American University of Beirut
Institute of Financial Economics

Lecture and Working Paper Series No. 1, 2005

Sovereign Credit Ratings: Guilty Beyond Reasonable Doubt?[1]

Nada Mora[2]

American University of Beirut

1. This paper is based on a longer version available through the Social Science Research Network, at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=561141

2. I wish to thank the director of the Institute of Financial Economics, Dr. Samir Makdisi, for his support in producing this paper as part of the Institute of Financial Economics Working Paper Series. I would also like to thank participants in the 2004 European Finance Association Annual Meeting in Maastricht. I am also grateful to an anonymous referee and Roberto Rigobon for many valuable suggestions. This paper is a revision of an earlier paper on which useful comments were received from Jerry Hausman and participants at the MIT international workshop.

IFE Lecture and Working Paper Series

This series of guest lectures and working papers is published by the Institute of Financial Economics (IFE) at the American University of Beirut (AUB) as part of its role in making available ongoing research, at the University and outside it, related to economic issues of special concern to the developing countries. While financial, monetary and international economic issues form a major part of the institute's work, its research interests are not confined to these areas, but extend to include other domains of relevance to the developing world in the form of general analysis or country specific studies.

Except for minor editorial changes, the lectures are circulated as presented at public lectures organized by the institute, while working papers reflect on-going research intended to be polished, developed and eventually published. Comments on the working papers, to be addressed directly to the authors, are welcome.

Sovereign Credit Ratings: Guilty Beyond Reasonable Doubt?

Nada Mora
American University of Beirut

Abstract

This paper questions the view that credit rating agencies aggravated the East Asian crisis by excessively downgrading those countries. I find that ratings are, if anything, sticky rather than excessively procyclical. Assigned ratings exceeded predicted ratings before the crisis, mostly matched predicted ratings during the crisis period, and did not increase as much as predictions in the period after the crisis. Ratings are also found to react to nonmacroeconomic factors such as lagged spreads and a country's default history. Therefore it is questionable that ratings exacerbate the boom-bust cycle if they are simply reacting to news, whether macroeconomic or market.

Introduction

In the aftermath of the Asian crisis of 1997, followed by Russia in 1998, and Brazil in 1999, much attention by the media and policymakers has turned to sovereign credit ratings. Rating agencies have been criticized for failing to predict the Asian crisis and for exacerbating the crisis when they downgraded the countries in the midst of the financial turmoil. The IMF (1998) highlights how the rating agencies reacted late when they downgraded the Asian countries. No sovereign credit rating was downgraded throughout 1996 or the first half of 1997 for the East Asian countries, with the exception of Thailand by Moody's in April 1997. During the crisis Indonesia, Korea, and Thailand were downgraded to below investment-grade. This was also the case for the Mexican crisis. Moody's downgraded Mexico from investment-grade Baa1 in early 1995 to speculative-grade Ba2 by mid-May (apparently at a time when Mexico was beginning to recover).

A recent paper by Ferri, Liu, and Stiglitz (1999), hereafter FLS, argues that in addition to failing to predict the Asian crisis, the credit rating agencies unduly amplified the crisis when they excessively downgraded the countries later. This would occur if, as a result, the cost of borrowing increased and if the potential pool of investors declined due to statuary requirements. The question motivating this paper therefore is whether ratings have "tremendous power to influence market expectations on a country" (as pointed out by FLS) or whether they are simply reacting, without contributing, to news.

Despite their alleged non-effectiveness and possible guilt, the Basel II proposals give increased prominence to ratings in bank capital requirements. The most recent proposals will allow banks to use ratings by either credit

rating agencies or their own internal assessment in determining the amount of capital they need to put aside for different types of loans. The rationale seems to be that although credit ratings have performed worse than their aim, they are still a second-best solution. Because they are readily available, their use will probably improve the current Basel standards that do not accurately account for risk. More generally, sovereign credit ratings have important implications for international capital flows and for the linkages between company ratings to sovereign ratings.

The objective of this paper is to investigate the behavior of sovereign credit ratings. I focus on the East Asian crisis and whether the rating agencies aggravated the crisis by excessively downgrading those countries. The key advantage of this study is the extension of the period of analysis to the post-crisis period, 1999–2001, allowing for a comparison of pre- and post-crisis rating behavior.

In order to understand how ratings behave, it is first important to ask what the sovereign ratings assigned by the agencies really capture. The two main rating agencies are US-based, Moody's and Standard & Poor's, that have been publishing ratings since 1909 and 1923, respectively. Only recently have they begun rating sovereigns and, only in this past decade, for many developing countries. Clearly the ratings they offer must be of some value since investors pay to subscribe to their credit reports. Rating agencies are neither specific about what determines their ratings nor their rating procedure. Ratings are meant to provide an estimate of the probability that borrowers will not fulfill the obligations in their debt issues. While this all seems plausible at the level of say US firms, it is problematic at the sovereign level.

If we take the statement that ratings are supposed to capture the

default probability of sovereign debt at face value, the question is whether this is true ex post. Do countries with lower ratings default more? The problem is that countries, unlike firms, rarely default because of the availability of international emergency credit (most of their problems being liquidity and not solvency related) and the high cost of future credit should they default. Therefore it is difficult to empirically assess this proposition. Table 1 illustrates the defaults on foreign currency bond debt as compiled by Standard & Poor (S&P). It shows the minimum rating assigned in the year prior to default and the ensuing rating during default. There are a number of striking features. First, only 15 countries have defaulted on foreign currency bond debt since 1975 (although many more have defaulted on bank debt). Second, the paucity of rating figures (only 7 were rated by Moody's and 5 by S&P at the time of default) shows that defaulting countries were not even likely to be rated, which would seem to defy the point of being rated. Third, while all the ratings were below investment-grade prior and during the default period, only Argentina, Ecuador, Pakistan, and Russia were downgraded by S&P to "default status" during the supposed default period. In fact, Pakistan and Venezuela were not even downgraded by Moody's (although this might reflect dating problems since Moody's downgraded Pakistan from a B2 to a B3 on 28 May 1998 and to a Caa1 on 23 October 1998, while the default year was recorded as 1999. Likewise, Moody's downgraded Venezuela from a Ba1 to a Ba2 on 8 April 1994). Nonetheless, it is troubling that even in the periods of default many of the ratings were not systematically rated as such. The other important feature of Table 1 is what it does not contain: all those countries rated at similar and even lower ratings that did not subsequently default. It all goes back to the question of what sovereign ratings try to capture.

This paper explores this question by revisiting the FLS results. I find that ratings are sticky rather than excessively procyclical. Assigned ratings exceeded predicted ratings in the run up to the crisis, mostly matched predicted ratings during the crisis period, and did not increase by the amount suggested by predictions in the recent period after the crisis.[1] This study includes data from the post-crisis period, which helps reveal the inertia in ratings. Therefore they capture the crisis but are overconservative in the period following the crisis. It takes a sufficient amount of either bad (or good) news to change in the direction of the news. Ratings are also found to react to nonmacroeconomic factors such as lagged spreads and a country's default history.

Although ratings appear to lag financial markets, this does not imply that they have considerable market impact. While the first part of the statement is true when ratings are sticky (and I find evidence of this), the next part does not follow. Ratings are probably not contributing much to new information in the market. Unfortunately, without a decent instrument for ratings, the discussion will not be closed. This is not to say that statuary requirements that extend the use of sovereign ratings will be harmless. Biases can occur if ratings are used in ways that distort capital allocations.

The previous literature on sovereign credit ratings can be divided into two main categories, one that focuses on their determinants and behavior (to which this paper belongs) and a second category that focuses on causality. An early example of research focusing on the factors explaining credit ratings is a study by Cantor and Packer (1996) that examines the determinants of credit ratings and looks at their impact on bond spreads. They find that high per capita income is a good explanatory variable, as well as inflation, external

1. This is largely the case when not controlling for country fixed effects. When controlling for country fixed effects, there is an even distribution of ratings relative to predictions across time periods.

debt, economic development, and history of default. FLS (1999) follow the Cantor and Packer method and find that the Asian countries were excessively downgraded relative to the predicted rating based on macroeconomic variables. This paper will be discussed in more detail in section 2. The remainder of this paper is organized as follows: section 2 describes the data and empirical techniques and presents empirical results, and section 3 concludes.

Empirical Framework

Data

Data on Moody's and Standard and Poor's sovereign credit ratings for long-term foreign-currency denominated debt were collected for all countries rated. Note that ratings for many developing countries were only assigned during the 1990s. There were 35 countries in the sample rated in 1989 compared to 105 by 1999. Rating changes were delayed during the East Asian crisis. As mentioned in the introduction, with the exception of Thailand (downgraded on 8 April 1997 from an A2 to an A3), none of the Asian crisis countries were downgraded prior to the crisis. The Mexican crisis occurred at the end of 1994 but Mexico was not downgraded until 6 January 1995. However this is not so much the case for the more recent crises. Russia was first downgraded on 11 march 1998 and later on in May prior to its crisis in August 1998. Argentina was repeatedly downgraded beginning in late 2000.

Two alternative measures of spreads were used. The first is the stripped yield spread obtained from JP Morgan's EMBI Global bond indices covering 33 countries. The alternative is the spread on sovereign eurobonds over comparable US denominated Treasury bonds. Figure 1 shows Moody's ratings and spreads for Korea and Mexico, two prominent crisis countries. An

interesting feature is that ratings lag spreads in many cases, exhibiting sticky behavior. For example, Moody's first downgraded Korea on 27 November 1997, but the spread had increased from 67 basis points on 30 June 1997 to 212 basis points on 26 November 1997. Similarly, Korea's spread declined from a high of 940 basis points on 31 August 1998 to 287 points on 11 February 1999, prior to the upgrade on 12 February 1999. Mexico was downgraded on 16 May 1995 at a time when its spread was falling, and was later upgraded only in 1999.

In addition to data on ratings and spreads, macroeconomic data was obtained mostly from the World Bank's *World Development Indicators 2003* database. The choice of variables derives from the FLS paper (in turn based on the Cantor and Packer variables). Among the variables used are PPP GDP per capita, real GDP growth, CPI inflation, overall budget deficit as % of GDP, current account balance as % of GDP, and the external debt to exports of goods and services. Missing data was supplemented from data in the International Monetary Fund's IFS database, the IMF Public Information Notice reports, as well as data from the EIU country profiles and reports. Refer to Appendix Table 1 for further definition details.

Finally indicators on development and default history were included. FLS use a dummy for whether a country is a member of the OECD as a measure of development. Alternatively, a dummy for whether a country is an industrial country (as classified by the IMF) is constructed. Default history data was obtained from Standard & Poor's publications and discussed in Table 1 in the introduction. The dummy is defined as 1 beginning in the year of default and thereafter. In addition to the data on default on foreign currency bond debt reported in Table 1, S&P also reports data on default on foreign currency bank debt.

Empirical Estimation

Was the Asian downgrading too late and too much? To answer this question we need to know the benchmark against which to judge the ratings. FLS propose an empirical strategy for determining the extent to which credit ratings were procyclical in the Asian crisis. Rating agencies do not offer a transparent list of what goes into determining ratings and their changes; therefore, FLS model ratings as a function of macroeconomic variables and then compare the predictions to the actual ratings assigned. They attribute the difference to the qualitative judgment component of rating agencies. The following sections will elaborate on the FLS results as well as on additional robustness tests.

The Ferri, Liu, and Stiglitz Specification and Predictions

FLS model Moody's rating as a function of explanatory variables derived from the Cantor and Packer 1996 study. They convert the alphanumeric Moody's ratings to a numerical scale using two different cardinalizations. The first is a linear match of the twenty grading notches to a scale from 100 to 5, with 100 being an Aaa and 5 being Ca. This is a similar approach to previous papers such as Cantor and Packer. The second method they use is based on a nonlinear conversion according to the relative increase in spreads with each notch change. I will focus on the linear conversion because the basic results and point they make is unchanged. Moreover, it is problematic to use a nonlinear conversion based on spreads because of potential direct endogeneity in construction of the dependent variable.

I include the FLS results in column (1) of Table 2 for comparison purposes. They use a sample of 17 countries for a period of ten years: 1989 to

1998.[2] They run an unbalanced panel with random effects of the form (where the variables are as defined in the previous data section):

(1)

$$Rating_{i,t} = \alpha + GDP_{i,t} + Growth_{i,t} + Inflation_{i,t} + Deficit_{i,t} + CA_{i,t} + Debt_{i,t} + OECD_{i,t} + \varepsilon_{i,t}.$$

FLS find that all the explanatory variables have the expected signs and most are significant with the exception of GDP per capita and the inflation rate. I first attempt to replicate their results in column (2) of Table 2 for their set of countries and the period from 1989 to 1998. The estimates are comparable for the most part. I obtain a more significant and positive GDP per capita, a less significant real growth and budget deficit, a stronger effect of external debt, and a comparable negative effect from inflation. The two discrepancies are the current account and the economic development indicator.[3]

FLS then proceed to show figures for four East Asian countries (Indonesia, Korea, Malaysia, and Thailand) like those in Figure 3, plotting model predictions along with actual ratings. They attribute the difference to the rating agency's judgment; therefore their approach to determining if ratings were procyclical is if actual ratings prior to the crisis were above the predicted ratings and were then downgraded excessively below the predicted at the time of the crisis. They find that prior to the Asian crisis, actual ratings were consistently higher than what the model predicts. Similarly, during the crisis in 1997, the actual ratings dropped much more than the model-predicted ratings (except for the case of Malaysia where actual ratings continued to be

2. They are Argentina, Australia, Brazil, China, Greece, Hungary, Iceland, India, Indonesia, Korea, Malaysia, Mexico, New Zealand, Philippines, Portugal, Thailand, and Venezuela.

3. For a discussion on the robustness of the FLS regression, refer to the longer version of this paper.

higher than predicted). They also find that in 1998 the error was smaller because model-predicted ratings converged to the actual ratings. They hypothesize that this reflects the endogeneity of macroeconomic fundamentals to ratings through the ratings effect on investors' capital outflow and the subsequent damage to fundamentals; therefore, they suggest that ratings were excessively procyclical for the East Asian economies compared to the model predicted ones.

A more useful technique to evaluate this claim in detail is to tabulate how many of the observations were correctly classified or not as done by Hu, et al (2002). Over all the years the correlation between the predicted ratings and the actual ratings is 0.96. However, this misses the important cross-year fit in the ratings which underlies the FLS argument; therefore, Figure 2 reports for each year the share of observations that were correctly predicted, the share with predicted ratings lower than actual, and the share with predicted ratings higher than actual. The FLS hypothesis would imply that the share of predicted ratings higher than actual exceeds the share of predicted ratings lower than actual at the time of the crisis, and the opposite prior to the crisis. This is indeed observed in the first graph based on the FLS model reported in Table 2(2). During the crisis period from 1997 to 1998, 72.7% of the observations were predicted to be higher than actually rated. In contrast, the figure is only 40.5% for the period from 1989 to 1996.

In short, the FLS model and predictions provide support for excessively procyclical sovereign credit ratings. The following sections discuss some technical limitations to the FLS model and report results for other specifications. More important, the availability of an updated data set allows for a closer look into the rating-through-the-cycle behavior, particularly for the post-crisis period.

Robustness of the Linear Specification

From an econometrics viewpoint, it is questionable whether random effects are preferred. Given that the panel is composed of countries, it is hard to justify that the country individual effects are uncorrelated with the explanatory variables. The downside of using fixed effects is that it reduces the degrees of freedom and is not efficient, but using random effects when the true model is fixed effects leads to inconsistent estimates. A Hausman specification test offers a simple way to test the two models. This test is reported in column (2) of Table 2 where $\chi^2(7)=26.4$, rejecting random effects.

A second point concerns the use of the minimum assigned Moody's rating in the original study. This contributes to the difference between model generated ratings and actual ratings. Large errors are more likely when the dependent variable is the minimum rating for the year while the explanatory variables reflect values over the entire year. This is particularly the case for East Asia when Korea and Indonesia experienced their crises and were downgraded late in 1997. For example, the average (time-weighted) Moody's rating for Korea in 1998 falls to 78 from 80, compared to a minimum of 50 in December 1997.

The subsequent columns of Table 2 therefore use the average rating as the dependent variable (the average of Moody and S&P and averaged over the year) and extend the sample to the full sample of available countries, 88 countries from 1989 through 2001. Columns (3) and (4) report random effects and fixed effects models similar to column (2), omitting the OECD dummy. A Hausman specification test continues to favor the fixed effect model. An alternative specification (not reported because of similarity) accounts for economic development with a dummy for industrial countries and default histories on foreign currency bonds and bank debt, respectively. This is a

similar approach to that taken in the original Cantor and Packer paper. The results are not substantially different, while the dummies enter with the expected significant sign. This suggests an inertia in a country's rating which is influenced by its history.

What may also be contributing to the large prediction errors may be an omitted variable, the market. To investigate this a sample of developing countries with available Eurobond spreads (29 countries)[4] or available EMBI spreads (24 countries)[5] is used to proxy for market sentiment. It is quite plausible to view rating agencies as just reacting to market movements rather than inducing them. Since the results are very similar, and for the purpose of brevity, I report the results for the Eurobond spreads in column (5). They are included as one year lagged spreads to account (as best as possible) for the possibility that spreads are endogenous to ratings. Because of the small sample size, 150, the fixed effect coefficients (not reported) are reduced in significance. However, what stands out is the significant and negative effect of lagged spreads on ratings, suggesting that the hypothesis that ratings react passively to the market is indeed a plausible one. Further evaluation of the model predicted ratings is discussed in the section below.

In short, this section has raised some specification problems. Among which are the use of the minimum Moody's rating in a year instead of the average, the use of random effects, and omitting to account for the potential influence of nonmacroeconomic variables. The latter are market sentiment

4. They are Argentina, Brazil, China, Colombia, Costa Rica, Croatia, Ecuador, Hungary, Indonesia, Israel, Kazakhstan, Korea, Lebanon, Malaysia, Mexico, Pakistan, Panama, Peru, Philippines, Poland, Romania, Russia, South Africa, Slovenia, Slovakia, Thailand, Tunisia, Turkey, and Uruguay.

5. They are Argentina, Brazil, Bulgaria, Chile, China, Colombia, Croatia, Ecuador, Hungary, Korea, Lebanon, Malaysia, Mexico, Morocco, Panama, Peru, Philippines, Poland, Russia, South Africa, Thailand, Turkey, Ukraine, and Venezuela.

and a country's default history. A further problem may be the choice of a linear specification. The ordered probit methodology is preferable to the linear model on theoretical grounds. The ordered probit model accounts for the ordinal nature of the dependent variable without the restrictive nature of the linear model, which treats the difference between an Aaa and an Aa1 the same as that between a Baa3 and a Ba1 (or whatever numerical scale is chosen). Ordered probit estimates are reported in Table 2, columns (6) through (8), and the results confirm that macroeconomic variables matter for ratings.

Evaluating Model Predictions

The previous discussion has shown that macroeconomic variables matter for ratings. However it remains to be determined how well the ratings are predicted and how this depends on the cycle. Is it the case that they are excessively procyclical as pointed out by FLS and demonstrated earlier in the text. This section attempts to evaluate the predicted ratings on various benchmarks to better evaluate the validity of this claim.

Comparing Predicted Ratings to Assigned Ratings: Higher or Lower

The first panel of Figure 2 tabulated how many of the observations were correctly or incorrectly classified for the FLS model. Results supporting the FLS claim were discussed earlier. The remaining panels show the alternative linear and ordered probit specifications respectively. The FLS result is not particularly robust to the different specifications. In the simple linear model without country fixed effects, default history and spreads (Model 2(3)), an even 53.4% of predicted ratings were higher than actual during the crisis

period. What is interesting is that in the period prior to the crisis, only 38% of observations were predicted higher, while in the period after the crisis (1999 to 2001) the figure is 60.7%. This lends support to an inertia view of ratings if they are predicted solely based on a country's macroeconomic factors. Ratings adjust during crisis times to their predicted levels and not excessively so. However, as predicted ratings improve during the post-crisis period (a period that was not yet analyzed at the time of the FLS paper), ratings do not adjust as quickly upward. This is comparable to their behavior in the run up to the crisis when it takes enough bad news to cause a downgrade in ratings.

The comparable ordered probit model (Model 2(6)) also exhibits sticky behavior. Note that the difference between the linear and ordered probit models is that the average rating used as a dependent variable is rounded to the nearest rating notch when running the ordered probit (and so are the predicted ratings from that model). The simple ordered probit model is able to estimate 18 cutoffs, that is ratings from Caa3 to Aaa, and is not able to estimate the cutoff for ratings below Caa3 due to no incidence. Therefore this results in a considerable number of observations (20% over the sample) rated the same as predicted. Nonetheless the pattern remains. In the period prior to the crisis, 62% of observations were predicted lower than actual and 23.6% were predicted higher, compared to a more even 35.1% predicted lower and 42.6% predicted higher during the crisis period. The post-crisis period is characterized by 24.2% predicted lower compared to 50% predicted higher.

An interesting result occurs once country fixed effects are included. First, including country dummies considerably improves the fit of the model. Sixty percent of observations are now correctly predicted by the ordered probit model (Model 2(7)) compared to 20% for the model without fixed effects.

Second, the cross-year variation in predicted ratings from actual ratings is largely reduced.[6] For example, the linear model (Model 2(4)) results in 48.2% of observations predicted higher in the pre-crisis period, 50% in the crisis period, and 51.6% in the post-crisis period. A similar pattern is observed for the ordered probit model. One explanation is that the introduction of fixed country effects would leave the macroeconomic variables with little variation to explain across the countries. Because that is the major source of variation, then it would not be surprising that the distribution does not change much from year to year.

The sample of observations is reduced (and limited to mainly developing countries) when including lagged spreads to the explanatory variables. For brevity, results are shown for the ordered probit model with Eurobond spreads and default history indicators. The pre-crisis sample now begins only in 1994, but the pattern remains. Sixty-two percent of observations are predicted lower than actual during the pre-crisis period, which evens out during the crisis period to 40% of observations predicted lower than actual compared to 33.3% predicted higher. During the post-crisis period, 38% are predicted lower than actual, albeit a less marked share than in the model with only macroeconomic variables.

In short, there is little support for the claim that actual ratings were excessively conservative during the crisis period from 1997 to 1998 based on the preliminary evidence in this section. When not including country fixed effects, ratings are found to be sticky rather than excessively procyclical. While there is support for the FLS finding that predicted ratings were lower than assigned ratings during the period prior to the crisis, there is no support for their being

6. This is not due to the entry of many developing countries towards the later part of the sample because of new sovereign ratings assignments. Even if the sample is restricted to the set of countries with ratings throughout the period, the difference remains.

predicted higher than assigned during the crisis period. The additional advantage of this study is that it extends the sample period to the post-crisis period from 1999 to 2001. This helps reveal the inertia pattern. When not including country fixed effects, predicted ratings are found to be higher than assigned ratings during this period. Under the hypothesis that ratings are excessively procyclical, assigned ratings should be higher than predicted during this upturn. Finally, when including country fixed effects, there is no evidence of either procyclicality or stickiness.[7]

Comparing Predicted Ratings to Assigned Ratings: Country Specific Evidence

This section focuses on the country-specific evidence as was done by FLS. Figure 3 shows the assigned ratings and the predicted ratings for Argentina, Brazil, Indonesia, Korea, Malaysia, Mexico, Russia, Thailand, and Turkey. These have comprised countries with crises over the past decade. The top panel shows the assigned ratings and the predicted ratings from the linear model based on only macroeconomic variables.[8] In the linear models without spreads, the errors are not as pronounced as those in FLS. However, as in FLS, predicted ratings do not fall as much as assigned ratings for Korea and Indonesia during 1998 (though they do and more so in 1997). In contrast, assigned ratings track the predicted ratings for Malaysia and Thailand quite well during the crisis period. More generally, it is evident that for many of the countries (not just the East Asian countries) predicted ratings recover faster than assigned ratings after a crisis.

The lower panel shows the predicted ratings from the ordered probit model with spreads and default history. Overall, the assigned ratings tracked

7. For further details and evaluation of model predictions, refer to the longer version of this paper.

8. Both the random and fixed effects linear models produce similar predictions for this set of countries.

predicted ratings for the Asian countries during the crisis but remained lower after the crisis and were higher before the crisis. It is interesting to note that in the case of Turkey, while the predicted rating based on macroeconomic variables has been higher than the actual rating from 1995 to 2001, the predicted rating that includes spreads has matched the "pessimistic" rating more closely. A similar result holds for Russia.

To summarize, the evidence does not much support the FLS view of excessive procyclicality. The results appear to be more consistent with inertia in rating decisions. There may be fixed costs to the agencies when changing an assigned rating and it takes a sufficient amount of either bad or good news to change in the direction of the news. Ratings also appear to be not only affected by macroeconomic news but also by market sentiment and default histories. When coming from a period of apparent overexcitement, it then looks like ratings are overreacting during the crisis. When, to the most part, they are realigning to their predicted values.

The Inertia of Sovereign Ratings

This section provides a more formal approach to evaluate the extent of inertia. I regress the change in a rating on the lagged error term, where the error term is defined as the predicted rating—the assigned rating:

(2)

$$\Delta Rating_{i,t} = \alpha + \beta\, error_{i,t-1} + \varepsilon_{i,t}\ .$$

If ratings are nonsticky, then ß should be equal to zero. In contrast, if ratings are sticky, ß should be positive. That is, when the divergence between the predicted and the assigned rating becomes sufficiently large,

there will be a rating change in the direction of the predicted rating.

Results are reported in the longer version of this paper. Overall, there is strong support for sticky ratings in all specifications. The coefficient on the lagged error term is positive and strongly significant and is robust to controlling for country and year effects. Moreover, results are robust to separately including lags of variables such as the current account balance, budget deficit, and bond spreads. Similar results are found for any of the specifications (linear and ordered probit) estimated, including ones that initially included country fixed effects.

The results suggest that rating agencies convey, with a lag, bad news in bad times and good news in good times. Therefore, ratings may not be very informative to investors participating in the market for sovereign debt. This view is best put by one rating agency veteran who is quoted as saying "we may be incompetent but we're not dishonest" (see *The Economist*, 2/6/2003).

Concluding Remarks

In short, I argue that the case for the guilt of sovereign credit rating agencies is not tenable. This paper was written to investigate the view of Ferri, Liu, and Stiglitz (1999) and others that credit rating agencies have a strong impact on market expectations, thereby damaging the macroeconomic fundamentals of a country and aggravating crises. This paper suggests a more cautious view.

There is little support for assigned ratings being excessively conservative during the crisis period from 1997 to 1998. When not including country fixed effects, ratings are found to be sticky rather than excessively procyclical. While there is support for the FLS finding that predicted ratings

were lower than assigned ratings during the period prior to the crisis, this is consistent with both the FLS and inertia views. That ratings are not found to be predicted higher than assigned during the crisis period weakens the FLS view. The additional advantage of this study is that it extends the sample period to the post-crisis period from 1999 to 2001. This helps reveal inertia in ratings. Predicted ratings are found to be higher than assigned ratings during this period (when not including country fixed effects). Therefore they capture the crisis but remain overconservative after the crisis. Ratings adjust only when there is a sufficiently large divergence of predicted ratings from assigned ratings. I also find that in a model where ratings react passively to market sentiment (as proxied by spreads), then the drying up of credit might be attributed to excessive downgrading, even though it would have occurred regardless. Ratings are also found to be influenced by a country's default history.

Although ratings may by the most part be reacting to macroeconomic and market news, this does not imply that their use in statuary requirements is benign. This particularly deserves attention with the introduction of ratings into the latest Basel II proposal for determining risk weights for bank loans. This point is well made by Kashyap and Stein (2004) who argue that a single time-invariant risk curve that maps credit-risk measures (such as ratings) into capital charges is sub optimal. They make a case for having a family of risk curves such that the capital charge is reduced when economy-wide bank capital is scarce, such as during recessions. Others, such as *The Economist* (2/6/2003), call for taking credit ratings completely out of financial regulation and using instead the spreads at which secondary market debt securities trade. More careful study is needed prior to 2007 when Basel II is expected to take effect so that the new capital standards do not exacerbate

business cycle fluctuations. There are good reasons for capital requirements to be countercyclical, *but* this does not imply that ratings can be expected to be countercyclical. Instead, their use for regulatory reasons should be carefully evaluated.

Table 1 Sovereigns: Years in Default 1975–2002 on Foreign Currency Bond Debt

	Years	Rating in previous year		Rating during period of default	
		Moody's	S&P	Moody's	S&P
Argentina	1989; 2001–02	Ba3; B1	na; BB–	B3; Ca	na; SD
Bolivia	1989–97				
Costa Rica	1984–85				
Ecuador	1999–2000	B3		Caa2	SD
Former Yugoslavia	1992–2002				
Guatemala	1989				
Ivory Coast	2000–02				
Molodova	1998; 2002	Ba2; Caa1		B2; Caa1	
Nigeria	1986–88; 92				
Pakistan	1999	Caa1	CC	Caa1	SD
Panama	1987–94				
Russia	1998–2000	Ba2	BB–	B3	SD
Ukraine	1998–2000	B2		Caa1	
Venzuela	1995–97	Ba2	B+	Ba2	B
Zimbabwe	1975–80				

Source: Standard & Poor's *Credit Week,* December 22, 1999, "Sovereign Defaults: Hiatus in 2000?" Updated with S&P publication September 24, 2002, "Sovereign Defaults: Moving Higher Again in 2003?" and S & P and Moody's Ratings Services for ratings.

Note:
Ratings refer to the minimum ratings given during the period. For Ukraine, the previous year is the new rating issued on February 6, 1998.

Table 2 Annual Rating Regressions: Extending Ferri-Liu-Stiglitz

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Linear specifications					Ordered probit specifications[2]		
Dependent variable	FLS model[1] Min rating	FLS rerun Min rating	Ran effects AV rating	Fix effects AV rating	Ran effects AV rating	w/ctry effects Av rating	w/ctry effects AV rating	w/ctry effects AV rating
Regressors								
GDP (PPP) per capita	0.0002	0.0010***	0.0012***	0.0003***	0.0010***	0.0002***	0.0002**	0.0001***
	(0.0004)	(0.0003)	(0.0001)	(0.0001)	(0.0004)	(0.0000)	(0.0001)	(0.0000)
GDP growth rate	31.2***	0.3049**	0.0121	-0.0013	-0.2532**	0.0452**	-0.0035	0.0151
	(7.6)	(0.1555)	(0.0730)	(0.0640)	(0.1167)	(0.0199)	(0.0203)	(0.0237)
Inflation rate	-0.001	-0.0114	-0.0677	-0.1121	-6.3970	-0.0593***	-0.0154	-1.7097**
	(0.0009)	(0.1378)	(0.1085)	(0.0947)	(4.2086)	(0.0125)	(0.0144)	(0.7435)
Overall budget balance, % GDP	0.4**	0.4019*	0.1054	0.2207***	0.7290***	-0.0032	0.0652**	0.0728***
	(0.2)	(0.2429)	(0.0859)	(0.0769)	(0.1958)	(0.0174)	(0.0290)	(0.0256)
Current account balance, %GDP	-0.5***	-0.6001***	-0.5928***	-0.6203***	-0.7271***	-0.0119	-0.1384***	-0.0430*
	(0.2)	(0.1733)	(0.0609)	(0.0548)	(0.1103)	(0.0137)	(0.0304)	(0.0252)
External debt to exports of g & s	-0.0001*	-3.5458***	-1.3521***	-0.7641*	-2.8697***	-0.1999***	-0.0999	-0.5239***
	(0.00006)	(1.3295)	(0.3993)	(0.4147)	(1.0430)	(0.0553)	(0.1713)	(0.1540)
OECD dummy	5.2***	-1.9961						
	(1.5)	(2.5778)						
Default history dummy (on bonds)					-3.5709			-0.4781
					(3.0167)			(0.4273)
Default history dummy (on banks)					-14.5678***			-0.8263**
					(2.9209)			(0.4082)
Spread (on eurobonds), first lag					-0.2335***			-0.0660**
					(0.0815)			(0.0280)
Constant	51.2***	57.8659***	48.9355***	62.9317***	58.9213***			
	(3.6)	(4.7942)	(1.8597)	(1.5340)	(4.4416)			
Observations	161	144	705	705	150	705	705	150
Number of countries	17	17	88	88	29	88	88	29
R^2 (pseudo if ordered probit)	0.30	0.57	0.68	0.21	0.64	0.24	0.58	0.25
Hausman specification test		chi2(7)=26.4	chi2(6)=621.7		chi2(9)=23.8			

Source: Author's estimates.

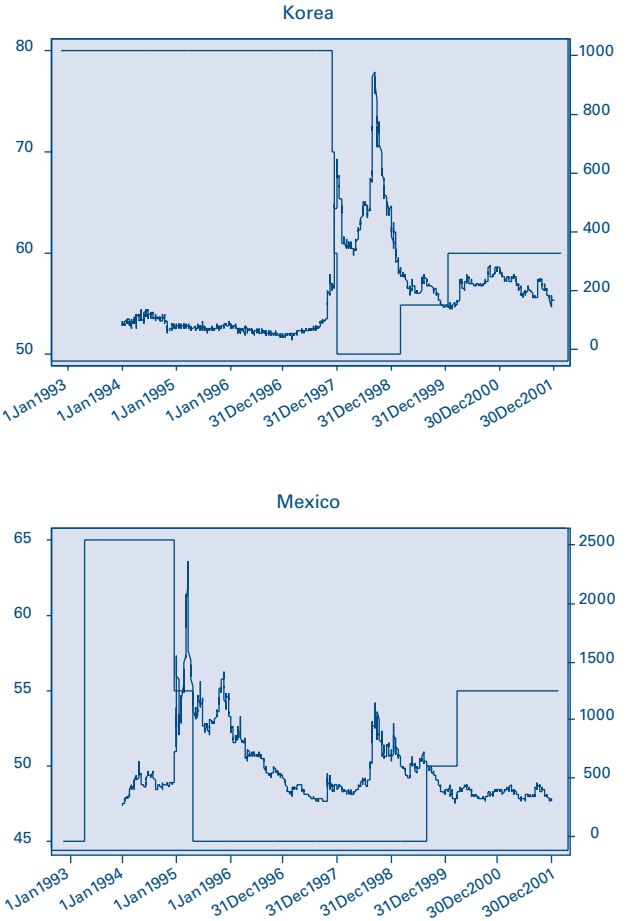
Notes:
1. The FLS model refers to Table 3 in FLS "Linear Numerical Conversion Regression Results" (FLS report t statistics instead of s.e. in their paper).
2. Ordered probit models round the dependent, average rating, to the nearest rating notch (maximum possible rating categories are 21). 18 rating cutoffs (not reported) are estimated for the ordered probit models without spreads (from 100 to 10, that is Aaa to Caa3 in Moody's scale). 15 rating cutoffs are estimated for the model with Eurobond spreads (from 85 to 10) and 13 for the model with EMBI (from 85 to 15, skipping 75). The ordered probit standard errors are clustered by country. Pseudo R-squared for the ordered probit are defined as 1- (ln L)/(ln L_0), where ln L is the log like lihood and ln L_0 is the log likelihood if estimated on only a constant.
3. Standard errors in parentheses.
4. An * indicates a significance of 10%, ** a significance of 5%, and *** a significance of 1%.

Appendix Table 1 Data Definitions and Sources

Variable	Description	Source
Moody's ratings	Long-term foreign currency sovereign rating for bonds and notes	Moody's website: www.moodys.com
S&P ratings	Long-term foreign currency sovereign rating	S&P website: www2.standardandpoors.com
Average rating	Average of (weighted annual) Moody's and S&P	
Eurobond spreads	Mainly US$ denominated Eurobonds yield minus equivalent risk-free US treasury bills' yield	Datastream and Bloomberg
EMBI Global spreads	Stripped yield spread on index comprising US$ denominated sovereign Brady bonds, Eurobonds, etc.	Datastream (from J.P. Morgan Chase originally)
GDP per capita	GDP per capita, PPP (current international $)	WB World Development Indicators 2003 CD (WDI)
Real GDP growth	GDP growth (annual %)	WDI (and EIU and IMF for some missing obs)
Inflation rate	Based on consumer price index (1995=100), in %	WDI (and EIU and IFS for some missing obs)
Current account	Current account balance (% of GDP)	WDI (and IMF and EIU for some missing obs)
External debt	Ratio of external debt, total (DOD, current US$) to exports of goods and services (BOP), current US$)	WDI; IFS International Investment Position sum of series 79ledzf...and 79lfdzf...for developed countries
Short-term debt	Short-term debt outstanding (DOD US$)	WDI
Budget deficit	Overall budget balance, including grants (% of GDP)	WDI, IMF International Financial Statistics (IFS), IMF PINs, Economist Intellegence Unit Profiles, and country central banks
Development indicator	1 if a country belongs to OECD in given year and 0 otherwise (30 ctries as of 2 OECD) or 1 if a country belongs to a set of industrial countries and 0 otherwise (24 countries IFS)	
Default histories	Equal to 1 once a country defaults on its foreign currency bond debt (since 197 S&P)	

Note:
Data and programs available upon request.

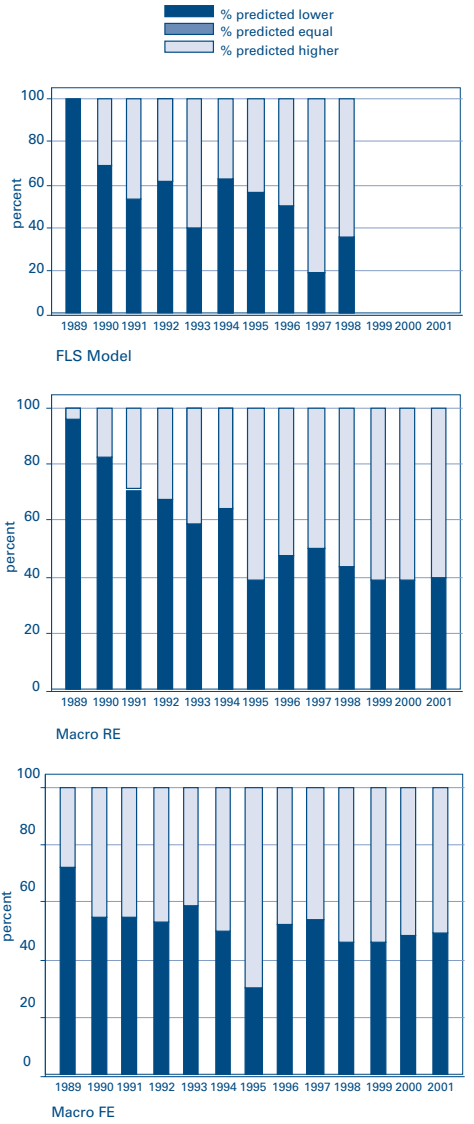
Figure 1 EMBI Spreads (right axis) versus Moody's Rating (left axis)



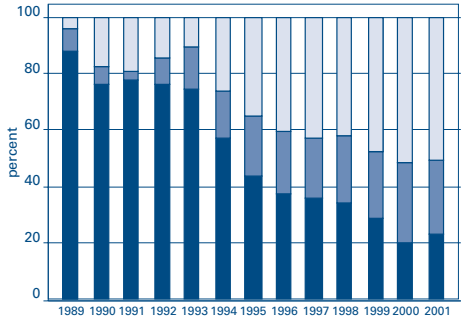
Sources: Moody's ratings from Moody's website (http://www.moodys.com)
EMBI Spread from Datastream Thomson Financial database based on JP Morgan Chase Emerging Market Bond Index stripped yield spreads.

Figure 2 Predicted Ratings Compared to Actual Ratings (Higher/Equal/Lower)

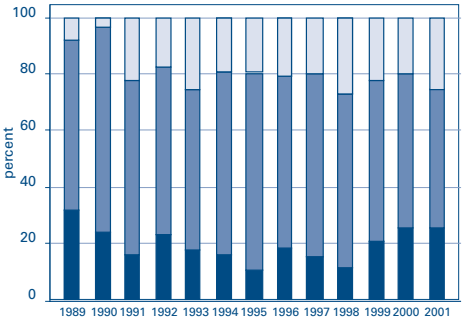
Linear



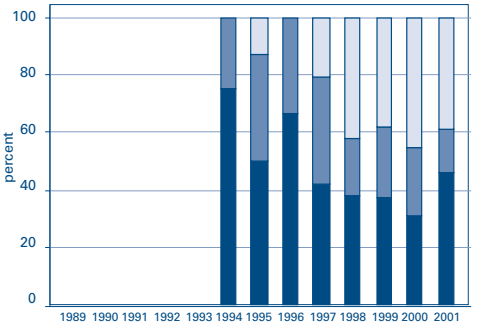
Ordered Probit



Macro



Macro with country effects

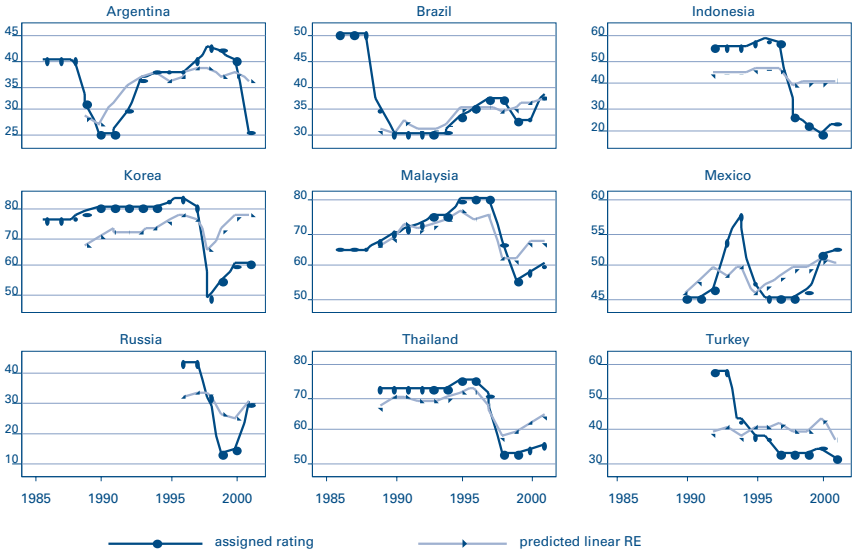


Including default history and Eurobond spreads

Source: Author's estimates.

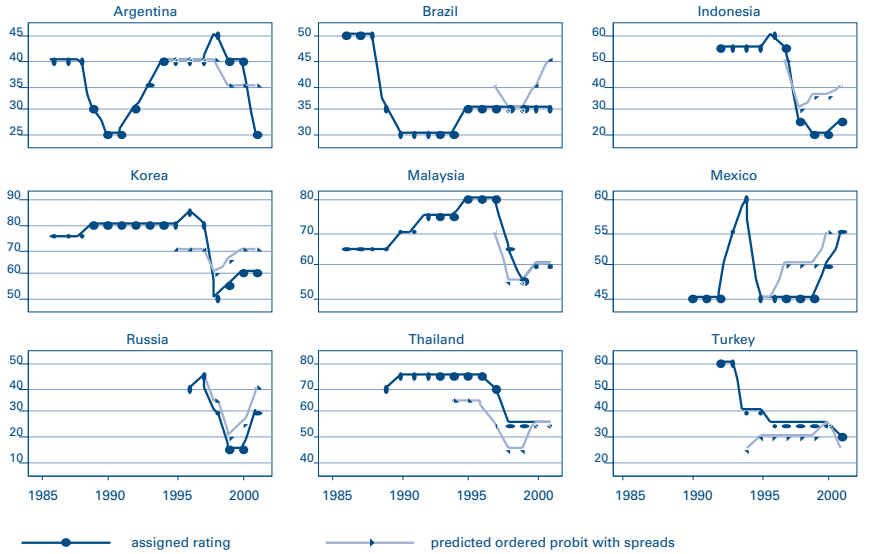
Figure 3 Assigned and Predicted Ratings by Country

Linear Macro Model



Graphs by COUNTRY

Ordered Probit with Default History and Spreads



Graphs by COUNTRY

Source: Author's estimates.

References

Cantor, R., and F. Packer. 1996. Determinants and impact of sovereign credit ratings. *Federal Reserve Bank of New York Economic Policy Review* 2, no. 2: 37–53.

Conover, W. J. 1999. *Practical Nonparametric Statistics*. New York: John Wiley & Sons.

Eichengreen, B., and A. Mody. 1998. What explains changing spreads on emerging-market debt: fundamentals or market sentiment? National Bureau of Economic Research Working Paper 6408.

Ferri, G., L. Liu, and G. Majnoni. 2001. The role of rating agency assessments in less developed countries: impact of the proposed Basel guidelines. *Journal of Banking and Finance* 25, no. 1: 115–48.

Ferri, G., L. Liu, and J. E. Stiglitz. 1999. The procyclical role of rating agencies: evidence from the east Asian crisis. *Economic Notes* 28, no. 3: 335–55.

Greene, W. H. 1997. *Econometric Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Hand, J. R., R.W. Holthausen, and. R. W. Leftwich. 1992. The effect of bond rating changes on bond and stock prices. *Journal of Finance* 47: 733–52.

Haque, N., N. Mark, and D. Mathieson. 1996. The economic content of indicators of developing country creditworthiness. IMF Staff Papers 43: 688–724.

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

International Monetary Fund. 1998. International capital markets: developments, prospects, and key policy issues. Washington, DC.

Kaminsky, G., and S. L. Schmukler. 2002. Emerging market instability: Do sovereign ratings affect country risk and stock returns? *World Bank Economic Review* 16: 171–95.

_____.1999. What triggers market jitters? A chronicle of the Asian crisis. *Journal of International Money and Finance* 18: 537–60.

Kashyap, A., and J. Stein. 2004. Cyclical implications of the Basel II capital standards. Federal Reserve Bank of Chicago Economic Perspectives Q1, 18–31.

Larrain, G., H. Reisen, and J. von Maltzan. 1997. Emerging market risk and sovereign credit ratings. Organization for Economic Cooperation and Development, Development Center Technical Paper 124.

Liu, P., F. J. Seyyed, and S. D. Smith. 1999. The independent impact of credit rating changes—The case of Moody's rating refinement on yield premiums. *Journal of Business Finance and Accounting* 26, no. 3: 337–62.

Monfort, B., and C. Mulder. 2000. Using credit ratings for capital requirements on lending to emerging market economies: Possible impact of a new Basel accord. International Monetary Fund Policy Working Paper, WP/00/69.

Moody's Investment Service. 1994. Note to the SEC on the disclosure of security ratings.

Nayar, N., and M. S. Rozeff. 1994. Ratings, commercial paper, and equity returns. *Journal of Finance* 49: 1431–49.

Reinhart, C. M. 2002. Default, currency crises, and sovereign credit ratings. *World Bank Economic Review* 16, no. 2: 151–70.

_____. 2001. Do sovereign credit ratings anticipate financial crises? Evidence from emerging markets. Mimeo, School of Public Affairs, University of Maryland, College Park.

Reisen, H. 2002. Ratings since the Asian crisis. Organization for Economic Cooperation and Development, web document 2.

Reisen, H., and J. von Maltzan. 1998. Sovereign credit ratings, emerging market risk and financial market volatility. *Intereconomics* 33, no. 2: 73–82.

Rigobon, R. 1997. Informational speculative attacks: good news is no news. Mimeo, Sloan School of Management, MIT.

Standard & Poor's. 1999. Sovereign defaults: Hiatus in 2000? Standard & Poor's *Credit Week*.

Stiglitz, J. E., and A. Bhattacharya. 1999. Underpinnings for a stable and equitable global financial system: from old debates to a new paradigm. Annual Bank Conference on Development Economics, discussion paper.

Sy, A. N. R. 2002. Emerging market bond spreads and sovereign credit ratings: Reconciling market views with economic fundamentals. *Emerging Markets Review* 3, no. 4, 380—408.