

IMF Working Paper

Are Rating Agencies Powerful?
An Investigation into the Impact
and Accuracy of Sovereign Ratings

*John Kiff, Sylwia Nowak, and Liliana
Schumacher*

IMF Working Paper

Monetary and Capital Markets Department

Are Rating Agencies Powerful? An Investigation into the Impact and Accuracy of Sovereign Ratings

Prepared by John Kiff, Sylwia Nowak, and Liliana Schumacher¹

Authorized for distribution by Laura Kodres

January 2012

This Working Paper should not be reported as representing the views of the IMF.

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Abstract

We find that Credit Rating Agencies (CRA)'s opinions have an impact in the cost of funding of sovereign issuers and consequently ratings are a concern for financial stability. While ratings produced by the major CRAs perform reasonably well when it comes to rank ordering default risk among sovereigns, there is evidence of rating stability failure during the recent global financial crisis. These failures suggest that ratings should incorporate the obligor's resilience to stress scenarios. The empirical evidence also supports: (i) reform initiatives to reduce the impact of CRAs' certification services; (ii) more stringent validation requirements for ratings if they are to be used in capital regulations; and (iii) more transparency with regard to the quantitative parameters used in the rating process.

JEL Classification Numbers: G20, G24, G38

Keywords: Sovereign ratings, credit rating agencies, credit default swap

Author's E-Mail Address: jkiff@imf.org; lschumacher@imf.org; snowak@imf.org

¹ Yoon Sook Kim and Ryan Scuzzarella provided technical assistance. The purpose of this paper is to expand on some of the empirical analysis in Chapter 3 of the October 2010 Global Financial Stability Report (IMF, 2010).

Contents	Page
I. Introduction	3
II. Do Ratings Matter?	5
A. The issues	5
B. Literature Review	6
C. Data Description and Methodology	8
D. Results	11
III. Accuracy and Stability of Ratings	15
A. What do Ratings Measure	15
B. An Examination of Rating Accuracy	17
C. An Examination of Rating Stability	21
IV. Some Policy Recommendations	25
References	27
Tables	
1. Univariate Regressions of Log Spreads on Rating Categories	10
2. The Empirical Tests of Rating Information Value	12
3. Multi-Agency Rating Dynamics	15
4. Sovereign Rating “Failures” during the 1997–1998 Asian Crisis	24
5. Sovereign Rating “Failures” during the 2008–2010 Crisis	25
Figures	
1. Average CDS Spread and Ratings for Countries Rated by Moody’s, 2005–2010	10
2. Impact of Change in Sovereign Ratings and Outlooks on CDS Spread	13
3. Ratings One Year Prior to Sovereign Default, 1975–2009	18
4. Sovereign Foreign-Currency Rating Performance by Standard & Poor’s	20
5. Average Proportion of S&P Sovereign Ratings Unchanged Over One Year	22
6. Average Proportion of S&P Sovereign Ratings Downgraded More Than Two Notches Over One Year	23
7. Asian Crisis: Sovereigns Rated by Moody’s between July 31, 1997 and December 31, 1998	23
8. Current Crisis: Sovereigns Rated by Moody’s between July 31, 2007 and June 30, 2010	24
Appendixes	
1. Analysis of One-year Fitch and S&P Sovereign Rating Changes	31
2. Sovereign Downgrades and Upgrades During the Asian and Global Crisis	33

I. INTRODUCTION

1. **Recent rating activities by the Credit Rating Agencies (CRAs) have induced some to ask whether ratings represent accurate risk assessments and to question how influential they are.** The contention that ratings represent accurate default risk metrics was brought into question by the sheer volume and intensity of the multiple downgrades to U.S. mortgage-related structured finance securities in the wake of the crisis. Voices have also been raised against the timing of recent downgrades of European sovereigns amidst criticism that these downgrades promoted uncertainty in financial markets, leading to “cliff effects” and as a consequence affect their ability to fund themselves. Rating agencies have also been accused of behaving oligopolistically.

2. **These criticisms are not new.** CRAs’ downgrading actions have been accused before of not being timely but instead procyclical. It was argued that “the Mexican crisis of 1994-95 brought out that credit rating agencies, like almost anybody else, were reacting to events rather than anticipating them” (Reisen, 2003). During the late 1990s Asian crisis, CRAs were also blamed of downgrading East Asian countries too late and more than the worsening in these countries’ economic fundamentals justified, exacerbating the cost of borrowing.²

3. **The goal of this paper is to assess these concerns.** We first examine CRAs’ role and whether CRAs are influential or just lag the market once new information is available and priced into fixed income securities. This is an important point. If CRAs influence the market, their opinions are important from a financial stability perspective. If they do not and just reflect information available to the market, their actions are not relevant and there is no policy concern. In this regard, we test three hypotheses regarding the services that CRAs provide to the market: information, certification and monitoring. We find evidence that CRAs’ opinions are influential and favor the information and certification role. We then attempt to determine what ratings actually measure and how accurate they are. We conclude with some policy recommendations based on these findings. This study is limited to sovereign ratings (of emerging markets and advanced economies) and covers the period January 2005-June 2010.

4. **Our analysis of the interaction between the market and CRAs indicates that ratings have information value beyond the information already publicly available to the market.** Specifically, the following results were evident:

- **An event study shows that negative credit warnings (i.e., “reviews,” “watches,” and “outlooks”) have a significant impact on CDS spreads.** This evidence is also supported by a Granger-causality test that finds that negative credit warnings Granger-cause changes in CDS spreads. These findings are consistent with the view that rating agencies do provide additional information to the markets, in addition to what is publicly available and used by markets to price fixed income securities.

²Ferri, Liu, and Stiglitz (1999) compared actual ratings with a rating model based on economic fundamentals. In their view, the difference represents the agency’s judgments. They conclude that rating agencies assigned higher weights to their judgment than to fundamentals. See also IMF (1999).

- **Although upgrades and downgrades in general do not have a significant impact on CDS spreads, upgrades and downgrades in and out of investment grade categories are statistically significant.** This supports the view that the certification services provided by rating agencies do matter and likely create a purely liquidity effect (e.g. purchases and sales of assets forced by regulations or other formal mandates and not based on the additional information already in the market).
- **We do not find evidence in favor of the most important testable implication of the monitoring services theory.** The impact of downgrades preceded by an outlook review in the same direction is not statistically significant.
- **From an informational point of view, the market appears to discriminate more than rating agencies among different kinds of issuers—in particular at lower rating grades and during crisis periods.** This finding may indicate that ratings need to incorporate more granularity and leads to the second question of what ratings measure and how accurate they are.³
- **A common element among ratings by the major CRAs is that they represent a rank ordering of credit risk.** This ordering is based on qualitative and quantitative inputs such as default probabilities, expected losses, and downgrade risk. However, there is no one-to-one mapping between any of these quantitative measures of credit risk and credit ratings. There is also no disclosure of the quantitative parameters that characterize each rating grade. For this reason, validation tests undertaken by outsiders can only apply to the ability of ratings to differentiate potential defaulters and non defaulters, but not to estimating cardinal measures such as default probabilities.
- **The point highlighted above implies that—in spite of playing a similar role to internal ratings in the Basel II internal ratings-based (IRB) approach—ratings produced by the CRAs are subject to lower validation standards than are the banks using the IRB approach.** In the Basel II IRB approach, financial institutions use measures of default probabilities (PD), losses given default (LGD) and exposure at default (EAD) to produce their internal ratings and are subject to calibration tests. Although validation is foremost the responsibility of banks, both bank risk managers and bank supervisors need to develop a thorough understanding of validation methods in evaluating whether banks' rating systems comply with the operating standards set forth by Basel II.
- **Ratings produced by the major CRAs perform reasonably well when it comes to rank ordering default risk among sovereigns, i.e. defaults tend to take place among the lowest rated issuers.** Accuracy ratios (AR) indicate that agencies are more successful at sorting out potential defaulters among sovereign issuers (average ARs in the 80 to 90 percent range) than among corporate and structured finance issuers (average ARs in the 63 to 87 percent range), the latter ones having suffered a strong deterioration

³ For example, one might put another rating notch between AA+ and AA, and so on.

over the global financial crisis. For all classes of products though, the ARs indicate that sovereign rating accuracy deteriorates as the evaluation horizon increases.

- **In general, long-term credit transition matrices show that higher ratings are more stable than lower ones, and tend to remain unchanged.** But this was not the case during the global crisis period, when there has been a tendency to see heavier downgrade activity among higher-rated sovereigns than among lower-rated ones. There has been evidence of significant rating failure (defined here as three or more rating changes in one year) during the recent global financial crisis, although less than in the Asian crisis.

II. DO RATINGS MATTER?

A. The issues

5. **This section discusses the role that CRAs play in the market and their impact.** CRAs are perceived as powerful institutions that can influence issuer survival by affecting their access to funding markets, and their funding costs. At the same time, there have been claims that they just follow the market and react to events rather than anticipate them. Whether CRAs anticipate or lag the market is important. If CRAs influence the market, their opinions are important from a financial stability perspective. If they do not and just reflect information available to the market, it means that their actions are not relevant and there is no policy concern. This study contributes to the literature on the impact of sovereign ratings on the funding costs of governments by incorporating the recent period of sovereign risk turmoil and attempting to formally test alternative theories explaining the role of rating agencies.

6. **Three theories have been proposed in the literature to explain whether CRAs add value to markets and how they exercise their power: via information provision, certification services and monitoring services.**

- The traditional approach about why CRAs add value is that in a world of asymmetric information, they produce and disseminate information on the risk of fixed income securities which is used by investors to price information-sensitive securities. By reducing information costs, they increase the pool of potential borrowers and promote liquid markets. Millon and Thakor (1985) provide a rationale for the existence of institutions that produce costly information, whose payoffs depend on noisy ex post monitors of information quality. They find that the formation of information gathering agencies enables screening agents to share information and diversify their risky payoffs.
- In addition to providing information, CRAs also provide certification services. For example ratings are used to classify securities as either investment or non-investment grade, which influences institutional demand and market liquidity, and serve as triggers in investment decisions and regulatory oversight.⁴ Another notable regulatory example is the use of ratings to calculate Basel II risk-based capital requirements under the standardized approach. Other examples are the limits set by central banks to determine

⁴ On the use of ratings in regulations see the BCBS (2009).

acceptable collateral and margin requirements, and to manage their reserves, and the use of ratings in investment mandates that dictate the behavior of various investment funds.

- In addition to the information and certification theories, Boot, Milbourn and Schmeits (2006) highlight the role of CRAs as monitors and providers of valuable coordination of beliefs in situations where multiple equilibria can obtain.⁵ In this approach, the change in the issuer's credit quality is observed by everybody—so CRAs do not provide any informational services. Instead, CRAs derive their value mainly from their monitoring role, most apparent in their credit watch procedures. Boot, Milbourn, and Schmeits (2006) suggest that the credit watch procedure is not just the initial information dissemination but an action taken by the CRAs by which a monitoring regime is put in place. The credit watch allows for an implicit contract between the issuer and the CRA where the former implicitly promises to undertake specific actions—recovery efforts—to mitigate the possible deterioration of its credit standing (and rating). The credit watch procedure also plays an economically meaningful role because it helps fix the desired equilibrium.

7. **These theories have testable implications that we exploit in this section.** The information theory would be compatible with evidence pointing to a significant market reaction to rating actions (either changes in ratings or credit warnings). The certification theory would point to a significant market reaction to upgrades and downgrades as well as a higher significant impact of rating actions involving issuers going in and out of the investment grade categories vs. other rating changes. If, instead, the monitoring services theory is a better explanation of the role played by rating agencies, we should see a significant market reaction to rating confirmation and change—conditional on a credit watch because the CRAs had access to private information during the credit watch procedure. Since credit watches are more likely to precede downgrades, the empirical evidence should be the strongest for downgrades.⁶

B. Literature Review

8. **Early empirical studies of credit ratings, such as Liu and Thakor (1984), and Ederington, Yawitz, and Roberts (1987), focused primarily on corporate ratings as an explanatory variable for cross-sectional differences in corporate yield spreads.** These studies, based on regression analysis, found that, as expected, ratings are correlated with spreads since ratings and credit risk are obviously related.

9. **More interestingly, later studies seek to discover the empirical impact of corporate rating changes and the adjustments in the prices of the firms' bonds, equities, and commercial paper.** Studies such as Hand, Holthausen, and Leftwich (1992); Dichev and Piotroski (2001); Hull, Predescu, and White (2004); Micu, Remolona, and Wooldridge (2006); and Norden (2008), have focused on the issue of whether changes in ratings convey information not already incorporated into prices from other sources. In general, they all find evidence that

⁵ The multiple equilibria arise from the fact that investors cannot see the choices made by the issuer. Depending on investors' beliefs, an issuer might be induced to choose a high-risk strategy rather than a more viable low-risk one. The market belief then can be self fulfilling.

⁶ This is the most important testable implication. The model has others that we do not discuss here.

ratings are “sticky”, i.e., rating changes tend to happen after some anticipation took place already in the market with regard to changes in the issuer’s credit quality. However, they all find that one or more rating actions (or some combination of them, such as aggregated rating announcements) do cause spreads to change in a significant manner, although there are mixed results about which actions are more influential.

10. **In their seminal paper, Cantor and Packer (1996) model the determinants of sovereign ratings and ask the question of whether ratings add to public information.** Their study, based on sovereign bond spreads for advanced and emerging economies, finds that the single rating variable explains 92 percent of the variation in spreads. While most of the correlation appears to reflect similar interpretations of publicly available information by the rating agencies and by market participants, their event study finds evidence that the rating agencies’ opinions independently affect market spreads, especially in the case of non-investment grade sovereigns. In addition, the impact of one agency’s announcement is greater if the announcement confirms the other agency’s rating or a previous rating announcement.

11. **Larrain, Reisen, and von Matzlan (1997), in the aftermath of the Mexican crisis, assess whether S&P and Moody’s ratings lead or lag market events.** Their event study of 78 events during 1987-1996 concludes that a change in the risk assessment by the two leading rating agencies is preceded by a similar change in the market’s assessment of sovereign risk, especially when countries have been put on review. This makes them caution against overestimating the independent long-run impact that sovereign credit ratings exert on the financial market assessment of sovereign risk.

12. **Reisen and von Maltzan (1999) associate downgrades with wider sovereign yield spreads, especially for emerging markets and the aggregated rating announcements of the three CRAs.** Their Granger causality test shows a two-way causality between Moody’s ratings and yield spreads but no significant link between S&P’s ratings and spreads. Correcting for joint determinants of bond yields and sovereign ratings, they show that rating decisions may be considered as an important signal of credit worthiness; and that this impact may be reinforced by prudential regulations that bar investors from holding securities below certain rating categories.

13. **Kaminsky and Schmuckler (2002) examined data from 16 emerging markets covering the period January 1990-June 2000 to assess spillover effects.** Specifically they assess whether ratings of assets from one country trigger contagious fluctuations in other countries and whether changes in ratings of one type of security affect other asset markets. They conclude that changes in sovereign debt ratings and outlooks affect financial markets in emerging economies. They affect not only the instrument being rated (bonds) but also stocks. As well, they directly affect the markets of the countries rated and generate cross-country contagion. The effects of rating and outlook changes are stronger during crises in non-transparent economies and in neighboring countries. They also conclude that upgrades tend to take place during equity market rallies, whereas downgrades occur during economic downturns, providing support to the idea that CRAs contribute to the instability in emerging financial markets.

14. **Ismailescu and Kazemi (2010) use an event study methodology to examine the effects of sovereign credit rating change announcements on the CDS spreads of the event countries and their spillover effects on other emerging economies’ CDS premiums.** They report that ratings announcements appear to reveal new information that affects CDS spreads,

with stronger reaction to positive events. Positive events display some spillover effects, but negative credit rating announcements have no impact on CDS spreads of other emerging economies. The spillover effect of positive events is however only marginally significant. The transmission channels of these spillover effects are a common lending center and competition in trade markets.

C. Data Description and Methodology

15. **We assess the impact of sovereign rating events on CDS spreads to gauge the importance of these services to market participants.** The main analysis covers 72 sovereigns for which both Moody's ratings and the relevant CDS data from Markit are available from January 2005 to July 2010.⁷ This includes 101,176 daily spread observations and 194 credit rating events: 26 downgrades, 57 upgrades, 71 positive outlooks/reviews for upgrade, and 40 negative outlooks/reviews for downgrade. In addition, we review the impact of rating actions by S&P and Fitch. Both event studies and causality tests are used.

16. **Stylized facts.** Figure 1 presents some stylized facts about the relationship between CDS spreads, which measure the market price of creditworthiness, and Moody's ratings. For this figure, for each year and each country, the average CDS spread and average credit rating is calculated (assuming a linear conversion, with 21 = AAA and 1 = D). This provides a total of 71 observations each year. An exponential trend line is then imposed. The data confirm important points raised by previous literature:

- There is a close association between spreads and ratings with lower credit categories being reflected in higher spreads. The relationship between spreads and credit ratings is clearly nonlinear. This non linearity is in line with historical default probabilities and credit rating agencies' estimates of "idealized default probabilities" by rating categories, both of which show PDs increasing more steeply at the lowest credit grades.
- The average spreads in 2006 and 2007 were low (63 and 60 basis points across all countries), even compared with 2005 (100 basis points on average). They increased to 212, 325 and 200 basis points in 2008, 2009 and H1 2010 respectively (the 2010 data does not include the downgrade of Greece).
- It is notable that not only spreads increased across the board during the recent crisis, but since 2007 the dispersion of spreads at the lowest rating grades has widened. This suggests that the market discriminates more among different risk profiles than the CRAs and that this additional discrimination takes place mainly among the worst credits. This

⁷ These countries are Argentina, Australia, Austria, Belgium, Bulgaria, Brazil, Tunisia, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Germany, Denmark, Dominican Republic, Egypt, El Salvador, Estonia, Finland, France, Greece, Guatemala, Iceland, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Korea, Latvia, Lebanon, Lithuania, Malaysia, Malta, Mexico, Morocco, Netherlands, Norway, New Zealand, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Qatar, Hungary, Romania, Russia, Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sweden, Thailand, Trinidad and Tobago, Turkey, United Arab Emirates, United Kingdom, Ukraine, Uruguay, United States, Venezuela, and Vietnam.

additional discrimination during crises has also been recognized in other studies (Sy, 2002).⁸

- Table 1 is the outcome of univariate regressions of log spreads on rating categories. Interestingly, the significance of ratings decreases over time, indicating the increasing impact of other determinants of spreads as the crisis developed.

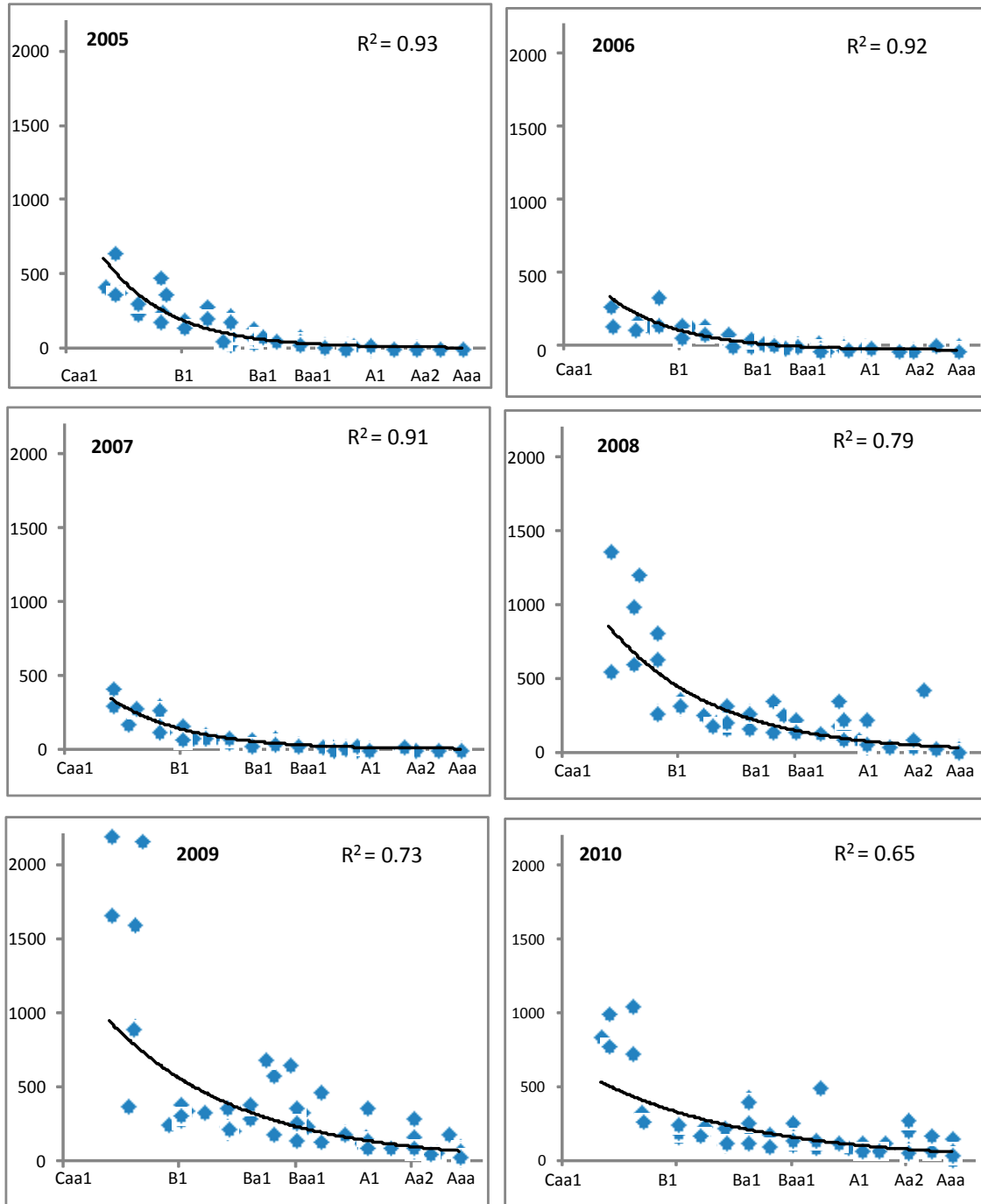
Table 1. Univariate Regressions of Log Spreads on Rating Categories

	Coefficient	Std. Error	t-Statistic	R-squared
2005				
Intercept	8.34	0.18	45.72	0.93
Coefficient on rating	-0.34	0.01	-29.17	
2006				
Intercept	7.83	0.18	42.64	0.92
Coefficient on rating	-0.32	0.01	-27.56	
2007				
Intercept	7.67	0.18	41.49	0.91
Coefficient on rating	-0.31	0.01	-26.12	
2008				
Intercept	8.06	0.21	37.78	0.79
Coefficient on rating	-0.22	0.01	-16.28	
2009				
Intercept	7.90	0.20	39.53	0.73
Coefficient on rating	-0.17	0.01	-13.74	
2010				
Intercept	7.10	0.20	35.63	0.65
Coefficient on rating	-0.14	0.01	-11.43	

Source: IMF staff estimates.

⁸ The increasing influence of factors other than credit risk in spreads is evident in the declining explanatory power of ratings during the global crisis. For a decomposition of credit and noncredit components of spread changes see Annaert and others (2010).

Figure 1. Average CDS Spread and Ratings for Countries Rated by Moody's, 2005-2010
(CDS spread in basis points)



Sources: Markit; Moody's; and IMF staff calculations.

Note: CDS = credit default swap. Each data point represents different sovereign's average credit rating and average CDS spread throughout the year. R^2 shows the proportion of the average CDS spreads' movement that can be attributed to credit ratings. Data for 2010 are through June 30, 2010.

D. Results

17. **The event study measures the impact of rating changes and credit warnings (i.e. “reviews,” “watches,” and “outlooks”) by averaging cumulative changes in CDS spreads across individual rating events.** Changes in CDS spreads are analyzed within an *event window* of 41 days, starting from 20 days before the event until 20 days after the event. Our choice of the event window is influenced by Larrain, Reisen, and von Matzlan (1997) and Reisen and von Maltzan (1999), who show that market response precedes rating actions and then is sustained for 20 days after the rating event, especially for emerging market sovereigns. Spread changes are calculated against the level at the beginning of the event window ($t = 0$) to make them comparable across events and sovereigns. For tractability, the sovereigns are divided into advanced economies and emerging markets. Hence, within each group, the relationship between credit ratings and CDS spreads can be assumed to be linear, so that a two-notch downgrade leads to spread widening that is twice as large as that associated with a one-notch downgrade. However, the results are robust to relaxing this assumption, as discussed below. The significance is tested using a standard one-sided t-test, with the hypothesis that downgrades should lead to an increase in spreads, and upgrades to a decrease.

18. **The baseline results for Moody’s confirm the agencies’ traditional role as information providers (Table 2).** However, most of the incremental information value is transmitted through negative credit warnings (i.e., “outlooks,” “reviews,” and “watches,” see Figure 2), rather than actual rating changes.⁹ Yet even though rating changes in general have little market impact, crossing the investment-grade classification boundary is associated with a statistically significant widening of CDS spreads. This suggests that some of the market impact associated with rating changes is related to their “certification” value. In contrast, the results provide no support for the monitoring theory.

- Only 45 and 51 percent of the rating events for advanced and emerging economies, respectively, are associated with changes in spread in the expected direction. This suggests that spreads are as likely to decrease as they are to increase following a downgrade, in contrast to the empirical evidence of Cantor and Packer (1996) that links about 63 percent of rating announcements to changes in spread in the expected direction.
- In case of outlook changes, the proportion of changes in spread in the expected direction increases to 67 percent for advanced economies and 63 percent for emerging economies.

⁹ The predictive power of S&P “watch” and “outlook” notices has been documented in Hessol, Erturk, and Ontko (2007) and Chambers (2010). Hamilton and Cantor (2005) and Hartelius, Kashiwase, and Kodres (2008) document similar results for Moody’s “outlook” notices.

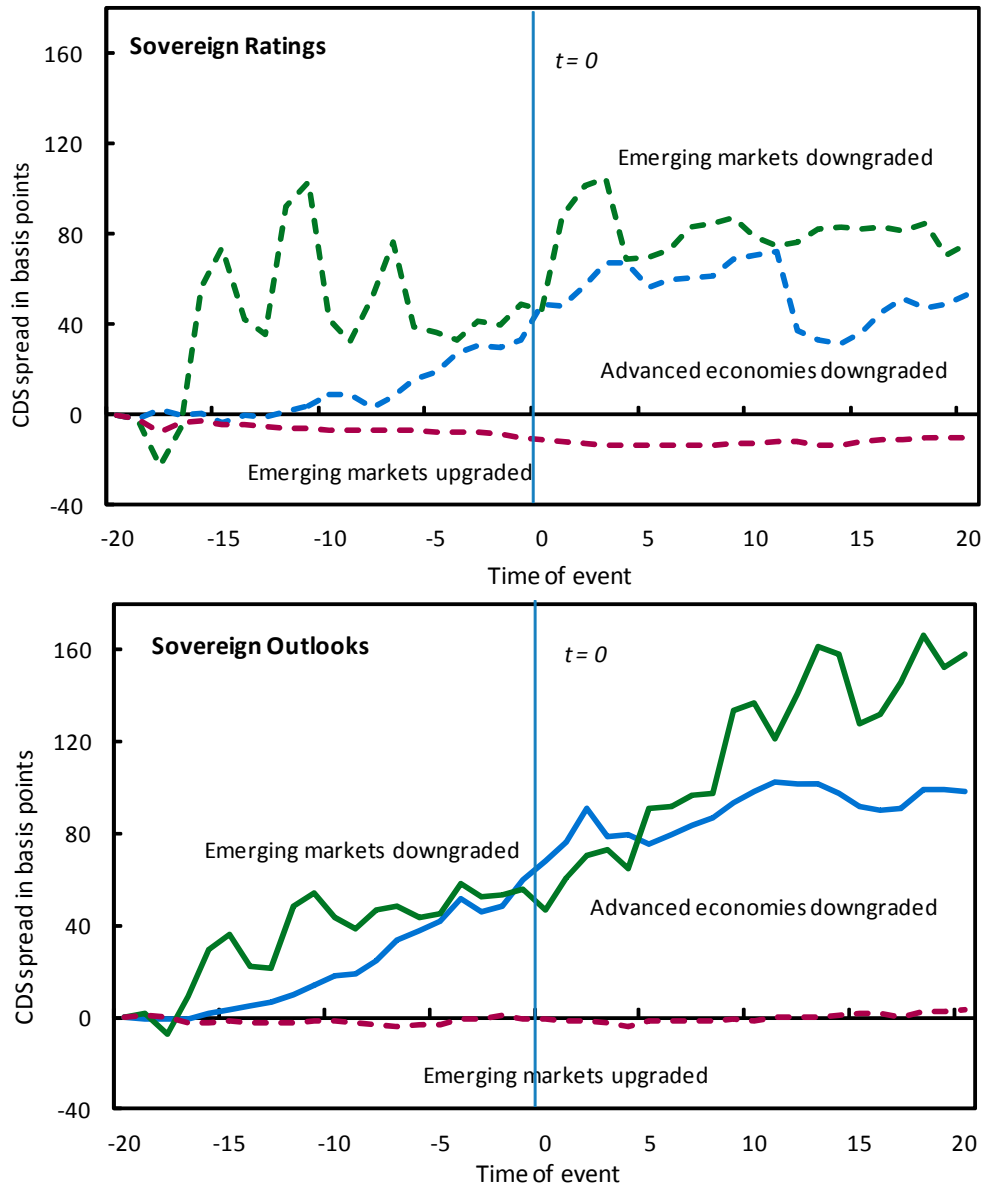
Table 2. The Empirical Tests of Rating Information Value

	Information Services	Certification Services	Monitoring Services
Theory	Ratings include new information \Rightarrow market prices react to rating actions	Classifying securities into investment-grade and high-yield grade \Rightarrow market prices react to rating changes in and out of the	Rating warnings influence issuers to take corrective actions to avert downgrades \Rightarrow market prices react to rating confirmations
Test	Event study of the impact of credit rating and outlook changes on CDS spreads	Event study of the impact of rating changes and downgrades/upgrades in and out of investment grade on CDS spreads	Event study of the impact of downgrades/upgrades preceded/not preceded by a matching warning on CDS spreads
Results	Negative outlook announcements are followed by statistically significant spread widening; 100 basis points for advance economies and 160 basis points for emerging markets. The impact of rating changes is insignificant (see Figure 2).	Downgrades through the investment grade threshold lead to statistically significant CDS spread widening of 38 basis points.	No evidence

Sources: Fitch; Moody's; and IMF staff estimates.

- Negative outlook announcements are followed by statistically significant spread widening; 100 basis points for advanced economies and 160 basis points for emerging markets. The impact of rating changes is insignificant (see Figure 2 where the dashed lines indicate statistical insignificance). This supports the view of Millon and Thakor (1985) that ratings have an information value beyond the information already available in the market. The larger impact on emerging economies is in line with the findings of Larrain, Reisen, and von Matzlan (1997) that rating actions have a significant impact—even to a degree of market overshooting—on investment-grade, emerging market securities only.
- In contrast to Cantor and Packer (1996), Reisen and von Maltzan (1999), and Ismailescu and Kazemi (2010), who associate only upgrades with significant declines in yield spreads, there is very little market reaction to upgrades.
- Although in general upgrades and downgrades do not have a statistically significant impact, specific upgrades/downgrades out of the investment grade categories (Baa3 or better) to high yield categories (Ba1 or lower) lead to a statistically significant widening of the CDS spreads of 37 basis points. This supports the view that certification matters through in and out of the investment grade categories. This result, albeit based on a few events only, suggests that some of the ratings' impact comes from the certification services and may be a purely liquidity effect (e.g. purchases and sales of assets forced by regulations and not based on the additional information already in the market).
- In order to find out if the monitoring theory holds, we disaggregate announcements into downgrades that were preceded and not preceded by a matching credit warning. As in Hill and Faff (2008), we find no evidence of monitoring services.

Figure 2. Impact of Change in Sovereign Ratings and Outlooks on CDS Spread



Sources: Markit; Moody's; and IMF staff estimates.

Note: Solid lines indicate rating events that have statistically significant impact of CDS spreads; dashed lines indicate insignificant events.

19. **The event study results are broadly robust.** The robustness checks include (1) shortening or lengthening the event window to 10 days before and after or 45 days before and after, from the base of 20 days before and after; (2) controlling for global volatility (as proxied by the S&P 500 VIX), the liquidity risk premium (as proxied by the U.S. LIBOR), and the daily business conditions in the United States, as tracked by the Aruoba-Diebold-Scotti business conditions index (Aruoba, Diebold, and Scotti, 2009); (3) using natural logarithms of spreads to account for the nonlinear relationship between the CDS spreads and rating/outlook events; (4) splitting the sample into the pre- and during-crisis periods; and (5) disaggregation of

announcements for investment/noninvestment grade sovereigns as opposed to advanced economies and emerging markets.

20. Statistical causality tests using Moody's ratings yield results consistent with the event study. In contrast to older studies that find rating events to be preceded by a similar change in the market's assessment of sovereign risk (Cantor and Packer, 1996; Larrain, Reisen, and von Matzlan, 1997; and Reisen and von Maltzan, 1999), our Granger-causality tests (Granger, 1969) suggest that information about preceding negative outlooks helps predict changes in CDS spreads of the emerging economies, even when the past values of CDS spreads are accounted for. As in Reisen and von Maltzan (1999), the causality tests are panel Granger-causality tests (Granger, 1969; and Hurlin and Venet, 2001), estimated using the generalized method of moments (GMM) with fixed effects. The instruments used are lagged S&P 500 VIX, lagged U.S. LIBOR, lagged Aruoba-Diebold-Scotti business conditions index for the United States, and CDS spreads lagged two periods.

21. An extension of the analysis to sovereign rating actions by S&P and Fitch largely confirms the above findings while pointing to two important differences.

- First, Moody's most frequently precedes its rating changes with credit reviews/watches (27 percent of rating changes are preceded by corresponding credit watches, compared with 14 percent in the case of S&P and only 8 percent in the case of Fitch). This suggests different information value of downgrades and credit watches across the CRAs.
- Second, Moody's tends to lag behind Fitch's and S&P's. An analysis of the dynamics of rating adjustments, as summarized in Table 3, reveals that Moody's and Fitch tend to follow S&P negative rating actions more often than S&P follows the others. This is consistent with conclusions of Güttler (2009) and Alsakka and ap Gwilym (forthcoming) that Moody's is more likely to adjust its rating given a rating change by S&P. In addition, on occasions when Moody's leads the rating adjustment cycle, it tends to issue credit watches rather actual downgrades. Consequently, Moody's downgrades have no significant impact on financial markets but Fitch's and S&P's do. This result, consistent with Brooks and others (2004), indicates that markets react to new information but not all CRAs convey new information through the same channels.

Table 3. Multi-Agency Rating Dynamics

	The first credit rating agency to take a negative rating action			The last credit rating agency to take a negative rating action		
	Fitch	Moody's	S&P	Fitch	Moody's	S&P
All countries	28	13	59	36	36	28
Advanced economies	22	4	74	30	52	17
Emerging economies	32	17	51	39	27	34

Sources: Fitch; Moody's; S&P; and IMF staff estimates.

Note: The table reports the frequencies with which each credit rating agency (CRA) either acts first or last within each rating adjustment cycle, defined as either until three months following the initial rating adjustment or until the originating CRA takes another action, depending on which event comes first.

III. ACCURACY AND STABILITY OF RATINGS

A. What do Ratings Measure

22. **In order to assess the accuracy of ratings, we first ask what ratings intend to measure.** This is what the agencies explain:

- *“Standard & Poor’s credit ratings are designed primarily to provide relative rankings among issuers and obligations of overall creditworthiness; the ratings are not measures of absolute default probability. Creditworthiness encompasses likelihood of default and also includes payment priority, recovery and credit stability”* (S&P, 2009).
- *“Credit ratings express risk in relative rank order, which is to say they are ordinal measures of credit risk and are not predictive of a specific frequency of default or loss. Fitch Ratings’ credit ratings do not directly address any risk other than credit risk, ratings do not deal with the risk of a market value loss on a rated security due to changes in interest rates, liquidity and other market considerations”* (Fitch Ratings, 2010).
- *Moody’s credit ratings are opinions of the credit quality of individual obligations or of an issuer’s general creditworthiness... Moody’s... ratings are opinions of the relative credit risk of fixed-income obligations.... They address the possibility that a financial obligation will not be honored as promised. Such ratings... reflect both the likelihood of default and any financial loss suffered in the event of default”* (Moody’s Investors Service, 2009).

23. **There are several common elements implicit in these definitions.** Specifically;

- Ratings represent a rank ordering of credit risk;
- Only credit risk is included;
- Ratings are not only a statement about credit risk but also about potential changes in credit risk (i.e. downgrading risk or default volatility/stability);

- Ratings are medium-term assessment that is not supposed to fluctuate following cyclical changes in the macroeconomic environment.

24. **Ratings represent a rank ordering of credit risk based on default and downgrade risks.** This rank ordering takes into account default probabilities, expected losses, and possible rating actions and warnings.¹⁰ However, there is no one-to-one mapping between any of these quantitative measures of credit risk and credit ratings since the agencies just promise to deliver a rank ordering of credit risk for obligors or obligations.

25. **In this regard, the ratings produced by the CRAs are different from the internal ratings used by banks in the context of the IRB approach to Basel II.** In the Basel approach, financial institutions use measures of PD, LGD and EAD. To assess the appropriateness of internal credit risk measurement, banks using the IRB approach are required to “have in place a process for vetting data inputs into a statistical default or loss prediction model which includes an assessment of the accuracy, completeness and appropriateness of the data specific to the assignment of an approved rating” (BCBS, 2006). Based on this requirement specific methodologies have been proposed to validate internal rating systems, see BCBS (2006).

26. **The idea of ratings as a rank ordering of credit risk is not new.** However, more recently, the agencies have made a special effort in clarifying this point, both because of criticisms made to their quantitative models as well as because some uses or ratings by investors and regulations are not fully compatible with this risk ordering idea. In this regard, the agencies have emphasized that ratings are not precise statements about default probabilities or other features of a credit loss distribution but just an assessment of relative credit risk. Moreover, ratings only measure credit risk. They do not measure liquidity, market, or interest rate risks.

27. **Ratings are not only a statement about credit risk but also about changes in credit risk—or credit stability.** As was clear from the opening quotes, rating stability (probability of being downgraded) is one dimension used by rating agencies for rating allocation. Higher rating categories are associated with a lower downgrading probability. In terms of a transition matrix, this means that the diagonal should be closer to 100 for the highest categories meaning that the highest categories will remain in the same category in the future. Stability is also related to risk-horizon. One-year transition matrices are typically more stable than 5-year (or longer) since the default probability increases with the risk horizon.

28. **Ratings are medium-term assessments and should not change due to the impact of cyclical components.** Traditionally, CRAs have tried to minimize rating volatility by rating through the cycle (TTC)—as opposed to allocating point in time (PIT) ratings. Ratings are based on the ability of an issuer to survive a trough. The rating is allocated by assessing credit risk under a stress scenario characterized by a shock to the transitory component. Once the rating is allocated, the rating is changed only in a response to changes in fundamental factors, such

¹⁰ Traditionally Moody’s ratings have been associated with expected losses while Standard & Poor’s and Fitch’s ratings have been associated with default probabilities. However, all three agencies have similar philosophical approach to ratings and use elements of default probabilities, losses given default and stability to produce their scores.

secular trends or unanticipated policy responses. Under this approach, a recession or tightening of global liquidity should not, in itself, be an occasion for a downgrade. Thus, vulnerability to cycles affects the rating decision, whereas the current position in the cycle does not. PIT rating systems tend to focus on the current conditions of an obligor.¹¹

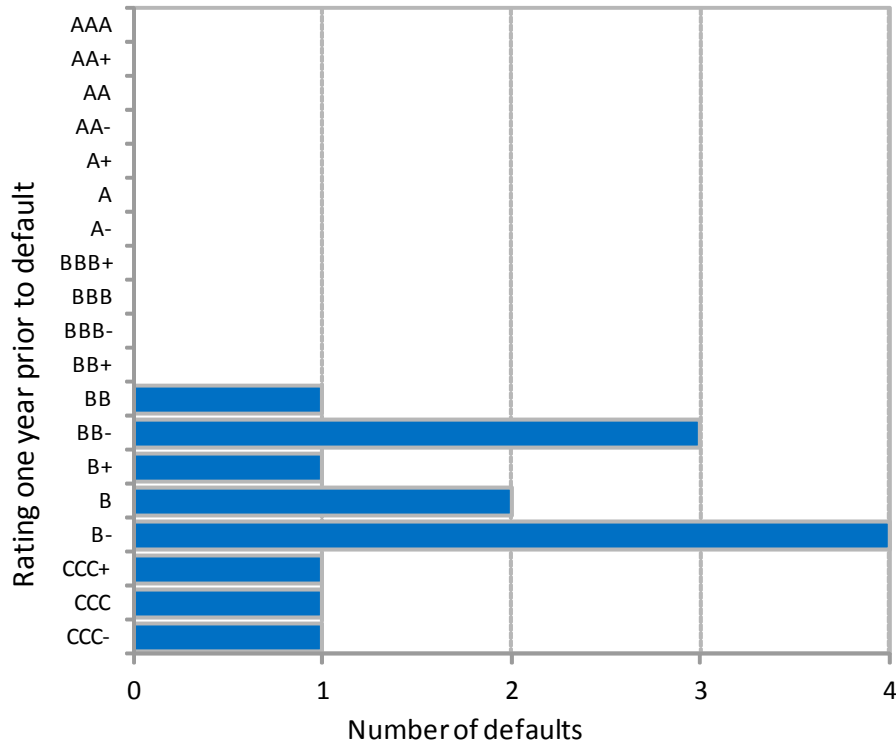
29. **More recently and faced with extreme downgrades, CRAs have started to develop new methodologies that intend to extend the TTC criteria to “through crisis criteria.”** For example Standard & Poor’s new credit stability criteria uses hypothetical stress scenarios as benchmarks for calibrating their criteria across different sectors and over time. Each scenario broadly corresponds to one of the rating categories. The scenario for a particular category reflects the level of stress that issuers or obligations rated in that category should be able to withstand without defaulting. In contrast to TTC rating, this new stability criterion allows for hypothetical scenarios affecting trend components. In this way, ratings become measures of risk conditional on the realization of extreme scenarios.

B. An Examination of Rating Accuracy

30. **Given the ordering nature of ratings produced by CRAs and the relevance given to stability, validation tests should be based on the discriminatory power of their rating system (i.e., the power to sort out ex-ante between potential defaulters and non-defaulters) as well as on the stability of ratings.** Tests based on the validation of the calibration of PD (or other parameters) are not appropriate for ratings produced by the rating agencies although they are required to be used by banks when they allocate capital on the basis of their internal rating systems. For ratings provided by rating agencies, our tests are only intended to assess if defaults take place among the lowest rating categories.

31. **The discriminatory power of sovereign ratings is validated to some extent by the fact that all of the defaults are among noninvestment grade sovereigns.** Taking into account the difficulty in predicting rare events, Figure 3 shows that all 14 sovereign defaulters involving 12 countries between 1975 and 2009 had (S&P) ratings of BB- or below, that is, all were noninvestment-grade one year prior to default and none were in the investment-grade categories.

¹¹ A change in the rating based on changes in long-term fundamentals faces the problem of recognition of fundamental versus cyclical elements. The problem is similar to those trying to maintain a fixed exchange rate regime in order to ensure exchange rate stability.

Figure 3. Ratings One Year Prior to Sovereign Default, 1975–2009

Source: Standard & Poor's.

32. **The point made above can be tested more formally with cumulative accuracy profile (CAP) curves and accuracy ratios (ARs), as done by Moody's (Sobehart, Keenan, and Stein, 2000) and S&P (S&P, 2010a).** The CAP curve is derived by plotting out the cumulative proportion of issuers by rating grade (starting by the lowest grade on the left) against the cumulative proportion of defaulters by rating grade. “Ideal” CAP curves (the red lines in Figure 4) look almost like vertical lines because all the defaulters should be among the lowest rated issuers.¹² In the “random” curve, all defaults occur randomly throughout the rating distribution (admittedly an unrealistically low bar for a CRA), so it lies along the diagonal (the green lines in Figure 4). The closer the CAP curve is to the ideal curve, the better is the discriminatory power of that CRA's ratings. The AR is the ratio of two areas: (i) the area bounded by the CAP curve and the random curve; and (ii) the area bounded by the ideal curve and the random curve. An AR

¹² In spite of the CAP curve's general use, it is not obvious that this ideal curve is a good estimate of “best” rating performance. Even good credits have positive default probabilities. This means that the ideal curve (in which all defaults happen among the worst credits) does not measure “best” rating performance.

equal to 1 is equivalent to perfect discriminatory power while an AR equal to 0 implies no discriminatory power.¹³

33. **The S&P CAP curves for sovereigns at different risk horizons tend to suggest that, as expected, discriminatory power is a function of time, since some power is lost as the risk horizon increases** (Figure 4). This loss of power is indicated by the ARs: 92 percent over the one-year horizon, 82 percent over three years, 80 percent over five, and 84 percent over 10. Hence, the CRAs are better at predicting defaults over short horizons than long ones. ARs that were calculated for S&P corporate ratings suggest that they are better at pinpointing potential defaulters among sovereign issuers than among corporates. The corporate rating ARs over a one-year horizon are 77 percent for financials and 81 percent for nonfinancials (versus 92 percent for sovereigns), and 63 percent and 71 percent over five years (versus 80 percent).¹⁴

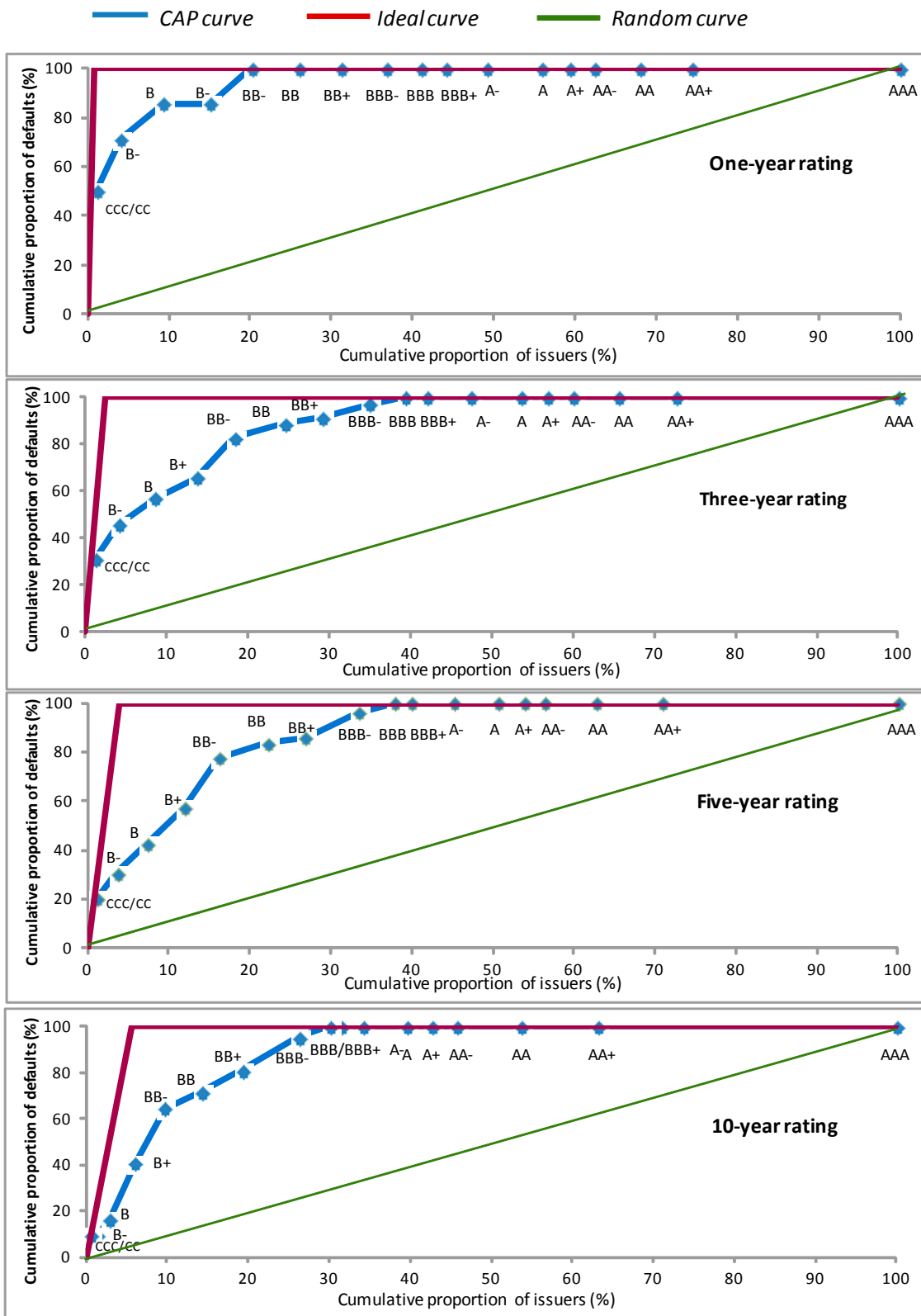
34. **A welcome contribution of the CRAs would be the calibration of ratings to target credit risk metrics, such as default probabilities, and the publication of validation tests against such standards.** This would make the message embedded in a rating more transparent and would also allow more relevant tests of accuracy and—in the end—review of the quality of the work produced by CRAs. The Basel Committee on Banking Supervision (BCBS, 2005) suggests a number of cardinal accuracy tests, but they presuppose specific default probabilities for individual rating grades. However, the CRAs put little emphasis on such tests, preferring to focus on ordinal accuracy tests, because they do not target specific risk metrics for their ratings.¹⁵

¹³ Engelmann, Hayden, and Tasche (2003) propose a statistical test of the accuracy ratios that measures the quality of credit rating models. However, the test requires that the sample contains at least 50 defaults and thus cannot be applied to the sovereign dataset.

¹⁴ The S&P sovereign ARs were estimated based on 1975–2009 data, and the corporate ARs based on 1981–2009 data.

¹⁵ Nevertheless, Moody's, for example, conducts some cardinal accuracy assessments—investment-grade default rates over multiple horizons and average ratings of defaulting issuers up to three years prior to default (Cantor and Mann, 2003). Alternatively, cardinal accuracy tests could be conducted against the target default probabilities (in the case of Fitch and S&P) or expected losses (Moody's), which have been used to rate structured credit products (Fender and Kiff, 2005).

Figure 4. Sovereign Foreign-Currency Rating Performance by Standard & Poor's



Source: Standard & Poor's.

Note: CAP = cumulative accuracy profile.

C. An Examination of Rating Stability

35. **There is an inevitable tradeoff between rating accuracy and stability, but users have expressed a preference for stability.** This is driven by the certification role played by ratings, and the transaction costs induced by trading when ratings change frequently. In an attempt to satisfy their users, CRAs use several mechanisms to promote stability. Though of less interest to CRAs, rating stability is important at a systemic level, since rating downgrades can be associated with forced sales.

36. **In practice, rating agencies seek to ensure rating stability by focusing on rank ordering credit risk, instead of rating to specific credit metrics.** Hence, the impact of cyclical fluctuations on ratings is automatically muted. Also, CRAs generally try to rate “through the cycle,” as opposed to rating based on “point in time” information. Conceptually, this means that they rate based on the issuer’s ex-ante perceived ability to survive cyclical troughs, which provides a cushion against the impact of economic downturns. Lastly, the CRAs apply smoothing rules that, for example, change ratings only if (1) the anticipated rating change is expected to be persistent, and/or (2) the prescribed change is more than one notch (Moody’s, 2006).

37. **Figures 5 and 6 show that higher ratings are more stable than lower ones, and that sovereign ratings are more stable than corporate ones.**¹⁶ Using S&P data from 1975, Figure 5 shows the average percent of ratings that remain at the same level over one-year horizons by rating grade. For example, 82.1 percent of A- rated sovereigns were still A- rated at the end of each year, versus 77.5 percent of the corporates. It clearly shows that the percent of unchanged ratings decreases as rating categories decline. Figure 6 shows the average percent of ratings that are downgraded by three or more notches, a definition of “rating failure” suggested by Bhatia (2002) and S&P (2010b).¹⁷ For example, 1.8 percent of A-rated sovereigns were downgraded three or more notches, versus 6.9 percent of corporates.

38. **However, Figures 7 and 8 show that these long-term averages hide signs of significant rating instability during periods of market stress, suggesting that smoothing techniques work less well during such times.** The figures summarize Moody’s upgrades and downgrades during the Asian crisis and the recent financial crisis. Sovereigns on the 45 degree line maintained their ratings while those below (above) were downgraded (upgraded).¹⁸ The figures show that only 68 percent of ratings remained unchanged during the Asian crisis and 63

¹⁶ Figures 5 and 6 are based on S&P ratings, but similar results were found for Fitch and Moody’s ratings (see Appendix 1).

¹⁷ The three or more notch stability threshold is also in line with a practice used by CRAs in their stability studies (Moody’s, 2010).

¹⁸ Some caution must be exercised in interpreting Figures 7 and 8. Although the equivalent Fitch and S&P figures are very similar (see Appendix 2), the number of rated sovereigns has grown significantly. For example, Moody’s rated 69 sovereigns in the late 1990s, but now rates 104.

percent during the current crisis (so far). However, the proportion of sovereigns that started each period as investment grade credits but finished as high yield credits fell from 2.9 percent to 1.9 percent. Similar conclusions would be drawn from a similar analysis of Fitch and S&P sovereign rating transitions (Appendix 2).

39. The size of the average rating change also fell between the Asian and the current crises. For Moody's, the average rating change fell from a downgrade by less 0.4 notch during the Asian crisis to a 0.07 notch during the current crisis, though this difference is not statistically significant. Focusing on the magnitude of the large rating changes, Tables 4 and 5 show that big downgrades (three or more notches) have been concentrated in the investment-grade categories, even though these higher rating categories are supposed to be more stable than lower rating categories by design. All of this highlights the fact that risk is a forward-looking measure that is difficult to estimate when conditions change drastically.

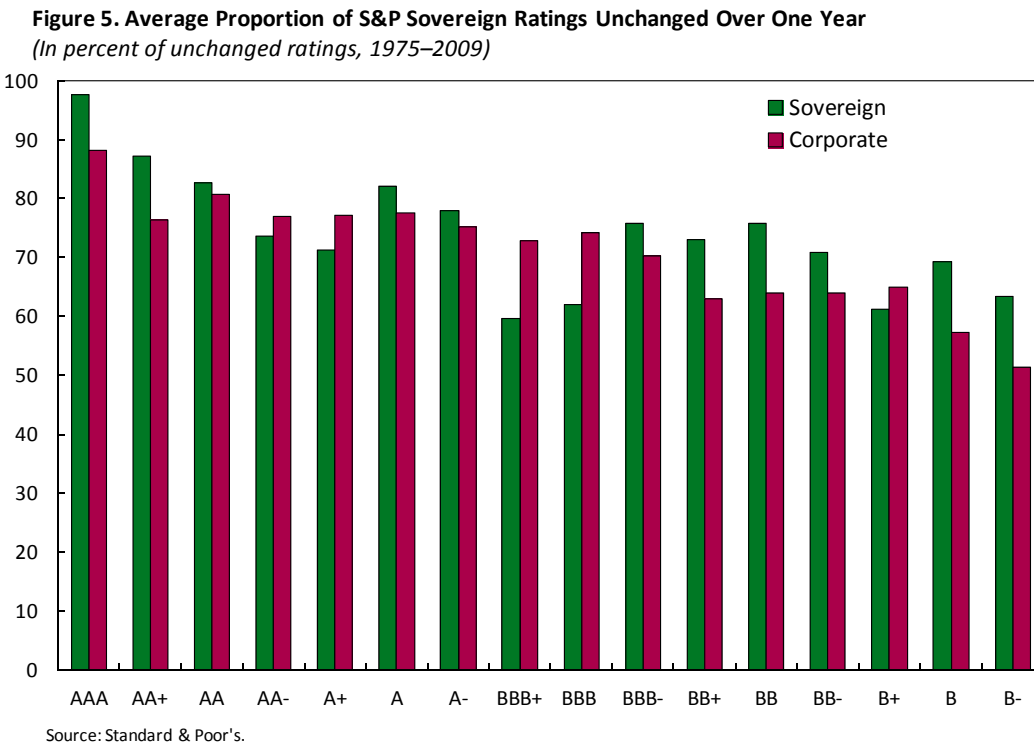
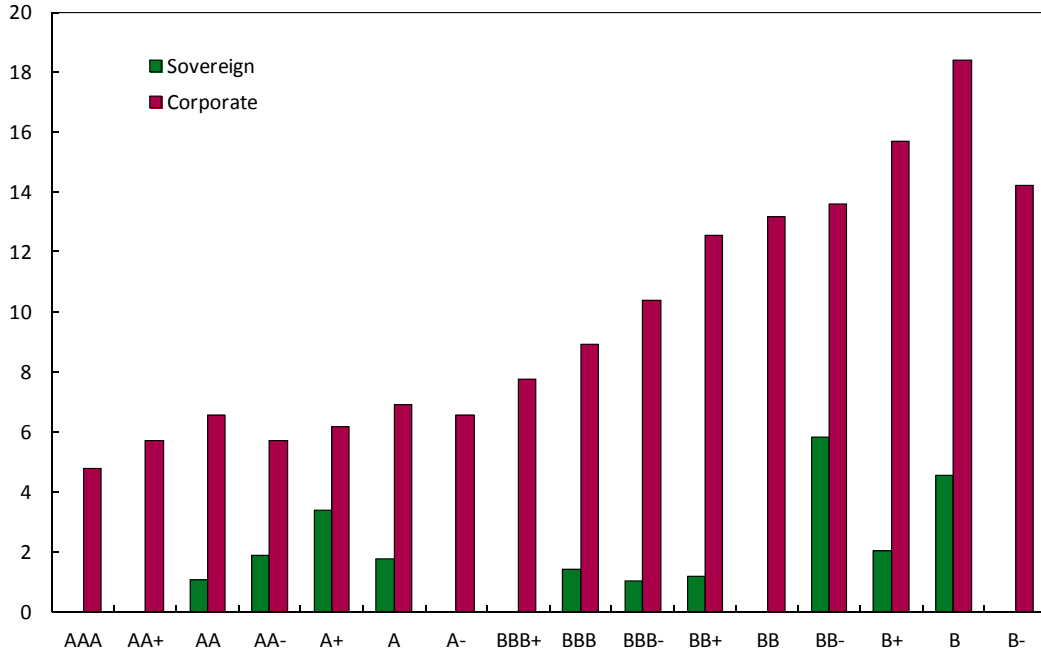
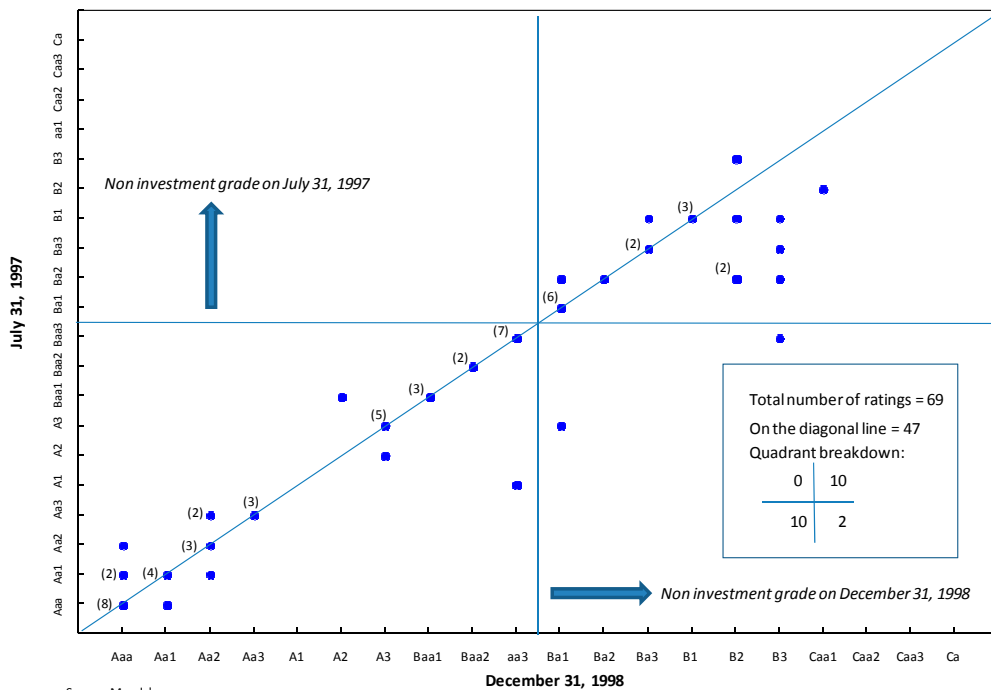


Figure 6. Average Proportion of S&P Sovereign Ratings Downgraded More Than Two Notches Over One Year
(In percent of ratings downgraded more than two notches, 1975–2009)



Source: Standard & Poor's.

Figure 7. Asian Crisis: Sovereigns Rated by Moody's between July 31, 1997 and December 31, 1998



Source: Moody's .

Note: The numbers in the brackets indicate the number of issuers associated with that rating.

Figure 8. Current Crisis: Sovereigns Rated by Moody's between July 31, 2007 and June 30 ,2010

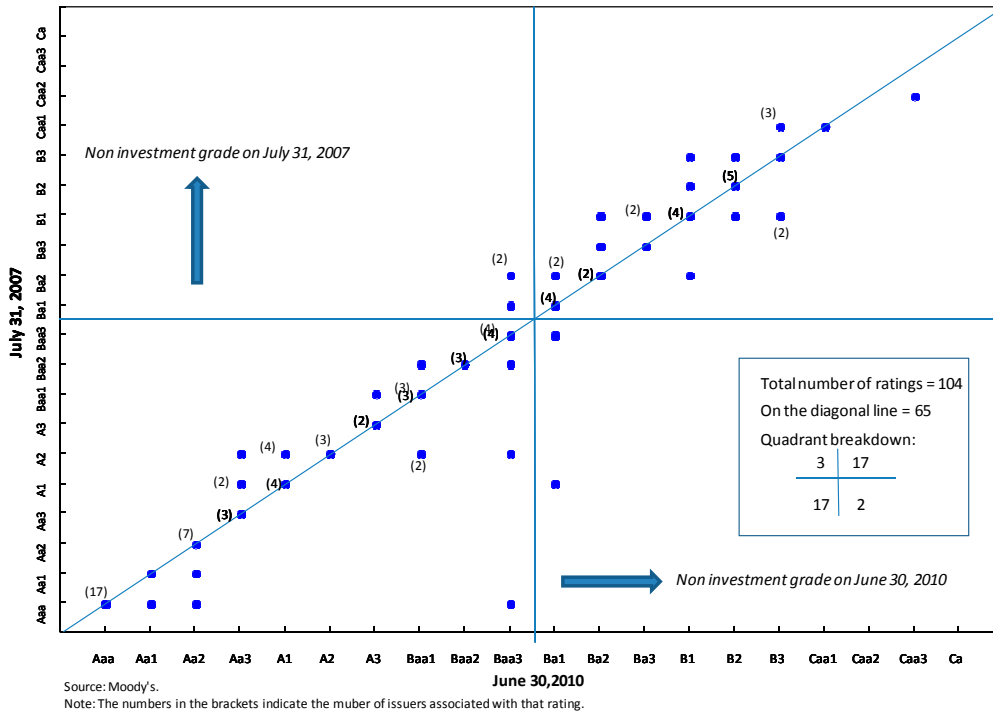


Table 4. Sovereign Rating "Failures" During the 1997-1998 Asian Crisis

	Fitch			Moody's			S&P		
	Start	End	Notches	Start	End	Notches	Start	End	Notches
Indonesia	BBB-	B-	-6	Baa3	B3	-6	BBB	B-	-7
Korea ¹	AA-	B-	-12				AA-	B+	-10
Korea ¹	B-	BBB-	6				B+	BBB-	4
Malaysia				A1	Baa3	-5	A+	BBB-	-5
Romania				Ba3	B3	-3	BB-	B-	-3
Russia	BB+	B-	-5	Ba2	B3	-4	BB-	B-	-3
Thailand				A2	Ba1	-5	A	BBB-	-4
Venezuela				Ba2	B2	-3			

Sources: Fitch; Moody's; and Standard & Poor's.

Note: Table shows successive downgrades or upgrades by three or more notches in aggregate during any rolling 12-month period, excluding downgrades or upgrades into, out of, within, or between rating categories below B- or B3; 1997 through January 1999.

¹Korea was downgraded by Fitch and S&P in October 1997–December 1997; and then a series of upgrades occurred between February 1998 and January 1999.

Table 5. Sovereign Rating "Failures" During the 2008-2010 Crisis

	Fitch			Moody's			S&P		
	Start	End	Notches	Start	End	Notches	Start	End	Notches
Greece	A	BBB-	-4	A1	Ba1	-6	A-	BB+	-4
Iceland ¹	A+	BBB-	-5	Aaa	Baa1	-7	A+	BBB-	-5
Iceland ¹				A1	Baa3	-5			
Ireland	AAA	AA-	-3						
Latvia	BBB+	BB+	-3	A2	Baa3	-4	BBB+	BB	-4
Lithuania	A	BBB	-3						
San Marino	AA	A	-3						

Sources: Fitch; Moody's; and Standard & Poor's.

Note: Table shows successive downgrades or upgrades by three or more notches in aggregate during any rolling 12-month period, excluding downgrades or upgrades into, out of, within, or between rating categories below B- or B3; 2007 through June 2010.

¹The Iceland downgrades by Moody's involve overlapping periods. The first period includes downgrades from May 2008 through end-December 2008, while the second period includes downgrades from December 1 2008 through end-November 2009. That is, both periods include the 3 notch downgrade on December 4, 2008.

40. Confronted with these evident “failures,” CRAs have been developing rating procedures that would condition a rating to the country’s ability to survive a crisis.

However, it is not clear whether this methodology enhances stability beyond that provided by the TTC approach. Kiff, Kisser, and Schumacher (2010) show that although TTC ratings are inherently more stable than point-in-time assessments, some of this stability is undermined by the aforementioned rating change smoothing rules. If the smoothing rule’s prescribed downgrade turns out to be persistent, and the creditworthiness of the issuer has possibly worsened, a more abrupt downgrade is required. In other words, by attempting to smooth ratings, typical TTC methodologies run the danger of actually exacerbating procyclical cliff effects. Even “through the crisis” methodologies will be prone to these smoothing-induced cliff effects, although perhaps with less frequency, because of the more severe ex-ante stress tests.

IV. SOME POLICY RECOMMENDATIONS

- Based on the evidence of the impact of the CRA’s certification services, the removal of the excessive reliance of regulations on ratings is warranted. This will not affect the information value of ratings—that appears to work mostly through outlook reviews—and will help lessen the additional liquidity impact due to the need to meet regulations, reducing potential cliff effects.
- To the extent that ratings continue to play a significant role in regulations, an issue arises as to whether CRAs should be more transparent about the quantitative measures they calibrate in the rating process (PDs, LGD, and stability assumptions), how these measures are mapped into ratings, and whether the final ratings can be used to infer the parameters used to obtain these measures. This is particularly relevant in the use of external ratings by banks employing the standardized approach in Basel II since internal ratings systems are subject to rigorous back testing.

- Moreover, recent heavy downgrade activity suggests that ratings should embed the notion that risk is a forward looking dimension conditional on the macroeconomic scenario. In this regard, ratings should be better tied to macroeconomic conditions, including their resilience to stress scenarios.

REFERENCES

- Alsakka, Rasha, and Owain ap Gwilym, 2010, “Leads and Lags in Sovereign Credit Ratings,” *Journal of Banking & Finance*, Vol. 34, No. 11, pp. 2614–26.
- Alsakka, Rasha and Owain ap Gwilym, forthcoming, “Rating Agencies’ Credit Signals: An Analysis of Sovereign Watch and Outlook,” *International Review of Financial Analysis*.
- Annaert, Jan, Marc De Ceuster, Patrick Van Roy, and Cristina Vespro, 2010, “What Determines Euro Area Bank CDS Spreads?” Working Paper 190 (Brussels: National Bank of Belgium).
- Aruoba, S. Borağan, Francis X. Diebold, and Chiara Scotti, 2009, “Real-Time Measurement of Business Conditions,” *Journal of Business and Economic Statistics*, Vol. 27, pp. 417–27.
- Basel Committee on Banking Supervision (BCBS), 2005, “Studies on the Validation of the Internal Rating Systems,” BCBS Working Papers No. 14 (Basel: Basel Committee on Banking Supervision).
- _____, 2006, “International Convergence of Capital Measurement and Capital Standards: A Revised Framework—Comprehensive Version” (Basel: Basel Committee on Banking Supervision).
- _____, 2009, “Stocktaking on the Use of Credit Ratings” (Basel: Basel Committee on Banking Supervision).
- Bhatia, Ashok Vir, 2002, “Sovereign Credit Ratings Methodology: An Evaluation,” IMF Working Paper 02/170 (Washington: International Monetary Fund).
- Boot, Arnoud W. A., Todd T. Milbourn, and Anjolein Schmeits, 2006, “Credit Ratings as Coordination Mechanisms,” *Review of Financial Studies*, Vol. 19, pp. 81–118.
- Brooks, Robert, Robert Faff, Joe Hillier, and David Hillier, 2004, “The National Market Impact of Sovereign Rating Changes,” *Journal of Banking and Finance*, Vol. 28, pp. 233–50.
- Cantor, Richard, and Christopher Mann, 2003, “Measuring the Performance of Corporate Bond Ratings,” Moody’s Investors Services Special Comment (April).
- Cantor, Richard, and Frank Packer, 1996, “Determinants and Impact of Sovereign Credit Ratings,” *FRBNY Economic Policy Review*, October, pp. 37–54 (New York: Federal Reserve Bank of New York).
- Chambers, John, 2010, “Outlooks: The Sovereign Credit Weathervane, 2009/2010 Update,” Standard & Poor’s RatingsDirect, April 20.
- Dichev, Iliia, and Joseph Piotroski, 2001, “The Long-Run Stock Returns Following Bond Ratings Changes,” *Journal of Finance*, Vol. 56, pp. 173–203.

- Ederington, Louis, Jess Yawitz, and Brian Roberts, 1987, "The Informational Content of Bond Ratings," *Journal of Financial Research*, Vol. 10, pp. 211–26.
- Engelmann, Bernd, Evelyn Hayden, and Dirk Tasche, 2003, "Testing Rating Accuracy," *Risk*, Vol. 16, pp. 82–86.
- Fender, Ingo, and John Kiff, 2005, "CDO Rating Methodology: Some Thoughts on Model Risk and its Implications," *Journal of Credit Risk*, Vol. 1, pp. 37–58.
- Ferri, Giovanni, Li-gang Liu, and Joseph E. Stiglitz, 1999, "The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis," *Economic Notes*, Vol. 28, pp. 335–55.
- Fitch Ratings, 2010, "Definitions of Ratings and Other Forms of Opinion," Fitch Ratings, August.
- Granger, Clive W. J., 1969, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, Vol. 37, pp. 424–38.
- Güttler, Andre, 2009, "Lead-lag Relationships and Rating Convergence among Credit Rating Agencies," Research Paper 09-14 (London: European Business School).
- Hand, John, Robert Holthausen, and Richard Leftwich, 1992, "The Effect of Bond Rating Agency Announcements on Bond and Stock Prices," *Journal of Finance*, Vol. 47, pp. 733–52.
- Hamilton, David T., and Richard Cantor, 2005, "Rating Transitions and Defaults Conditional on Rating Outlooks Revisited: 1995–2005," Moody's Investor Service Special Comment, December.
- Hartelius, Kristian, Kenichiro Kashiwase, and Laura Kodres, 2008, "Emerging Market Spread Compression: Is it Real or Is It Liquidity?" IMF Working Paper 08/10 (Washington: International Monetary Fund).
- Hessol, Gail I., Erkan Erturk, and Jason Ontko, 2007, "Credit Watch and Rating Outlooks Provide Powerful Warning Signals," Standard & Poor's RatingsDirect, August 7.
- Hill, Paula, and Robert Faff, 2008, "Do Credit Watch Procedures Affect the Information Content of Sovereign Credit rating Changes? Unpublished working paper.
- Hull, John, Mirela Predescu, and Alan White, 2004, "The Relationship between Credit Default Swap Spreads, Bond Yields, and Credit Ratings Announcements," *Journal of Banking and Finance*, Vol. 28, pp. 2789–2811.
- Hurlin, Christophe, and Baptiste Venet, 2001, "Granger Causality Tests in Panel Data Models with Fixed Coefficients, Working Paper Eurisco 2001-09 (Paris: Université Paris IX Dauphine).
- International Monetary Fund (IMF), 1999, *International Capital Markets* (Washington, July).

- _____, 2010, *Global Financial Stability Report*, World Economic and Financial Surveys (Washington, October).
- Ismailescu, Iuliana, and Hossein Kazemi, 2010, "The Reaction of Emerging Market Credit Default Swap Spreads to Sovereign Credit Rating Changes," *Journal of Banking and Finance*, forthcoming.
- Kaminsky, Graciela, and Sergio Schmukler, 2002, "Emerging Market Instability: Do Sovereign Ratings Affect Country Risk and Stock Returns?" *The World Bank Economic Review*, Vol.16, pp. 171–95.
- Kiff, John, Michael Kisser, and Liliana Schumacher, forthcoming, "An Inspection of the Through-the-Cycle Rating Methodology," IMF Working Paper (Washington: International Monetary Fund).
- Larrain, Guillermo, Helmut Reisen, and Julia von Maltzan, 1997, "Emerging Market Risk and Sovereign Credit Ratings" *OECD Development Centre*, Working Paper 124, April 1997
- Liu, Pu, and Anjan Thakor, 1984, "Interest Yields, Credit Ratings and Economic Characteristics of State Bonds," *Journal of Money, Credit and Banking*, Vol. 16, pp. 344–51.
- Micu, Mirian, Eli Remolona, and Philip Wooldridge, 2006, "The Price Impact of Rating Announcements: Which Announcements Matter?" Working Paper 207 (Basel: Bank of International Settlements).
- Millon, Marcia, and Anjan Thakor, 1985, "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies," *Journal of Finance*, Vol. 40, pp. 1403–22
- Moody's Investors Service, 2009, "Rating Symbols and Definitions," Moody's Investors Service, November.
- Moody's Investors Service, 2010, "Rating Symbols and Definitions," Moody's Investors Service, July.
- Norden, Lars, 2008, "Credit Derivatives. Corporate News, and Credit Ratings," University of Mannheim Working Paper, October 18.
- Reisen, Helmut, 2003, "Ratings since the Asian Crisis," OECD Development Centre Working Paper 214 (Paris: OECD Development Centre).
- Reisen, Helmut, and Julia von Maltzan, 1999, "Boom and Bust and Sovereign Ratings," *International Finance*, Vol. 2, p. 273–93.
- Sobehart, Jorge R., Sean C. Keenan, and Roger M. Stein, 2000, "Benchmarking Quantitative Default Risk Models: A Validation Methodology," Moody's Investors Service Rating Methodology, March.
- Standard & Poor's (S&P), 2009, "Understanding Standard & Poor's Rating Definitions," Standard & Poor's RatingsDirect, June 3.

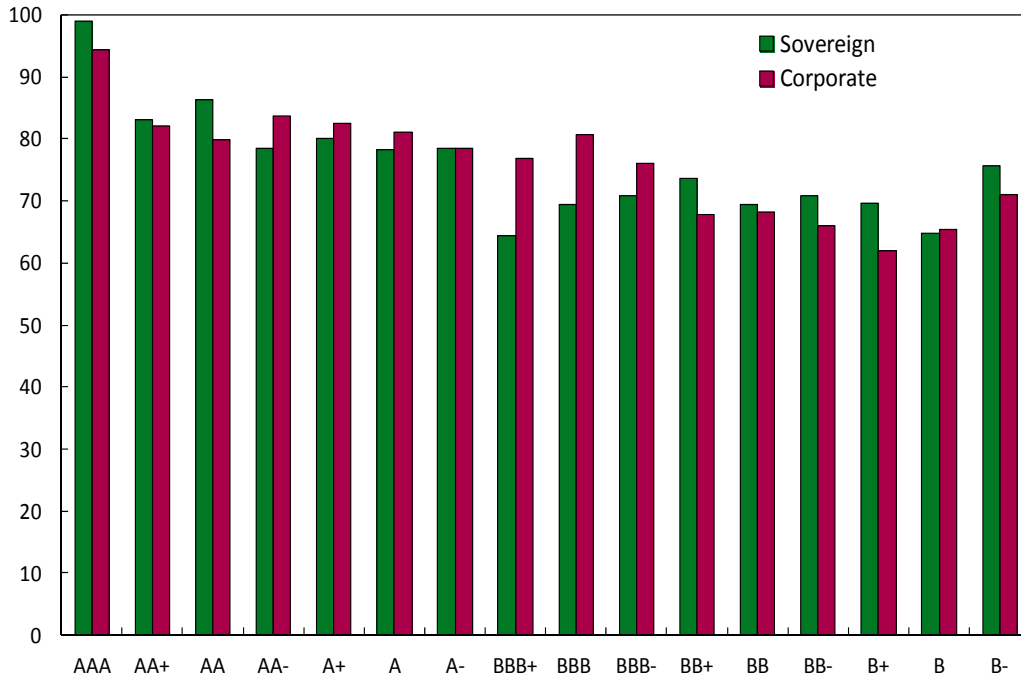
_____, 2010a, “Another Perspective on Rating Comparability and Performance,” Standard & Poor’s RatingsDirect, April 9.

_____, 2010b, “Methodology: Credit Stability Criteria,” Standard & Poor’s RatingDirect, May 3.

Sy, Amadou, 2002, “Emerging Market Bond Spreads and Sovereign Credit Ratings: Reconciling Market Views with Economic Fundamentals,” *Emerging Markets Review*, Vol. 3, pp. 380–408.

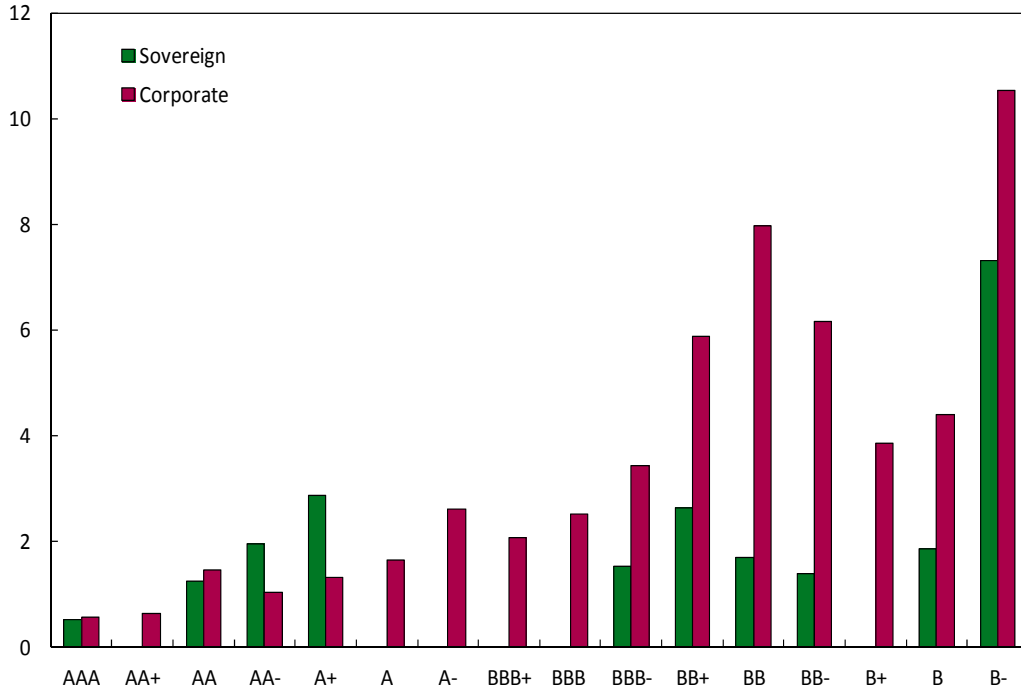
APPENDIX 1: ANALYSIS OF ONE-YEAR FITCH AND S&P SOVEREIGN RATING CHANGES

Figure A1.1. Average Proportion of Fitch Sovereign Ratings Unchanged Over One Year
(In percent of unchanged ratings)



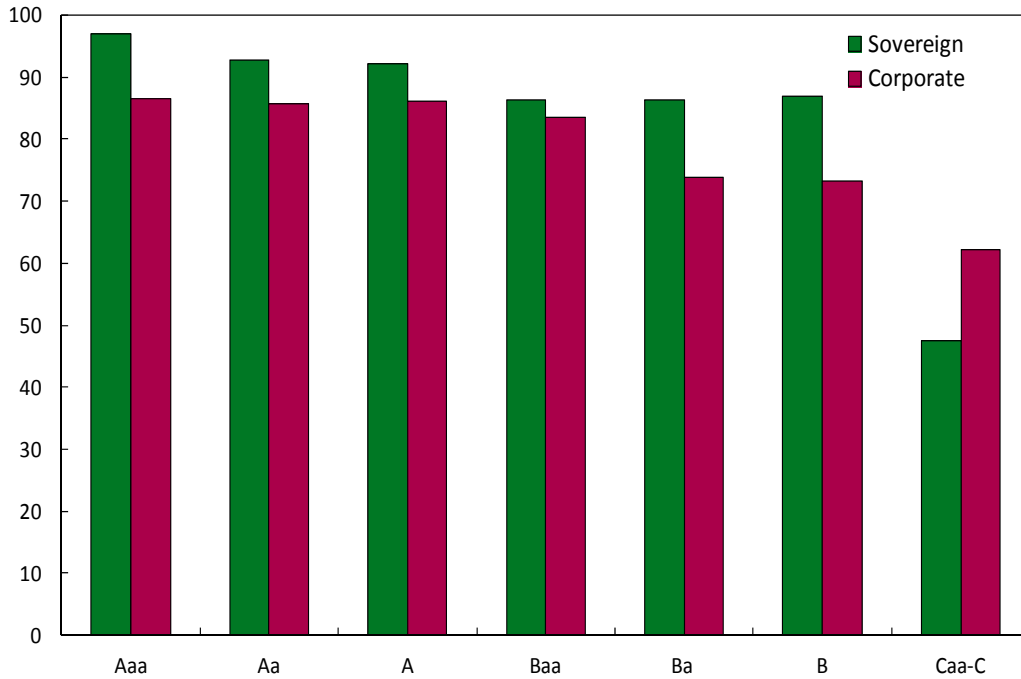
Source: Fitch Ratings.
Note: Sovereign ratings for 1995-2009 and corporate ratings for 1990-2009.

Figure A1.2. Average Proportion of Fitch Sovereign Ratings Downgraded More Than Two Notches Over One Year
(In percent of ratings downgraded more than two notches)



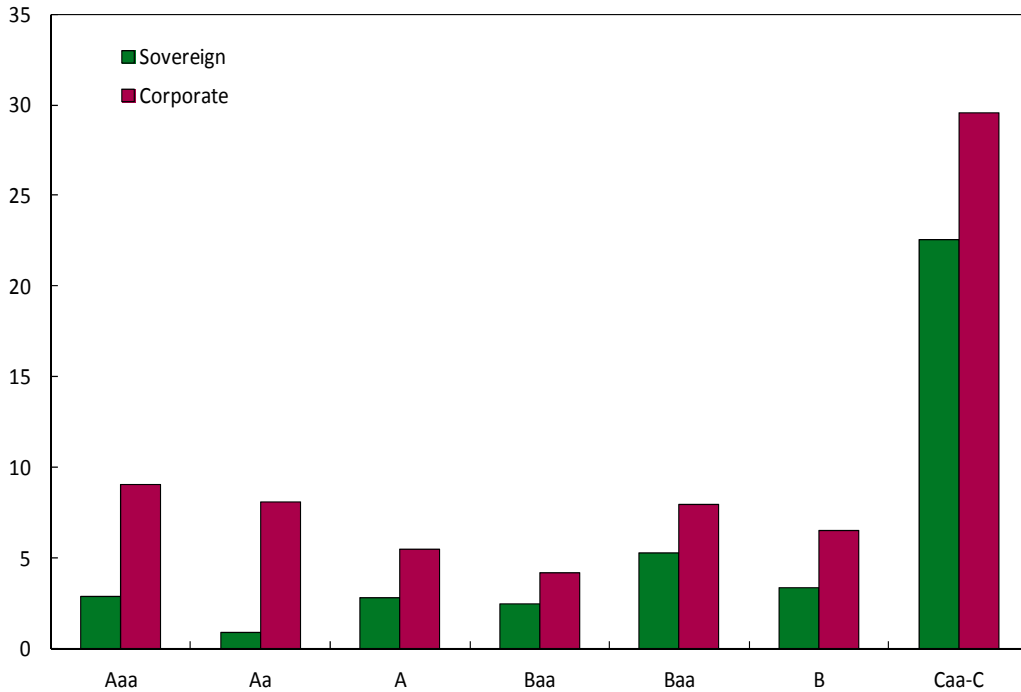
Source: Fitch Ratings.
Note: Sovereign ratings for 1995-2009 and corporate ratings for 1990-2009.

Figure A1.3. Average Proportion of Moody's Sovereign Ratings Unchanged Over One Year
(In percent of unchanged ratings)



Source: Moody's Investors Service.
 Note: Both sovereign and corporate ratings for 1983-2009.

Figure A1.4. Average Proportion of Moody's Sovereign Ratings Downgraded Through at Least One Grade Over One Year
(In percent of ratings downgraded at least one grade)



Source: Moody's Investors Service.
 Note: Both sovereign and corporate ratings for 1983-2009.

APPENDIX 2: SOVEREIGN DOWNGRADES AND UPGRADES DURING THE ASIAN AND GLOBAL CRISIS

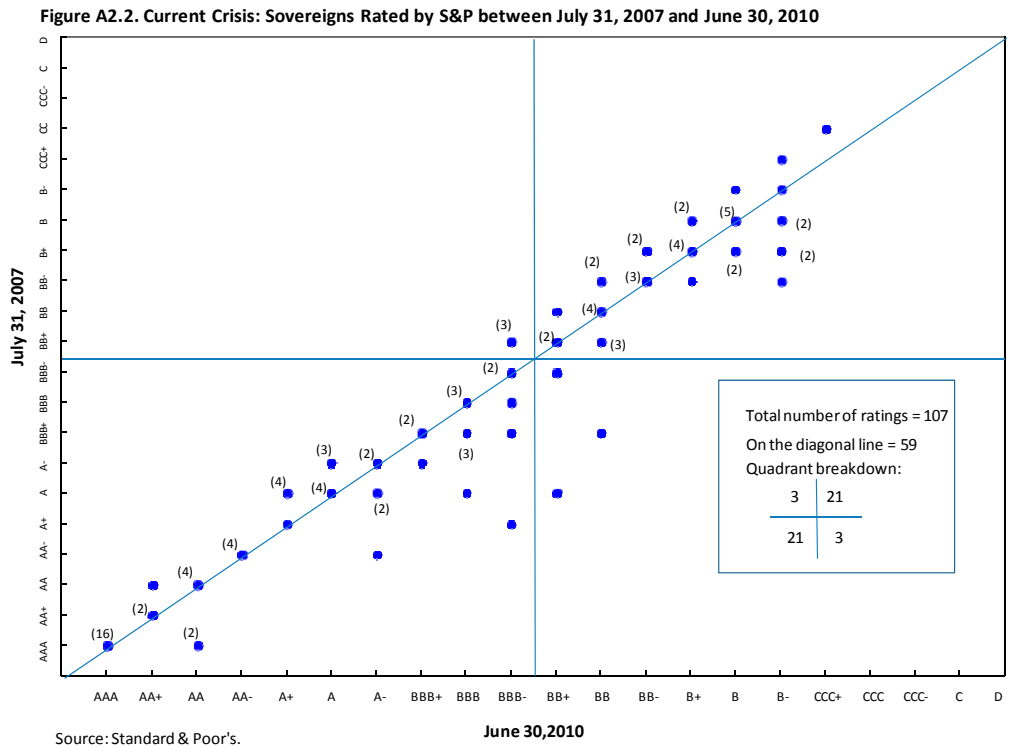
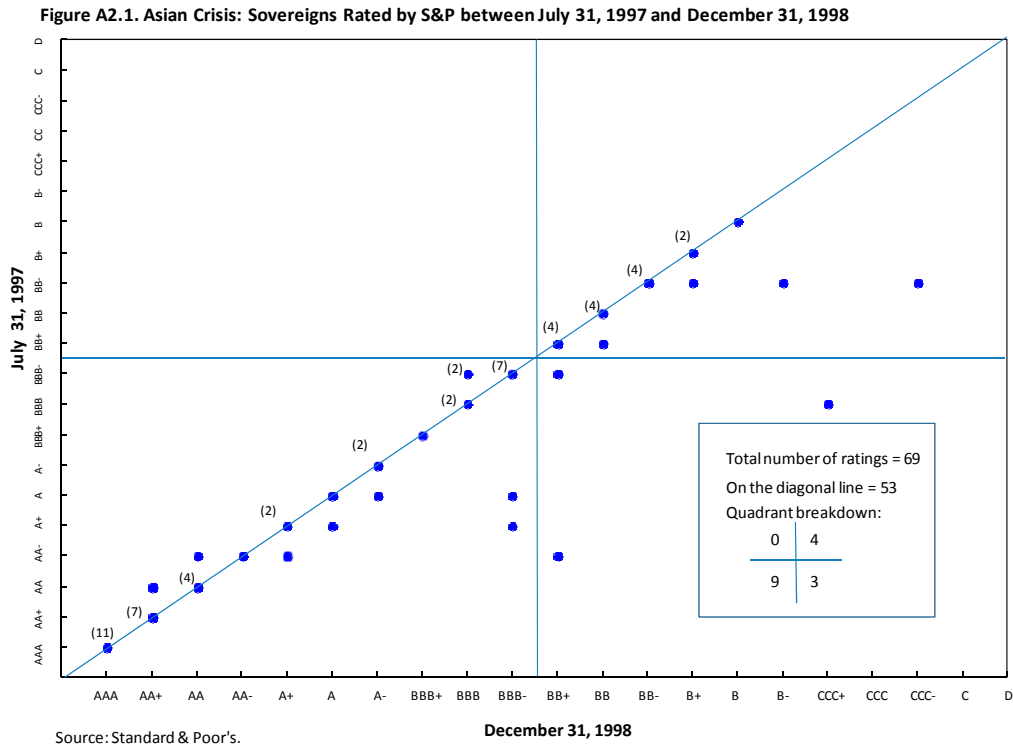
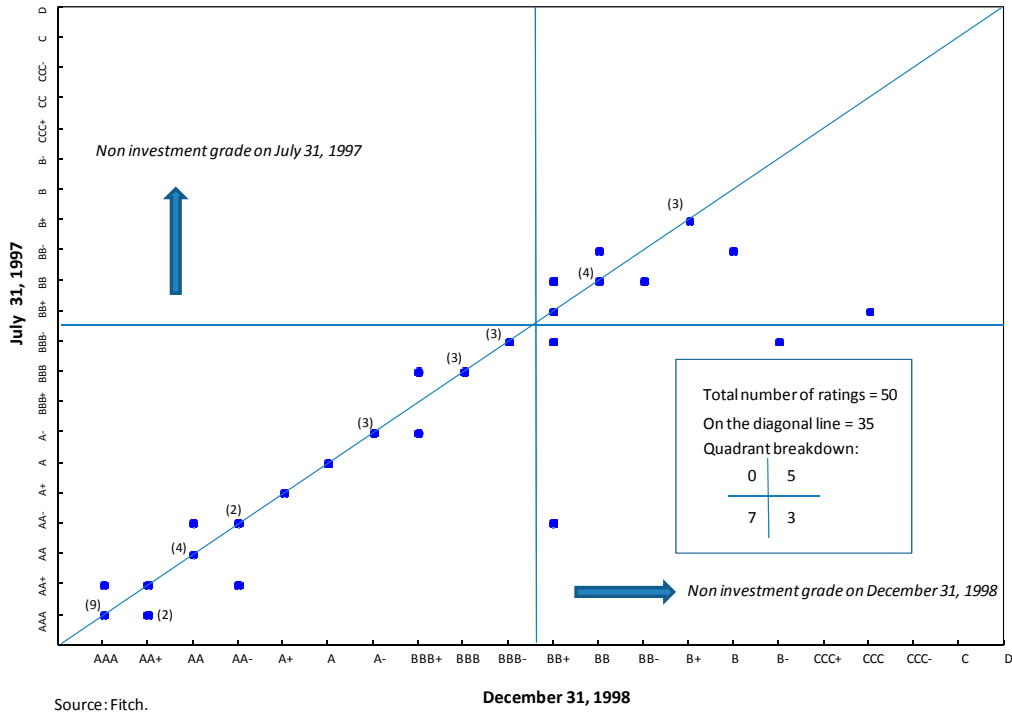
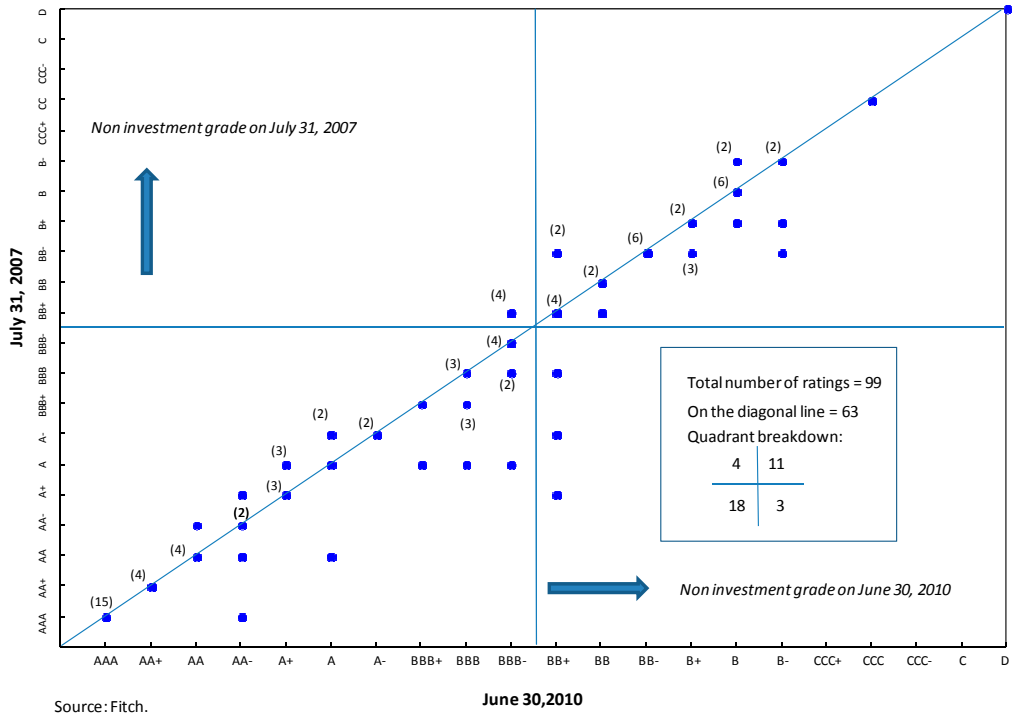


Figure A2.3. Asian Crisis: Sovereigns Rated by Fitch between July 31, 1997 and December 31, 1998



Source: Fitch.

Figure A2.4. Current Crisis: Sovereigns Rated by Fitch between July 31, 2007 and June 30, 2010



Source: Fitch.