On sovereign ratings: observations and implications

Andrew Powell

Introduction

Rating agencies have come in for much criticism over the years. Most recently, they have been criticized for downgrading high rated sovereigns. It has been claimed that these actions are counter-productive and might even destabilize crucial government bond markets. While such criticisms have surfaced again in the wake of the Lehman crisis, the Great Recession and the European crisis, they have been widely known and discussed within emerging economies at least since the emerging economy crises of the 1990’s.

There is something of a love-hate relationship between the official sector and the rating agencies: on the one hand the agencies are held to blame when things go awry, perhaps they may even serve as useful scapegoats to deflect attention from more fundamental concerns, but on the other hand the official sector, especially in emerging economies, courts the agencies when a new bond issue is planned. Moreover, international regulatory authorities continue to entertain the notion that ratings are useful for regulatory purposes.

Given such controversial issues it seems useful to take various steps back to try to gain perspective. The idea in this brief note is to make some, mostly empirical, observations about sovereign ratings. In doing this I draw on a set of recent papers that mostly focus on emerging economies. As there are more emerging than advanced economy sovereigns and their ratings have varied more over time, focusing on emerging country sovereign ratings may be more revealing. Furthermore, as discussed below, the recent crises in advanced economies have had a distinct emerging country flavor. In the final section, I attempt to draw some conclusions to shed light on these recent debates and I even dare to offer some recommendations for policy makers.

The first three observations are descriptive in nature, simply noting certain properties of ratings from the different agencies. The fourth draws on a set of papers that attempt to model ratings as a function of explanatory variables. These econometric models try to mimic how agencies actually model ratings and it is found that agencies’ behavior is in general not very difficult to disentangle. Having

---

1 Principal Advisor, Research Department, Inter-American Development Bank.

Acknowledgements and disclaimer: I would like to thank my co-authors of various papers that relate to sovereign ratings, namely Eduardo Cavallo, Giovanni Majnoni, Juan Francisco Martinez and Roberto Rigobon, as much of this note draws on that body of work. I would also like to thank John Chambers (S&P) and Mauro Leos (Moody’s) and my IDB colleagues Ed Bartholemew, Eduardo Borensztein and my ex colleague Ugo Panizza for useful discussions on ratings, as well as the participants in the Sovereign Risk Seminar, held at the BIS January 8/9 2013 and particularly Bob McCauley at the BIS for his helpful and constructive suggestions. Remaining mistakes however are my own. All the opinions expressed are solely my own and do not necessarily reflect those of the Inter-American Development Bank, those of its Board of Directors, or those of the countries they represent.
said that, there are some tricky methodological issues with this work that merit discussion. Observation five highlights one such problem. Observations six through eight relate to the debate as to whether ratings add anything over and above the information that is already to be found in market variables. Observation nine focuses on the issue of multiple downgrades and observation ten relates to the cost structure of the information and ratings business. With this background, the final section then provides a discussion and recommendations.

Observation 1: rating agencies disagree about as much as they agree

Considering a standard mapping between Moody’s and Standard and Poor’s ratings, Moody’s and S&P ratings are in agreement and in disagreement about 50% of the time each over a period of about 10 years – Powell and Martinez (2008). Considering the distribution at the time of writing there are about 54% of disagreements within emerging economies. At times there are also disagreements of two notches and more although these are rarer (disagreements of at least two notches were about 10% of the sample). In theory there may be differences regarding which variables are important to establish ratings, the definitions of those variables and their relative weightings. Higher fiscal surplus, higher tax revenue and higher reserves and lower inflation were all associated with higher Standard and Poor’s ratings relative to Moody’s – Powell and Martinez (2008). In general, similar variables are significant in standard regressions of ratings but some weights differ.

Observation 2: S&P ratings relate to default probability, Moody’s to total return

A further difference is that S&P ratings relate to default probability (PD) while Moody’s prefer a focus on total loss – PD*LGD where LGD is loss given default. I have often thought that this difference could or should be exploited more in research on ratings.

For example, one potential hypothesis stemming from the 2002 Argentine default and eventual restructuring was that sovereigns could drive home deeper haircuts in than previously thought – see Sturzenegger and Zettelmeyer (2007) and Cruces and Trebesch (2011), but in some ongoing (preliminary) work we do not find strong evidence for lower expected recovery values considering the changes in Moody’s ratings versus those of S&P. Having said that, the recent legal actions in New York may be changing recovery expectations from sovereign defaults once again – see Allen & Overy LLP (2012) for a discussion of the recent NY legal actions and potential implications.
Observation 3: ratings vary over economic cycles, are pro-cyclical and exhibit serial correlation

Ratings change over economic cycles in a pro-cyclical fashion. The conditional correlation between rating changes and growth is statistically significant (value = 0.25) and there is pro-cyclicality with real exchange rate cycles. Nickell et al (2000) show ratings are not stable over time and vary with economic variables. Munford and Mulder (2000) argue that ratings vary pro-cyclically with emerging economy real exchange rates. Standard and Poor's (2012) documents that an upgrade (downgrade) is followed by a further upgrade (downgrade) within two years, in 35% (54%) of cases in 37 years of S&P sovereign ratings history, whereas it is followed by a downgrade (upgrade) for only 6% (10%) of the sample.

Observation 4: sovereign ratings are relatively easy (too easy?) to model

Panel regressions of ratings against a relatively small set of economic and institutional variables deliver significant coefficients and high R-squared statistics – see Cantor and Packer (1996), Afonso et al (2007) and Powell and Martinez (2008). This is actually not that surprising, especially given that a Principal Component analysis of all emerging country ratings over 15 years shows that two Principal Components explain over 80% of the variation in ratings and three factors can explain about 90%. Moreover, just three factors explain about 70% of the variation in seven leading economic indicators often used to explain ratings – Powell and Martinez (2008).

Considering standard regressions of ratings however, some methodological questions emerge. Perhaps the most serious ones relate to endogeneity and the fact that history clearly matters. One particular aspect of endogeneity can be seen considering the results regarding the inclusion of the current account as an explanatory variable. If it is included it is normally significant but with a negative sign. This suggests reverse causality. Countries with higher ratings tend to be able to finance higher current account deficits. Also a measure of income (such as GDP per capita) is often included but again perhaps a higher rating, by giving greater access to cheaper finance, allows a country to become richer? There are further problems of endogeneity; I come back to this in the next observation.

If history is important, this suggests a dynamic panel approach might be more appropriate but that then raises a set of other issues, particularly as the rating is not a continuous, cardinal variable but rather ordinal in nature. Cavallo et al (2012) discuss these issues and opt for a very different approach when trying to assess whether ratings add value to other market variables.

There are also, perhaps more surprisingly, some issues with the definition of critical variables. For example, debt is multi-faceted and the homogeneity and quality of debt statistics across countries has been poor to say the least – see IDB (2006) for a discussion. It’s not clear what debt definition should be employed nor what denominator to use; GDP, tax revenues, exports etc. Debt structure may be as important as debt quantity. It is likely rating agencies take several definitions into
account and also consider the combination of different variables, suggesting interaction effects or some type of cluster analysis may also be relevant.

**Observation 5: overrated?**

Government Effectiveness or other institutional indicators are often included in rating regressions. Indeed ratings are highly correlated with many other country indicators. Table 5 of Powell and Martinez (2008) includes 12 country indicators ranging from the ranking of the Global Competitiveness Survey of the World Economic Forum to the World Bank's “Doing Business” and Voice and Accountability indices as well as Standard and Poor's and Moody's ratings – see the appendix. Correlations between the ratings and the other indicators range from a minimum of 0.75 to a maximum of 0.93, although these are largely derived from cross-section, not time series variation. A further problem is whether these types of variables are truly exogenous. Many of these variables rely on surveys, indeed some are surveys of surveys. Perhaps respondents have been influenced by the rating and statements of the rating agencies? In the end the rating also appears to reflect something deeper about the country, its institutions or its level of development, but again it's possible that the rating history itself may play a role in affecting a country's development trajectory.

**Observation 6: rating agencies were not (wholly) convinced by the Great Moderation**

While modeling ratings may be problematic methodologically, the fact that they may be modeled relatively easily by a small set of economic variables suggests that they may be a useful summary measure. What happens if then ratings are used as an indicator of country fundamentals to, say, analyze how spreads have moved over time?

The Great Moderation period, from, say, January 2000 to July 2007, resulted in a considerable improvement in ratings for emerging economies and a significant reduction in spreads. However standard regressions cannot explain the spread reduction given the improvement in ratings. Powell and Martinez (2008) conclude that given the estimated relation between spreads and ratings, out of sample predictions of spreads result in emerging country spreads some 150–170 basis points higher than the actuals. The reduction in spreads can be explained by including indicators of global liquidity or risk aversion in the regressions (e.g.: the VIX, US High Yield, US interest rate). Interestingly, if these variables are included in regressions of ratings on economic variables, they are not significant. This suggests that the rating agencies may have allowed higher global liquidity and lower global risk aversion to affect ratings in so far that these developments helped to improve “fundamentals” but not more than that.

Martinez and Powell (forthcoming) suggests that much of the decompression of spreads as a result of the Lehman crisis can also be explained largely by liquidity factors and not by the movement in country fundamentals, while the compression of emerging economy spreads thereafter has been due to a mixture of liquidity and fundamentals.
Observation 7: ratings add value (not all information is in market prices)

A non-trivial question is whether ratings matter. A series of papers using a variety of techniques suggest that they do – see Cantor and Packer (1996), Eichengreen and Mody (1998) and Powell and Martinez (2008). But there are strong methodological assumptions made in these analyses in order to obtain identification. The essential problem is how to control for other variables also publicly known, such as fundamentals and market prices that reflect those fundamentals.

Cavallo et al (2012) provide new evidence on this old question by treating both ratings and available market variables such as spreads, stock market prices and exchange rates as noisy signals of economic fundamentals (i.e.: all subject to measurement error), and, given that framework, by then developing a methodology to test whether there is information in ratings that affect future market prices that is not already contained within existing market prices. The conclusion is that sovereign ratings do add value in the sense that they add information to the noisy signals contained in popular market signals.

Observation 8: but rating changes cannot be considered as “events”, anticipated changes add no value

It is tempting to consider rating changes as “events” and then employ an event study type methodology. But rating agencies announce outlooks and credit watches and frequently advertise their future actions, in that they outline why a rating move may be forthcoming and even suggest what a country should do to obtain an upgrade or avoid a downgrade. This hardly conforms to the assumptions of the classic event study frequently employed in the corporate finance literature – see Campbell et al (1996). Steiner and Heinke (2001) argue rating changes may have effects due to regulation rather than information per se. Cavallo et al (2012) divide rating changes into ones that are more anticipated versus ones which are not (using outlook changes and rating watches as indicators of anticipation), and find anticipated changes in sovereign ratings contain no new extra information over and above market prices.

Observation 9: multiple downgrades are not so uncommon, liquidity and sudden stops imply significant uncertainties, multiple equilibria are possible

Available S&P sovereign ratings data indicate that of 285 total downgrades there were 64 downgrades of at least 2 notches – 43 double downgrades and 21 of more than 2 notches. There is a clustering of multiple downgrades around crises such as the Asian crisis and subsequent Russian default and the more recent European crisis. A multiple downgrade might come about for various reasons including some large exogenous shock, some big policy change (or no policy change when a large
positive one was expected), as well as a mistake in terms of underestimating a 
potential risk factor. Such events are then not necessarily a signal of incompetence.

Interestingly, there is also a clustering of Sudden Stop events around the 
aforementioned crises – Calvo et al (2004, 2008), Cavallo and Frankel (2007) and 
Forbes and Warnock (2011). These crises have been labeled crises of the capital 
rather than the current account, related more to shifts in financial stocks than flows 
for payments of goods and services. One possibility is that the risks and potential 
effects of Sudden Stops have been underestimated and when they do occur, 
multiple downgrades then occur ex post rather than ex ante.

Consider the case of Spain that has suffered a dramatic Sudden Stop of private 
capital flows. Plugging Spain’s end of 2011 parameters into a Sudden Stop debt 
sustainability model with assumptions on certain elasticities reveals a significant 
required real devaluation but debt is denominated in nominal euros – see Powell 
and Ruiz (2012) and Borensztein et al (2010). The required real devaluation is 
engineered through recession and the real devaluation implies a shift in relative 
prices, which both interact to worsen Spain’s debt sustainability indicators, and the 
required fiscal retrenchment hurts growth yet further. The relationship among 
required current account adjustment, the relative price changes, growth and debt 
sustainability is not obvious and fraught with problems to estimate precisely. It 
seems quite likely that rating actions in such cases may arrive later once these 
processes are well-advanced rather than anticipating them precisely.

Moreover in the case of Spain, due to official Eurosystem financing, the private 
sector Sudden Stop has not been converted to a fully fledged Calvo et al (2004) 
type Sudden Stop, requiring full current account adjustment as suffered by many 
emerging economies. This financing has also maintained interest rates at a lower 
level. Simulations reveal that if this financing were to be interrupted, and interest 
rates rise, then the combination of higher rates and the Sudden Stop would be even 
more devastating for the Spanish economy. Rating Spain then involves not only an 
analysis of the implications of the Sudden Stop on debt sustainability but also a 
forecast of how Eurosystem financing will evolve.

Indeed it remains uncertain exactly how the Spanish adjustment will play out 
and what that will imply for debt sustainability. As this does not depend solely on 
actions by Spain but also on those of the European authorities, it depends also on 
the intricate nature of the politics of the European Union. There may be significant 
good news or other news over the course of the next few years that may have 
consequences for ratings.

Given the importance of liquidity there is also the possibility of multiple 
equilibria – see for example the discussion in Blanchard (2011). Put in the simple 
language of debt sustainability if interest rates remain low, the debt of a nation 
saddled with a relatively high debt/GDP ratio may be sustainable, but if interest 
rates were to rise then the risk of that debt burden becoming unsustainable may be 
high, justifying those higher interest rates. The Sudden Stop literature has amply 
illustrated the link between shifts in the stock of financial assets, the required 
adjustment in the current account and the potentially sharp effects on growth. 
Assuming there are multiple equilibria and not just significant uncertainty invokes a 
set of deeper issues, and I come back to them in the final section.
Observation 10: there are many banks, not that many sovereigns and even fewer rating agencies

Basel II’s Standardized Approach introduced the possibility of using external ratings for bank regulatory capital calculations. An average sized bank may have thousands of small corporate loans to medium sized companies. Moreover, several banks may be lending to the same medium sized company. There are then large cost advantages of using such external opinions versus generating a system of internal ratings for every average sized bank.

However, there are relatively few sovereigns. And while many banks may lend to the same sovereign, the cost advantage of using an external versus an internal rating for a bank lending to a sovereign is surely smaller. Moreover, it is generally agreed that rating corporates is more about competent analysis and projections of balance sheets and income statements while rating a sovereign includes many more subjective elements such as assessing political as well as economic risk, and the relation between the two. Rating corporates might be then considered something more of a science, rating sovereigns more of an art. If this is correct, rating corporates may then be more about obtaining the right information and pursuing the right analysis while in rating sovereigns, subjective opinions may differ.

Rating agencies and regulation: a discussion

At least since Akerlof (1970) and Mirlees (1975), it has been known that information structure is a critical determinant for how markets function and may strongly affect economic outcomes. More recently, Morris and Shin (2003, 2004) advanced a “global game” framework to analyze the impact of the quality of public versus private information signals. Consider an application to bond markets. Following these latter authors, one possibility is that there are three regions: one where the fundamentals are so strong (e.g.: debt low, fiscal situation sound) that the debt of a country would be sustainable whatever the interest rate, a second region where the debt would be unsustainable whatever the interest rate (high debt, large fiscal deficit) and an intermediate region where there might be a multiple equilibrium as discussed above. A unique equilibrium however emerges in a global game framework if the private signal that agents have is of sufficiently high quality given the quality of a public signal. In an interesting twist to the global games story, Angeletos and Werning (2004, 2006) argue that in a context where the market price aggregates private information the multiple equilibrium regime may be expected to dominate.

While these papers do not model the information market and the existence of credit rating agencies explicitly, and more research seems warranted, it seems appropriate to consider credit ratings as a further public signal in addition to the price. In the spirit of Cavallo et al (2012), prices and ratings are both noisy signals of underlying fundamentals.

It might then be argued, extrapolating from these recent theoretical contributions, that if ratings are of poor quality and essentially add noise, then the region in which a unique equilibrium might emerge may be larger in the parameter space, whereas if credit ratings are of high quality then multiple equilibria would tend to prevail. But if they were of poor quality (and so there was a unique
equilibrium) then a downgrade might be expected to have little impact. On the other hand if they were of high quality (and so there were multiple equilibria) then a downgrade might be expected to have a significant effect, particularly if they acted to coordinate the private sector to a good or to a bad equilibrium. A change in ratings could even conceivably prompt a Sudden Stop. Given potential multiple equilibria, it is very hard to test causality empirically but either way it is easy to see why Sudden Stops might be accompanied by multiple downgrades.

Ratings may then play a quite subtle role in the game of information aggregation. It’s interesting to tease out the implications of the above arguments. If ratings are of low quality relative to private signals, there may be a unique equilibrium, but, since they are of low quality, we would not expect downgrades (or upgrades) to have too much impact. If they are of high quality there might be multiple equilibria and downgrades might be significant for other agents to update their views. Of course it may not be just one or the other; the quality of the signal may vary over country or over time. These considerations may lie behind some of the results obtained by Cavallo et al (2012) for different cases.

Should regulators then regulate rating agencies? This question has many aspects and I will only give a partial view here. If regulation could improve the quality of ratings (let’s suppose the market may have some race to the bottom characteristic and regulation might force the agencies to use more of their rents to invest in better research for example), then this might push us to the multiple equilibria story, and ratings might then become more influential, which is not perhaps what some regulators have in mind. On the other hand, some regulators may wish to constrain rating agencies’ actions, which would presumably lower the quality of their ratings relative to an unconstrained case, and hence render them less important.

Rather than regulating existing agencies directly, perhaps policy makers should try to work to make the information market function more efficiently, which may also have the effect of reducing the influence of the small number of existing agencies. One idea would be to minimize any pure information rents that agencies may have. To some extent improvements in information and transparency go in this direction but there remains some way to go, for example to improve the quality and timeliness of debt and other statistics.

A further step would be to harness the internal ratings that banks already make regarding sovereigns. Bank regulators could collect such internal ratings and publish the mean, median, and percentiles of the relevant distributions. In fact this would yield new information to the market, while maintaining confidentiality. As there are only three main rating agencies it is hard to gauge the precision regarding the information provided. An agency may state a country is an A, but how sure is it that it is an A? While rating watches and outlooks give some qualitative assessment that a rating may change, a statistical analysis of banks’ internal ratings published on a regular basis would be much more adequate to the task. The analogy would be a Central Bank’s survey of inflation; many Central Banks publish the statistics of the distribution of such surveys that then yields information on the uncertainty of any inflation forecast.

Indeed, Majnoni and Powell (2005) suggested that banks should not be allowed to use external ratings for sovereigns. We suggested that if a bank wished to lend beyond its border to a sovereign it should have the capacity to analyze the risks involved. As argued above the extra costs of doing this do not appear unwarranted. Of course the ratings of the leading agencies may be used as an input to that
decision but the bank should have its own internal rating system. This should harness the opinions of the relevant senior managers of that institution and be reviewed by the bank’s regulator. Moreover, we suggested that banks could report those ratings to the BIS, or another central body that could convert them to a standardized scale and publish aggregate statistics regarding the distributions of banks’ opinions. It is unlikely that this would remove the problem of multiple equilibria but it may reduce the dependence on the opinions of a very few agencies.

A second recommendation is to sponsor research specifically on the information structure of these markets. There is fascinating recent research on the role of market players and prices and how information is transmitted but the research to date does not include rating agencies explicitly. Rating agencies may be considered as a combination of information aggregators and opinion formers and the ratings are perhaps best considered as a further noisy signal of the actual “fundamentals” regarding probability of default or total expected loss. I developed some arguments in the above extrapolating from existing theoretical models, and combined them with arguments regarding the cost structure of the information business but they should definitely be read with health warnings attached. Further research would no doubt contribute to our understanding regarding the benefits and dangers of the sovereign rating business.

References


Panizza, U., A. Powell and P. Tavella (ongoing), pls. contact Andrew Powell at andrewp@iadb.org for details.


