# Subjectivity in sovereign credit ratings

De Moor, Lieven; Luitel, Prabesh; Sercu, Piet; Vanpée, Rosanne

*Published in:*
Journal of Banking and Finance

*DOI:*
10.1016/j.jbankfin.2017.12.014

*Publication date:*
2018

# Subjectivity in sovereign credit ratings

**Abstract**

A sovereign credit rating is a function of hard and soft information that should reflect the creditworthiness and the probability of default of a country. We propose an alternative characterisation for the subjective component of a sovereign credit rating — the parts related to the ratee's lobbying effort or its familiarity from a United States point of view — and apply it to S&P, Moody's and Fitch ratings, using both traditional ordered-logit panel models and machine learning techniques. This subjective component turns out to be large, especially for the low-rated countries. Countries that are rated as investment grade tend to be positively influenced by it, and *vice versa*. Subjective judgment in credit ratings does have predictive value: it helps in identifying chances of sovereign defaults in the short-term. Still, the impact of subjectivity in sovereign ratings on borrowing costs is very limited on average.
**JEL Classification**: G24, G15, O16
**Keywords**: Credit ratings, rating agencies, artificial intelligence

# 1 Introduction

Sovereign credit ratings result from the analysis of quantitative, qualitative and judgmental factors that affect the creditworthiness of a country. Quantitative economic, political and fiscal variables are modeled by credit rating analysts who provide the results of their analysis to a credit rating committee. The credit rating committee assesses the creditworthiness of the borrowers, trying to be forward looking in its final rating decision. In all of this, there inevitably is model uncertainty and fuzziness, though; also, some inputs are missing, idiosyncratic, or subjective while other aspects are utterly non-quantifiable. All this implies that credit ratings are known to contain a considerable part of subjective judgment, even though the relative importance of the hard *v.* soft inputs is unclear (Bruner and Abdelal, 2005, Vernazza and Nielsen, 2015, Ozturk et al., 2016, Amstad and Packer, 2015).

Subjectivity, while inevitable, may lead to noisy or even biased opinions (Zheng, 2012). Vernazza and Nielsen (2015) conclude that while the objective component of sovereign credit ratings is able to predict sovereign defaults, the subjective component does not seem to help to assess default risks over a horizon of three years and even falsely forecasts defaults for certain time horizons. Their bottom line is that the subjective component in credit ratings is detrimental because it seems to be unrelated to the country's true credit risk.

If so, emerging markets might be the first ones to suffer. For one, modeling issues and data problems are likely to arise especially w.r.t. emerging markets, which creates extra room for qualitative biases (Luitel et al., 2016). Second, any such bias is likely to be downward: the well-documented familiarity bias implies that credit rating analysts will be more prone to favor issuers that are 'closer' to them ('foreign bias'). Recent studies by Dalsgaard and Hirth (2014), Fuchs and Gehring (2016), Gültekin-Karakaş et al. (2011) and Ozturk (2014) demonstrate that sovereign credit ratings assigned by U.S.-based rating agencies exhibit an upward bias for their home country and similar developed countries. Gültekin-Karakaş et al. (2011) likewise show that credit ratings of emerging markets are biased downward. Compounding all this, not all foreign bias may be spontaneous and unconscious. Subjective judgments can be influenced, Luitel et al. (2016) note; and the use of lobby groups, combined with the issuer-pays business model under which rating agencies operate, tend to favor richer countries.

The lack of transparency has become one more source of criticism of the rating process, especially in the wake of the global financial crisis of 2008, in response to which regulators have stepped up. In the U.S., the Dodd–Frank act, signed in July 2010, requires rating agencies to be more transparent about the rating events and rating methodology. In Europe, a regulatory framework for rating agencies was gradually set up as of 2009.[1] One objective is to increase transparency by setting disclosure requirements, including the publFication of a yearly transparency report. Amstad and Packer (2015) state that since 2010 rating agencies do rely more on quantitative inputs in their methodologies that should make sovereign credit ratings more transparent and replicable. However, these methodological improvements are not convincingly confirmed by the empirical analysis they perform.

The central themes of this paper are the degree of replicability of sovereign credit ratings, and the size and role and economic costs of their subjective component. Specifically, we revise the modeling of ratings in general by, first, exploring the link between the subjective component in ratings and generalised 'distance' (or unfamiliarity) and, second, using AI methods next to regression. In addition, we re-evaluate the economic effectiveness and costs of subjective judgments.

Prior work on subjective elements in ratings tends to define them as residuals from a regression of (cardinalised) ratings on the familiar macroeconomic and political variables. This assumes that the researcher knows the true model (which then turns out to be quite tractable) and has all the data. But reality is notoriously non-linear and defies complete modeling. Also, the stand-in models' residuals may reflect objective but idiosyncratic facts rather than errors and inconsistencies. Third, objective covariates measuring economic development are cross-sectionally correlated with familiarity, which means that part of foreign bias is bound to be mislabeled as objectively justified.

In our attempt to further improve this literature, we accordingly decompose sovereign credit ratings into three, not two, components. The first of these has the standard characteristics of an objective ingredient, *i.e.* an association with the standard macroeconomic and political variables that an unbiased, knowledgeable rater would rationally consider. The second component positively exhibits characteristics of a bias, in the sense that it is being related to either the ratee's lobbying power or the rater's lack of familiarity w.r.t. the ratee. That lack of knowledge can be detected,

[1] EC Regulation No 1060/2009 Credit Rating Agencies, known as CRA I and followed by CRA II and CRA III in 2011 and 2013 respectively. For more, see e.g. Amtenbrink and Haan (2009), or, for the U.S., Dimitrov et al., 2015.

in turn, via an empirical link between the rating and familiarity covariates. Unlike the standard measure of subjective judgment, it is defined positively (via an empirical link with unfamiliarity) rather than negatively (the left-over variability in a regression). By adding familiarity covariates, we also reduce the risk that measures of economic development pick up familiarity effects.

About the residual part of the rating, lastly, we prefer to remain agnostic: next to truly irrational judgments, effects of modelling imperfections, and unsystematic errors and inconsistencies it may contain both idiosyncratic events and other missing data. If our aim is to study bias in subjective judgments, as distinct from mere noisiness, it makes sense to focus on the part that is systematic rather than occasional, i.e. positively defined rather than residual. It remains an empirical question whether adding the residual (on top of the explicit part) further sharpens the picture or not.

In short, when identifying subjective judgments we reorient subjectivity from a residual towards also a positively identified component; we verify whether the remaining noise still adds value; and we reduce the room for objective covariates to poach familiarity effects. This fresh look at subjectivity represents one methodological contribution of this paper. A second is that we compare a standard parametric technology (a random-effects ordered-logit panel) with a machine-learning approach based on decision tables, which turns out to be much better not just at replicating or fitting the agencies' opinions but also at predicting actual defaults.[2] This suggests the approach does get us closer to the true model than an ordered-logit classification. Not surprisingly, in that light, the ordered-logit occasionally produces some bizarre implications, especially regarding the role of subjective factors in the rating or the cost of subjectivity. The machine-learning model does not come with any inference apparatus, though; so if the aim is to identify statistically significant determinants of credit ratings, a traditional parametric model remains the only option.

A third contribution of our work is that we measure the economic costs of the subjective part of the credit rating. We estimate the difference between the sovereign spread associated with the actual rating and the spread for a purely objective rating, and apply this to the amount of foreign sovereign debt outstanding. We find that these economic costs are marginal: our calculated cost is quite non-systematic in terms of sign and exhibits no correlation with sovereign spreads. Fourth, we assess whether the alternative measure of subjective judgment of the rating committee does

[2] Unless stated otherwise, in this text 'prediction' means predicting default, while we use 'fitting' for computing a model value for the rating (the dependent variable). In standard regression where the left-hand-side is not a probability or an odds ratio, that distinction is not necessary.

impair the prediction accuracy of the credit rating, as some surmise. We find that, actually, our measure of subjectivity in ratings positively predict sovereign defaults within one and two years, and in one of our model specifications, the subjective component even has a predictive value both in the short and the long-term. Fifth, we show that the size of our measure of subjective component in credit ratings varies across rating notches and over time. The subjective component turns out to be especially large (and downward) for the low-rated countries, but that phenomenon has become less pronounced as of 2010. Puzzlingly, though, at the same time the upward bias for top-rated countries has increased.

The paper is structured as follows: Section 2 reviews the literature on sovereign credit ratings. In Section 3 we describe our data, the variables used to measure the objective and subjective component of a sovereign credit rating and the methods applied to model the ratings. Section 4 presents the empirical results. In Section 5, we measure the subjectivity in credit ratings and assess its economic implications. Section A.4 contains the robustness checks and section 6 concludes.

## 2 Prior literature

There is a vast literature on the sources of variation in credit ratings. We provide an overview of this literature in chronological order in Table 1, and discuss only the articles most closely related to our work.

Early studies mainly attempt to identify the determinants of a country's credit risk. The initial focus tended to be on quantitative macroeconomic variables (McFadden et al., 1985, Saini and Bates, 1984, Cantor and Packer, 1996). An exception is Cosset and Roy (1991), who also add political instability but conclude that this variable has no effect on sovereign ratings. As of the early 2000's, though, researchers tend to find that next to macroeconomic indicators, political variables and institutional quality do play an important role in explaining sovereign credit ratings (see Alexe et al. (2003), Connolly (2007), Mellios and Paget-Blanc (2006) and others). Ozturk (2014), for instance, uses six different governance indicators to capture the total political risk of the sovereign. He finds that institutional quality, as captured by government effectiveness and the quality of the regulatory framework, is an important determinant of sovereign credit ratings.

Recent work, for instance Fuchs and Gehring (2016) and Ozturk (2014), also reviews the sovereign rating methodology itself rather than just its objective determinants. They document that sovereign credit ratings are prone to a qualitative bias: countries that are close to the United States (as measured by geographical, cultural, political and economic distance) are assigned higher credit ratings, reflecting a foreign bias in credit ratings. Ozturk (2014), specifically, finds that sharing a common language influences sovereign credit ratings upwards. This is in line with earlier work by Zheng (2012), who compares the ratings assigned by U.S.-based S&P and Chinese-based Dagong and shows that Dagong assigns higher ratings to non-Western countries than S&P. Because this rating discrepancy cannot be explained by economic factors, Zheng concludes that qualitative factors and subjective criteria must play an important role in the rating decisions of at least one of the two agencies. The home bias in credit ratings is not observed for corporate bonds, though (Güttler and Wahrenburg, 2007).

The literature closest to our own work studies the qualitative component of sovereign credit ratings and investigates the link between credit ratings and borrowing costs, like Fuchs and Gehring (2016), Gande and Parsley (2005) and Vernazza and Nielsen (2015). Gande and Parsley (2005) find that cultural linkage, geographical distance and rule of law do not influence sovereign spreads movements significantly. Fuchs and Gehring (2016), in contrast, conclude that although geographical distance does not explain differences in sovereign credit ratings, there is a culturally driven home bias in credit ratings. Our work is most closely related to Vernazza and Nielsen (2015) who innovatively decompose sovereign credit ratings into an objective and a subjective component and assess the effectiveness of each separately. They define the subjective part as the residual from a standard parametric model.

Relying on subjective judgment in the assessment of credit risk is not specific to the evaluation of sovereign debt. Credit to corporations, especially to small businesses, depends to a great extent on firm-specific subjective information collected by loan officers. Agarwal and Hauswald (2010), for instance, show that the physical distance between the borrower and the lender negatively affects the amount of soft information that is collected by the lender. Given that the inclusion of soft information enhances credit decisions, borrowers that are located farther away pay higher interest rates and are more likely to default. Another strand of this literature considers the impact of securitization, where the initiator does not bear the risk and the ultimate investors buy on the basis of hard information only. Ignoring soft data, interest rates no longer reflect the overall quality of the borrower and statistical default models increasingly fail to predict defaults (Rajan et al., 2015).

In sum, and unlike surmises in early work on sovereign-bond ratings, these studies on corporate

Table 1: Overview of the literature on sovereign credit rating determinants

| Authors | Methodology | CRAs | Dataset | Sample period | Significant variables |
|---|---|---|---|---|---|
| Feder and Uy (1985) | Logistic regression & simulation model | Institutional Investors | 55 developing countries | 1979-83 | International reserves/imports, external debt/GNP, terms of trade, GDP growth, export growth, export vulnerability, GNP per capita/Avg. GNP per capita for the country group, Political dummy variables, Debt service difficulty, Oil dummy variable |
| Brewer and Rivoli (1990) | OLS | Institutional Investors & Euromoney ratings | 30 developing countries | 1967-1986 | Regime change |
| Cosset and Roy (1991) | OLS | Institutional Investors & Euromoney ratings | 77 countries | September 1987 | GNP per capita, propensity to invest & ratio of external debt to exports |
| Lee (1993) | OLS | Institutional Investors | 29 developing countries | Sovereign ratings as of 1986 | Foreign debt/GNP, GDP per capita growth |
| Haque et al. (1996) | Principal component analysis, Logit | II, EIU, Euromoney | 60 developing countries | 1980 -1993 | CA Balance/GDP, International reserves/Imports, GDP growth and Inflation |
| Cantor and Packer (1996) | OLS | S&P & Moody's | 49 countries (28 developed and 21 emerging) | Sovereign ratings as of September 29, 1995 | Per capita income, GDP growth, inflation, external debt, level of economic development and default history |
| Ul Haque et al. (1998) | Stepwise regression | II, EIU, Euromoney ratings | 60 developing countries | 1980 - 1993 | CA Balance/GDP, International reserves/Imports, GDP growth & Inflation |
| McNamara and Vaaler (2000) | Ordinal Logistic regression | S&P, Moody's, Fitch, Duff and Phelps, Thomson Bank Watch, Investment Bank Credit Analysis | 52 countries | 1987 - 1996 | Per Capita Income, GDP growth, Inflation, Fiscal Balance, External Balance, External Debt, Default history, Insurgent agency indicator, Regional specialization & Market structure |
| Monfort and Mulder (2000) | OLS | II, S&P, Moody's | 20 emerging countries | 1995 - 1999 | Debt/Exports, Debt rescheduling dummy, Fiscal balance, GDP growth, Inflation, TOT, Export growth rate, Investment/GDP |
| Mulder and Perrelli (2001) | Pooled OLS, Feasible GLS | S&P and Moody's | 25 emerging countries | 1992-1999 | Investment/GDP, Debt/Exports, Rescheduling history dummy, Fiscal balance, GDP growth, Inflation and Short-term debt/Reserves |
| Hu et al. (2002) | Ordered probit | S&P | 71 countries | 1981-1998 | Previous year default dummy, Debt/Exports, Debt/GNP, Reserves/Imports, Inflation rate, Non-industrial countries dummy |
| Afonso (2003) | OLS | S&P and Moody's | 81 countries (29 developed and 52 developing) | 2001 | GDP per capita, external debt to exports ratio, development indicators, Default indicators, GDP real growth rate & inflation rate |
| Alexe et al. (2003) | Panel regression | S&P | 69 countries | 1998 | Financial depth and efficiency, GDP per capita, Debt/GDP, Political stability & Government effectiveness |
| Block and Vaaler (2004) | GMM | S&P | 19 developing countries | 1987 - 1999 | Political business cycle, inflation, fiscal balance, external debt, previous default |
| Butler and Fauver (2006) | OLS, 2SLS | Institutional Investors, S&P and Moody's | 69 countries | 2009 | Legal environment, GDP per capita, inflation, foreign debt to gdp, underdevelopment index, emerging market dummy |
| Mellios and Paget-Blanc (2006) | Ordered logistic | S&P, Moody's, Fitch | 86 countries | 2003 | Per capita income, Government revenue, Real exchange rate, Inflation rate, Default history & Corruption index |
| Archer et al. (2007) | Panel regression | S&P, Moody's, Fitch | 55 developed countries | 1987-2003 | Trade commitment, GDP growth, inflation, bond default |
| Connolly (2007) | Random effect, Pooled 2SLS | S&P | 52 | 1993-2002 | GDP per capita, GDP growth rate, Previous default, Absolute latitude, Former British colony, %Hindu, %Muslim |
| Remolona et al. (2008) | Fixed effect | S&P, Moody's and Institutional Investors | 27 emerging countries | 2002-2006 | Nominal GDP, GDP per capita, Inflation, External debt/GDP, Original sin, Currency mismatch |
| Afonso et al. (2009) | Ordered probit and ordered logit | S&P and Moody's | 66 | 1996-2005 | GDP per capita, GDP growth, inflation, unemployment, government debt, fiscal balance, government effectiveness, external debt, foreign reserves, current account balance, default history, regional dummies |
| Gaillard (2009) | Ordered probit | Moody's | 105 subnational entities | December 2005 | Sovereign default history, gdp per capita, net direct and guaranteed debt to operating revenue ratio, interest payment to operating revenue ratio |
| Afonso et al. (2011) | Panel regression, ordered probit and random effect ordered probit | S&P, Moody's and Fitch | | 1995-2005 | Gdp per capita, gdp growth, government debt, government deficit, government effectiveness, foreign reserves, sovereign default dummies |
| Ratha et al. (2011) | Pooled OLS | S&P, Moody's and Fitch | 44 emerging countries | Dec 2006 | Gdp growth, GNI per capita, Reserves/(Imports + short term debt), gdp volatility, rule of law |
| Erdem and Varli (2014) | Fixed effects, random effects, ordered probit and random effects ordered probit | S&P | 8 emerging countries | 2002-2011 | Budget balance/gdp, gdp per capita, governance indicators, reserves/gdp |
| Ozturk (2014) | Ordered probit | Moody's | 93 countries | 1999-2010 | Nominal GDP, government expenditures/GDP, government financial balance/GDP, primary balance/GDP |
| Vernazza and Nielsen (2015) | Fixed effects and random effects | Moody's | 99 countries | 1996-2013 | Gdp, gdp growth, public debt, current account, past default, advanced country, government effectiveness and rule of law |

debt ratings conclude that the inclusion of subjective information in the debt evaluation process is positive rather than negative. If subjective judgment leads to better default predictions, including it in the rating process is undeniably valuable. Empirically evaluating the informative value of soft information for sovereigns is precisely what this study is about. In the next section, we present our dataset and describe the variables that we use to measure the objective and the subjective components in credit ratings. We also discuss the methodologies used to model sovereign credit ratings.

## 3 Data and estimation techniques

The sovereign credit rating industry is dominated by three global rating agencies: Moody's, S&P and Fitch. We collected sovereign ratings for 103 countries from 1995Q2 to 2014Q1 for which S&P, Moody's and Fitch issue long term foreign currency sovereign credit ratings. The countries included in our sample are shown in Table 2. The list contains 23 developed and 80 emerging countries.[3] By and large, the developed countries are rated as investment grade (BBB- or higher), while most of the emerging countries are rated as speculative grade (below BBB-). Some exceptions occurred during the global financial crisis (2007-2008) and the European sovereign debt crisis (2009-2011) when certain developed countries were rated below investment grade.

Initially, and in line with earlier work, credit ratings are simply transformed into ordinal numbers with AAA being the highest (AAA = 21) and C the lowest (C = 1) ranked rating. The 'linear' transformation of sovereign credit ratings (*first pass*) does not take into account the yield jump between certain consecutive rating classes, especially at the investment grade border, and more generally assumes that a one-notch change is equally important regardless of the rating. Yet yields tell us this is not the case: at the high end, a one-notch demotion has a much smaller impact than at the low end. To introduce this effect we regress the sovereign 5-year CDS spread on the first-pass rating scale, the squared first-pass rating and an investment-grade dummy:

$$Spread = \alpha + \beta_1 First\ pass + \beta_2 First\ pass^2 + \beta_3 Investment + \epsilon, \quad (1)$$

which compounds the credit risk increases, and especially so at notches where spreads rise fast. All regressors are very significant. We then calculate a fitted value for each first-pass credit

[3]The developed countries are labeled on the basis of the United Nation's country classification (UN 2014: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK and USA); the remaining 80 countries are labeled emerging countries.

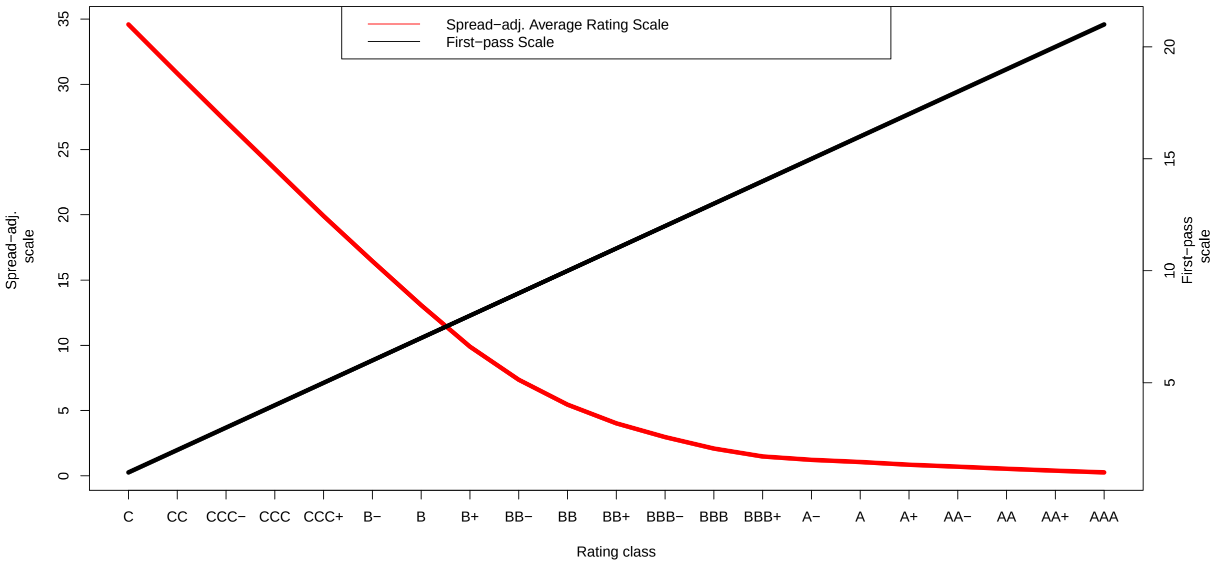



Figure 1: Ordinal transformation of sovereign credit ratings

rating ranging from the highest (AAA = 21) to the lowest (C = 1). A detailed description of this procedure to transform the credit ratings and the resulting fitted and smoothed values are reported in the Appendix (see Section A.1). Figure 1 shows the linear and spread-adjusted rating transformations of the average rating assigned by Moody's, Standard and Poors' and Fitch ratings. In the graph, we smooth away the steps in this fitted value using a locally linear approximation. The numbers are also presented in Table 15 in the Appendix.

### 3.1 Variables associated with sovereign credit ratings

Our choice of macroeconomic and political variables to pick up the variation in credit ratings starts from the literature on the determinants of sovereign credit ratings (see Table 1). We classify our explanatory variables into three categories: (*i*) macroeconomic fundamentals and fiscal strength, (*ii*) political risk and governance and (*iii*) subjective judgment of the rating committee. Table 3 contains a description of each variable together with its source.[4]

Regarding groups (*i*) and (*ii*), the variables are familiar from prior literature, so we refer to the brief definitions in Table 3 and the summary statistics in Table 4.[5] Winsorisation at 90%

[4]Table 16 and 17 list the Pearson correlations and collinearity test statistics between the variables.

[5]Some of the annual data (GDP per capita, GDP growth, Current account, Budget balance and External debt) are converted to quarterly data using Denton-Cholette interpolation using "tempdisagg" R package (Denton, 1971, Dagum and Cholette, 2006, Sax and Steiner, 2013). This technique is briefly explained in the online appendix

Table 2: Sample countries

| Country | Country | Country | Country | Country |
|---|---|---|---|---|
| Albania | Cyprus | Italy | Netherlands | Slovakia |
| Angola | Czech Republic | Jamaica | New Zealand | Slovenia |
| Argentina | Denmark | Japan | Nicaragua | South africa |
| Armenia | Dominican Republic | Jordan | Nigeria | Spain |
| Australia | Ecuador | Kazakhstan | Norway | Sri Lanka |
| Austria | Egypt | Kenya | Oman | Sweden |
| Azerbaijan | El Salvador | Korea (South) | Pakistan | Switzerland |
| Bahrain | Estonia | Kuwait | Panama | Thailand |
| Bangladesh | Finland | Latvia | Papua New Guinea | Trinidad and Tobago |
| Belarus | France | Lebanon | Paraguay | Tunisia |
| Belgium | Germany | Lithuania | Peru | Turkey |
| Bolivia | Ghana | Luxembourg | Philippines | Uganda |
| Brazil | Greece | Madagascar | Poland | Ukraine |
| Bulgaria | Guatemala | Malaysia | Portugal | United Arab Emirates |
| Burkina Faso | Honduras | Mali | Qatar | United Kingdom |
| Canada | Hungary | Malta | Romania | United States |
| Chile | Iceland | Mexico | Russia | Venezuela |
| China | India | Moldova | Saudi Arabia | Vietnam |
| Colombia | Indonesia | Mongolia | Senegal | Zambia |
| Costa rica | Ireland | Morocco | Serbia | |
| Croatia | Israel | Namibia | Singapore | |

and, for some variables, log transformation are adopted to dampen outlier effects. Financial risk, institutional quality and governance are measured as the difference of the log of the original data compared to that for the U.S.. Group (*iii*), next, contains covariates that are meant to pick up any systematic subjectiveness, an approach motivated in the Introduction. For these, namely proximity of the country to the United States and the lobbying power of the sovereign, a longer discussion is provided in the subsections that follow.

### 3.1.1 Proximity to the United States

We measure the proximity to the U.S. in three different dimensions: cultural proximity, economic proximity and geographical distance. Cultural affinity between the U.S. and other countries, first, is proxied for by measures of commonality in language or religion. Language data is obtained from the CEPII; from that, we construct a dummy variable that is equal to unity if English is the mother tongue of at least 20% of the population, and zero otherwise.[6] Religious proximity (or religious similarity) is defined as the log probability that two randomly chosen individuals in

containing the dataset descriptions.

[6]CEPII provides an open-source database. The data can be extracted from http://www.cepii.fr/cepii/en/bdd_modele/bdd.asp.

Table 3: Definition and source of the variables explaining sovereign credit ratings

| Variables | Expected sign | Definition | Source |
|---|---|---|---|
| **Macroeconomic fundamentals and fiscal strength** | | | |
| Log GDP per capita | – | Natural logarithm of nominal GDP in US$ divided by mid year population | Thomson Reuters Eikon |
| GDP growth | – | Quarterly real GDP growth rate | Thomson Reuters Eikon |
| Current account/GDP | – | Current account balance in million US$ (as % of nominal GDP) | Thomson Reuters Eikon |
| Budget balance/GDP growth | – | Growth rate of Budget surplus or deficit balance in million US$ (as % of nominal GDP) | Thomson Reuters Eikon |
| External debt/GDP | + | Gross external debt position in million US$ (as % of nominal GDP) | Thomson Reuters Eikon, World Bank QEDS |
| International reserves | – | Natural logarithm of foreign currency reserves of the government | Thomson Reuters Eikon |
| Trade | – | External trade of the country in million US$ (as % of nominal GDP) | Thomson Reuters Eikon |
| Previous default | + | Exponential Decay variable. 1 for defaulted year $t_0$ and exponentially decay at the rate of 20% till year $t+4$* | Standard and Poors and Moody's Default Database |
| Unemployment rate | + | Quarterly unemployment rate | Thomson Reuters Eikon |
| Financial risk | + | Difference of natural logarithm value of financial risk between particular country and the USA** | International country risk guide Table 5B |
| **Political Risk and Governance** | | | |
| Institutional quality | – | Difference of natural logarithm value of institutional quality between particular country and the USA** | International country risk guide Table 3B |
| Governance | – | Difference of natural logarithm value of governance between particular country and the USA** | International country risk guide Table 3B |
| **Subjective judgment of the rating committee** | | | |
| Lobbying power | – | Natural logarithm of total lobby amount spent by a particular country for lobbying in the USA. | Foreign Agents Registration Act Reports, U.S. Department of Justice |
| Economic proximity: | | | |
| Trade proximity | – | Trade intensity of a country with the USA | Thomson Reuters Eikon |
| Cultural proximity: | | | |
| Common language | – | Dummy variable: 1 if more than 20% of population speaks English as mother tongue, zero otherwise | CEPII |
| Religious proximity | – | The probability that two randomly chosen individuals in the US and particular country share the same religion | World Religion Database |
| Nearest geographical distance | + | Physical distance (in km) based on latitude and longitude from the New York City (U.S.) to the nearest border city of a country divided by 100 | MaxMind |

*If the country has defaulted in 2002-Q1 then we assign 1 to the particular country in 2002-Q1, 0.80 in 2002-Q2, 0.64 in 2002-Q3 and so on till fourth quarter of 2006. Default includes both foreign currency and domestic currency sovereign default reported by Moody's and Standard and Poors'.
**Higher numbers of International country risk guide (ICRG) ratings imply a better situation of the country. Therefore, we multiply Financial risk by -1 for the ease of interpretation,i.e, a lower value of Financial Risks reflects less risk in the country. Financial risk includes: foreign debt service/exports, net international liquidity excluding foreign liabilities/average imports costs and exchange rate stability. Institutional quality includes: law and order, bureaucracy quality, democratic accountability and corruption. Governance includes: government stability, socio-economic conditions and investment profile.

the U.S. and the foreign country share the same religion. We work with four broad groups; in alphabetical order: Christianity, Islam, Judaism and Other Religion/atheist. If $p(r,k)$ denotes the fraction of the population in country $k$ that has religion $r$, our measure for country pair $(k,l)$ is defined as

$$Rprox(k,l) := \ln\big[\sum_{r=1}^{4} P(r,k)\cdot P(r,l)\big]. \tag{2}$$

Following in the portfolio-choice and corporate-bond literatures, we next conjecture that, if there is any foreign bias in credit ratings, sovereign ratings assigned by US-based agencies would be lower the more geographically distant the country is from the U.S. We measure geographical distance as the physical distance from New York City to the nearest city of the respective country.[7]

As our key ingredient for economic proximity, lastly, we choose the intensity of the trade relation between the U.S. and the respective country. If the U.S. has a more intense trade with country $k$, we hypothesize, $k$'s sovereign credit rating will be more favorable. More specifically, trade proximity is measured as follows:

$$Trade\ proximity_{j,t} := \frac{Imports_{USA,j,t} + Exports_{j,USA,t}}{Total\ Trade_{USA,t}} \tag{3}$$

where $Imports_{USA,j,t}$ is the total value of imports in the USA from country $j$ in quarter $t$; $Exports_{j,USA,t}$ is the total value of exports from the USA to country $j$ in quarter $t$ and $TotalTrade_{USA,t} = Total\ Imports_{USA,t} + Total\ Exports_{USA,t}$.

### 3.1.2 Lobbying power

Besides a familiarity effect, also the lobbying activities of the borrowers may influence the subjective judgment of the rating committee. Both Fuchs and Gehring (2016) and Borensztein et al. (2013) describe how governments put pressure on rating agencies to obtain a better credit rating and reduce the high economic costs of sovereign default risk. We proxy the lobbying power of a country by the amount a country spends on lobbying in the U.S..

## 3.2 Model Framework

Sovereign credit ratings are usually modeled by applying an ordered logit model or by fixed- or random-effects panel models. The merits of such a parametric technique to model sovereign credit ratings include its simplicity and the ease of interpretation of the estimation results. Unlike an AI-type model like the one discussed below, for each potential determinant of a country's credit risk

[7] The physical distance is calculated using Haversine great circle distance formula. The latitude and longitude data for the world cities can be extracted from https://www.maxmind.com/en/free-world-cities-database.

Table 4: Summary Statistics

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| Log GDP per capita | 5,213 | 7.554 | 1.353 | 4.325 | 10.262 |
| GDP growth | 5,213 | 0.009 | 0.008 | −0.036 | 0.069 |
| Current account/GDP | 5,213 | −0.006 | 0.104 | −1.230 | 0.762 |
| Budget balance/GDP | 5,213 | −0.028 | 0.594 | −17.272 | 0.231 |
| Budget balance/GDP growth | 5,213 | −0.208 | 10.476 | −701.756 | 14.972 |
| External debt/GDP | 5,213 | 1.272 | 3.861 | 0.000 | 53.443 |
| International Reserves | 5,213 | 9.200 | 1.877 | −1.741 | 15.031 |
| Trade | 5,213 | 0.673 | 0.384 | 0.126 | 3.567 |
| Previous default | 5,213 | 0.017 | 0.099 | 0.000 | 1.000 |
| Unemployment rate | 5,213 | 0.081 | 0.046 | 0.003 | 0.376 |
| Financial risk | 5,213 | −0.126 | 0.154 | −0.485 | 1.730 |
| Institutional quality | 5,213 | −0.332 | 0.293 | −1.204 | 0.165 |
| Governance | 5,213 | −0.201 | 0.189 | −0.859 | 0.196 |
| Lobbying power | 5,213 | 4.056 | 2.493 | 0.000 | 7.881 |
| Trade proximity | 5,213 | 0.011 | 0.025 | 0.000 | 0.197 |
| Common language | 5,213 | 0.192 | 0.394 | 0 | 1 |
| Religious proximity | 5,213 | −1.039 | 1.101 | −4.435 | −0.206 |
| Nearest geographical distance | 5,213 | 73.397 | 33.942 | 0.000 | 153.203 |
| First-pass Average Ratings | 5,213 | 13.507 | 5.012 | 2 | 21 |
| Spread adj. Average Ratings | 5,213 | 4.529 | 5.383 | 0.259 | 30.844 |
| First-pass Moody's Ratings | 4,992 | 13.748 | 5.120 | 1 | 21 |
| Spread adj. Moody's Ratings | 4,992 | 4.453 | 5.886 | 0.330 | 37.255 |
| First-pass Standard Poors' Ratings | 4,850 | 13.768 | 4.971 | 1 | 21 |
| Spread adj. Standard Poors' Ratings | 4,850 | 4.543 | 6.129 | 0.379 | 39.974 |
| First-pass Fitch Ratings | 4,222 | 14.324 | 4.843 | 1 | 21 |
| Spread adj. Fitch Ratings | 4,222 | 3.427 | 3.954 | 0.070 | 26.538 |

the logit model delivers an explicit coefficient estimate, allowing us to assess the relative importance of each factor and its statistical significance in explaining credit ratings. An important downside of linear estimation techniques is their mediocre fit, both in-sample and out-of-sample. McFadden R-squares (or pseudo R-squares) are generally around 50 to 55 percent and the percentage of correctly fitted ratings is relatively low (see also Section 4.3).

This mediocre fit is linked to the fact that credit ratings contain a specific form of nonlinearity that is rather hard to model with parametric statistical models. More specifically, the switch from speculative grade to investment grade is only one notch of a difference in terms of credit rating, but the economic impact of this switch is enormous (Eijffinger, 2012). We control for this to some extent through non-linear transformation of the first-pass score and by including an investment grade dummy to the rating class' concurrent spread (see Section 3), but it is not *a priori* clear whether this suffices. Apart from nonlinearity issues, a sovereign credit rating also incorporates qualitative and judgmental factors, which are hard to incorporate into a parametric model. If machine learning techniques have become very popular for the modeling of credit risk, their ability to handle nonlinearity and qualitative data is the prime reason.

Machine learning is, specifically, often applied for calculating credit scores for banking purposes (Bahrammirzaee, 2010), for modeling corporate bond ratings (Huang et al., 2004) and for assessing country risk (Van Gestel et al., 2006). Widely used machine learning methods for classification include boosting, random forests and support vector machines. A survey of the plethora of methods is beyond the scope of this paper; we will briefly explain only the method that we selected, a decision-trees-based classification approach called 'random trees' or 'random forests' (Section 3.2.2). The random forests algorithm handles imbalanced data and is robust to outliers (Chen et al., 2004, Trevor Hastie, 2009). It aggregates the conclusions from multiple decision trees, maintaining the advantages of individual decision trees but increasing prediction accuracy via averaging. But, as mentioned, the approach produces no regression coefficients, so that an assessment of economic and statistical significance is out of the question.

As parametric linear estimation models and nonparametric machine learning techniques each have manifest strengths as well as weaknesses, we apply both methods for modeling sovereign credit ratings and compare the estimation results and the computed subjective component based on each methodology. The specific methods applied are discussed in the subsections below.

#### 3.2.1 Random effects ordered logit regression

Ordered logit or proportional odds models are based on the logit of cumulative probabilities of each response, which is a linear function of covariates (McCullagh, 1980). An ordered logit model with $c$ categories of credit ratings will construct an optimal scoring rule $cr^*$, linear in the observed characteristics $X$,

$$cr^*_{it} = \alpha + \beta' X_{it} + \epsilon_{it} \tag{4}$$

with the property that the level of the score $cr^*$ tells us which rating is most likely:

$$cr_{it} = \begin{cases} C, & \text{if } cr^*_{it} \leq \mu_1 \\ CC, & \text{if } \mu_1 < cr^*_{it} \leq \mu_2 \\ \vdots \\ AAA, & \text{if } \mu_c < cr^*_{it} \end{cases} \tag{5}$$

Random-effects ordered logit is an extension of the proportion odds model, also called the random-intercept cumulative logit model. It is appropriate for the analysis of correlated ordinal responses (Agresti and Natarajan, 2001, Hedeker, 2008). The error terms are possibly autocorrelated and/or correlated across countries. The random-effects model is able to take into account these non-iid features.

The baseline regression analysis is estimated as a linear random effects model for each rating agency separately and for the average rating of the three credit rating agencies:

$$logit[P(cr_{it} \leq c)] = \sum_{k=1}^{K} \beta_k x_{kit-1} + \sum_{m=1}^{M} \gamma_m z_{mi} + \tau_i + \epsilon_{it}. \tag{6}$$

where $x$ and $z$ are, respectively, time-varying and time-invariant variables, $\epsilon_{it}$ is an error term, $\beta$ and $\gamma$ represent the estimated parameters, and $\tau_i$ is the random intercept for country $i$. The odds ratio of observing sovereign credit rating $(cr_{it})$ of country $i$ at time $t$ is projected on the following $K$ time-varying lagged variables and $M$ time invariant variables: DUMMIED OUT

$x$ : (Log GDP per capita, GDP growth, Current account balance/GDP,
Budget balance/GPD growth, External debt/GDP, International reserves,
Trade, Previous default, Unemployment, Financial risk,
Institutional quality, Governance,
Lobbying power, Trade proximity, Religious proximity )

$z$ : (Common language, Geographical distance)

$i$ : $\{1 .. 103\}$, 103 sample countries from Table 2

$t$ : $\{1995Q2 .. 2014Q1\}$

As an alternative to the random effects model, we also estimate the model using fixed effects and panel corrected standard errors (PCSE). The estimation results are comparable and are shown in the Appendix.

#### 3.2.2 Random forests classification

In addition to linear models, we also model credit ratings using a supervised machine learning algorithm, random decision forests, recently introduced in this literature by Ozturk et al. (2016). The random decision forests algorithm employs an ensemble of de-correlated trees, and it predicts a variable by averaging the predictions of these independent trees (Breiman, 2001). This ensemble method is based on randomization where each tree in the forest is built randomly with the same distribution. Each tree is trained in isolation. The basic procedure involves three steps (Trevor Hastie, 2009, p.588), illustrated in Figure 2. First, a bootstrap sample is selected from

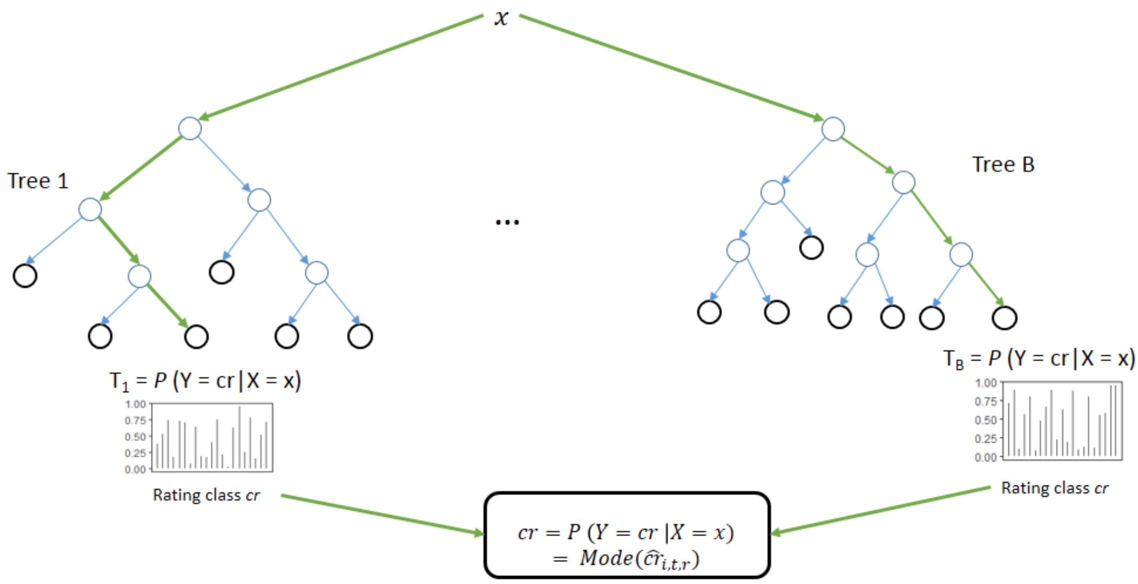



Figure 2: The random forest algorithm - multiple decision trees
*Note* This figure shows graphical representation of the random forests machine learning algorithm.

the training data. Second, we grow a tree to the bootstrapped data, and select a set of variables $(m)$ randomly from all listed variables $(p)$. In that sample we pick the best variables from $m$ and keep splitting each node into two daughter nodes until the minimum node size is reached. That second step is repeated for each of the, in all, $B$ random trees we have decided to grow. Every tree develops its own best classification rule, but across trees the rules obviously differ, as they each work with a different set of objects to classify and a different set of characteristics. So when we let each tree classify *all* observations, they will never agree. But the set of classifications of object $(i, t)$ across all rules $r$, $\widehat{cr}_{i,t,r}$ can be assembled into a probability distribution, whose mode then gives us the random-forest's proposed rating. That third step is in line with the logit approach: logit gives us a distribution, and we pick the most likely value:

$$cr_{i,t,r} = \text{Mode}(\widehat{cr}_{i,t,r}). \tag{7}$$

# 4 Empirical Results

In this section we present the estimation results of each of the above two approaches, as applied to four panels. Three of these contain the ratings by each credit rating agency, while the fourth data-set contains the average of the ratings of the three agencies. For conciseness, we only present the estimation results of spread-adjusted ratings in the main text. Tables and figures containing estimation results of the unscaled ratings (i.e., AAA = 21 and C = 1) are reported in the Appendix (see Table 18). We discuss the differences between the rating agencies and scaling methods in the text, when relevant.

## 4.1 Ordered logit estimation results

The estimation results of the random-effects ordered logit regression are reported in Table 5. In the first output column we provide the results for the average ratings with on the right-hand side just the macroeconomic fundamentals, political risks and governance (column 2). Gradually we include the measures of the subjective components and exclude insignificant variables based on a forward step-wise selection approach, finally giving us the second output column. The only variable that is deleted based on this approach is budget balance to GDP. That also applies for the firm-by-firm results shown in the last three columns.

As can be seen in Figure 1, the spread-based rating transformation results in high values for low rating classes and vice versa. Empirically, the *objective component* of a sovereign credit rating loads on GDP per capita (expected –, observed –), external debt/GDP (expected +, observed +), trade (expected –, observed –), previous default (expected +, observed +), unemployment (expected +, observed +), financial risks (expected +, observed +), institutional quality (expected –, observed –) and governance (expected –, observed –). Most signs are as expected, except for GDP growth (for which four of the five coefficients are also significant) and current-account balance/GDP (three coefficients significant). As for the rating's *subjective component*, we find evidence for both a familiarity effect and successful lobbying activities: for each of the four estimations, the rating variable loads on lobbying power (expected –, observed –), common language (expected –, observed –), trade proximity (expected –, observed –), religious proximity (expected –, observed –) and nearest geographical distance from the New York City (expected +, observed +). All signs are as expected, this time. The marginal effects at the mean of the independent variables for each rating class is shown on Figure 9 in an appendix. We conclude that there probably is more to rating than objective information. Before asking whether that is bad, we seek further evidence via the random-tree approach.

## 4.2 Random forests classification results

Random forests classification does not allow a discussion of coefficient estimates because it is not possible to measure how much each variable has contributed to the final split-up into all final

Table 5: Random Effects ordered logit estimation results for the determinants of sovereign credit ratings

| | | *Dependent: Spread adj. scaled ratings* | | | | |
|---|---|---|---|---|---|---|
| | Expected sign | Average | Average | Moody's | S&P | Fitch |
| Log GDP per capita | − | −1.91*** | −2.17*** | −1.90*** | −2.89*** | −2.45*** |
| | | (0.05) | (0.05) | (0.05) | (0.06) | (0.06) |
| GDP growth | − | 7.04* | 5.28 | 8.55** | 12.64*** | 8.45* |
| | | (3.89) | (4.21) | (3.97) | (4.41) | (4.60) |
| Current account/GDP | − | 0.40 | 0.50* | 0.29 | 2.30*** | 1.98*** |
| | | (0.29) | (0.29) | (0.28) | (0.42) | (0.44) |
| Budget balance/GDP growth | − | 0.00 | | | | |
| | | (0.00) | | | | |
| External debt/GDP | + | 0.22*** | 0.27*** | 0.10*** | 0.34*** | 0.39*** |
| | | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) |
| International Reserves | − | −0.07*** | 0.07*** | −0.07*** | 0.49*** | 0.16*** |
| | | (0.02) | (0.03) | (0.03) | (0.03) | (0.03) |
| Trade | − | −1.00*** | −0.78*** | −1.02*** | −0.37*** | −1.06*** |
| | | (0.10) | (0.10) | (0.11) | (0.11) | (0.11) |
| Previous default | + | 3.55*** | 3.59*** | 2.62*** | 3.14*** | 3.53*** |
| | | (0.34) | (0.34) | (0.32) | (0.38) | (0.45) |
| Unemployment rate | + | 22.94*** | 19.74*** | 17.12*** | 26.02*** | 30.92*** |
| | | (0.77) | (0.78) | (0.76) | (0.91) | (0.99) |
| Financial risk | + | 3.48*** | 3.20*** | 2.05*** | 3.31*** | 4.48*** |
| | | (0.22) | (0.22) | (0.20) | (0.23) | (0.27) |
| Institutional quality | − | −4.97*** | −4.07*** | −3.76*** | −3.15*** | −3.09*** |
| | | (0.16) | (0.17) | (0.16) | (0.18) | (0.20) |
| Governance | − | −7.46*** | −8.29*** | −7.29*** | −8.87*** | −7.57*** |
| | | (0.27) | (0.29) | (0.27) | (0.30) | (0.31) |
| Lobbying power | − | | −0.09*** | −0.10*** | −0.10*** | −0.11*** |
| | | | (0.01) | (0.01) | (0.01) | (0.02) |
| Trade proximity | − | | −16.72*** | −23.14*** | −14.69*** | −8.06*** |
| | | | (1.64) | (1.64) | (1.81) | (1.52) |
| Common language | − | | −4.04*** | −1.12*** | −1.58*** | −1.60*** |
| | | | (0.12) | (0.10) | (0.10) | (0.11) |
| Religious proximity | − | | −0.22*** | −0.42*** | −1.19*** | −0.39*** |
| | | | (0.03) | (0.03) | (0.04) | (0.03) |
| Nearest geographical distance | + | | 0.01*** | 0.01*** | 0.01*** | 0.02*** |
| | | | (0.00) | (0.00) | (0.00) | (0.00) |
| AAA\|AA+ | | −25.63*** | −23.62*** | −21.44*** | −24.41*** | −23.89*** |
| | | (0.53) | (0.54) | (0.50) | (0.58) | (0.63) |
| AA+\|AA | | −23.48*** | −21.47*** | −20.57*** | −22.03*** | −21.83*** |
| | | (0.51) | (0.51) | (0.49) | (0.57) | (0.61) |
| AA\|AA− | | −21.14*** | −19.00*** | −18.59*** | −19.89*** | −18.81*** |
| | | (0.49) | (0.49) | (0.48) | (0.55) | (0.59) |
| AA−\|A+ | | −19.86*** | −17.64*** | −17.62*** | −18.00*** | −17.16*** |
| | | (0.48) | (0.49) | (0.47) | (0.53) | (0.57) |
| A+\|A | | −18.75*** | −16.48*** | −16.04*** | −16.57*** | −15.51*** |
| | | (0.47) | (0.48) | (0.46) | (0.52) | (0.56) |
| A\|A− | | −17.15*** | −14.79*** | −14.76*** | −14.52*** | −13.99*** |
| | | (0.46) | (0.47) | (0.45) | (0.51) | (0.55) |
| A−\|BBB+ | | −15.74*** | −13.26*** | −13.94*** | −12.94*** | −12.34*** |
| | | (0.45) | (0.46) | (0.44) | (0.50) | (0.54) |
| BBB+\|BBB | | −14.65*** | −12.07*** | −12.69*** | −11.62*** | −10.81*** |
| | | (0.44) | (0.45) | (0.43) | (0.49) | (0.53) |
| BBB\|BBB− | | −13.16*** | −10.48*** | −11.72*** | −9.85*** | −9.05*** |
| | | (0.43) | (0.44) | (0.43) | (0.47) | (0.51) |
| BBB−\|BB+ | | −11.07*** | −8.38*** | −9.59*** | −8.11*** | −6.83*** |
| | | (0.42) | (0.42) | (0.41) | (0.45) | (0.50) |
| BB+\|BB | | −9.18*** | −6.57*** | −7.82*** | −6.49*** | −4.56*** |
| | | (0.40) | (0.41) | (0.40) | (0.44) | (0.48) |
| BB\|BB− | | −7.26*** | −4.75*** | −6.39*** | −4.31*** | −3.14*** |
| | | (0.39) | (0.41) | (0.39) | (0.44) | (0.49) |
| BB−\|B+ | | −5.62*** | −3.15*** | −5.31*** | −2.18*** | −1.00** |
| | | (0.39) | (0.41) | (0.39) | (0.45) | (0.50) |
| B+\|B | | −2.84*** | −0.37 | −3.00*** | 0.24 | 1.33*** |
| | | (0.39) | (0.42) | (0.40) | (0.45) | (0.51) |
| B\|B− | | −0.65* | 1.85*** | −1.12*** | 2.12*** | 3.68*** |
| | | (0.39) | (0.42) | (0.41) | (0.47) | (0.52) |
| B−\|CCC+ | | 1.87*** | 4.28*** | 1.02** | 4.38*** | 8.03*** |
| | | (0.40) | (0.43) | (0.41) | (0.48) | (0.54) |
| CCC+\|CCC | | 3.67*** | 5.97*** | 4.99*** | 5.80*** | 8.55*** |
| | | (0.43) | (0.46) | (0.47) | (0.50) | (0.55) |
| CCC\|CCC− | | 4.64*** | 6.96*** | 6.13*** | 5.97*** | |
| | | (0.46) | (0.49) | (0.50) | (0.51) | |
| CCC\|CC | | | | | | 10.91*** |
| | | | | | | (0.71) |
| CCC−\|CC | | 7.27*** | 9.64*** | 7.59*** | 6.00*** | |
| | | (0.72) | (0.74) | (0.61) | (0.51) | |
| CC\|C | | | | 7.84*** | 6.26*** | 12.53*** |
| | | | | (0.63) | (0.52) | (1.14) |
| McFadden $R^2$ | | 0.521 | 0.527 | 0.490 | 0.529 | 0.555 |
| *Note:* | | | | | | ***p<0.01; **p<0.05; *p<0.1 |

nodes – not well put.[8] What we can do, though, is study the impact on the number of correct classifications after randomly permuting the observations for one particular variable, say log GDP per capita, and applying the classification rule using these partially garbled data. We lastly divide that revised success number by the success number when the rule is fed correct information. Figure 3 depicts this measure of relative importance of each predictor. The variables included to pick up subjectivity are marked in bold. We observe that the rating agencies assign slightly different weights to the variables under consideration, but there seems to be a consensus about the most important determinants of sovereign ratings.[9] Below we discuss the two groups (objective versus subjective).

The covariates that are designed to pick up the *objective* component are clearly the most important. In each of the four series, the three to six top-ranking variables are of this type. Among these, GDP per capita, institutional quality, governance, external debt and international reserves turn out to be the influencing determinants of a sovereign credit rating. If we randomly permute the GDP per capita data, the mean decrease in the number of correct classifications of credit rating classification over all trees is above 35%, which represents a substantial drop in quality.

That said, the variables proxying for the subjective judgment of the rating committee are still quite influential. With minor permutations across rating agencies, the highest-ranking are geographical distance, trade proximity, religious proximity and lobbying power, with mean decreases in successful classifications of at least 7%, and about 20% for the highest ranking. Thus, the ranking of the variables in Figure 3 indicates that both the objective and the subjective variables (in that order) are important components of the overall sovereign credit rating. We next show that the same conclusion holds when we let the models try their hand at actual failure predictions rather than published ratings.

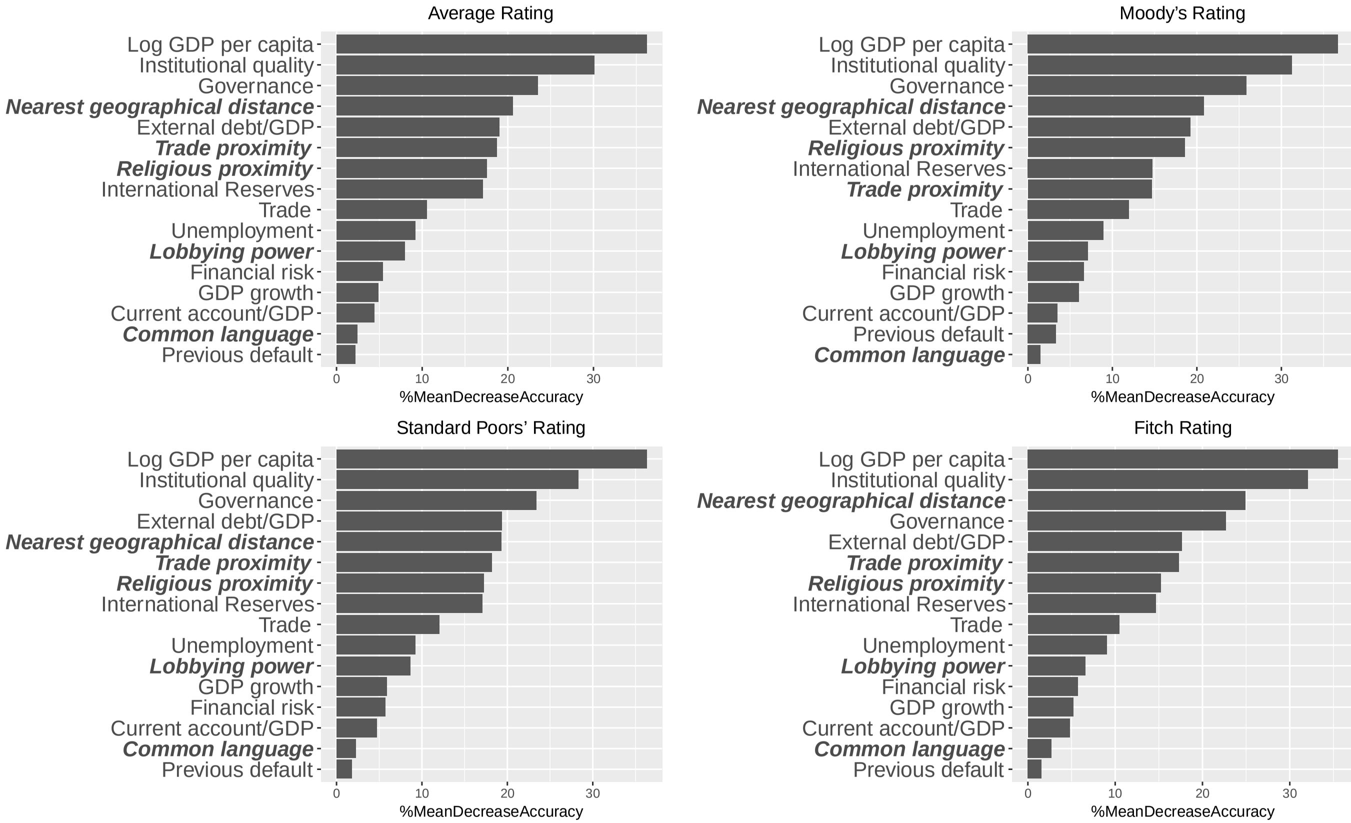



Figure 3: Variable importance for the random forest model

This figure shows the percentage decrease in the mean prediction accuracy if the respective variable is omitted from the model. Covariates that should pick up subjectivity in the rating are printed in boldface.

Table 6: Predictive validity of the competing models

| CRAs | Type | *Panel A: Whole sample* Nobs | % predicted within $n$ notches $n=0$ | $n=1$ | $n=2$ | $n=3$ | *Panel B: Out of sample (70/30% split)* Nobs | % predicted within $n$ notches $n=0$ | $n=1$ | $n=2$ | $n=3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Moodys | OL | 4992 | 51.08 | 84.17 | 95.79 | 98.02 | 1487 | 49.09 | 81.44 | 95.29 | 97.98 |
| Std.Poors | OL | 4850 | 57.22 | 90.23 | 96.41 | 98 | 1447 | 55.91 | 89.15 | 95.72 | 97.86 |
| Fitch | OL | 4222 | 57.7 | 91.21 | 97.49 | 98.74 | 1257 | 55.61 | 90.61 | 96.82 | 98.73 |
| Avg.Rating | OL | 5213 | 54.36 | 89.28 | 96.82 | 98.64 | 1554 | 52.64 | 89.06 | 96.4 | 98.71 |
| Moodys | RF | 4992 | 93.09 | 98.26 | 99.46 | 99.72 | 1487 | 91.53 | 97.85 | 99.19 | 99.6 |
| Std.Poors | RF | 4850 | 92.68 | 98.58 | 99.32 | 99.57 | 1447 | 90.53 | 97.65 | 98.76 | 99.31 |
| Fitch | RF | 4222 | 92.21 | 98.48 | 99.36 | 99.55 | 1257 | 90.61 | 98.09 | 99.28 | 99.28 |
| Avg.Rating | RF | 5213 | 91.85 | 98.93 | 99.46 | 99.65 | 1554 | 91.25 | 98.39 | 99.03 | 99.55 |

**Note:** This table shows the percentage of correctly predicted ratings for each model. For the logit models, we used model specification 2, which includes only significant explantory variables based on a stepwise selection approach. OL refers to random effects ordered logit model scaled on the fitted sovereign CDS spread as a function of the first-pass score and the investment grade dummy (see Figure 1 and Table 15). RF stands for random forests. Panel A shows the predictive performance for the whole sample while Panel B shows the out-of-sample predictive performance. For the out-of-sample test, we randomly selected 70% of our data to train the model. The predictions are based on the remaining 30% of the data.

Table 7: Prediction accuracy reported by previous studies

| Study | Method | CRA | Test | % predicted within $n$ notches $n=0$ | $n=1$ | $n=2$ |
|---|---|---|---|---|---|---|
| Erdem and Varli (2014) | Linear RE panel | S&P | whole sample | 42.81% | 51.25% | 64.06% |
| Ozturk et al. (2016) | Random Forests | Moody's | whole sample | 79.26% | 91.88% | 97.36% |
| Van Gestel et al. (2006) | Support Vector Machine | Moody's | out-of-sample | 50.68% | 81.02% | 95.11% |
| Bennell et al. (2006) | Neural Network | Average | out-of-sample | 36.70% | 73.50% | 89.80% |

**Note:** This table shows the prediction accuracy reported by previous studies that applied linear models or machine learning techniques to predict sovereign credit ratings.

## 4.3 Models' abilities to fit the actual ratings

In the preceding section we used the number of correct classifications as a scaler; we now study that performance measure for its own sake. We also rely on slightly broader measures, by counting how often the rule was off-mark by no more than one (or two or three) notches. The results are assembled in Table 6.

A first observation from that table is that the percentage of correctly fitted ratings is familiarly non-stellar for the random effects ordered logit models, ranging from 50% to 58%. That proportion is much more impressive for the random forests models, though: in the in-sample test, the exact fits

[8]For the random decision trees classification, we grow between 30 and 1000 trees and use 3 to 6 variables at each split. The sole purpose of using limited variables to split a node is to create de-correlated trees. We find that using between 30 and 80 trees and using 5 variables to create a node split, the global misclassification error keeps decreasing from 9% to 8%. Beyond the 80$^{\text{th}}$ tree there is no further reduction in the misclassification error. Therefore, we stop at the 80$^{\text{th}}$ tree to avoid over-fitting.

[9]The Spearman rank correlation of variable importance between all rating agencies are above 98% and significant at the 99% confidence level.

of the machine learning models amount to 93.09% for Moody's, 92.68% for S&P, 92.21% for Fitch and 91.85% for the average ratings. Those numbers are in-sample, though, so they may be too optimistic about the performance out of sample. For the purpose of calculating the out-of-sample fit's precision, we randomly select 70% of the data from each rating class as training data and the remaining observations are used as test data. The resulting out-of-sample accuracies hardly differ from the in-sample ones. The random forests models continue to outperform the logit models also out-of-sample, with correctly fitted ratings equal to 91.53% (Moody's), 90.53% (S&P), 90.61% (Fitch) and 91.25% (Average ratings). We conclude that random forests models perform much better than logit models both in the whole-sample tests and out of sample. Within one notch off, their fitting ability is above 97% for both the whole-sample and out-of-sample predictions, up from 81%-91% for the logit model.[10]

The predictive validity of our models can be compared with the prediction precision reported in earlier work shown on Table 7. Based on a linear random effects panel for determining credit ratings issued by S&P, Erdem and Varli (2014) are able to predict 42.8% of the ratings correctly while 51.3% and 64.1% are correctly predicted within one and two notches, respectively. For comparison, our random effects model applied to S&P's credit ratings predicts 57.22% of the ratings correctly and 90.23% and 96.41% are correctly predicted within a range of one and two notches respectively. For the machine learning methods, our work can be best compared with Ozturk et al. (2016), who also apply a random forests classification to Moody's credit ratings. Ozturk et al. (2016) predict 79.2% of the ratings correctly, 91.9% and 97.4% within a range of one and two notches in an in-sample test, while we obtain a prediction precision of 91.53%, 97.85% and 99.19%, respectively, in an out-of-sample test. A plausible explanation is that, with our wider sample, the classification rule is better trained.

# 5 Subjectivity in sovereign credit ratings: A closer look

## 5.1 The magnitude of the subjective component of sovereign credit ratings

To measure the subjective component of a sovereign credit score of country $i$ at time $t$, we compute how score $(i, t)$ would have changed if familiarity and lobbying had played no role. The change

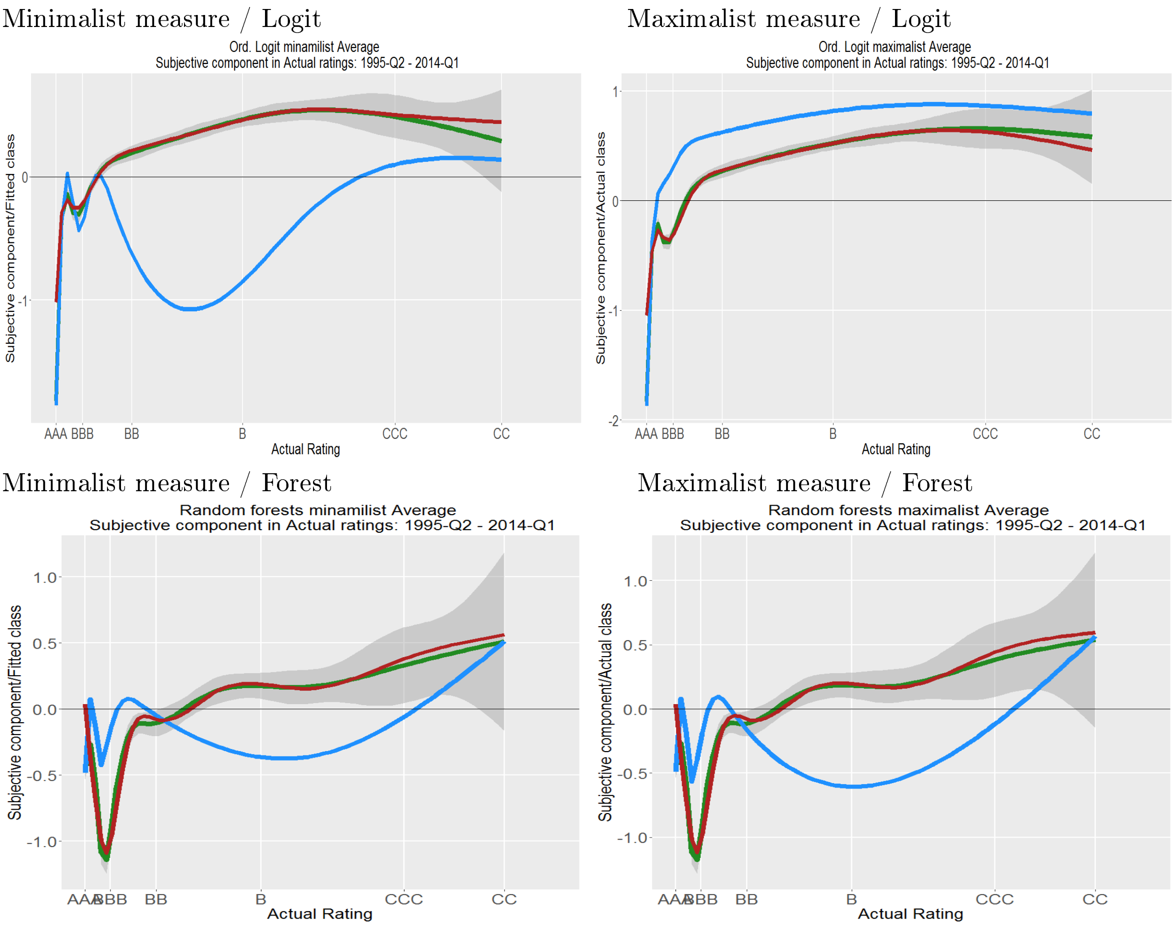
[10]It is reasonable to still count predictions that are one notch off as correct: rating agencies often keep a rating under review for upgrading/downgrading notch(es) even when objective and subjective components already herald a higher/lower rating.

– the part reflecting the effect of the familiarity variables and lobbying power, *i.e.* the estimated subjective part – is averaged across all observations in rating class $j$ at time $t$, to get a bias figure for each of the 21 notches in year $t$. For much of the discussion we also average over time, to obtain 21 unconditional bias figures. Complementarily, we produce averaged objective scores per class, the fitted value of the model's objective variables (logit) or the random forest classifications only using these variables. This objective score is then projected back onto the categorical AAA–C scale.

As acknowledged in the Introduction, the above approach is minimalist in that it focuses on the systematic part of the subjective element, the part with a positively demonstrated association with familiarity/lobbying variables. This puts the random errors and inconsistencies into the residuals, alongside objective effects (omitted variables, idiosyncratic events). Following Vernazza and Nielsen (2015), we can also adopt a maximalist definition, lumping the entire residual into the subjective part. After averaging, much of the residuals are bound to disappear, but not all, as averaging done is per rating class.

Figure 4 provides plots of the subjective component relative to the fitted rating (minimalist) and relative to the actual rating (maximalist) against the actual rating class for the two alternative specifications of subjectivity using the ordered logit model and the random forests model. The averages are computed from a scatter plot of the individual subjective components and are smoothed using cubic splines. Each plot shows these patterns for two subgroups of countries, *Advanced* (blue line) and *Emerging* (red), and also for the all-countries sample (green). The shaded areas show a two-sigma band around the smoothed average of all-countries sample. We evaluate the subjectivity in credit ratings on two grounds: *(i)* general pattern and *(ii)* influence at the border between speculative and investment grades. Complementing the graphs, Table 8 shows the average subjective part of the score per rating class, once expressed in the same numerical units as the total score and once as a the number of rating notches between the fitted values with and without the familiarity and lobbying variables.

In general, the level of the subjective component is positively related to the credit rating. This is observed for both ways of calculating the subjective component and for both estimation models. The positive slope in the subjectivity plot means that if ratings were based purely on objective variables, the total dispersion in credit ratings would be smaller: creditworthy countries would receive lower ratings and financially risky countries would be rated higher. To some extent, this is a tautology: if $y = x + z + \epsilon$, then $y$ is likely to be positively correlated with $z$ (or $x$, for that




emerging markets (red), developed (blue) and all (green)

Figure 4: Subjective component in sovereign credit ratings per rating class

*Note:* This figure provides plots of the average subjective component per rating class relative to the actual (maximalist) or fitted (minimalist) rating estimated based on a random ordered logit model (upper half) and a random forests model (lower half). The subjective component is calculated in two alternative ways, once without the residual ('minimalist') and once with ('maximalist'). The estimates are based on the average of the credit rating of the three agencies. We show the smoothed averages plus/minus two standard deviation for the entire sample (green).

matter).[11] So the more interesting questions are how much of the variation in the total score comes from the subjective part, and how this varies across the spectrum and country subgroups.

The relative importance of the subjective components across rating classes depends crucially on the estimation method. According to the machine learning (Panel B in Table 8), the subjective component is small for the highest rating classes but increases significantly as we move towards the investment grade border. Countries with a fair rating of BB+ can be favored with a 3-notch upward effect from the subjective component, lifting the rating over the investment grade border. At the same time, the subjective component is the smallest just below investment grade. For the

[11] unless the regression coefficient of $z$ on $x$ is below $-1$.

Table 8: Average subjective component per rating class and weight of subjectivity

| *Panel A: RE Ordered Logit* | | | | | | *Panel B: Random Forests* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Rating | score | Subjective part units | Subjective part notch(es) | objective score | fair rating | Rating | score | Subjective part units | Subjective part notch(es) | objective score | fair rating |
| AAA | 0.259 | –0.477 | +4 | 0.738 | A+ | AAA | 0.259 | –0.115 | +1 | 0.374 | AA+ |
| AA+ | 0.390 | –0.414 | +3 | 0.835 | A+ | AA+ | 0.390 | –0.282 | +2 | 0.672 | AA– |
| AA | 0.537 | –0.289 | +3 | 0.867 | A | AA | 0.537 | 0.030 | 0 | 0.514 | AA |
| AA– | 0.696 | –0.251 | +2 | 1.016 | A | AA– | 0.696 | –0.254 | +2 | 0.949 | A |
| A+ | 0.847 | –0.100 | +2 | 1.063 | A– | A+ | 0.847 | –0.618 | +3 | 1.468 | BBB+ |
| A | 1.055 | –0.089 | +1 | 1.160 | A– | A | 1.055 | –0.263 | +2 | 1.324 | BBB+ |
| A– | 1.221 | –0.296 | +2 | 1.569 | BBB | A– | 1.221 | –1.145 | +3 | 2.370 | BBB– |
| BBB+ | 1.480 | –0.806 | +2 | 2.486 | BBB– | BBB+ | 1.480 | –2.301 | +3 | 3.822 | BB+ |
| BBB | 2.088 | –0.212 | +1 | 2.342 | BBB– | BBB | 2.088 | –1.632 | +2 | 3.759 | BB+ |
| BBB– | 2.965 | –0.166 | +1 | 2.977 | BB+ | BBB– | 2.965 | –1.031 | +1 | 4.015 | BB+ |
| BB+ | 4.021 | 0.418 | 0 | 3.492 | BB+ | BB+ | 4.021 | –0.849 | +1 | 4.886 | BB |
| BB | 5.453 | 1.148 | –1 | 3.822 | BB+ | BB | 5.453 | –0.358 | +1 | 5.812 | BB– |
| BB– | 7.361 | 1.864 | –1 | 4.609 | BB | BB– | 7.361 | 0.224 | 0 | 7.124 | BB– |
| B+ | 9.898 | 3.594 | –1 | 5.704 | BB– | B+ | 9.898 | 0.924 | 0 | 8.934 | B+ |
| B | 13.075 | 5.614 | –2 | 6.228 | BB– | B | 13.075 | 3.018 | 0 | 9.934 | B |
| B– | 16.449 | 8.037 | –3 | 6.428 | BB– | B– | 16.449 | 2.334 | 0 | 13.945 | B– |
| CCC+ | 19.910 | 9.734 | –4 | 7.112 | BB– | CCC+ | 19.910 | 3.800 | –1 | 15.247 | B– |
| CCC | 23.532 | 8.170 | –4 | 9.384 | B+ | CCC | 23.532 | 10.124 | –3 | 10.190 | B |
| CCC– | 27.153 | 7.808 | –5 | 8.126 | B+ | CCC– | 27.153 | 8.362 | –2 | 17.545 | CCC+ |
| CC | 30.844 | 8.669 | –4 | 15.017 | B– | CC | 30.844 | 17.064 | –5 | 11.943 | B |

*Note:* This table shows the average subjective component per rating class its impact on the actual rating score. For example, '+4' in the fourth column implies that for the ordered logit model, the subjective component increases the fair rating with four notches. The sixth column shows the fitted objective rating (rounded to the closest rating scale). The results for the ordered logit model are presented in Panel A and Panel B shows the results for the random forests classification.

lowest rating classes, the subjective component has a negative effect, which can lead a 5 notches downward adjustment of the fair rating, on average.

When we look at the results of the ordered logit model (Panel A Table 8 and upper part of Figure 4), the impact of the subjective component is concentrated at the extremes: high-rated countries benefit from a positive impact of the subjective component (up to four notches, on average), while lower rated countries are penalized with a downgrade up to five notches. Presumably, the difference between the two estimation methodologies reflects the fact that the machine learning algorithm is better able to capture nonlinear effects between rating notches. In all four graphs, the blue line (developed countries) exhibits far more variability across the ratings spectrum. This partly reflects a lower number of observations: data below investment grade are scarce, especially in the B–CCC range where the available observations refer to just Greece and Ireland for a few years. Second, we note that for the logit model, with its large residual variance, the distinction between minimalist and maximalist makes a big difference in, again, the sample of developed countries (the blue line). In short, this raises questions about reliability of that particular pattern of developed countries.

All this is about the averaged ratings, across the three agencies. When we investigate the difference in subjectivity among rating agencies, according to the machine learning algorithm, we find that the subjective component of the credit rating is quite similar across agencies. For the highest rating class, the subjective components favor countries with one notch (Moody's), with three notches (S&P) and with zero notch (Fitch). Moving towards the investment grade border, BBB– rated countries are favored by two to three notches. For the lowest rating classes, the subjective component has a negative effect, which can downgrade by four to seven notches compared to the fair ratings. The only exception is Moody's, for which we find no impact of the subjective component in rating class CC.

### 5.2 Subjective judgment over time

In response to the changed regulatory framework after the financial crisis of 2008, the three credit rating agencies implemented changes to their procedures to determine the credit rating of a sovereign.[12] In general, rating agencies now claim to give more weight to quantitative inputs, leaving less room for subjective judgment in the final rating decision. Amstad and Packer (2015) investigate whether the revised methodologies have a statistically detectable impact in empirical models of credit ratings. Their results are mixed. They find that ratings have become increasingly sensitive to certain quantitative (objective) variables, but find no confirmation for a strong methodological change in the credit rating model. They do show that simple linear models are no longer suited to model credit ratings with precision. This might reflect the application of non-linear and non-parametric models by the rating agencies. In a press release, Fitch states that they apply, amongst others, a decision tree data mining methodology to assign credit ratings.[13] Time-variation in credit rating methodologies is also demonstrated by Reusens and Croux (2017) who show that rating agencies changed their sovereign rating assessment shortly after the start of the European debt crisis by allocating a higher weight to financial balance, GDP growth and external debt measures.

If credit rating agencies rely less on the subjective judgment of the rating committee, this implies that the subjective component of the credit rating should decrease over time, especially in the post-crisis period. As part of the robustness check in Section A.4, we allow for time variation

[12] Fitch (2014) Sovereign rating criteria; Moody's Investors Service (2013): Sovereign bond ratings: Rating Methodology; Standard & Poor's Rating Service (2014): Sovereign rating methodology

[13] Fitch Solutions (2013): Financial Implied Ratings Model Methodology – Special report

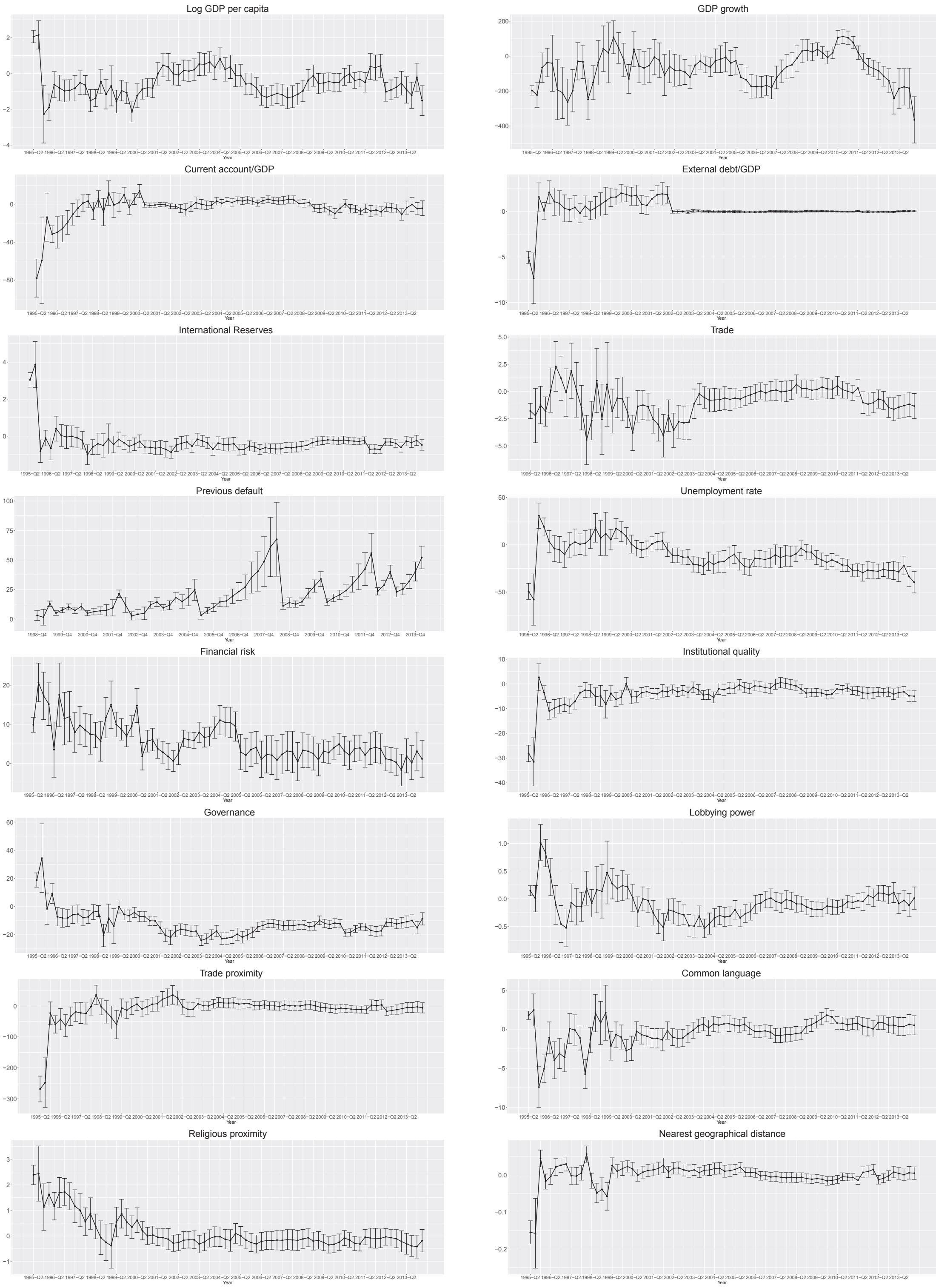



Figure 5: Time-varying coefficient estimates
This figure show the coefficient estimates of quarter-by-quarter OLS regressions that explain the spread-adjusted average rating with the variables of the objective and subjective components. Because the low number of observations in the early period, we were unable to run logit regressions.

Table 9: Subjectivity over time. $t$-tests on average subjective component per rating class, 2010–2014, compared to average for two preceding five-year periods

| Year | Rating | O_logit | R_forests | Rating | O_logit | R_forests |
|---|---|---|---|---|---|---|
| 2001-2006 | AAA | 1.147 | 4.522*** ↗ | BB | 0.041 | −3.151*** ↙ |
| 2007-2009 | AAA | 2.517** ↗ | 3.844*** ↗ | BB | 1.584 | 0.77 |
| 2001-2006 | AA | 2.046** ↗ | 4.317*** ↗ | B | −0.53 | −3.587*** ↗ |
| 2007-2009 | AA | 4.316*** ↗ | 4.209*** ↗ | B | 2.581 | −0.173 |
| 2001-2006 | A | −2.741*** ↙ | −1.11 | CCC | 4.069*** ↙ | −0.56 |
| 2007-2009 | A | 0.274 | 1.896* ↗ | CCC | 3.348*** ↙ | −0.729 |
| 2001-2006 | BBB | −2.938*** ↙ | −6.404*** ↙ | | | |
| 2007-2009 | BBB | −1.029 | −0.784 | | | |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

*Note:* This table shows respectively the t-statistic for the difference in average subjective component for the period '01 to '06 and the period '10 to '14 and the t-statistic for the difference in the average subjective component for '07-'09 and the period '10 to '14. Because the subjective component can be both positive or negative, the arrows indicate whether the subjective component increased or decreased (in absolute value) over time.

in the coefficients of the explanatory variables by estimating the model quarter-by-quarter.[14] We start with an eyeball inspection of the 16 time series of quarter-by-quarter estimates per coefficient. In Figure 5 we have plotted these coefficient estimates, alongside a two-sigma band. We do not observe a clear rise as of 2010 in the plots for the top nine graphs (which shows the coefficients for the economic/fiscal variables), nor do the loadings for the lower five (the familiarity covariates) seem to fall. What did happen, though, is that the high variability in the loadings from quarter to quarter, which was quite pronounced in the first third of the data period, seems to be gone. To at least some extent this reflects the low number of observations in the early years.[15]

This lack of evidence is not necessarily convincing. The impact of variable is best summarized by the product of the slope and the standard deviation of the associated regressor and the graphs consider only one component. To test for time variation in a more conclusive way, now look at the entire subjective component, whose variability reflects the slopes times their associated covariates, all duly summed. We estimate the ordered-logit and random-forest models for each of the three most recent five-year periods, and calculate for each period the average subjective component per rating class. The underlying data are still the mean ratings across agencies, not including the residuals.[16] Table 9 reports $t$-tests when the 2010-2014 average subjective component

[14] Actually, the coefficients are from an OLS regression rather than a logit: there are too many gaps in the yearly cross-sections to run a logit.

[15] Another part of the explanation may be that rating agencies have become more systematic and predictable, possibly because they rely more on a rather stable model than before. While that does sound like a positive evolution, if true, the change still happened long before 2010, so it is not a response to the post-crisis new regulation.

[16] The estimates per rating agency are similar; they are available upon request. The results for a 'maximal-

is compared to each of the two preceding periods. For the higher rating classes, the average subjective component is negative, so a positive value for the t-statistic reflects that the subjective component has increased over time (in absolute value), while for the lower rating classes this is the other way around because the subjective component is generally positive. For the ease of interpretation, we added arrows to indicate an absolute sense of increase or decrease in the subjective component.

The tests provide a mixed picture. The average subjective component in the highest rating classes (AAA and AA), where the bias was already upward, has *in*creased as of 2010. For BBB and below, where the bias is generally downward, subjectivity now has a less negative effect. So overall we find that after the global financial crisis, subjectivity in sovereign credit ratings has become more moderate for the countries with BBB or less, as the regulators had hoped, but it has increased for the high-rated ones (AAA and AA). The latter observation may be explained by the overall weakening of macroeconomic fundamentals (rising debt and unemployment figures) for the developed countries post-crisis.

## 5.3 Subjective judgment and default prediction

A next relevant question is whether the subjective component of a credit rating is informative in the sense that it helps in predicting sovereign defaults. Following Vernazza and Nielsen (2015), we estimate the following random effects ordered logit model for the probability of default of country $i$ in year $t$:

$$\begin{aligned} default^*_{it} &= a_i + b_1(objective_{i,t-\tau}) + b_2(subjective_{i,t-\tau}) + u_{it} \\ default_{it} &= 1[default^*_{it} > 0]; i = 1, ..., N; t = 1, ..., T; \tau \in \{1, 2, 3, 4, 5\}, \end{aligned} \tag{8}$$

where $default^*_{it}$ is the latent propensity to default for country $i$ in year $t$; $objective_{i,t-\tau}$ and $subjective_{i,t-\tau}$ are the objective and subjective components of the rating for country $i$ in year $t-\tau$, respectively; $default_{it}$ is the observed default event that takes the value one if country $i$ defaults in year $t$ and zero otherwise; $u_{it}$ is an error term; and 1[.] denotes the indicator function.

We investigate the predictive validity of both the objective and subjective components 1 year

ist'computation of the subjective component for the average rating series are reported in the Appendix (see Table 22).

Table 10: Defaults observed as of 1998 in the sample countries.

| Country | Year | Country | Year | Country | Year | Country | Year | Country | Year |
|---|---|---|---|---|---|---|---|---|---|
| Russia | 1998-Q3 | Russia | 1999-Q1 | Argentina | 2001-Q4 | Venezuela | 2005-Q1 | Jamaica | 2010-Q1 |
| Venezuela | 1998-Q3 | Dominican R | 1999-Q2 | Indonesia | 2002-Q2 | Dominican R | 2005-Q2 | Greece | 2012-Q1 |
| Indonesia | 1999-Q1 | Indonesia | 2000-Q2 | Moldova | 2002-Q2 | Nicaragua | 2008-Q2 | Greece | 2012-Q4 |
| Pakistan | 1999-Q1 | Peru | 2000-Q3 | Nicaragua | 2003-Q3 | Ecuador | 2008-Q4 | Jamaica | 2013-Q1 |
| | | | | | | | | Cyprus | 2013-Q2 |

Table 11: Predictive power of the first pass scores and spread adjusted scaled ratings

| | *Dependent: Default dummy* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Panel A: First pass scores* | | | | | *Panel B: Spread adj. scaled ratings* | | | | |
| | 1yr | 2yr | 3yr | 4yr | 5yr | 1yr | 2yr | 3yr | 4yr | 5yr |
| Actual Rating | $-0.39^{***}$ | $-0.21^{***}$ | $-0.19^{***}$ | $-0.18^{**}$ | $-0.12$ | $0.16^{***}$ | $0.11^{***}$ | $0.11^{***}$ | $0.10^{***}$ | 0.06 |
| | (0.08) | $(0.07)^{***}$ | (0.07) | (0.07) | (0.08) | (0.03) | (0.04) | (0.03) | (0.04) | (0.05) |
| Groups (Country) | 103 | 103 | 102 | 98 | 96 | 103 | 103 | 102 | 98 | 96 |
| McFadden $R^2$ | 0.15 | 0.07 | 0.07 | 0.06 | 0.06 | 0.12 | 0.05 | 0.06 | 0.06 | 0.06 |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Note: This table shows ability of the first pass scores and spread adjusted scaled ratings to predict sovereign defaults calculated using random effects ordered logit model. Panel A, on the left, shows the coefficient estimates of the first pass scores lagged from 1 year to 5 years. Panel B, on the right, shows the coefficient estimates for the spread adjusted ratings.

Table 12: Predictive power of the subjective and objective component of spread adj. scaled ratings

| | *Dependent: Default dummy* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Panel A: RE Ordered Logit* | | | | | *Panel B: Random Forests* | | | | |
| | 1yr | 2yr | 3yr | 4yr | 5yr | 1yr | 2yr | 3yr | 4yr | 5yr |
| Objective | $0.12^{*}$ | 0.11 | $0.14^{**}$ | 0.05 | $-0.02$ | $0.13^{***}$ | $0.11^{**}$ | $0.14^{***}$ | $0.12^{**}$ | 0.07 |
| | (0.07) | (0.07) | (0.06) | (0.09) | (0.12) | (0.04) | (0.04) | (0.04) | (0.05) | 0.06 |
| Subjective | $0.14^{**}$ | $0.14^{**}$ | 0.07 | $0.16^{**}$ | $0.2^{**}$ | $0.16^{***}$ | $0.10^{*}$ | 0.03 | 0.06 | 0.07 |
| | (0.06) | (0.07) | (0.07) | (0.07) | (0.10) | (0.04) | (0.06) | (0.07) | (0.07) | 0.09 |
| Groups (Country) | 103 | 103 | 102 | 98 | 96 | 103 | 103 | 102 | 98 | 96 |
| McFadden $R^2$ | 0.07 | 0.06 | 0.06 | 0.06 | 0.08 | 0.10 | 0.06 | 0.07 | 0.05 | 0.06 |

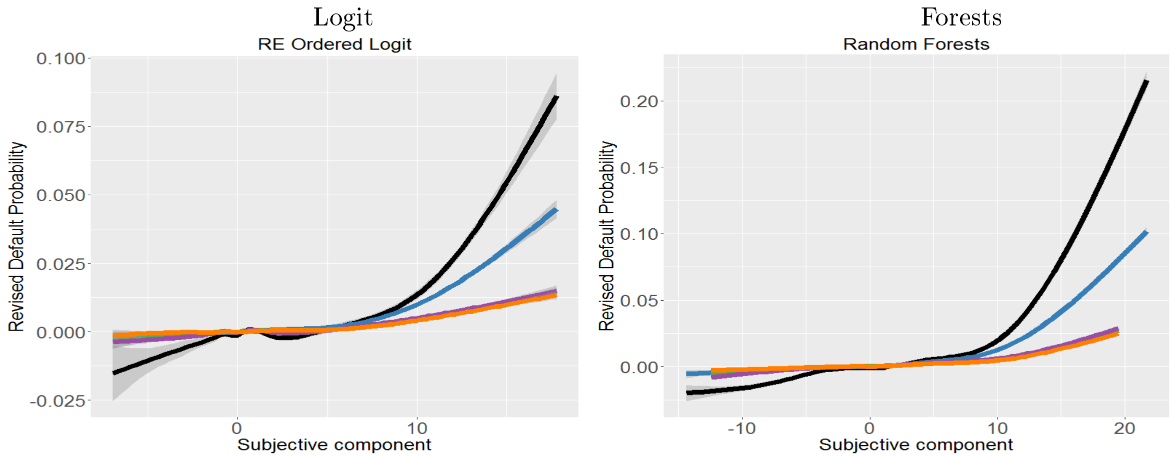
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Note: This table shows ability of the objective and subjective components to predict sovereign defaults. Panel A, on the left, shows the coefficient estimates of the objective and subjective components lagged from 1 year to 5 years and calculated based on the ordered logit model. Panel B, on the right, shows the coefficient estimates for the objective and subjective components of the random forests model.

to 5 years preceding the sovereign debt default.[17] Table 10 shows the occurrences of sovereign debt defaults in our sample. Table 11 and 12 present the results for the average rating of three rating agencies.[18] We first show in Table 11 how the total rating performs in predicting defaults. For the actual rating, both first pass scores and spread-adjusted ratings have predictive power for

[17] We work with quarterly data, so when we consider a default event in the third quarter of 2003, we check whether it was correctly predicted already in the third quarter of 2002, and likewise in that of 2001, 2000, 1999 and 1998.

[18] The estimates per rating agency are similar and are available on request.




Quarter Lag: 1Q (black), 4Q (blue), 8Q (green), 12Q (purple) and 16Q (orange)
Figure 6: Subjective component and default probability

Note: This crossplot shows on the X-axis the subjective component estimated by the RE ordered-logit model (left) and the random-forests model (right) for spread adjusted scaled ratings. The revised default probability (on the Y-axis) is the difference between default probability when the actual rating is considered in the model (see Table 11) and the default probability with objective components only. We show the smoothed line plus/minus two standard deviation for the entire sample lagged for 1 quarter to 16 quarters.

sovereign defaults up to four years before the event.[19] When we divide the rating in its objective and subjective component (Table 12) we find that for both the logit model and the random forests model the one-year lagged objective and subjective components are all significant, with the expected positive sign, which implies that they are both useful predictors of imminent sovereign default. The subjective component also has predictive power for defaults within two years. Over longer time frames, however, the models provide qualitatively similar conclusions regarding the predictive ability of the two generic components. When modeled via the ordered logit model, the subjective component maintains its ability to predict defaults all the way up to five years before the event, albeit with an unsteady performance in the third year, while the objective component's usefulness seems to peter out after third year. When we analyse the probability of default with the machine learning algorithm, we find a poor performance for the subjective component from the third year onwards, while the objective component has a predictive value for defaults within four years.

Further, we investigate the contribution of the subjective component to the revised probability of default. The revised default probability is equal to the difference in the default probability of the actual rating and the fair rating. The good news is that default prediction improves with the inclusion of subjective components in the credit ratings as shown in Figure 6. According to the

[19]Recall, the spread-adjusted ratings have high values for low rating classes and vice versa. Therefore, the coefficient estimates are positive for the spread-adjusted scaled ratings in Table 11 and 12.

random forests model, the subjective component results in a sharper default prediction for one to four quarters before the default event. At any rate, our conclusions are quite positive. They are in line with Agarwal and Hauswald (2010), Rajan et al. (2015) Qi et al. (2014)'s evidence from corporate borrowing and in sharp contrast to Vernazza and Nielsen (2015) who find that the subjective component (measured as the regression residuals) only helps to predict imminent defaults and even gives false warnings when predicting a default within three years.[20]

## 5.4 The economic costs of subjective judgment in credit ratings

Lastly, we asses the economic costs of the subjective component in credit ratings. Some researchers argue that the subjective judgment of the credit rating committee leads to a qualitative bias in credit ratings which is potentially detrimental to the country (Vernazza and Nielsen (2015), Zheng (2012), Luitel et al. (2016)). The evidence from corporate bond markets suggests otherwise: Agarwal and Hauswald (2010) and Rajan et al. (2015) not only show that omitting subjective information at the corporate level leads to worse credit scoring models, but also to higher borrowing costs.

Our way of measuring to what extent subjectivity in sovereign credit ratings affects borrowing costs is to compare sovereign CDS spreads for a country given its predicted rating with the average CDS spread that is associated with the country's counterfactual 'objective' rating.[21] Then we multiply the spread differential by the amount of public debt outstanding and scale this by the country's GDP. That is, the economic costs of the subjective component in the credit rating of country $i$ ($cost_{subj,i}$) is equal to:

$$cost_{subj,it} = \frac{(CDS_{it}^{pred} - CDS_{it}^{obj}) \times PD_{it}}{GDP_{it}}. \tag{9}$$

where $CDS_{it}^{pred}$ and $CDS_{it}^{obj}$ are the average CDS spreads that are associated with the predicted and objective rating class of country $i$ in quarter $t$, $PD_{it}$ and $GDP_{it}$ are the public debt and GDP of country $i$ in quarter $t$, respectively. In comparison, the economic cost associated with the total rating is equal to:

$$cost_{rating,it} = \frac{(CDS_{it}^{pred}) \times PD_{it}}{GDP_{it}}. \tag{10}$$

[20]The results for 'maximalist'prediction performance is reported in the Appendix (see Table 23). We find that for the machine learning model, including residuals in the subjective component enhances the prediction performance for the third year compared to the 'minimalist'method in Table 12.

[21]5-year CDS spreads are obtained from Thomson Reuters.

From Table 8 we know that the economic cost will vary across rating classes: in general, countries that are rated as investment grade will benefit from the subjective component, while subjectivity may create an economic cost for speculative grade countries. For instance in 2010-Q3, the average credit rating assigned by the three rating agencies to Greece was BBB. Based on the ordered logit model, the predicted and objective rating of Greece for that period are BBB and BBB+ respectively. In 2010-Q3, the average 5-year CDS spread associated with BBB and BBB+ rated sovereign was 296.85 and 123.42 basis points, which implies a premium of 173.43 basis points for Greece. [22] Whether this premium is economically important depends on the size of Greece's government debt. With a debt to GDP ratio then at 148.30 percent, the economic costs of the subjective component correspond to a premium of 2.57 percent of Greece's GDP which amounts to USD 7.89 billion.

The individual country computations for the period 2010-2014 are shown in Figure 7. The scatter plot puts the subjective component of each country's credit rating on the horizontal axis and the related economic costs on the vertical axis, with red dots representing emerging markets and blue dots developed markets. The noisiness is quite important, and especially so for the logit-based results (on the left). This is because, in terms of our three-way decomposition, we just removed the systemic subjective part, so the gap between the rate and the normalised rate contains the residual, which has a high variance in the logit case. The concentration of extreme cost estimates associated with negative subjectivity only applies to the logit model, so its reliability is unclear. But even in the figure on the left there is otherwise no pronounced relation between subjectivity and cost. A clearer pattern is that red dots (emerging markets) tend to be in the positive-cost zone and vice versa, but even to that rule there are many exceptions.[23]

Table 13 shows the average economic costs per rating class, still for the sub-period 2010-2014.[24] The results for the CCC class defies credibility, which can be explained by the high average economic costs of Greece and Cyprus during that sub-period. According to the random-

[22]For the 'maximalist'economic costs of Greece in 2010-Q3, we replace the $CDS_{it}^{pred}$ by $CDS_{it}$ , i.e. $cost_{subj,it} = \frac{(CDS_{it} - CDS_{it}^{obj}) \times PD_{it}}{GDP_{it}}$ where $CDS_{it}$ is the effective 5-year CDS-spread of country $i$ in quarter $t$. In 2010-Q3, the average 5-year CDS spread for the Greece was 830.18 basis points, which implies a premium of 10.48 percent of the GDP. The total borrowing cost for the Greece was 12.31 percent of the GDP which is calculated as $cost_{rating,it} = \frac{(CDS_{it}) \times PD_{it}}{GDP_{it}}$.

[23]The Pearson correlations between the average economic costs (as percent of GDP) and the subjective component are 32% and 53% for the ordered logit and random forests classification, respectively and they are significant at the 99% confidence level.

[24]Results for earlier periods are similar and are available upon request.

Table 13: Economic costs of the subjective component in sovereign credit ratings, 2010–2014

| | *Panel A: RE Ordered Logit* | | | | | *Panel B: Random Forests* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Minimalist** | | | | | | | | | | |
| | Rating costs | | Subjectivity costs | | | Rating costs | | Subjectivity costs | | |
| Ratings | Avg. costs (bps) | Avg. costs (% of GDP) | Avg. economic costs (bps) | Avg. economic costs (% of GDP) | Avg. economic costs (USD mill.) | Avg. costs (bps) | Avg. costs (% of GDP) | Avg. economic costs (bps) | Avg. economic costs (% of GDP) | Avg. economic costs (USD mill.) |
| AAA | 55.55 | 0.34 | –92.02 | –0.59 | –22,971.49 | 55.23 | 0.34 | –19.91 | –0.02 | –100.00 |
| AA | 145.72 | 1.20 | –26.67 | –0.17 | –4,125.07 | 136.58 | 1.17 | –15.30 | –0.06 | –2,103.29 |
| A | 174.73 | 0.83 | –7.21 | –0.02 | –68.56 | 138.27 | 0.67 | –0.56 | –0.00 | –6.42 |
| BBB | 229.94 | 1.04 | 22.00 | 0.06 | 326.15 | 226.32 | 1.11 | –37.82 | –0.06 | –405.27 |
| BB | 277.01 | 1.33 | 69.93 | 0.28 | 664.34 | 295.03 | 1.65 | –26.27 | –0.04 | –120.23 |
| B | 509.24 | 4.83 | 314.15 | 2.84 | 3,033.95 | 619.38 | 5.32 | 174.39 | 0.88 | 1,293.23 |
| CCC | 348.83 | 4.25 | 3,392.20 | 36.24 | 88,879.28 | 10,679.78 | 170.67 | 1,346.77 | 9.80 | 22,074.50 |
| **Maximalist** | | | | | | | | | | |
| | Rating costs | | Subjectivity costs | | | Rating costs | | Subjectivity costs | | |
| Ratings | Avg. costs (bps) | Avg. costs (% of GDP) | Avg. economic costs (bps) | Avg. economic costs (% of GDP) | Avg. economic costs (USD mill.) | Avg. costs (bps) | Avg. costs (% of GDP) | Avg. economic costs (bps) | Avg. economic costs (% of GDP) | Avg. economic costs (USD mill.) |
| AAA | 55.58 | 0.36 | –92.02 | –0.59 | –22,971.49 | 55.58 | 0.36 | –19.91 | –0.02 | –100.00 |
| AA | 136.90 | 1.18 | –38.73 | –0.27 | –7,022.69 | 136.90 | 1.18 | –14.75 | –0.06 | –2,051.46 |
| A | 148.52 | 0.86 | –23.05 | –0.07 | –322.84 | 148.52 | 0.86 | –0.56 | –0.00 | –6.42 |
| BBB | 235.37 | 1.44 | 26.05 | 0.08 | 406.63 | 235.37 | 1.44 | –36.21 | –0.06 | –403.20 |
| BB | 317.08 | 2.24 | 80.08 | 0.37 | 823.32 | 317.08 | 2.24 | 2.79 | 0.01 | 13.82 |
| B | 595.84 | 4.78 | 284.18 | 2.40 | 2,600.41 | 595.84 | 4.78 | 173.18 | 0.87 | 1,280.39 |
| CCC | 11,001.32 | 173.34 | 10,749.70 | 148.58 | 375,239.27 | 11,001.32 | 173.34 | 1,346.77 | 9.80 | 22,074.50 |

Note: Panel A shows the average economic costs of the subjective component in basis points, as percentage of GDP and in million US dollar based on the random effects ordered logit model. Panel B shows the same for the random random forests classification. The costs are calculated for the period 2010-Q1 to 2014-Q1.

forests model, economic costs of subjectivity vary from –6bp to 88bp of GDP. When we compare the economic costs of the total rating with the costs associated with the subjective component, we conclude that the impact of subjectivity on borrowing costs is low. The average cost of the total rating ranges from 0.3 percent of GDP for AAA rated countries to 5 of GDP for the lowest rating classes. This implies that a rating's objective component is the main determinant in a country's borrowing costs.

Figure 8 puts the average economic costs of the individual country on the horizontal axis and the related change in the expected dollar default cost/GDP on the vertical axis. The expected default cost is read as revised default probability times public debt/GDP. There is no clear pattern, since all countries are clustered close to zero. According to the machine learning algorithm, there is an exception blue dot on top-right (Greece) and red dot on mid-right (Cyprus) were loaded with economic costs caused by subjectivity.

# 6 Conclusion

In the process of assigning a credit rating to a sovereign debt issuer, a credit rating analyst takes into account several criteria that focus on the country's past economic and political situation as well as factors that suggest future economic, financial and political risks. Based on an initial analysis, the rating analyst provides his/her recommendation to the credit rating committee who decides

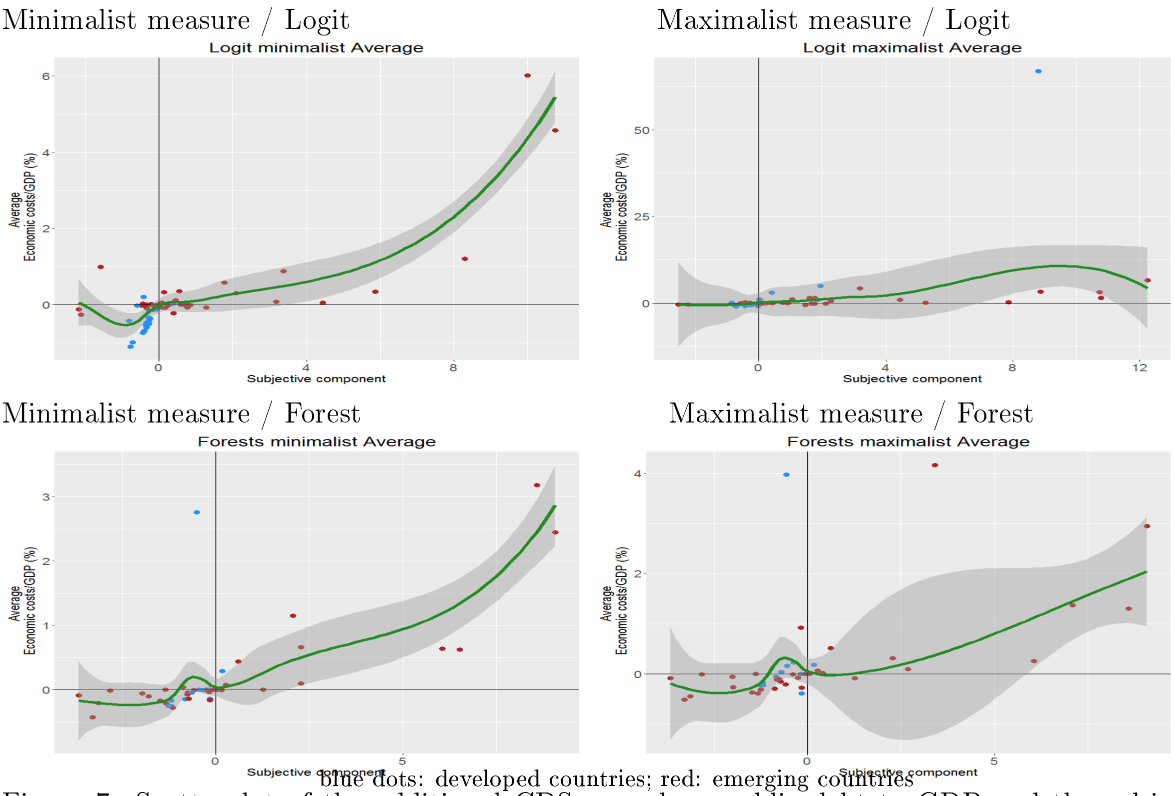
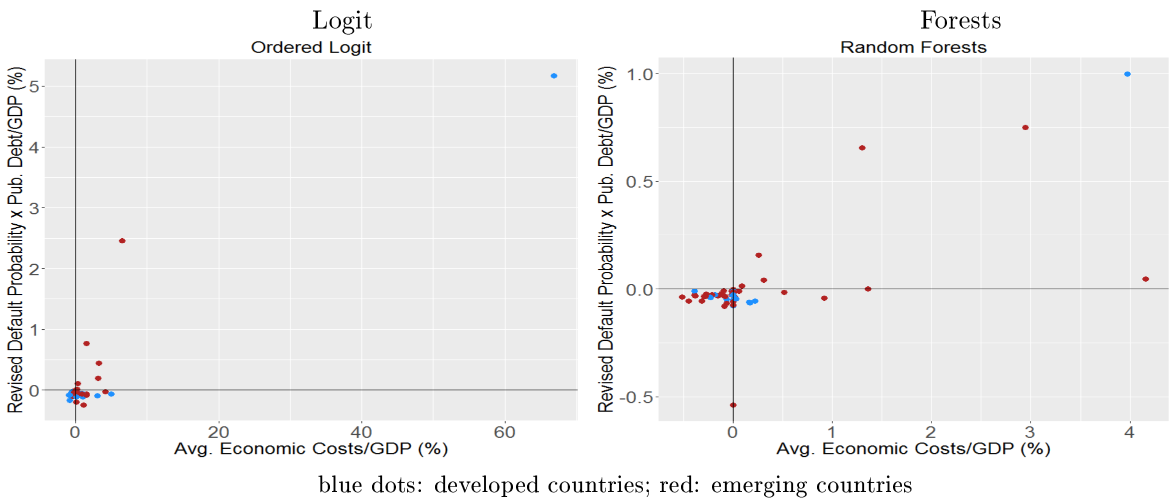



Figure 7: Scatterplot of the additional CDS spread on public debt to GDP and the subjective component of the credit rating

Note: This crossplot shows on the X-axis the average subjective component for each country estimated by the RE ordered-logit model (left) and the random-forests model (right). The economic cost (as % of GDP) is read off on the Y-axis — so a negative reading means a benefit. This economic cost *(i) minimalist*: is equal to the difference in the CDS-spread that is associated with the predicted rating of the country and the CDS-spread associated with the objective rating of the country taking public debt into account and *(ii) maximalist*: is the difference in the actual CDS-spread of the country and the CDS-spread that is associated with the objective rating of the country taking public debt into account. We then multiply this extra cost by the country's public debt level and scale it by its GDP. We show the smoothed averages plus/minus two standard deviation for the entire sample (green).

upon the final rating. The rating committee may adjust the analyst's recommendation upwards or downwards, and in this process may rely on certain qualitative factors and subjective judgment. The opaqueness about the relative importance of quantitative, qualitative and judgmental criteria in the credit rating process makes it impossible to fully replicate the published credit ratings on the basis of publicly available information. This has led to fierce criticism of the credit rating process from investors and bond issuers. In the aftermath of the global financial crisis, regulatory changes have been made both in the United States and Europe with the objective of increasing the transparency of the credit rating process. Since 2010, the leading credit rating agencies, S&P, Moody's and Fitch, claim to have changed their methods for assigning sovereign credit ratings: credit ratings should rely more on quantitative inputs and less on qualitative and judgmental factors.

This paper contributes to the discussion on the subjectivity in credit ratings and the transparency of the rating procedure, by a more careful disentanglement of the objective and the subjective components of sovereign credit ratings. On the basis of a machine learning algorithm, we find that although subjective factors like the proximity of a country to the U.S. and its lobbying power are relatively important in the credit rating process, the average economic impact of these variables is small. The subjective component in sovereign credit ratings leads to a downward adjustment of the objective rating up to five notches for the lowest-rated countries, to an upward adjustment of one to four notches for the highest-rated countries. Interestingly, the subjective component results in the highest upward adjustment for countries with a fair rating just below investment grade; that is, a country just below BBB is often nudged into investment-grade status, it seems. However, we find no evidence for the hypothesis that subjectivity in credit ratings would lead to substantially increased borrowing costs for these sovereigns; the subjective component in the credit ratings shows small correlation with credit spreads. The subjective component in credit ratings is uniform across rating agencies, and varies over time, although only mildly so, and without following clear trends. We find no manifest effect of the introduction of the Dodd–Frank act and the European Securities and Markets Authority, which had the intention of reducing the subjective component in credit ratings. Although the subjective part of the credit rating has decreased significantly for the lower ratings, typically emerging markets, the subjective component




blue dots: developed countries; red: emerging countries

Figure 8: Scatterplot of the additional CDS spread on public debt to GDP and the default probability of subjective component of the credit rating

Note: This crossplot shows on the X-axis the ‘maximalist’ average economic cost (as % of GDP) for each country. The economic cost (as % of GDP) is read off — so a negative reading means a benefit. The Y-axis puts the revised default probability times public debt/GDP. The revised default probability is the difference between default probability when actual rating is considered in the model (see Table 11) and default probability with objective components only.

has increased for the high-rated countries.

On the positive side, we find that the subjective component in sovereign credit ratings is actually informative. In the ordered logit model, the subjective component is a much better predictor of sovereign default than the objective component of the rating for one up to five years before the default. In the random forests classification, the subjective component helps to predict defaults within one and two years, while the objective component has predictive value over an horizon of one to four years.

Our findings have relevant implications for several financial market participants. First of all, we present a relatively simple framework that allows investors and researchers to better replicate and interpret sovereign credit ratings. We show that the random forest machine learning algorithm outperforms logit models in terms of rating prediction accuracy and the addition of subjectivity variables improves the predictive power of the model. They also do better than in prior studies that apply machine-learning models for the replication of credit ratings. Second, our work has relevance to regulatory bodies and financial market authorities because it helps understand the subjective part of the credit rating process. This may become a starting point for increased transparency.

## References

Afonso, A. (2003). Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies. *Journal of Economics and Finance*, 27(1):56–74.

Afonso, A., Gomes, P., and Rother, P. (2009). Ordered response models for sovereign debt ratings. *Applied Economics Letters*, 16(8):769–773.

Afonso, A., Gomes, P., and Rother, P. (2011). Short-and long-run determinants of sovereign debt credit ratings. *International Journal of Finance & Economics*, 16(1):1–15.

Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. *Review of Financial Studies*, 23(7):2757–2788.

Agresti, A. and Natarajan, R. (2001). Modeling clustered ordered categorical data: A survey. *International Statistical Review*, 69(3):345–371.

Alexe, S., Hammer, P., Kogan, A., and Lejeune, M. (2003). A non-recursive regression model for country risk rating. *RUTCOR-Rutgers University Research Report RRR*, 9:1–40.

Amstad, M. and Packer, F. (2015). Sovereign ratings of advanced and emerging economies after the crisis. *BIS Quarterly Review December*.

Amtenbrink, F. and Haan, J. D. (2009). Regulating credit ratings in the european union: A critical first assessment of regulation 1060/2009 on credit rating agencies. *Common Market Law Review*, 46(6):1915–1949.

Archer, C. C., 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(02):341–365.

Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8):1165–1195.

Bennell, J. A., Crabbe, D., Thomas, S., and ap Gwilym, O. (2006). Modelling sovereign credit ratings: Neural networks versus ordered probit. *Expert Systems with Applications*, 30(3):415–425.

Block, S. A. and Vaaler, P. M. (2004). The price of democracy: sovereign risk ratings, bond spreads and political business cycles in developing countries. *Journal of International Money and Finance*, 23(6):917–946.

Borensztein, E., Cowan, K., and Valenzuela, P. (2013). Sovereign ceilings "lite"? the impact of sovereign ratings on corporate ratings. *Journal of Banking & Finance*, 37(11):4014–4024.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Brewer, T. L. and Rivoli, P. (1990). Politics and perceived country creditworthiness in international banking. *Journal of Money, Credit and Banking*, 22(3):357–369.

Bruner, C. M. and Abdelal, R. (2005). To judge leviathan: Sovereign credit ratings, national law, and the world economy. *Journal of Public Policy*, 25:191–217.

Butler, A. W. and Fauver, L. (2006). Institutional environment and sovereign credit ratings. *Financial Management*, 35(3):53–79.

Cantor, R. and Packer, F. (1996). Determinants and impact of sovereign credit ratings. *Economic policy review–Federal Reserve Bank of New York*, 2(2):37–53.

Chen, C., Liaw, A., and Breiman, L. (2004). Using random forest to learn imbalanced data. *University of California, Berkeley*, 110:1–12.

Connolly, M. (2007). Measuring the effect of corruption on sovereign bond ratings. *Journal of Economic Policy Reform*, 10(4):309–323.

Cosset, J.-C. and Roy, J. (1991). The determinants of country risk ratings. *Journal of International Business Studies*, 22(1):135–142.

Dagum, E. B. and Cholette, P. A. (2006). *Benchmarking, temporal distribution, and reconciliation methods for time series*, volume 186e. Springer Science & Business Media.

Dalsgaard, J. C. G. and Hirth, S. (2014). Is there a home bias in soverign ratings? Working paper, Aarhus University.

Denton, F. T. (1971). Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66(333):99–102.

Erdem, O. and Varli, Y. (2014). Understanding the sovereign credit ratings of emerging markets. *Emerging Markets Review*, 20:42–57.

Feder, G. and Uy, L. V. (1985). The determinants of international creditworthiness and their policy implications. *Journal of Policy Modeling*, 7(1):133–156.

Ferri, G., Liu, L.-G., and Stiglitz, J. E. (1999). The procyclical role of rating agencies: Evidence from the east asian crisis. *Economic Notes*, 28(3):335–355.

Fuchs, A. and Gehring, K. (2016). The home bias in sovereign ratings. *Journal of the European Economic Association*, forthcoming.

Gaillard, N. (2009). The determinants of moody's sub-sovereign ratings. *International Research Journal of Finance and Economics*, 31(1):194–209.

Gande, A. and Parsley, D. C. (2005). News spillovers in the sovereign debt market. *Journal of Financial Economics*, 75(3):691–734.

Gültekin-Karakaş, D., Hisarcıklılar, M., and Öztürk, H. (2011). Sovereign risk ratings: Biased toward developed countries? *Emerging Markets Finance and Trade*, 47(2):69–87.

Güttler, A. and Wahrenburg, M. (2007). The adjustment of credit ratings in advance of defaults. *Journal of Banking & Finance*, 31(3):751–767.

Haque, N. U., Kumar, M. S., Mark, N., and Mathieson, D. J. (1996). The economic content of indicators of developing country creditworthiness. *Staff Papers (International Monetary Fund)*, 43(4):688–724.

Hedeker, D. (2008). *Multilevel Models for Ordinal and Nominal Variables*, pages 237–274. Springer–Verlag, New York.

Hu, Y.-T., Kiesel, R., and Perraudin, W. (2002). The estimation of transition matrices for sovereign credit ratings. *Journal of Banking & Finance*, 26(7):1383–1406.

Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., and Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems*, 37(4):543–558.

Lee, S. H. (1993). Relative importance of political instability and economic variables on perceived country creditworthiness. *Journal of International Business Studies*, 24(4):801–812.

Luitel, P., Vanpée, R., and Moor, L. D. (2016). Pernicious effects: How the credit rating agencies disadvantage emerging markets. *Research in International Business and Finance*, 38:286–298.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(2):109–142.

McFadden, D., Eckaus, R., Feder, G., Hajivassiliou, V., and O'Connell, S. (1985). Is there life after debt? an econometric analysis of the creditworthiness of developing countries. *International Debt and the Developing Countries*, pages 179 – 209.

McNamara, G. and Vaaler, P. M. (2000). The influence of competitive positioning and rivalry on emerging market risk assessment. *Journal of International Business Studies*, 31(2):337–347.

Mellios, C. and Paget-Blanc, E. (2006). Which factors determine sovereign credit ratings? *The European Journal of Finance*, 12(4):361–377.

Monfort, B. and Mulder, C. (2000). Using credit ratings for capital requirementson lending to emerging market economies: Possible impact of a new basel accord. IMF Working Papers 00/69, International Monetary Fund.

Mulder, C. and Perrelli, R. (2001). Foreign currency credit ratings for emerging market economies. IMF Working Papers 01/191, International Monetary Fund.

Ozturk, H. (2014). The origin of bias in sovereign credit ratings: Reconciling agency views with institutional quality. *The Journal of Developing Areas*, 48(4):161–188.

Ozturk, H., Namli, E., and Erdal, H. I. (2016). Modelling sovereign credit ratings: The accuracy of models in a heterogeneous sample. *Economic Modelling*, 54:469–478.

Polo, M. D. (2011). El comportamiento de los ratings crediticios a lo largo del ciclo. *Estabilidad financiera*, 20:71–91.

Qi, M., Zhang, X., and Zhao, X. (2014). Unobserved systematic risk factor and default prediction. *Journal of Banking & Finance*, 49:216 – 227.

Rajan, U., Seru, A., and Vig, V. (2015). The failure of models that predict failure: Distance, incentives, and defaults. *Journal of Financial Economics*, 115(2):237–260.

Ratha, D., De, P. K., and Mohapatra, S. (2011). Shadow sovereign ratings for unrated developing countries. *World Development*, 39(3):295–307.

Remolona, E. M., Scatigna, M., and Wu, E. (2008). A ratings-based approach to measuring sovereign risk. *International Journal of Finance & Economics*, 13(1):26–39.

Reusens, P. and Croux, C. (2017). Sovereign credit rating determinants: A comparison before and after the european debt crisis. *Journal of Banking & Finance*, 77:108 – 121.

Saini, K. G. and Bates, P. S. (1984). A survey of the quantitative approaches to country risk analysis. *Journal of Banking & Finance*, 8(2):341–356.

Sax, C. and Steiner, P. (2013). Temporal disaggregation of time series. *The R Journal*, 5:80–87.

Trevor Hastie, Robert Tibshirani, J. F. (2009). *The Elements of Statistical Learning*. Springer-Verlag New York.

Ul Haque, N., Mark, N. C., and Mathieson, D. J. (1998). The relative importance of political and economic variables in creditworthiness ratings. IMF Working Papers 98/46, International Monetary Fund.

Van Gestel, T., Baesens, B., Van Dijcke, P., Garcia, J., A.K. Suykens, J., and Vanthienen, J.

(2006). A process model to develop an internal rating system: Sovereign credit ratings. *Decision Support Systems*, 42(2):1131–1151.

Vernazza, D. R. and Nielsen, E. F. (2015). The damaging bias of sovereign ratings. *Economic Notes*, 44(2):361–408.

Zheng, L. (2012). Are sovereign credit ratings objective? a tale of two agencies. *Journal of Applied Finance & Banking*, 2(5):43–61.

# A Appendix

## A.1 Ordinal transformation of sovereign credit ratings

### Step 1

First, we run the following regression:

$$Spread_{i,t} = \alpha + \beta_1 First\ pass_{i,t} + \beta_2 First\ pass^2_{i,t} + \beta_3 Investment_{i,t} + \epsilon_{i,t} \tag{11}$$

where, $Spread_{i,t}$ is the sovereign 5-year CDS spread for country $i$ in quarter $t$, $First\ pass_{it}$ is the numeric scale of the sovereign credit rating ranging from 21 (AAA) to 1 (C) and $Investment$ is a dummy variable which is equal to unity for rating classes of BBB- and higher, and zero otherwise. We have CDS spread data from 2007 onwards, so the data for 2007 is used to adjust earlier ratings. Table 14 shows the estimation results of the regression specification.

$$First\ pass_{it} = \begin{cases} cr_{it} = C, & \text{then} \quad 1 \\ cr_{it} = CC, & \text{then} \quad 2 \\ \vdots & \\ cr_{it} = AAA, & \text{then} \quad 21 \end{cases}$$

Table 14: Non linear transformation of first-pass score

| | *Dependent variable:* | | |
|---|---|---|---|
| | Yield | | |
| | Moody's | S&P | Fitch |
| First pass | −6.273*** | −6.745*** | −3.917*** |
| | (0.385) | (0.414) | (0.465) |
| First pass square | 0.182*** | 0.195*** | 0.111*** |
| | (0.012) | (0.013) | (0.015) |
| Investment dummy | 7.208*** | 8.271*** | 2.809*** |
| | (0.950) | (0.963) | (0.995) |
| Constant | 46.228*** | 49.634*** | 32.010*** |
| | (2.425) | (2.655) | (3.022) |
| Observations | 1,546 | 1,582 | 1,492 |
| $R^2$ | 0.197 | 0.187 | 0.090 |
| *Note:* | | | *p<0.1; **p<0.05; ***p<0.01 |

### Step 2

We calculate the fitted value for each first-pass score from 1 to 21 and repeat the following process for each rating agency - *Moody's, Standard and Poor's and Fitch*:

for(i in 1:21){ $fit\_yield[i] \leftarrow$ fitted value when the first pass score is $i$ }

**Step 3**

We calculate the average fitted yield.

$$Average\_fit\_yield := \frac{Moodys\_fit\_yield + Std.\ Poors\_fit\_yield + Fitch\_fit\_yield}{3} \tag{12}$$

**Step 4**

We smooth the fitted yield for the individual rating agencies and the average fit using the smoothing package loess smooth with first degree. The results after smoothing are shown in the last four columns of the right-hand-side panel in Table 15. In R, the user can replicate the numbers shown in Table 15 by using the following command:

$first\_pass \leftarrow c(1:21)$
$smooth\_yield \leftarrow loess(fit\_yield \sim first\_pass, degree = 1)$

Table 15: Ordinal transformation of sovereign credit ratings

| | *Ratings* | | | *Scale* | | | | |
|---|---|---|---|---|---|---|---|---|
| Grade | Moody's | Std. Poors | Fitch | First-pass | Moody's | Std. Poors | Fitch | Average |
| Investment | Aaa | AAA | AAA | 21 | 0.330 | 0.379 | 0.070 | 0.259 |
| | Aa1 | AA+ | AA+ | 20 | 0.408 | 0.485 | 0.278 | 0.390 |
| | Aa2 | AA | AA | 19 | 0.504 | 0.606 | 0.502 | 0.537 |
| | Aa3 | AA- | AA- | 18 | 0.610 | 0.736 | 0.742 | 0.696 |
| | A1 | A+ | A+ | 17 | 0.707 | 0.856 | 0.978 | 0.847 |
| | A2 | A | A | 16 | 0.865 | 1.035 | 1.266 | 1.055 |
| | A3 | A- | A- | 15 | 0.978 | 1.169 | 1.517 | 1.221 |
| | Baa1 | BBB+ | BBB+ | 14 | 1.186 | 1.392 | 1.862 | 1.480 |
| | Baa2 | BBB | BBB | 13 | 1.772 | 1.997 | 2.497 | 2.088 |
| | Baa3 | BBB- | BBB- | 12 | 2.655 | 2.915 | 3.327 | 2.965 |
| Speculative | Ba1 | BB+ | BB+ | 11 | 3.730 | 4.024 | 4.310 | 4.021 |
| | Ba2 | BB | BB | 10 | 5.229 | 5.591 | 5.538 | 5.453 |
| | Ba3 | BB- | BB- | 9 | 7.273 | 7.754 | 7.056 | 7.361 |
| | B1 | B+ | B+ | 8 | 10.031 | 10.695 | 8.969 | 9.898 |
| | B2 | B | B | 7 | 13.519 | 14.437 | 11.269 | 13.075 |
| | B3 | B- | B- | 6 | 17.233 | 18.428 | 13.684 | 16.449 |
| | Caa1 | CCC+ | CCC+ | 5 | 21.047 | 22.530 | 16.152 | 19.910 |
| | Caa2 | CCC | CCC | 4 | 25.046 | 26.832 | 18.720 | 23.532 |
| | Caa3 | CCC- | CCC- | 3 | 29.041 | 31.131 | 21.286 | 27.153 |
| | Ca | CC | CC | 2 | 33.118 | 35.519 | 23.896 | 30.844 |
| | C | C,SD,D | C,DDD, DD,RD,D | 1 | 37.255 | 39.974 | 26.538 | 34.589 |

*Note:* *Fitted yield is smoothed using loess method with first degree.*

## A.2 Correlation matrix

Table 16 list the Pearson correlations between the regressors used in the empirical model. Table 17 shows the test statistics of three different multicollinearity tests - Variance Inflation Factor (VIF), Leamer's method (Leamer) and Corrected Variance Inflation Factor (VIF). The threshold used for the VIF and CVIF is 5 and all variables statistics are below the threshold. For Leamer's method, if $x_j$ is uncorrelated with other variables, $c_j$, would be one. All regressors' statistics are near to one. The above three tests failed to detect multicollinearity.

Table 16: Correlation matrix

| | GDP | ΔGDP | CA,% | Gsurpl | Ext D | Resrv | Trade | Default | Jobless | Finrisk | Qual'ty | Gov'nce | Lobbying | Trade | Tongue | Relig | Distan |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Log GDP per capita | 1.00 | | | | | | | | | | | | | | | | |
| GDP growth | −0.26 | 1.00 | | | | | | | | | | | | | | | |
| Current account/GDP | 0.22 | 0.01 | 1.00 | | | | | | | | | | | | | | |
| Budget balance/ΔGDP | 0.02 | −0.02 | −0.00 | 1.00 | | | | | | | | | | | | | |
| External debt/GDP | 0.31 | −0.11 | 0.08 | 0.00 | 1.00 | | | | | | | | | | | | |
| International Reserves | 0.20 | 0.07 | 0.24 | 0.02 | −0.27 | 1.00 | | | | | | | | | | | |
| Trade | 0.12 | 0.07 | 0.19 | −0.00 | 0.10 | −0.05 | 1.00 | | | | | | | | | | |
| Previous default | −0.15 | −0.08 | 0.01 | 0.00 | −0.03 | −0.12 | −0.09 | 1.00 | | | | | | | | | |
| Unemployment rate | −0.13 | −0.19 | −0.17 | −0.02 | −0.08 | −0.23 | −0.09 | 0.14 | 1.00 | | | | | | | | |
| Financial risk | −0.06 | −0.23 | −0.15 | 0.01 | 0.04 | −0.24 | −0.15 | 0.16 | 0.23 | 1.00 | | | | | | | |
| Institutional quality | 0.71 | −0.26 | −0.00 | 0.01 | 0.28 | 0.06 | 0.06 | −0.16 | −0.07 | 0.05 | 1.00 | | | | | | |
| Governance | 0.66 | 0.03 | 0.19 | 0.04 | 0.22 | 0.19 | 0.20 | −0.29 | −0.27 | −0.16 | 0.56 | 1.00 | | | | | |
| *Lobbying power* | 0.20 | −0.03 | 0.20 | 0.03 | −0.04 | 0.43 | −0.11 | 0.00 | −0.19 | −0.13 | 0.17 | 0.20 | 1.00 | | | | |
| *Trade proximity* | 0.15 | −0.01 | 0.09 | 0.01 | −0.04 | 0.42 | −0.10 | −0.05 | −0.19 | −0.08 | 0.10 | 0.16 | 0.32 | 1.00 | | | |
| *Common language* | 0.00 | −0.01 | 0.01 | −0.04 | 0.03 | 0.02 | −0.04 | −0.04 | 0.03 | 0.05 | 0.14 | 0.09 | 0.13 | 0.09 | 1.00 | | |
| *Religious proximity* | 0.24 | −0.17 | −0.14 | −0.01 | 0.12 | −0.13 | −0.07 | 0.03 | 0.05 | 0.12 | 0.30 | −0.01 | −0.07 | 0.07 | 0.03 | 1.00 | |
| *Nearest physical distance* | −0.17 | 0.20 | 0.14 | −0.02 | −0.08 | 0.21 | 0.28 | −0.08 | −0.15 | −0.12 | −0.12 | 0.06 | 0.05 | −0.14 | 0.27 | −0.37 | 1.00 |

Table 17: Multicollinearity Test

| | VIF | Leamer | CVIF |
|---|---|---|---|
| Log GDP per capita | 3.46 | 0.54 | -1.36 |
| GDP growth | 1.29 | 0.88 | -0.50 |
| Current account/GDP | 1.26 | 0.89 | -0.49 |
| Budget balance/GDP growth | 1.01 | 1.00 | -0.39 |
| External debt/GDP | 1.32 | 0.87 | -0.52 |
| International Reserves | 1.95 | 0.72 | -0.76 |
| Trade | 1.26 | 0.89 | -0.49 |
| Previous default | 1.14 | 0.94 | -0.45 |
| Unemployment rate | 1.24 | 0.90 | -0.49 |
| Financial risk | 1.20 | 0.91 | -0.47 |
| Institutional quality | 2.50 | 0.63 | -0.98 |
| Governance | 2.41 | 0.64 | -0.95 |
| Lobbying power | 1.38 | 0.85 | -0.54 |
| Trade proximity | 1.43 | 0.84 | -0.56 |
| Common language | 1.23 | 0.90 | -0.48 |
| Religious proximity | 1.37 | 0.85 | -0.54 |
| Nearest geographical distance | 1.71 | 0.76 | -0.67 |

## A.3 Panel regression estimation results per credit rating agency

The estimation results of random effects ordered logit regression for the first-pass score are reported in Table 18.

## A.4 Robustness checks and additional tests

We follow several test procedures to verify the robustness of our models and results. For the logit model, we allow for time variation in the coefficient estimates. For the machine learning models, we calculate the prediction accuracy for each rating class and rating forecast from 1 to 5 years ahead. We also check the subjective component per rating class for the first-pass score. We discuss each of these additional analyses in the following paragraphs.

### A.4.1 Linear model: Time varying coefficients

Ferri et al. (1999) and Polo (2011) demonstrate that credit rating models change through time and are in general dependent on the state of the economy. Credit rating agencies tend to be more conservative during downturns and more optimistic during boom phases. They assign a greater importance to economic and political fundamentals shortly after a crisis. To check for the potential importance of time-varying coefficient estimates, we re-run the cross-sectional regressions for each quarter from 1995-Q2 to 2014-Q1. Figure 5, already discussed when we checked for the effect of the new regulation, plots the quarterly OLS coefficient estimates for each variable included in the baseline regression model 2 (see Table 5). We conclude that most coefficient estimates have remained relatively stable since 2002. The jumpy pattern in the graphs for the first five years of our sample is at least partly caused by the low number of observations for this period, which is also the reason why we applied OLS in stead of a logit regression. We do not observe a significant impact due to the global financial crisis of 2008–2010.

### A.4.2 Machine learning: Out-of-sample prediction accuracy per rating class

We compute the prediction accuracies per rating class and report them in Tables 24 to 27 in the Appendix. When we compare the deviation(s) in notch(es), then for investment-grade countries and junk-bond countries our predictions are relatively cautious; that is, they tend to be below the actual rating for investment-grade countries, and above them for the other countries.

Table 18: Random Effects ordered logit estimation results for the determinants of Average, Moodys's, S&P and Fitch sovereign credit ratings

| | *Dependent: First-pass score* | | | |
|---|---|---|---|---|
| | Average | Moody's | S&P | Fitch |
| Log GDP per capita | 2.01*** | 1.85*** | 2.67*** | 2.47*** |
| | (0.05) | (0.05) | (0.06) | (0.06) |
| GDP growth | −2.33 | −1.76 | −4.95 | −5.49 |
| | (4.18) | (4.13) | (5.09) | (4.58) |
| Current account/GDP | −0.69** | −0.64** | −0.66 | −3.52*** |
| | (0.31) | (0.31) | (0.41) | (0.44) |
| External debt/GDP | −0.28*** | −0.06** | −0.40*** | −0.37*** |
| | (0.04) | (0.03) | (0.03) | (0.04) |
| International Reserves | 0.03 | 0.16*** | −0.35*** | −0.20*** |
| | (0.03) | (0.02) | (0.03) | (0.03) |
| Trade | 0.85*** | 0.17* | 0.26** | 0.93*** |
| | (0.16) | (0.10) | (0.10) | (0.11) |
| Previous default | −4.06*** | −2.48*** | −3.41*** | −3.63*** |
| | (0.35) | (0.33) | (0.40) | (0.44) |
| Unemployment rate | −17.89*** | −14.52*** | −26.27*** | −25.65*** |
| | (0.77) | (0.75) | (0.93) | (0.91) |
| Financial risk | −2.73*** | −2.02*** | −3.51*** | −4.04*** |
| | (0.22) | (0.21) | (0.23) | (0.26) |
| Institutional quality | 3.50*** | 2.34*** | 1.69*** | 2.08*** |
| | (0.24) | (0.17) | (0.18) | (0.20) |
| Governance | 7.85*** | 8.15*** | 8.91*** | 8.57*** |
| | (0.28) | (0.28) | (0.28) | (0.31) |
| Lobbying power | 0.09*** | 0.06*** | 0.08*** | 0.11*** |
| | (0.02) | (0.02) | (0.01) | (0.02) |
| Trade proximity | 13.79*** | 13.84*** | 8.47*** | 8.13*** |
| | (1.66) | (1.63) | (1.67) | (1.50) |
| Common language | 1.99*** | 1.31*** | 2.58*** | 1.69*** |
| | (0.10) | (0.10) | (0.10) | (0.12) |
| Religious proximity | 1.01*** | 0.32*** | 0.38*** | 0.10*** |
| | (0.04) | (0.03) | (0.03) | (0.03) |
| Nearest geographical distance | −0.01*** | −0.01*** | −0.00** | −0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) |
| C\|CC | | −7.76*** | −5.10*** | −10.01*** |
| | | (0.70) | (0.56) | (1.12) |
| CC\|CCC- | −9.78*** | −7.51*** | −4.85*** | |
| | (0.71) | (0.67) | (0.56) | |
| CC\|CCC | | | | −8.39*** |
| | | | | (0.68) |
| CCC-\|CCC | −7.22*** | −5.99*** | −4.81*** | |
| | (0.48) | (0.56) | (0.55) | |
| CCC\|CCC+ | −6.24*** | −4.84*** | −4.65*** | −6.14*** |
| | (0.45) | (0.53) | (0.55) | (0.53) |
| CCC+\|B- | −4.52*** | −0.95* | −3.18*** | −5.67*** |
| | (0.42) | (0.52) | (0.54) | (0.52) |
| B-\|B | −2.05*** | 1.14** | −0.83 | −1.89*** |
| | (0.41) | (0.49) | (0.53) | (0.51) |
| B\|B+ | 0.10 | 2.97*** | 1.08** | 0.31 |
| | (0.41) | (0.47) | (0.52) | (0.50) |
| B+\|BB- | 2.87*** | 5.22*** | 3.54*** | 2.50*** |
| | (0.40) | (0.46) | (0.50) | (0.49) |
| BB-\|BB | 4.47*** | 6.33*** | 5.68*** | 4.54*** |
| | (0.40) | (0.46) | (0.49) | (0.49) |
| BB\|BB+ | 6.28*** | 7.83*** | 7.86*** | 5.95*** |
| | (0.40) | (0.46) | (0.50) | (0.48) |
| BB+\|BBB- | 8.07*** | 9.66*** | 9.46*** | 8.20*** |
| | (0.41) | (0.47) | (0.51) | (0.50) |
| BBB-\|BBB | 10.13*** | 11.86*** | 11.17*** | 10.43*** |
| | (0.43) | (0.49) | (0.53) | (0.51) |
| BBB\|BBB+ | 11.67*** | 12.90*** | 12.91*** | 12.28*** |
| | (0.44) | (0.50) | (0.54) | (0.53) |
| BBB+\|A- | 12.81*** | 14.23*** | 14.18*** | 13.91*** |
| | (0.45) | (0.51) | (0.55) | (0.54) |
| A-\|A | 14.30*** | 15.06*** | 15.70*** | 15.65*** |
| | (0.46) | (0.52) | (0.56) | (0.55) |
| A\|A+ | 16.01*** | 16.37*** | 17.75*** | 17.16*** |
| | (0.47) | (0.53) | (0.57) | (0.57) |
| A+\|AA- | 17.23*** | 17.93*** | 19.21*** | 18.82*** |
| | (0.48) | (0.55) | (0.58) | (0.58) |
| AA-\|AA | 18.66*** | 18.90*** | 21.18*** | 20.49*** |
| | (0.49) | (0.55) | (0.60) | (0.59) |
| AA\|AA+ | 21.18*** | 20.91*** | 23.43*** | 23.40*** |
| | (0.51) | (0.57) | (0.62) | (0.62) |
| AA+\|AAA | 23.27*** | 21.77*** | 25.84*** | 25.43*** |
| | (0.53) | (0.57) | (0.63) | (0.64) |
| McFadden $R^2$ | 0.524 | 0.492 | 0.531 | 0.553 |
| *Note:* | | | | $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ |

Table 19: Forecasting validity of the random forest model

| Year | Nobs | *Sovereign rating forecast* % predicted within $n$ notches | | | |
|---|---|---|---|---|---|
| | | $n = 0$ | $n = 1$ | $n = 2$ | $n = 3$ |
| 2010–2010 | 377 | 78.78 | 95.49 | 96.82 | 98.67 |
| 2010–2011 | 745 | 66.85 | 90.34 | 94.36 | 96.78 |
| 2010–2012 | 1102 | 58.62 | 84.85 | 91.56 | 94.56 |
| 2010–2013 | 1410 | 53.69 | 80.92 | 88.94 | 93.12 |
| 2010–2014 | 1481 | 52.8 | 80.35 | 88.59 | 92.78 |

**Note:** This table shows the percentage of correctly predicted ratings for the random forest model. We used training data from 1995-Q2 to 2009-Q4 and predict the data from 2010 onwards.

### A.4.3 Sovereign rating forecast

We investigate the accuracy of rating forecasts of the random forests model over a horizon of 1 year to 5 years ahead. Table 19 presents the results for the average rating of the three rating agencies. The forecasting accuracy decreases over longer horizons, which is obvious since different stochastic scenarios can play a role in the credit ratings. Within two notches of deviations, the decision tree model correctly forecasts at least 94% of the test sample for two years.

### A.4.4 First-pass score: The magnitude of subjective components

We compute the average objective and subjective components per class for the first-pass score and report the results in Table 20 and find that the results are comparable to our preferred version of the adjusted rating scale. The impact of the subjective component is positively related to the credit rating. According to the machine learning algorithm, the subjective component is most influential for countries with a fair rating just below investment grade and for the most risky-looking countries. According to the ordered logit model, there is no jump in subjectivity at the investment grade border, on the contrary, the impact of subjective judgment is considered to be the lowest for these rating classes.

## A.5 Alternative estimation results for the spread adjusted ratings

As an alternative to the random effects ordered logit model, we estimate the model by applying fixed effects panel regression and panel corrected standard errors. The estimation results are reported respectively in Table 21.

Table 20: Average subjective component per rating class and weight of subjectivity for first pass score

| | | *Panel A: RE Ordered Logit* | | | | | *Panel B: Random Forests* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Rating | score | Subjective part units | Subjective part notch(es) | objective score | fair rating | Rating | score | Subjective part units | Subjective part notch(es) | objective score | fair rating |
| AAA | 21 | 3.074 | +4 | 17.916 | A+ | AAA | 21 | 0.189 | +1 | 20.810 | AA+ |
| AA+ | 20 | 2.444 | +3 | 17.344 | A+ | AA+ | 20 | 1.119 | +2 | 18.815 | AA− |
| AA | 19 | 1.548 | +2 | 17.189 | A+ | AA | 19 | −0.588 | 0 | 19.561 | AA |
| AA− | 18 | 0.756 | +2 | 16.863 | A | AA− | 18 | 0.500 | +1 | 17.512 | A+ |
| A+ | 17 | 0.116 | +1 | 16.418 | A | A+ | 17 | 0.667 | +1 | 16.317 | A |
| A | 16 | −0.054 | 0 | 16.047 | A | A | 16 | −0.088 | 0 | 16.082 | A |
| A− | 15 | 0.085 | +1 | 14.933 | BBB+ | A− | 15 | 1.493 | +2 | 13.504 | BBB |
| BBB+ | 14 | 0.397 | +1 | 13.507 | BBB | BBB+ | 14 | 3.073 | +4 | 10.868 | BB |
| BBB | 13 | −0.458 | 0 | 13.393 | BBB | BBB | 13 | 1.782 | +2 | 11.184 | BB+ |
| BBB− | 12 | −0.411 | 0 | 12.698 | BBB− | BBB− | 12 | 1.050 | +2 | 10.923 | BB |
| BB+ | 11 | −0.846 | −1 | 12.059 | BBB− | BB+ | 11 | 1.013 | +2 | 9.992 | BB− |
| BB | 10 | −1.356 | −1 | 11.792 | BB+ | BB | 10 | 0.274 | +1 | 9.726 | BB− |
| BB− | 9 | −1.749 | −2 | 11.289 | BB+ | BB− | 9 | 0.259 | +1 | 8.734 | B+ |
| B+ | 8 | −2.389 | −2 | 10.659 | BB | B+ | 8 | −0.022 | 0 | 8.034 | B+ |
| B | 7 | −3.025 | −3 | 10.463 | BB | B | 7 | −0.940 | 0 | 7.986 | B |
| B− | 6 | −3.704 | −4 | 10.291 | BB | B− | 6 | −0.673 | 0 | 6.695 | B− |
| CCC+ | 5 | −3.911 | −4 | 9.848 | BB− | CCC+ | 5 | −1.608 | −1 | 6.848 | B− |
| CCC | 4 | −3.222 | −4 | 8.889 | B+ | CCC | 4 | −2.741 | −3 | 7.519 | B |
| CCC− | 3 | −2.783 | −6 | 9.304 | BB− | CCC− | 3 | −1.783 | −2 | 5.130 | CCC+ |
| CC | 2 | −3.500 | −5 | 7.250 | B | CC | 2 | −3.500 | −3 | 5.750 | CCC+ |

***Note:*** This table shows the average subjective component per rating class and the weight of the subjective component. The sixth column shows the fitted objective rating (rounded to the closest rating scale). The results for the ordered logit model are presented in Panel A and Panel B shows the results for the random forests classification.

Table 21: Fixed Effects and PCSE

| | *Dependent: Spread adjusted scaled ratings* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Panel A: Fixed Effects* | | | | *Panel B: PCSE* | | | |
| | Average | Moody's | S&P | Fitch | Average | Moody's | S&P | Fitch |
| Log GDP per capita | −0.95*** | −0.43*** | −1.53*** | −0.84*** | −1.461*** | −1.145*** | −1.046*** | −1.127*** |
| | (0.09) | (0.11) | (0.13) | (0.07) | (0.155) | (0.15) | (0.3) | (0.156) |
| GDP growth | −13.22*** | −11.84** | −18.87*** | −14.51*** | −18.489** | −13.027* | −49.605*** | −15.602* |
| | (4.49) | (5.28) | (6.90) | (3.39) | (7.84) | (7.118) | (13.384) | (8.454) |
| Current account/GDP | 1.07*** | 1.12*** | 2.06*** | 0.50 | 0.167 | −0.177 | 0.831 | −0.082 |
| | (0.35) | (0.40) | (0.78) | (0.40) | (0.287) | (0.289) | (0.508) | (0.264) |
| External debt/GDP | 0.05* | 0.05 | 0.04 | 0.04* | 0.011 | −0.011 | 0.022 | −0.009 |
| | (0.03) | (0.03) | (0.04) | (0.02) | (0.018) | (0.015) | (0.032) | (0.013) |
| International Reserves | −0.14*** | −0.37*** | −0.12* | −0.13*** | −0.549*** | −0.498*** | −0.599*** | −0.264*** |
| | (0.04) | (0.05) | (0.06) | (0.03) | (0.077) | (0.068) | (0.141) | (0.049) |
| Trade | 0.36 | 0.09 | 0.84** | −0.12 | 0.173 | −0.216 | −0.087 | −0.345 |
| | (0.24) | (0.28) | (0.36) | (0.18) | (0.201) | (0.207) | (0.354) | (0.215) |
| Previous default | 8.81*** | 8.76*** | 13.30*** | 5.04*** | 3.609*** | 1.61*** | 5.683*** | 2.515*** |
| | (0.30) | (0.35) | (0.47) | (0.25) | (0.528) | (0.472) | (1.102) | (0.613) |
| Unemployment rate | 10.03*** | 16.18*** | 16.83*** | 10.37*** | 0.659 | −0.545 | 10.802*** | 2.924* |
| | (1.39) | (1.66) | (2.14) | (1.01) | (1.601) | (1.712) | (3.126) | (1.541) |
| Financial risk | 3.24*** | 1.70*** | 5.10*** | 2.48*** | 1.007** | −0.266 | 1.149* | 1.309*** |
| | (0.22) | (0.26) | (0.34) | (0.18) | (0.456) | (0.397) | (0.687) | (0.368) |
| Institutional quality | −3.38*** | −4.00*** | −3.46*** | −1.70*** | −4.368*** | −6.021*** | −7.237*** | −2.797*** |
| | (0.34) | (0.40) | (0.53) | (0.29) | (0.473) | (0.458) | (0.876) | (0.453) |
| Governance | −5.23*** | −5.27*** | −7.08*** | −3.01*** | −5.693*** | −5.286*** | −6.166*** | −4.652*** |
| | (0.30) | (0.35) | (0.45) | (0.23) | (0.536) | (0.526) | (0.842) | (0.643) |
| Lobbying power | −0.08*** | −0.07*** | −0.13*** | −0.04*** | −0.043*** | −0.052*** | −0.066*** | −0.032* |
| | (0.02) | (0.02) | (0.03) | (0.01) | (0.015) | (0.015) | (0.022) | (0.018) |
| Trade proximity | 6.51 | 2.21 | 13.47* | 7.69* | 3.111 | −1.421 | 2.925 | −8.138*** |
| | (5.52) | (6.37) | (7.81) | (4.02) | (3.005) | (2.901) | (4.053) | (2.81) |
| Common language | | | | | −0.307 | 0.539 | 0.101 | −0.563** |
| | | | | | (0.285) | (0.341) | (0.487) | (0.238) |
| Religious proximity | −0.89*** | −0.30 | −1.91*** | −1.40*** | −0.362*** | 0.124 | −0.726*** | 0.145 |
| | (0.33) | (0.38) | (0.47) | (0.30) | (0.132) | (0.142) | (0.194) | (0.109) |
| Nearest geographical distance | | | | | −0.016*** | −0.01** | −0.008 | −0.005 |
| | | | | | (0.004) | (0.005) | (0.008) | (0.004) |
| Intercept | | | | | 19.578*** | 16.887*** | 15.109*** | 14.521*** |
| | | | | | (1.296) | (1.327) | (2.207) | (1.41) |
| $R^2$ | 0.42 | 0.36 | 0.42 | 0.41 | 0.58 | 0.47 | 0.45 | 0.51 |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

## A.6 Maximalist: Subjective judgment over time

Table 22: Subjectivity over time. t-tests on average subjective component per rating class, 2010-2014, compared to average for two preceding five-year periods (***maximalist***)

| Year | Rating | O_logit | R_forests | Rating | O_logit | R_forests |
|---|---|---|---|---|---|---|
| 2001-2006 | AAA | 1.149 | 4.522*** ↗ | BB | −3.606*** ↗ | −3.373*** ↙ |
| 2007-2009 | AAA | 2.417** ↗ | 3.863*** ↗ | BB | 2.531** ↙ | 0.697 |
| 2001-2006 | AA | 3.347*** ↗ | 4.422*** ↗ | B | −3.915*** ↗ | −3.176*** ↗ |
| 2007-2009 | AA | 5.492*** ↗ | 4.297*** ↗ | B | 0.637 | −0.061 |
| 2001-2006 | A | −1.149 | −1.117 | CCC | −5.526*** ↗ | −1.372 |
| 2007-2009 | A | 2.391** ↗ | 1.903* ↗ | CCC | −4.004*** ↗ | −1.52 |
| 2001-2006 | BBB | −3.865*** ↙ | −6.397*** ↙ | | | |
| 2007-2009 | BBB | −2.629*** ↙ | −0.8 | | | |

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

*Note:* This table shows respectively the t-statistic for the difference in average subjective component for the period '01 to '06 and the period '10 to '14 and the t-statistic for the difference in the average subjective component for '07-'09 and the period '10 to '14. Because the subjective component can be both positive or negative, the arrows indicate whether the subjective component increased or decreased (in absolute value) over time.

## A.7 Maximalist: Subjective judgment and default prediction

We revisit the predictive validity of both the objective and subjective components by including the residuals in the calculation of subjectivity. For the random forests model, including the residuals enhances the default prediction of the third year, while the performance of the objective components mirrors the ability of the minimalist's objective components.

Table 23: Predictive power of the subjective and objective component (***maximalist***)

| | *Dependent: Default dummy* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Panel A: RE Ordered Logit* | | | | | *Panel B: Random Forests* | | | | |
| | 1yr | 2yr | 3yr | 4yr | 5yr | 1yr | 2yr | 3yr | 4yr | 5yr |
| Objective | 0.11* | 0.12* | 0.14** | 0.07 | 0.03 | 0.13*** | 0.11** | 0.14*** | 0.12** | 0.07 |
| | (0.06) | (0.07) | (0.06) | (0.08) | (0.1) | (0.04) | (0.04) | (0.04) | (0.05) | (0.06) |
| Subjective | 0.18*** | 0.09* | 0.08 | 0.12** | 0.08 | 0.18*** | 0.09* | 0.05 | 0.08 | 0.05 |
| | (0.03) | (0.05) | (0.05) | (0.06) | (0.08) | (0.04) | (0.05) | (0.06) | (0.06) | (0.09) |
| Groups (Country) | 103 | 103 | 102 | 98 | 96 | 103 | 103 | 102 | 98 | 96 |
| McFadden $R^2$ | 0.13 | 0.06 | 0.06 | 0.06 | 0.06 | 0.13 | 0.06 | 0.07 | 0.06 | 0.06 |

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Note: The signs of estimates are multiplied by minus one for the ease of interpretation. Panel A, on the left, shows the random effects ordered logit regression to predict defaults based on the subjective and objective components of the ordered logit model reported in Table 5 and lagged from 1 year to 5 years. Panel B, on the right, shows the random effects ordered logit regression to predict defaults based on the subjective and objective components of the random forests model.

## A.8 Marginal effects

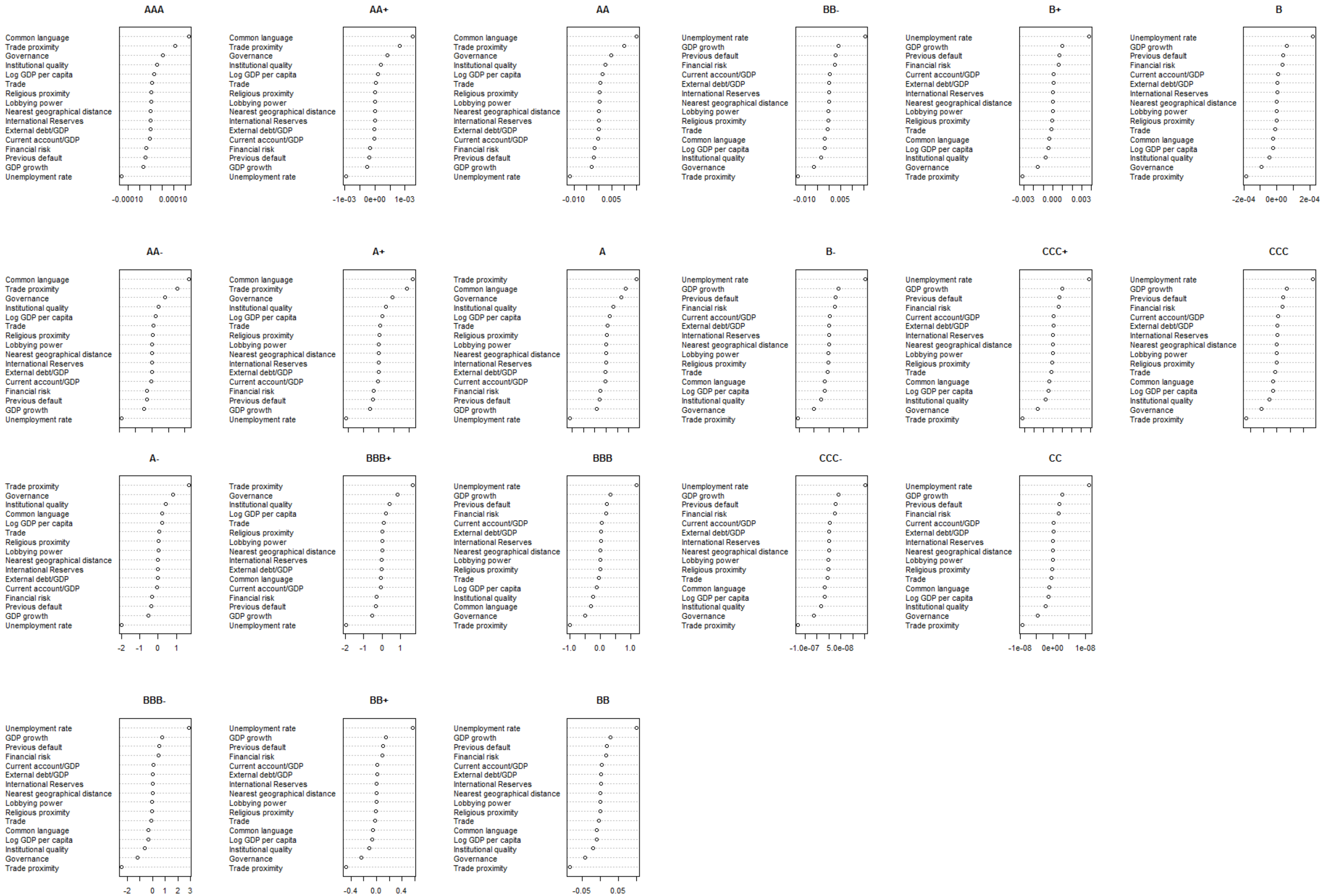



Figure 9: Marginal Effects for spread adjusted ratings

Table 24: Moody's Rating class out-of-sample predictions %

| Predicted/Actual Rating | AAA | AA+ | AA | AA- | A+ | A | A- | BBB+ | BBB | BBB- | BB+ | BB | BB- | B+ | B | B- | CCC+ | CCC | CCC- | CC | C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AAA | 100 | 8.33 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA+ | - | 83.33 | 5.71 | 2.5 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA | - | 4.17 | 92.86 | 2.5 | - | 1.19 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA- | - | - | 1.43 | 85 | 1.16 | 1.19 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| A+ | - | - | - | 10 | 96.51 | 10.71 | - | 1.18 | - | - | 0.88 | - | - | - | - | - | - | - | - | - | - |
| A | - | - | - | - | 2.33 | 85.71 | 1.85 | 4.71 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| A- | - | - | - | - | - | - | 88.89 | 1.18 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB+ | - | 4.17 | - | - | - | 1.19 | 5.56 | 87.06 | 3.03 | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB | - | - | - | - | - | - | 1.85 | 3.53 | 93.94 | 2.76 | - | - | - | - | - | - | - | - | - | - | - |
| BBB- | - | - | - | - | - | - | 1.85 | 1.18 | 3.03 | 93.79 | 7.02 | 2.35 | - | - | - | 1.67 | - | - | - | - | - |
| BB+ | - | - | - | - | - | - | - | - | - | 2.76 | 89.47 | 3.53 | - | - | - | - | - | - | - | - | - |
| BB | - | - | - | - | - | - | - | 1.18 | - | 0.69 | 2.63 | 90.59 | - | 0.95 | - | - | - | - | - | - | - |
| BB- | - | - | - | - | - | - | - | - | - | - | - | 3.53 | 92.73 | 4.76 | - | - | - | - | - | - | - |
| B+ | - | - | - | - | - | - | - | - | - | - | - | - | 3.64 | 88.57 | 7.89 | 1.67 | - | - | - | - | - |
| B | - | - | - | - | - | - | - | - | - | - | - | - | - | 2.86 | 90.79 | 10 | 2.38 | - | - | - | - |
| B- | - | - | - | - | - | - | - | - | - | - | - | - | - | 2.86 | - | 83.33 | 9.52 | - | 50 | - | - |
| CCC+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.32 | 3.33 | 80.95 | 50 | - | - | - |
| CCC | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 2.38 | 50 | - | - | - |
| CCC- | - | - | - | - | - | - | - | - | - | - | - | - | 3.64 | - | - | - | 2.38 | - | 50 | - | - |
| CC | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 2.38 | - | - | 100 | - |
| C | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 100 |
| Num_obs | 283 | 24 | 70 | 40 | 86 | 84 | 54 | 85 | 66 | 145 | 114 | 85 | 55 | 105 | 76 | 60 | 42 | 6 | 4 | 1 | 2 |

Table 25: Standard Poors' Rating class out-of-sample predictions %

| Predicted/Actual Rating | AAA | AA+ | AA | AA- | A+ | A | A- | BBB+ | BBB | BBB- | BB+ | BB | BB- | B+ | B | B- | CCC+ | CCC | CCC- | CC | C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AAA | 100 | 5.45 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA+ | - | 92.73 | 4.55 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA | - | 1.82 | 84.09 | 1.82 | 1.79 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA- | - | - | 4.55 | 89.09 | 1.79 | 0.96 | 2.33 | - | - | - | - | 0.81 | - | - | - | - | - | - | - | - | - |
| A+ | - | - | 6.82 | 5.45 | 92.86 | 2.88 | 1.16 | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| A | - | - | - | 3.64 | 3.57 | 90.38 | 5.81 | - | 0.98 | - | - | - | - | - | - | - | - | - | - | - | - |
| A- | - | - | - | - | - | 1.92 | 87.21 | 2.9 | - | - | 2.33 | - | - | - | - | - | - | - | - | - | - |
| BBB+ | - | - | - | - | - | - | 3.49 | 81.16 | 4.9 | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB | - | - | - | - | - | 1.92 | - | 14.49 | 90.2 | 7.45 | 1.16 | - | - | - | - | - | - | - | - | - | - |
| BBB- | - | - | - | - | - | - | - | 1.45 | 3.92 | 89.36 | 2.33 | - | - | - | 1.59 | - | - | - | - | - | - |
| BB+ | - | - | - | - | - | - | - | - | - | 2.13 | 91.86 | 4.03 | 1.09 | - | - | - | - | - | - | - | - |
| BB | - | - | - | - | - | - | - | - | - | - | 2.33 | 93.55 | 4.35 | - | - | - | - | - | - | - | - |
| BB- | - | - | - | - | - | - | - | - | - | - | - | 1.61 | 84.78 | 1.08 | 1.59 | 1.96 | - | - | - | - | - |
| B+ | - | - | - | - | - | 1.92 | - | - | - | 1.06 | - | - | 8.7 | 92.47 | 6.35 | 3.92 | - | - | - | - | - |
| B | - | - | - | - | - | - | - | - | - | - | - | - | - | 6.45 | 85.71 | 5.88 | 5.88 | - | - | - | - |
| B- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 4.76 | 84.31 | 23.53 | - | - | - | 18.18 |
| CCC+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.96 | 64.71 | - | - | - | - |
| CCC | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.96 | - | - | - | - | - |
| CCC- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| CC | - | - | - | - | - | - | - | - | - | - | - | - | 1.09 | - | - | - | 5.88 | - | - | 100 | - |
| C | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 100 | - | - | 81.82 |
| Num_obs | 243 | 55 | 44 | 55 | 56 | 104 | 86 | 69 | 102 | 94 | 86 | 124 | 92 | 93 | 63 | 51 | 17 | 1 | 0 | 1 | 11 |

Table 26: Fitch Ratings class out-of-sample predictions %

| Predicted/Actual Rating | AAA | AA+ | AA | AA- | A+ | A | A- | BBB+ | BBB | BBB- | BB+ | BB | BB- | B+ | B | B- | CCC+ | CCC | CCC- | CC | C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AAA | 99.59 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA+ | 0.41 | 92.11 | 1.37 | 4.35 | - | 1.67 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA | - | 7.89 | 93.15 | 6.52 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA- | - | - | 5.48 | 86.96 | 3.28 | - | - | - | - | - | - | - | - | - | - | 1.92 | - | - | - | - | - |
| A+ | - | - | - | 2.17 | 91.8 | 5 | 1.47 | - | - | 1.69 | - | - | - | - | - | - | - | - | - | - | - |
| A | - | - | - | - | 4.92 | 90 | 1.47 | 3.03 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| A- | - | - | - | - | - | 3.33 | 92.65 | 4.55 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB+ | - | - | - | - | - | - | 2.94 | 86.36 | 7.06 | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB | - | - | - | - | - | - | 1.47 | 6.06 | 87.06 | 4.24 | - | - | - | - | - | - | - | - | - | - | - |
| BBB- | - | - | - | - | - | - | - | - | 3.53 | 88.14 | 5.45 | 1.72 | - | - | - | - | - | - | - | - | - |
| BB+ | - | - | - | - | - | - | - | - | 2.35 | 5.93 | 90 | 3.45 | 1.45 | - | - | - | - | - | - | - | - |
| BB | - | - | - | - | - | - | - | - | - | - | 0.91 | 89.66 | 2.9 | - | - | - | - | - | - | - | - |
| BB- | - | - | - | - | - | - | - | - | - | - | 0.91 | 3.45 | 86.96 | 10.53 | - | - | - | - | - | - | - |
| B+ | - | - | - | - | - | - | - | - | - | - | - | - | 7.25 | 80.7 | 9.76 | - | - | - | - | - | - |
| B | - | - | - | - | - | - | - | - | - | - | 2.73 | - | - | 7.02 | 78.05 | 3.85 | - | - | - | - | - |
| B- | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.75 | 12.2 | 92.31 | 25 | 22.22 | - | - | - |
| CCC+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 75 | - | - | - | - |
| CCC | - | - | - | - | - | - | - | - | - | - | - | 1.72 | - | - | - | 1.92 | - | 77.78 | - | - | - |
| CCC- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| CC | - | - | - | - | - | - | - | - | - | - | - | - | 1.45 | - | - | - | - | - | - | 100 | - |
| C | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Num_obs | 241 | 38 | 73 | 46 | 61 | 60 | 68 | 66 | 85 | 118 | 110 | 58 | 69 | 57 | 41 | 52 | 4 | 9 | 0 | 1 | 0 |

Table 27: Average Ratings class out-of-sample predictions %

| Predicted/Actual Rating | AAA | AA+ | AA | AA- | A+ | A | A- | BBB+ | BBB | BBB- | BB+ | BB | BB- | B+ | B | B- | CCC+ | CCC | CCC- | CC | C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AAA | 99.6 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA+ | 0.4 | 95.56 | 1.47 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA | - | 4.44 | 91.18 | 8 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| AA- | - | - | 7.35 | 92 | 3.57 | - | - | - | - | - | 0.9 | - | - | - | - | - | - | - | - | - | - |
| A+ | - | - | - | - | 91.07 | 7.37 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| A | - | - | - | - | 5.36 | 91.58 | 4.76 | - | 3.12 | - | 0.9 | - | - | - | - | - | - | - | - | - | - |
| A- | - | - | - | - | - | 1.05 | 92.86 | 6.15 | 1.04 | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB+ | - | - | - | - | - | - | 2.38 | 84.62 | 6.25 | - | - | - | - | - | - | - | - | - | - | - | - |
| BBB | - | - | - | - | - | - | - | 7.69 | 84.38 | 3.79 | - | - | 2.56 | - | - | - | - | - | - | - | - |
| BBB- | - | - | - | - | - | - | - | 1.54 | 4.17 | 93.94 | 6.31 | 1.83 | - | - | - | - | - | - | - | - | - |
| BB+ | - | - | - | - | - | - | - | - | 1.04 | 2.27 | 90.09 | 3.67 | - | - | - | 1.52 | - | - | - | - | - |
| BB | - | - | - | - | - | - | - | - | - | - | 1.8 | 87.16 | 7.69 | - | - | - | - | - | - | - | - |
| BB- | - | - | - | - | - | - | - | - | - | - | - | 5.5 | 83.33 | 0.81 | 1.19 | - | 4.35 | - | - | - | - |
| B+ | - | - | - | - | - | - | - | - | - | - | - | 0.92 | 6.41 | 95.16 | 4.76 | - | 4.35 | - | - | - | - |
| B | - | - | - | - | - | - | - | - | - | - | - | - | - | 4.03 | 89.29 | 4.55 | 8.7 | 12.5 | - | - | - |
| B- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 3.57 | 92.42 | 8.7 | 12.5 | 33.33 | - | - |
| CCC+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.52 | 65.22 | - | - | - | - |
| CCC | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.19 | - | 8.7 | 75 | - | - | - |
| CCC- | - | - | - | - | - | - | - | - | - | - | - | 0.92 | - | - | - | - | - | - | 66.67 | 50 | - |
| CC | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 50 | - |
| C | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Num_obs | 252 | 45 | 68 | 50 | 56 | 95 | 84 | 65 | 96 | 132 | 111 | 109 | 78 | 124 | 84 | 66 | 23 | 8 | 6 | 2 | 0 |