU and IMF institutional constraints. When considered separately, these indicators demonstrate the additive effect of changes in discretionary authority on incoherence; when interacted, however, the interaction term captures the possibility that different combinations of discretionary authority have compound effects that are greater than the sum of their separate effects. For example, in cases where fiscal space is increasing while the excessive deficit procedure holds, the EU is constrained while the IMF enjoys discretion. In such cases, we would expect a synergy that produces more incoherence than if both organizations enjoyed the same amount of discretion. However, the interaction term in column 4 which captures this possibility is not statistically significant while the constitutive term – the excessive deficit procedure – is significant and substantively important. Thus, we find no evidence of compound effects in our sample. Rather, the evidence points towards the importance of the excessive deficit procedure as the most important factor in determining the extent to which the sentiment of both institutions is aligned or not, followed by fiscal space.

In column 5, we introduce alternative rule measures, including separate debt and deficit dummy variables to record years where a state has breached one of the critical thresholds, and an alternative measure of fiscal space. These alternative measures are not statistically significant and our earlier findings regarding the excessive deficit procedure and fiscal space hold.[10]

This section has established our headline finding, that incoherence is driven by changes is discretionary authority. It does not, however, deal with a key step in our

argument - our expectation that the degree of pessimism expressed by the EU or IMF changes as the rules underpinning their respective surveillance regimes begin to bite. We address this question in the next section.

EU and IMF panel data tests

To unpack these results, we explore the determinants of the pessimism scores produced separately by each organization using panel data techniques. These tests confirm that changes in the rules which limit discretionary authority are associated with greater pessimism and that the EU's excessive deficit, in particular, is likely to be responsible for much of the divergence between EU and IMF assessments. The data consists of separate EU and IMF unbalanced panels of 15 countries between 1997 and 2012. There were 397 episodes of surveillance during this period, including 206 EU assessments and 191 IMF assessments. We are missing data on some independent variables, described in Table 2, which reduces our IMF sample to 171 assessments across 14 countries.

For our EU sample, we use the EU Penrose Index as our power variable and the excessive deficit procedure as our discretionary authority variable. For our IMF sample, we use the IMF Penrose Index and our two measures of fiscal space. Both of these latter measures are narrow, theoretically-informed proxies for the importance of discretionary authority. However, the literature also suggests that IMF assessments are broad and flexible and that they change dynamically (see Heller, 2003). Therefore, we complement our narrow measures with a broad measure of macroeconomic performance based on the World Economic Outlook (WEO) database, the IMF's own formal assessment of economic performance. To construct our broad indicator, we aggregate all of the WEO's 43 sub-indicators. First, we standardize the WEO to

obtain Z scores for each of the 43 sub-indicators. Then we extract the first component using principal components analysis. The first component is loaded primarily with positive developments in the economy, including growth, government revenue, and employment. Therefore, an increase in the WEO measure can be read as a broad improvement in a country under surveillance, according to the IMF's own methodology for measuring economic performance.

Finally, in both samples we control for general macroeconomic conditions, including consumer price inflation, GDP growth, the fiscal deficit, government debt, the current account balance, and the VIX Volatility Index, a measure of market expectations of near-term stock market volatility.

We detected heteroskedasticity using the Breusch-Pagan/Cook Weisberg test and therefore estimate Prais-Winsten models with panel-corrected standard errors. The Wooldridge test indicates autocorrelation, so we use the AR1 correction. We do not include country fixed effects because of the time invariant nature of some of our key independent variables.[11] The Wald test indicates no need to include dummy variables for each year.[12] Inspection of the correlation matrix and variance inflation factors show that multicollinearity is not a cause for concern.

Table 4 presents our findings. Columns 1-5 relate to EU sentiment. The first column presents our base specification, which includes proxies for decisional power, discretionary authority, and economic controls. The next column adds a variable measuring whether a state has exceeded the 3 percent deficit threshold. Column 3 adds a variable measuring whether a state has exceeded the 60 per cent debt threshold. Column 4 substitutes the log of GDP for the Penrose Index. Column 5 introduces the fiscal deficit to our base specification, reducing our sample by 9 observations.

Columns 6-9 relate to IMF sentiment. Column 6 is the base specification,

which includes proxies for decisional power, discretionary authority, and economic controls. Column 7 adds our measure of variation in the IMF's own WEO Database. Column 8 substitutes the log of GDP for the Penrose Index. Column 9 adds the fiscal deficit, reducing our sample by 7 observations.

Our results show that several factors are driving pessimism in this regime complex. The economic control variables suggest that the EU attaches greater weight to past economic performance than the IMF. A one standard deviation increase in economic performance, as measured by growth, is associated with a 0.24 standard deviation reduction in EU pessimism, according to Model 1. The current account balance and government debt are also associated with small spikes in pessimistic language.[13] Neither the fiscal deficit as a percentage of GDP, inflation, nor global economic conditions as captured by the VIX index are associated with the tone of EU or IMF assessments.

In line with our earlier findings, there is no support for the argument that differences in the power of member states matter. The EU Penrose Index and the log of GDP are statistically insignificant. The IMF Penrose Index is significant but, contrary to expectations, an increase is associated with more pessimistic language in IMF surveillance. All other things being equal, EU member states with high degrees of economic and decision making power do not receive more favorable assessments under international economic surveillance, with the Fund but not the EU more likely to take a tougher line against such countries.

Table 4 is consistent with our previous findings which use the difference in tone as the dependent variable. The triggering of the excessive deficit procedure is associated with a large and statistically significant spike in pessimistic language. The effect is large in the first year and persists for the time that a member state is under the

procedure. Substantively, it is associated with a 0.94 standard deviation increase in pessimistic language, all else being equal. The evidence linking IMF surveillance to discretionary authority is weaker but not inconsistent with our earlier findings regarding discretionary authority. Neither of our primary or secondary indicators of fiscal space are statistically significant, suggesting that the IMF may take a less rigid interpretation of surveillance.[14] However, our measure of variation in the WEO indicator is statistically significant and correlated with a reduction in pessimistic language (See Column 7). This measure captures primarily good economic performance across multiple indicators as defined by the IMF's own criteria for measuring performance.[15] Substantively, an increase in this indicator is associated with a 0.41 standard deviation decrease in pessimistic language. By contrast, as the indicator falls, the IMF has less discretion and tends to use more pessimistic language. While dwarfed by the substantive effect of the excessive deficit procedure, the substantive effect of the WEO indicator increases by almost 50 per cent when we drop our economic control variables from the specification.[16]

Finally, we performed a series of further checks, available as part of the replication materials and explained in an online appendix, which reinforce the robustness of our findings.

Table 4. The Determinants of Pessimism in EU and IMF surveillance

	(1) EU	(2) EU	(3) EU	(4) EU	(5) EU	(6) IMF	(7) IMF	(8) IMF	(9) IMF
Δ Debt	0.00**	0.00**	0.00	0.00**	0.00	0.00	0.00	0.00	0.00
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Δ Current account	0.03***	0.03***	0.03***	0.03***	0.03***	0.01	0.01	0.01	0.01
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.015)	(0.015)
Δ Inflation	-0.05	-0.05	-0.06	-0.05	-0.05	0.08	0.09	0.09	0.07
	(0.039)	(0.039)	(0.041)	(0.039)	(0.052)	(0.081)	(0.079)	(0.082)	(0.081)
Growth t-1	-0.15**	-0.14**	-0.15**	-0.15**	-0.15	0.13**	0.06	0.10	0.13**
	(0.068)	(0.068)	(0.067)	(0.068)	(0.089)	(0.063)	(0.064)	(0.067)	(0.064)
VIX	0.05	0.05	0.05	0.05	0.05	0.02	0.01	0.02	0.02
	(0.035)	(0.036)	(0.035)	(0.035)	(0.046)	(0.031)	(0.027)	(0.031)	(0.031)
Penrose Index	1.97	1.84	2.06		1.28	7.45***	15.86***		8.21***
	(1.766)	(1.708)	(1.727)		(1.966)	(2.146)	(3.988)		(2.263)
Ex. Deficit Pro.	1.57***	1.55***	1.57***	1.59***	1.74***				
	(0.346)	(0.342)	(0.344)	(0.344)	(0.458)				
Deficit > 3		0.22							
		(0.226)							
Debt > 60			0.11						
			(0.147)						
GDP (log)				0.11				0.33	
				(0.110)				(0.255)	
Deficit					-0.11				-0.05
					(0.062)				(0.067)
Fiscal space						0.01	0.01	0.02	0.01
						(0.035)	(0.030)	(0.038)	(0.035)
WEO							-0.19**	-0.07	
							(0.075)	(0.106)	
Observations	206	206	206	206	197	171	171	171	164
R-squared	0.305	0.312	0.307	0.305	0.367	0.108	0.147	0.074	0.117
No. of countries	15	15	15	15	15	14	14	14	14

Panel-corrected standard errors in parentheses (AR1)
***p<0.01,**p<0.05

Conclusion

Over the last two decades, the field of international relations has moved on from the study of individual institutions and regimes in isolation to study the interaction between them. Regime complexity is at the cutting edge of efforts to understand this new politics of global governance although further work needs to be done to understand the character and causes of incoherence within regime complexes. This article has examined the extent of – and factors driving – incoherence between two key players in the regime complex surrounding international economic surveillance: the IMF and the EU. Our findings – based on a sentiment analysis of more than 400 surveillance documents between 1997 and 2014 – points to the presence of incoherence in this regime complex. Our results show that the IMF was more pessimistic on average than the EU but that the EU was less pessimistic in it surveillance efforts before the global financial crisis hit and more so thereafter. Such incoherence is explained not by organizational politics but by technocratic rules. We find little evidence that differences in the distribution of power within the elemental components of this regime was decisive. Neither the IMF nor the EU treaded softly when dealing with economically or politically powerful states. Instead, our results point towards differences in discretionary authority as the key driver of incoherence. When the rules underpinning EU and IMF fiscal surveillance bite and reduce these institutions' room for discretion, each institution tends to be more pessimistic. But the two institutions are responding to different rules with differing degrees of intensity and from different starting points, leading to incoherent assessments of member states' economic policies.

These results shed new light on the EU and IMF as institutions as well as the interaction between them. First, our findings challenge the idea that these institutions

are dominated by large states (see Hallerberg and Bridwell, 2008). Second, our analysis, confirms that rules matter for EU and IMF surveillance while encouraging a more nuanced interpretation thereof. The Stability and Growth Pact emerges in our account not as a rigid instrument of austerity (Scharpf, 2015) but one that swings between more and less pessimistic assessments of member states' economic policies. The IMF, though critical of the EU's approach to surveillance, is more beholden to rules than it looks (Annett et al, 2005).

The focus of this article is on the causes rather than the consequences of incoherence in regime complexes but our findings invite reflection about the latter. The lack of coherence between EU and IMF surveillance, the regime complexity literature conjectures, will impede the effectiveness of economic surveillance by making it easier for EU member states to wriggle out of their obligations. Precisely what EU member states could and should have done prior to the global financial crisis is a matter of economic judgement but the IMF's greater pessimism during this period lends weight to the view that it was the more vigilant economic watchdog and that the EU's fiscal rules, as originally conceived, made the EU more rather than less attentive to the policy mistakes made by member states at this time. Since the global financial crisis hit, the EU has become bad cop to the IMF's good cop. It is too soon to say whether this constitutes a case of over-vigilance by the EU or under-vigilance by the IMF but what is clear is that the problem of incoherence in relation to international economic surveillance persists.

Our analysis is also relevant for wider debates in EU studies and international relations. The methodology and findings of this article respond to Henning's call for a deeper study of regime complexity and institutional interaction, particularly as it relates to the role of non-European forces in European integration and governance

(Henning, 2017: 28 and 258). It also chimes with ongoing debates about the importance of state size for influence in EU policy-making (Thorhallsson, 2017) and the constraining power of institutions in this domain (Heipertz and Verdun, 2011). Turning to the wider international relations literature, our findings may help to refine explanations of why some regime complexes enhance the effectiveness of international cooperation while others do not (see also Gomez-Mera, 2016). A central message of this article is that assumptions about the degree of coherence within a regime complex is a matter for empirical investigation rather than something to assume *a priori* or explore through case studies alone. Finally, the methods employed in this article have wider application for EU and international relations scholars alike. While content analysis is now part of the political science tool kit, our analysis demonstrates the potential of sentiment analysis to generate new empirical insights concerning textually rich areas of international cooperation such as economic surveillance.

References

Abbott, K. W., Keohane, R.O., Moravcsik, A., Slaughter A. M. and Snidal D. (2000) The Concept of Legalization. *International Organization* 54 (03):401-19.

Agca, S., and Igan, D. (2013) Fiscal consolidation and the cost of credit: Evidence from syndicated loans, IMF Working Paper 13(36).

Alter, K. J. & Meunier, S. (2009). The Politics of International Regime Complexity. *Perspectives on Politics* 7(1):13–24.

Annett, A.M., Decressin, J. and Deppler, M. (2005) Reforming the Stability and Growth Pact, IMF Policy Discussion Paper No. 05/2.

Ban, C. (2015) Austerity Versus Stimulus? Understanding Fiscal Policy Change at the International Monetary Fund since the Great Recession. *Governance: An International Journal of Policy Administration and Institutions* 28(2):167–83.

Baerg, N. R. and Hallerberg, M. (2016) Explaining Instability in the Stability and Growth Pact The Contribution of Member State Power and Euroskepticism to the Euro Crisis. *Comparative Political Studies* 49(7): 968-1009.

Botev, J. et al. (2016) A Re-assessment of Fiscal Space in OECD Countries. *OECD Economic Working Papers ECO/WKP* (2016)76

Broome, A., Homolar, A., and Kranke, M. (2017) 'Bad science: International organizations and the indirect power of global benchmarking', *European Journal of International Relations*.

Clery, C. (2001) IMF Report to Praise Irish Tax Reforms. *Irish Times*, 3 August.

Council of Ministers of the EU (2001) 2329th Council meeting (ECOFIN) Brussels, 12 February 5696/01 (Presse 35).

Doleys, T. (2009) Incomplete contracting, commission discretion and the origins of EU merger control. *Journal of Common Market Studies* 47(3): 483-506.

Drezner, D. W. (2009) The Power and Peril of International Regime Complexity. *Perspectives on Politics* 7(1):65–70.

Faude, B. and Gehring, T. (2017) 'Regime complexes as governance systems' in W. Sandholtz and C.A. Whytock (eds) Research Handbook on the Politics of International Law (Cheltenham: Edward Elgar): 176-206.

Fratzscher, M. and Reynaud, J. (2011) IMF Surveillance and Financial Markets–a Political Economy Analysis. *European Journal of Political Economy* 27(3):405–22.

G7 (2018) The Charlevoix G7 Summit Communique. Full text available at: https://g7.gc.ca/en/official-documents/charlevoix-g7-summit-communique/

G20 (2017) G20 Hamburg Climate and Energy Action Plan for Growth. Full text available at: http://unepinquiry.org/wp-content/uploads/2017/07/Climate_and_Energy_Action_Plan_for_Growth.pdf

Gehring, T. and Faude, B. (2014) A Theory of Emerging Order within Institutional Complexes: How Competition among Regulatory International Institutions Leads to Institutional Adaptation and Division of Labor. *The Review of International Organizations* 9(4): 471–98.

——— (2013) The Dynamics of Regime Complexes: Microfoundations and Systemic Effects. *Global Governance* 19(1):119–30.

Gómez-Mera, L. (2016) Regime complexity and global governance: The case of trafficking in persons. *European Journal of International Relations* 22 (3):566-95.

Hallerberg, M. and Bridwell, J. (2008) Fiscal Policy Co-ordination and Discipline: The Crisis of the Stability and Growth Pact and Domestic Fiscal Regimes. In K. Dyson, ed., *The Euro at Ten: Europeanization, Power and Convergence*. Oxford: Oxford University Press: 69–86.

Hawkins, D. G., Lake, D.A., Nielson, D.L. and Tierney, M.J., eds, (2006) Delegation under Anarchy: States, International Organizations, and Principal-Agent Theory. In *Delegation and Agency in International Organizations*, Darren G. Hawkins, David A. Lake, Daniel L. Nielson and Michael J. Tierney, eds., pp. 1–38. Cambridge: Cambridge Univeristy Press.

Heipertz, M. and Verdun, A. (2011) *Ruling Europe : The Politics of the Stability and Growth Pact*. Cambridge: Cambridge University Press.

Heller, P.S. (2003) Considering the IMF's Perspective on a"Sound Fiscal Policy'. Public Finance Analysis 59, 1: 141-161.

Henning, R. C. (2017) *Tangled Governance: International Regime Complexity, the Troika, and the Euro Crisis*. Oxford: Oxford University Press.

IMF (2001) Ireland: 2001 Article IV Consultation – Staff Report and Public Information Notice on the Executive Board Discussion. IMF Country Report No. 1/139.

Keohane, R. O., and Victor D. G. (2011) The Regime Complex for Climate Change. *Perspectives on Politics* 9(1):7–23.

Krippendorff, K. (2012) *Content Analysis; An Introduction to its Methodology.* 3rd Edition. Thousand Oaks, CA: Sage.

Leech, D. and Leech, R. (2005) Voting Power Implications of a Unified European Representation at the International Monetary Fund. *LSE Research Online*.

Lombardi, D. and Woods, N. (2008) T*he politics of influence: An analysis of IMF surveillance*. Review of International Political Economy 15(5): 711-739.
Loughran, T. and McDonald, B. (2011) When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66(1):35–65.
Newmann, A.L., and Posner, E. (2018) *Voluntary Disruptions: International Soft Law, Finance, and Power* Oxford: Oxford University Press.
Ostry, J. D., Ghosh, A.R., Kim, J.I., Mahvash, S.Q. (2010) 'Fiscal space' *IMF Staff Position Note 10/11*.
Raustiala, K. and Victor, D. G. (2004) The Regime Complex for Plant Genetic Resources. *International Organization* 58(2): 277–309.
Savage, J. D., and Verdun, A. (2016) Strengthening the European Commission's budgetary and economic surveillance capacity since Greece and the euro area crisis: a study of five Directorates-General *Journal of European Public Policy*, 23(1), 101-118.
Scharpf, F. W. (2015). Political legitimacy in a non-optimal currency area. Democratic politics in a European Union under stress. In O. Cramme and S. Binzer Hobolt (eds) *Democratic Politics in a European Union Under Stress* (Oxford: Oxford University Press): 19-47.
Seabrooke, L. and Tsingou E. (2014) 'Distinctions, affiliations, and professional knowledge in financial reform expert groups' *Journal of European Public Policy* 21(3): 389-407.
Simmons, B. A. (2000) The legalization of international monetary affairs. I*nternational Organization*, 54(3), 573-602.
Tallberg, J. (2003) European governance and supranational institutions: making states comply Abindgon: Routledge.
Task Force on IMF Issues of the International Relations Committee of the European System of Central Banks (2015) 'IMF Surveillance in Europe (Frankfurt A.M.: IMF).
Tetlock, P. C. (2007) Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance* 62(3):1139–68.
Thorhallsson, B. (2017) *The role of small states in the European Union*. London: Routledge.

Notes

[1] Our definition builds on the definition of coherence provided by Keohane and Victor (2011, 16) according to whom "A regime whose components are compatible and mutually reinforcing is coherent."

[2] The closest EU policy-makers have come to imposing such penalties was in August 2016 when the Council of Ministers cancelled proposed fines against Portugal and Spain.

[3] Since July 2013, PINs have been replaced by a Press Release that serves much the same purpose.

[4] One reason is that, before 1997, the number of published surveillance documents was limited.

[5] The whole dictionary (which includes 85,131 words) can be downloaded from Loughran and McDonald's website at http://www3.nd.edu/~mcdonald/Word_Lists_files/LoughranMcDonald_MasterDictionary_2014.xlsx As per this dictionary, positive words include 'always' and 'definitely'. Weak words include 'almost' and 'might'. Uncertain words also capture no sentiment but they are neither positive nor negative.

[6] To calculate this figure we pool all of our documents and perform a principal component analysis of sentiment across our categories of language. A separate principal components analysis on each institution's documents produces an almost identical figure.
[7] This finding resonates with Baerg and Hallerberg (2016: 975), who find that during the euro crisis the Council's tendency to moderate the Commission's assessment of Stability and Convergence Programme reduced.
[8] The EU reports begin in 1998 and the IMF reports begin in 1997. There were 19 country-years where EU surveillance was delayed or deferred between 1998 and 2012, and 49 country-years between 1997 and 2012 where IMF surveillance was delayed or deferred.
[9] Protocol (No 12) on the Excessive Deficit Procedure, Treaty on the Functioning of the European Union.
[10] Although we do not theorise interactions between power and discretionary authority, we tested models that involve interactions between our power and rules variables. In all cases, the interaction terms were not statistically significant and had weak explanatory power.
[11] The Hausman test rejects the use of fixed effects in a specification that excludes time invariant independent variables.
[12] However, our findings are robust to the inclusion of time dummies.
[13] A one standard deviation increase in the current account balance as a percentage of GDP is associated with a 0.08 standard deviation increase in pessimistic language.
[14] Our secondary measure of fiscal space is not displayed in Table 3. Results available in the replication set.
[15] More specifically, the WEO variable is the first component from principal components analysis (PCA) of the WEO dataset. The first component explains 23.8 per cent of the variation in this dataset. Approximately 10 of the 43 WEO variables contribute the most to this component particularly growth, government revenue and employment.
[16] Our measure of fiscal space is not robust to the substitution of the Penrose index by the log of GDP (see Column 8). Models without economic controls are available in the replication set.