




RESEARCH ARTICLE

# The politics of creditworthiness: political and policy commentary in sovereign credit rating reports

Zsófia Barta<sup>1</sup>  and Kristin Makszin<sup>2</sup>

<sup>1</sup>Department of Political Science, University of Albany, USA E-mail: [zbarta@albany.edu](mailto:zbarta@albany.edu) and <sup>2</sup>Leiden University College The Hague, Netherlands E-mail: [k.m.makszin@luc.leidenuniv.nl](mailto:k.m.makszin@luc.leidenuniv.nl)

(Received 01 January 2019; revised 03 December 2019; accepted 06 January 2020; first published online 06 February 2020)

## Abstract

How much do politics and politically sensitive policy choices matter for sovereign credit ratings? We contend that while policy is consistently important for rating decisions, attention to politics varies with perceived uncertainty. Quantitatively analysing the text of 635 sovereign rating reports issued by Standard and Poor's (S&P) between 1999 and 2012 for 40 European countries, we find that S&P scrutinises *policy* with similar intensity across countries, but *political* scrutiny was less intense in developed countries and prospective European Union members (categories formerly associated with lower uncertainty) than in emerging countries until the crisis dispelled illusions of lower uncertainty in these categories. Our findings nuance the common notion that financial market actors allow countries perceived to belong to low-risk categories more “room-to-move” in their political and policy choices, by showing that in rating decisions such permissiveness only applied to politics – but not to policy – and it ended with the global financial crisis.

**Keywords:** content analysis; “golden straightjacket”; political and policy constraints; sovereign credit ratings; sovereign rating reports; uncertainty and risk

## Introduction

To what extent do credit rating agencies (CRAs) focus on politics and policy in a country when they pass judgement on that country's creditworthiness? How heavily do they weigh political factors like elections, intragovernmental dynamics or relations with unions when they decide to raise, lower or affirm a country's ratings? How much emphasis do they place on politically sensitive policy issues like welfare reform, liberalisation or privatisation? The answers to these questions are important for our understanding of the way countries get rewarded or penalised for their political choices and policy decisions by financial market actors. As gatekeepers to financial markets, CRAs influence the price of government debt through the ratings

they award (IMF 2010; Afonso et al. 2015; Barta and Johnston 2018 and 2019<sup>1</sup>). What is more, upon each rating decision, CRAs publicly explain – in rating reports published in the financial press – what factors influenced their decision *and* what their future decisions will likely be if current domestic trends continue. By doing so, CRAs’ communications send messages to domestic decision-makers about the market consequences of their choices in ways that few (if any) other market actors do, explicitly assigning costs and benefits (in terms of negative and positive changes in credit ratings) to various domestic developments on which they comment. Therefore, exploring rating reports to see whether they go beyond the analysis of economic and fiscal developments to comment on politics and politically sensitive policy choices helps us better understand the extent to which CRAs assign pecuniary costs and benefits to issues most closely associated with democratic choice.

By analysing CRAs’ official communications about specific rating decisions, this article taps into a so far unexploited resource for understanding the way CRAs pass judgement on countries’ creditworthiness. While it has long been clear that macro-economic and fiscal developments are important for credit ratings (Cantor and Packer 1996; Afonso 2003; Afonso et al. 2007), scholars have repeatedly claimed that CRAs hold their cards too close to their chest to allow us to know what role politics and policy play in influencing their decisions (Bruner and Abdelal 2005; Archer et al. 2007; Paudyn 2013), although some analyses of rating scores suggested that government partisanship and welfare policies affect sovereign ratings (Vaaler et al. 2006; Barta and Johnston 2018 and 2019). In this article, we call attention to and harness the wealth of information about CRAs’ rationale provided by rating reports.

Studying the communications of sovereign CRAs with a focus on the commentary on politics and policy contributes to a long-standing debate in international political economy about the constraints financial markets place on democratic politics. Scholars have long assumed that globalisation and the internationalisation of the markets for government debt place countries in a “golden straightjacket” with investors penalising certain political developments and policy choices and rewarding others (Garrett and Lange 1991; Strange 1996; Rodrik 2000 and 2011; Streeck 2014). Mosley nuanced this expectation by contending that investors only paid attention to policy and politics in emerging countries. She claimed that investors allowed developed nations “room to move” in their politics and supply-side policies (as long as their fiscal and macroeconomic performance stayed within reasonable limits) in an effort to reduce information costs in cases where default was unlikely (Mosley 2000 and 2005). Following the same logic, other scholars pointed out that investors use a whole range of simplifying categorisations as information shortcuts to replace in-depth analysis of a country’s credit risk: not just levels of development, but also geographic location or membership in “clubs” of countries (like trade alliances or other forms of inter- and supranational integration) associated with given

<sup>1</sup>The IMF (2010) study referenced here confirms the presence of a significant effect of ratings on sovereign credit-default swap spreads via event studies as well as causality models. Afonso et al. (2015) and Barta and Johnston (2018 and 2019) explore the effect of sovereign ratings on government bond yield spreads via regression analyses. While Afonso et al. only control for fiscal and economic indicators that have been shown to influence bond spreads, Barta and Johnston also control for the effect of partisanship and several supply-side policy indicators that might affect bond spreads. All four studies confirm the independent effect of ratings on sovereign bond spreads.

levels of risk (Gray 2009 and 2013; Brooks, Cunha and Mosley 2015). Yet others implicitly question Mosley's and others' information-economising hypothesis. Breen and McMenamin (2013) and Sattler (2013) argued that investors' seeming indifference towards conspicuous political developments (like shifts in the partisan composition of government) in certain countries is better explained by the presence of checks and balances in those countries. This implies that investors are well informed about the political contexts of countries where they invest.

Our analysis of CRAs' justification for their rating decisions casts doubt on the notion that country categories and "clubs" associated with lower uncertainty have more "room to move" in their *policy* choices. In our sample of European countries, we find that CRAs consistently scrutinise policy choices when making rating decisions in all countries across all time periods. It is only when it comes to monitoring *politics* that country categories and "clubs" mattered, and they only made a difference prior to the global financial and economic crisis. Up to the crisis, the weight of political commentary in rating reports was significantly heavier in emerging countries than in developed ones, except in those emerging countries that were transitioning towards European Union (EU) membership. The latter received similar treatment as developed countries, suggesting the presence of a significant "club"-effect. This leeway from political scrutiny for developed and transitioning countries disappeared as the global financial and economic crisis dispelled the illusion of lower uncertainty in these two privileged categories.

Our empirical strategy relies on the quantitative text analysis of 635 rating reports issued by Standard and Poor's (S&P) between 1999 and 2012 for 40 European countries. We focus on the relative frequency of terms related to politics, policy issues, macroeconomic conditions and fiscal indicators within the texts. We investigate the variation in the patterns of analysis across different country groups and time, contrasting developed with developing countries and exploring the impact of EU membership within the developing category. We also compare patterns before, during and after the onset of the global financial and economic crisis. Of the "big three" rating agencies (Fitch, Moody's and S&P), we focus on S&P, because it is the most visible and influential agency of the three.<sup>2</sup> Our geographic sample is chosen to enable our investigation to be sensitive to potential differences in S&P's attitudes towards different categories and "clubs" of countries, as Europe has an even mix of developed and emerging countries, and has important "clubs," like the EU. The years between 1999 and 2012 allow us to capture the impact of the global financial and economic crisis on the degree to which S&P incorporates political and policy factors into its ratings.

The article proceeds in five steps. The next section explores CRAs' motivations to engage in or avoid scrutiny of domestic developments and explains why credit rating reports provide the best empirical resource to use when trying to understand CRAs' attitudes towards politics and policy. The third section explains the design of our content analysis. The fourth section discusses the results. The fifth section concludes by underlining that the "golden straightjacket" is real, because some of the most central and influential actors in financial markets do not constrain

<sup>2</sup>For details on S&P's leading role in the sovereign credit rating market, see p. 13.

themselves to purely technical analysis of macroeconomic and fiscal indicators, but systematically comment (and act) on politics and policy choices of democratically elected governments. It also points out that the global financial crisis eliminated the comparative immunity from political interference that developed countries previously enjoyed relative to their less developed counterparts.

### **Terra incognita: the role of politics and politically sensitive policy choices in sovereign rating decisions**

Despite two decades of ever-intensifying scholarly interest in sovereign credit ratings, it remains incompletely understood to what extent CRAs' assessment of the creditworthiness of countries goes beyond the analysis of macroeconomic and fiscal indicators to scrutinise politics and politically sensitive policy choices. Monitoring headline macroeconomic and fiscal indicators (like Gross Domestic Product (GDP), GDP/capita, growth, unemployment, inflation, trade balance, public debt or deficit) has direct and unambiguous relevance for assessing a government's ability and willingness to pay because it allows for gauging the trajectory and current affordability of debt and the inclination of a government to prioritise fiscal balance over other policy goals. In contrast, the scrutiny of politically charged policy choices (like social policy, taxation, regulation, industrial policy, public investment, etc.) and political developments (such as electoral shifts, interest-group activity, or intragovernmental tensions) is a much less self-evident dimension of judging creditworthiness because the effect of such factors on a country's ability and willingness to pay is indirect, ambiguous and often only arises in the distant and uncertain future.

Politically sensitive policy areas – like the size and focus of social spending, the types and size of taxation, the regulation of product, labour and financial markets, subsidies, or public investment projects – are often referred to as supply-side policy because they significantly influence human- and physical capital accumulation, the availability and quality of the labour force, innovation and personal incentives (Feldstein 1986; Mosley 2000). Since these policies affect the economy's long-term productive potential as well as the government's future spending and revenue streams, they arguably affect long-term credit quality. Following a similar logic, political developments can be linked to future creditworthiness if one expects certain political developments to be reliably associated with certain fiscal, economic and supply-side policy choices. For example, the ideological composition of a new government is sometimes taken to be an indicator of what policy agenda will be pursued, while the strength of unions or lobbies is often seen to limit policy choice. Yet, drawing firm conclusions about future creditworthiness from current policy choices and political developments is a controversial practice. In light of the intense and unabating scholarly and ideological controversy about what concrete supply-side policy choices yield better economic and fiscal outcomes (e.g. Feldstein 1986; Lucas 1990; Easterly and Rebelo 1993; Engen and Skinner 1996; Carlin and Soskice 2009) and the uncertain correlation between political factors and certain policy choices (Cusack 1999; Ross 2000; Potrafke 2009), judging policy choices and political developments (even from the perspective of their effects on future credit quality) is an inherently value-laden, ideologically loaded and

contentious exercise, which implies taking sides on politically sensitive issues at the very heart of democratic contestation.

CRA's face conflicting incentives when it comes to scrutinising policies and political developments. The potential advantages of such analysis in terms of additional information to predict the future might be more than counterbalanced by risks to CRA's authority from taking sides in contentious and ideologically loaded political and policy debates. Ratings need to provide *forward-looking* indicators of credit quality (Carruthers 2013). Since current fiscal and macroeconomic performance only conveys information about current governments' current willingness and ability to pay, rating agencies might use political and policy analysis to bolster their ability to credibly predict future fiscal and macroeconomic performance and, thus, future credit quality. However, judging politics and policy might be risky for CRA's *epistemic authority*, which the rating business is founded on. Epistemic authority guarantees the trustworthiness of the risk assessment ratings represent (Sinclair 2005). Since incorporating the scrutiny of politics and policy into rating decisions involves a number of contentious, normative and ideologically loaded assumptions about how politics and policies affect the economy and fiscal performance, entering this terrain could be risky for the epistemic authority needed to undergird the credibility of ratings. This leads some scholars to expect CRA's to steer clear of such issues and focus strictly on relatively uncontroversial quantitative analysis of economic and fiscal indicators, in order to shroud their inherently subjective judgments in the "objectifying cloak of economic and financial analysis" (Sinclair 2005, p. 35; Paudyn 2013).

A possible way to reconcile the contradicting requirements of needing to avoid controversy that could undermine the epistemic authority of ratings and the need to integrate additional information on which forward-looking projections for the future could be based is to conform closely to market conventions and widely accepted mental models in assessing the tenuous links of politics and supply-side policies to creditworthiness. Indeed, scholars have argued that CRA's retain their authority by converging closely on market sentiment, going as far as claiming that ratings "codify what the market already knows" (Abdelal and Blyth 2015, p. 40, see also Sinclair 2005).

However, what this implies concretely for CRA's approach towards political and policy analysis is somewhat obscured by the ambiguity about the question to what extent markets themselves monitor politics and policy. This issue is at the heart of the "golden straitjacket" debate. Some scholars assume that markets pay close attention to – and reward and penalise – politics and policy (Garrett and Lange 1991; Strange 1996; Rodrik 2000 and 2011; Streeck 2014). Others argue that investors use shortcuts and heuristics to economise on information costs, and therefore, they employ different levels of political and policy scrutiny to different country categories and country "clubs" (like trade alliances or other forms of inter- and supranational integration) associated with different levels of perceived political, economic and fiscal uncertainty. In developed countries and members of "clubs" associated with low levels of uncertainty, like the EU, politics and policy receive little attention, whereas developing countries and countries belonging to less prestigious "clubs" are subject to extensive political and policy scrutiny (Mosley 2000 and 2005; Gray 2009 and 2013; Brooks, Cunha and Mosley 2015).

This view of information-economising investor behaviour stands in marked contrast with conceptions of more sophisticated decision-making by investors. For example, Breen and McMenamin (2013) and Sattler (2013) contend that market reactions to political developments (like the outcomes of elections) are modulated by the political institutions (specifically the number of veto points) in a country. Such nuanced assessment of the consequences of conspicuous political events presupposes intimate knowledge of domestic political relationships by investors. Given this theoretical ambiguity, it is difficult to determine how CRAs should be expected to approach political and policy analysis, if their goal is to replicate the mental models of markets. Whether they only monitor politics and policy closely in categories and “clubs” associated with greater uncertainty, or they engage in nuanced analysis of these issues across the board, is an issue to be empirically decided.

Empirically, the existing evidence is too scarce to draw definitive conclusions. Quantitative studies have mostly focused on confirming the importance of fiscal and macroeconomic indicators for credit rating scores (Cantor and Packer 1996; Afonso 2003; Afonso et al. 2007), and insofar as they incorporated political factors, they attempted to capture the impact of the broadest political institutional frameworks such as regime type or of major political upheavals like coups and revolutions (Haque et al. 1998; Archer et al. 2007; Beaulieu et al. 2012; Biglaiser and Staats 2012). The effects of day-to-day politics and policy have been less investigated, although two studies provide evidence that elections and the partisan colour of governments affect sovereign ratings (Vaaler et al. 2006; Barta and Johnston 2018), while a third shows that ratings react negatively to generous entitlement systems but are neutral towards other types of social policies (Barta and Johnston 2019).

Qualitatively oriented studies have emphasised how difficult it is to glean unambiguous information from CRAs themselves about the role of politics and policy in their decisions about ratings. Archer et al. (2007, p. 357) report major obstacles in obtaining definite answers about the role of politics in rating decisions from interviews with sovereign rating analysts. In a similar vein, Bruner and Abdelal (2005, pp. 199–200) point out that even though CRAs publish detailed methodologies to explain the thought processes that underlie rating decisions, these documents discuss policy at a level of such generality that it is impossible to dispel the ambiguities that surround what is essentially a highly subjective assessment process and to fully understand to what extent and how rating decisions incorporate policy developments.

In the decade since the publication of these studies, CRAs have only become more taciturn about the standards along which they assess the impact of politics and policy choices on creditworthiness. Although sovereign rating methodologies have been significantly expanded and revised in the wake of the global financial crisis and the European sovereign debt crises (as CRAs sought to restore their credibility by increasing the transparency of their practices), recent editions reveal even less about CRAs’ attitude towards day-to-day politics and policy than before. Whether out of a desire to maintain epistemic authority by avoiding mention of contentious assumptions about the long-term impact of political factors and policy on the economy and the budget, or because CRAs in fact have little interest in day-to-day politics and in policy, the newest editions of the sovereign methodologies eschew these issues altogether. This stands in marked contrast with the extensive



discussion on the importance of lasting economic and monetary institutions and current macroeconomic and fiscal outcomes. Potentially politically sensitive policy areas (like welfare arrangements, the size of the public sector, the characteristics of tax systems, regulatory regimes, etc.) are not even mentioned.<sup>3</sup> Nor do the methodologies comment on the effect of elections or the partisan colour or structure of the government in power on the assessment of sovereign risk.

In contrast to their reticence about their general principles of assessing politics and policy, however, CRAs are remarkably forthcoming about the factors that influence their *specific* rating decisions. Each time a CRA officially announces its decision to raise, lower or affirm a country's rating and to issue negative, positive or stable outlooks about future rating changes, a report is published in which the CRA spells out the reasons that motivated its decision.<sup>4</sup> These reports usually range between 300 and 600 words, and their frequency depends on how often a CRA deems it necessary to revisit a country's rating. They are usually published in their entirety by outlets like Reuters or Bloomberg, whereas other financial news sources quote lengthy excerpts in their reporting of new rating decisions. These documents provide a rich source of information about how CRAs arrive at, justify and communicate their decisions, but no one has systematically analysed them to draw conclusions about the role of politics and policy choices in influencing ratings.

The stark contrast between the secrecy surrounding general principles and the openness about specific cases in rating reports is consistent with the tension, explored above, between CRAs' incentives to avoid taking an explicit stance in theoretical political or policy controversies and their need to bolster the credibility of their forward-looking assessment of credit quality. Rating reports make it possible for CRAs to inspire trust in their decisions by elaborating on all the factors that influenced the final verdict. At the same time, CRAs can eschew taking sides in abstract theoretical debates through high-profile blanket pronouncements about what policies and politics generate optimal results for creditworthiness in general, which could damage their epistemic authority. Arguably, comments on a concrete country's concrete policy choice or political developments are less likely to stir a highly publicised controversy than making generalised statements in defining documents like rating methodologies.

Anecdotal evidence gleaned from reading a random sample of reports suggests that CRAs complement the assessment of economic and fiscal trends with scrutiny of domestic politics and policy. The quotations below provide an assortment of illustrative examples of political and policy analysis from reports from each of the "big three" agencies, about a broad range of countries, justifying negative, neutral and positive rating decisions:

<sup>3</sup>Even generic references to policy ideals such as liberalised trade, labour and finance, efficient tax systems and sustainable welfare regimes that were present in previous editions of methodologies (S&P 2006) are absent from the latest editions (S&P 2011, 2013 and 2016, Moody's 2016, Fitch Ratings 2016).

<sup>4</sup>To our best knowledge, these reports do not have a uniform designation. S&P has often, but not always, called them "Research Updates," whereas Moody's reports mostly appear under the heading "Rating Action," but the standard element in naming these reports across agencies and time is that the title of the document always states the rating action being taken (e.g. "Fitch Affirms Germany at 'AAA'; Outlook Stable").

*S&P, Venezuela, downgrade, outlook stable:* The downgrade reflects the setback to structural reform and fiscal discipline stemming from Venezuela's new constitution, approved Dec. 15, 1999, and from other measures taken during the autumn by the Chavez administration. The constitution may well reverse last year's severance payment reform and scuttle prospects for private sector participation in management of pension funds. It may also further compromise the independence of the central bank, raise protective barriers for agriculture, and abrogate the right for international arbitration of contracts in the "public interest." (21 December 1999, p. 1)

*Fitch, Germany, rating affirmed, outlook stable:* [L]abor market institutions are largely responsible for the persistence of high unemployment, especially amongst the less-skilled. Though wage-bargaining is gradually becoming less centralized, tax and welfare arrangements continue to mitigate against rapid job creation and flexibility [...] The party financing scandal that has embroiled the CDU and former Chancellor, Helmut Kohl [...] has strengthened the position of Chancellor Schroder and the ruling Sozialdemokratische Partei Deutschlands (SPD)/Green coalition that is likely to remain in office until general elections in 2002. It may also provide the Schroder government with an opportunity to pursue a bolder economic reform agenda. (28 January, 2000, p. 1)

*Moody's, Ukraine, rating affirmed, positive outlook:* Ukraine's new coalition government provides the country with more political stability than did the previous governing coalition and the controversial 'Orange Revolution' campaign that preceded it. [...] Authorities will take steps necessary to reduce the general tax burden, introduce important pension and health care reform, and follow policies designed to stimulate the capital investment necessary to modernize steel and other important sectors of the economy. (10 November 2006, p. 1)

Evincing the desire to provide forward-looking assessments of risk, the reports also spell out possible future political and policy scenarios to alert investors to possible rating changes (and, in the process, also signalling to governments the penalties and rewards associated with different political and policy choices):

*S&P, Germany, rating affirmed, outlook stable:* The stable outlook reflects S&P's expectation that the reform process will continue, albeit subject to likely delays prior to the general elections in 2006. One area of reform where progress would be attainable prior to the elections is on the federal system [...]. Conversely, failure to push the reform process further, or a more expansive fiscal stance, would put increasing downward pressure on the ratings. (1 March, 2005)

These excerpts indicate that CRAs analyse politics and supply-side policy in AAA-rated Germany just as much in-depth as in (at the time) B-rated Venezuela and (at the time) CCC-rated Ukraine. They also suggest that commentary on policy is, at least occasionally, informed by the neoliberal paradigm, demanding, for example, the privatisation of pension funds, independent central



banks, trade liberalisation (in Venezuela), labour market liberalisation and the reform of tax and welfare arrangements (in Germany), the reduction of the tax burden and health care and pension reform (in the Ukraine). Casual further reading of more than 100 reports reinforces this impression, with reports regularly talking about the need for structural reforms in labour markets, welfare and health in ways that are consistent with a neoliberal view of policy. Interestingly, reform also features repeatedly in discussions of politics (e.g. the need to reform the federal system in Germany in reports from the early 2000s, or the need for constitutional reform in Italy in the first half of the 2010s), besides discussions of day-to-day politics.

The excerpts also document complementarity between political and policy analysis as CRAs seek to provide projections about future economic and fiscal developments. Policy is scrutinised in the context of assessing the longer term economic and fiscal health of a country, whereas political analysis is employed to predict potential changes in policy. Thus, these excerpts suggest that policy analysis and political scrutiny represent complementary components of CRAs' attempts to reduce uncertainty about future macroeconomic and fiscal performance.

The rest of this article systematically explores how the assessment of politics and policy features in CRAs' rating reports. We entertain the possibility that categories and "clubs" of countries associated with different levels of uncertainty are treated differently by CRAs, mirroring the behaviour of investors documented by Mosley and others. Earlier versions of CRAs' methodologies provided indication that CRAs associate certain categories and "clubs" of countries with lower levels of (political and policy) uncertainty and, therefore, treat them differently. Fitch stated that developed sovereigns constitute a privileged category (Fitch Ratings 2009, p. 5), whereas Moody's and S&P confirmed that prestigious "clubs" like the EU are seen more favourably when creditworthiness is evaluated (Moody's 2008, S&P 2006, p. 7). Such statements have been removed from more recent editions of the methodologies issued after the global financial and economic crisis (S&P 2011, 2013 and 2016, Moody's 2016, Fitch Ratings 2016), suggesting that the crisis might have changed CRAs' perception of lower uncertainty in developed sovereigns and members of certain "clubs."

Therefore, we test five related hypotheses about the variation in the weight of political and policy scrutiny in CRAs' assessment of creditworthiness across countries and time. We form separate hypotheses about political and policy scrutiny, in case they might be applied differently across country groups associated with different levels of uncertainty. We also form separate hypotheses about country categories (e.g. developed/emerging) and membership in "clubs" (like the EU) because they represent slightly different theoretical propositions. Finally, we hypothesise that the global financial and economic crisis represents a trend break in CRAs' approach to uncertainty. The crisis represented a major shock to CRAs' authority, which led them to re-evaluate some of their theoretical propositions (which is reflected in major revisions in all of the three agencies' methodologies), while the series of bank failures and sovereign debt crises in advanced economies also called attention to the mistakeness of markets' earlier complacency about the creditworthiness of developed sovereigns. Therefore, we expect the weight of political and policy scrutiny across countries to change after the crisis. On the one hand, the perception of lower uncertainty in developed countries might disappear. On the other hand, the relative

weight of political and policy scrutiny might change if a generalised sense of elevated uncertainty causes CRAs to move towards political analysis in an effort to better predict changes in policy and in economic and fiscal performance.

- *Hypothesis 1a (politics and categories)*: The relative weight of political commentary is significantly heavier in emerging economies than in developed countries.
- *Hypothesis 1b (politics and “clubs”)*: Within the emerging category, the relative weight of political commentary is significantly lighter in countries that belong to organisations (“clubs”) associated with positive connotations, like the EU.
- *Hypothesis 2a (policy and categories)*: The relative weight of policy commentary is significantly heavier in emerging economies than in developed countries.
- *Hypothesis 2b (policy and “clubs”)*: Within the emerging category, the relative weight of policy commentary is significantly lighter in countries that belong to organisations (“clubs”) associated with positive connotations, like the EU.
- *Hypothesis 3 (trend break)*: The financial crisis of 2008 causes a trend break in the pattern of political and policy commentary, affecting the relative emphasis on politics and policy, or the treatment of different country groups, or both.

### **The empirical strategy: interrogating country rating reports about the relative importance of politics and policy in rating decisions**

Our empirical investigation of the above hypotheses is based on a quantitative text analysis of 635 credit rating reports issued by Standard & Poor’s Financial Services LLC between 1 January 1999 and 30 June 2012 for European countries, broadly understood. We gain a better understanding of the weight that politics and policy receive relative to the economy and fiscal performance in rating reports by measuring the relative frequency of words associated with each topic. We investigate the patterns of variation across country groupings and time periods by using linear mixed-effects models fitted with residual maximum likelihood.

#### ***The sample***

Of the “big three” rating agencies, we choose to focus on S&P because S&P is arguably the most visible and materially consequential CRA on the market. It has been shown that the “big three” are generally in approximate agreement about the creditworthiness of countries, assigning credit scores within one notch of one another’s ratings (Afonso 2003). At the same time, it has also been documented that S&P is frequently the first-mover in initiating rating adjustments, especially negative ones.<sup>5</sup> This makes S&P particularly influential in setting the common trends in sovereign rating, in shaping the discourse about the implications of a country’s policies and politics on creditworthiness and, ultimately, in determining the conditions under which a country can borrow.

<sup>5</sup>S&P led the other two CRAs 74% of the time in downgrading developed economies between 2005 and 2010 (IMF, 2010, p. 115).

In sampling countries, we focus our attention on Europe, because Europe provides a fairly varied and balanced sample across the country categories relevant for our hypotheses (developed and emerging). Europe also has well-known exclusive “clubs,” like the EU, whose membership might have an impact on the intensity of political and policy scrutiny. To test our hypotheses, we divide European countries in three groups. Old member states of the EU, as well as Iceland, Norway and Switzerland constitute our “developed” category.<sup>6</sup> Countries that transitioned to EU membership in 2004 and 2007 constitute our “transitioning” category.<sup>7</sup> Countries that were not part of the fifth and sixth waves of EU enlargement make up our “emerging” category.<sup>8</sup> We create a separate bracket for developing countries that were scheduled to become members of the EU during the period under consideration to capture any potential “club-effect” of impending accession on levels of political and policy commentary *within the developing category*. Based on their level of economic development, “transitioning” countries were closer to “emerging” than to “developed” countries throughout the 1990s and 2000s, and the gap relative to developed countries persisted well after they joined the EU (Makszin et al. 2020). Yet, given definite plans for accession to the EU from the late 1990s and actual membership in the second half of the 2000s, these countries were differentiated from the rest of the developing countries within Europe by being associated with the prestigious “club” that the EU was considered to be at the time. Should we detect differential treatment between our “transitioning” and “emerging” categories, this would point to the presence of a “club” effect.

From the vantage point of the late 2010s, these categories and “club” memberships might seem less meaningful than the labels (like core versus periphery or PIIGS<sup>9</sup>) that have emerged in the financial market discourse – and which have arguably better characterised the economic and financial fortunes of European countries – during and since the global financial and economic crisis and the wave of sovereign debt crises that shook Europe. We choose to forego the benefit of hindsight and focus primarily on the developed versus emerging distinction and the EU as a club<sup>10</sup> because these designations allow us to better explore the theoretical proposition that CRAs seek to converge on the mental models of the markets and, therefore, apply the same heuristics, when deciding whether and how thoroughly to scrutinise politics and

<sup>6</sup>This category includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Switzerland, Sweden and United Kingdom.

<sup>7</sup>Transitioning countries include those that joined the EU between 2004 and 2007, specifically: Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia and Slovenia.

<sup>8</sup>This category includes Albania, Belarus, Bosnia, Croatia, Macedonia, Montenegro, Russia, Serbia, Turkey and Ukraine.

<sup>9</sup>PIIGS refers to some of the countries worst hit by the crisis, specifically Portugal, Italy, Ireland, Greece and Spain.

<sup>10</sup>It is beyond the scope of this paper to explore what mental models of creditworthiness explain the widespread use by investors in the 1990s and 2000s of the developed/developing distinction and “club” membership as shortcuts. Whether such labels are to proxy for the level and age of democracy, or the stability of the political system (whatever the regime), or they are to serve as a heuristic for the strength and stability of the economy is a subject for future research. What matters for the purposes of this paper is whether CRAs observe this distinction, whose use by investors in the 1990s and 2000s is so well documented (Mosley 2000; Brooks et al. 2015).

policy. These categories and “clubs” were the most dominant mental shortcuts before the crisis, that is, for *most of the duration of the time period that our sample covers* (as documented not only by Mosley’s work, but also by CRAs’ methodologies, as discussed above).

That said, it is important to identify the potential effect of the crisis on the mental shortcuts in operation. This is why we hypothesise a trend break in the usage of these categories and “clubs” around 2008 (see hypothesis 3 above). Furthermore, we explore – in an auxiliary regression analysis, whose results are presented in Table A6 in the Appendix and briefly analysed in the next section – whether the new categories that emerged with the crisis explain changes or persistence in the patterns of political and policy scrutiny among our original categories after the crisis.

In terms of temporal sampling, we chose to start our analysis in 1999 because many countries in the “transitioning” and “emerging” categories were not rated before the end of the 1990s, as S&P (just like other agencies) gradually expanded its portfolio of rated countries from developed to developing countries in the 1980s and 1990s. The 1999–2012 timeframe conveniently allows us to compare the periods before and since the onset of the global financial crisis and the European debt crises. Ideally, we would have liked to extend our analysis to the years beyond 2012 in order to explore any potential changes in the patterns of emphasis on politics, policy, fiscal and economic factors after the turbulence of the crisis had fully subsided, but we do not have access to reports beyond 2012.

### ***The quantitative text analysis***

To explore the relative emphasis S&P places on politics, policy, the economy and fiscal conditions in its assessments of creditworthiness and its official discourse about its decisions, we analyse the relative frequency of terms associated with these four issue areas in the rating reports. We derived the word frequencies in three steps. First, we generated a list of words – excluding stop words – from all of the rating reports using the software “tm: Text Mining Package in R” (Feinerer and Hornik 2017). We did not use word stemming to ensure that we could interpret the terms correctly for the qualitative coding and we kept the top 97% of words according to frequency. Dropping the least frequent 3% of words from our database is appropriate in quantitative text analysis to exclude outliers (Benoit and Herzog 2017, p. 146). At the same time, it leads to the exclusion of country-specific words like proper nouns such as names of political personalities, policy makers and political or policy-making institutions, which (as the quotations in the previous section show) repeatedly feature in reports, but (due to their specificity) do not exceed the 3% threshold in the joint sample of 40 countries over 14 years. Given that the use of such proper nouns likely suggests political or policy commentary, this omission should bias our results about the weight of political and policy-related issues in rating reports slightly downwards, that is, we are likely to slightly underreport the weight of politics and policy in the rating discourse.

As a second step, we conducted two independent qualitative codings of the resulting list of words to determine whether they relate to politics, policy, the

economy or fiscal developments, or they are neutral. Both coders read an extensive sample (approximately 100 random reports each) prior to coding the word list, in order to gain familiarity with the terminology applied in the reports. We reconciled the two codings where any discrepancies existed.<sup>11</sup>

We labelled words “Economic” if they can be directly linked to long- and short-term economic issues on which we know (from S&P’s methodologies as well as from sampling reports) S&P places significant emphasis. These include long-term overarching characteristics of the economy (e.g. diversification, export-orientation, GDP per capita, etc.) and indicators of short-term performance (e.g. headline macroeconomic indicators like growth, inflation, employment, trade balance or the exchange rate, but also short-term sectoral issues like banking or construction performance, household consumption, etc.). We also used our knowledge of the terminology of reports when coding terms like “resilient,” “inefficient” or “recovery” under this label. Under “Fiscal,” we coded words referring to the usual headline fiscal indicators (e.g. debt to GDP, deficits, revenues, spending, etc.) or to the financing and management of the outstanding debt (e.g. rollover, bonds, maturity, etc.). All of these aspects of fiscal policy feature prominently in S&P’s methodology. We also coded words under this label that we recognised as recurring phrases in the terminology of reports (like “sustainable trajectory,” “tightening” or “prudence”).

As previously explained, CRAs’ methodologies provide little help in pinning down the political and policy-related aspects of the rating assessment. In the absence of clear cues, we included under “Politics” words that unambiguously relate to governing (like coalition, government, minister, parliament, etc.), to interest representing organisations (e.g. agenda, party, union), electoral issues (e.g. election, referendum), and to conflict and its resolution (agreement, consensus, summit, tensions, etc.). This also reduced the concern about exclusion of proper nouns, as they would usually be coupled with terms that were counted (such as “Chavez administration” and “SPD/Green coalition”). Under “Policy,” we included terms that relate directly to a specific policy area (e.g. health, pensions), or policy actions (e.g. austerity, programme, package, reform), but our codings also relied on our knowledge of the terminology of reports surrounding politically sensitive policy issues (e.g. when classifying “challenge,” “commitment,” “credible,” “resolve” under this label). Words that did not unambiguously relate to one of the four issue areas, we coded as “Neutral.” Table A7 in the Appendix displays the list of words in each in our four coding labels (plus the residual “Neutral” label).

Our final step was counting the frequency of the coded words and recording the total word count in each report (using the same text-mining package). We work with the relative frequency of terms as a share of the total non-neutral word count for the report, rather than the absolute numbers, because reports vary greatly in length (expanding pronouncedly in the years since the global financial and economic crisis, see Figure 1). The average number of words per report is 296.6 and the standard deviation is 105.5. Using relative frequencies allows us to eliminate potential confounding effects such as different writing styles of different analysts

<sup>11</sup>The intercoder reliability score (Cohen’s kappa) was 0.8494 for economic terms, 0.8031 for fiscal terms, 0.8797 for political terms and 0.7395 for policy terms.

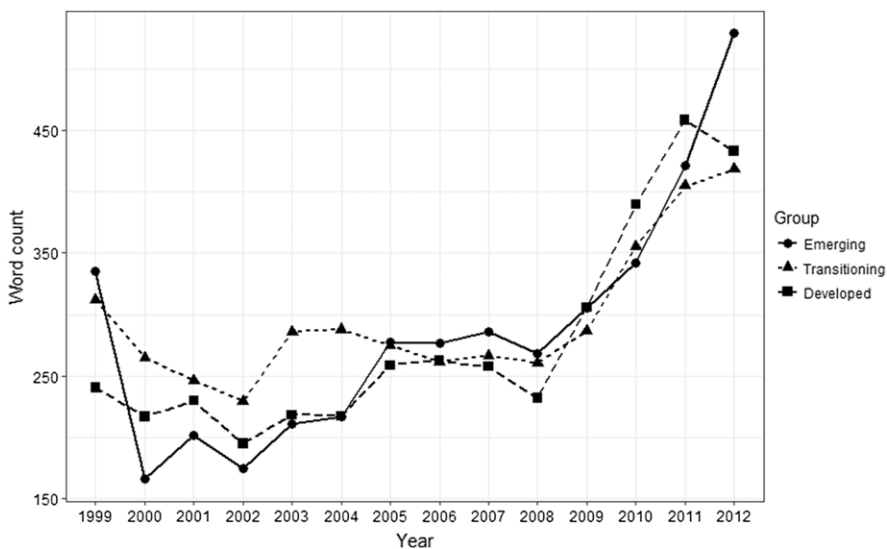


Figure 1. Word count per report over time by country group.

assigned to specific countries (which could lead to shorter/longer reports, and therefore fewer/more words, in the case of specific countries) or the general lengthening of reports in the wake of the crisis. As a robustness check, we also ran our models using absolute word count of politics- and policy-related terms, including a control variable for report length. The results using absolute word count (presented in Table A3 of the Appendix) are consistent with the findings using relative word frequencies (presented in Table 2).

Since using relative frequencies makes it possible that our two variables of interest (the share of political commentary and the share of comments related to supply-side policies) covary in ways that variation in one variable confounds variation in the other, we check the correlation of our four categories. We find that politics- and policy-related terms have negligible correlation with each other ( $r = 0.13$ ). The strongest correlations are between the relative weight of “economic” terms and the relative weight of each of the three other codes ( $-0.46$ ,  $-0.56$  and  $-0.63$  with “fiscal,” “politics” and “policy,” respectively), suggesting that the weight of economic analysis expands and contracts at the expense of the analysis of the other three areas, while the other three areas are relatively unrelated to each other.

Having measured the relative frequency of words in all four issue areas in all of the 635 credit rating reports, we analyse the variation in the relative frequency of terms related to politics and policy across country groupings (“developed,” “transitioning” and “emerging”) and time relative to crisis (precrisis from 1999 to 2007; during the onset of crisis in 2008 and 2009, and after the onset of the crisis from 2010 to 2012), in accordance with our hypotheses specified above. To do so, we record the country and the year associated with each rating report. Individual reports serve as our unit of analysis. Our dependent variables are the relative frequency of terms associated with politics and policy in rating reports. Our independent variables are country



group, time period and the interaction of the two. Given the varying length of reports, we include the word count of the report as a control variable.

The aim of our empirical analysis is to test whether there are significant differences in the extent to which politics and policy are discussed across country groups and over time. We are also interested in the interaction of country groups and time periods. We apply linear mixed-effects models (fitted with residual maximum likelihood and random intercepts) instead of analysis of variance (ANOVA) because our sample is unbalanced, with a varying number of reports among countries and among years, which makes both time-series cross-sectional and factorial ANOVA analysis unreliable in capturing the effect of country groupings and period on our dependent variables. As Barr et al. (2013) explain, the strength of this method (in contrast with ANOVA) is in the flexibility of the model design with regard to specifying fixed and random effects, and the fact that it can accommodate an unbalanced sample. Benefits of linear mixed-effects models over factorial ANOVA also include the possibility to establish the direction and magnitude of the differences across categories.

The aim of the analysis is to measure variation in the dependent variables across categories rather than to establish complete causal explanations for the dependent variable. This is why our models do not control for the objective economic factors but focus on variation across country groups, time period and the interaction of the two. The decision to focus on comparison across these categories is also based on the limited number of country level units (40) and high covariance between objective economic factors, such as GDP-levels and country groups. Our models estimate the relative frequency of terms associated with politics and policy, treating country grouping, time period relative to the crisis and the interaction between the two as fixed effects, meaning that we fit a model with a single coefficient for these variables as we do not expect the slopes to vary across countries. The models include a random intercept for countries to account for the lack of independence between reports for the same country. We also control for the length of the report as a fixed effect in each model given that this varies, as shown in Figure 1. The regression equation used in our analysis is

$$y_{ij} = b_0 + b_1 * (\text{country group})_{ij} + b_2 * (\text{time period})_{ij} + b_3 * (\text{country group}) * (\text{time period})_{ij} + b_4 * (\text{world count})_{ij} + v_{i0} + e_{ij}$$

for  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, m\}$ , where  $i$  represents the country and  $j$  the number of the report per country.  $y$  is the share of non-neutral terms in the report corresponding to the category being analysed.  $b_{1-4}$  represents the coefficients for the fixed effects.  $v_{i0}$  are the random intercepts for countries.

The analysis was done in R using lme4 and lmerTest packages (Bates et al. 2015; Kuznetsova et al. 2017). The models are fitted using restricted maximum likelihood, and p-values are estimated using the Satterthwaite approximation in the lmerTest package, which is shown to produce acceptable type 1 error rates even with limited sample size (Luke 2017). The estimated means for categories and plots were generated using emmeans package (Lenth 2019).

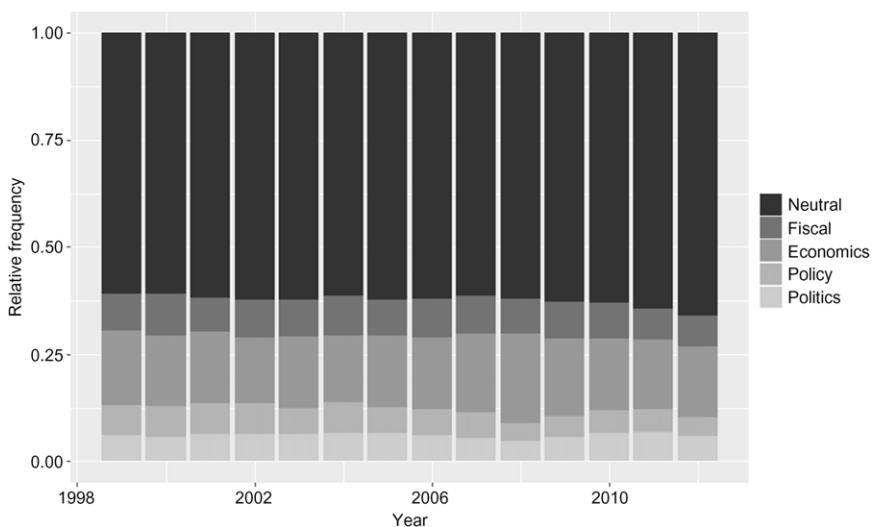
**Table 1.** Summary statistics of relative word frequencies (all countries, 1999–2012)

Issue Area	Mean	Standard Deviation	Minimum	Maximum
Economic	45.7	10.9	16.7	100
Fiscal	22.3	6.7	0	47.7
Politics	16.5	6.6	0	52.6
Policy	15.5	6.5	0	38.5

### The results: the emphasis on politics and policy in rating reports

Table 1 presents the summary statistics of the relative frequency of terms in all four issue areas as a percentage of all words coded under one of our four substantive labels for all 40 countries over the entire period between 1999 and 2012. Figure 2 presents the weight of the different issue areas over time. We excluded the residual neutral words from our analysis of the relative weight of the four issue areas to avoid confounding effects of potential changes in the share of “Neutral” terms. As these words could not be assigned theoretical meaning in the coding process, predicting their share within the reports in a model would not produce theoretically interpretable results. (However, as a robustness check, we replicate our models including “neutral” words in the total word count and the substantive findings are consistent. See Table A5 in the Appendix.)

Table 1 shows that although the analysis of economic and fiscal conditions dominates S&P’s discourse on countries’ creditworthiness, terms referring to politics and policy also have notable weight. Words clearly relating to economic and fiscal performance account for an average of 45 and 22%, respectively, of the total words coded under one of our substantive labels in all the countries across the years. Terms

**Figure 2.** Relative word frequencies per issue-area over time.

unambiguously referring to politics and policy make up 16 and 15% of the words of the reports, respectively, accounting for a third of the non-neutral words.

Table 2 presents the results of our analysis of the variation of the weight of political and policy scrutiny across country groups (developed/transitioning/emerging) and periods (before/during/after the onset of the crisis), using linear mixed-effects models. (Results for economic and fiscal terms can be found in Table A2 in the Appendix.) It shows that country groups, time period relative to the crisis and their interaction statistically significantly influence the weight of *political commentary* in reports, suggesting that (1) S&P approached politics in different country groups differently; (2) the crisis increased the average weight of political scrutiny for the whole sample; and (3) the crisis diminished the differential treatment of country groups. In the case of *policy commentary*, we find no significant variation in the weight of policy-related words across country groups, only across time periods, with attention to policy declining after the crisis. We elaborate on these results below with the help of graphs that visualise the interaction between country group and time period by displaying the estimated marginal means with confidence intervals for ease of interpretation.

The models for politics demonstrate substantial variation in the share of political terms across country groups and over time. To interpret the main findings, we refer to Model 2, which includes the interaction between country groups and time. Our findings are robust across models that include or exclude the interaction term or use the absolute instead of relative word count (see Table A3 in the Appendix). Overall, we detect significantly more scrutiny of politics in the emerging group before the crisis ( $B = 0.086$ ;  $p < 0.001$ ) compared with the developed group (base category) and no significant difference between developed and transitioning ( $B = 0.012$ ;  $p > 0.05$ ). This suggests that on average, before the crisis, emerging countries have 8.6 percentage points more substantive words related to politics in their reports when compared with developed countries. For the time dimension, we detect an overall heightened scrutiny of politics for developed countries after the crisis ( $B = 0.068$ ;  $p < 0.001$ ) compared to before (base category), suggesting an average increase of 6.8 percentage points in the substantive words related to politics. The interaction terms suggest a convergence between developed and emerging group after the crisis ( $B = -0.113$ ;  $p < 0.001$ ). Transitioning countries do not differ from developed countries before the crisis, but the interaction term suggests that after the crisis, countries in the transitioning category have a slightly more moderate increase in political scrutiny compared to developed countries ( $B = -0.040$ ;  $p < 0.01$ ).

In order to display our results more clearly, we estimate the marginal mean values for each country group and time period (as recommended by Brambor et al. 2006).<sup>12</sup> The interactions are shown in Figure 3, which depicts the estimated mean percentage of *political commentary* across country groups and periods based on Model 2 in Table 2, with 95% confidence intervals. It shows that in the precrisis years, the weight of political commentary differed among developed, transitioning and emerging groups with politics receiving more emphasis in the emerging group

<sup>12</sup>Table A3 in the Appendix displays the estimated means, standard errors and degrees of freedom that are the basis for the quantities of interest displayed in Figures 3 and 4. We used emmeans (Lenth 2019) for these calculations.

**Table 2.** Results of the linear mixed-effects models. The impact of country groups and timing relative to the crisis on the relative frequency of words related to politics and policy in country rating reports

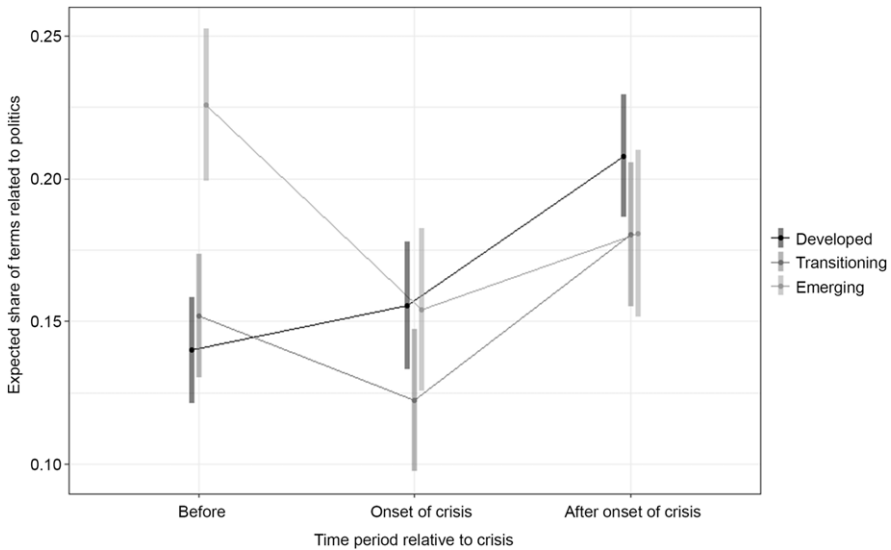
	Dependent Variables			
	Politics		Policy	
	Model 1 Without Interaction	Model 2 With Interaction	Model 3 Without Interaction	Model 4 With Interaction
<b>Fixed effects</b>				
<b>Country groups<sup>†</sup> (base: developed)</b>				
Transitioning	-0.006 (0.012)	0.012 (0.014)	-0.012 (0.128)	-0.018 (0.0135)
Emerging	0.032* (0.014)	0.086*** (0.016)	-0.015 (0.014)	-0.007 (0.015)
<b>Period<sup>‡</sup> (base: before crisis)</b>				
Onset of crisis	-0.022*** (0.006)	0.016 (0.009)	-0.042*** (0.006)	-0.036*** (0.009)
After onset of crisis	0.031*** (0.007)	0.068*** (0.009)	-0.024*** (0.007)	-0.033*** (0.009)
<b>Interaction: group*period (base: developed*before crisis)</b>				
Transition*onset of crisis		-0.045*** (0.013)		-0.001 (0.013)
Emerging*onset of crisis		-0.087*** (0.014)		-0.028 (0.014)
Transition*after crisis		-0.040** (0.012)		0.024* (0.012)
Emerging*after crisis		-0.113*** (0.014)		-0.004 (0.014)
Word count of report		-0.000 (0.000)		-0.000 (0.000)
<b>Random effects</b>				
	Variance		Variance	
Country	0.001	0.001	0.001	0.001
Residuals	0.003	0.003	0.003	0.003
<b>Pseudo R-squared</b>				
Marginal (for fixed effects)	0.104	0.195	0.100	0.109
Conditional (for entire model)	0.304	0.426	0.338	0.349

\*\*\*p < 0.001, \*\*p < 0.01, \*p < 0.05.

<sup>†</sup>Country groups are developed/transitioning/emerging.

<sup>‡</sup>Periods are before/during/after the financial crisis. Coefficients are unstandardised. Standard errors are in parentheses. p-values are estimated using the Satterthwaite approximation. The analysis includes 635 reports across 40 countries. The intraclass correlations for the models are 0.286 and 0.269 for politics and policy, respectively, which represents the amount of variance that is accounted for by between-country differences. Model diagnostics were checked, and residuals were normally distributed and cases of high leverage did not influence the substantive findings.<sup>13</sup> Pseudo R-squared calculated using MuMIN (Barton 2019).

<sup>13</sup>The reports for Albania (n = 2) exhibited high leverage. The analysis was repeated removing them for the sample and the substantive results were the same.

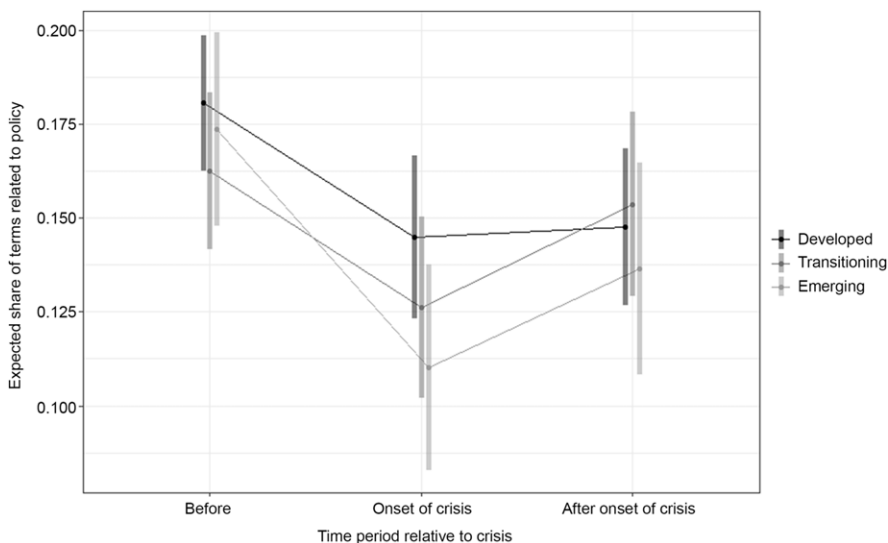


**Figure 3.** Percentage of politics terms by time period and country groups, 1999–2012 (estimated marginal means based on model 2).

than in either the developed or the transitioning one. The distinction between different country groups disappears with the onset of the crisis. At the start of the crisis in 2008, attention to politics plummets in the emerging group, as economics gains supreme importance in the middle of the economic and financial turmoil. After the crisis, attention to politics is heightened in all country groups and the discrepancy between the country groups disappears. The estimated mean share of political terms in the reports for developed countries increased from 14 to over 20.8%.

Cross-country variation in the weight of political scrutiny in the precrisis years confirms our hypotheses 1a and b, in conformity with our expectation that the intensity of scrutiny varies with perceived levels of uncertainty across different categories and “clubs” of countries. Politics was scrutinised significantly more lightly in developed countries than in emerging ones, and new EU-members benefited from association with a prestigious “club,” even though their level of economic development was much closer to emerging than to developed countries in the 1990s and 2000s.

The temporal pattern also lends support to our uncertainty explanation for variation in emphasis on political analysis. Attention to politics significantly intensifies in developed countries and new EU-member states after the crisis dispels the illusion that such countries have more predictable economic and fiscal trajectories than emerging countries (which makes the differentiation between-country groups disappear). We conducted auxiliary analysis of only eurozone countries (presented in Table A6 in the Appendix) to test whether the heightened emphasis on politics in our developed group is due to increased scrutiny of “periphery” countries, referring to the core-periphery distinction that emerged during the eurozone crisis. We detect no difference between core and periphery at the onset of the crisis. In the later years, the estimates increase dramatically for both core and periphery countries, and the



**Figure 4.** Percentage of policy terms by time period and country groups, 1999–2012 (estimated marginal means based on model 4).

estimated share of political terms is actually higher in the core than the periphery by approximately 3 percentage points (see Figure A12 in the Appendix). This finding demonstrates that the escalation in political scrutiny in the developed category is not driven by increased attention to politics in the “periphery” of the eurozone but affects the entire group in an undifferentiated fashion.

Figure 4 visualises our results for average levels of *policy commentary* across country groups and periods. The share of policy-related terms is similar in all country categories in each period (differences are not statistically significant). The weight of policy-related words fell by 3.6 percentage points during the onset of the crisis due to the temporary surge in emphasis on economic factors as a result of the shock ( $B = -0.036$ ,  $p < 0.001$ ) and remained at the lower level after the crisis ( $B = -0.033$ ,  $p < 0.001$ ). Lack of variation in policy scrutiny across country groups disconfirms hypotheses 2a and b: S&P has never allowed countries associated with lower uncertainty “more room to move” in policy and the analysis of such policies is a standard component of the assessment of creditworthiness for any type of country. It appears that even prior to the crisis no country category or “club” was deemed to have predictable enough economic and fiscal trajectories to warrant a more relaxed attitude about policy. At the same time, policy analysis is to some extent “crowded out” by greater emphasis on political scrutiny in the context of increased uncertainty after the crisis.

## Conclusion

Our results show that the analysis of political and policy developments plays an important role in S&P’s rating decisions. In its highly publicised official rating



reports, S&P systematically comments on both politics and policy, thereby linking these issues – which only have indirect and ambiguous connections to governments' ability and willingness to pay – to creditworthiness, ratings and, implicitly, a country's access to credit. While official methodologies deemphasise the role of politics and policy, we detect a meaningful presence of these topics in the justifications of specific credit rating decisions. At the same time, the relative emphasis on these issues varies across countries and time, with political scrutiny more selectively applied than the analysis of policy choices. *Political* commentary was significantly less important in the case of developed countries and new EU-member states before the global financial and economic crisis, although this relative immunity evaporated in the wake of the crisis. Commentary on *policy developments*, on the other hand, affected all country groups equally throughout the entire period.

The lack of alignment between patterns of scrutiny applied to politics and policy issues across countries and time sheds light on the way CRAs use political and policy analysis to get a handle on uncertainty and gauge sovereign credit risk. The contrast between the universality of policy analysis across countries and time and the differentiation in political analysis according to perceived uncertainty (specifically, its increasing relevance with growing uncertainty) is consistent with our theoretical expectation that political and policy analysis represent complementary steps in reducing uncertainty about future macroeconomic and fiscal performance. When uncertainty is perceived to be limited, policy analysis largely suffices to evaluate longer term economic and fiscal scenarios. When policy trajectories are themselves uncertain, political analysis is relied on to a greater extent to supplement policy analysis.

At the same time, this misalignment complicates conclusions about the constraints that credit ratings impose on democratic choice. Evidence of universal emphasis on the scrutiny of policy in S&P's rating decisions supports the classical "golden straightjacket" thesis, which claims that market actors consistently reward and penalise domestic policy choices. At the same time, varying emphasis on political analysis across country groups associated with different levels of uncertainty is consistent with Mosley's and others' findings that the level of development and association with prestigious "clubs" matter for how much "room to move" countries have, although this applied only to their political choices and the leeway that developed countries and members of prestigious "clubs" enjoyed disappeared with the crisis. Now, all countries are similarly closely scrutinised.

In a similar vein, the findings lead to complex conclusions about the degree to which ratings might emulate mental models applied by the market. Our results about differences in emphasis on politics across country groups are consistent with the expectation that CRAs apply the same shortcuts and heuristics as investors – levels of development and membership in "clubs" of countries – when categorising countries according to perceived levels of uncertainty. At the same time, when it comes to policy scrutiny, CRAs do not seem to differentiate according to the conventional categories.

This first foray into analysing S&P's sovereign credit rating reports demonstrates the possibilities offered by the wealth of information contained in rating reports for better understanding the decisionmaking processes of sovereign CRAs. Previously thought to be unfathomable, CRAs' thinking is laid bare in these texts. Obviously,

the bag-of-words approach applied in this article is only a first step in understanding the way CRAs' decisions might constrain democratic choices. It allows for a systematic investigation of consistent patterns of emphasis on politics and politically sensitive policy choices across countries and time, which was thus far not possible based on analysis of methodologies or interviews with analysts. Nevertheless, it provides limited understanding of the substantive content of the commentary that CRAs publish, of the political and policy ideals that they advocate, or of the exact degree to which the incorporation of these issues into rating decisions translates into tangible differences in rating scores. Exploring these questions requires further research, including detailed analysis of the discourse applied in rating reports and quantitative analysis of how the factors identified by content analysis influence rating scores.

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/S0143814X20000033>

**Acknowledgement.** We would like to thank Doro Bohle, Gregory Fuller, Alison Johnston, Waltraud Schelkle, Miklós Sebők and three anonymous reviewers for their invaluable comments on earlier versions of this article.

**Data Availability Statement.** Replication materials are available in the *Journal of Public Policy* Dataverse at <https://doi.org/10.7910/DVN/FYXNIM>.

## References

- Abdelal R and Blyth M** (2015) Just Who Put You in Charge? We Did: Credit Rating Agencies and the Politics of Ratings. In Cooley A and Snyder J (eds.), *Ranking the World: Grading States as a Tool of Global Governance*. Cambridge: Cambridge University Press, 39–59.
- Afonso A** (2003) Understanding the Determinants of Sovereign Debt Ratings. *Journal of Economics and Finance*, 27(1): 56–74.
- Afonso A, Arghyrou MG and Kontonikas A** (2015) The Determinants of Sovereign Bond Yield Spreads in the EMU. ECB Working Paper 1781.
- Afonso A, Gomes P and Rother P** (2007) What 'Hides' Behind Sovereign Debt Ratings? ECB Working Paper Series (711).
- Archer CC, Biglaiser G and DeRouen K** (2007) Sovereign Bonds and the Democratic Advantage: Does Regime Type Affect Credit Rating Agency Ratings in the Developing World?. *International Organization*, 61(2): 341–365.
- Barr DJ, Levy R, Scheepers C and Tily HJ** (2013) Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal. *Journal of Memory and Language*, 68(3): 255–278.
- Barta Z and Johnston A** (2018) Rating Politics? Partisan Discrimination in Credit Ratings in Developed Economies. *Comparative Political Studies*, 15(5): 587–620.
- Barta Z and Johnston A** (2019) Entitlements in the Crosshairs? How Sovereign Credit Ratings Shape Market Reactions to Policy Choice in Advanced Industrial Countries. Paper presented at the 29th Conference of Europeanists in Madrid.
- Barton K** (2019) MuMIn: Multi-Model Inference. R package version 1.43.6. <https://CRAN.R-project.org/package=MuMIn>.
- Bates D, Maechler M, Bolker B, and Walker S** (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1): 1–48. doi:10.18637/jss.v067.i01.
- Beaulieu E, Cox GW and Saiegh S** (2012) Sovereign Debt and Regime Type: Reconsidering the Democratic Advantage. *International Organization*, 66(4): 709–738.
- Benoit K and Herzog A** (2017) Text Analysis: Estimating Policy Preferences from Written and Spoken Words. In Bachner J, Ginsberg B and Wagner Hill K (eds.), *Analytics, Policy and Governance*. New Haven: Yale University Press, 137–159.

- Biglaiser G and Staats JL** (2012) Finding the 'Democratic Advantage' in Sovereign Bond Ratings: The Importance of Strong Courts and the Rule of Law. *International Organization*, **66**(3): 515–535.
- Brambor T, Clark W and Golder M** (2006) Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, **14**(1): 63–82.
- Breen M and McMenamin I** (2013) Political Institutions, Credible Commitment, and Sovereign Debt in Advanced Economies *International Studies Quarterly*, **57**(4): 842–854. doi:10.1111/isqu.12079.
- Brooks SM, Cunha R and Mosley L** (2015) Categories, Creditworthiness, and Contagion: How Investors' Shortcuts Affect Sovereign Debt Markets. *International Studies Quarterly*, **59**(3): 587–601.
- Bruner CM and Abdelal R** (2005) To Judge Leviathan: Sovereign Credit Ratings, National Law, and the World Economy. *Journal of Public Policy*, **25**(2): 191–217.
- Cantor R and Packer F** (1996) Determinants and Impact of Sovereign Credit Rating. *Economic Policy Review*, **2**(2): 37–53.
- Carlin W and Soskice D** (2009) German Economic Performance: Disentangling the Role of Supply-side Reforms, Macroeconomic Policy and Coordinated Economy Institutions. *Socio-Economic Review*, **7**(1): 67–99.
- Carruthers BG** (2013) From Uncertainty Toward Risk: The Case of Credit Ratings. *Socio-Economic Review*, **11**(3): 525–551.
- Cusack T** (1999) Partisan Politics and Fiscal Policy. *Comparative Political Studies*, **32**(4): 464–486.
- Easterly W and Rebelo S** (1993) Fiscal Policy and Economic Growth: An Empirical Investigation. NBER Working Paper No. 4499.
- Engen EM and Skinner J** (1996) Taxation and Economic Growth. NBER Working Paper No. 5826.
- Feinerer I and Hornik K** (2017) Tm: Text Mining Package (version 0.7-1). <https://cran.r-project.org/web/packages/tm/index.html>.
- Feldstein M** (1986) Supply-side Economics Old Truths and New Claims. NBER Working Paper 1792.
- Fitch Ratings** (2009) Sovereign Rating Methodology.
- Fitch Ratings** (2016) Sovereign Rating Methodology.
- Garrett G and Lange P** (1991) Political Responses to Interdependence: What's Left For the Left? *International Organization*, **45**(04): 539–564.
- Gray J** (2009) International Organization as a Seal of Approval: European Union Accession and Investor Risk. *American Journal of Political Science*, **53**(4): 931–949.
- Gray J** (2013) *The Company States Keep. International Economic Organization and Sovereign Risk in Emerging Markets*. Cambridge: Cambridge University Press.
- Haque NU, Mathieson N and Mark DJ** (1998) The Relative Importance of Political and Economic Variables in Creditworthiness Ratings. IMF Working Paper, WP/98/46.
- International Monetary Fund** (2010) The Uses and Abuses of Sovereign Credit Ratings. In *Global Financial Stability Report*. Washington, DC: International Monetary Fund, 85–123.
- Kuznetsova A, Brockhoff P and Christensen R** (2017) lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, **82**(13): 1–26.
- Lenth R** (2019) *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R Package Version 1.4.3.01. <https://CRAN.R-project.org/package=emmeans>.
- Lucas E** (1990) Supply-Side Economics: An Analytical Review. *Oxford Economic Papers*, **42**(2): 293–316.
- Luke S** (2017) Evaluating Significance in Linear Mixed-effects Models in R. *Behavior Research Methods*, **49**(4): 1494–1502.
- Makszin K, Medve-Bálint G and Bohle D** (2020) North and South, East and West: Is It Possible to Bridge the Gap?. In Crespy A, Coman R and Schmidt V (eds.), *Governance and Politics in the Post-Crisis European Union*. Cambridge: Cambridge University Press.
- Moody's** (2008) Rating Methodology: Sovereign Bond Ratings.
- Moody's** (2016) Rating Methodology: Sovereign Bond Ratings.
- Mosley L** (2000) Room to Move: International Financial Markets and National Welfare States. *International Organization*, **54**(4): 737–773.
- Mosley L** (2005) Globalisation and the State: Still Room to Move? *New Political Economy*, **10**(3): 355–362.
- Paudyn B** (2013) Credit Rating Agencies and The Sovereign Debt Crisis: Performing the Politics of Creditworthiness Through Risk and Uncertainty. *Review of International Political Economy*, **20**(4): 788–818.

- Potrafke N** (2009) Did Globalization Restrict Partisan Politics? An Empirical Evaluation of Social Expenditures in a Panel of OECD Countries. *Public Choice* **140**: 105. doi:[10.1007/s11127-009-9414-2](https://doi.org/10.1007/s11127-009-9414-2).
- Rodrik D** (2000) How Far Will International Economic Integration Go?. *Journal of Economic Perspectives*, **14**(1): 177–186.
- Rodrik D** (2011) *The Globalization Paradox: Why Global Markets, States, and Democracy Can't Coexist*. Oxford: Oxford University Press.
- Ross F** (2000) 'Beyond Left and Right': The New Partisan Politics of Welfare. *Governance World Politics*, **13**(2): 155–183. doi:[10.1111/0952-1895.00127](https://doi.org/10.1111/0952-1895.00127).
- Sattler T** (2013) Do Markets Punish Left Governments?. *The Journal of Politics*, **75**(2): 343–356.
- Sinclair TJ** (2005) *The New Masters of Capital*. Ithaca, NY: Cornell University Press.
- Standard and Poor's** (2006) Sovereign Credit Ratings: A Primer.
- Standard and Poor's** (2009) Understanding Standard & Poor's Rating Definitions.
- Standard and Poor's** (2013) Sovereign Government Rating Methodology And Assumptions.
- Standard and Poor's** (2016) Rating Stability Criteria.
- Strange S** (1996) *The Retreat of the State: The Diffusion of Power in the World Economy*. Cambridge: Cambridge University Press.
- Streeck W** (2014) *Buying Time: The Delayed Crisis of Democratic Capitalism*. London and Brooklyn: Verso Books.
- Vaaler PM, Schrage BN and Block SA** (2006) Elections, Opportunism, Partisanship and Sovereign Ratings in Developing Countries. *Review of Development Economics*, **10**(1): 154–170.