Sovereign Credit Ratings: An Assessment of Methodologies and Rating Biases[*]

CAFRAL[†]

December 25, 2020

Abstract

We scrutinize and evaluate the rating methodologies of the big three credit rating agencies (CRAs): S&P, Moody's, and Fitch. We examine the factors that drive sovereign ratings, using a common regression framework, principal component analysis, and machine learning techniques with a panel of 162 countries covering ratings from 2000–2018. While all three CRAs employ complex rating methodologies based on qualitative and quantitative inputs, only a handful of variables can account for a significant proportion of the rating variation. 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. We also analyze CRA rating performance and show that CRA rating changes, especially during crisis periods, are poor predictors of sovereign defaults, particularly for CRAs that rely on more subjective information (e.g., Moody's). Finally, using machine learning techniques, we show that while the parsimonious factors in the baseline analysis have good explanatory power when retro-fitted to past defaults, they are poor predictors of *future* defaults. Our findings suggest that the over-reliance of market participants on CRA ratings to assess sovereign creditworthiness may be unwarranted, particularly during crisis periods.

JEL Classification: G24, G28, G31, G32, H63

Keywords: Credit Ratings, Sovereign Default

[*]All the opinions expressed in this paper reflect those of the authors and not necessarily those of CAFRAL or the Reserve Bank of India.

[†]Authors: Nirupama Kulkarni, Kaushalendra Kishore, Gautham Udupa, Ilisa Goenka, Kavya Ravindranath, Rahul Singh Chauhan. We thank Sitikantha Pattanaik (Reserve Bank of India) for invaluable guidance and support during this project. We also thank Harendra Kumar Behera (Reserve Bank of India) for providing us data and for helpful discussions. We gratefully acknowledge Sangeetha Matthews (Reserve Bank of India) and Silu Muduli (Reserve Bank of India) for useful comments.

1 Introduction

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

Despite these downgrades, there has been a 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, 2010*a*,*b*). 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 downgraded. Rating downgrades may also lead to negative feedback loops, as rating downgrades can worsen economic conditions, leading to further downgrades (Aizenman, Binici and Hutchison, 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. Even such systematic biases and arbitrary factors in rating downgrades can trigger self-fulfilling prophecies, driving even relatively healthy countries to default (Gärtner, Griesbach and Jung, 2011).

[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

The above reasons underscore the need to study what factors 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 examine the quantitative and qualitative factors that drive individual CRA ratings. Second, we narrow down to a parsimonious set of factors and examine whether these parsimonious factors 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 (a) EMEs, and (b) rating downgrades during crises periods. Finally, we use Machine Learning techniques to narrow down to the variables that predict defaults and evaluate whether retro-fitting data to past defaults is a good predictor of future defaults.

The three CRAs use complex rating methodologies using 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 is 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 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 Fitch at 91%, 73% for S&P, and 80% for Moody's. 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. Remaining significant factors driving sovereign credit ratings include: macroeconomic health, measured in terms of fiscal balance, general government debt to GDP, CPI, and broad money to GDP for Fitch; GDP per capita, general government debt to GDP, and GDP per capita growth rate for S&P; unemployment rate, the current account to GDP, and general government debt to GDP for Moody's. Since this model captures only quantitative variables, it performs best for ratings based on more quantitative factors, such as for Fitch. 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.

We then compare actual ratings using the parsimonious model to predicted ratings, especially focusing on India. The predicted ratings for Fitch in our model is the closest to its actual ratings for India. S&P performs the second-best, and Moody's performs the worst. To explain the factors driving these differences, we then compare predicted ratings to actual ratings across the cross-section of countries in 2016. We find that for Fitch, our predicted ratings are very close to their actual ratings for both advanced and emerging market economies. On the other hand, our model for Moody's (and to a lesser extent S&P) accurately predicts advanced economies' ratings but under-predicts ratings for EMEs. We conjecture that Moody's relies more on qualitative adjustments for EMEs, whereas the ratings for advanced economies load more heavily on quantitative factors explaining the discrepancy between actual and predicted ratings. Together, the first step of the analysis highlights that the CRAs rely on more qualitative adjustments for the lower- and middle-income economies, but there are significant differences across the CRAs in the way these factors enter the models.

Next, to determine what variables are important *across* the three CRAs, we build a common model. We pick variables that are common to the three agencies and regress them against ratings. Thus, our independent variables here are the common set of regressors across all three CRAs. Again, we find that institutional health is the most critical variable for all three CRAs. A one standard deviation rise in the percentile rank of institutional quality — measured using the World Governance Indicators (WGI) —is associated with a 3-notch higher ratings for Fitch, S&P and Moody's. Other significant variables are GDP per capita, broad money, years since default, general government debt to GDP, and current account to GDP.

Results are similar using a Principal Component Analysis (PCA). The PCA helps us succinctly summarize the main pillars driving rating variation by reducing the dimensionality of the data. We divide all our variables into 4 pillars and run a PCA within each pillar. We retain the first component within each pillar, and regress on CRA ratings. Again, we find that the institutional pillar (comprising of the WGI indicators, the Gini index, and years since default) and the fiscal pillar (comprising general government debt to GDP, fiscal balance, and interest payments to revenue) are significant in affecting ratings. A one standard deviation higher institutional quality is associated with a three-notch higher rating, while a one standard deviation higher fiscal factor principal component is associated with a one-notch higher ratings. The one striking finding, across specifications, is that institutional quality is the single most important determinant of CRA ratings.

We now examine five factors in detail: Institutional quality, government debt to GDP, broad money, GDP per capita, and GDP growth rate. Across all models, a consistent theme is that institutional quality is a primary driver of CRA ratings. To examine whether this reliance on institutional quality is justified, we examine whether this factor is a good predictor of sovereign default in the short-term (one-year ahead default) and long-term (eight-year ahead default). We find that institutional quality is an important predictor for both short-term and long-term default, justifying the heavy reliance in CRA methodologies. Likely, strong institutions augur well for a sovereign's debt repayment capability; hence the CRAs place the most weight on this factor.

Government debt to GDP is also a significant factor driving ratings in the common model and the PCA analysis. Excess accumulation of government debt is a tax burden on future generations. Reinhart and Rogoff (2009) also stress that large debts can hurt countries in the short-run if the market thinks that the government will not be able to finance the debt in the long-term. We find that debt to GDP is important in predicting short-term default, and a one standard deviation rise in government debt to GDP is associated with a 9% higher incidence of default. Surprisingly, debt to GDP is less critical for predicting long-term default suggesting that while institutional quality is a more reliable measure of long-run debt repayment capacity, a high accumulation of debt likely signals that sovereign default is imminent near-term. Broad money to GDP, which serves as a proxy for financial intermediation in the economy, is important in predicting near-term default justifying its significance in some CRA models.

The fourth variable we examine is GDP per capita. GDP per capita is a significant variable in the common regression and the individual regression for S&P. In the past, commentators have highlighted the CRA over-reliance on GDP per capita, arguing that this unfairly disadvantages poorer countries (Government of India, 2017). GDP per capita is an essential factor for crises episodes, arguably, justifying their use in CRA methodology (Primo Braga and Vincelette, 2010). However, GDP per capita is not a significant variable determining either near- or long-term default. Additionally, even when we compare the default rate of India's peer countries (in terms of GDP per capita), India is an outlier with an impeccable default history. The years since default for India is nearly 35 years, compared to only 12 years on average for its peer countries.

GDP growth rate enters into credit ratings in a more complicated manner. Despite its anecdotal importance in ratings, we find that GDP growth is a noisy determinant of sovereign ratings. This is because the CRAs adjust for GDP growth in their rating methodologies based on whether they think the growth is sustainable. While for all three CRAs, GDP growth enters as a base variable to calculate the initial rating, all agencies make further adjustments. Fitch adjusts ratings upwards if the country performs well relative to its peers. S&P makes a negative adjustment to ratings if the unproductive household sector fuels GDP growth. Moody's adjusts the initial score based on GDP growth sustainability and takes into account factors such as female labor force participation, labor market laws, and export diversification.

Next, we analyze 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, Griesbach and Jung, 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-income and low/middle-income countries. Crisis periods refer to the global financial crisis of 2007–09 and the EU sovereign debt crisis of 2010–14. We find that in the sample of all countries, Moody's performs poorly in predicting sovereign defaults during both crises, while S&P and Fitch perform well. For high-income countries, only Fitch performs well during crises, whereas all three CRAs perform poorly during crises for low and middle-income countries. Together, these findings suggest that CRAs that rely less on subjective information (e.g. Fitch) are better able to predict sovereign defaults, especially during crisis periods. Further, 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 credit rating agencies along with three additional factors influencing sovereign default probability hypothesized in recent literature (Chari, Dovis and Kehoe, 2020; Eberhardt, 2018; Perez et al., 2015). We find that the set of predictors commonly used by the CRAs can only explain 29.81% of the variation in 1-year ahead default incidences, and 45% of the variation in 5-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 90% of the variation in some CRA ratings. The supervised learning analysis suggests that the existing CRA methodologies suffer from survivor-ship 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, optimizing 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 minimized by using only 4-6 predictors and *increases* as more predictors are included. Therefore, we find that selection biases and model complexity can dent the overall prediction accuracy of sovereign rating models in predicting near-term and long-term sovereign default incidence and calls for caution in relying exclusively on CRA ratings.

Our paper is organized as follows. Section 2 explains the motivation for this study. Section 3 provides a detailed summary of the CRA methodologies. Section 4 discusses the primary empirical exercise and Section 5 examines the main determinants of the CRA ratings in detail. Section 6 examines CRA rating performance by relating CRA ratings to sovereign defaults. Section 7 uses machine learning techniques to assess model suitability, and finally Section 8 concludes.

2 Motivating the Need to Reassess CRA Methodologies

In this section, we examine India's rating over the years and the recent rating downgrades by the Credit Rating Agencies (CRAs, henceforth). We then discuss rating biases as highlighted in previous literature and motivate the need to reassess CRA rating methodologies.

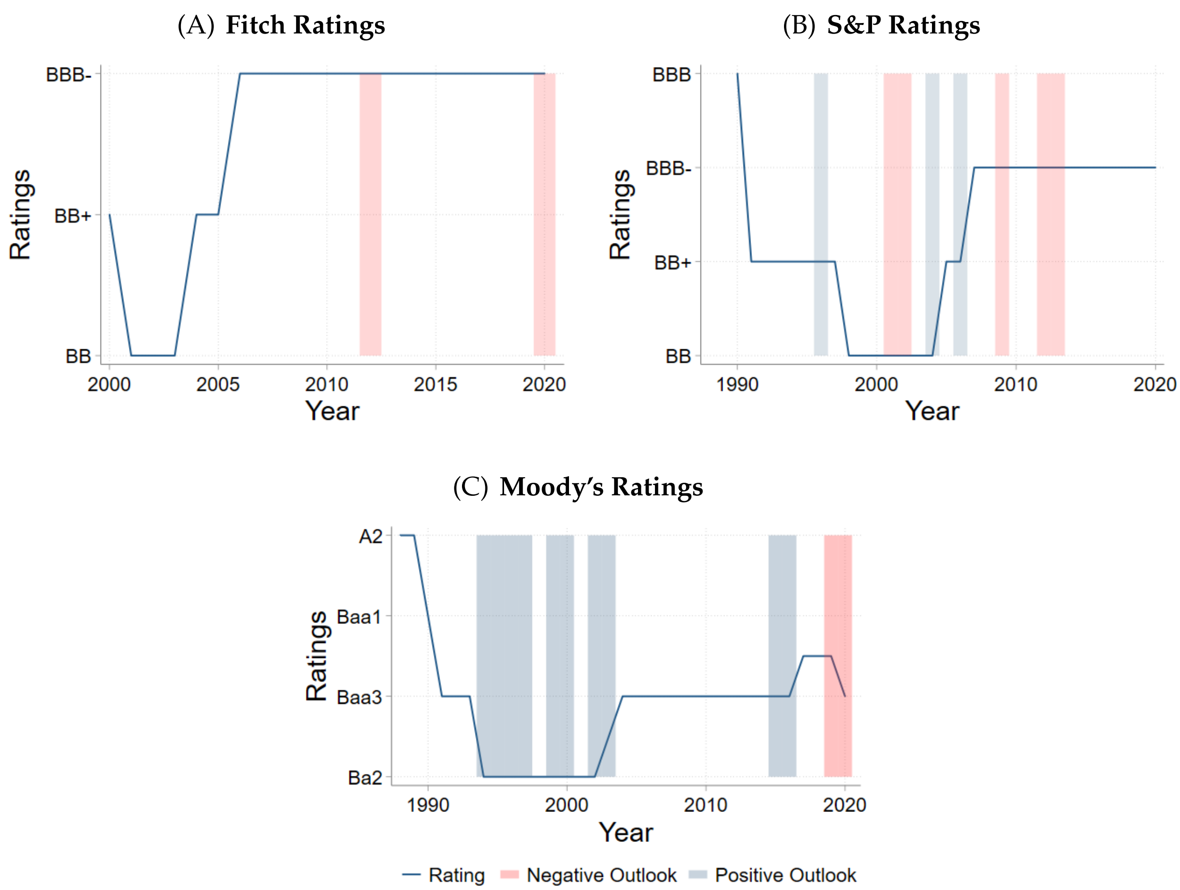
2.1 India's Ratings

Figure 1 shows India's ratings time-series for the three major rating agencies: Fitch, Standard & Poor and Moody's. Most rating agencies upgraded their rating outlook in the early 2000s and barring the year of 2019–2020, ratings have remained relatively stable.

The ongoing Covid-19 pandemic has resulted in countries adopting large fiscal stimulus packages and unconventional monetary measures to combat the economic fallout of the pandemic. This has raised questions of the limitations in fiscal capacity and the sustainability of debt of the emerging economies (EMEs). As a result, CRAs have down-

Figure 1: Time series of India's credit ratings

This figure presents a time series of the three rating agencies' sovereign ratings for India. The light red bars indicate a negative outlook and the light blue bars indicate a positive outlook, as assigned by the rating agency.

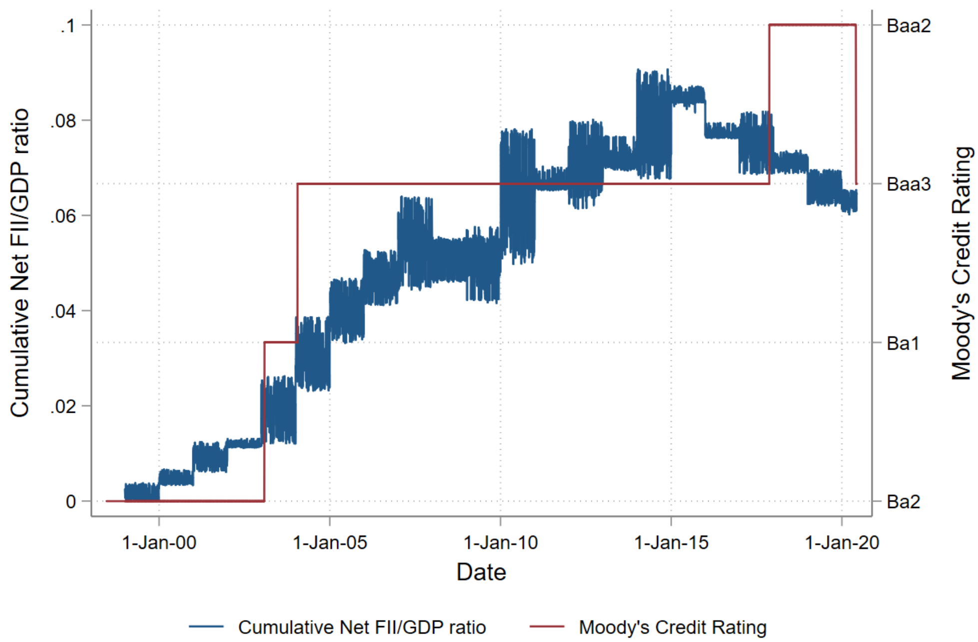


graded several EMEs, India included. Moody's downgraded India's rating from Baa2 (negative) to Baa3 (negative) on June 1st, 2020. It quoted the weak implementation of economic reforms since 2017, relatively low economic growth over a sustained period of time, deterioration in the fiscal position of union and states governments and lastly, rising stress in the financial sector (Moody's, 2020) as the reason for the downgrade. Fitch changed its outlook on India's rating from BBB- (stable) to (negative) on June 18th, 2020. In addition to the rationale for revision stated in Moody's (2020), Fitch (2020*a*) also noted the government's response to the pandemic, as well as growing geopolitical risks with China as major reasons for revisions in outlook.

This recent rating changes have raised concerns of India's vulnerability to downgrades. India is currently at the last notch of an investment-grade rating, making it especially vulnerable to rating downgrades since the cliff effects of dropping from investment to non-investment grade can potentially lead to large Foreign Institutional Investor (FII) outflows and a balance of payments crises. One reason for this is that India India's persistently low level of ratings, even before the COVID-19 pandemic. Commentators have attributed the low ratings to heavy reliance by the CRAs on per-capita GDP, arguably, biasing ratings downwards for lower middle income countries (Government of India, 2017). However, a low rating in the pre-pandemic era may not be justified given India's strong economic performance. Figure 2 highlights the lack of CRA rating upgrades for India, despite long periods of massive foreign investment flows in recent years. [3]

Figure 2: FII

This figure shows a time series of Moody's credit ratings for India against the net FII flows to GDP ratio.

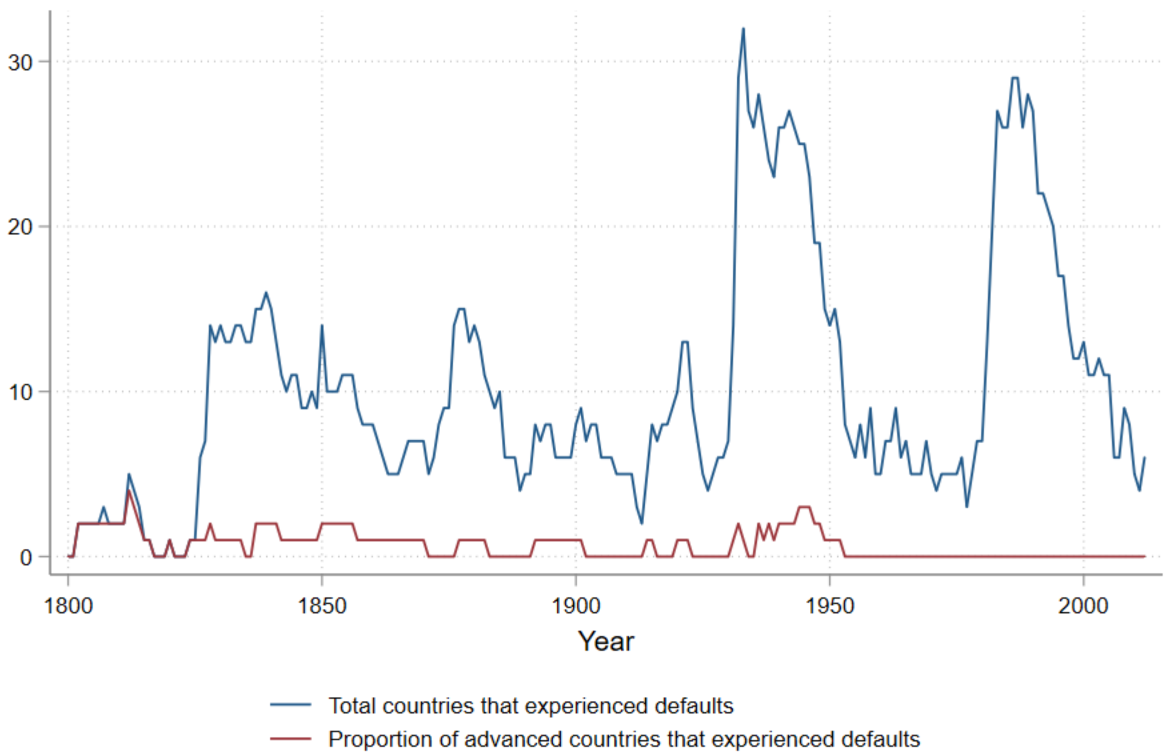


India's ratings are especially notable considering its sound default history and high "willingness to pay" (Government of India, 2017). Figure 3 shows the total number of countries that experienced defaults since 1800s using data from Reinhart and Rogoff (2009). While there have been periods of high sovereign defaults over the years (blue line), advanced economies have relatively low default rates. Plausibly, India's low ratings may be driven by the high default incidence within it peer group countries. However, given India's sound default history, such comparisons may likely be unwarranted.

[3]See remarks by Secretary-General of OECD in Feb. 2017 pointing to to potential for rating upgrades given India's strong economic performance (Outlook, 2017).

Figure 3: India's comparison to peer countries

This figure compares the number of sovereign defaults in advanced and emerging economies, through time. The red line represents the number of advanced economies that experienced defaults, and the blue line is the total number of countries that experienced defaults. Data is from Reinhart and Rogoff (2009).



2.1.1 Limited Market Reaction to Downgrades

Despite the rating downgrades, there has been a limited adverse impact on the market in India. Figure 4 shows that the downgrade of India's rating by Moody's to Baa3 on June 1st, 2020, had a very limited immediate impact on the market. Panel 4(A) suggests that markets shrugged off the downgrade and stock prices went up in the aftermath of the announcement. Possibly, ratings are backward looking and markets may have already priced in sovereign credit risk. Further, Moody's maintained India's investment-grade status potentially assuaging market concerns of further rating downgrades. The intra-day impact of the 10-year government bond's benchmark yield was also muted. Despite an immediate 5 basis points fall post-announcement, 10-year government bond yields returned to pre-downgrade levels withing 2 hours post-announcement (4(B)). The INR/USD foreign exchange rate also showed a similar pattern (panel 4(C) as it returned to pre-downgrade levels after depreciating by about 20 paise post-announcement.

Figure 4: Intra-day Market Impact

This figure shows the intra-day market impact of the Moody's downgrade announcement in June 2020. Panel (A), (B) and (c) show the impact on the Nifty 50 index, the Government 10-year benchmark yield and the INR vs USD bid rate, respectively.

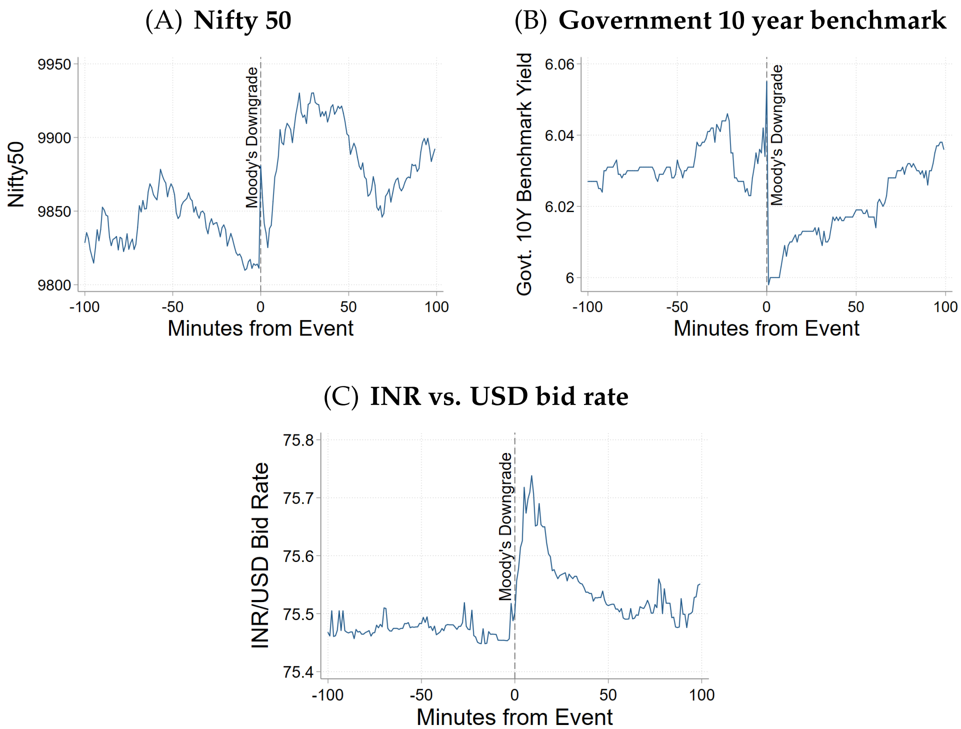


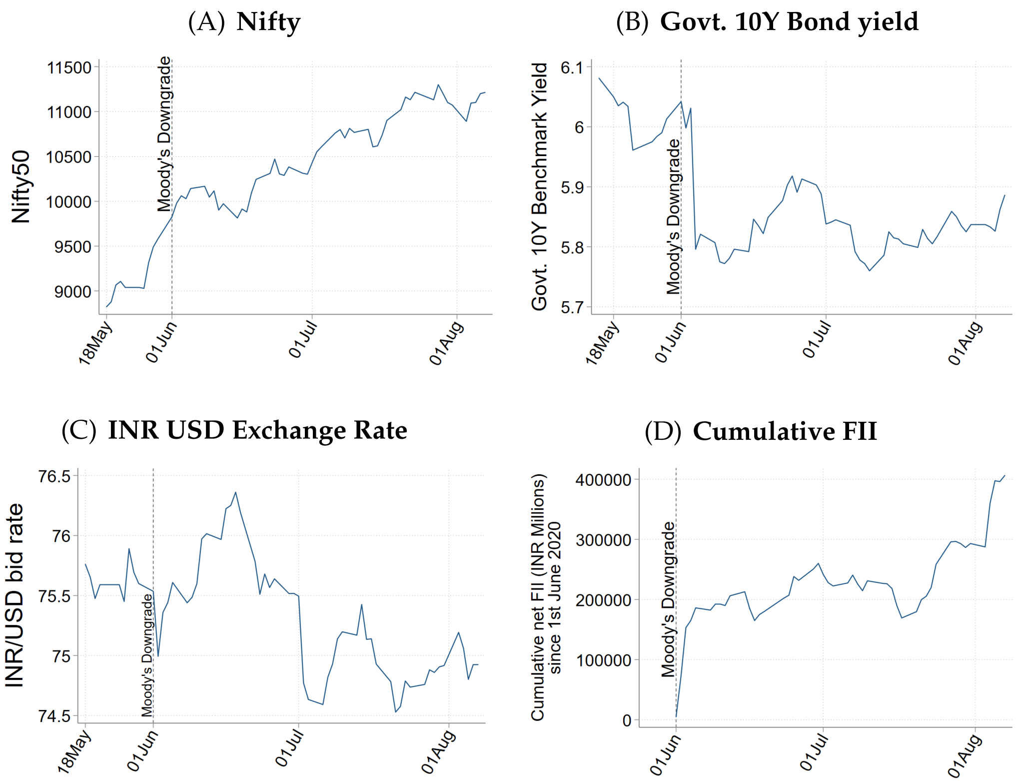
Figure 5 shows the impact on four indicators over two months. Panel 5(A) suggests that the stock market has largely ignored the downgrade. However, the date of announcement also coincided with a nation-wide easing of the COVID-19 related lock-down measures and hence the market impact may be confounded. The benchmark yield of the Indian government 10-year bond fell by 20 basis points after the downgrade, as shown by panel 5(B). In panel 5(C), we see that even though the Indian rupee depreciated against the dollar in mid-June due to geopolitical tensions, the downgrade did not have a lasting impact. Cumulative foreign institutional investment flows increased significantly in the two months post-downgrade (panel 5(D)), further underscoring the limited market impact of the downgrades.

2.2 CRA ratings still matter

The previous section called into question whether CRA rating matter given their lack of informativeness as market-based measures may be a more timely indicator of sovereign creditworthiness. However, CRA ratings can still matter despite the lack of informative-

Figure 5: Market Impact till date

This figure shows a longer-term market impact of the Moody's downgrade announcement in June 2020. Panel (A), (B), (c) and (D) show the impact on the Nifty 50 index, the Government 10-year benchmark yield, the INR vs USD bid rate and cumulative FII flows, respectively.



ness as shown by the market reaction. Mechanistic reliance by market participants can result in large effects of CRA rating changes. Rating thresholds are often integrated into laws, regulations, and market practices, and can result in herding and cliff effects. Foreign exchange reserves and asset managers rely on CRA ratings as the main source for credit assessments (Muller and Bourque, 2017). Further, bank capital requirements are often based on CRA ratings and even marginal changes in ratings can significantly alter financial stability.

CRA rating downgrades can also affect corporate borrowing costs. CRA policies prevent them from rating corporate firms in the country above the sovereign rating. Thus a sovereign rating downgrade results in an automatic rating downgrade of corporate firms, especially those that are the highest rated firms in the economy. Sovereign rating downgrades can thus can have real economic and financial consequences (Adelino and Ferreira, 2016; Almeida et al., 2017). For example, the recent Moody's downgrade in India was followed by downgrades of 6 PSUs.[4]

Rating downgrades can also amplify downturns due to negative feedback loops. Downgrades may trigger speculative selling, further reducing asset prices and resulting in a contagion. A recent example is the EU sovereign debt crisis of 2010-12, where some commentators speculate that a series of CRA downgrades amplified the sovereign debt crisis.[5] Sovereign credit default-swap spreads (CDS) for Greece increased in tandem with downgrades potentially suggesting negative feedback loops (Aizenman, Binici and Hutchison, 2013).

2.3 Biases in CRA ratings

Despite their importance, CRA ratings are often inconsistent and subject to numerous biases as documented by prior literature. Fuchs and Gehring (2017) document that CRAs have a home country bias and the CRA's home country gets one category better rating, on average. They also find that if a CRA's home-country banks are more invested in another country's assets, then the country's debt has a better rating. Additionally, cultural biases arise when similar language and culture compared to a CRA's home country leads to a higher rating.

A case in point highlighting the home-country bias is the rating downgrade of France post the Global Financial Crisis, whereas there were no similar downgrades for UK and

[4]Namely, Indian Oil Corporation, Hindustan Petroleum Corporation, Oil India, Petronet LNG, Bharat Petroleum Corporation and Oil and Natural Gas Corporation

[5]Wolfgang Schaeuble, the then German Finance Minister on Portugal downgrade (6 July 2011) said, "Yesterday's decisions by one rating agency do not provide more clarity. They rather add another speculative element to the situation" (Reuters, 2011).

USA with similar fundamentals.[6] Reluctance to upgrade countries like India are also problematic. For example, China's credit rating was upgraded from A+ to AA- in December 2010 while India's rating was at BBB-, despite China's soaring debt and growth slowdown (Government of India, 2017).

However, the lack of transparency in rating methodologies makes it difficult for market participants to assess and correct for such biases. The big three rating agencies - Moody's, Fitch and S&P - control 95% of the market (Hill, 2004), making it difficult to keep a check on such inherent biases. As documented in the previous subsection, CRA rating matter and even downgrades due to systematic bias and arbitrary factors can trigger self-fulfilling prophecies, driving even relatively healthy countries to default (Gärtner, Griesbach and Jung, 2011).

The above reasons underscore the need to study what factors drive CRA ratings. In particular, given the home-country bias in CRA ratings, we need to assess their suitability for developing countries. Finally, our goal is to answer the question: are CRA ratings able to predict sovereign defaults? We first start by describing the rating methodologies of the largest three CRAs, Fitch, S&P, and Moody's, in the next section.

3 CRA Rating Methodologies

In this section we discuss the individual CRA sovereign rating methodologies. Figure 6 gives an overview of the methodology of the credit rating agencies. Each CRA clubs factors into broad pillars. The figure shows the key pillars under each CRA and their respective ratings scale. The degree of qualitative and quantitative factors are also shown. Fitch has a 16-point ratings scale, which is 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, which is determined using a mix of qualitative and quantitative measures. The detailed methodology for each CRA is described below.

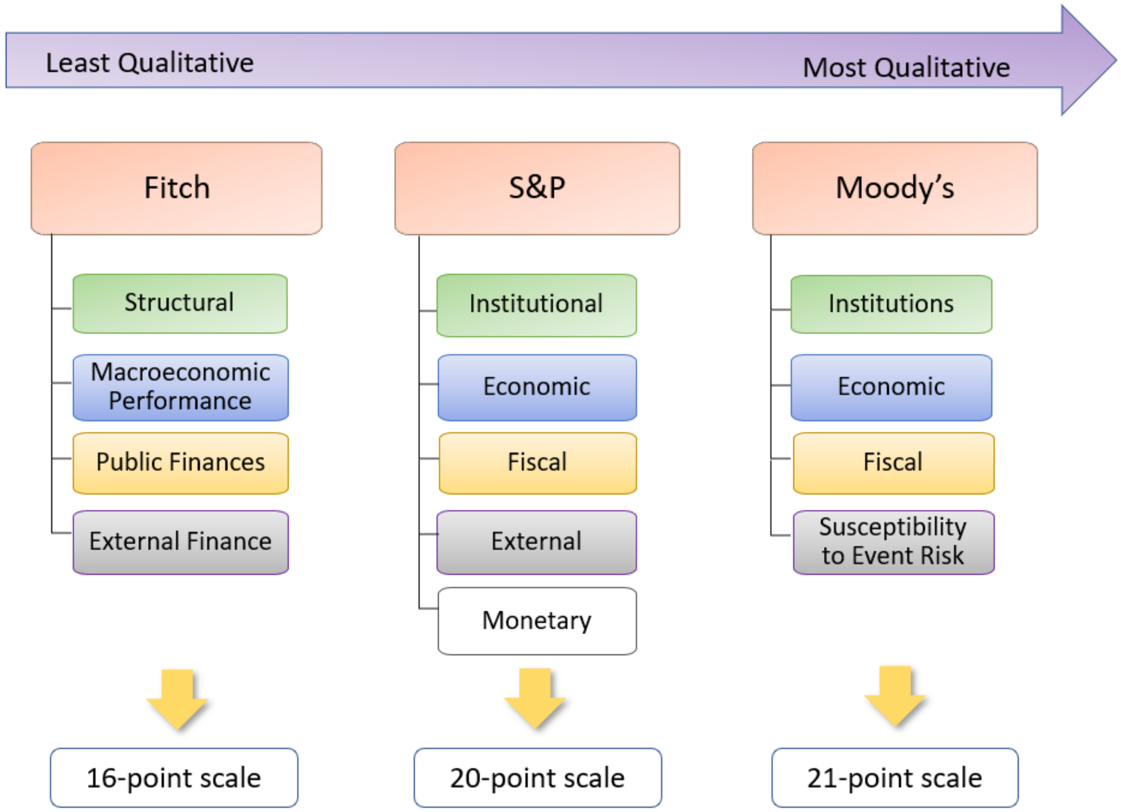
3.1 Fitch

Fitch arrives at the sovereign long-term foreign currency issuer default rating through a two-step approach. It first estimates a baseline rating score using a multivariate regression based Sovereign Ratings Model (SRM) of 18 variables representing four key pillars of the sovereign's credit profile - institutions, macroeconomic performance, public finances and external finances. The sovereign ratings model generates a predicted rating for every sovereign that is then scrutinized subjectively by the agency in its Qualitative Overlay

[6] See https://www.cfr.org/backgrounder/credit-rating-controversy

Figure 6: 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*b*), S&P (2017) and Moody's (2019)



(QO) (Fitch, 2020*b*). Qualitative adjustments, based on pre-determined 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. Fitch's variable specifications and detailed summary of the qualitative overlay can be found in the appendix in Section A.1 We describe below each of the four pillars as shown in Figure 6.

Structural features evaluate governance quality, wealth, flexibility of the economy, political stability and financial sector risks. The sovereign ratings model takes into account five variables that represent the institutional and structural features of the sovereign being rated. Firstly, "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 multi-dimensional institutional quality of the sovereign. Secondly, the percentile rank of GDP per capita in US dollars at market exchange rates measures individual income and savings capacity. Thirdly, 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. Fourthly, years since default or restructuring enters the regression as a non-linear 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. Lastly, broad money supply as a percentage of GDP enters the regression as the natural log of the percentage ratio. This variable proxies the level of financial intermediation in the sovereign, as it takes into account bank deposits, sovereign treasury bonds and other highly liquid financial instruments. Fitch (2020*b*) states that the overall post-estimation weight of this pillar in the model is 53.7%. Barring year since default, all variables should have a positive impact on the model output i.e. an increment in them results in a linear increment in the ratings score. In its qualitative overlay, Fitch ratings assesses metrics of political stability and capacity, financial sector risks and other structural factors not captured in the SRM.

Fitch's macroeconomic performance, policy and prospects pillar evaluates the macroeconomic stability, policy credibility, GDP growth outlook and inflation of the rated sovereigns. The first variable it takes into acccount is real GDP growth volatility which is measured as the natural logarithm of an exponentially weighted standard deviation of historical percent changes in real GDP. Secondly, the pillar encompasses consumer price inflation as the three-year centered average of annual percent change in consumer price index, truncated between 2% and 50%. Lastly, real GDP growth is used as a variable. It enters the regression as the three-year centered average of annual percent change in real GDP.

This pillar contributes to 10% post-estimation weight in the sovereign rating model. In its qualitative overlay, Fitch looks at the five-year GDP growth outlook and the sovereign's relative performance, both across time and its peers in the rating group, to make qualitative adjustments to the baseline rating score.

In its analysis of public finances, Fitch studies government debt, fiscal balance, public debt dynamics and fiscal policy. Firstly, gross general government debt enters the regression as the three-year centered average of debt as a percentage of GDP. This is used as the leading indicator for the debt burden of the sovereign. Secondly, the three-year centered average of gross government interest payments expressed as the percentage of general government revenues, is used to denote the annual fiscal burden of servicing the sovereign's debt. Thirdly, general government fiscal balance again enters as the three-year centered average of gross general government budget balance expressed as a percentage of GDP. It too gives an estimation of general government borrowing position. Lastly, the three-year centered average of public foreign currency denominated and indexed debt expressed as a percentage of gross general government debt is used to denote the gross external currency exposure to the sovereign.

The post-estimation weight of the "Public Finances" pillar in the sovereign ratings model is 18%. In its qualitative rationale, the agency looks at the sovereign's fiscal financing flexibility - its record of market access, depth of capital markets, potential sources of financing, public debt sustainability and fiscal structure.

Fitch analyzes the sovereign's balance of payments, external balance sheet and external liquidity under the fourth and final pillar: external finances. The first variable it includes under this pillar is reserve currency flexibility which is the natural logarithm of the share of that country's currency in global foreign-exchange reserve portfolios (plus a technical constant), as reported by the IMF in its COFER database. The second variable used is commodity dependence which is the share of non-manufactured merchandise exports in current account receipts. Thirdly, the model also uses official international reserves for non-reserve currency sovereigns. It is defined as the year-end stock of international reserves, including gold, expressed as months' cover of current external payments. The fourth variable under this pillar is sovereign net foreign assets expressed as three-year centered average percent of GDP. Fifthly, the model uses the sum of current account balance and net FDI inflows expressed as the three-year centered average percentage over GDP. Lastly, the three-year centered average of external interest service expressed as percentage of current account receipts is used under this pillar to measure the financial burden of servicing external debt.

External finances makes up for the remainder 17.4% weight in the sovereign ratings model. The qualitative overlay under this pillar looks at metrics of external financing flexibility, external debt sustainability and the sovereign's vulnerability to external shocks.

3.1.1 S&P

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 6. 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.

The Institutional factor comprises an analysis of how a government's institutions and policy-making affect its credit fundamentals. The initial institutional assessment combines two main sub-factors: effectiveness, stability, and predictability of policy making, political institutions and civil society; and transparency and accountability of institutions, data, and processes. To assess effectiveness, stability and predictability of policy-making, they consider the track record of a sovereign in managing past political, economic, and financial sector crises; maintaining prudent policy-making; and delivering balanced economic growth. Additionally, they consider predictability in the overall policy framework; Actual or potential challenges to political institutions; and cohesiveness of civil society. To assess transparency and accountability, they consider the existence of checks and balances between institutions; perceived level of corruption in the country; respect for the rule of law; and independence of statistical offices and the media. The initial institutional score can be negatively adjusted by taking into consideration a sovereign's debt payment culture and external security risks. A sovereign's debt payment culture results in a negative adjustment when it has significant and sustained arrears on bilateral official debt (i.e., debt owed to foreign governments and government-owned entities) or has an odious debt or there has been no material policy change since the last default on commercial debt.

The key drivers under the Economic factor are income levels, growth prospects, and economic diversity and volatility. The initial score for Economic factor is based on a country's income level, as measured by its current-year estimate for GDP per capita, converted to U.S. dollars. The initial score can receive a positive or negative adjustment by up to two categories based on the economy's growth prospects, its potential concentration or volatility, and the potential material data inconsistencies, gaps, or discontinuities. A country's growth prospect is measured using the average growth in a country's real per capita GDP over a 10-year period, to cover generally at least one economic cycle. More specifically, the measure of real per capita GDP trend growth is the average of six years of historical data, the current-year estimate, and three year forecasts. The latest historical year, current-year estimate, and forecasts are weighted 100%, while previous years are assigned a lower weight, to avoid a steep drop or increase when an exceptional year drops out of the 10-year average. A country's initial score of the Economic factor would generally be one category worse when GDP growth seems to be fueled mostly by a rapid increase in depository corporation claims on the resident non-government sector. Subsequently, a country receives a negative adjustment if it carries significant exposure to a single cyclical industry or its economic activity is vulnerable because of constant exposure to natural disasters or adverse weather conditions. Finally, a country receives a negative adjustment in cases where national accounts data display material data inconsistencies, gaps, or discontinuities, or where there is reason to believe that the quality of national accounts data is hampered by technical or administrative shortcomings or political interference.

The main sub-factors that determine a country's External factor are the status of its currency in international transactions, the country's external liquidity, and its external position. To measures the currency's status in international transactions, S&P assesses whether a country has a "reserve currency" or an "actively traded currency." Sovereigns with reserve currency are the ones with a currency that accounts for more than 3% of the world's total allocated foreign exchange reserves based on the IMF report "Currency Composition of Official Foreign Exchange Reserves." Subsequently, a country with an actively traded currency is one with a currency that is bought or sold in more than 1% of global foreign exchange market turnover, based on the Bank for International Settlement (BIS) report "Triennial Central Bank Survey," that is not a reserve currency as defined above. The key measure of a country's external liquidity is the ratio of gross external financing needs to the sum of CAR plus usable official foreign exchange reserves, which is an average of the current-year estimate and forecasts for the next two to three years. Finally, the measure of a country's external indebtedness is the ratio of narrow net external debt to current account receipts (CAR) (or current account payments (CAP) if external liquid assets exceed external debt). A sovereign receives the initial score for the External factor after taking into account the above aforementioned sub-factors as described in S&P (2017). The initial score can receive positive adjustments if a country displays a significantly stronger net external position or a country with actively traded currency runs consistent current account surpluses. Subsequently, negative adjustments by one notch are made when a country is exposed to a risk of marked deterioration in external financing or to significant volatility in terms of trade, or a country has low external debt reflects debt constraints or material data inconsistencies or actively traded currencies running high current account deficits. Finally, negative adjustments by two notches are made when sovereigns with actively traded currencies run very high current account deficits.

The Fiscal factor reflects the sustainability of a sovereign's fiscal balances and debt burden. It considers fiscal flexibility, long-term fiscal trends and vulnerabilities, debt structure and funding access, and potential risks associated with contingent liabilities. The analysis of the Fiscal factor is the analysis is divided into two segments: Fiscal performance and flexibility, and Debt burden. The overall score is the average of the two.

To determine a sovereign's fiscal performance and flexibility, they first derive an initial assessment based on the prospective change in net general government debt calculated as a percentage of GDP and assign score as give in S&P (2017). Change in net general government debt is an average of the current-year estimate and forecasts for the next two or three years. Positive adjustments are made for governments with large liquid financial assets or for governments with greater ability to increase general government revenues or cut general government expenditures in the short term compared with governments in countries with a similar level of development. Similarly, negative adjustments are made when a country has unsustainable or volatile revenue base that may boost fiscal performance over the period average, or for a government with limited ability to raise general government revenues in the short term compared with sovereigns with a similar level of development or a country has shortfalls in basic services and infrastructure or for countries with unaddressed medium-term pressure due to age-related expenditure. A sovereign's debt burden assessment reflects its prospective debt level. Factors underpinning the assessment are debt relative to GDP, the interest cost of the debt relative to general government revenue, debt structure and funding access, and the magnitude of and likelihood that contingent liabilities may become government debt. The initial score for debt burden takes into account average of the current-year estimate and forecasts for the next two or three years of net general government debt and general government interest expenditure, and is obtained by the method described in S&P (2017). For sovereigns in a net general government debt position and benefiting from concessional lending, the debt assignment is generally one category better than the initial assessment if it is assessed that a government's borrowing needs are likely to be covered by official funding during the next two to three years. Debt burden score is negatively adjusted when more than 40% of gross government debt is denominated in foreign currency, or the average maturity is typically less than three years or nonresidents hold consistently more than 60% of government commercial debt or the debt service profile is generally subject to significant variations or the banking sector's exposure to the government is typically above 20% of its assets.

The initial score for the Monetary factor is derived by combining the assessments of the exchange-rate regime (weighted 40%), and the monetary policy credibility (weighted 60%). S& P uses exchange-rate regime definition from the IMF System Annual Report on Exchange Arrangements And Exchange Restrictions. The credibility of its monetary policy is measured using different factors such as monetary authority independence, monetary authority tools and effectiveness, price stability, lender of last resort, and development level of local financial system and capital markets. The initial score is derived as given in S&P (2017). Negative adjustments are made if a country's transmission mechanisms are weak or are significantly weakening, thereby impeding monetary flexibility, or resident deposits or loans in foreign currency (dollarization) exceed roughly 50% of total, or extensive exchange restrictions are applied (as informed by compliance with IMF Article VIII obligations).

The "indicative rating level" obtained using the rating matrix is subject to the following supplemental adjustments in order to obtain the final foreign currency ratings: extremely weak external liquidity, extremely high fiscal debt burden, very high institutional risk and high debt burden, and event risk. S& P uses a mix of publicly available and propriety data along with their internal forecasts.

3.1.2 Moody's

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 6, and each of these pillars have a host of sub factors and sub-subfactors. All these sub-factors are pre-assigned weights and these weights are used while 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). Out of these four pillars, 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.

Economic Strength consists of four main sub factors. These are – a 10 year average of real GDP growth centered on the current year, a 10 year average of GDP growth volatility for t-9 time periods, nominal GDP in USD billions and per capita GDP in PPP terms. The qualitative adjustments to this factor are based on four criteria, namely– Flexibility, Diversity, Productivity and labour supply challenges. To assess flexibility they consider indicators such as the WEF Global Competitive Index, the WEF Financial Market Development Index and other components that measure labour and goods market efficiency. Diversity is assessed using indicators such as the UNCTAD Products Export Diversification Index, the Observatory of Economic Complexity's Economic Complexity Index and the WDI indicator for goods exports to high-income countries. They make upward adjustments if countries are especially diverse for their size, and also if a country has a large amount of untapped natural resources. Conversely, they make downward adjustments if a country's growth is largely dependant on the export of one particular commodity. To asses productivity, they consider WEF's Infrastructure, Innovation and Higher Education and Training Indexes. They also consider estimates of long-term changes in productivity based on the average growth of real GDP per capita over 10 years. To assess labour supply challenges they consider trends in migration and female labour force participation, indicators measuring the degree of ageing, estimates of working age population growth. Apart from these 4 criteria, they also consider structural breaks in the economy, macro-prudential frameworks that curb excessive credit growth and climate change effects.

Fiscal Strength has two sub-factors – Debt burden and debt affordability. Debt burden considers general government debt to GDP and general government debt to revenue. Debt affordability considers general government interest payments to revenue and general government interest payments to GDP. Qualitative adjustments are based on the government's debt trend, the government's exposure to debt denominated in foreign currencies, magnitude of non-financial public sector debt that is government guaranteed and sovereign wealth funds.

Institutions and Governance Strength consists of two main sub-factors – Quality of Institutions, and Policy Effectiveness. Quality of Institutions further has two sub-sub-factors – the quality of legislative and executive institutions, and the strength of civil society and the judiciary. The quality of legislative and executive institutions is primarily based on the World Governance Indicators for Regulatory Quality and Government Effectiveness. It also considers forward-looking views on the efficiency of government and public administration, the reporting of data, the capacity to translate policy into law, skill of the public sector workforce and the voice independent bodies have in policymaking. The strength of civil society and the judiciary is primarily measured using the Worldwide Governance Indicators for voice and accountability, rule of law, and control of corruption. It also considers forward-looking on law enforcement, the separation of power between the judiciary and the government, the effectiveness of judicial processes and the civil society's capacity to act as a check on the exercise of government power (Moody's, 2019). Policy Effectiveness also has two sub-sub-factors – fiscal policy effectiveness and monetary policy effectiveness. Fiscal policy effectiveness considers factors such as the historical and anticipated debt to GDP levels, adherence to fiscal targets and expenditure ceilings, the trajectory and flexibility of budget balances,the existence of non-partisan bodies that take part in the budget-making process transparency of government accounts and the existence of robust risk-mitigating and medium-term policy planning processes. Monetary and Macroeconomic Policy Effectiveness considers factors such as the level of inflation relative to targets, the implied effectiveness of monetary policy, the effectiveness of public policy response to adverse economic or social shocks, price stability, the capacity and willingness to address macroeconomic imbalances, central bank independence, imbalances in the financial and banking system, and effective banking regulations. Negative adjustments to this score are made if government's default on debts owed to the private sector.

Susceptibility to Event Risk enters the calculation using a minimum function. It takes into account 4 kinds of risks. political risk, government liquidity risk, banking sector risk and external vulnerability risk. Political risk is assessed using the World Governance Indicator for voice and accountability, and political stability; the gini index; unemployment; geopolitical relationships; political transitions and the degree of armed conflict within the country. Government liquidity risk is based on the governments ease of access to three categories of borrowing – local currency borrowing from domestic creditors, local currency borrowing from external creditors and foreign currency borrowing. Other considerations include "the government's record of having access to these types of funding, their cost and maturity relative to peers, the diversity of each sovereign's investor base for different types of debt instruments, the reliance on borrowing from official lenders and the existence of material foreign currency ". Banking sector risk is based on the country's total domestic bank assets to GDP and the risk of Banking Sector Credit Event (BSCE). The BSCE is based on the Moody's Baseline Credit Assessments. Adjustments to this score are made if the domestic banking system in concentrated in a few banks or if there is a sharp rise in funding costs. External vulnerability risk is based on the current account balance, FDI inflows, and a diversification of the export base structure. It is also based on the net international investment position, the ratio of gross external dent to current account receipts, the composition of foreign liabilities, and the external vulnerability indicator ratio.

The numeric scores resulting from Economic Strength and Institutions and Governance Strength is averaged to give an Economic Resiliency score. This score is averaged with the Fiscal Strength score to result in a Governance and Fiscal Strength score. This score is converted to an alpha numeric score. Using a matrix of alphanumeric scores, a final score is arrived at by combining the Governance and Fiscal Strength Score and the Susceptibility to Event Risk score. This results in an initial alphanumeric rating. There could be further minor adjustments made to this initial score, leading to a final score for the sovereign. Moody's uses data from a lot of public sources, however a lot of its data is also based on internal forecasts and proprietary information.

4 Examining the Factors that Determine CRA Ratings

We now build a parsimonious model to determine the major *quantitative* factors that drive CRA ratings. We then compare actual ratings to and predicted ratings, with a special focus on India. 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.

4.1 Individual Rating Models Based on Quantitative Factors

4.1.1 Empirical Specification

We select variables based on data availability and ensure that we include at least one variable from each pillar so that all pillars are represented for each CRA model. The empirical specification is:

$$Y_{kit} = \alpha + \beta * X_{kit} + \epsilon_{kit} \tag{1}$$

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. The rating score of each CRA is converted to a numerical scale, with the highest rating getting higher values. The Fitch rating score ranges from B- to AAA converted to a 16-point scale, with AAA corresponding to the highest rating. Similarly, the S&P rating scale ranges from CC to AAA converted to a 20-point scale and the Moody's rating score ranges from C to Aaa converted to a 21-point scale. The panel for Fitch consists of 41 countries for the period 1994-2020. The panel for S&P consists of 87 countries for the time period 1998–2020, and finally the panel for Moody's consists of 83 countries for the time period 2000–2020.[7]

X_{kit} is a vector of explanatory variables of rating agency 'k', for country 'i' and for time 't'. The explanatory variables of Fitch are: Institutional Score, GDP per capita , Log share in global GDP, Broad Money, Debt to GDP, Years since default, Real GDP growth, Commodity Dependence, Growth Volatility, CPI, Fiscal Balance, Total Reserves, (CAB+FDI). We winsorize the three-year centered average of CPI inflation at 1% and 99% of its distribution to remove bias due to outliers.

The explanatory variables of S&P are: Transparency of Institutions, GDP Per Capita, GDP Per Capita Growth Rate, Actively Traded Currency, Gross Financing Needs by (Current Account Receipt + Total Reserves), Government Gross Debt (% GDP), Central Bank

[7]The difference in number of observations arises from difference in data-availability.

Independence. Transparency of institutions is a country percentile rank constructed using transparency of policy-making, control of corruption, rule of law, and voice and accountability, from WGI and GDP per capita and GDP per capita growth rate are in nominal terms. Actively traded currency is a dummy variable which takes the value 1 when a sovereign's currency is bought or sold in more than 1% of global foreign exchange market turnover. Gross financing needs ratio and Government Debt as a percentage of GDP are an average of the current year and two years forecast. Central Bank Independence is a categorical variable ranging from 1 to 5, with 5 being the most independent.

Finally, the explanatory variables of Moody's are: GDP Per Capita, GDP Growth, Regulatory Quality, Rule Of Law, Debt to GDP, Unemployment Rate, Current Account. GDP per capita is in PPP terms and growth in real GDP is a 10 year average, centered on the current year. It considers a 5 years of forecast data, 4 years of historical data, and current year data. Regulatory quality and rule of law are indicators from the WGI, and are country percentile ranks. Debt to GDP and unemployment rate are both in percentage terms, and current account balance is a percentage of GDP. All variables, except actively traded currency are standardised and standard errors are clustered at the country level.

Figure A1 in the Appendix summarizes the factors under each pillar used in the regression specification. A detailed list of data sources can be found in Section A of the appendix.

We now discuss the results from running the specification in Equation 1 for each of the CRAs below. Our main results come from the regressions using a simple OLS. However, in an OLS, we have to assume that the dependent variable is cardinal. For instance, we assume that the difference between a AAA and a Aa1 is the same as that between a Baa1 and a Baa2. As this may not be reasonable assumption, as a robustness check, we repeat all regressions using an Ordinal Probit. However, as we will see, our results are robust to this assumption.

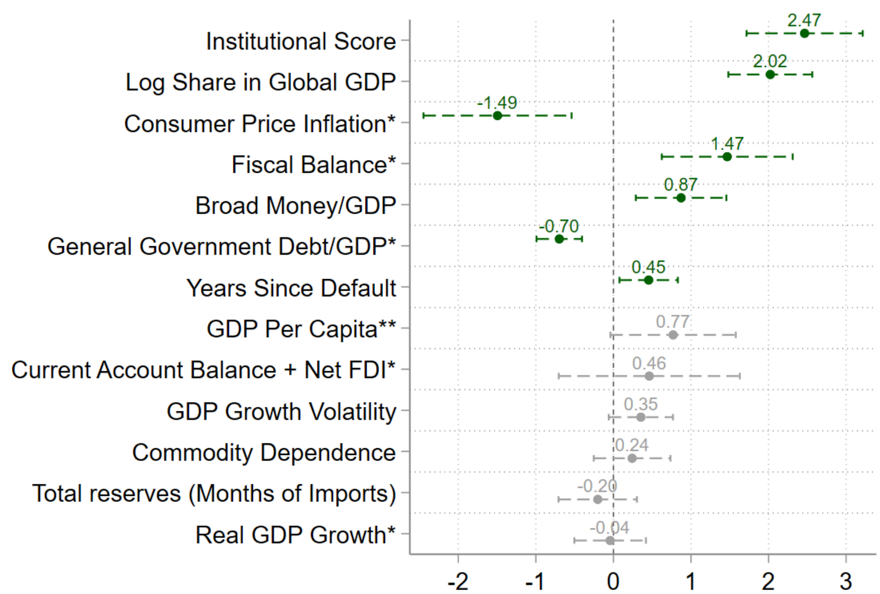
We also use the World Bank's 2016 classification of countries by income and group our sample into high and low/middle-income countries. The countries classified as high-income by the World Bank are also high-income in our sample, while remaining countries are clubbed together as low/middle-income (low-income, middle-income and upper-middle income) countries.

4.1.2 Fitch

Figure 7 and Table 1 present the results of the analysis for Fitch. Institutional score, which is the composite average of six governance indicators is the most important factor for Fitch. A one standard deviation change in the score is associated with a 2.5 higher rating score. The share in global GDP measures the relative size of the economy to the rest of the world. A one standard deviation higher log of share in global GDP results in a 2 point higher rating. Next, inflation and fiscal balance have an opposite but similar impact on ratings in terms of magnitude for a one standard deviation change. Fitch uses broad money as a proxy for the level of financial inter-mediation in the economy. A one standard deviation higher broad money to GDP ratio translates to a 0.87 point higher rating score. General government debt, which is a leading indicator for near-term debt stress, has an expected negative and significant coefficient of -0.7. Fitch also takes into account the default history of sovereigns in its baseline model. A one standard deviation increase in years since default is associated with a 0.45 point higher rating score. Thus, rarer a default event in a sovereign's history, the better its credit rating is. GDP per capita, GDP growth and Current Account Balance have insignificant coefficients, which suggests that structural features and fiscal indicators can predominantly explain variation in sovereign ratings for Fitch.

Figure 7: Coefficient Plot: Fitch

This figure presents the coefficient plot of the OLS regression for Fitch. Variables are as defined in 3.1



Column (2) presents the results using an ordinal probit specification. The results are similar to column 1. We run the OLS regression separately for high-income and low/middle-income countries in columns 3–4 to check whether results are driven by these-sub-group of countries. However, columns 3–4 show similar loading on each factor, assuaging such concerns.

To summarize, institutional quality is the main determinant of Fitch sovereign ratings.

Table 1: Fitch

This table presents the results of the regression for Fitch. The independent variables used are as mentioned in the Fitch (2020*b*) report. Variables are as defined in Section 3.1. Columns (1) and (2) present results using OLS and ordered probit, respectively. Columns (3) and (4) present the results for high-income and low/middle-income countries, respectively. Standard errors are clustered at country-level and all variables are standardized for ease of interpretation.

	(1) OLS	(2) Ordered Probit	(3) High – Income	(4) Low/Middle – Income
Institutional Score	2.465***	1.859***	2.771***	1.546***
	(0.370)	(0.303)	(0.363)	(0.373)
Log Share in Global GDP	2.023***	1.470***	1.908***	0.787**
	(0.268)	(0.293)	(0.302)	(0.334)
Consumer Price Inflation*	-1.494***	-1.215***	-0.825**	-1.552**
	(0.472)	(0.377)	(0.280)	(0.710)
Fiscal Balance*	1.467***	1.074***	0.954**	1.033*
	(0.418)	(0.333)	(0.421)	(0.553)
Broad Money/GDP	0.872***	0.650***	0.773***	1.332***
	(0.289)	(0.207)	(0.233)	(0.313)
General Government Debt/GDP*	-0.698***	-0.539***	-0.860***	-0.904*
	(0.145)	(0.124)	(0.0682)	(0.479)
Years Since Default	0.454**	0.355***	0.132	0.542***
	(0.186)	(0.135)	(0.242)	(0.185)
GDP Per Capita Percentile	0.769*	0.551*	2.782***	-0.103
	(0.400)	(0.309)	(0.773)	(0.332)
Current Account Balance + Net FDI*	0.463	0.505	-0.0730	0.818
	(0.578)	(0.426)	(0.395)	(0.548)
GDP Growth Volatility	0.354*	0.374**	0.422	0.472*
	(0.205)	(0.158)	(0.332)	(0.239)
Commodity Dependence	0.241	0.228	-0.480*	0.791***
	(0.245)	(0.186)	(0.228)	(0.237)
Total reserves (Months of Imports)	-0.202	-0.0191	-0.229	0.428**
	(0.250)	(0.186)	(0.436)	(0.191)
Real GDP Growth*	-0.0422	-0.0291	-0.0659	0.150
	(0.228)	(0.209)	(0.218)	(0.282)
No. of Obs.	609	609	255	354
R squared	0.915		0.959	0.825
Pseudo R squared		0.429		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

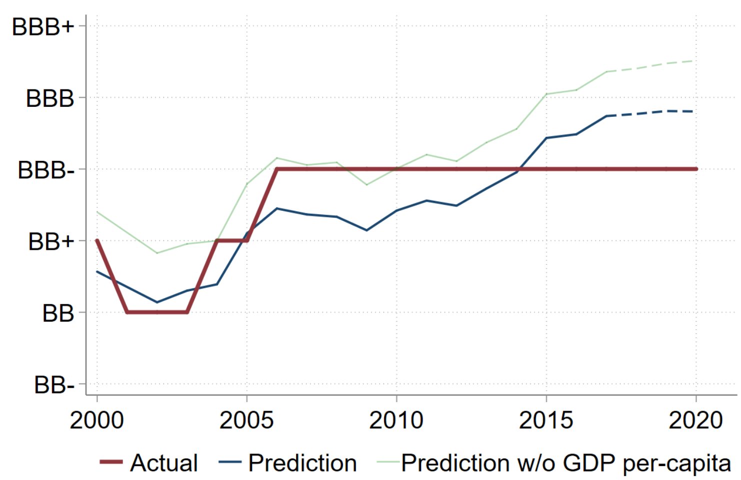
Fiscal health measured in terms of government debt to GDP and fiscal balance are also important. This is a theme that will be common across the CRAs.

We next examine how well our simple model parsimonious model is able to account for the variation in sovereign credit ratings for Fitch. The R^2 in Table 1, column 1 suggests that a handful of factors account for nearly 91% of the variation in ratings. Further, the model has a better fit for high-income countries (R^2=96% in column 3) compared to low/middle income countries (R^2=83% in column 4). This difference perhaps highlights the lower suitability of this one-size fits all approach to rating methodology, especially for low/middle-income countries.

We next plot the predicted rating across time for India using our simple OLS model in Figure 8 along with actual ratings. We also plot the prediction without GDP per capita in the regression specification, given prior criticism that GDP per capita can unfairly bias ratings for low/middle-income countries (Government of India, 2017).

Figure 8: Predicted Ratings for India: Fitch

This figure presents the predicted ratings for India using the OLS regression for Fitch. Dashed lines represent rating prediction with estimated data.

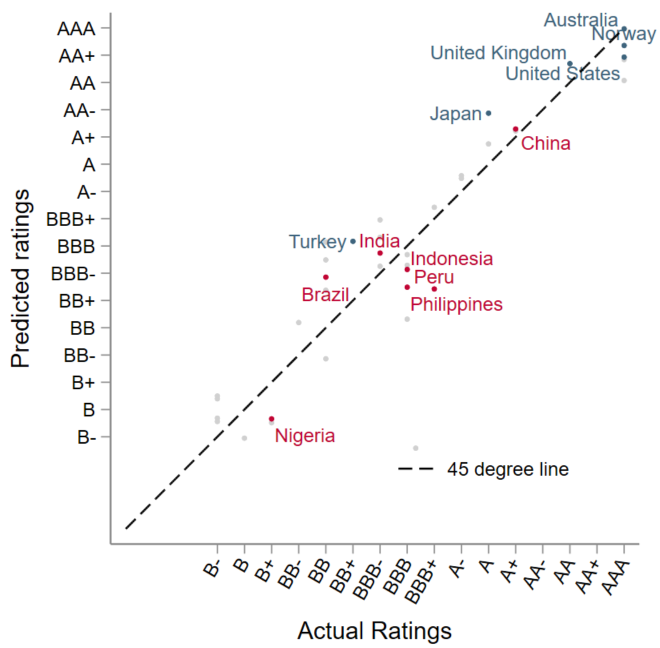


We see that the prediction across time is very close to the actual ratings, and that the prediction without GDP per capita is generally one-notch higher than the one with.

Figure 9 plots the scatter plot of actual and predicted ratings for the cross-section of countries in 2016. The 45 degree line indicates countries for which the predicted rating (using our simple model) is exactly equal to actual ratings. Points below (above) the line refer to countries where the CRA rating from the parsimonious model is under(over)-predicted relative to the actual Fitch rating. Strikingly, observations are as likely to be above the 45 degree line as below for both India's peers (highlighted in red) as well as advanced economies (highlighted in blue). To sum, we are able to match the Fitch ratings both in the cross-section and the time-series. This is particularly striking considering that our model is parsimonious and extremely simple. The model for Fitch seems to be equally accurate for both EMEs and advanced economies, as there does not seem to be a pattern of under or over-predicting in the scatter plot. This could plausibly be due to Fitch's reliance on hard, quantitative factors.

Figure 9: Predicted vs Actual: Fitch

This figure presents a scatter plot of predicted versus actual ratings for the year of 2016, using the OLS regression. Countries in red are EMEs while countries in blue are developed economies.



4.1.3 S&P

The results for the regression for S&P are presented in Table 4.1.3, and the corresponding coefficient plot is in Figure 10.

Table 4.1.3, column 1 shows that Transparency of Institutions, GDP per capita and GDP per capita growth rate are all positive and significant. General Government Debt as a percentage to GDP is significant and negative. A one standard deviation increase in the percentile rank for transparency of institutions is associated with almost a 3 notch increase in ratings. Similarly, a one standard deviation increase in GDP per capita leads to a nearly 2 notch higher rating. A one standard deviation greater General government gross debt has a 1 notch decrease in ratings. Finally, a one standard deviation higher GDP

Table 2: S&P

This table presents the results of the regression for S&P. The independent variables used are as mentioned in the S&P (2017) report. Variables are defined in Section 3.1.1. Columns (1) and (2) present results using an OLS and an ordered probit, respectively. Columns (3) and (4) present the results for high-income and low/middle-income countries, respectively. Standard errors are clustered at country-level and all variables are standardized for ease of interpretation.

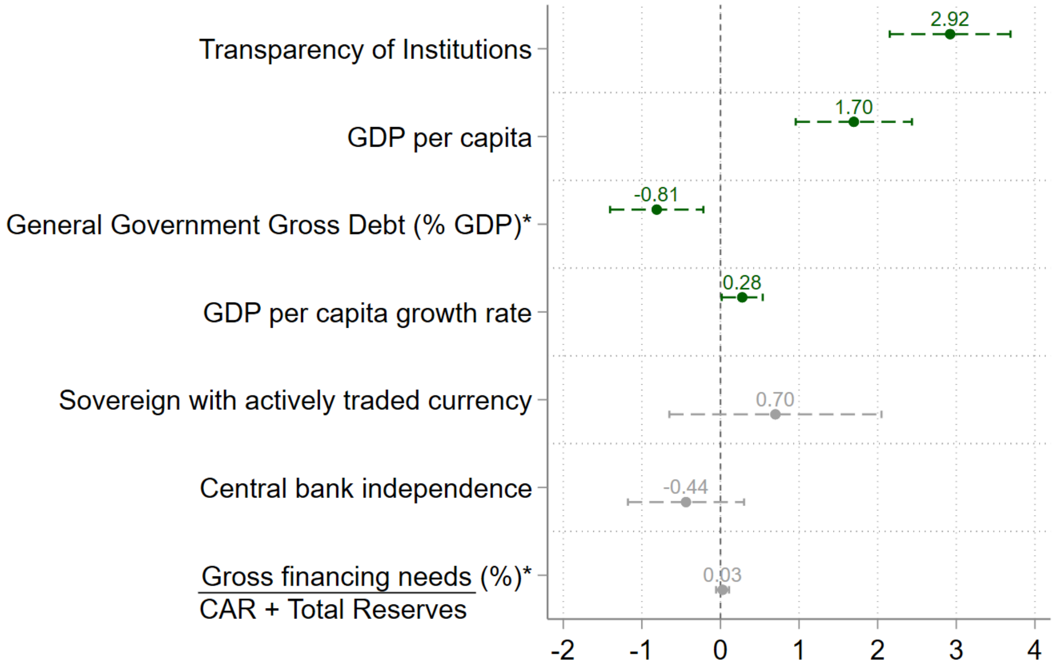
	(1) OLS	(2) Ordered Probit	(3) High – Income	(4) Low/Middle – Income
Transparency of Institutions	2.921***	0.956***	3.156***	2.357***
	(0.386)	(0.157)	(1.083)	(0.433)
GDP per capita	1.697***	1.258***	1.315***	1.946
	(0.372)	(0.274)	(0.387)	(1.714)
GDP per capita growth rate	0.277**	0.136***	0.132	0.180
	(0.131)	(0.0509)	(0.178)	(0.149)
$\frac{\text{Gross Financing Needs}}{\text{CAR + Total Reserves}}\%$	0.0268	0.0278	0.0822***	-17.12***
	(0.0415)	(0.0403)	(0.0273)	(3.994)
Sovereign with actively traded currency	0.699	0.213	0.171	3.755***
	(0.679)	(0.297)	(0.629)	(0.994)
Govnt Gen Gross Debt (%GDP)*	-0.811***	-0.473***	-0.576*	-0.881**
	(0.299)	(0.115)	(0.302)	(0.345)
Central bank independence	-0.439	-0.209	-0.805**	-0.260
	(0.372)	(0.144)	(0.393)	(0.335)
No. of Obs.	960	960	458	487
R squared	0.736		0.574	0.541
Pseudo R squared		0.256		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 10: Coefficient Plot: S&P

This figure presents the coefficient plot of the OLS regression for S&P. Variables are as defined in 3.1.1.



per capita growth rate is associated with a 0.28 points higher rating. All these effects are in the expected direction. Column 2 of shows the results using an ordered probit and most effects stay the same, except Transparency of Institutions which has a much more muted effect. In columns (3) and (4) we examine effects by dividing the sample into high-income and low/middle-income countries. Notably, we see that GDP per capita loses significance for low/middle-income countries.

The R^2 of 74% for S&P is smaller than what we observed for Fitch, with marginally lower explanatory power for low/middle-income countries. Figure 11 shows the time-series plot for India's predicted vs actual ratings for S&P. The thick red line shows the actual S&P rating from 2001 to 2020. We predict ratings with and without GDP per capita. As shown, the predicted ratings with GDP per capita is lower (dark blue line) than the predicted ratings without GDP per capita (light blue line). The dashed lines show predicted ratings with the estimated data. Our predicted ratings from these regressions for S&P are as accurate compared to Fitch. We conjecture that S&P's methodology contains a mix of both quantitative and qualitative factors and the slight under-prediction could be because our model (by design) captures only the quantitative factors.

Figure 12 shows the scatter plot for predicted vs actual ratings for the cross-section of countries in 2016. The correlation between predicted and actual ratings is 0.872 (for

Figure 11: Predicted Ratings for India: S&P

This figure presents the predicted ratings for India using the OLS regression for S&P. Dashed lines represent rating prediction with estimated data.

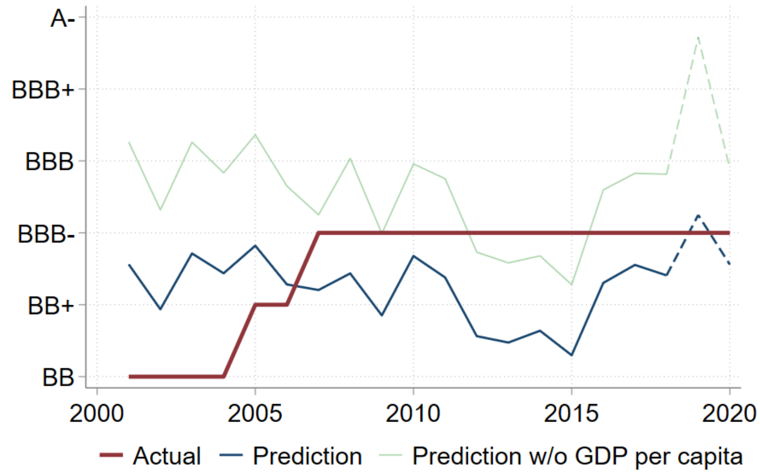
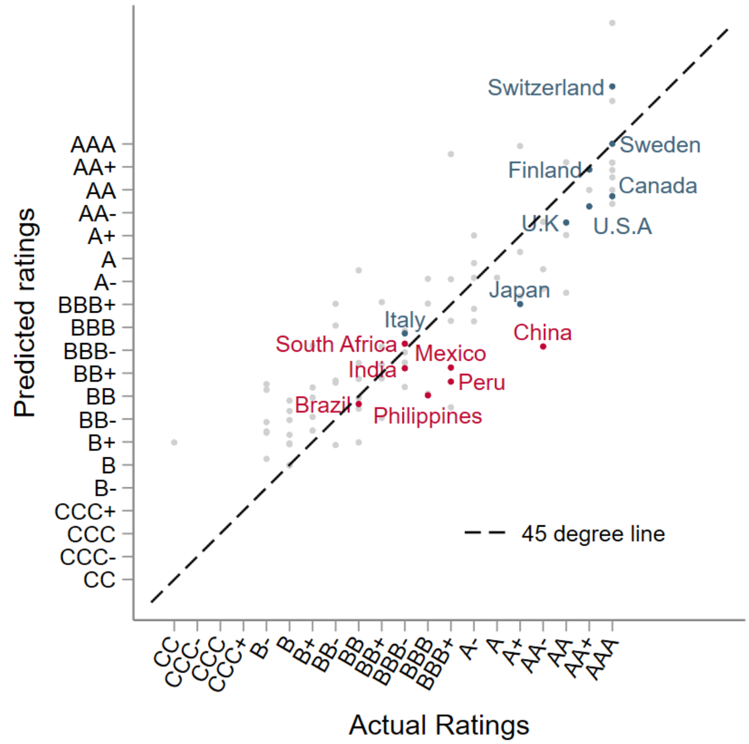


Figure 12: Predicted vs Actual: S&P

This figure presents a scatter plot of predicted versus actual ratings for the year of 2016, using the OLS regression. Countries in red are EMEs while countries in blue are developed economies.



the cross-section in 2016), slightly lower than that of Fitch in as seen in Figure 9, though still notable given our sparse specification. The predicted ratings for India's peers (highlighted in red) lie below the 45 degree line suggesting that the quantitative factors under-predict ratings for the low/middle income countries for S&P. Upward adjustments, likely based on qualitative factors, can match predicted ratings to actual ratings. This is in contrast to the advanced economies (highlighted in blue) which are as likely to lie above or below the 45 degree line suggesting that there is no systematic over- or under-bias in rating estimates.

4.1.4 Moody's

The results for Moody's are presented in Table 3 and the corresponding coefficient plot in Fig 13. Table 3, column 1 shows that regulatory quality, the rule of law, unemployment rate, current account balance, and general government debt to GDP are all economically and statistically significant determinants of Moody's ratings. Out of these, regulatory quality and the rule of law are proxies for institutional quality. A one standard deviation higher percentile rank of regulatory quality is associated with a 3-notch higher sovereign credit rating. Similarly, a one standard deviation higher percentile rank in the rule of law is associated with a 2-notch increase in rating. A one standard deviation increase in the unemployment rate and general government debt to GDP will lead to a one-notch increase in ratings. Similarly, a one standard deviation increase in current account balance (a current account surplus) will lead to a 1-notch higher rating. Real GDP is also a determinant but is significant at only the 1 percent level and is small in magnitude. Like the previous results for Fitch and S&P, this regression also highlights the importance of institutional quality in determining sovereign credit ratings. An improvement in institutions can positively impact ratings.

Fig 14 shows India's predicted ratings using the OLS regression specification. The thick red line is the actual Moody's rating for India and the dark blue line is the predicted rating. The light blue line is this predicted rating without GDP per capita. The dashed lines show uses predicted rating with estimated data (back-filled when some variables are missing for later years). This regression under-predicts India's ratings starting 2004. For instance, in 2015 India's actual rating was a Baa3, but the regression predicts it as a Ba2, 2 notches lower. As noted before, Moody's methodology involves several qualitative factors and proprietary data. This information may not be captured by our quantitative metrics, leading to an under-prediction based on quantitative factors but a subsequent upward adjustment based on qualitative factors.

As before, 15 plots the predicted and actual ratings for across countries in 2016. We

Table 3: Moody's

This table presents the results of the regression for Moody's. The independent variables used are as mentioned in the Moody's (2019) report. Variables are as defined in Section 3.1.2. Columns (1) and (2) present results using an OLS and an ordered probit, respectively. Columns (3) and (4) present the results for high-income and low/middle-income countries, respectively. Standard errors are clustered at country-level and all variables are standardized for ease of interpretation.

	(1) OLS	(2) Ordered Probit	(3) High – Income	(4) Low/Middle – Income
Regulatory Quality	3.219***	1.431***	4.208***	2.900***
	(0.493)	(0.225)	(1.441)	(0.556)
Rule of Law	1.903***	0.690***	1.261	1.237
	(0.511)	(0.224)	(1.368)	(0.735)
Unemployment Rate	-1.061***	-0.424***	-1.788***	-0.548
	(0.306)	(0.136)	(0.568)	(0.342)
Current Account (% of GDP)	0.975***	0.390**	0.688	1.353**
	(0.341)	(0.163)	(0.442)	(0.515)
General Govt Debt to GDP	-0.773**	-0.371**	-0.697**	-1.011**
	(0.340)	(0.148)	(0.333)	(0.448)
Real GDP Growth*	0.135	0.111	-0.916*	0.842
	(0.430)	(0.187)	(0.490)	(0.583)
GDP per Capita (PPP)**	0.629*	0.390*	0.419	1.454
	(0.343)	(0.210)	(0.405)	(1.177)
No. of Obs.	1250	1250	661	589
R squared	0.802		0.655	0.546
Pseudo R squared		0.284		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 13: Coefficient Plot: Moody's

This figure presents the coefficient plot of the OLS regression for Moody's. Variables are as defined in 3.1.2.

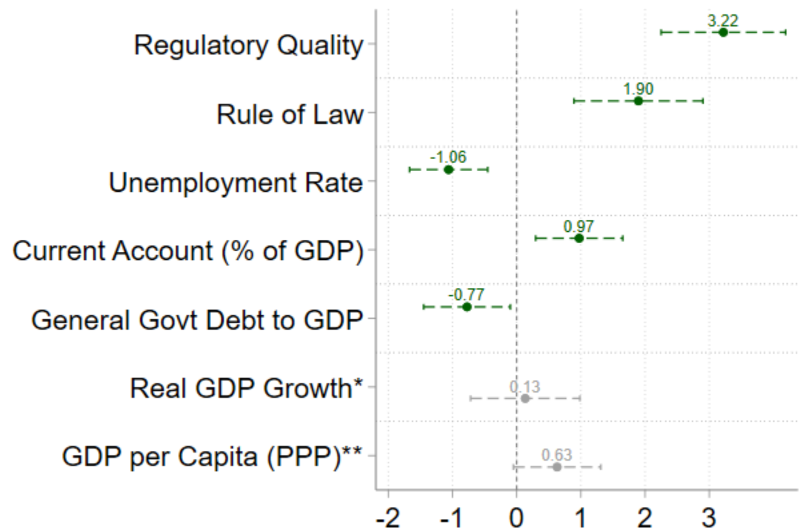
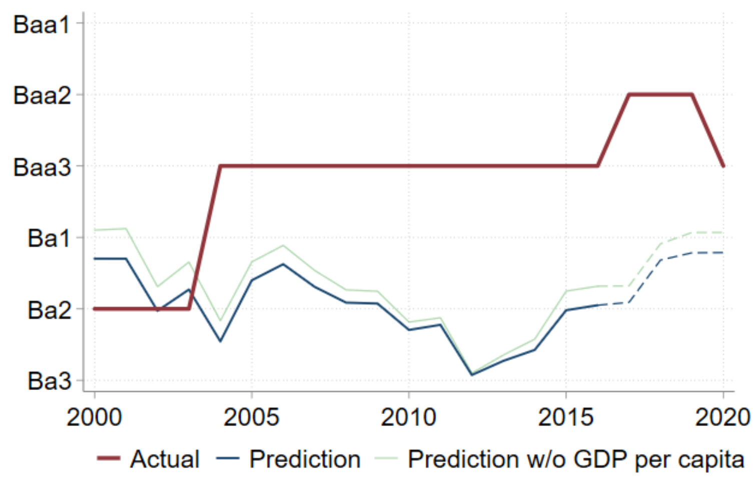


Figure 14: Predicted Ratings for India: Moody's

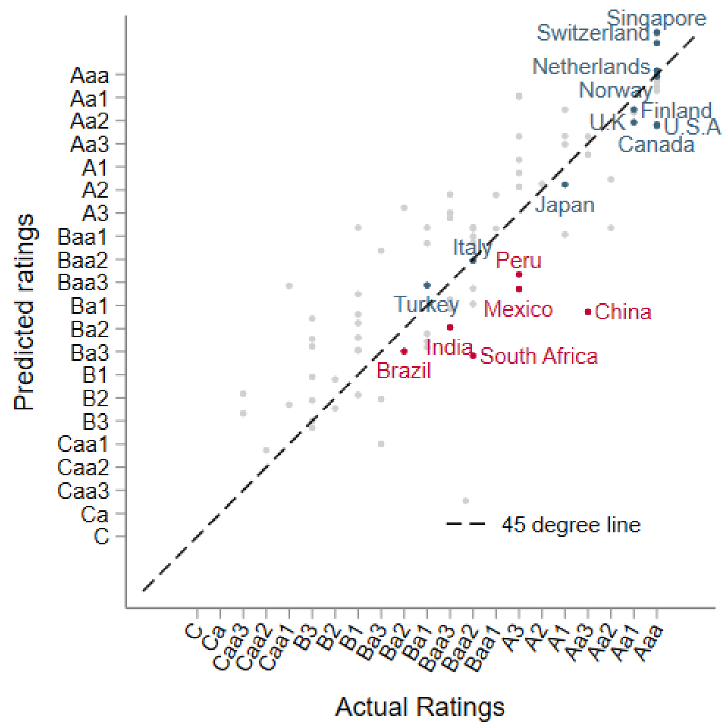
This figure presents the predicted ratings for India using the OLS regression for Moody's. Dashed lines represent rating prediction with estimated data.



see that while the quantitative factors are able to predict Moody's rating for advanced economies (shown in blue), there is significant under-prediction for India and its peers (shown in red). We can see that this model does a good job of predicting ratings for advanced economies as they mostly lie on or close to the 45 degree line. Emerging economies on the other hand lie under this 45 degree line, thus giving some evidence to the hypothesis that all these economies get the benefit of some soft information in the actual rating.

Figure 15: Predicted vs Actual: Moody's

This figure presents a scatter plot of predicted versus actual ratings for the year of 2016, using a OLS regression. Countries in red are EMEs while countries in blue are developed economies.



4.2 Common Model

The analysis in the previous section analyzed the main quantitative factors that were driving the CRA ratings. We also argued that while Fitch relies on more quantitative factors, S&P, and to a larger extent Moody's examines qualitative factors, especially for the EMEs, to adjust ratings upwards.

We now use a common regression specification to examine whether there are differences in the way each of these factors, in effect, determines the CRA ratings. While the previous section focused on the individual *quantitative* factors that drive CRA ratings, in this section we examine the relative importance of each factor in the model. The common model allows comparison across CRAs and determines importance of the factors that can enter the CRA model either as quantitative or qualitative inputs in the CRA methodologies. We use two approaches to examine commonality across CRA rating models: (i) a regression analysis using a common set of variables and (ii) a Principal Component Analysis (PCA) using all available variables.

4.2.1 Regression Framework

In the first method for the common model, we use variables that are are common across the three CRAs and regress them individually on each of these ratings. Our regression specification is as follows :

$$\begin{aligned}
\mathrm{Y}_{kit} = \alpha_{it} &+ \beta_1 * \text{Governance Indicators}_{it} + \beta_2 * \text{Ln(GDP per Capita)}_{it} + \\
&\beta_3 * \text{Ln(Broad Money)}_{it} + \beta_4 * \text{Years Since Default}_{it} + \\
&\beta_5 * \text{General Government Debt to GDP}_{it} + \beta_6 * \text{Current Account Balance}_{it} + \\
&\beta_7 * \text{GDP growth rate}_{it} + \beta_8 * \text{Inflation}_{it} + \beta_9 * \text{Net FDI}_{it} + \\
&\beta_{10} * \text{Interest Payments}_{it} + \beta_{11} * \text{Fiscal Balance}_{it} + \epsilon_{it}
\end{aligned} \tag{2}$$

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 standardized and errors are clustered at the country-level. Governance Indicators$_{it}$ refers to the average of country 'i' 's percentile rank across the six World Governance Indicators for year 't'. The World Governance Indicators are Voice and Accountability, Rule of Law, Political Stability, Government Effectiveness, Control of Corruption and Regulatory Quality. All other variables are as defined in previous sections.

Figure 16: Common Regression: All countries

This figure presents the coefficient plot using the common model for Fitch, S&P and Moody's. All coloured bars are significant at the 5% level. Fitch R-Squared = 0.85; S&P R-Squared = 0.84; Moody's R-Squared = 0.81

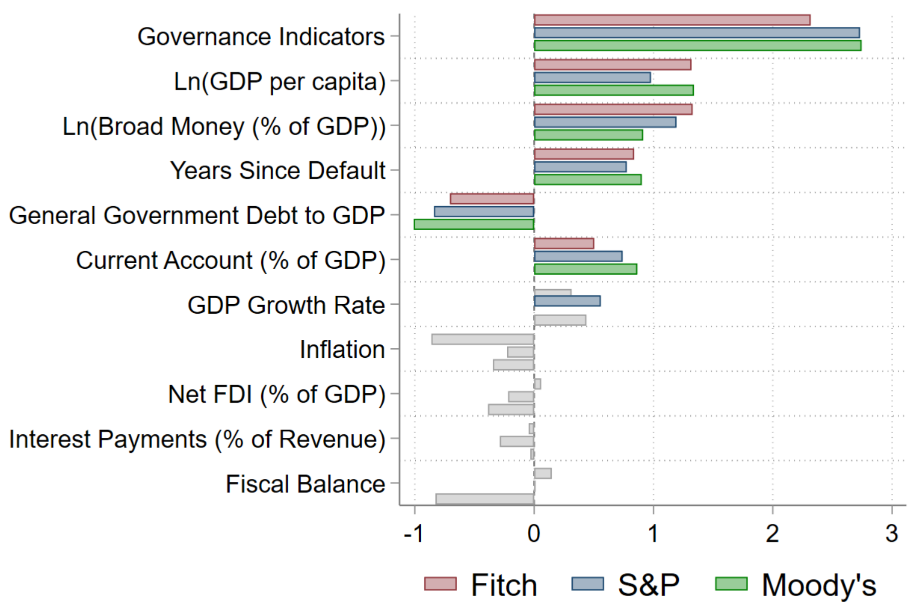


Table 4: Combined

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. Variable definitions can be found in 4.2.1. Standard errors are clustered at country-level and all variables are standardized for ease of interpretation.

	OLS			Ordered Probit		
	(1) Fitch	(2) S&P	(3) Moody's	(4) Fitch	(5) S&P	(6) Moody's
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)
No. of Obs.	802	915	694	802	915	694
R squared	0.842	0.817	0.853			
Pseudo R squared				0.318	0.286	0.329

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results of the regression specification in Equation 2 are shown in Table 4. Fig 16 shows the corresponding coefficient plots for easy readability and comparison. Columns 1–3 in Table 4 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

Figure 17: Common Regression: High Income

This figure presents the coefficient plot using the common model for only high-income countries, as defined by the World Bank's classification. All coloured bars are significant at the 5% level.

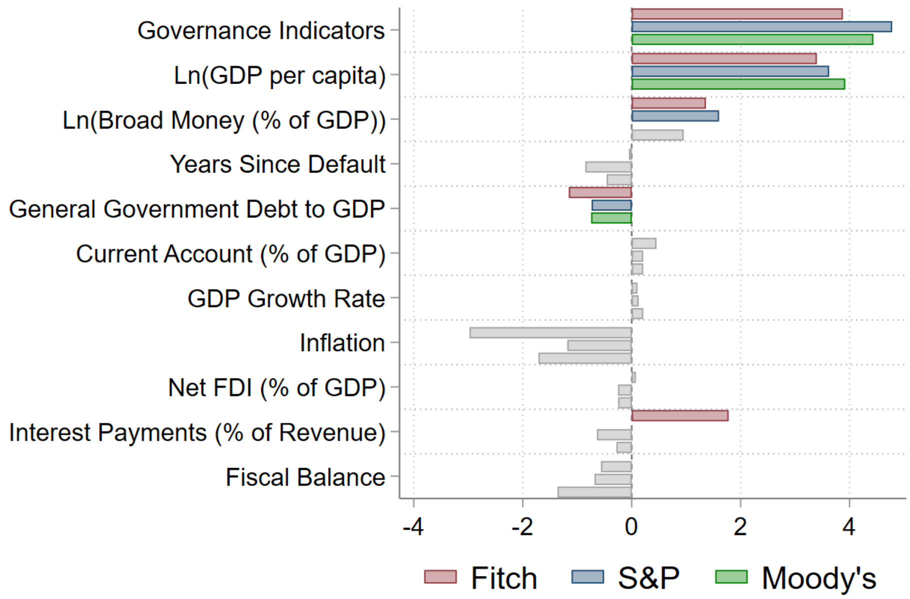
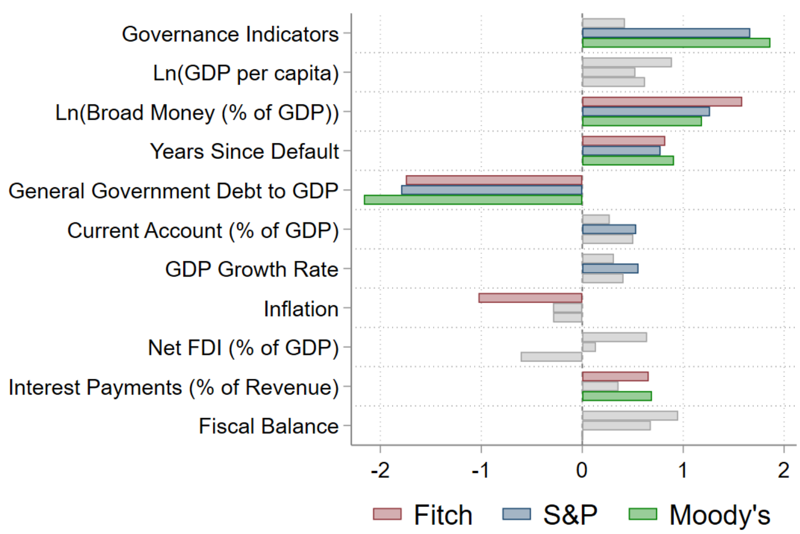


Figure 18: Common Regression: Low/Medium Indcome

This figure presents the coefficient plot using the common model for only low/medium-income countries, as defined by the World Bank's classification. All coloured bars are significant at the 5% level.



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 and consistent with the previous analysis in Section 4.1. A 1 SD higher governance indicators leads to a 2-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 1-notch higher rating. Similarly, a 1 SD higher debt to GDP ratio is associated with a one-notch lower rating. Columns 4–6 in Table 4 present the results using an ordinal probit and results are robust to this alternate specifications, consistent with prior literature (Fuchs and Gehring, 2017).

Figures 17 and 18 also show the results of Equation 2 separately for the sample of high-income 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 in Section 3, 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. We will discuss whether reliance on these factors is justified in Section 5.

4.3 Principal Component Analysis

We next turn to a principal component analysis (PCA). We have a number of independent variables across the three agencies, and the PCA helps us in effectively summarizing the drivers of CRA ratings. The PCA reduces the dimensionality of the data by expressing all of the data in indices known as the "principal components". Each principal component is a linear combination of the independent variables, where the first component accounts for the largest possible variance. In other words, the first principal component contains most of the information contained across all the variables. Within each component we can also examine how all variables load on the first principal component to get a sense of where the variation in the data is coming from. PCA is also a useful tool in our case as our variables all have different units and scales, allowing us to easily succinctly determine how various pillars load into the CRA models.

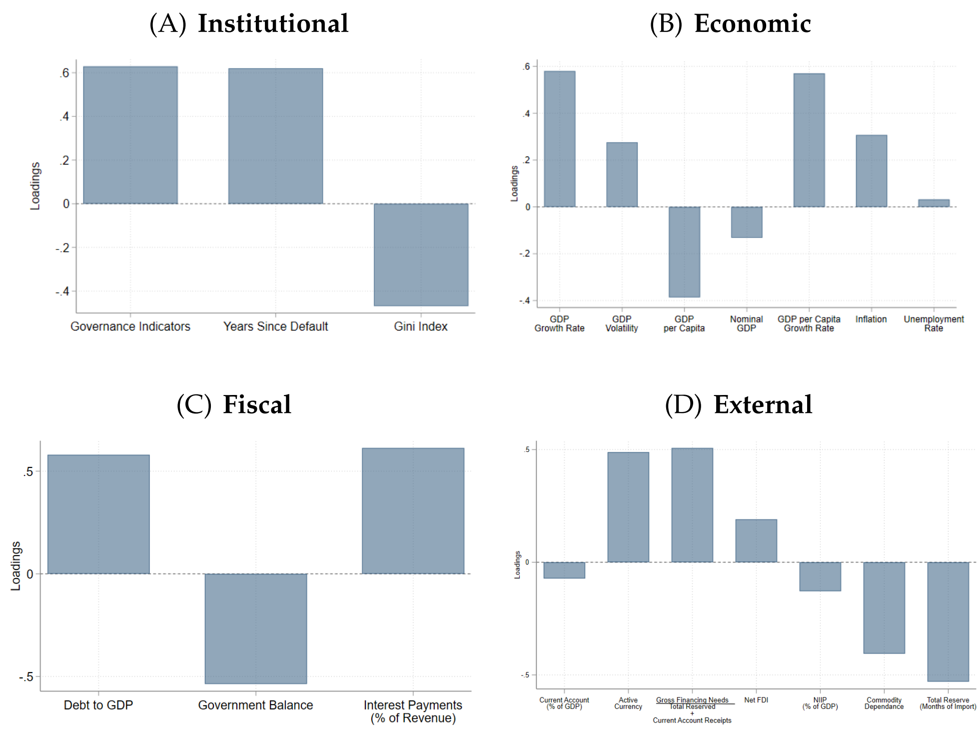
To conduct the PCA analysis, we do as follows. First, we narrow down to the universe of all independent variables available to us. In case of variables that were aggregated across time periods, we dis-aggregate it and use the simplest definition. We then divide these variables into four pillars: Institutional, Fiscal, Economic and External. We then conduct a PCA within each pillar and retain only the first principal component, giving us four first principal components for each of the pillars. We ensure that the first principal component has an eigenvalue greater than 1. Finally, we run a regression of the four principal components on the CRA ratings using the specification below:

$$\begin{aligned} \mathrm{Y}_{kit} = \alpha_{it} \ + \beta_1 * \text{Institutional PC-1}_{it} \ + \beta_2 * \text{Fiscal PC-1}_{it} + \\ \beta_3 * \text{Economic PC-1}_{it} \ + \beta_4 * \text{External PC-1}_{it} \ + \epsilon_{it} \end{aligned} \tag{3}$$

Figure 19 shows the loading of the first component within each pillar. Panel A shows the Institutional pillar. Governance Indicators and Years Since Default both affect institutions positively, while the Gini Index has a negative effect. This is in the expected direction, as an improvement in governance indicators as well as not defaulting for many years, signal healthy institutions.

Figure 19: PCA - Loadings

This figure represents the loadings of the first principal component for each of the 4 pillars - Institutional, Economic, Fiscal and External.



Panel B shows the Economic pillar. Here, GDP growth rate, GDP volatility, GDP per capita growth rate, unemployment and inflation all load positively, while GDP per capita and nominal GDP load negatively. Panel C shows the Fiscal pillar. Debt to GDP and Interest Payments as a percentage of Revenue load positively, while government balance loads negatively on the fiscal pillar. Panel D shows the External pillar. Here, active currency, gross financing needs as a percentage of total reserves and net foreign direct investment all load positively. Current account as a percentage of GDP, Net IIP as a percentage of GDP, commodity dependence and total reserves all load negatively. For ease of interpretation, the loading on the economic, fiscal and external pillars are multiplied by -1, so that a positive loading indicates a positive change on CRA (consistent with the principal component for institutional quality).

Table 5: PCA - Regression

This table presents the results using a PCA. Details on the PCA methodology can be found in 4.3. Columns (1), (2) and (3) present results for S&P, Moody's and Fitch, respectively. The first principal components on the Fiscal, Economic and External pillars were multiplied by -1 so that positive loadings correspond to improvements.

	(1) S & P	(2) Moody's	(3) Fitch
Institutional	3.378***	3.383***	3.091***
	(0.348)	(0.342)	(0.360)
Fiscal	0.752***	0.733***	0.745***
	(0.218)	(0.218)	(0.215)
Economic	0.263	0.410*	0.377*
	(0.171)	(0.212)	(0.202)
External	0.371	0.490	0.240
	(0.347)	(0.333)	(0.324)
N	648	667	609
R squared	0.743	0.728	0.745

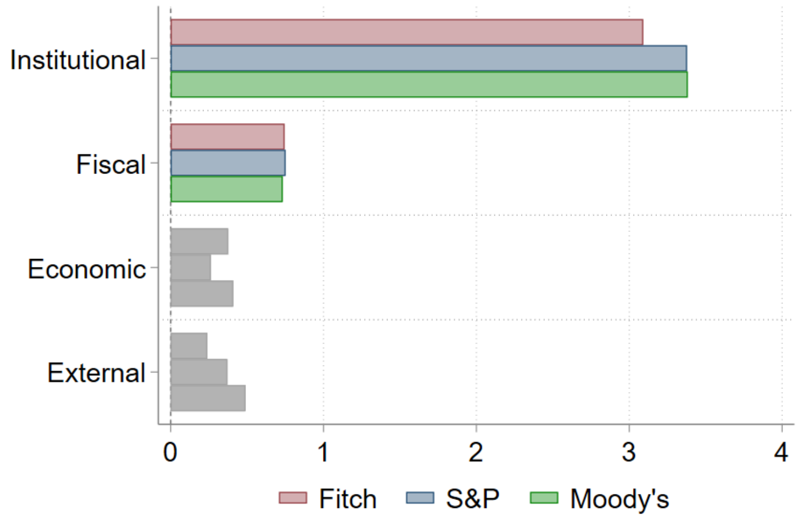
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 presents the results using the specification in Equation 3. Columns 1–3 show that out of the 4 pillars, the institutional and fiscal pillars are significant and have the highest positive effect on ratings. These effects are similar across the three agencies. A one standard deviation improvement in institutions leads to a 3-notch higher rating. Similarly a one standard deviation improvement in fiscal factors, leads to a nearly 1-notch higher rating.

Fig 20 presents the corresponding coefficient plot for easy comparison. Institutional factors are the single most important determinant of CRA ratings followed by fiscal health. In terms of magnitude a 1 standard deviation increase in institutional quality has a far greater impact on CRA ratings compared to the remaining pillars. Strikingly, the remaining pillars, as indicated by the coefficients on the first principal component of Economic factors and External factors, have a more limited influence on credit ratings. The point estimates on these coefficients are smaller in magnitude and are statistically insignificant.

Figure 20: Principal Component Analysis

This graph shows the coefficient plot for 3. All coloured bars are significant at the 5% level, while the grayed out bars are insignificant. The loading on the Fiscal, Economic and External pillar are flipped so that a positive change corresponds to a ratings increase.



5 Discussion of Determinants of CRA Ratings

We now examine important factors in detail with a focus on how much the CRAs rely on these factors and whether it is justified. The five important factors we focus on are: (i) institutional quality, (ii) government debt to GDP, (iii) broad money, (iv) GDP per capita, and (v) GDP growth.

5.1 Institutional Quality

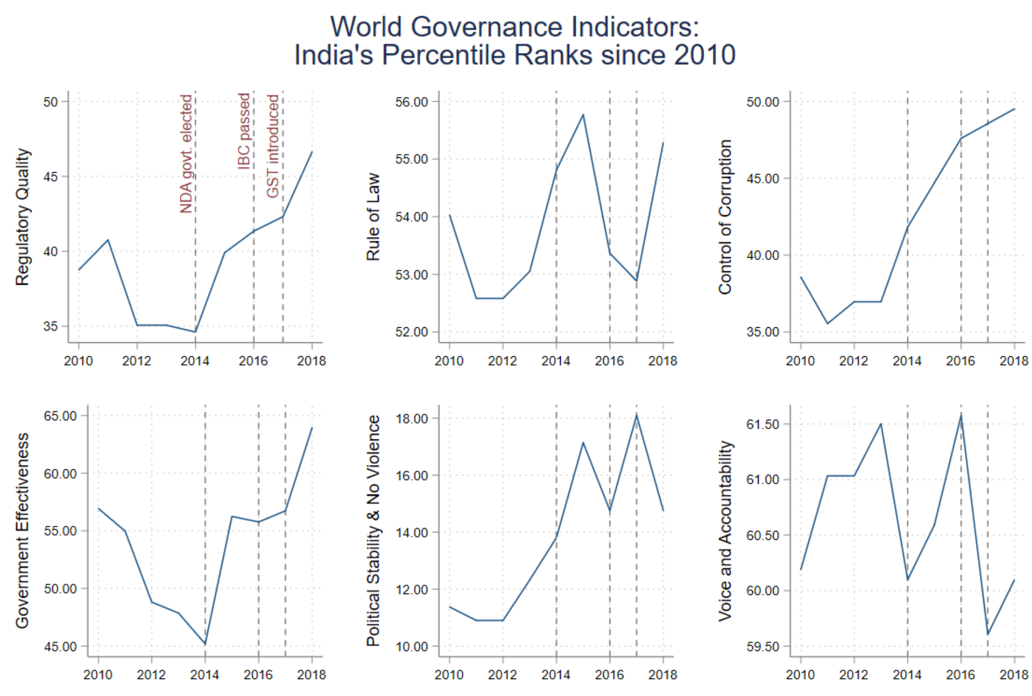
We have seen that Institutional quality is the most important variable in the regressions of all three agencies as well as in the combined regression. While some specifications use only some elements of institutional quality (such as Rule of Law, Regulatory Quality), some use all. For Fitch, we use the mean of the 6 World Bank Governance Indicators (WGI) in the quantitative model. For S&P, we use the Transparency in Policymaking indicator as well as the WGI. Finally, for Moody's we use two of the WGI's sub-indicators - Regulatory Quality and Rule of Law.

All the variables that we use to represent institutions thus come from the World Governance Indicators (WGI) and the World Economic Forum. These governance indicators capture governance perceptions as reported by survey respondents, NGOs, commercial business information providers and public sector organizations worldwide (Kraay, Kaufmann and Mastruzzi, 2010). The Voice and Accountability measure from the WGI also includes a democracy index in it's calculation. As a result, on average, democratic countries are 5 times more likely to get a higher rating relative to countries with authoritarian regimes. Thus, by default, CRAs tend to rate democratic countries higher as also emphasized in prior literature (Archer, Biglaiser and DeRouen Jr, 2007).

We also examine these indicators to see how responsive they are to Indian policy changes. Figures 21 show that the six governance indicators do respond to important political and policy developments in India, and are thus largely reliable measures. For example, we find our percentile ranking for regulatory quality went up both when the Insolvency and Bankruptcy Code (IBC) was passed in 2016, as well as when the General Service Tax (GST) was introduced in 2017.[8] In the common regression, as seen in Fig 16, a 1 standard deviation increase in the percentile ranks for governance indicators leads to a nearly 3 notch increase in ratings. Similarly, we see in the PCA that the Institutional component has the largest impact on ratings.

Figure 21: World Governance Indicators: India

This graph represents the trends in the World Governance Indicators for India's percentile rank. The subraphs are for each of WGI's sub-indicators - Regulatory Quality, Rule of Law, Control of Corruptions, Government Effectiveness, Political Stability and Voice and Accountability.



We now focus on a different question. Is institutional quality a good predictor of

[8]The IBC was a bankruptcy law passed to reform the process of resolving insolvencies and bad loans. The GST was a law passed to subsume all the other indirect taxes in India, thus attempting to unify the tax code.

default? To test the importance of a particular variable in predicting long-term default we use the following specification:

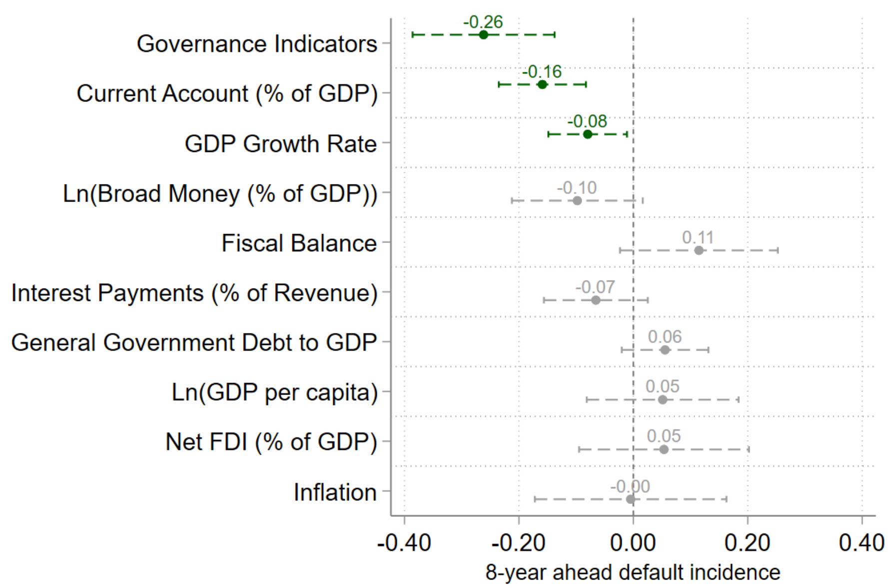
$$\begin{aligned}
\mathrm{Y}_{it} = \alpha_{it} &+ \beta_1 * \text{Governance Indicators}_{it} + \beta_2 * \text{Ln(GDP per capita)}_{it} + \\
&\beta_3 * \text{Ln(Broad Money)}_{it} + \beta_4 * \text{General Government Debt to GDP}_{it} + \\
&\beta_5 * \text{Current Account Balance}_{it} + \beta_6 * \text{GDP growth rate}_{it} + \beta_7 * \text{Inflation}_{it} + \\
&\beta_8 * \text{Net FDI}_{it} + \beta_9 * \text{Interest Payments}_{it} + \beta_{10} * \text{Fiscal Balance}_{it} + \epsilon_{it}
\end{aligned} \tag{4}$$

where,Y_{it} measures 8-year ahead default incidence, where it takes a value 1 if for year 't' country 'i' has defaulted in the next 8 years, and 0 otherwise. We use this variable as a measure of long-term default. We also measure short-term default using the 1-year ahead default incidence. All variables are standardized and errors are clustered at the country-level. Governance indicators is the same as defined in Equation 2.

The results of Equation 4 using the long-term incidence of default are shown in the coefficient plot in Fig 22. The results of Equation 4 using the short-term incidence of default are shown in the coefficient plot in Fig 23.

Figure 22: Long-Term Incidence of default

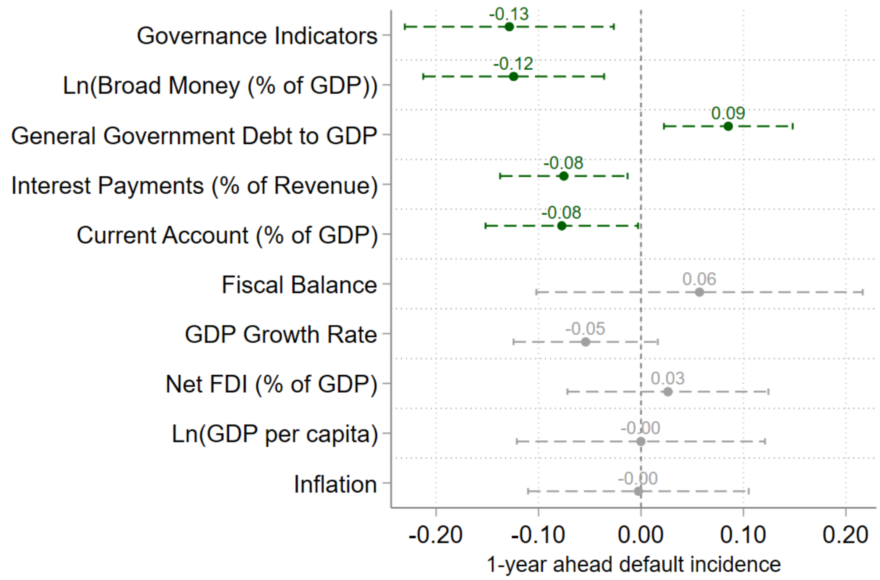
This graph presents the results of the regression for Equation 4. Here, long-term refers to 8-year ahead default. All variables are significant at the 5% level and are standardized for ease of interpretation.



A 1 standard deviation lower Governance Indicators is associated with a 0.26 higher default incidence. Plausibly, good institutions are a good indicator of whether a sovereign is able to stave off default incidence and hence, CRA reliance on institutional quality is justified. Good institutions are associated with quick and efficient law enforcement, a robust political class, transparent budget accounts, all of which are good predictors of a country being able to repay its debt in time. This finding also suggests that investment in building strong independent institutions is necessary to increase sovereign borrowing capacity as it is an indicator of a sovereign's creditworthiness.

Figure 23: Short-Term Incidence of default

This graph presents the results of the regression for Equation 4. Here, short-term refers to 1-year ahead default. All variables are significant at the 5% level and are standardized for ease of interpretation



5.2 Government Debt to GDP

General Government Debt to GDP is significant and an important variable in all three separate regressions for the CRAs in Section 4.1, and also in the common regression. Government debt (including bank debt) accumulates very quickly, and is often unchecked by markets. According to Reinhart and Rogoff (2009), this debt is a large cause of financial crisis and episodes of default. According to traditional economic literature, public debt is only thought of as having a long term tax burden. That is, if the government borrows a lot today, the burden of repayment falls on future time periods, thus taxing the long term. However, Reinhart and Rogoff (2009) stress that large debts can also hurt countries in the short term if the market thinks that the government will not be able to finance the debt in the long term. Greater transparency in government budget accounts, including clear records of government guaranteed debt, are good signals for institutional quality. Government guaranteed debt is often used to keep debt off from government books, and can ultimately lead to crises. Reinhart and Rogoff (2009) make a case for keeping debt low for extended periods while governments undertake structural reforms.

We next examine whether high levels of government debt are a good predictor of near-term and long-term default. Fig 23 shows that a 1 standard deviation rise in general government debt to GDP leads to a 0.09 rise in default incidence in the short-term. However, Fig 22 suggests that higher general government debt to GDP is not as strong a predictor of long-term default. Together, this analysis suggests that high levels of government debt implies that sovereign default is imminent in the near-term.

5.3 Broad Money

The ratio of broad money to GDP is used as a proxy for the level of financial intermediation in the economy. The richer the economy, the more monetary assets it holds which could indicate the amount of public debt that the economy can tolerate. As broad money covers highly liquid securities such as treasury bills, commercial papers, certificate of deposits, all domestic and foreign checkable, time and savings deposits apart from currency in circulation, it represents the broad monetary wealth present in the sovereign. Therefore comparing it with Gross General Government Debt can indicate how the sovereign's creditworthiness compares to its internal and external borrowing. Fitch (2020*b*) uses this variable under the structural features of sovereigns, and it is prominently significant in both the CRA-level analysis as well as the common analysis across all CRAs. In the one-year ahead incidence of default regressions the coefficient on broad money is the second-highest in magnitude among all regressors (in Fig 23). A one standard deviation increase in broad money to GDP ratio reduces incidence of near-term default by 12%. However, the coefficient is insignificant at the five-percent level for long term (8-years ahead) default incidence as seen in Figure 22.

5.4 GDP per Capita

Another determinant that is statistically significant in the common regression (Figure 16) and in the regression for S&P (Figure 10) is GDP per capita. Previous commentators have also questioned the over-reliance of CRAs on GDP per capita, which is an extremely slow moving variable and severely biases sovereign ratings downwards for low/middle-income countries. In the case of India, despite its strong fundamentals, a low GDP per capita can explain its low ratings.[9] In Reinhart and Rogoff (2009), countries that attain

[9] A report by the Government of India (Government of India, 2017) makes the claim that: "Lower middle income countries experienced an average growth of 2.45% of GDP per capita (constant 2010 dollars) between 1970 and 2015. At this rate, the poorest of the lower middle income countries would take about 57 years to reach upper middle income status. So if this variable is really key to ratings, poorer countries

and maintain their investment grade status, are countries that are defined as 'graduating countries'. One of the measures used to define this graduation could be per capita income. Primo Braga and Vincelette (2010) also find that GDP per capita is a huge risk factor in episodes of distress.

Our analysis in 22 and 23 suggests that GDP per capita is only a noisy determinant of both short-term and long-term default. This calls into question whether CRAs should focus so heavily on GDP per capita in their rating methodologies.

A related point is that India is an outlier in terms of default compared to its peers. Fig 24, panel (A) make this point by examining default incidence and GDP per capita. Panel (A) shows the percentage of countries defaulted in each decile of GDP per capita in recent history (post-2000). While the highest GDP per capita countries have low defaults, the relationship breaks down as we move to lower deciles. Thus, plausibly CRA reliance on GDP per capita may not be fully justified.

We make this point with an alternate analysis in panel (B). We plot the mean years since default for each decile of GDP per capita. India's years since default is overlaid in the third decile and shown in pink. This graphical illustration allows us to compare India with its peer group of countries. When viewed this way, it does seem that there is a steady increase in years since default as FDP per capita rises. For India's peer countries (based on GDP per capita decile), nearly 50% of the countries have defaulted at least once between 2000–2012. However, India's mean years since a default is much higher than the mean of its GDP per capita decile. The years since default is calculated based on data from Reinhart and Rogoff (2009) on sovereign external debt defaults and restructuring.[10] Countries that have similar GDP per capita as India have defaulted more that India has. Thus India's ratings and perception of default is likely driven by its peers. The evidence suggests that GDP per capita is a bad predictor of default despite its outsize role in CRA rating models.

5.5 GDP Growth

We consistently find that GDP growth is not an important determinant of CRA ratings, despite its anecdotal importance on ratings. However, in 2017, when India's rating was

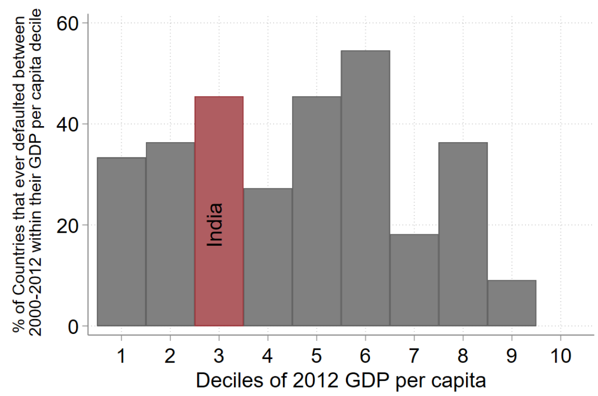
might be provoked into saying, Please don't bother this year, come back to assess us after half a century."

[10] Lindert and Morton (1989), Babbel and Bertozzi (1996) and Reinhart and Rogoff (2009) cite World Debt Tables of the World Bank as evidence of restructuring of official loans to India from 1969 to 1976. Gulhati (1972) states that from FY1968-69 to FY1971-72, India received a $300 million debt relief from a World Bank Led consortium enabling it to postpone nearly 1/5th of its scheduled debt service for these years. In subsequent years till 1976, the Aid-to-India consortium, a group of thirteen creditor nations led by the World Bank provided an annual debt relief of more than $100 million each year. These were $153 million in 1972, $187 million in 1973, $194 million in 1974, $228 million in 1975 and $169 million in 1976.

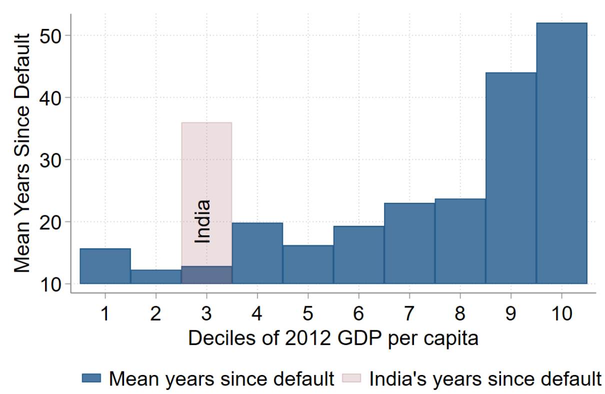
Figure 24: Default and GDP per Capita

In this graph all bars represent GDP per capita deciles (based on 2012 data). India lies in the third decile and is highlighted in both panels. The y-axis is default incidence for (A) and years since default for (B).

(A) **Default Incidence by GDP per capita decile**



(B) **Years Since Default by GDP per capita decile**



upgraded by Moody's from Baa3 to Baa2, its main rationale was that "Reforms will foster sustainable growth" [11]. The word "sustainable" is key to understanding the role of GDP growth in increasing ratings. Indeed, the insignificant role of GDP growth in predicting CRA ratings can be explained by the way CRAs adjust for GDP growth. For all the three CRAs, GDP Growth enters as a base variable in calculating the initial rating. However, all agencies later make adjustments to the score based on GDP growth based on whether this is "good" or "bad" growth. Fitch adds an adjustment if a country does well compared to its peers. S&P makes a *negative* adjustment to ratings if growth is fueled by "depository corporation claims on the resident non-government sector", that is, if growth is fueled by the unproductive household sector. Moody's makes adjustments to

[11]See: https://www.moodys.com/research/Moodys-upgrades-Indias-government-bond-rating-to-Baa2-from-Baa3–PR374998

the initial score based on the sustainability of GDP growth. It considers factors such as investment in education, infrastructure, female labour force participation, flexible labour market laws, export diversification, a structural break in the economy, etc, and according makes upward or downward adjustment. Thus, only GDP growth fueled by investment in productive capacity or structural factors enter positively into the CRA ratings. Despite popular rhetoric that focuses on GDP growth, unsustainable GDP growth is viewed negatively by the CRAs explaining its low explanatory power in the previous analysis. This also suggests that investment in structural factors as opposed to focus on short-term GDP growth numbers may be more useful in promoting investor confidence and increasing CRA ratings.

6 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 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 ahead and five-year ahead incidence of default, against levels of Fitch's sovereign credit ratings. Next, we analyze how well rating levels as well as rating actions predict default during normal times and in periods of crisis.

6.1 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-term 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*b*). In fact short-term ratings have much less granularity for corresponding rating-levels, and any discrepancies - different short-term and long-term ratings - are rare. In Fig 25(A), we plot the Fitch rating levels along 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+. On the other hand sovereigns with ratings above A- see no one-year ahead defaults or debt-restructuring events.

Figure 25(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. We see that 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: We examine how rating agencies perform in predicting default, comparatively and inter-temporally, during crisis and non-crisis periods.

6.2 Default Prediction During Periods of Crisis

How good are rating changes in predicting near-term and long-term sovereign default? In Section 2 we argued that 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 biases or arbitrariness, can trigger self-fulfilling prophecies, driving even relatively healthy countries to default (Gärtner, Griesbach and Jung, 2011).

This motivates us to study how well do sovereign credit rating changes, especially during crisis periods, predict defaults. We use a simple regression of near-term 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 (2000-2006), the global financial crisis (2007-2009) and the European sovereign debt crisis (2010-2014). We use the following specification:

$$Y_{it} = \alpha + \beta \times X_{it} + \epsilon_{it} \tag{5}$$

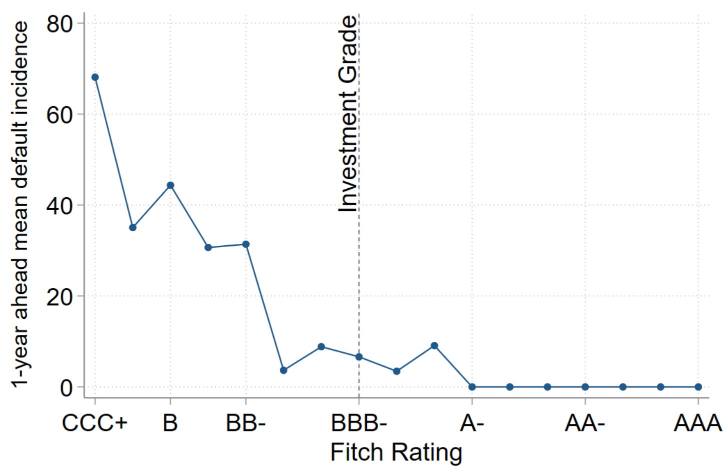
where Y_{it} is the default indicator and X is the level of credit rating or change in credit ratings

Figure 26(B) and 26(A) present the results for the level of rating against the five-year ahead and one-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

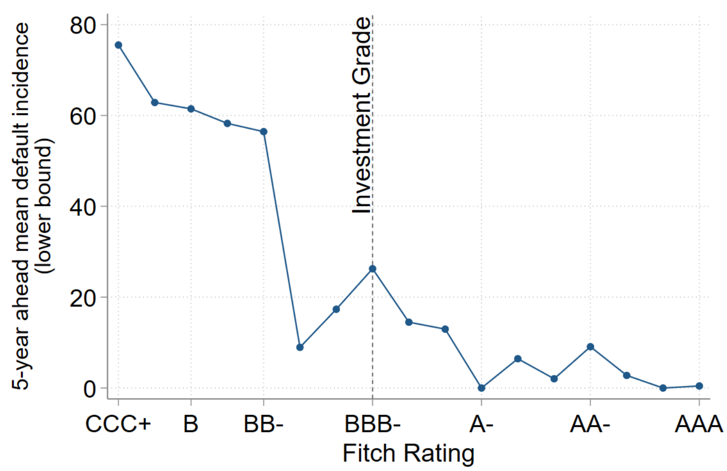
Figure 25: Default incidence against Fitch Ratings

This graphs plots Fitch ratings against 1-year and 5-year ahead default incidence for all countries. The dotted line shows the line above which all countries are considered as investment grade.

(A) **1-year ahead**



(B) **5-year ahead**



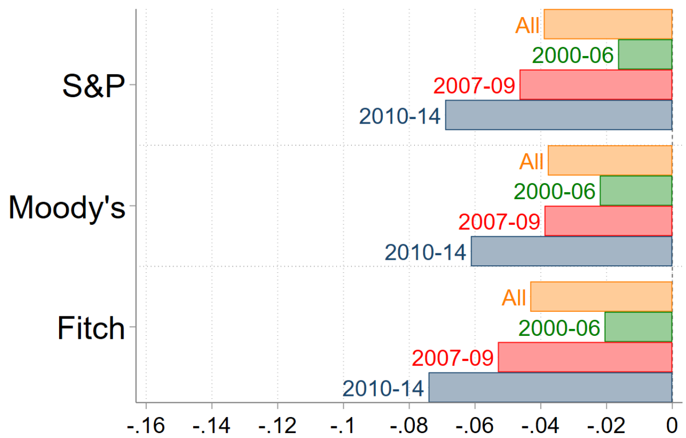
relation between rating levels and incidence of default. Figure 26(A) examines near-term, one-year ahead defaults. Sovereign rating-levels across agencies perform very similarly during the non-crisis and global financial crisis in predicting one-year ahead defaults. A one notch lower rating is correlated with a 2 percent higher default incidence during the non-crisis period and between 5 to 7 percent higher one-year ahead default probability during the financial crisis and the EU sovereign debt crisis. Rating agencies likely internalize the events during these crises which likely results in a higher coefficient for the sub-samples for the years 2007-2014.

For the five-year ahead term, as seen in Figure 26(B), the coefficients on level of ratings

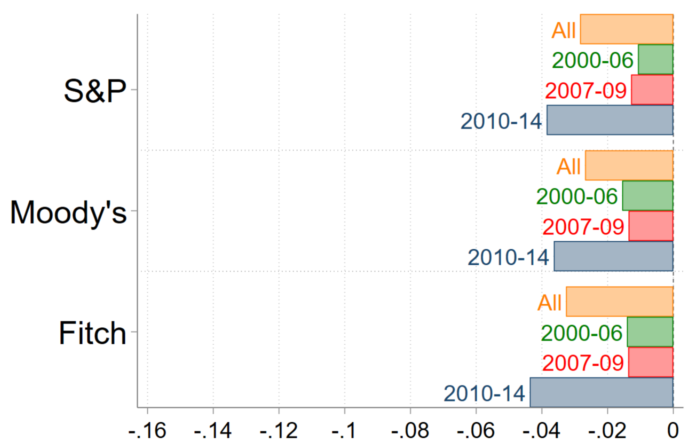
Figure 26: Rating Levels as Predictors of Sovereign Default

This graph plots default incidence against rating levels for all the three CRAs. Here 2007-09 represents the global financial crisis and 2010-14 represents the sovereign debt crisis. 2000-05 is a non-crisis period

(A) **1-year ahead**



(B) **5-year ahead**



across rating agencies are similar in scale for each corresponding period. While in non-crisis periods, a unit lower rating is correlated with a probability of default incidence in the next five years by one to three percent, during the financial crisis and the sovereign debt crisis a reduction in ratings predicted 5-year default incidence between four and five percent.

As the factors driving sovereign default change along the non-crisis and crisis periods, we analyze how well do rating actions such as downgrades, rather than rating levels, predict sovereign defaults? We regress, the near-term 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 sub-divide the samples into High income and non-High (low/middle) income economies.

Figure 27(A) represents the results for one-year ahead default incidence for all countries. Almost all agencies have statistically insignificant coefficients during both crisis periods. Figure 27(B) points that Standard and Poor performs relatively worse for low and middle income sovereigns while figure 27(C) suggests that all credit rating agencies have statistically insignificant correlation with default for high-income countries.

Figure 28(A) represents the results for all countries. The coefficient on Moody's is not-significant for both crisis periods which indicates that this rating agency performs relatively poorly in predicting medium-term default compared to its competitors. All rating agencies have an insignificant coefficient during the sovereign debt crisis. Figure 30 points out that only Fitch ratings has a significant coefficient against 5-year ahead incidence of default during the great-financial crisis 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. Figure 28(B) suggests that all rating agencies do poorly in predicting five-year ahead default among low and middle income countries, especially during periods of crises.

Figure 27: Rating Changes as Predictor of 1-year ahead Sovereign Default

This graph plots near-term default incidence against rating levels for all the three CRAs. Here 2007-09 represents the global financial crisis and 2010-14 represents the sovereign debt crisis. 2000-05 is a non-crisis period. Coloured bars represent coefficients significant at the 5% level, whereas a grey bar represents a statistically insignificant coefficient.

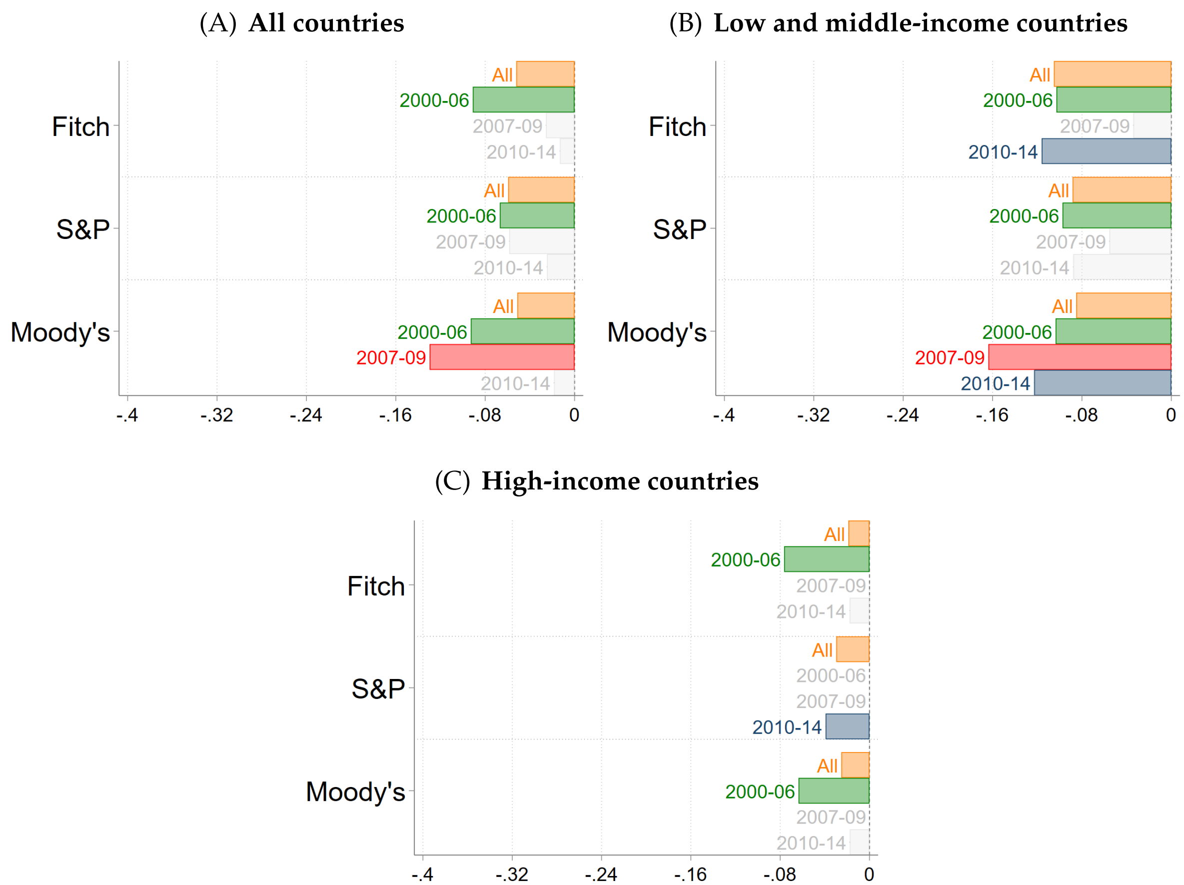
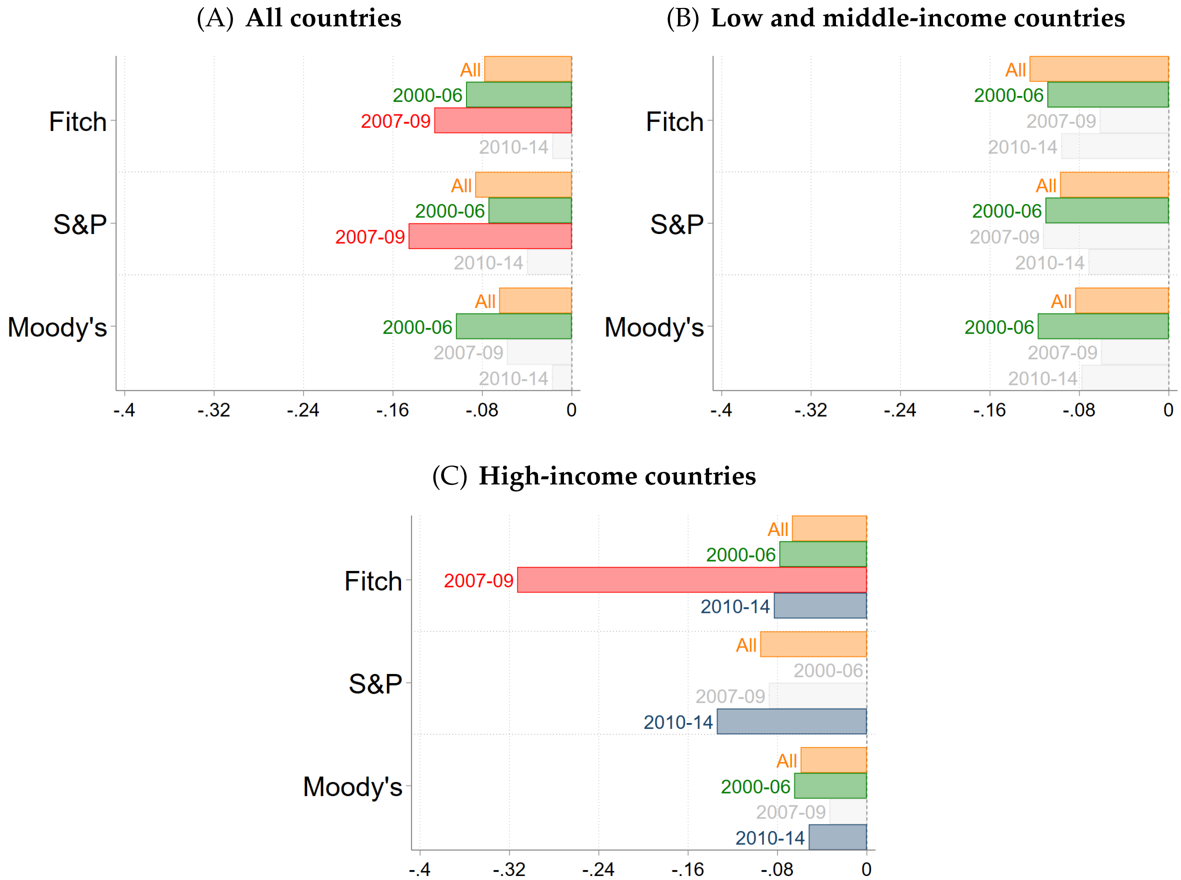


Figure 28: Rating Changes as Predictor of 5-year ahead Sovereign Default

This graph plots medium-term default incidence against rating levels for all the three CRAs. Here 2007-09 represents the global financial crisis and 2010-14 represents the sovereign debt crisis. 2000-05 is a non-crisis period. Coloured bars represent coefficients significant at the 5% level, whereas a grey bar represents a statistically insignificant coefficient.



7 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. Figure25 suggests default incidence remains flat for high rating levels and increases non-linearly as ratings go below investment-grade thresholds. Secondly, these predictions often have low statistical power due to the lack of default incidences across time in the cross-section of countries. Thirdly, 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 analyze the set of ten existing predictors used in section 5 with three new factors as proxies for financial repression, banking system vulnerability, and regional economic integration, which can also influence sovereign default probability.

7.1 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, Dovis and Kehoe (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 act as a credible commitment device. Perez et al. (2015) notes 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 behavior to recover 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 capitalization 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). Secondly, 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 1-year and 5-years ahead defaults with the set of thirteen predictors.

7.2 Random Forest Methodology

At the core of our supervised machine learning methods are regression-trees which allows 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-minimizing 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 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 29(A) and 29(B) represent the test and out-of-bag (OOB) error [12] for the number of predictors included in the random forest model of 1-year ahead and 5-years ahead default prediction respectively.

7.3 Results of Random Forest Model

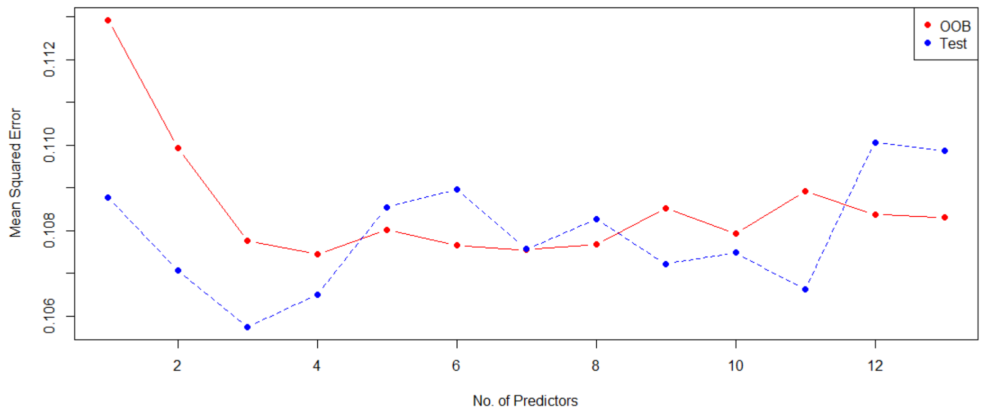
The random forest model only explains 29.81% and 45.1% of variance in one-year and five-year ahead default incidence. This is in stark contrast to the linear regression frame-

[12] The Leave-One-Out cross validation error from bootstrap aggregation. Trees are repeatedly fit to bootstrapped 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 i^{th} observation using each of the trees in which that observation was out-of-bag. This will yield around one-third of the number of sample predictions for the i^{th} observation, which we average

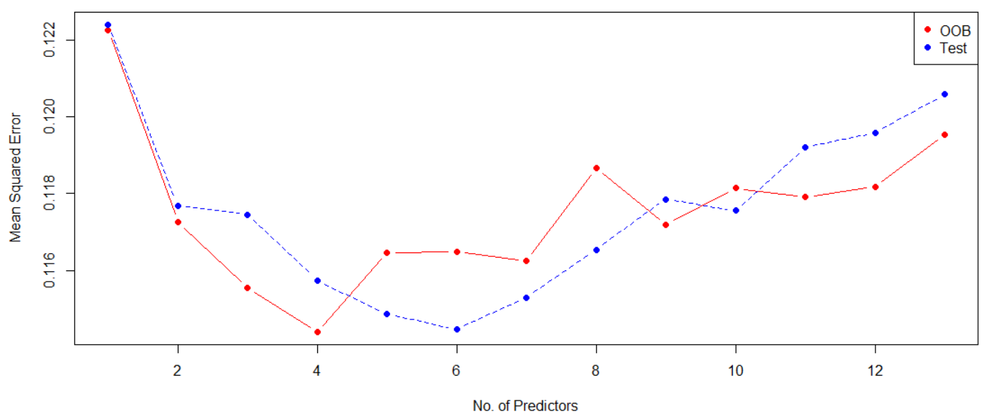
Figure 29: Random Forests: Mean squared errors of predictors

This graph plots the Mean Squared Errors of the 1-year and 5-year ahead default prediction. The red line is the Out of Bag Error and the blue line is the Test error.

(A) **MSE of 1-year ahead default prediction**



(B) **MSE of 5-years ahead default prediction**

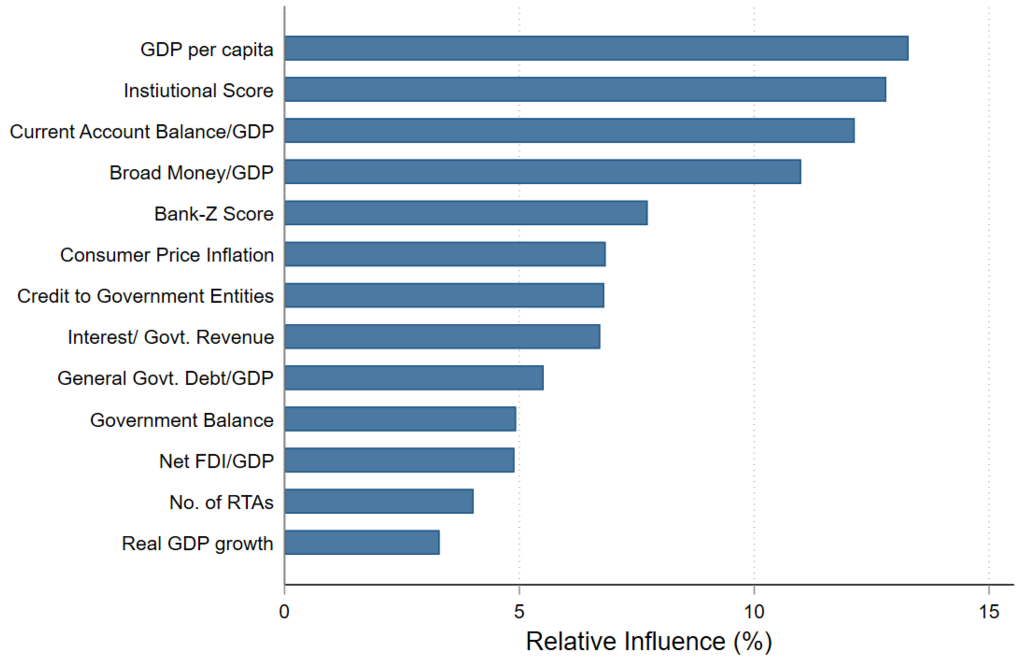


works with high R-squares (see Section 4.1) reflecting the bias in the data. This means that these thirteen fundamentals are not very powerful in *predicting* future default, but fit well with default data *ex-post*. Figure 30(A) explains the relative influence of predictors in explaining one-year ahead sovereign default incidence.

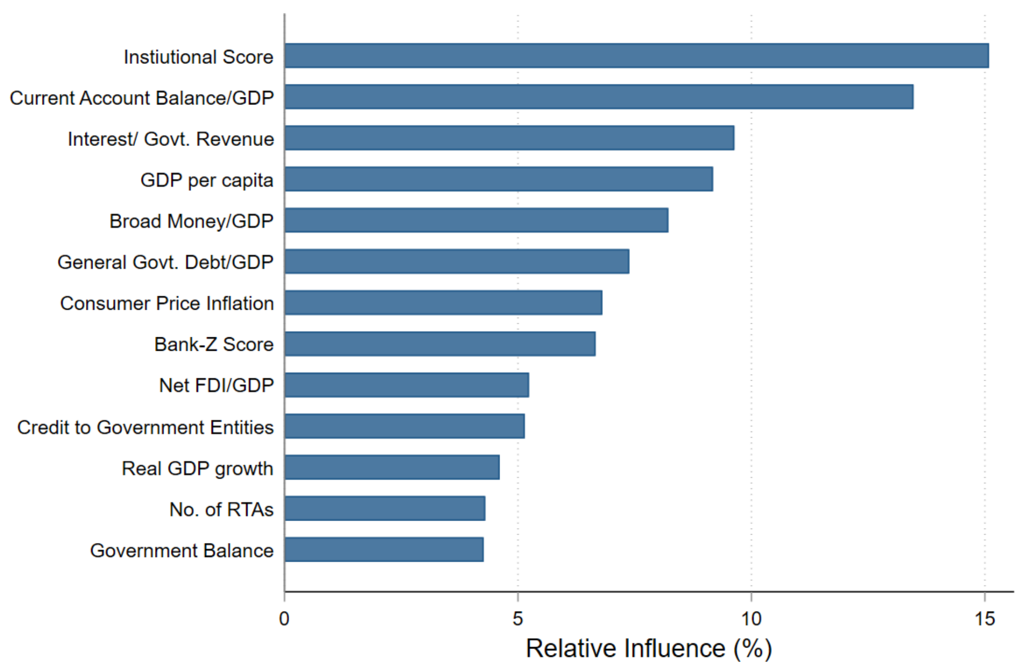
Relative influence is calculated as the amount of RSS reduction due to splits over a given predictor, averaged over all bootstrapped trees. A large value indicates an important predictor. GDP per capita, institutional score, current account balance and broad money explain more than 10% of RSS reduction in predicting short-term default. However, interestingly bank-Z score and credit to government entities are the 5^{th} and 7^{th} most

Figure 30: Relative influence of factors in predicting incidence of default

(A) **Relative Influence on 1-year ahead default**



(B) **Relative Influence on 5-years ahead default**



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 5-years ahead. Figure 30(B) gives a similar picture for five-years 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 credit rating agencies explain much less of the variation in default incidence, when adjusted for data-fitting bias (ii) 4-6 set of predictors can minimize 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), (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 5-years ahead default.

8 Conclusion

In this paper, we motivate the need to reassess CRA rating methodologies and accuracy in predicting sovereign default. Our analysis shows that Fitch uses more quantitative factors, with S&P and Moody's relying on a mix of quantitative and qualitative factors as inputs in their rating models. Subjective adjustments to the quantitative baseline score determine the final sovereign credit rating. 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 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 ahead 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.[13] Countries with weak institutional fundamentals suffer a two-fold 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 don't 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.[14] In 2010, a G20 resolution acknowledged the over-dependence on the CRAs and suggested that central banks and banks independently conduct their own rating (Financial Stability Board, 2010*a*).[15] Motivated by above, Bank of Canada produces internal sovereign credit ratings for use in its management of Canada's foreign exchange reserves.[16] Our paper makes a case for India, too, building alternative internal rating models to assess sovereign credit risk instead of relying on the CRAs.

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

[14] 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, 2010*b*).

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

[16] 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." *The Review of Financial Studies* 29(7):1709–1746.

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* 29(3):582–609.

Almeida, Heitor, Igor Cunha, Miguel A Ferreira and Felipe Restrepo. 2017. "The real effects of credit ratings: The sovereign ceiling channel." *The Journal of Finance* 72(1):249–290.

Archer, Candace C, Glen Biglaiser and Karl DeRouen Jr. 2007. "Sovereign bonds and the" democratic advantage": Does regime type affect credit rating agency ratings in the developing world?" *International organization* pp. 341–365.

Babbel, David F and Stefano Bertozzi. 1996. *Insuring sovereign debt against default*. World Bank.

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

Chari, VV, Alessandro Dovis and Patrick J Kehoe. 2020. "On the optimality of financial repression." *Journal of Political Economy* 128(2):710–739.

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

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

Financial Stability Board. 2010*b*. "Principles for Reducing Reliance on CRA Ratings.".

Fitch. 2020*a*. "Fitch Revises India's Outlook to Negative, Affirms IDR at 'BBB-'.".

Fitch. 2020*b*. "Sovereign Ratings Criteria: Master Criteria.".

Fuchs, Andreas and Kai Gehring. 2017. "The home bias in sovereign ratings." *Journal of the European Economic Association* 15(6):1386–1423.

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* 17(3):288.

Government of India. 2017. "Economic Survey 2016–17.".

Gulhati, Ravi I. 1972. "India's External Debt." *India Quarterly* 28(1):3–11.

Hill, Claire A. 2004. "Regulating the Rating Agencies." *Washington University Law Quarterly* 82:43.

Kraay, Aart, Daniel Kaufmann and Massimo Mastruzzi. 2010. *The worldwide governance indicators: methodology and analytical issues*. The World Bank.

Lindert, Peter H and Peter J Morton. 1989. How sovereign debt has worked. In *Developing Country Debt and Economic Performance, Volume 1: The International Financial System*. University of Chicago Press pp. 39–106.

Livemint. 2017. "Modi, Putin agree to develop 'independent' credit rating industry.".

Moody's. 2019. "Sovereign Ratings Methodology.".

Moody's. 2020. "Moody's downgrades India's ratings to Baa3, maintains negative outlook.".

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

Outlook. 2017. "India deserves a better rating: OECD.".

Perez, Diego et al. 2015. "Sovereign debt, domestic banks and the provision of public liquidity." *Manuscript, New York University* 3:14.

Primo Braga, Carlos A and Galli A Vincelette. 2010. *Sovereign debt and the financial crisis: will this time be different?* The World Bank.

Reinhart, Carmen M and Kenneth S Rogoff. 2009. *This time is different: Eight centuries of financial folly*. Princeton university press.

Reuters. 2011. "EU attacks credit rating agencies, suggests bias.".

S&P. 2017. "Sovereign Ratings Methodology.".

The World Economic Forum. 2020. "Countries are piling up record amounts of debt amid Covid-19.".

Sovereign Credit Ratings – An Assessment of Methodological Flaws and Rating Biases

Internet Appendix

A Model specification, Variables Used, and Data Sources

Below we describe the variables used in our analysis. Figure A1 summarizes the variables used for each model in Section 4.1 The following sub-section lists the sources for each data variable.

A.1 Rating Methodology: Variables Used

Structural Factors: Structural features evaluate governance quality, wealth, flexibility of the economy, Political stability and financial sector risks. The sovereign ratings model takes into account five variables that represent the institutional and structural features of the sovereign being rated. They variables and how they enter the sovereign ratings model are as follows:

1. Composite Governance Indicators: simple average percentile rank of world bank governance indicators: rule of law; government effectiveness; control of corruption, and voice & accountability; regulatory quality; political stability & absence of violence

2. GDP per capita: percentile rank of GDP per capita in US dollars at market exchange rates

3. Share in world GDP: Natural logarithm of percentage share in world GDP in US dollars at market exchange rates

4. Years since default or restructuring: non-linear function of the time since the last event; 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.

5. Money supply: natural logarithm of broad money expressed as a percentage of GDP

Fitch (2020*b*) states that the overall post-estimation weight of this pillar in the model is 53.7%. Barring year since default, all variables should have a positive impact on the model output i.e. an increment in them results in a linear increment in the ratings score. In its qualitative overlay, Fitch ratings assesses metrics of political stability and capacity, financial sector risks and other structural factors not captured in the SRM.

Macroeconomic Performance and Prospects: Fitch's Macroeconomic pillar evaluates the macroeconomic stability, policy credibility, GDP growth outlook and inflation of the rated sovereigns. The variables included in the sovereign rating model are:

1. Real GDP growth volatility: natural logarithm of an exponentially weighted standard deviation of historical percent changes in real GDP
2. Consumer price inflation: three-year centered average of annual percent change in consumer price index, truncated between 2% and 50%
3. Real GDP growth: three-year centered average of annual percent change in real GDP

This pillar contributes to 10% weight in the sovereign rating model. In its qualitative overlay, Fitch looks at the five-year GDP growth outlook and the sovereign's relative performance, both across time and its peers in the rating group, to make qualitative adjustments to the baseline rating score.

Public Finances and Government: In its analysis of public finances, Fitch studies government debt, fiscal balance, public debt dynamics and fiscal policy. The variables used under this pillar are:

1. Gross general government debt: three-year centered average of gross general government debt as a percentage of GDP
2. Interest Payments:three-year centered average of gross government interest payments expressed as the percentage of general government revenues
3. General Government Fiscal Balance: three-year centered average of gross general government budget balance expressed as a percentage of GDP
4. Public Foreign Currency Debt: Three-year centered average of public foreign currency denominated and indexed debt expressed as a percentage of gross general government debt.

Public finances' weight in the sovereign ratings model is 18%. In its qualitative rationale, the agency looks at the sovereign's fiscal financing flexibility - its record of market access, depth of capital markets, potential sources of financing, public debt sustainability and fiscal structure.

External Finances: Fitch analyzes the sovereign's balance of payments, external balance sheet and external liquidity under the fourth and final pillar: external finances. The variables that go into the sovereign ratings model from this pillar are:

1. Reserve Currency Flexibility: natural logarithm of the share of that country's currency in global foreign-exchange reserve portfolios (plus a technical constant), as reported by the IMF in its COFER database

2. Commodity Dependence: Non-manufactured merchandise exports as a share of current account receipts

3. Official international reserves (for non-reserve currency sovereigns): Year-end stock of international reserves (including gold) expressed as months' cover of current external payments

4. Sovereign Net Foreign Assets: three-year centered average of sovereign net foreign assets as percent of GDP

5. Current Account Balance plus net FDI inflows: three-year centered average of current account balance and net FDI (% of GDP)

6. External Interest Service: three-year centered average of external interest service expressed as percentage of current account receipts

External finances makes up for the remainder 17.4% weight in the sovereign ratings model. The qualitative overlay under this pillar looks at metrics of external financing flexibility, external debt sustainability and the sovereign's vulnerability to external shocks.

Figure A1: Model Specification

This table summarizes the variables, number of observations, and time-period for the analysis for the individual models described in Section 4.1.

Factors determining CRA's Rating Score			
Pillars	Fitch	S&P	Moody's
Institutional	Institutional score; Years since default	Transparency of Institutions	Regulatory Quality; Rule of Law
Economic	GDP Per capita; GDP Growth; Growth volatility; Share in Global GDP; CPI Inflation; Broad money/GDP ratio;	Growth in GDP Per capita; GDP Growth	GDP Per capita; GDP Growth; Unemployment rate
Fiscal	Debt to GDP ratio; Fiscal balance/GDP ratio	Debt to GDP ratio	Debt to GDP ratio
External	Import cover; (CAB+FDI)/GDP ratio; Commodity dependence;	Actively traded currency; Gross external financing need	Current Account Deficit/GDP ratio
Monetary		Central bank independence	
No. of countries	41	87	83
Sample period	1994-2020	1998-2020	2000-2020
Total Obs.	609	960	1251

A.2 Data Sources

Below we list the data sources for the variables used for each CRA.

Moody's

- **GDP per Capita**: IMF World Economic Outlook (WEO) April 2020 & June 2020 update, World Bank World Development Indicators (WDI)
- **Current Account/GDP**: World Bank World Development Indicators (WDI)
- **Growth of Real GDP**: IMF World Economic Outlook (WEO) April 2020, World Bank World Development Indicators(WDI)
- **Unemployment Rate**: IMF World Economic Outlook (WEO), April 2020
- **General Government Debt to GDP ratio**: IMF World Economic Outlook (WEO) October 2019
- **Regulatory Qualtity** World Bank Governance Indicators
- **Rule of Law**: World Bank Governance Indicators

Fitch

- **Institutional Score**: World Bank Governance Indicators
- **GDP per capita**: IMF World Economic Outlook (WEO) April 2020 & June 2020 update, World Bank World Development Indicators (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 (2015) and Bank of Canada's Credit Rating Assessment Group Database of Sovereign Defaults 2019
- **Real GDP growth**: IMF WEO April & June 2020 update

- **Commodity Dependence**: UN Conference of Trade and Development Statistics, World Bank WDI
- **GDP growth volatility**: IMF WEO April & 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

S&P

- **Transparency of Institutions**: World Bank Governance Indicators; Global Competitiveness Index (World Economic Forum)
- **GDP per capita**: World Bank World Development Indicators (WDI), IMF World Economic Outlook (October 2019)
- **Actively Traded Currency**: Bank for International Settlement (BIS) report "Triennial Central Bank Survey"
- **Gross Financing Needs**: Consists of Current Account Payments from World Development Indicators, 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 World Development Indicators(WDI)
- **Total Reserves**: World Bank World Development Indicators (WDI)
- **Government Gross Debt (%GDP)**: IMF World Economic Outlook (October 2019)
- **Central Bank Independence**: Institutional Profiles Database