Understanding Sovereign Ratings and Their Implications for Emerging Economies

RAHUL S CHAUHAN, ILISA GOENKA, KAUSHALENDRA KISHORE, NIRUPAMA KULKARNI, KAVYA RAVINDRANATH, GAUTHAM UDUPA

The rating methodologies of the big three credit rating agencies—S&P, Moody's, and Fitch—are scrutinised and evaluated. The factors driving sovereign ratings are examined using a regression framework and machine learning techniques with a panel of 162 countries covering ratings from 2000 to 2018. Across all models, institutional quality is the most significant factor driving sovereign ratings, suggesting that building more vital institutions can lower a sovereign's borrowing costs by improving sovereign ratings. Additionally, only sustainable GDP growth propelled by strong structural reforms and productive investment increase CRA ratings. The findings suggest that the over-reliance of market participants on CRA ratings to assess sovereign creditworthiness may be unwarranted, particularly during crisis periods.

All opinions expressed in this paper reflect those of the authors and not necessarily those of CAFRAL.

The authors wish to thank Sitikantha Pattanaik for invaluable guidance and support during this project. They also thank Harendra Kumar Behera for providing data and for helpful discussions. They gratefully acknowledge Sangeetha Matthews and Silu Muduli for useful comments.

Rahul S Chauhan (*rahul.chauhan@chicagobooth.edu*) is a research associate at Chicago Booth School of Business. Ilisa Goenka (*ilisa.goenka@cerge-ei.cz*) is a PhD candidate at CERGE-EI. Kaushalendra Kishore (*kaushalendra.kishore@cafral.org.in*) is a research director and Nirupama Kulkarni (*nirupama.kulkarni@cafral.org.in*) is senior research director at CAFRAL. Kavya Ravindranath (*kavyaravi@gwu.edu*) is a PhD candidate at George Washington University. Gautham Udupa (*gautham.udupa@cafral.org.in*) is a research director at CAFRAL.

The COVID-19 pandemic saw countries adopt large fiscal stimulus packages and unconventional monetary measures to combat the pandemic's economic fallout. These measures raised questions of sovereigns' fiscal capacity and debt sustainability, especially for emerging market economies (EMEs). In turn, credit rating agencies (CRAs) downgraded several EMEs, including India.[1] Moody's downgraded India's rating from Baa2 (negative) to Baa3 (negative) on 1 June 2020. Fitch, too, changed its outlook on India's rating from BBB- (stable) to (negative) on 18 June 2020.[2]

Despite these downgrades, there has been limited adverse impact on capital markets in India, possibly indicating that the sovereign ratings themselves have limited new information, and market-based measures may be more timely indicators of a sovereign's creditworthiness. The question then arises, if CRA ratings are not informative, do they still matter? Mechanistic reliance by market participants leads to large effects of CRA rating changes as rating thresholds are often integrated into laws, regulations, and market practices, often leading to herding and cliff effects (Financial Stability Board 2010a, 2010b). CRA policies also prevent them from rating firms in a country above the sovereign rating, and thus sovereign ratings determine firms' rating and costs of borrowing (Almeida et al 2017; Adelino and Ferreira 2016). For instance, after India's downgrade by Moody's, six major public sector entities were also downgraded. Rating downgrades may also lead to negative feedback loops, as rating downgrades can worsen economic conditions, leading to further downgrades (Aizenman et al 2013). India's rating is just above non-investment grade status, and even a one-notch downgrade can trigger large foreign capital outflows.

Despite the importance of CRA ratings, prior literature has highlighted biases and inconsistencies in CRA ratings. Fuchs and Gehring (2017) document that CRAs display a positive bias in ratings for their home country and countries culturally similar to the CRA's own country. Additionally, ratings are higher for countries to which the home-country banks have greater risk exposure. Further, CRA methodologies are not transparent, making it difficult for market participants to assess and correct for such biases. Such systematic biases and arbitrary factors in rating downgrades can even trigger self-fulfilling prophecies, driving even relatively healthy countries to default (Gärtner et al 2011). These reasons underscore the need to study the factors that drive CRA ratings, assess their suitability for developing countries, and evaluate their ability to predict sovereign defaults.

In this paper, we examine the ratings of the largest three CRAs, Fitch, S&P, and Moody's. We structure our study as follows. First, we examine the rating methodologies of the CRAs and the quantitative and qualitative factors that drive individual CRA ratings. Second, we narrow down to a parsimonious set of factors and examine whether these can explain the variation in ratings across time and across countries. Third, we examine whether the emphasis on these factors by CRAs is justified. Fourth, we evaluate the performance of CRA sovereign ratings by examining their ability to predict sovereign default with a particular focus on (i) EMEs, and (ii) rating downgrades during crises periods. Finally, we use machine learning techniques to narrow down to the variables that predict defaults and evaluate whether retrofitting data to past defaults is a good predictor of future defaults.

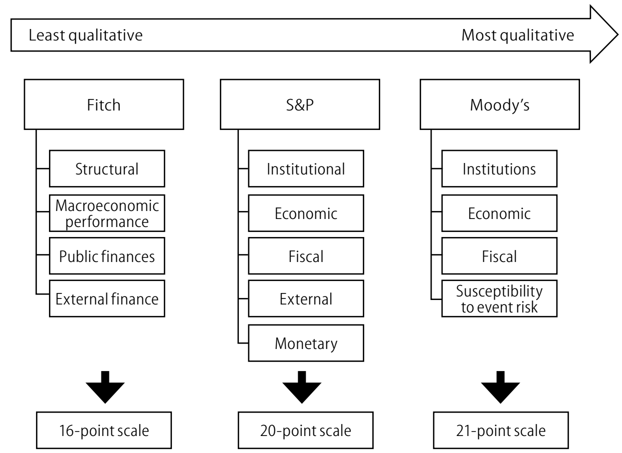
The three CRAs use complex rating methodologies based on both quantitative and qualitative factors as inputs. The input factors fall under four or five main pillars, representing a country's credit health, namely institutional, fiscal, monetary, and external factors. Fitch uses four pillars: structural, macroeconomic performance, public finances, and external finance; S&P uses five pillars, namely institutional, economic, fiscal, external, and monetary; and Moody's uses four pillars, namely institutional, economic, fiscal, and susceptibility to event risk. The CRAs also differ in their reliance on qualitative versus quantitative factors. Fitch's model is the most quantitative as it largely depends on variables that are strictly defined. Moody's is the least quantitative; while it defines a large number of factors and variables, its methodology also depends on a large number of qualitative factors and soft adjustments as inputs in the final stages. Each rating agency also varies in the final rating scale; Fitch rates on a 16-point scale, S&P on a 20-point scale, and Moody's on a 21-point scale. All three CRAs have 10 scales for investment-grade ratings and the differences in scale are in the lower non-investment grade ratings.

In the first step of the analysis, we build a parsimonious model to determine the significant quantitative factors affecting sovereign ratings. The goal is to determine whether (i) a handful of factors can explain the variation in CRA ratings, and (ii) highlight the methodological differences between the CRAs and distinguish between quantitative and qualitative factors that feed into each CRA's rating model. We use a simple ordinary least squares (OLS) specification, including select quantitative variables from each CRA's methodology report. This simple, parsimonious model explains a large proportion of the variation in CRA ratings with R^2 for 80% across CRAs. The model also identifies important variables that drive the ratings. Across the rating agencies, institutional factors are the most crucial in determining CRA ratings. Institutional factors measure the quality of a sovereign's institutions, which is likely a good predictor of a sovereign's ability to take the necessary actions to repay its debt. A one-standard deviation higher percentile rank of institutional quality—measured using the World Governance Indicators (WGI)—is associated with a three-notch higher ratings for Fitch, S&P, and Moody's. Other significant variables are gross domestic product (GDP) per capita, broad money, years since default, general government debt to GDP, and current account to GDP. Though the baseline uses a simple OLS specification that assumes cardinality of the dependent variables, our analysis is also robust to using an ordered probit.

Next, we analyse how well rating changes predict sovereign default for high- and low-middle-income countries. In particular, we focus on CRA performance during periods of crisis as rating downgrades can lead to self-fulfilling prophecies and can cause even relatively healthy countries to default (Gärtner et al 2011) due to negative feedback loops. We regress rating changes on default incidence in the near and long term, during crises and non-crisis periods, and for high-and low-middle-income countries. Crisis periods refer to the global financial crisis (GFC) of 2007–09 and the European Union (EU) sovereign debt crisis of 2010–14. We find that in the sample of all countries, Moody's and Fitch perform poorly in predicting sovereign defaults during the 2010–14 crisis, while S&P perform better. Unconditionally, Moody's performs poorly during the GFC. On analysing within-country changes, the performance for Moody's improves, which is likely driven by a few outliers. All three CRAs perform poorly during crises for low- and middle-income countries, though default prediction is relatively better for high-income countries. Together these findings suggest that, across the CRAs, rating changes do a poor job predicting default during crisis periods for low- and medium-income economies.

To conclude, we implement a supervised learning design to assess some common predictors of default used by the CRAs along with three additional factors influencing sovereign default probability hypothesised in recent literature (Chari et al 2020; Eberhardt 2018; Perez 2015). We find that the set of predictors commonly used by the CRAs can only explain 29.81% of the variation in one-year ahead default incidence, and 45% of the variation in five-year ahead default incidence. These findings are in contrast to the first stage of the analysis that showed that the handful of factors could explain nearly 80% of the variation in some CRA ratings. The supervised learning analysis suggests that the existing CRA methodologies suffer from survivorship bias as they retrofit rating criteria using characteristics of sovereigns that typically do not default. The exercise helps us evaluate predictors of sovereign default in a non-linear random forest framework, optimising the bias caused by fitting economic and financial fundamentals on sovereign default occurrences. The relative importance of each of the factors are largely in line with the CRA weightings. GDP per capita is important for predicting near-term default, whereas institutional score, external sector, and government fiscal health are relatively more important in predicting longer-term default incidences. Banking system health and financial repression also play a role in determining near-term default probability. Importantly, the model finds that mean squared error of such predictions is minimised by using only 4–6 predictors and increases as more predictors are included.

Figure 1: Overview of Sovereign Ratings Methodology



This figure presents a visual overview comparing the three methodologies. Details on these methodologies can be found in Fitch (2020), S&P (2017), and Moody's (2019).

Therefore, we find that selection biases and model complexity can dent the overall prediction accuracy of sovereign rating models (SRMs) in predicting near- and long-term sovereign default incidence and calls for caution in relying exclusively on CRA ratings.

CRA Rating Methodologies

Figure 1 gives an overview of the methodology of the CRAs. Each CRA clubs together factors into broad pillars. The figure shows the key pillars under each CRA, their respective ratings scales, and the degree of qualitative and quantitative factors. Fitch has a 16-point ratings scale, evaluated using relatively more quantitative measures, whereas Moody's has a 21-point scale evaluated using relatively more qualitative measures. S&P has a 20-point ratings scale, determined using a mix of qualitative and quantitative measures.

Fitch arrives at the sovereign long-term foreign currency issuer default rating through a two-step approach. A baseline rating score is assigned using a multivariate regression based SRM of 18 variables representing four key pillars of the sovereign's credit profile—institutions, macroeconomic performance, public finances, and external finances. The SRM generates a predicted rating for every sovereign that is then scrutinised subjectively by the agency in its qualitative overlay (QO) (Fitch 2020). Qualitative adjustments, based on predetermined metrics as well as analyst opinion are made under each of the four pillars. The final rating is the sum of the predicted baseline rating from the SRM and adjustments made under QO. Structural features evaluate governance quality, wealth, flexibility of the economy, political stability and financial sector risks. The SRM takes into account five variables that represent the institutional and structural features of the sovereign being rated. First, "Composite Governance Indicators," created as a simple average percentile rank of World Bank Governance Indicators—rule of law, government effectiveness, control of corruption, and voice and accountability, regulatory quality, political stability and absence of violence—measure the multidimensional institutional quality of the sovereign. Second, the percentile ranks of GDP per capita in US dollars at market exchange rates measure individual income and savings capacity. Third, the share of the country's nominal GDP in world GDP measures the sovereign's global reputation and size. It enters the regression as a natural logarithm of percentage share in world GDP in US dollars at market exchange rates. Fourth, years since default or restructuring enters the regression as a nonlinear function of the time since the last event and the indicator is zero if there has been no such event after 1980. For each year that elapses, the impact on the model output declines. It is also the only variable that updates the model on default history of the sovereign. Last, broad money supply as a percentage of GDP enters the regression as the natural log of the percentage ratio. These variables proxy the level of financial intermediation in the sovereign, and incorporate bank deposits, sovereign treasury bonds and other liquid financial instruments. The overall post-estimation weight of this pillar in the model is 53.7% (Fitch 2020). Barring the year since default, all variables should have a positive impact on the model output, that is, an increment in them results in a linear increment in the ratings score. In its qualitative overlay, Fitch ratings assess metrics of political stability and capacity, financial sector risks and other structural factors not captured in the SRM.

S&P arrives at a final foreign currency credit rating using a two-step approach. The initial score is calculated based on five factors shown in Figure 1. Each factor is assessed on a six-point numerical scale from "1" (strongest) to "6" (weakest). Both quantitative and qualitative considerations form the basis for these forward-looking assessments. While calculating the initial score for these factors, adjustments can be made to the score, as described in the methodology (S&P 2017). These factors are averaged into two profiles, and then an "indicative rating level" is derived from these profiles using a rating matrix. The rating matrix defines two broad profiles: institutional and economic profile and flexibility and performance profile. Institutional and economic profile is an average of institutional and economic factors, while flexibility and performance profile is an average of external, fiscal, and monetary factors. Subsequent adjustments can be made to the indicative rating to get the final foreign currency credit rating.

Moody's arrives at a final credit rating score using a two-step approach. It first calculates an initial alphanumeric score using a scorecard-based approach. It then makes changes to this score based on other considerations to arrive at a final rating. The Moody's scorecard defines four pillars shown in Figure 1, and each of these pillars have a host of sub-factors. All sub-factors have pre-assigned weights and used to aggregating up to the initial score. While calculating the initial score, adjustments to the score may also be made within each pillar based on factors as stated in the Moody's methodology report (Moody's 2019). Economic strength and fiscal strength are evaluated using quantitative metrics while institutions and governance strength and susceptibility to event risk are evaluated using quantitative factors. Overall, while Fitch relies on more quantitative factors, S&P, and to a larger extent

Moody's examine qualitative factors, especially for the EMEs, to adjust ratings upwards.

Data Source

For our analysis, we focus on 149 countries for which CRA data is available. Table A1 (p 104) lists all the countries in the analysis. Data on ratings is from Country Economy website (https://countryeconomy.com/ratings) and from the individual CRA websites. Each of these countries has a rating from at least one CRA. One hundred and six countries have a rating from all the three CRAs, indicating a good overall CRA coverage for the countries in our sample. Only 18 countries have a rating from one of the three CRAs and 26 countries have been rated by exactly two CRAs.

The remaining data on macroeconomic variables used for the individual ratings is from various sources as listed in Table 1. It provides the details of the data sources used in this analysis. Data are collected as of August of 2020. All countries for which the data is available were used in the analysis.

Common Model

Using a common regression specification, we examine whether there are differences in the way different factors, in effect, determine the CRA ratings. The common model allows comparison across CRAs and determines the importance of the factors that can enter the CRA model either as quantitative or qualitative inputs in the CRA methodologies.

Regression Framework

Our regression specification is as follows:

$$\begin{aligned} Y_{kit} &= \alpha_{it} + \beta_1 * \text{Governance Indicators}_{it} \\ &+ \beta_2 * \text{Ln(GDP per Capita)}_{it} + \beta_3 * \text{Ln(Broad Money)}_{it} \\ &+ \beta_4 * \text{Years since Default}_{it} \qquad \ldots (1) \\ &+ \beta_5 * \text{General Government Debt to GDP}_{it} \\ &+ \beta_6 * \text{Current Account Balance}_{it} \\ &+ \beta_7 * \text{GDP growth rate}_{it} + \beta_8 * \text{Inflation}_{it} + \beta_9 * \text{Net FDI}_{it} \\ &+ \beta_{10} * \text{Interest Payments}_{it} + \beta_{11} * \text{Fiscal Balance}_{it} + e_{it} \end{aligned}$$

where Y_{kit} is the rating of rating agency "*k*," for country "*i*" and for time "*t*." Thus, *k* equals Fitch ratings, S&P ratings or Moody's ratings. All variables are standardised and errors are clustered at the country level. *Governance Indicators*$_{it}$ refers to the average of country "*i*"'s percentile rank across the six WGI for year "*t*." The WGI are voice and accountability, rule of law, political stability, government effectiveness, control of corruption and regulatory quality. The baseline is estimated for the full sample of countries. We also separately look at the sub-sample of high- and low-middle-income countries, as defined using the World Bank classification.

The baseline regression does not include country-fixed effects. Since the goal is to account for how the individual factors affect ratings differentially across countries, we do not want to solely focus on the within-country variation. Nonetheless, as robustness, we also show results with country-fixed effects but maintain that the specification without country-fixed effects captures the variation we are interested in.

The results of the regression specification in equation 1 are shown in Table 2 (p 97). Columns 1–3 in Table 2 Panel A show that all three CRAs place similar weights on the same factors. Governance indicators, GDP per capita, broad money, years since default and current account balance all receive significant and positive weights, while general government debt to GDP receives a negative and significant weight. These coefficients are in the expected direction. A 1 SD higher governance indicators lead to a two-notch higher ratings across all the three CRAs, and a 1 SD increase in per capita GDP, broad money, years since default and current account balance all lead to approximately a one-notch higher rating. Similarly, a 1 SD higher debt to GDP ratio is associated with a one-notch lower rating.

Table 1: Data Sources

Variable	Data Source
A: Moody's	
GDP per capita	IMF World Economic Outlook (WEO), April 2020 and June 2020 update, World Bank World Development Indicators (WDI)
Current account/GDP	World Bank WDI
Growth of real GDP	IMF WEO, April 2020, World Bank WDI
Unemployment rate	IMF WEO, April 2020
General government debt to GDP ratio	IMF WEO, October 2019
Regulatory quality	World Bank Governance Indicators
Rule of law	World Bank Governance Indicators
B: Fitch	
Institutional score	World Bank Governance Indicators
GDP per capita	IMF WEO, April 2020 and June 2020 update, World Bank WDI
Log share in global GDP	IMF WEO, April 2020
Broad money/GDP	World Bank WDI
General government gross debt/GDP	IMF Global Debt Database
Years since default	Rogoff and Reinhart (2009) and Bank of Canada's Credit Rating Assessment Group Database of Sovereign Defaults 2019
Real GDP growth	IMF WEO April and June 2020 update
Commodity dependence	UN Conference of Trade and Development Statistics, World Bank WDI
GDP growth volatility	IMF WEO, April and June 2020 update
Consumer price inflation	IMF WEO, April 2020
Fiscal balance	IMF WEO, April 2020
Total reserves (months of imports)	World Bank WDI
Current account balance	IMF WEO, April 2020
Net FDI inflow	World Bank WDI
C: S&P	
Transparency of institutions	World Bank Governance Indicators; Global Competitiveness Index (World Economic Forum)
GDP per capita	World Bank WDI, IMF WEO, October 2019
Actively traded currency	Bank for International Settlement (BIS) report "Triennial Central Bank Survey"
Gross financing needs	Consists of current account payments from WDI, short-term external debt and long-term external debt maturing within the year from Quarterly External Data Statistics (SDDS)
Current account receipts (CAR)	World Bank WDI
Total reserves	World Bank WDI
Government gross debt (% GDP)	IMF WEU, October 2019
Central bank independence	Institutional Profiles Database

Columns 4–6 in Table 2 present the results using an ordinal probit and results are robust to this alternate specifications, consistent with prior literature (Fuchs and Gehring 2017).

For completeness, the regressions repeat the regression with country-fixed effects as robustness (Table 2, Panel B). Qualitatively, the results are similar. The GDP per capita has a larger coefficient, though broad money and current account are not significant anymore. Inflation, too is a stronger factor for S&P and Moody's. General government debt to GDP continues to be an important predictor of ratings across the board.

Tables 3 and 4 (p 98) also show the results of equation 1 separately for the sample of high- and low-income countries. While institutional factors, particularly the governance indicators, are an important determinant for CRA ratings for low-middle- and high-income countries, GDP per capita is noisier (insignificant at the 5% level) for low-middle-income countries. Effectively, through qualitative adjustments, as described earlier, the GDP per capita is a much noisier determinant of CRA ratings for low-income countries. S&P is the only agency for which GDP growth rate is positive and significant.

How Well Do Credit Ratings Predict Default Incidence?

Estimating sovereign default risk has been one of the *raison d'etre* of credit rating agencies. We compile default and restructuring data from Reinhart and Rogoff (2009) and Beers and de Leon-Manlagnit (2019) for 162 sovereigns for the time

Table 3: Common Regression—High Income

	OLS		
	(1)	(2)	(3)
	S&P	Moody's	Fitch
Governance indicators	4.781***	4.438***	3.877***
	(0.575)	(0.612)	(0.752)
Ln(GDP per capita)	3.624***	3.920***	3.398***
	(0.980)	(1.061)	(0.765)
Ln(broad money [% of GDP])	1.603***	0.954*	1.362***
	(0.422)	(0.458)	(0.402)
Years since default	-0.851	-0.459	-0.0512
	(0.514)	(0.536)	(0.561)
General government debt to GDP	-0.735***	-0.744**	-1.154***
	(0.229)	(0.268)	(0.207)
Current account (% of GDP)	0.211	0.212	0.454
	(0.261)	(0.312)	(0.281)
GDP growth rate	0.128	0.212	0.106
	(0.170)	(0.282)	(0.215)
Inflation	-1.181	-1.713	-2.977
	(2.021)	(2.688)	(2.244)
Net FDI (% of GDP)	-0.250	-0.248	0.0769
	(0.367)	(0.577)	(0.0917)
Interest payments (% of revenue)	-0.637	-0.280	1.779***
	(1.036)	(1.220)	(0.560)
Fiscal balance	-0.683	-1.361*	-0.563
	(0.413)	(0.691)	(0.357)
Number of obs	267	297	267
R squared	0.854	0.783	0.868

This table presents the results of the regression with a combined set of independent variables for only high-income countries, as defined by the World Bank's classification. Columns (1)–(3) show the results using an OLS. Variable definitions can be found in 2.1 Standard errors are clustered at country level and all variables are standardised for ease of interpretation.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01.

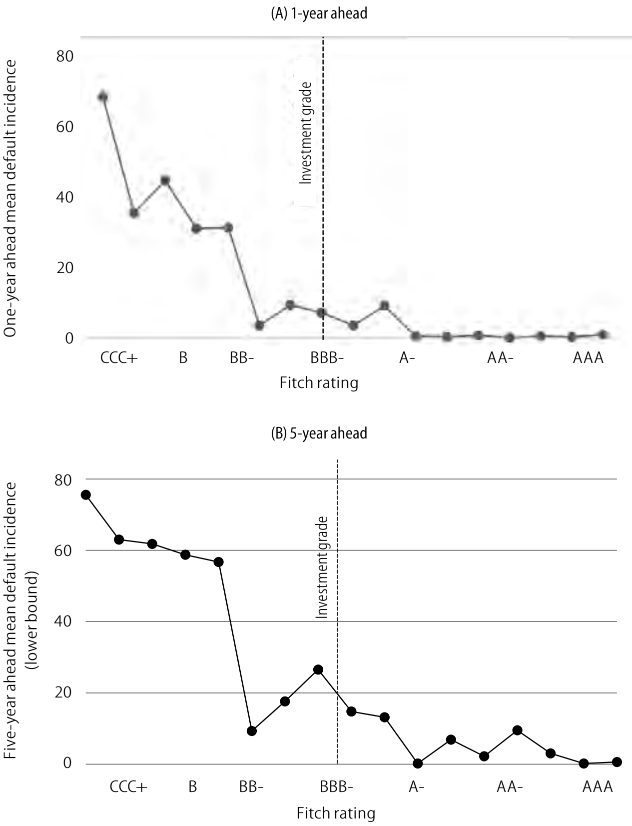
Table 2: Common Regression Framework

	OLS			Ordered Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
	S&P	Moody's	Fitch	S&P	Moody's	Fitch
Panel A: Baseline						
Governance indicators	2.731***	2.744***	2.317***	1.418***	1.230***	1.169***
	(0.362)	(0.388)	(0.427)	(0.231)	(0.209)	(0.284)
Ln(GDP per capita)	0.979**	1.339***	1.319**	0.446**	0.573***	0.775***
	(0.385)	(0.424)	(0.507)	(0.191)	(0.185)	(0.248)
Ln(broad money [% of GDP])	1.193***	0.914**	1.328***	0.562***	0.408**	0.712***
	(0.370)	(0.370)	(0.411)	(0.176)	(0.169)	(0.217)
Years since default	0.776***	0.902***	0.838***	0.403***	0.422***	0.433***
	(0.233)	(0.231)	(0.254)	(0.124)	(0.109)	(0.136)
General government debt to GDP	-0.839***	-1.009***	-0.706***	-0.504***	-0.526***	-0.458***
	(0.301)	(0.369)	(0.200)	(0.175)	(0.184)	(0.112)
Current account (% of GDP)	0.741***	0.865***	0.503**	0.479***	0.401***	0.372***
	(0.198)	(0.232)	(0.224)	(0.113)	(0.105)	(0.126)
GDP growth rate	0.559**	0.438*	0.314	0.386***	0.199*	0.239
	(0.248)	(0.246)	(0.263)	(0.131)	(0.110)	(0.156)
Inflation	-0.227	-0.346	-0.862	-0.278	-0.178	-0.684**
	(0.541)	(0.456)	(0.516)	(0.269)	(0.213)	(0.326)
Net FDI (% of GDP)	-0.218	-0.386	0.0581	-0.0372	-0.156	0.125
	(0.316)	(0.464)	(0.213)	(0.181)	(0.224)	(0.147)
Interest payments (% of revenue)	-0.287	-0.0319	-0.0461	-0.163	0.0155	-0.0124
	(0.210)	(0.260)	(0.240)	(0.113)	(0.119)	(0.127)
Fiscal balance	0.0121	-0.827	0.148	-0.226	-0.403	-0.157
	(0.639)	(0.635)	(0.612)	(0.358)	(0.301)	(0.370)
Number of obs	802	915	694	802	915	694
R^2	0.842	0.817	0.853			
Pseudo R^2				0.318	0.286	0.329

	OLS			Ordered Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
	S&P	Moody's	Fitch	S&P	Moody's	Fitch
Panel B: With country fixed effects						
Governance indicators	2.737***	2.748***	2.323***	1.422***	1.233***	1.173***
	(0.361)	(0.387)	(0.426)	(0.232)	(0.209)	(0.285)
Ln (GDP per capita)	0.963**	1.330***	1.308**	0.438**	0.570***	0.771***
	(0.385)	(0.423)	(0.506)	(0.191)	(0.184)	(0.247)
Ln(broad money [% of GDP])	1.193***	0.913**	1.327***	0.563***	0.409**	0.715***
	(0.371)	(0.370)	(0.411)	(0.177)	(0.169)	(0.218)
Years since default	0.784***	0.906***	0.841***	0.408***	0.424***	0.435***
	(0.232)	(0.230)	(0.252)	(0.124)	(0.109)	(0.136)
General government debt to GDP	-0.848***	-1.014***	-0.709***	-0.510***	-0.530***	-0.461***
	(0.302)	(0.369)	(0.198)	(0.176)	(0.185)	(0.112)
Current account (% of GDP)	0.751***	0.870***	0.509**	0.484***	0.404***	0.375***
	(0.199)	(0.232)	(0.223)	(0.114)	(0.105)	(0.125)
GDP growth rate	0.547**	0.434*	0.310	0.378***	0.198*	0.236
	(0.243)	(0.244)	(0.260)	(0.128)	(0.109)	(0.153)
Inflation	-0.232	-0.348	-0.868*	-0.281	-0.178	-0.690**
	(0.539)	(0.456)	(0.515)	(0.268)	(0.213)	(0.325)
Net FDI (% of GDP)	-0.215	-0.385	0.0605	-0.0356	-0.156	0.127
	(0.316)	(0.464)	(0.213)	(0.182)	(0.224)	(0.148)
Interest payments (% of revenue)	-0.273	-0.0226	-0.0381	-0.154	0.0200	-0.00785
	(0.209)	(0.257)	(0.231)	(0.113)	(0.118)	(0.122)
Fiscal balance	0.0000177	-0.835	0.142	-0.232	-0.407	-0.160
	(0.637)	(0.633)	(0.611)	(0.356)	(0.300)	(0.369)
Number of obs	805	918	698	805	918	698
R^2	0.842	0.817	0.853			
Pseudo R^2				0.318	0.286	0.329

This table presents the results of the regression with a combined set of independent variables. Columns (1)–(3) show the results using an OLS and columns (4)–(6) show the results using an ordered probit. Panel A shows the baseline regressions. In Panel B, county fixed effects are also included. Variable definitions can be found in Table 1. Standard errors are clustered at country-level and all variables are standardised for ease of interpretation.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01.

Figure 2: Default Incidence against Fitch Ratings



This graphs plots Fitch ratings against one-year and five-year ahead default incidence for all countries. The dotted line shows the line above which all countries are considered as investment grade.

period 1960–2018 and merge it with cross-country credit ratings data to analyse how well ratings predict sovereign default. We first plot data for the one-year and five-year ahead incidence of default, against levels of Fitch's sovereign credit ratings. Next, we analyse how well rating levels as well as rating actions predict default during normal times and in periods of crisis.

Default Prediction in the Near and Medium Term

As a preliminary analysis, we first study how well sovereign credit rating levels and changes explain near- and long-term incidence of sovereign default. While rating agencies also offer short-term sovereign ratings, they are mapped from the long-term issuer default ratings (Fitch 2020). In fact, short- ratings have much less granularity for corresponding rating levels, and any discrepancies—different short- and long-term ratings are rare. In Figure 2(A), we plot the Fitch rating levels against the number of incidences of one-year ahead defaults against each rating bracket. We see that while sovereigns with Fitch sovereign credit ratings above the investment grade level of BBB- see low or zero default incidence in the near-term, default rates of sovereigns with ratings below the investment grade increase non-linearly. One-year ahead default incidence jumps to 31% at BB-, 44% at B and more than 68% at ratings below CCC+. Sovereigns with ratings above A- see no one-year ahead defaults or debt-restructuring events. Figure 2(B) points to a similar pattern for medium-term default incidence. It plots Fitch credit ratings against five-year ahead incidence of default, which is an indicator for whether a sovereign defaults at least once in the next five years. We calculate whether a sovereign ever defaulted within the five-year ahead period from the time a rating is assigned. Default incidence levels are higher in comparison to the one-year ahead counterpart scenarios. Even though the increase in default incidence is drastic for sovereigns rated below BB-, default incidence is not monotonically decreasing for higher ratings in the five-year term. This simple analysis indicates that Fitch ratings are not accurate in estimating the size of default probabilities, especially for countries close to the investment-grade threshold. The bunching of defaults at the investment-grade threshold also indicates that once a country drops below the investment-grade threshold, credit rating agencies are reluctant to give an investment-grade rating (and vice versa). This simple finding motivates our next analysis and we examine how rating agencies perform in predicting default, comparatively and inter-temporally, during crisis and non-crisis periods.

Default Prediction during Periods of Crisis

How good are rating changes in predicting near- and long-term sovereign default? Rating downgrades, especially during crises periods, can have significant cliff or herding effects. These effects could be more pronounced for countries close to the investment-grade rating threshold. In addition, credit rating downgrades, even if triggered by systematic

Table 4: Common Regression—Low/Medium Income

	OLS		
	(1)	(2)	(3)
	S&P	Moody's	Fitch
Governance indicators	1.664***	1.863***	0.423
	(0.578)	(0.584)	(0.582)
Ln(GDP per capita)	0.527	0.621	0.888*
	(0.394)	(0.410)	(0.449)
Ln(broad money [% of GDP])	1.265***	1.187***	1.585***
	(0.418)	(0.432)	(0.428)
Years since default	0.776***	0.909***	0.823***
	(0.250)	(0.247)	(0.233)
General government debt to GDP	-1.795***	-2.164***	-1.748***
	(0.334)	(0.360)	(0.337)
Current account (% of GDP)	0.534***	0.506*	0.272
	(0.194)	(0.270)	(0.269)
GDP growth rate	0.558**	0.409*	0.314
	(0.228)	(0.231)	(0.239)
Inflation	-0.290	-0.289	-1.028**
	(0.546)	(0.451)	(0.434)
Net FDI (% of GDP)	0.136	-0.611	0.642
	(0.649)	(0.729)	(0.669)
Interest payments (% of revenue)	0.360	0.691***	0.658**
	(0.220)	(0.212)	(0.248)
Fiscal balance	0.679	0.00211	0.950
	(0.707)	(0.752)	(0.726)
Number of obs	538	621	431
R squared	0.652	0.629	0.670

This table presents the results of the regression with a combined set of independent variables for only low/medium-income countries, as defined by the World Bank's classification. Columns (1)–(3) show the results using an OLS. Variable definitions can be found in 2.1 Standard errors are clustered at country level and all variables are standardised for ease of interpretation.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01.

biases or arbitrariness, can trigger self-fulfilling prophecies, driving even relatively healthy countries to default (Gärtner et al 2011).

This motivates us to study how well do sovereign credit rating changes predict defaults, especially during crisis periods. We use a simple regression of near- and medium-term default incidence against the levels of credit rating as well as changes in credit rating for three sub-samples representing a period of no crisis (2000–06), the GFC (2007–09) and the European sovereign debt crisis (2010–14). We use the following specification:

$$Y_{it} = \alpha + \beta \times X_{it} + e_{it} \quad \dots (2)$$

where Y_{it} is the default indicator. For the five-year ahead regressions, Y_{it} takes the value of 1 if there was at least 1 default in the next five years for that country. Similarly, for the one-year ahead regressions, Y_{it} takes the value of 1 if there was at least 1 default in the next one year for that country, and is 0 otherwise. X is the level of credit rating or change in credit ratings.[3] The baseline is estimated for the full sample of countries. We also separately look at the sub-sample of high- and low/middle-income countries, as defined using the World Bank classification.

Table 5 presents the results for the level of rating against the one-year and five-year ahead default incidence, respectively, for each rating agency. We see that all rating agencies have a negative and significant coefficient pointing to the inverse relation between rating levels and incidence of default. Panel A examines near-term one-year ahead defaults. Sovereign rating-levels across agencies perform very similarly during the non-crisis and GFC in predicting one-year ahead defaults. A one notch lower rating is correlated with a 2% higher default incidence during the non-crisis period and between 1% and 4% higher one-year ahead default probability during the financial crisis and the EU sovereign debt crisis.

For the five-year ahead term, as seen in Panel B in Table 5, the coefficients on level of ratings across rating agencies are similar in scale for each corresponding period. While in non-crisis periods, one-unit lower rating is correlated with a probability of default incidence in the next five years by 2%, during the financial crisis and the sovereign debt crisis a reduction in ratings predicts five-year default incidence between 4% and 7%.

As the factors driving sovereign default change along the non-crisis and crisis periods, we analyse how well do rating actions such as downgrades, rather than rating levels, predict sovereign defaults? We regress, the near- and medium-term incidence of default on change in credit ratings, to account for rating actions. We keep the sub-sample divisions across the different time-periods as mentioned previously, and further subdivide the samples into high- and non-high (low/middle)-income economies.

Table 6 (p 100) represents the results for one-year ahead default incidence for all countries (Panels A and B) and the sub-sample of high- and low-middle-income countries as classified by the World Bank. Panel A shows that almost all agencies have statistically insignificant coefficients during both crisis periods, except Moody's during the 2007–09 crisis. Panel C points that S&P performs relatively worse for low- and middle-income sovereigns while Panel B suggests that S&P outperforms the other CRAs during the most recent crisis between 2010 and 2014 for high-income countries. Note, however, given the limited rate changes across CRAs for advanced economies, some of the coefficients for the sub-sample periods cannot be estimated.

Table 7 (p 101) represents the results for the five-year ahead default. For the full sample in Panel A, the coefficient on Moody's is not significant for both crisis periods which indicate that this rating agency performs relatively poorly in predicting

Table 5: Rating Levels as Predictors of Sovereign Default

	All	2000–06	2007–09	2010–14	All	2000–06	2007–09	2010–14	All	2000–06	2007–09	2010–14
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: one-year ahead default												
S&P	-0.0284***	-0.0108***	-0.0129***	-0.0386***								
	(0.00144)	(0.00223)	(0.00396)	(0.00372)								
Moody's					-0.0269***	-0.0156***	-0.0136***	-0.0364***				
					(0.00139)	(0.00237)	(0.00390)	(0.00358)				
Fitch									-0.0327***	-0.0141***	-0.0137***	-0.0437***
									(0.00153)	(0.00199)	(0.00335)	(0.00356)
Number of obs	1927	431	167	471	1907	442	161	431	1718	427	172	449
R squared	0.167	0.0517	0.0603	0.187	0.165	0.0894	0.0714	0.194	0.210	0.106	0.0900	0.252
Panel B: five-year ahead default												
S&P	-0.0391***	-0.0165***	-0.0465***	-0.0691***								
	(0.00183)	(0.00251)	(0.00651)	(0.00391)								
Moody's					-0.0379***	-0.0220***	-0.0389***	-0.0613***				
					(0.00174)	(0.00269)	(0.00614)	(0.00398)				
Fitch									-0.0432***	-0.0206***	-0.0530***	-0.0741***
									(0.00191)	(0.00241)	(0.00577)	(0.00346)
Number of obs	1533	431	167	374	1559	442	161	344	1332	427	172	356
R squared	0.230	0.0916	0.236	0.456	0.233	0.132	0.201	0.409	0.278	0.147	0.332	0.565

This table presents the results of the OLS regression of default incidence against rating levels for all the three CRA for the full sample (columns 1, 5, and 9) and for the sub-sample periods. 2007–09 represents the global financial crisis and 2010–14 represents the sovereign debt crisis. 2000–06 is the non-crisis period. The dependent variable is the default indicator. For the five-year ahead regressions (Panel A), it takes the value of 1 if there was at least 1 default in the next five years for that country. Similarly, for the one-year ahead regressions (Panels B), it takes the value of 1 if there was at least one default in the next one year for that country, and is 0 otherwise. Standard errors are clustered at the country level.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01.

medium-term default compared to its competitors. All rating agencies have an insignificant coefficient during the sovereign debt crisis. Panel B shows that only Fitch ratings have a significant coefficient against the five-year ahead incidence of default during the GFC for high-income countries. S&P and Moody's reliance on qualitative factors likely explains the inability to predict sovereign default during crisis periods. All rating agencies have negative and significant coefficients during the sovereign debt crisis for high-income countries. Panel C suggests that all rating agencies do poorly in predicting the five-year ahead default among low- and middle-income countries, especially during periods of crises.

Supervised Learning Models for Predicting Default

While a linear default prediction model based on economic and financial fundamentals is easy to interpret, the linear functional form assumption for the relationship between the incidence of default and the predictors may not reflect reality. Figure 2 suggests that the default incidence remains flat for high rating levels and increases non-linearly as ratings go below investment-grade thresholds. These predictions often have low statistical power due to the lack of default incidences across time in the cross-section of countries. The prediction is based on a limited set of conventional economic and financial variables, which may not sufficiently explain variation in default and restructuring events across time. Therefore, we analyse the set of 10 existing predictors used earlier with three new factors as proxies for financial repression, banking system vulnerability, and regional economic integration, which can also influence sovereign default probability.

Additional Factors Influencing Default

Financial repression takes place when the government, through covert duress or overt policy action, forces banks to hold government debt. Chari et al (2020) argue that the government's willingness to repay debt endogenously and credibly increases when it issues debt without commitment. This deters sovereigns from defaulting on their debt and thus acts as a credible commitment device. Perez et al (2015) note that default deterrence can originate from the banks' balance sheet as sovereign defaults reduce their ability to raise funding and lend to productive investments. Further, defaults undermine the liquidity available at the banks as treasury securities get replaced by less productive investments. The effect, if present, shall be more pronounced when the banking sector is more vulnerable to defaults as riskier banks can engage in risk-shifting

Table 6: Rating Changes as Predictor of One-year Ahead Sovereign Default

	All	2000–06	2007–09	2010–14	All	2000–06	2007–09	2010–14	All	2000–06	2007–09	2010–14
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All countries												
Change in Standard and Poor ratings	-0.0594*** (0.0112)	-0.0669*** (0.0190)	-0.0585 (0.0399)	-0.0247 (0.0220)								
Change in Moody's ratings					-0.0512*** (0.00922)	-0.0929*** (0.0167)	-0.130*** (0.0372)	-0.0184 (0.0161)				
Change in Fitch ratings									-0.0522*** (0.0118)	-0.0911*** (0.0169)	-0.0256 (0.0275)	-0.0132 (0.0244)
Number of obs	1,850	417	162	458	1,849	436	160	425	1,619	401	172	442
R squared	0.0151	0.0291	0.0133	0.00277	0.0165	0.0666	0.0715	0.00307	0.0119	0.0676	0.00508	0.000660
Panel B: High-income countries												
Change in Standard and Poor ratings	-0.0298*** (0.00518)			-0.0393*** (0.0131)								
Change in Moody's ratings					-0.0253*** (0.00483)	-0.0635*** (0.00565)		-0.0177* (0.01000)				
Change in Fitch ratings									-0.0189*** (0.00575)	-0.0764*** (0.00685)		-0.0178 (0.0165)
Number of obs	858	198	66	171	856	198	66	165	723	194	66	169
R squared	0.0371			0.0507	0.0310	0.392		0.0188	0.0147	0.394		0.00692
Panel C: Low/middle-income countries												
change in Standard and Poor ratings	-0.0884*** (0.0192)	-0.0974*** (0.0292)	-0.0554 (0.0558)	-0.0878* (0.0465)								
Change in Moody's ratings					-0.0851*** (0.0169)	-0.104*** (0.0268)	-0.164*** (0.0525)	-0.123*** (0.0414)				
Change in Fitch ratings									-0.105*** (0.0216)	-0.103*** (0.0290)	-0.0340 (0.0399)	-0.116** (0.0510)
Number of obs	992	219	96	287	993	238	94	260	880	202	104	268
R squared	0.0209	0.0489	0.0104	0.0124	0.0249	0.0595	0.0956	0.0330	0.0263	0.0591	0.00708	0.0191

This table presents the results of the OLS regression of one-year ahead default incidence against change in ratings for all the three CRA for the full sample (columns 1, 5, and 9) and for the sub-sample periods in the remaining columns. 2007–09 represents the global financial crisis and 2010–14 represents the sovereign debt crisis. 2000–06 is the non-crisis period. Panel A for the full sample, Panel B is for the high- and low/middle-income country samples, respectively. High- and low/middle- income sample is from the World Bank classification. When sample sizes are limited due to lack of variation in the data, coefficients have been left blank (Panel B). The dependent variable is the default indicator for the one-year ahead regressions and takes the value of 1 if there was at least 1 default in the next one year for that country, and 0 otherwise. Standard errors are clustered at the country level.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01.

behaviour to re-cover expected losses from future sovereign default. This hence raises the *ex post* cost of defaulting and can help reduce the probability of default. We proxy these factors by the proportion of credit directed towards the government sector within a country in each year and the country's Bank-z score which indicates banking system vulnerability as the ratio of the country's banking system capitalisation and return on assets to the volatility of those returns.

Regional economic integration also engenders a commitment device as trading partners fear the negative spillovers due to sovereign default from each other (Eberhardt 2018). Second, regional trading agreements also implicitly incorporate favourable financing measures and/or fiscal and economic goals enabling fruitful integration. These factors lead to the prior that more integrated countries should have lesser defaults. We quantify this result, as per Eberhardt (2018), by measuring the number of regional trade agreements a country is part of, for each year.

As a consequence, not only does the ability of the new and existing predictors need to be tested, we also seek to know how many and which predictors can explain most of the variation in predicting defaults. We therefore use a supervised learning framework to predict the incidence of one-year and five-year ahead defaults with the set of 13 predictors.

Random Forest Methodology

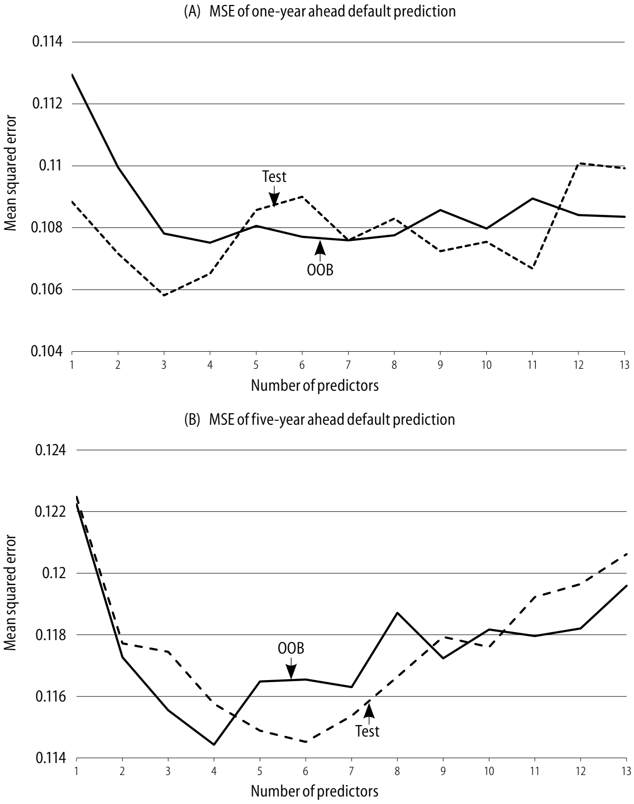
At the core of our supervised machine learning methods are regression trees that allow us to sequentially and randomly stratify the predictor space. This enables us to delineate (i) which predictors reduce the residual sum of squares the most when they are sequentially included in the prediction regression and (ii) how many predictors reduce the test error of the predictions, without substantially increasing the bias induced by their inclusion. The learning design picks a subset of all country-year observations and trains the data by running several iterations of the regression tree technique on that subset. Subsequently, we cross-validate the accuracy of those predictions with a test data, which is the data outside the training subset to obtain the mean squared error of the predictions. In all, this enables us to pin down the mean-squared-error-minimising predictors which have the highest relative influence in predicting the incidence of default. In order to achieve this, we use a random forest technique which randomly chooses a set of four predictors (the closest to the

Table 7: Rating Changes as Predictor of Five-year Ahead Sovereign Default

	All	2000–06	2007–09	2010–14	All	2000–06	2007–09	2010–14	All	2000–06	2007–09	2010–14
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All countries												
Change in Standard and Poor ratings	-0.0864***	-0.0745***	-0.146**	-0.0400								
	(0.0142)	(0.0219)	(0.0711)	(0.0272)								
Change in Moody's ratings					-0.0650***	-0.104***	-0.0581	-0.0176				
					(0.0115)	(0.0196)	(0.0650)	(0.0196)				
Change in Fitch ratings									-0.0783***	-0.0946***	-0.123**	-0.0175
									(0.0143)	(0.0208)	(0.0547)	(0.0288)
Number of obs	1,460	417	162	362	1,501	436	160	338	1,242	401	172	352
R squared	0.0249	0.0272	0.0257	0.00597	0.0207	0.0603	0.00503	0.00240	0.0236	0.0495	0.0290	0.00105
Panel B: High-income countries												
Change in Standard and Poor ratings	-0.0955***		-0.0878	-0.134***								
	(0.00897)		(0.129)	(0.0190)								
Change in Moody's ratings					-0.0593***	-0.0650***	-0.0335	-0.0520***				
					(0.00788)	(0.00910)	(0.106)	(0.0159)				
Change in Fitch ratings									-0.0671***	-0.0783***	-0.313***	-0.0832***
									(0.0100)	(0.0110)	(0.0727)	(0.0233)
Number of obs	718	198	66	136	724	198	66	132	587	194	66	135
R squared	0.137		0.00721	0.272	0.0727	0.207	0.00156	0.0757	0.0710	0.208	0.225	0.0874
Panel C: Low/middle-income countries												
Change in Standard and Poor ratings	-0.0973***	-0.110***	-0.113	-0.0718								
	(0.0235)	(0.0332)	(0.0807)	(0.0495)								
Change in Moody's ratings					-0.0838***	-0.117***	-0.0606	-0.0781*				
					(0.0208)	(0.0307)	(0.0773)	(0.0447)				
Change in Fitch ratings									-0.125***	-0.109***	-0.0618	-0.0962*
									(0.0254)	(0.0347)	(0.0638)	(0.0505)
Number of obs	742	219	96	226	777	238	94	206	643	202	104	213
R squared	0.0226	0.0484	0.0202	0.00930	0.0205	0.0583	0.00664	0.0147	0.0362	0.0468	0.00912	0.0169

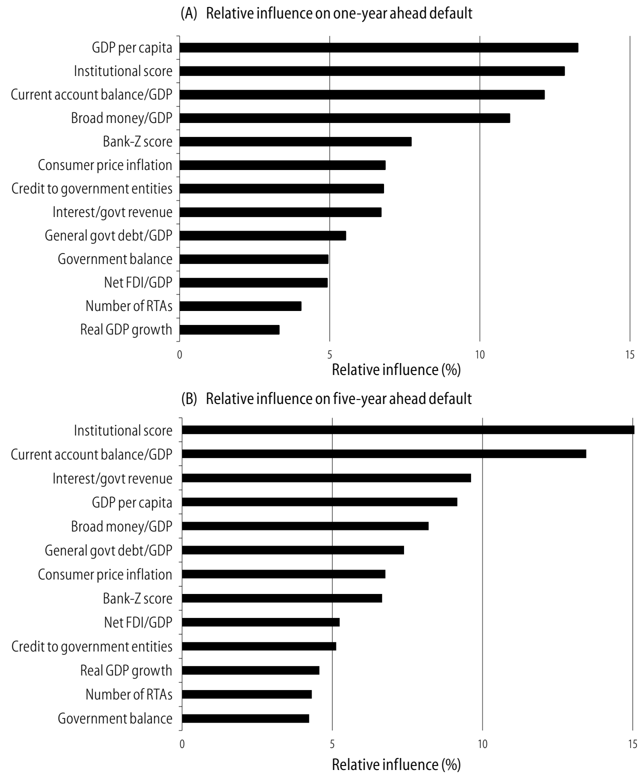
This table presents the results of the OLS regression of five-year ahead default incidence against change in ratings for all the three CRA for the full sample (columns 1, 5, and 9) and for the sub-sample periods in the remaining columns. 2007–09 represents the global financial crisis and 2010–14 represents the sovereign debt crisis. 2000–06 is the non-crisis period Panel A for the full sample, Panel B is for the high- and low/middle-income country samples, respectively. High- and low/middle- income sample is from the World Bank classification. The dependent variable is the default indicator for the five-year ahead regressions and takes the value of 1 if there was at least 1 default in the next five years for that country, and 0 otherwise. Standard errors are clustered at the country level.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01.

Figure 3: Random Forests—Mean Squared Errors of Predictors



This graph plots the mean squared errors of the one-year and five-year ahead default predictions. The black line is the out of bag error and the dotted line is the test error.

Figure 4: Relative Influence of Factors in Predicting Incidence of Default



square-root of the total number of predictors) for each iteration of a default-predicting regression tree. It generates thousand trees randomly and thus bootstrap-aggregates the predictions from each regression tree. This also helps us to evade statistical power issues in our predictions, by randomly simulating the combinations of training and testing data sets in each iteration.

Figures 3(A) and 3(B) represent the test and out-of-bag (OOB) error[4] for the number of predictors included in the random forest model of one-year and five-year ahead default prediction respectively.

Results of Random Forest Model

The random forest model only explains 29.81% and 45.1% of variance in one-year and five-year ahead default incidences. This is in stark contrast to the linear regression framework.

Figure 4(A) explains the relative influence of predictors in explaining the one-year ahead sovereign default incidence. Relative influence is calculated as the amount of RSS reduction due to splits over a given predictor, averaged over all bootstrapped trees. A large value indicates an important predictor. GDP per capita, institutional score, current account balance and broad money explain more than 10% of RSS reduction in predicting short-term default. However, interestingly Bank-z score and credit to government entities are the fifth and seventh most important factors after them. This aligns with the hypothesis that financial repression is an important channel tying sovereign default and the banking system. Historical experience also suggests that sovereign default crises are followed by a banking crisis which reinforces the theory that government's issuance of debt and willingness to pay takes into consideration the health of the banking sector.

Economic integration does not have a very significant impact on predicting defaults—both near term as well as five-year ahead. Figure 4(B) gives a similar picture for five-year ahead default. However, the precedence at the top changes: institutional fundamentals and external sector strength take precedence over banking system vulnerabilities and financial repression. Real GDP growth, often a feature of emerging economies, has a low relative importance across the two default variables. Therefore, one can infer from the random forest exercise that: (i) the existing set of common predictors used by CRAs explain much less of the variation in default incidence, when adjusted for data-fitting bias; (ii) 4–6 set of predictors can minimise the standard error of predictions as opposed to 13 (or even more if one accounts for all the factors that enter into CRA rating models); and (iii) while prosperity, external sector and institutional fundamentals play an important role in predicting default, financial repression and banking sector risks are more important variables explaining near-term default than general government fiscal health. However, government fiscal health is a more important factor in explaining five-year ahead default.

Conclusions

In this paper, we motivate the need to reassess CRA rating methodologies and accuracy in predicting sovereign default. We narrow down to a parsimonious set of factors and find that these factors can explain a large proportion of the variation in ratings across time and countries. Across all models, we find that institutional quality is the most significant factor driving sovereign ratings. GDP growth does not influence sovereign ratings unless sustainable, and GDP growth that is fuelled by investment in unproductive sectors receives a negative weight. However, GDP per capita is an important determinant in some specifications, suggesting a negative methodological bias towards emerging economies. CRA ratings are better predictors of sovereign default for advanced economies but perform relatively poorly for low- and middle-income countries, especially countries near the minimum investment-grade rating threshold (such as India). Additionally, rating downgrades are poor indicators of subsequent sovereign default, especially for rating agencies such as Moody's that rely on more qualitative factors.

We assess the factors that influence default independent of ratings in a supervised learning framework and show that while the parsimonious set of factors have good explanatory power when fitted to past defaults, they are not very powerful in predicting future defaults. The conventional economic, fiscal, and external sector variables can explain less than 50% of one-year and five-year ahead default occurrences in the past 60 years. Also, adding more variables to the prediction—as CRA methodologies do—increases the predictions' bias.

A crisis like the COVID-19 pandemic has affected economies worldwide and debilitated demand and employment. In such times, government stimulus and relief measures require massive public debt-issuance.[5] Countries with weak institutional fundamentals suffer a twofold setback as economic contraction can accompany worsening sovereign debt sustainability. Rating downgrades at such a time, especially if this pushes a country's rating to below-investment-grade, will lead to negative feedback loops that can be devastating for the economy. Biased rating methodologies retrofitted on past experiences of developed economies should be used with caution, especially when they do not reflect true sovereign creditworthiness.

Our findings suggest that the over-reliance of market participants on CRA ratings to assess sovereign creditworthiness may be unwarranted, particularly during crisis periods. There has been a growing recognition that sovereign credit ratings of the major rating agencies are biased and dependence on CRA ratings need to reduce.[6] In 2010, a G20 resolution acknowledged the overdependence on the CRAs and suggested that central banks and banks independently conduct their own ratings (Financial Stability Board 2010a).[7] Motivated by above, the Bank of Canada produces internal sovereign credit ratings for use in its management of Canada's foreign exchange reserves.[8] Our paper makes a case for India, too, building alternative internal rating models to assess sovereign credit risk instead of relying on the CRAs.

NOTES

1 See https://timesofindia.indiatimes.com/business/india-business/india-not-alone-to-get-moodys-downgrade-tag/articleshow/76166388.cms.

2 See https://www.livemint.com/news/india/fitch-ratings-downgrades-india-outlook-from-stable-to-negative-11592513114717.html.

3 We do not include country-fixed effects since we are interested in how ratings across countries determine default rates. Results are robust to including country-fixed effects, but not shown in the interest of brevity. Results are available on request.

4 The leave-one-out cross validation error from bootstrap aggregation. Trees are repeatedly fit to boot-strapped subsets of the observations. One can show that on average, each bagged tree makes use of around two-thirds of the observations. The remaining one-third of the observations not used to fit a given bagged tree are referred to as the out-of-bag (OOB) observations. We can predict the response for the ith observation using each of the trees in which that observation was OOB. This will yield around one-third of the number of sample predictions for the ith observation, which we average works with high r-squares reflecting the bias in the data. This means that these 13 fundamentals are not very powerful in predicting future default, but fit well with default data *ex post*.

5 Public debt to GDP ratio has surpassed 140% in the US and is expected to touch 90% in India in the next fiscal year (World Economic Forum 2020).

6 The Financial Stability Board (formed after the 2008 financial crisis) says "Reducing reliance [on CRAs] in this way will reduce the financial stability-threatening herding and cliff effects that currently arise from CRA rating thresholds being hard-wired into laws, regulations and market practices" (Financial Stability Board 2010b).

7 Motivated by the large CRA bias towards certain countries, Prime Minister Narendra Modi and President Vladimir Putin, in 2017, discussed developing an independent credit rating agency (*Livemint* 2017).

8 BoC makes the methodology (but not the actual ratings) publicly available (Muller and Bourque 2017).

REFERENCES

Adelino, Manuel and Miguel A Ferreira (2016): "Bank Ratings and Lending Supply: Evidence from Sovereign Downgrades," *Review of Financial Studies,* Vol 29, No 7, pp 1709–46.

Aizenman, Joshua, Mahir Binici and Michael Hutchison (2013): "Credit Ratings and the Pricing of Sovereign Debt During the Euro Crisis," *Oxford Review of Economic Policy,* Vol 29, No 3, pp 582–609.

Almeida, Heitor, Igor Cunha, Miguel A Ferreira and Felipe Restrepo (2017): "The Real Effects of Credit Ratings: The Sovereign Ceiling Channel," *Journal of Finance,* Vol 72, No 1, pp 249–90.

Beers, David and Patrisha de Leon-Manlagnit (2019): "The BoC-BoE Sovereign Default Database: What's New in 2019?"

Chari, V V, Dovis Alessandro and J Kehoe Patrick (2020): "On the Optimality of Financial Repression," *Journal of Political Economy,* Vol 128, No 2, pp 710–39.

Eberhardt, Markus (2018): "(At Least) Four Theories for Sovereign Default."

Financial Stability Board (2010a): "Overview of Progress in the Implementation of the G20 Recommendations for Strengthening Financial Stability, Report of the Financial Stability Board to G20 Leaders."

— (2010b): "Principles for Reducing Reliance on CRA Ratings."

Fitch (2020): "Sovereign Ratings Criteria: Master Criteria."

Fuchs, Andreas and Kai Gehring (2017): "The Home Bias in Sovereign Ratings," *Journal of the European Economic Association,* Vol 15, No 6, pp 1386–423.

Gärtner, Manfred, Björn Griesbach and Florian Jung (2011): "PIGS or Lambs? The European Sovereign Debt Crisis and the Role of Rating Agencies," *International Advances in Economic Research,* Vol 17, No 3, pp 288.

Livemint (2017): "Modi, Putin Agree to Develop 'Independent' Credit Rating Industry."

Moody's (2019): "Sovereign Ratings Methodology."

Muller, Philippe and Jérôme Bourque (2017): "Methodology for Assigning Credit Ratings to Sovereigns," Bank of Canada.

Perez, Diego et al (2015): "Sovereign Debt, Domestic Banks and the Provision of Public Liquidity," Manuscript, New York University 3:14.

Reinhart, Carmen M and Kenneth S Rogoff (2009): *This Time is Different: Eight Centuries of Financial Folly,* Princeton University Press.

S&P (2017): "Sovereign Ratings Methodology."

World Economic Forum (2020): "Countries Are Piling Up Record Amounts of Debt Amid COVID-19."

Table A1: List of Countries

Albania	Andorra	Angola	Argentina	Armenia	Aruba
Australia	Austria	Azerbaijan	Bahamas	Bahrain	Bangladesh
Barbados	Belarus	Belgium	Belize	Benin	Bermuda
Bolivia	Bosnia and Herzegovina	Botswana	Brazil	Bulgaria	Burkina Faso
Cabo Verde	Cambodia	Cameroon	Canada	Cape Verde	Chile
China	Taiwan	Colombia	Congo Rep	Costa Rica	Croatia
Cuba	Cyprus	Czech Republic	Denmark	Dominican Republic	Ecuador
Egypt	El Salvador	Estonia	Ethiopia	Fiji	Finland
France	Gabon	Gambia	Georgia	Germany	Ghana
Greece	Guatemala	Honduras	Hong Kong	Hungary	Iceland
India	Indonesia	Iran	Iraq	Ireland	Isle of Man
Israel	Italy	Jamaica	Japan	Jordan	Kazakhstan
Kenya	Korea Rep	Kuwait	Latvia	Lebanon	Lesotho
Libya	Liechtenstein	Lithuania	Luxembourg	Macedonia	Malawi
Malaysia	Maldives	Mali	Malta	Mauritius	Mexico
Moldova	Mongolia	Montenegro	Morocco	Mozambique	Namibia
Netherlands	New Zealand	New Zealand	Nicaragua	Nigeria	North Macedonia
Norway	Oman	Pakistan	Panama	Papua New Guinea	Paraguay
Peru	Philippines	Poland	Portugal	Qatar	Romania
Russian Federation	Rwanda	San Marino	Saudi Arabia	Senegal	Serbia
Seychelles	Singapore	Slovak Republic	Slovenia	South Africa	Spain
Sri Lanka	Suriname	Sweden	Switzerland	Taiwan	Tajikistan
Thailand	Togo	Trinidad and Tobago	Tunisia	Turkey	Turkmenistan
Uganda	Ukraine	United Arab Emirates	United Kingdom	United States	Uruguay
Uzbekistan	Venezuela	Vietnam	Zambia		