INTERNATIONAL JOURNAL OF FINANCE AND ECONOMICS
Int. J. Fin. Econ. (2012)
Published online in Wiley Online Library
(wileyonlinelibrary.com). DOI: 10.1002/ijfe.1461

DO CREDIT RATING AGENCIES ADD VALUE? EVIDENCE FROM THE SOVEREIGN RATING BUSINESS

EDUARDO CAVALLO[a,*,†], ANDREW POWELL[a] and ROBERTO RIGOBON[b]
[a]*Inter-American Development Bank, 1300 New York Ave. N.W., Washington, DC 20577, USA*
[b]*Sloan School of Management, MIT, Room E52-442 50 Memorial Drive, Cambridge, MA 02142-1347, USA*

ABSTRACT

The debt crisis in several European Union nations has resulted in a set of downgrades in sovereign ratings, sparking a lively debate whether these opinions actually matter. Ratings and bond spreads may both be considered as noisy signals of fundamentals. Ratings only add value if, controlling for spreads and observable country fundamentals, they help explain other market variables. We employed a unique dataset of over 75 000 daily observations on emerging countries around rating actions by the three major agencies. We found that ratings do indeed add information, and this finding is robust to a variety of different tests.

Received 10 June 2010; Accepted 17 February 2012

JEL CODE: F37; G14; G15; C23

KEY WORDS: ratings; spreads; information economics; event studies

1. INTRODUCTION

Rating agencies were one part of the jigsaw explaining the development of the global financial crisis that erupted in the summer of 2007.[1] Moreover, the crisis in Europe put the spotlight on their role in rating sovereigns. Between the months of April and May 2010, the three leading rating agencies (Standard & Poor's (S&P), Moody's, and Fitch) downgraded Greece a cumulative total of six notches, and S&P has downgraded Spain and Portugal by one notch each (Figure 1).

But do investors learn anything from rating agencies actions? One of the charges made against rating agencies in the market for sovereign instruments is that they tend to lag market developments and behave reactively.[2] Indeed, this is oftentimes taken as evidence that sovereign ratings are uninformative.[3] This is the benign interpretation of the pro-cyclicality. In this interpretation, the credit rating agencies are irrelevant—they simply reflect information already available to investors. Others have argued that rating agencies do matter and indeed may even amplify financial crises and distress.[4] We take this issue seriously and ask, do rating agencies provide information beyond what is already priced in the market?

Using a unique dataset of over 75 000 daily observations of financial data from emerging countries—where most of the action in the sovereign rating business used to be focussed prior to the current spate of activity in continental Europe—and robust empirical tests, we scrutinized the actions of the three leading rating agencies in the sovereign debt market and found that they add information. Whereas other studies have sought to answer this question before, our contribution is to explicitly test—using high-frequency data—if rating changes contain new information even *after* controlling for sovereign bonds spreads (i.e., taking into account the information already embedded in market prices).

In our view, sovereign bond ratings and spreads are both noisy signals of the true and perhaps unknowable deep economic fundamentals. Although spreads provide direct market signals—sovereign debt is traded everyday in

*Correspondence to: Eduardo Cavallo, Inter-American Development Bank. 1300 New York Ave. N.W., Washington, DC, 20577, USA.
†E-mail: cavalloe@iadb.org

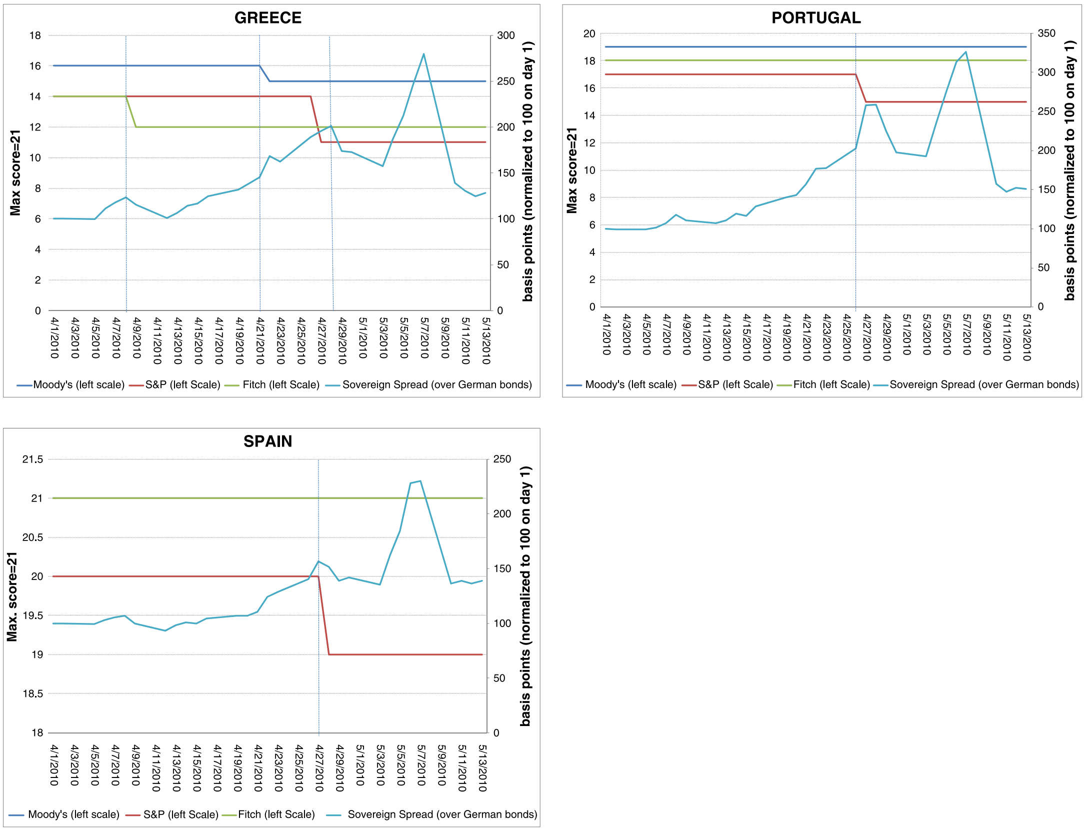


Figure 1. The European downgrades.

secondary markets thereby providing up-to-date information on spreads over riskless debt—ratings are the result of an assessment that rating agencies make on the underlying probability of default of a given country. Credit ratings and spreads appear then to be capturing the same thing, but the important question is if there is value added in ratings, that is, do rating agencies' "opinions" provide some incremental information about unobservable countries' fundamentals that would otherwise not be readily available?

Several recent articles have considered the role of rating agencies in the sovereign debt market. Cantor and Packer (1996) and Afonso *et al.* (2007) showed that ratings level can be modelled fairly successfully by a set of observable economic fundamentals. However, these models tend to be less successful in explaining rating dynamics.[5] In addition, Kaminsky and Schmukler (2002), Brooks *et al.* (2004), and Ferreira and Gama (2007) showed that sovereign ratings affect financial asset prices, but they do not assess the extent to which the information embedded in rating changes reflects information already available to investors. The unanswered question is whether rating changes affect market variables controlling for observable fundamentals and for current bond spreads. The stumbling block that has prevented the finding of an answer is an identification problem: how to isolate the exogenous component of ratings or rating agencies' "opinion."

Eichengreen and Mody (1998) and Dell'Ariccia *et al.* (2006) regressed ratings on observable fundamentals and interpreted the error as the rating agencies' "opinion." They then showed that this residual is highly significant in explaining bond spreads. Powell and Martínez (2007) replicated these analyses; they also employed a system of equation approach and further argued that the rating agencies' differences in opinion are informative. In other words, when one agency changes a rating and the others do not, then this is associated with a change in spreads.

Despite these efforts, it is not yet possible to argue that a definitive answer has been given to the question of the informational content of ratings. Each methodology employed to date has its particular drawbacks.[6] Moreover, there may be more information in markets than is captured in these models, and the above approaches do not

DOI: 10.1002/ijfe

control for the current information in markets, but only observable variables. This article raises the bar with respect to the articles cited by testing if credit ratings influence spreads and other asset prices over and above the information that is already aggregated in market variables.

Another tack would be to attempt an event study as in the corporate finance literature—see Campbell *et al*. (1997) for a discussion.[7] However, rating agencies appear to try to signal when rating changes may occur. Sovereign debt is either given a positive or negative outlook (suggesting an upgrade or a downgrade may be the next change, respectively) and additionally may be placed on a "rating watch" (indicating that a decision may be about to be made). Moreover, agencies publish what a particular sovereign would have to do to improve its rating, and although targets may not be precise, the information required to make a judgment is generally public and indeed may become a focus of market research and analysis. All this implies that the classic event study methodology may not be appropriate as rating changes may be anticipated.

This means that it is a real challenge to answer the question as to whether ratings agencies add value. If rating agency actions are fully anticipated, then we would see no incremental effect of upgrades or downgrades on market variables such as bond spreads, stock market indices, or even nominal exchange rates in emerging countries. However, seeing no effect would not mean that rating agencies do not add value. In our view outlined earlier that both ratings and spreads are noisy signals of fundamentals, it just implies that whatever effect ratings had on market variables may have already been incorporated into bond spreads and asset prices by the time of the announcement. Alternatively, the fact that market variables react before actual rating changes could be interpreted as evidence that rating agencies behave reactively, deciding to downgrade (upgrade) a country when the prices of its financial instruments go down (up).[8] This is consistent with the view that ratings simply follow the market providing no value added.[9] The elusive identification problem suggests that we need to seek other methods to tackle this question.

For this purpose, we first devised a simple and robust specification test to evaluate whether ratings explain a portion of the variation in other macro variables in addition to what can be explained by sovereign bond spreads. The null hypothesis is that the spread is a sufficient statistic for the unobserved fundamental—that is, that the rating does not add information beyond what the spread already captures. In a well-specified regression, we could test for this by just running a horse race between spreads and ratings in explaining a third macro variable.[10] However, if the variables are endogenous or if they are measured with error, as we argue they are in this case,[11] then this simple procedure might not produce the correct inference. Instead, the proposed specification test is shown to be robust to the most typical forms of misspecification, in particular, the anticipation effect of rating changes that is observed in the data.

In the second part of this article, we also considered a type of horse race between ratings and spreads as to how well they are correlated to other macroeconomic variables using daily data. The value added here is the use of high-frequency data as it is conceivable that if ratings have any informational content beyond spreads, we expect this information to be incorporated into macro variables within days; therefore, monthly or lower frequency data will be unable to disentangle the spread and the ratings informational components.[12] The results from the horse race exercises suggest that ratings usually enter these regressions with a statistically significant sign after controlling for the spread. This is consistent with the results of the specification tests, thereby suggesting that ratings add information beyond what is already incorporated into market prices.

However, we also argue that outlook changes give interesting information on how anticipated rating changes have been. If the outlook is changed just a few days before the rating, then it seems reasonable to suggest that the rating change is largely unanticipated before that date. We found that unanticipated rating changes have a bigger impact on asset prices, which is additional supportive evidence about the informational content of ratings.

We also conducted tests on whether certain rating changes are more important than others. In particular, if a debt issue obtains an investment grade rating, this may allow different classes of investors to purchase those issues, and hence the instrument may be said to have changed asset class. We tested whether rating changes in and out of investment grade are more important than other changes. We found that rating changes between asset classes have no additional explanatory power vis-à-vis all the other rating changes. These results are reassuring as they suggest that rating changes do not drive market movements for purely technical reasons that are unrelated to the underlying informational content of ratings.[13]

Our results across several methods and for the three main credit rating agencies are strong and highly consistent. We found that we cannot reject the view that rating agencies add value. We found that this is true for both changes in asset classes and other rating changes, and we found that less anticipated rating changes have even more significant effects.

DOI: 10.1002/ijfe

2. ORGANIZING FRAMEWORK

What we were interested in was evaluating the informational content that the rating has in addition to the observed spread on marketable sovereign debt (henceforth spread).[14] In other words, the null hypothesis that all the information in the rating is already reflected in the spread is equivalent to saying that the spread is a sufficient statistic. The alternative hypothesis, on the other hand, implies that spreads and ratings are imperfect measures of the unobservable fundamentals of the economy, and therefore, ratings provide information above and beyond what spreads reflect.

In this section, we organize our thoughts regarding ratings and spreads in a simple error-in-variables (EIV) framework. The goal was to devise a simple specification test to evaluate whether ratings are informative.

2.1. Preliminary considerations

Some considerations are necessary to clarify before devising an empirical strategy. First, this article is studying sovereign ratings, and in this context, rating agencies are concerned with evaluating countries' probability of default, or country risk. This is important because in this environment, if the spread of the sovereign debt is observed, then it is reasonable to assume that the spread and the rating are supposedly capturing the same aspect.

It is impossible to evaluate the informational content in the rating only using ratings and spreads. We need other variables. Fortunately, country risk not only affects the spread but also impacts other macroeconomic variables. For instance, an increase in the probability of default of a country should have a negative impact on all asset prices, particularly stock prices. Therefore, if we observe a downgrade, we should expect a drop in the stock market index. If the spread is a sufficient statistic for the rating, then if we were to run a regression where the spread and the rating are included on the Right Hand Side (RHS), the rating should be insignificant after controlling for the spread. In fact, we studied three macro variables: the spread one period ahead, stock market prices, and the nominal exchange rate.

The second consideration is that we concentrated on high-frequency (i.e., daily) data. This explains our choice of macro variables. The main reason we look at daily data is that, if ratings have any informational content beyond the spread, we expect this information to be incorporated into macro variables within days; therefore, monthly data will be unable to disentangle the spread and the ratings informational components.

Third, if ratings and spreads are imperfectly measuring the fundamental—default probability—then we can interpret them as noisy versions of an unobservable fundamental. However, the rating, because of its discrete nature, is then a version of the fundamental whose noise is not of classical form. In other words, the rating can be interpreted as a discretization of the fundamental, and the noise implied in this measure is serially correlated and correlated with the fundamental—hence making it a nonclassical EIV problem. Our methodology testing for the informational content has to be robust to this property of the data. Furthermore, ratings are very sticky, in the sense that they change very infrequently when observed daily. This means that the EIV problem in the rating is probably more severe than in the spread estimation.

Fourth, exchange rates, spreads, stock prices, and ratings are all endogenous. The methodology we devise has to take into consideration that linear regressions might be misspecified. The test has to be meaningful even in the presence of other forms of misspecification (not just the EIV interpretation). More importantly, a crucial form of endogeneity is the fact that credit rating changes are indeed anticipated by market participants. This affects not only the interpretation but also the implementation of the estimation. We return to the point of anticipation later in the Results section.

2.2. Specification test

With these four considerations at hand, we now proceed to explain our empirical strategy. We assume that spreads (i_t) and ratings (r_t) are noisy versions of an unobserved fundamental,

$$i_t = i_0 + \theta x_t + \varepsilon_t \tag{1}$$

$$r_t = r_0 + f(x_t, \eta_t) \tag{2}$$

where the idea is thatx_t is the unobserved fundamental that not only affects the probability of default of the country (and its spread) but also affects the exchange rate, stock markets, and future spreads. We assumed that the rating is a nonlinear function of the fundamental—trying to emphasize the discreteness of the variable. We assumed a simple linear function for the spread, although that is not restrictive.

DOI: 10.1002/ijfe

Assume that another macroeconomic variable y_t(which for expositional simplicity let us assume it is the stock market) is affected by the same fundamental,

$$y_t = y_0 + \beta x_t + \mu_t \tag{3}$$

The null hypothesis is that the spread is a sufficient statistic—that is, that the rating does not add information beyond what the spread already captures. In a well-specified regression, we could test for this by just running a horse race between spreads and rating. However, if the variables are endogenous or if they are measured with error, then this simple procedure might not produce the correct inference.

To resolve this problem, we took several steps in the estimation procedure. First, we concentrated on the relationship between macro variables, spreads, and ratings around the periods in which the rating changes. Our preferred specification looks at the window 10 days before and after a credit rating is modified. Second, we computed the cumulative return on all the variables over the events windows. This means that if the movement in the rating is anticipated, spreads and macro variables will adjust before the rating actually changes. Hence, all will be endogenously determined. Third, in this environment, we regressed the cumulative change in the macro variables on the spread and compared the estimates when the spread was instrumented by the rating. If the spread is a sufficient statistic for the rating, the two coefficients should be similar. If the spread and the rating summarized different sets of information—that is, both are imperfect measures of the fundamentals—then the two coefficients will be statistically different.

This procedure is robust to misspecification of the macro variable on the spread regression. In other words, when we say that the spread is a sufficient statistic for the rating, technically what we are saying is that the change in the rating is captured by the movement of the spread, and everything else in the rating is just noise. Instrumenting the spread with the rating around the window in which the rating is changing therefore implies that both capture the same change in fundamentals. By concentrating on the window around the rating change, we are minimizing the EIV in the rating measure and providing the best chance to the rating to provide additional information.[15]

What does it mean that the spread is a sufficient statistic for the rating? The simple model that follows highlights a case in which the spread is indeed a sufficient statistic.

$$i_t = i_0 + \theta x_t \tag{4}$$

$$r_t = r_0 + f(x_t, \eta_t) \tag{5}$$

$$y_t = y_0 + \beta x_t + \mu_t \tag{6}$$

2.2.1. Uncorrelated EIV and exogenous fundamentals

Let us start by studying the case when all the residuals are uncorrelated. Because the spread captures the information in the fundamental perfectly, when we estimate the regression,

$$y_t = c_0 + b i_t + \varphi_t \tag{7}$$

the Ordinary Least Squares (OLS) estimate is consistent. Because the rating is a noisy version of the same fundamental, and its noise is uncorrelated with the residual in the stock market equation, then if we instrument the spread with the rating, we also estimate a consistent coefficient. Importantly, the instrumental variable estimate is inefficient under the null hypothesis, and OLS is efficient.

Under the alternative hypothesis, the spread is a noisy version of the fundamental. This means that the OLS estimate is inconsistent and biased; the bias comes exactly from the noise. This estimate can be improved, however, and the rating is a perfect instrument for doing so. First, it is correlated with the spread because both are measures of the same fundamental. Second, their noises are different, and such noises are uncorrelated with the fundamentals. This means that the rating is uncorrelated with the residual in the stock market regression. In other words, the rating is a valid instrument for the spread, and the Instrumental Variables (IV) estimates are going to be a consistent estimate of the true parameter.

This is a standard specification test. Under the null hypothesis, OLS is consistent and efficient, whereas IV is consistent but inefficient. On the other hand, under the alternative hypothesis, OLS is inconsistent, but IV continues to be consistent (see Hausman, 1978).

DOI: 10.1002/ijfe

2.2.2. Uncorrelated EIV and endogenous fundamentals

The most important source of possible misspecification in this model is when the fundamentals are not exogenous—in other words, when $\mathrm{cov}(x_t, \mu_t) \neq 0$.

The methodology we have described has no problems dealings with this form of misspecification. Let us assume that the measured fundamental and the residual in the stock market equation are correlated. The implication of this assumption is that OLS is biased, but because in our window the rating is proportional to the fundamental x_t, then the IV will be equally biased if and only if the spread is a sufficient statistic. In other words, if the spread is a sufficient statistic but the fundamentals are correlated with the residual in the macro equation, OLS and IV are equally biased. In the alternative hypothesis, when the spread is not a sufficient statistic, then both coefficients are biased, but they are biased differently.

The simplest way to understand the intuition behind this test is to assume that both the spread and the rating are linear functions of the fundamental:

$$i_t = i_0 + \theta x_t \tag{8}$$

$$r_t = r_0 + \alpha x_t + \eta_t \tag{9}$$

$$y_t = y_0 + \beta x_t + \mu_t \tag{10}$$

The OLS estimate of the stock market on the spread is equal to

$$\hat{b}_{\mathrm{OLS}} = \frac{\mathrm{cov}(i_t, y_t)}{\mathrm{var}(i_t)} = \frac{\beta\theta\,\mathrm{var}(x_t) + \theta\,\mathrm{cov}(x_t, \mu_t)}{\theta^2\,\mathrm{var}(x_t)} = \frac{\beta}{\theta} + \frac{\mathrm{cov}(x_t, \mu_t)}{\theta\,\mathrm{var}(x_t)} \tag{11}$$

where, just for clarification, the bias arises from the correlation between the fundamental and the residual in the stock market regression. It is needless to say that the OLS estimate—when consistent—is an estimate of the ratio between $\frac{\beta}{\theta}$.

In this environment, the IV estimate is (using the rating as the instrument)

$$\hat{b}_{\mathrm{IV}} = \frac{\mathrm{cov}(r_t, y_t)}{\mathrm{cov}(r_t, i_t)} = \frac{\beta\alpha\,\mathrm{var}(x_t) + \alpha\,\mathrm{cov}(x_t, \mu_t)}{\theta\alpha\,\mathrm{var}(x_t)} = \frac{\beta}{\theta} + \frac{\mathrm{cov}(x_t, \mu_t)}{\theta\,\mathrm{var}(x_t)} \tag{12}$$

where the source of the misspecification $\mathrm{cov}(x_t, \mu_t) \neq 0$ is exactly the same in both regressions. Notice that both estimates are numerically the same.

Under the alternative hypothesis, the two estimators are going to differ from each other. The OLS estimator has two forms of bias: one from misspecification and one from the EIV. On the other hand, the IV estimate will have only bias from misspecification. In the end, the test is roughly the same: The coefficients should be the same under the null hypothesis but different in the alternative hypothesis. The main difference is the interpretation of the coefficients, but not the validity of the test.

This is an important characteristic of our design because, certainly, changes in ratings, spreads, and financial variables are endogenous, they are driven by common shocks that are unobservable, and rating changes might be anticipated.[16] Our test will be able to deal with these aspects.

This example highlights the form of specification that we can solve analytically. It is the one in which the fundamental and the residual of the economy are correlated but the EIV are still orthogonal to everything else. In other words, this solves the most basic (and possibly important) form of misspecification: the fact that the fundamentals and the residuals in the stock market are correlated. For instance, this covers omitted variable biases and endogeneity. In particular, this includes the anticipation of rating changes.

2.2.3. Correlated EIV

Assume that the errors in the rating equation are also correlated with the fundamental; then, the estimate of the IV is slightly different from the OLS:

$$\hat{b}_{\mathrm{IV}} = \frac{\mathrm{cov}(r_t, y_t)}{\mathrm{cov}(r_t, i_t)} = \frac{\beta\{\alpha\,\mathrm{var}(x_t) + \mathrm{cov}(x_t, \eta_t)\} + \alpha\,\mathrm{cov}(x_t, \mu_t)}{\theta\{\alpha\,\mathrm{var}(x_t) + \mathrm{cov}(x_t, \eta_t)\}} = \frac{\beta}{\theta} + \frac{\alpha\,\mathrm{cov}(x_t, \mu_t)}{\theta\{\alpha\,\mathrm{var}(x_t) + \mathrm{cov}(x_t, \eta_t)\}} \tag{13}$$

In this case, the estimates (IV and OLS) will be different because the noise of the rating is correlated with the fundamental. Interestingly, in this case, the rating is indeed providing information above and beyond that contained

DOI: 10.1002/ijfe

in the spread, and therefore, a rejection should be found. However, in this case, the information is not necessarily contained in the actual change in the rating but in its noise. This is important because we will be able to conclude with our method whether the rating contains information, although we do not know—or will not be able to disentangle—its source.

In summary, if the spread is a sufficient statistic, then it captures all the relevant fluctuation of x_t that is contained in the rating. Because the stock market (or exchange rate) equation is likely to be misspecified, the test can be performed, but the coefficients cannot be interpreted—structurally speaking. If the spread is a noisy measure of the fundamental (add noise to the first equation of our model) or the noise of the rating is correlated with the fundamental or the residual, then the rating is indeed providing information beyond the one contained in the spread, and we have shown that the estimation of the OLS and IV coefficients will differ from each other.[17]

2.3. Error-in-variables framework

Finally, before discussing the estimation and results, we devote our attention to the EIV interpretation we are providing to the spread and the rating. In Figure 2, we have depicted the fundamental, the spread, and the rating. In general, we assumed that the spread differs from the fundamental and that those differences can be captured with a standard classical EIV. *A priori*, there is no reason to have a different view on the discrepancy between the fundamental and the spread. In fact, most will argue that there is no difference and that the spread indeed captures the fundamental.

The difference between the fundamental and the rating is what we interpret as the EIV. The idea is that the rating is trying to capture the fundamental, but it is a discretized version of it. If the fundamental increases, the rating increases, but it does so in a "sticky" way. This implies that the EIV in the rating clearly is nonclassical.

In other words, the EIV are serially correlated. When the rating is below the fundamental, it is very likely to continue to be below the fundamental in the following period. A classical error is serially uncorrelated. Second, and probably more importantly, when the rating remains the same and the fundamental increases, the EIV increases, which means that the EIV is correlated with the fundamental. Finally, around the credit rating changes, the EIV are serially negatively correlated. The reason is that if there is a trend in the fundamental, and the rating moves up, then the errors prior to the change in the rating were negative, and they are likely to be positive afterward.

When the spread is a sufficient statistic, we are assuming that the spread measures the fundamental without error, and therefore, the spread captures x_t perfectly, whereas the rating does not. When the spread is not a sufficient statistic, we assume that the EIV for the spread is classical, whereas the one for IV is not.

One question that should arise immediately is what assumptions are needed for the IV strategy to be valid. This is very simple: We just need the EIV of the rating to be uncorrelated with the EIV of the spread—which we assume is trivially satisfied under the null hypothesis (given that the error is exactly zero for the spread under the null).

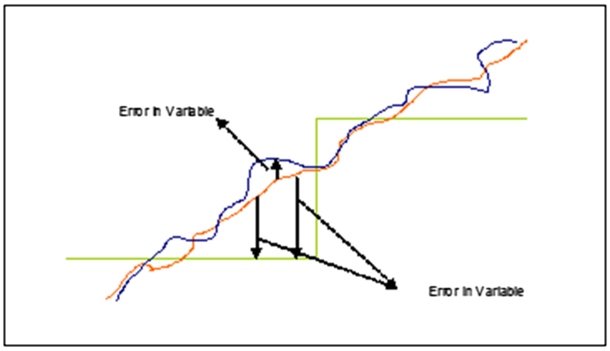


Figure 2. Error-in-variables framework.

3. DATA

3.1. Dataset and methodology

The raw data for this study came from the Bloomberg database and from rating industry sources. From Bloomberg, we collected daily information available for 32 emerging market economies between 1 January 1998 and 25 April 2007.[18] In particular, we collected data on the following macroeconomic variables: sovereign spreads, nominal bilateral exchange rates (domestic currency units vis-à-vis the US dollar), and local stock market indices.[19] We also collected data on the so-called volatility index (VIX), a widely used measure of market risk.[20] From the three main rating agencies (Fitch, Moody's, and S&P), we collected data on ratings and outlooks for the same dates, and we tabulated the days of rating and outlook changes.[21] The resulting dataset is an unbalanced panel with 77 760 observations.

The ratings from the three agencies are transformed into a numeral scale (between 1 (*lowest*} and 21 (*highest*)) using the scale proposed by Afonso *et al*. (2007) (Table 1).

The next step consisted of rearranging the master dataset to make it amenable to the analysis. For this purpose, first we defined *events* as changes in the ratings for each of the three rating agencies. Rating changes are either upgrades or downgrades of one notch or more. Table 2 summarizes the resulting events per rating agency.

For each of these events, we defined a 21-day window[22] centred on the day of event. Thus, the rating becomes a step variable within each window: It has a starting value for the first 10 days, then jumps on day 11 (either upgrade or downgrade), and then remains at the new value for the subsequent 10 days.[23]

Next, to make the rest of the data comparable across countries and events, we normalized the variables so that the starting point for every series in each event window is the same. The normalization consists of taking, for every day t in the window, the following transformation:

$$y_t = \log(X_t) - \log(X_0) \tag{14}$$

where X is, alternatively, the sovereign spread, the stock market index, the nominal exchange rate, and the VIX; X_0 is the value of the corresponding variable on the first day of the window; and y_t is the transformed variable, which is

Table 1. Rating Scale

Fitch rating	Number	Moody's rating	Number	Standard & Poor's rating	Number	
AAA	21	Aaa	21	AAA	21	Investment grade
AA+	20	Aa1	20	AA+	20	
AA	19	Aa2	19	AA	19	
AA–	18	Aa3	18	AA–	18	
A+	17	A1	17	A+	17	
A	16	A2	16	A	16	
A–	15	A3	15	A–	15	
BBB+	14	Baa1	14	BBB+	14	
BBB	13	Baa2	13	BBB	13	
BBB–	12	Baa3	12	BBB–	12	
BB+	11	Ba1	11	BB+	11	Speculative grade
BB	10	Ba2	10	BB	10	
BB–	9	Ba3	9	BB–	9	
B+	8	B1	8	B+	8	
B	7	B2	7	B	7	
B–	6	B3	6	B–	6	
CCC+	5	Caa1	5	CCC+	5	
CCC	4	Caa2	4	CCC	4	
CCC–	3	Caa3	3	CCC–	3	
CC	2	Ca	2	CC	2	
C	2	`C	1	SD	1	
DDD	1			D	1	
DD	1					
D	1					

Source: Afonso *et al*. (2007).

DOI: 10.1002/ijfe

Table 2. Number of Events by Rating Agency

	Number of events	Downgrades	Upgrades
Standard & Poor's	145	62	83
Fitch	111	44	67
Moody's	90	39	51

simply the cumulative return. Thus, the initial value for these variables in each event window (y_0) is normalized at zero. Table 3 reports the summary statistics for the normalized variables grouped by rating agencies.

The first panel shows that for the case of S&P ratings, where we have 145 events, we end up with 3045 observations for the rating (i.e., 145 events $\times$ 21 days per event). We also report the mean and the standard deviation of the rating for all the events. In the rows, we report the summary statistics for the other variables of interest, where, for example, a value of 0.01 for the mean indicates that the average value of the corresponding variable for all the available days, across all events, is 1% higher than the average value on the first day of the window. The other two panels replicate the same exercise but for events based on the data from the other two rating agencies.

3.2. Relationship between spreads and ratings

As discussed in the Introduction section, several recent articles consider the relationship between spreads and ratings. Eichengreen and Mody (1998) argued that ratings are important in explaining spreads. They regressed ratings on fundamentals and then introduced the residual of that regression together with fundamentals in a regression to explain spreads. They argued the residual reflects the rating agency opinion and found that it is highly significant. González Rozada and Levy Yeyati (2008) suggested that a large component of individual country spreads is driven by global factors such as the overall EMBI spread or the US high-yield spread. In one specification, they included the rating as a control for country fundamentals and found it to be significant with the expected sign. Powell and Martínez (2007) started with a simple regression of spreads against ratings and suggested that a simple log–log relationship works reasonably well to capture how an improvement in the rating may lead to a reduction in spreads. They suggested, though, that the reduction in spreads to June 2007 levels is only partially explained by the improvement in ratings. They replicated the results of Eichengreen and Mody (1998) and also suggested a system of equations with similar results, suggesting that ratings may matter. They also exploited the differences between rating agencies' opinion and showed that those differences may be informative in explaining spreads.

Table 3. Summary statistics

Variable	Observations	Mean	Standard deviation
Standard & Poor's			
Rating	3045	8.51	3.80
Spread	2533	0.01	0.17
Stock market	2438	−0.01	0.10
Exchange rate	2996	0.01	0.06
VIX	3045	−0.01	0.13
Fitch			
Rating	2331	9.14	3.36
Spread	2159	0.03	0.16
Stock market	1768	0.00	0.10
Exchange rate	2265	0.02	0.09
VIX	2331	0.00	0.13
Moody's			
Rating	1890	9.13	3.31
Spread	1718	0.03	0.18
Stock market	1582	−0.02	0.11
Exchange rate	1832	0.01	0.07
VIX	1890	0.02	0.14

DOI: 10.1002/ijfe

The differences in opinions between rating agencies can be represented in various ways. In this article, we focussed on rating changes as events. Later, we will present a Venn diagram that summarizes the distribution of events across the three rating agencies and their overlap. As explained, in the baseline, each event has a 21-day window. Thus, an overlap (or a potential agreement) occurs when rating changes for the different agencies happen within the same window. For example, out of 141 events for S&P,[24] 21 overlap with events of Fitch, 12 with events of Moody's, and 15 with the two rating agencies concurrently. The general message that emerges from Figure 3 is that the overlap is relatively small across the three rating agencies. This suggests that the rating agencies do not always act concurrently and hence that disagreements between agencies persist. In turn, this suggests that the informational content of the events across the agencies might be different. In particular, if the credit ratings are not perfectly correlated, then they all three cannot be fully explained by the exact same statistic (in this case, the spread). In other words, given how uncorrelated the actions of the ratings agencies are, it should be *a priori* clear that they provide different information among themselves. If one of these ratings is perfectly explained by the spread, then the other two cannot. Therefore, in the analysis that follows, we considered these differences and tested the validity of our results using the data from the three rating agencies.

4. RESULTS

4.1. Specification test

We applied a standard Hausman specification test. This is performed in two steps. First, we estimated the following models:

OLS Model

$$y_{i,t} = \alpha_{OLS} \times i_{i,t} + \theta \times VIX_{i,t} + \kappa_i + \varepsilon_{i,t}; \quad i = \text{ events, and } t = \text{days} \tag{15}$$

where $y_{i,t}$ is alternatively $i_{i,t+1}$ (i.e., the spread 1 day forward), $s_{i,t}$(i.e., the stock market index), and $ner_{i,t}$ (i.e., the nominal exchange rate); κ_i is an event-fixed effect; and $\varepsilon_{i,t}$ is the error term. The VIX is included to control for the effect of global factors.

We also run instrumental variables version of this regressions, where the only variant is that we instrument spreads with ratings:

IV Model

$$\begin{aligned} y_{i,t} &= \alpha_{IV} \times i_{i,t} + \theta \times VIX_{i,t} + \kappa_i + \varepsilon_{i,t} \\ i_{i,t} &= r^j_{i,t} \end{aligned} \tag{16}$$

where j is alternatively S&P, Moody's, or Fitch ratings.

For robustness checks purposes, we also run an error correction model for the case when the dependent variable is the spread. In this case, the estimated equation is as follows:

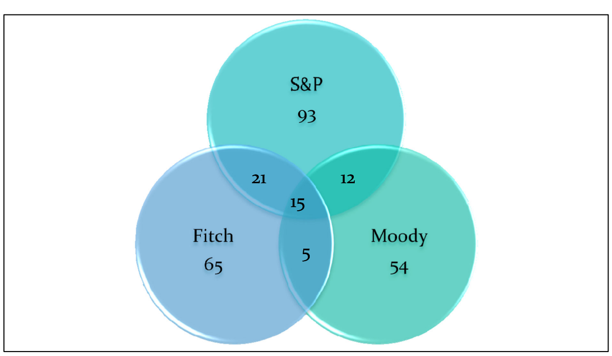


Figure 3. Venn diagram.

DOI: 10.1002/ijfe

Error Correction Model

$$\Delta i = \alpha \times i_{i,t} + \theta \times VIX_{i,t} + \phi \times \Delta VIX + \kappa_i + \varepsilon_{i,t} \tag{17}$$

$$\text{where } \Delta i = i_{i,t+1} - i_{i,t}, \text{ and } \Delta VIX = VIX_{i,t+1} - VIX_{i,t} \tag{18}$$

In the IV variant of the error correction model, we simply instrument the spread with the rating.

The second step consists of applying a specification test using the estimates from these models. Hausman (1978) proposed a test where a quadratic form in the differences between two vectors of coefficients, scaled by the matrix of the difference in the variances of these vectors, gives rise to a test statistic (chi-squared). Under the null hypothesis, OLS is consistent and efficient, whereas IV is consistent but inefficient. On the other hand, under the alternative hypothesis, OLS is inconsistent, but IV continues to be consistent.

Table 4 summarizes the results we obtain when we apply this test to our baseline specification (i.e., using all the events—upgrades and downgrades— from S&P and a window of 21 days per event).[25] Every column in the table is a different dependent variable, and the last column is the error correction model. In the first two rows, we report α_{OLS} and α_{IV}, respectively.[26] Thus, the coefficient reported in the first row under the first column is the OLS estimate for the effect of the current spread on the spread 1 day forward.

The OLS results suggests that increases in the spread have a positive effect on the spread forward (first and fourth columns), are related to decreases in the stock market index (second column), and are also related to depreciations of the nominal exchange rate vis-à-vis the US dollar (third column). The IV results (i.e., instrumenting spreads with ratings) are qualitatively similar. What the Hausman specification test reveals is, in essence, if these coefficients are also quantitatively the same.[27] If they are statistically different, the null hypothesis is rejected—that is, OLS is inconsistent. Quite importantly for our purposes, the rejection of the null hypothesis is evidence that the spread is not a sufficient statistic.

In the next two rows of Table 4, we report the Hausman statistic (chi-squared) and the corresponding *p*-value. The results are that the null hypothesis is rejected at standard confidence levels (10% or less) in three out of four cases. This suggests that, for these selected macro variables, the spread is not a sufficient statistic. In other words, not all the information in the rating is reflected in the spread, and thus the rating explains some of the variation in these macro variables. It is worth reemphasizing here that the test is valid even if the OLS regression is misspecified, at least for the most common forms of misspecification.

The next step consists of checking the robustness of these results: We wanted to evaluate if we get a high number of rejections for the specification test across different possible variants. In Table 5, we summarize the results of a series of robustness checks. The first set of checks consists of splitting the sample of events into upgrades and downgrades and running the regressions separately. Next, we repeated the same exercise for both the full and the split samples, using the events of Fitch and Moody's. Then, we went back to the S&P data and changed the event window to 11 days per event (i.e., 5 days around the rating change) and also to 41 days per event (i.e., 20 days around each rating change). Finally, we dropped the few events that occur within the same 21-day window (contemporaneous events).[28]

Table 4. OLS Versus IV

	Spread $t+1$	Stock market	Exchange rate	Spread $t+1$ (error correction)
OLS	0.906***	−0.217***	0.100***	−0.095***
	(0.010)	(0.009)	(0.007)	(0.010)
IV	1.008***	−0.280***	0.109***	0.008
	(0.025)	(0.024)	(0.017)	(0.025)
Hausman test (Ch^2)	20.13	8.03	0.33	20.48
P-value	0.001	0.018	0.848	0.001

Notes: S&P ratings are used for these regressions.
To perform these estimations, the data were arranged to allow a 10-day window around the day of the change in the rating.
The OLS coefficient is the estimated effect of the change in spread on the corresponding dependent variable.
The IV is the coefficient obtained when the spread is instrumented by the rating.
All these regressions include event fixed effects and the volatility index (VIX) as controls.
The null hypothesis in the Hausman test is that the OLS estimator is more efficient.

DOI: 10.1002/ijfe

Table 5. Hausman Test, *P*-Values

	Spread $t+1$	Stock market	Exchange rate	Spread $t+1$ (error correction)	Rejection rate[a]
Standard & Poor's (downgrades + upgrades)	0.001	0.018	0.848	0.001	75%
Standard & Poor's (downgrades)	0.010	0.800	0.436	0.018	50%
Standard & Poor's (upgrades)	0.001	0.140	0.001	0.015	75%
Fitch (downgrades + upgrades)	0.430	0.600	0.001	0.700	25%
Fitch (downgrades)	0.960	0.001	0.001	0.960	50%
Fitch (upgrades)	0.190	0.001	0.031	0.420	50%
Moody's (downgrades + upgrades)	0.066	0.061	0.082	0.160	75%
Moody's (downgrades)	0.355	0.053	0.001	0.725	50%
Moody's (upgrades)	0.078	0.009	0.001	0.154	75%
Standard & Poor's: 5-day window (all)	0.001	0.078	0.771	0.001	75%
Standard & Poor's: 5-day window (downgrades)	0.001	0.770	0.018	0.021	75%
Standard & Poor's: 5-day window (upgrades)	0.100	0.017	0.001	0.235	75%
Standard & Poor's: 20-day window (all)	0.001	0.660	0.850	0.001	50%
Standard & Poor's: 20-day window (downgrades)	0.001	0.001	0.670	0.001	75%
Standard & Poor's: 20-day window (upgrades)	0.001	0.068	0.001	0.001	100%
Standard & Poor's: without contemporaneous change in rating	0.000	0.091	0.953	0.002	75%
Rejection rate[b]	75%	69%	63%	56%	

Note: Every cell is the *P*-value of the Hausman test in the correspondent OLS versus IV regressions.
[a]Number of rejections of the null hypothesis in the Hausman test over the total regressions run per dependent variable.
[b]Number of rejections of the null hypothesis in the Hausman test over the total regressions run per specification.

For each of these alternative specifications, we run the OLS, IV, and error correction models and performed the corresponding Hausman test. In Table 5, we report the *p*-values. For comparability purposes, in the first row, we report the *p*-values from the previous regressions (Table 4). The last row and the last column in the table are the rejection rates, that is, the percentage of rejections of the null hypothesis for each row or column.

The results are very telling: The rejection rate varies between 56% and 75% in every column, which means that we reject a lot across many possible permutations of the dependent variable and also the estimation model. In the case of the rows, the rejection rate is below 50% only once, that is, Fitch upgrades and downgrades. The high rejection rates across the board reinforce the conclusion that the spreads are not a sufficient statistic. In other words, there seems to be some informational content in ratings that is not captured by the spreads.[29]

At this point, we can also evaluate the robustness of the test to the misspecifications that we are not fully able to solve analytically. In particular, recall from the methodological section that if the errors in the rating equation are also correlated with the fundamental, then the estimate of the IV is slightly different from the OLS. In this case, the estimates (IV and OLS) will be different under the null hypothesis. If this form of misspecification is significant, we expect more rejections the bigger the windows are. The reason is that the EIV implied in the rating grows with the window in which the rating is not changing. We did not find this in our tests. On the contrary, if anything, focusing on the case of the full sample (upgrades and downgrades for S&P), we found that the rejection rate was smaller when the width of the event window was increased to 20 days around the event.[30] At the same time, the rejection rate for the cases when the width of the event window was only 5 days around the event was 75%—the same as the baseline—hence indicating that the EIV introduced by the nonlinearity is not significantly large. Despite this, and even if this particular form of misspecification is significant and we did not find more rejections when we expanded the window simply because widening the window weakens the power of the test (because the instrument becomes noisier and hence weaker), the reader should rest assured that the validity of the test is not invalidated because, as explained in Section 2, we still expect to find rejections if the rating is providing information above and beyond the one contained in the spread. The only difference is that we cannot disentangle whether this information comes from the rating change itself or from the noise.

Having established that spreads and ratings are different, in the next section, we run a horse race between these variables. If ratings have informational content as we suggest, then we expect that when we run a regression where both variables are included on the RHS, the rating should be significant after controlling for the spread.

DOI: 10.1002/ijfe

4.2. Horse race

Having established that spreads are not a sufficient statistic, we turn now to estimating a new model in which we exploited the informational content of ratings to explain the variation in three macro variables using high-frequency data. These regressions are similar to the ones by Kaminsky and Schmukler (2002), but adding an additional control for current spreads.[31]

We estimated the following OLS model:

$$y_{i,t} = \alpha \times i_{i,t} + \beta \times r^{j}_{i,t} + \theta \times VIX_{i,t} + \kappa_i + \varepsilon_{i,t}; \quad i = \text{ events, and } t = \text{days} \tag{19}$$

where $y_{i,t}$ is alternatively $i_{i,t+1}$; $s_{i,t}$; $ner_{i,t}$; κ_i is an event-fixed effect; and $\varepsilon_{i,t}$ is the error term. The VIX is included to control for the effect of global factors.

For robustness checks purposes, we also run an error correction model for the case when the dependent variable is i_t. In this case, the estimated equation is as follows:

Error Correction Model

$$\Delta i = \alpha \times i_{i,t} + \beta \times r^{j}_{i,t} + \theta \times VIX_{i,t} + \phi \times \Delta VIX + \kappa_i + \varepsilon_{i,t} \tag{20}$$

$$\text{where } \Delta i = i_{i,t+1} - i_{i,t}, \text{ and } \Delta VIX = VIX_{i,t+1} - VIX_{i,t} \tag{21}$$

It is clear from the previous discussion on misspecification that we cannot interpret the magnitude of these coefficients in a structural way. Therefore, in what follows, we just focussed on the signs and their statistical significance. We wanted to test if ratings explain part of the variation in the cumulative returns of the macro variables over the selected event windows after we control for spreads, and also if rating and spreads are correlated to these macro variables in ways that make intuitive sense.

The results are reported in Table 6. The table is organized slightly different than the previous ones. The panel on the upper LHS has the results for the baseline regressions: S&P, all events, and a 21-day window for each event. Every row is a different regression: either a different dependent variable or the error correction model. Every column is the estimated coefficient for the corresponding RHS variable. The standard errors are reported in parentheses below every point estimate. To make the interpretation easier, we put asterisks next to the coefficients that are statistically significant.[32] Thus, the first row shows the results of estimating the model by OLS for the case in which the dependent variable is the spread 1 day forward. We found that, as expected, α is positive and statistically significant, meaning that increases in the spread today (i.e., a higher perceived probability of default) are correlated with increases in the spread tomorrow.

Interestingly, β enters with a negative sign and is also statistically significant, meaning that an increase in the rating (i.e., an upgrade) is correlated with a decrease in spreads 1 day forward. The fact that the rating is significant after controlling for the spread is additional evidence in favour of the hypothesis that spreads are not a sufficient statistic. The third RHS variable included in the regression, the VIX, is positive but not statistically significant.

Next, we changed the LHS variable to the stock market index. In this case, we found that increases in the spread are associated with decreases in the stock market indices, whereas an increase in the rating, controlling for the spread, is correlated to a statistically significant increase in the stock market. Finally, in this case, the coefficient estimate for the VIX is negative and statistically significant.

The next row presents the results for the case in which the LHS variable is the nominal exchange rate vis-à-vis the US dollar. The results are that increases in the spread are associated to nominal exchange rate depreciations, a result that we found consistent with what we would expect for emerging market economies: Higher probability of default is oftentimes associated with capital flight and a weakening of the domestic currency. At the same time, the estimated effect for changes in the rating, in this case, is not statistically significant. Note, incidentally, that this is the one case for which we did not reject the Hausman test for the baseline specification in Table 4. This is additional evidence in favour of the power of the test: In the case where we did not reject the specification test, we found that the rating is insignificant after controlling for the spread (i.e., the rating provides no additional information). Finally, we found that the coefficient estimate for the VIX is also positive and statistically significant.

In the last row, we report the results of the error correction model.[33] The results are reassuringly similar to those in the first row, which are based on the same dependent variable. The only difference is that the coefficient estimate for the VIX, although still positive, is now statistically significant at the 10% level.

DOI: 10.1002/ijfe

Table 6. OLS With Event Fixed Effects

	S&P upgrades & downgrades			S&P downgrades			S&P upgrades		
	Spread	Rating	VIX	Spread	Rating	VIX	Spread	Rating	VIX
Spread $t+1$	0.884***	−0.006***	0.006	0.894***	−0.006***	0.013	0.876***	−0.007***	−0.003
	(0.011)	(0.0014)	(0.015)	(0.014)	(0.001)	(0.017)	(0.017)	(0.001)	(0.022)
Stock market	−0.205***	0.004***	−0.104***	−0.484***	−0.002	−0.067***	−0.018**	0.002**	−0.085***
	(0.011)	(0.014)	(0.001)	(0.020)	(0.002)	(0.025)	(0.008)	(0.001)	(0.012)
Exchange rate	0.098***	−0.0005	0.045***	0.196***	−0.003	0.089***	0.007**	0.002***	−0.006
	(0.008)	(0.0009)	(0.010)	(0.018)	(0.002)	(0.022)	(0.003)	(0.0003)	(0.005)
ΔSpread	−0.117***	−0.006***	0.029*	−0.109***	−0.006***	0.030*	−0.124***	−0.006***	0.027
	(0.011)	(0.001)	(0.015)	(0.014)	(0.001)	(0.018)	(0.017)	(0.0019)	(0.024)
	Fitch upgrades and downgrades			Moody's upgrades and downgrades					
	Spread	Rating	VIX	Spread	Rating	VIX			
Spread $t+1$	0.863***	−0.002	0.036***	0.855***	−0.004**	0.040***			
	(0.010)	(0.001)	(0.011)	(0.013)	(0.002)	(0.015)			
Stock market	−0.404***	0.002	−0.132***	−0.297***	0.005**	−0.140***			
	(0.016)	(0.002)	(0.017)	(0.014)	(0.002)	(0.016)			
Exchange rate	0.225***	−0.009***	0.033**	0.190***	−0.003**	0.046***			
	(0.013)	(0.002)	(0.014)	(0.010)	(0.0014)	(0.012)			
ΔSpread	−0.139***	−0.001	0.064***	−0.147***	−0.004**	0.070***			
	(0.010)	(0.001)	(0.012)	(0.013)	(0.002)	(0.015)			
	S&P 5-day window			S&P 20-day window					
	Spread	Rating	VIX	Spread	Rating	VIX			
Spread $t+1$	0.742***	−0.005***	0.013	0.941***	−0.005***	0.021***			
	(0.019)	(0.001)	(0.019)	(0.006)	(0.0009)	(0.007)			
Stock market	−0.234***	0.003**	−0.040**	−0.271***	0.001	−0.173***			
	(0.018)	(0.001)	(0.017)	\(0.009)	(0.002)	(0.011)			
Exchange rate	0.048***	−0.0007	0.026*	0.168***	−0.0006	0.093***			
	(0.013)	(0.001)	(0.013)	(0.007)	(0.001)	(0.009)			
ΔSpread	−0.257***	−0.005***	0.029	−0.060***	−0.005***	0.037***			
	(0.019)	(0.001)	(0.021)	(0.006)	(0.0009)	(0.008)			

Notes: Regressions include event fixed effects as controls and estimated by OLS.
Every row is a different regression: either a different dependent variable or the error correction model.
Every column is the estimated coefficient for the corresponding RHS variable.
The standard errors are reported in parentheses below every point estimate.
S&P, Standard & Poor's; VIX, volatility index.

Next, we rerun the baseline specification, splitting the sample between upgrades and downgrades. The results for the case of downgrades are reported in the upper centre panel, whereas for the downgrades, they are reported in the upper right panel. The results are very similar to the previous ones, with only a couple of differences. When we focussed on downgrades, we found that the estimated effect of changes in the rating is no longer statistically significant when the LHS variable is the stock market. In the case of upgrades, the coefficient estimate for the effect of changes in rating on the nominal exchange rate is now positive and significant.

The middle panels of Table 6 report the results for the same exercise, but for the case of the events from the other two rating agencies. For concreteness, we concentrated only on the cases of the full sample (upgrades and downgrades stacked together). We found that in all cases, the coefficient estimate for α enters the regressions with the expected sign and is statistically significant: Increases in the spread today are associated with higher spreads tomorrow, decreases in the stock market, and nominal exchange rate depreciations. In the case of β, the estimated effect of changes in the rating, we found that for Fitch events, they typically have no explanatory power, except in the case when the macro variable is the exchange rate: In that case, we found that increases in the ratings (i.e., upgrades) are associated with nominal appreciations. This is interesting because this is the one case where we found an insignificant estimate for S&P. Also, it is consistent with the results of the specification test: In the case of Fitch, full sample, we rejected the null hypothesis only when the dependent variable was the nominal exchange rate. Instead, in the case of Moody's, the rating always enters the regressions with the expected sign and is statistically significant: Upgrades are associated with decreases in the spread forward, increases in the stock market, and nominal appreciations. In the case of VIX, the coefficient estimates are always significant in both samples and have the same signs: Increases in VIX are associated with higher spreads forward, lower stock market indices, and more depreciated nominal exchange rates.

Finally, in the lower panels of Table 6, we report the results for the cases in which we narrowed the width of the event window to 5 days around the event and, alternatively, expanded it to 20 days. We report the results based on the S&P sample only (upgrades and downgrades). The results are reassuringly similar to those of the baseline specification, with one exception. For the case of the expanded window, the effect of rating changes on the stock market is not statistically significant.

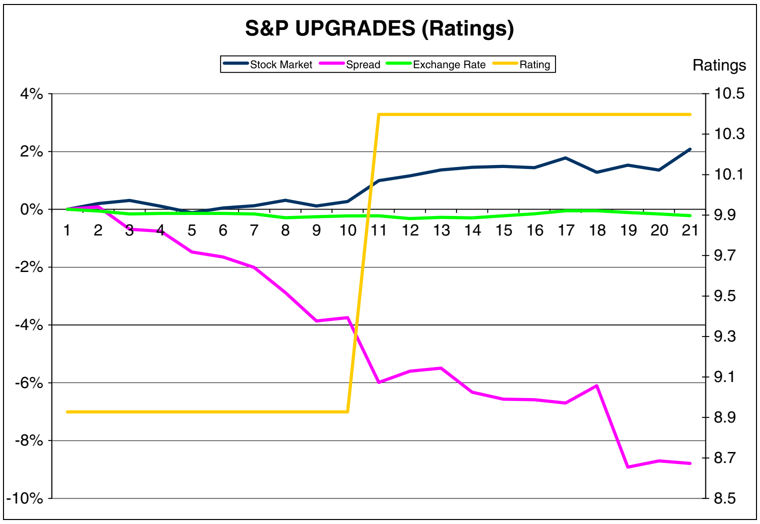
In summary, there are a few important takeaways from this section. First, the results from the horse race exercise suggest that ratings usually enter these regressions with a statistically significant sign after controlling for the spread. In the case of S&P ratings, this is true in three out of the four regressions; in the case of Moody's sample, it is always true; and in the case of Fitch ratings, it is true in just one case (which is incidentally the case when it is not true for S&P). This is consistent with the results from the specification tests (i.e., we rejected less for Fitch). At the same time, the high rate of rejections across the board suggests that spreads are not a sufficient statistic for the rating. Second, the additional robustness checks show that these results are consistent when we split the sample between upgrades and downgrades, and between expanding and narrowing the event window.

One interesting feature of the data is that the events are oftentimes anticipated by the market several days before the rating change. This is shown in Figures 3 and 4, where we plot the cumulative returns of spreads, stock market, and the nominal exchange rate around the days of the change in the rating. To facilitate the interpretation of the graphs, we separated between upgrades and downgrades. We present the results for the S&P sample only.

In the case of upgrades, we observed that the average change across all events is a one-notch increase in the rating (right scale, from 9 to 10). Although the rating change happens on day 11, by the time of the change, the stock market has already accumulated a 1% increase, and the spreads have fallen by approximately 4%. In both cases, this is roughly one half the cumulative changes over the entire event window. In the case of the nominal exchange rate, there are no noticeable effects, either before or after the rating change (recall that, for the S&P sample, the rating is not statistically significant in the regressions where the nominal exchange rate is the dependent variable).

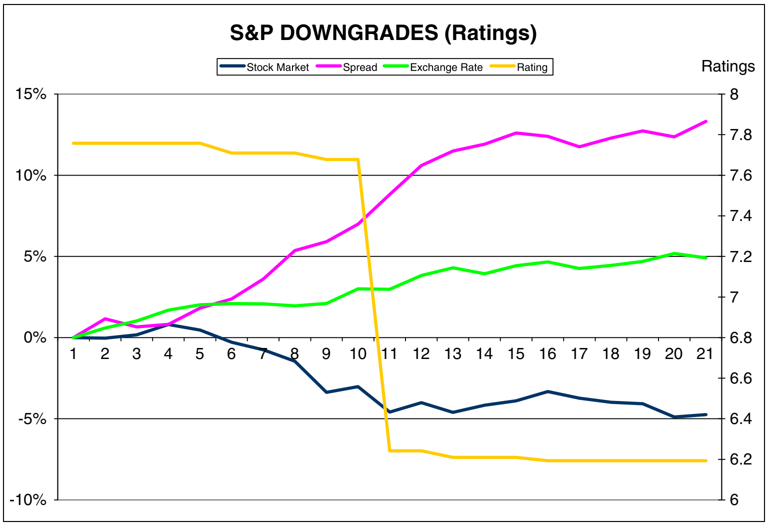
In the case of downgrades, the results are very similar. The average downgrade is 1.5 notches (from approximately 7.7 to 6.2). By the time the rating changes, the stock market has already declined—in this case, by almost the entirety of the total cumulative decline in the window. This might explain why, in the previous regressions, the coefficient estimate for β, in the case S&P downgrades, was not statistically significant for the stock market equation. Instead, spreads are on the rise before the downgrade, but they accumulate only about one half of the total increase by day 11. Finally, the nominal exchange rate appears to depreciate over the event window (Figure 5).

DOI: 10.1002/ijfe



Note: Rating Change occurs on day 11.

Figure 4. Standard & Poor's upgrades (ratings).



Note: Rating Change occurs on day 11.

Figure 5. Standard & Poor's downgrades (ratings).

The pattern of anticipation depicted in Figures 3 and 4 could be alternatively interpreted as evidence of reactivity in the behaviour of rating agencies. That is indeed the interpretation favoured by Kaminsky and Schmukler (2002). Instead, we interpret the results as evidence of an "anticipation" effect on the basis that in the specification tests of the first part of the article, we found that ratings add new information beyond that already reflected in market variables.

All in all, these results suggest that exchange rates, spreads, stock prices, and ratings are endogenous variables. Thus, as explained in Section 2, the methodology we devised has to take into consideration that linear regressions might be misspecified. We have already explained how the specification test is robust to this particular form of endogeneity. At the same time, this also affects how to implement the estimation and interpret the results. We return to the point of anticipation later, as we devote an entire section of this article to this issue. Before that, we perform some additional robustness checks.

DOI: 10.1002/ijfe

4.3. Asset class shift

In this subsection, we check if the results are robust to the introduction of nonlinearities for the case of rating changes. In particular, we wanted to explore if rating changes that happen between asset classes (i.e., investment to noninvestment grade and vice versa) explain more of the variation in the macro variables than do rating changes that happen within the same asset class. Even if rating changes are largely anticipated by the market by the time they are announced, there may be still some impact from rating news due to the fact that many institutional investors face limits in the amount of low-grade assets they can invest in or face clients' questions if they are exposed to a recently downgraded credit (note that a rating change may not surprise the specialist but may be news to the client who invests with him).[34] None of these technical reasons would be related to fundamental information and yet, at the high frequency studied, may have sometimes statistically significant correlations that may look like that. However, these technical reasons should be more prevalent in the data for cases when the rating change includes a shift in asset class.

To test if that is the case, we create a dummy variable that takes the value of 1 if the rating change is between asset classes and 0 otherwise. We interacted the new variable with the rating and included the interaction in the regression. The results, for the case of the S&P sample, are reported in Table 7. We found that rating changes between asset classes have no additional explanatory power vis-à-vis all the other rating changes: The interaction term is insignificant in all but one of the 12 regressions in Table 7. These results are reassuring as they suggest that

Table 7. Interaction With Dummy Variable of Change in Asset Class

	Spread	Rating	Rating*(Δ Asset Class)[a]	VIX
S&P upgrades and downgrades				
Spread $t+1$	0.812***	−0.004***	−0.014	−0.009
	(0.015)	(0.001)	(0.008)	(0.013)
Stock market	−0.101***	0.003**	0.001	−0.039***
	(0.015)	(0.001)	(0.008)	(0.013)
Exchange rate	0.018	0.002*	0.002	0.036***
	(0.013)	(0.001)	(0.007)	(0.011)
ΔSpread	−0.187***	−0.004***	−0.011	0.013
	(0.015)	(0.001)	(0.009)	(0.014)
S&P downgrades				
Spread $t+1$	0.894***	−0.005***	−0.016	0.012
	(0.014)	(0.002)	(0.012)	(0.017)
Stock market	−0.485***	−0.002	0.011	−0.065***
	(0.021)	(0.003)	(0.019)	(0.025)
Exchange rate	0.196***	−0.003	0.012	(0.022)
	(0.018)	(0.002)	(0.014)	−0.003
ΔSpread?	−0.109***	−0.006***	−0.015	0.029
	(0.014)	(0.002)	(0.011)	(0.018)
S&P upgrades				
Spread $t+1$	0.875***	−0.006***	−0.005	−0.004
	(0.017)	(0.002)	(0.012)	(0.022)
Stock market	−0.021**	0.003**	−0.023***	−0.088***
	(0.008)	(0.001)	(0.006)	(0.012)
Exchange rate	0.007**	0.002***	−0.004	−0.006
	(0.003)	(0.001)	(0.002)	(0.004)
ΔSpread	0.875***	−0.006***	−0.005	−0.004
	(0.017)	(0.002)	(0.012)	(0.022)

Notes: Regressions include event fixed effects as controls and estimated by OLS.
Every row is a different regression: either a different dependent variable or the error correction model.
Every column is the estimated coefficient for the corresponding RHS variable.
The standard errors are reported in parentheses below every point estimate.
[a]Interaction between the rating and a dummy that takes the value of 1 if the change in the rating implies a change between investment and noninvestment grade, and zero otherwise.
S&P, Standard & Poor's; VIX, volatility index.

DOI: 10.1002/ijfe

rating changes do not drive market movements for purely technical reasons that are unrelated to the underlying informational content of ratings. The fact that we did not find significant differential effects for some rating changes reinforces our view about their informational content.

To probe deeper on this issue, we also tried creating two dummy variables: one for changes between asset class and another for changes within one of the asset classes only—for example, investment grade to investment grade—and including the two interactions in the regression. In this case, we checked if there were significant differences when rating changes occur in either one of these categories vis-à-vis the omitted one. The results (not reported) were not significant either. Similarly, the results we obtained when we used the data from the other rating agencies were also weak. All in all, we did not find evidence that the effect of changes in ratings is different if they represent a shift in asset class.[35]

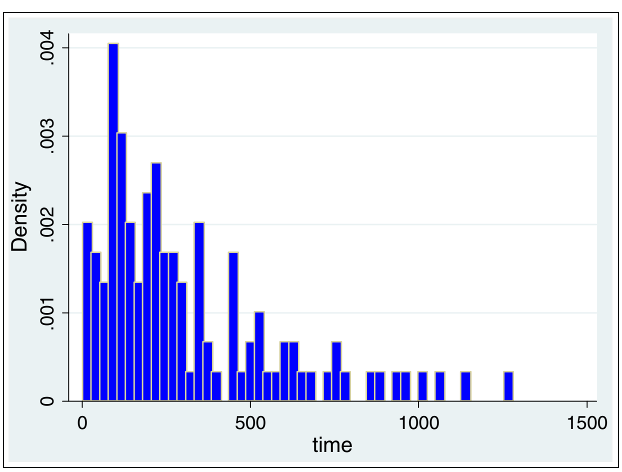
4.4. Change in outlook

Rating agencies publish outlooks as well as ratings. In particular, rating changes are typically preceded by changes in outlooks: either a positive outlook before an upgrade or a negative outlook before a downgrade. These outlook changes usually happen well in advance of the actual rating change.[36] Figure 6 is the plot of the distribution of the number of days between a change in the outlook and a change in the rating for the S&P sample. The distribution is highly skewed, and although the minimum number of days the outlook is changed before a rating change is 2, the mode is 98 and the mean number of days is a staggering 311. In other words, for most rating changes, the outlook was altered at least 1 year before.

In the next section, we exploit the discrepancies between outlook and rating changes to build a measure of the degree of anticipation of events. In this section, we take a different approach: We replace ratings with outlooks in the horse race regressions and check whether changes in the outlook also have explanatory power once we control for spreads.

We proceed as follows: we redefine *events* as episodes when there are changes in outlook and rearrange the data accordingly. *Outlook* is a step variable that can take only three values in every event window: –1 if there is a negative outlook, 0 if the outlook is stable, and +1 if the outlook is positive. Next, we rerun the baseline regressions using outlook as an RHS variable. The results for the baseline specification are reported in Table 8.

We found that changes in the outlook have very similar effects to changes in ratings. For the full sample (upgrades and downgrades together), we found that improvements in the outlook are associated with lower spread forward (although the coefficient estimate is not statistically significant), increases in the stock market, and appreciations of the exchange rate. The results are very similar when we split the sample between upgrades (i.e., favourable changes in the outlook) and downgrades, although the estimates tend to be more significant in the subsample of upgrades.



Time= Number of days between outlook change and the subsequent change in the rating.
Mean = 311 days.

Figure 6. Frequency distribution, Standard & Poor's ratings.

DOI: 10.1002/ijfe

Table 8. Benchmark Regressions Replacing Ratings With Outlooks

	Spread	Outlook	VIX
S&P upgrades & downgrades			
Spread $t+1$	0.856***	−0.0005	0.022**
	(0.009)	(0.002)	(0.009)
Stock market	−0.363***	0.007***	−0.009
	(0.011)	(0.002)	(0.010)
Exchange rate	0.083***	−0.005***	0.015***
	(0.006)	(0.0008)	(0.005)
*Δ*Spread?	−0.148***	−0.0009	0.0536***
	(0.009)	(0.002)	(0.0097)
S&P downgrades			
Spread $t+1$	0.876***	0.002	0.02
	(0.013)	(0.002)	(0.014)
Stock market	−0.400***	−0.001	−0.017
	(0.013)	(0.002)	(0.013)
Exchange rate	0.090***	−0.007***	0.020**
	(0.009)	(0.002)	(0.009)
*Δ*Spread?	−0.128***	0.002	0.059***
	(0.013)	(0.002)	(0.014)
S&P upgrades			
Spread $t+1$	0.808***	−0.003*	0.024*
	(0.016)	(0.001)	(0.012)
Stock market	−0.283***	(0.002)	0.0159***
	(0.020)	0.015***	(0.0023)
Exchange rate	0.054***	−0.002***	0.009**
	(0.004)	(0.001)	(0.003)
*Δ*Spread	−0.195***	−0.004**	0.045***
	(0.016)	(0.0018)	(0.013)

Notes: Regressions include event fixed effects as controls and estimated by OLS.
Every row is a different regression: either a different dependent variable or the error correction model.
Every column is the estimated coefficient for the corresponding RHS variable.
The standard errors are reported in parentheses below every point estimate.
S&P, Standard & Poor's; VIX, volatility index.

Interestingly, changes in the outlook also seem to be anticipated by the market. This is shown in the next two graphs (Figures 6 and 7). In the case of upgrades to the outlook, the markets seem to anticipate it only partially, as spreads continue to fall, and the stock market continues to increase, after the day of the change in the outlook. In particular, the cumulative change in these variables up to day 11 is roughly one half of the cumulative change over the entire window.

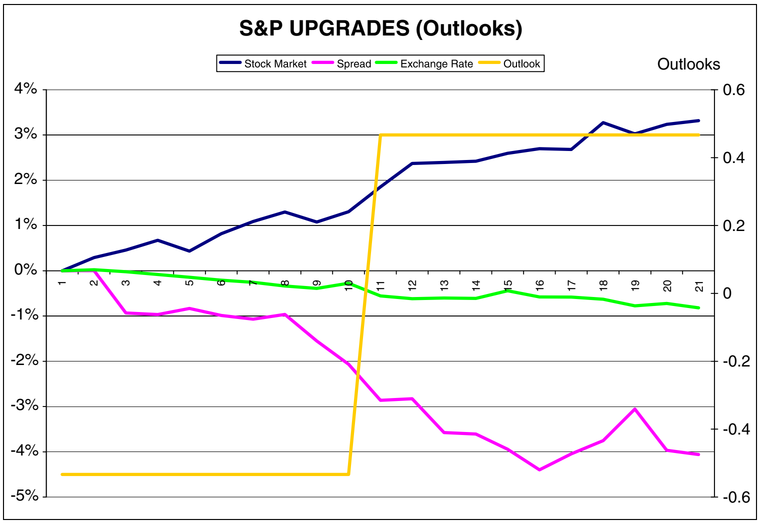
In the case of downgrades, the anticipation is even stronger, as we do not observe any discernible pattern in the spreads or the stock market after day 11 (Figure 8).

4.5. Anticipation and rating agency value added

We now return to the outlook changes as an indicator of the degree of anticipation of a rating change. As graphed earlier, the distribution of the number of days that outlooks change before ratings is skewed and has extremely wide dispersion. We suggest here that if the outlook change precedes the rating change by only a reasonably small number of days, then the rating change may not be fully anticipated. If the outlook change precedes the rating change by more than that, then it is likely that the rating change is fully anticipated. We also suggest that if the outlook change precedes the rating change by a very large number of days, then again the outlook change gives very little information on the rating change—or at least the timing thereof.

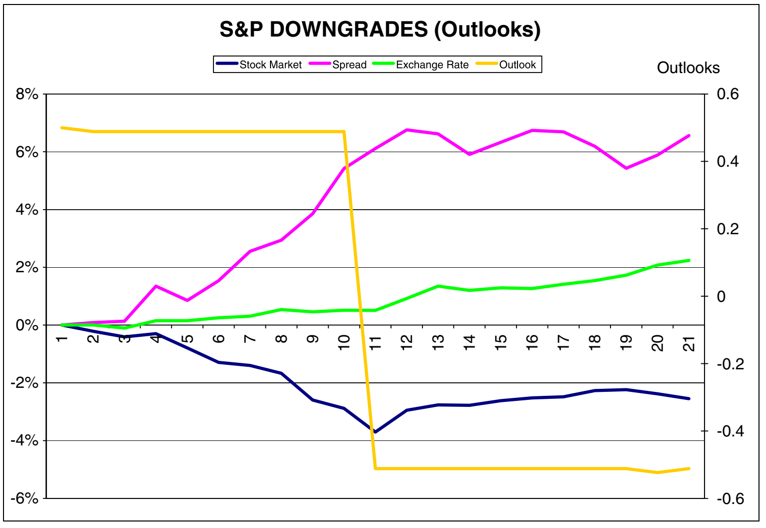
To motivate this analysis, in Figure 9, we plot the change in spreads around rating changes, dividing the sample into those where outlook changes occurred less than 60 days before the rating change, between 60 and 220 days

DOI: 10.1002/ijfe



Note: Outlook change occurs on day 11

Figure 7. Standard & Poor's upgrades (outlooks).



Note: Outlook change occurs on day 11.

Figure 8. Standard & Poor's downgrades (outlooks).

before the rating change, and more than 220 days. We normalized upgrades and downgrades to plot them on the same scale.[37] We also rescaled the series so that they are centred at zero on the day of the rating change. To allow for more variation in the graph, we plot the results for the case of the expanded window (i.e., 41 days per event). The graph is highly suggestive. As can be seen when the outlook change was between 60 and 220 days before the rating change, there appears to be no reduction in the spread after the change in rating, suggesting that the rating change was entirely anticipated. However, when the outlook change was more than 220 days before the rating change, and especially if it was less than 60 days before the change in rating, there was a reduction in spreads after the rating change. In these cases, it seems that the change in rating was only partially anticipated (as spreads decline before the event as well) and that the rating change appears to add information.

The graphs for the other macro variables used in this study show similar patterns for more and less anticipated events. These graphs are reported in Appendix A. A graph, however, does not constitute a statistical significance.

DOI: 10.1002/ijfe

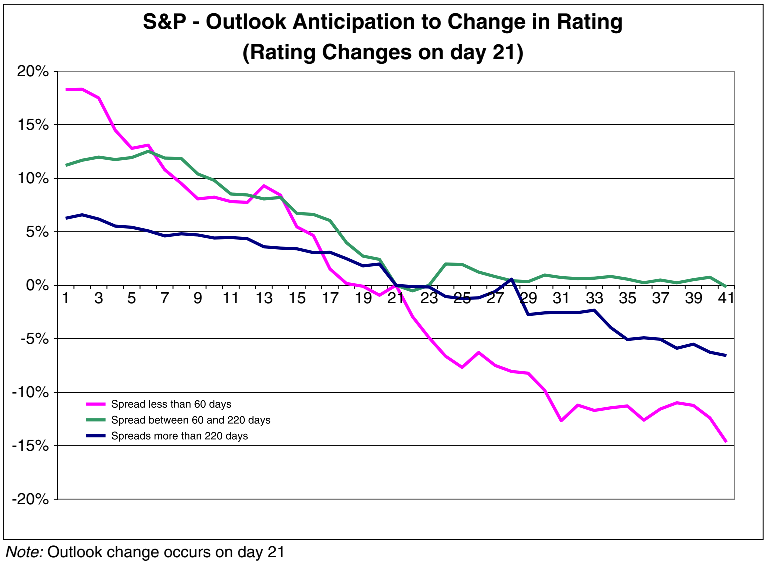


Figure 9. Standard & Poor's—outlook anticipation to change in rating (rating changes on day 21).

Table 9. Benchmark Regressions With Anticipation Effect: First Variant

	Spread	Rating	Rating * [Ln(number of days[a])]	VIX
S&P upgrades and downgrades				
Spread $t+1$	0.940***	0.008	−0.002**	0.021**
	(0.007)	(0.005)	(0.001)	(0.008)
Stock market	−0.178***	0.136***	−0.022***	−0.125***
	(0.008)	(0.007)	(0.001)	(0.010)
Exchange rate	0.086***	−0.137***	0.023***	0.064***
	(0.006)	(0.005)	(0.001)	(0.008)
*Δ*Spread?	−0.062***	0.008	−0.002**	0.036***
	(0.006)	(0.006)	(0.001)	(0.009)
S&P downgrades				
Spread $t+1$	0.932***	−0.018***	0.003**	0.032***
	(0.010)	(0.007)	(0.001)	(0.010)
Stock market	−0.516***	0.053***	−0.010***	−0.014
	(0.018)	(0.011)	(0.002)	(0.019)
Exchange rate	0.216***	−0.130***	0.022***	0.102***
	(0.016)	(0.011)	(0.002)	(0.017)
*Δ*Spread?	−0.072***	−0.019***	0.003**	0.045***
	(0.010)	(0.007)	(0.001)	(0.011)
S&P upgrades				
Spread $t+1$	0.930***	0.054***	−0.010***	0.008
	(0.009)	(0.013)	(0.002)	(0.013)
Stock market	−0.056***	0.044***	−0.007***	−0.119***
	(0.007)	(0.011)	(0.002)	(0.011)
Exchange rate	−0.001	0.011***	−0.002***	−0.007**
	(0.002)	(0.003)	(0.001)	(0.003)
*Δ*Spread	−0.071***	0.056***	−0.010***	0.027**
	(0.009)	(0.013)	(0.002)	(0.013)

Notes: Regressions include event fixed effects as controls and estimated by OLS.
Every row is a different regression: either a different dependent variable or the error correction model.
Every column is the estimated coefficient for the corresponding RHS variable.
The standard errors are reported in parentheses below every point estimate.
[a]Number of days between the day of the change in the outlook and the change in the rating.
S&P, Standard & Poor's; VIX, volatility index.

Table 10. Benchmark Regressions With Anticipation Effect: Second Variant

	Spread	Rating	Rating * T1[a]	VIX
S&P upgrades and downgrades				
Spread $t+1$	0.933***	−0.013***	0.008**	0.019**
	(0.007)	(0.003)	(0.003)	(0.009)
Stock market	−0.140***	0.103***	−0.106***	−0.120***
	(0.007)	(0.004)	(0.004)	(0.009)
Exchange rate	0.066***	−0.085***	0.085***	0.058***
	(0.006)	(0.003)	(0.003)	(0.007)
ΔSpread?	−0.069***	−0.013***	0.008**	0.035***
	(0.007)	(0.003)	(0.003)	(0.009)
S&P downgrade				
Spread $t+1$	0.924***	−0.015***	0.014***	0.030***
	(0.010)	(0.003)	(0.003)	(0.010)
Stock market	−0.437***	0.059***	−0.072***	−0.016
	(0.018)	(0.005)	(0.005)	(0.017)
Exchange rate	0.217***	−0.061***	0.057***	0.092***
	(0.017)	(0.005)	(0.005)	(0.017)
ΔSpread?	−0.079***	−0.015***	0.014***	0.043***
	(0.010)	(0.003)	(0.003)	(0.011)
S&P upgrade				
Spread $t+1$	0.934***	−0.008	0.001	0.007
	(0.009)	(0.013)	(0.013)	(0.013)
Stock market	−0.052***	0.002	0.002	−0.120***
	(0.007)	(0.010)	(0.011)	(0.011)
Exchange rate	0.0005	0.005*	−0.006**	−0.007**
	(0.002)	(0.003)	(0.003)	(0.003)
ΔSpread	−0.066***	−0.007	0.0001	0.025*
	(0.009)	(0.013)	(0.013)	(0.013)

Notes: Regressions include event fixed effects as controls and estimated by OLS.
Every row is a different regression: either a different dependent variable or the error correction model.
Every column is the estimated coefficient for the corresponding RHS variable.
The standard errors are reported in parentheses below every point estimate.
S&P, Standard & Poor's; VIX, volatility index.
[a]Interaction between the rating and a dummy that takes the value of 1 if the change in the outlook occurred more than 60 days before the change in the rating and 0 otherwise.

We therefore run regressions as earlier, but we added an additional term. Our first hypothesis was that the further the outlook change precedes the change in rating, then the more anticipated the rating change is. Nonetheless, we doubted that the relationship is linear. In particular, we suggested that the further the outlook change precedes the rating change, then the less the timing of the outlook change matters. We thus used the logarithm of the number of days the outlook was altered before the rating as an indicator of the potential lack of anticipation. We interacted this variable with the rating.

The second approach was to simply add a dummy interacted with the rating where the dummy takes the value of 1 if the outlook change precedes the rating change by more than a fixed number of days. We used the value of 60 days as this gave us a reasonable number of observations of rating changes that might be less than fully anticipated.[38] The results are given in Tables 9 and 10.

We found that both of these additional terms are significant and with the expected signs for virtually all of the cases detailed—across the different dependent variables, for upgrades and downgrades and for all changes. In particular, note that when the interaction term is significant, its sign is usually the opposite of the one of the coefficients for rating itself. This suggests that—whatever the impact of rating changes on these macro variables—the more anticipated the event, the smaller the effect. We conclude therefore that when the outlook change is closer to the rating change, then the rating change tends not to be fully anticipated and the rating change has a significant correlation with country variables. We suggest that this is further evidence that ratings matter.

DOI: 10.1002/ijfe

5. CONCLUSIONS

To facilitate the investment decisions of their clients, credit rating agencies monitor countries' fundamentals and assign individual (subjective) ratings and outlooks to sovereign debt. Given that they probably have more information than does the average investor, it is conceivable that these subjective ratings end up determining the level of sovereign spreads. What is not clear, and where there is considerable debate in the literature, is whether the opinion of rating agencies matter for the level of sovereign spreads even after controlling for countries' fundamentals and the current spread.

The objective of this article was to devise a test to evaluate the informational content that ratings have over what is already observed in bond spreads. We developed a simple Hausman specification test that is motivated in an EIV framework. The proposed test has the virtue of being robust to the most typical forms of misspecification, such as omitted variable bias, endogeneity, and in particular, the anticipation effect of rating changes that is observed in the data. The null hypothesis is that the spread is a sufficient statistic—that the rating does not add information beyond what the spread already captures. We applied this test to various alternative specifications and conclude that we can reject the null hypothesis. In other words, there seems to be some informational content in ratings that is not completely captured by spreads.

Next, we considered a type of horse race between ratings and spreads as to how well they are correlated to other macroeconomic variables using high-frequency data. We suggested that, given the possibility of full anticipation of rating changes, this is a better method for whether rating agencies add value. We found that they do, as the ratings typically explain part of the variation in the selected macro variables, even after controlling for the spread.

We also performed a battery of sensitivity tests, including using different windows for the regressions, using data from different rating agencies, using alternative estimation models, and also conducting tests on whether certain rating changes (i.e., changes in asset class or changes that are more anticipated) are more important than others. These additional tests reinforce the main conclusion that ratings add information.

NOTES

1. See Bolton *et al*. (2009), Ashcraft and Schuermann (2008), and Becker and Milbourn (2008).
2. See Portes (2008) for a discussion on a number of problems with the credit ratings agencies. For a comprehensive review of the credit rating agencies business, see Levich *et al*. (2002) and Cantor (2004).
3. See, for example, Ferri, Liu, and Majnoni (2001).
4. See, for example, Kaminsky and Schmukler (2002).
5. See Cantor (2004) for a discussion.
6. In the case of the technique used by Eichengreen and Mody (1998) and Dell'Ariccia *et al*. (2006), it is a heroic assumption that the error of the ratings equation represents the rating agencies' opinion and not that this equation is simply misspecified. In the system approach favoured by Powell and Martinez (2007), a different but also heroic assumption is needed to identify the system. In the approach employing rating agencies' differences of opinions, one rating agency may follow a spread change rather than actually affect the spread.
7. See, for example, Reisen and von Maltzan (1999) and Brooks *et al*. (2004) for applications of the event study's methodology to the sovereign ratings literature. Behr and Güttler (2008) applied the methodology to study unsolicited corporate credit ratings.
8. See, for example, Ferri, Liu, and Stiglitz (2001) and Kaminsky and Schmukler (2002) for a discussion on pro-cyclicality of sovereign ratings. Also, Reinhart (2002) found that sovereign rating changes are "lagging" (rather than "*leading*") indicators of currency crises, although they do better predicting defaults.
9. Reactivity is oftentimes taken as evidence that sovereign ratings are uninformative. For example, Ferri, Liu, and Majnoni (2001) argued against the use of banks' ratings for capital asset requirements in nonindustrialized countries (as proposed by Basel) on the basis that banks' ratings in these countries are not generalized, and sovereign ratings tend to be uninformative because they show pro-cyclical swings.
10. An example is stock market returns, which are conceivably also affected by the same macroeconomic fundamental.
11. Possible sources of misspecification include omitted variable bias, endogeneity, and in particular, the anticipation effect of rating changes that is observed in the data.
12. At the same time, the aforementioned misspecification problems are conceivably more important in low-frequency datasets (monthly or above).
13. Even if rating changes are largely anticipated by the market by the time they are announced, there may be still some impact from rating news due to the fact that many institutional investors face limits in the amount of low-grade assets they can invest in. However, this as well as other technical reasons that may drive asset prices after an announcement are not necessarily related to fundamental information. More on this in the Asset Class Shift section.
14. We focused the analysis on sovereign bonds spreads, which are computed as the difference of the yield-to-maturity of a bond minus the yield-to-maturity of a comparable riskless bond (i.e., US Treasuries). These are the most widely used proxies of risk by market observers.
15. In other words, when the rating is not changing, it is possible to argue that the default probability is changing little as well, and therefore, no change in ratings is imperfectly measuring small changes in fundamentals. However, when the rating is indeed changing, we expect in those windows for the fundamental to cross some threshold, and therefore, the increase in the rating indeed reflects an improvement in the fundamental.
16. In fact, anticipation of improvements in fundamentals implies that$\text{cov}(x_t, \mu_t) \neq 0$.
17. The test described here has discussed mostly the linear case, but the nonlinear case is exactly the same. For instance, take a nonlinear model and linearize it. The residuals in that model will be correlated with the unobservable fundamental exactly in the way we discussed cases 2 and 3.
18. The list of countries is in Appendix A.
19. Some countries have multiple stock market indices. The list of indices used in this study is in Appendix A.
20. The VIX is constructed using the implied volatilities of a wide range of S&P 500 index options. This volatility is meant to be forward-looking and is calculated from both calls and puts.
21. One contribution of this article is assembling a consistent dataset with precise dates for rating and outlook changes that have been cross-checked with industry sources.

DOI: 10.1002/ijfe

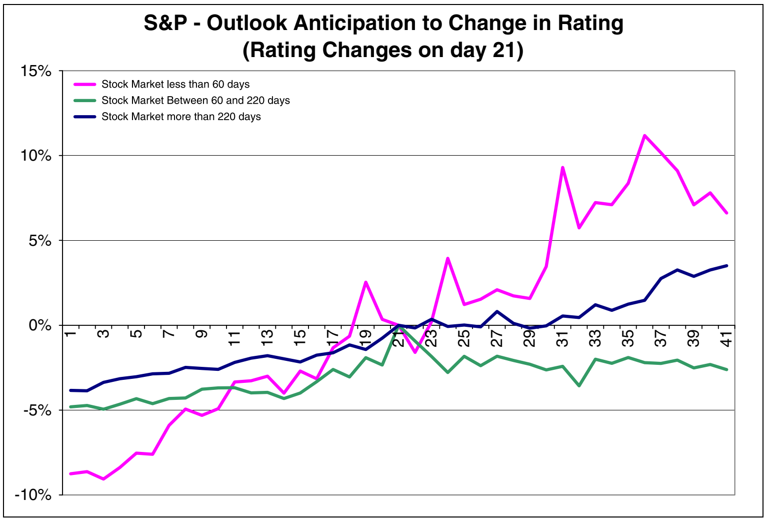
22. Alternatively, for robustness checks purposes, we defined 41-day and 11-day windows around the event.
23. In the cases where there are multiple rating changes within the same event window, we treated each rating change as an independent event. We alternatively dropped these events from the sample for robustness checks purposes, but the results remained unchanged.
24. We used 141 events, rather than the total of 145 events in Table 2, because there were four events that happened within the window of a previous event. Thus, we dropped these to avoid double counting when comparing with the other rating agencies. We did the same for Moody's and Fitch, where we dropped four and five events, respectively.
25. For this purpose, we stacked all the events (i.e., both upgrades and downgrades) together and run the regressions for the full sample of S&P events.
26. We omitted to report the coefficient for the VIX in the standard OLS and IV regressions and the rest of the coefficients in the error correction model, as they are not essential for explaining the test we performed in this section.
27. This is not technically correct, as the Hausman procedure uses all the estimated coefficients, and their variance matrix, to perform the test.
28. In the case of S&P, these are four events that happened within the window of a previous event.
29. We also run the same tests including a time trend in the regressions for each event window. We found somewhat lower rejection rates, although in most cases, they remained over 50%. It is hardly surprising that the rejection rates fell when we included a time trend, as many events were anticipated (more on this later) and the effect of the anticipation may be precisely a trend over the event window for the macro variables. Thus, we found it reassuring that we still found a high number of rejections even when we included a trend.
30. We rejected three of four times when the window is 10 days around the event, and only two times when the window is expanded.
31. We also explored an additional market outcome variable: nominal exchange rates.
32. *: significant at 10 percent, **: significant at 5%, and ***: significant at 1%.
33. We omitted the coefficient estimates for ϕ, as they were not essential.
34. We thank an anonymous referee for suggesting these examples to us.
35. Despite this, it is possible that we found no effect because there were relatively few events that represent changes in asset class in our sample. For example, for the S&P sample, only nine of the 145 events are changes between asset classes.
36. This is not surprising as normally the outlook change means that the rating upgrade/downgrade will be made between 6 and 12 months after. For example, according to S&P, a rating outlook assesses the potential direction of a rating change over the intermediate term (typically 6 months to 2 years). In determining a rating outlook, consideration is given to any changes in the economic and/or fundamental business conditions. An outlook is not necessarily a precursor of a rating change.
37. Thus, all the observations corresponding to downgrades are multiplied by -1 so they can be mapped into the same scale as upgrades.
38. Alternatively, we used a dummy variable that takes the value of one if the outlook change precedes the rating change by more than 60 days, but also less than 220 days. The results are unchanged.

ACKNOWLEDGEMENT

We thank Jeromin Zettelmeyer, John Chambers, Eduardo Fernandez-Arias, and the seminar participants at the XXVII Meeting of the Latin American Network of Central Banks and Finance Ministries for very useful comments and Francisco Arizala and Oscar Becerra for superb research assistance. All remaining errors are our own. This article represents the views of the authors and does not necessarily reflect the views of any institution including the Inter-American Development Bank, its executive directors, or the countries they represent.

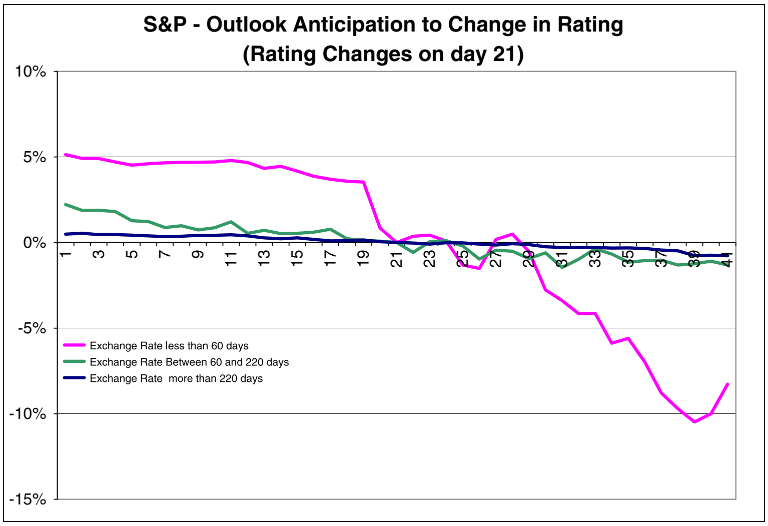
APPENDIX A

Less anticipated events have a bigger impact on macro variables ex post than do events that are more anticipated. In the text, we present the case of changes in the spread. Here, we report the cases of the stock market and the nominal exchange rate.



Note: Outlook change occurs on day 21.

Figure 10. S&P—outlook anticipation to change in rating (rating changes on day 21).



Note: Outlook change occurs on day 21.

Figure 11. S&P—outlook anticipation to change in rating (rating changes on day 21).Standard & Poor's downgrades (outlooks).

Table 11. Countries and Stock Market Indices Used

Country	Stock market index
Argentina	Argentina Merval Index
Bulgaria	Sofix Index
Brazil	Brazil Bovespa
Chile	Chile Stock Market General
China	Shanghai Stock Exchange Composite Index
Colombia	Colombia General Index - Bogota Stock Market Index
Dominican Republic	Not available in Bloomberg
Ecuador	Ecuador Guayaquil Stock Exchange Bolsa
Egypt	Egypt Hermes Index
Croatia	Croatia Zagreb Crobex
Hungary	Budapest Stock Exchange Index
Indonesia	Jakarta Composite Index
Korea	Kospi Index
Lebanon	Blom Stock Index
Morocco	Madex Free Float Index
Mexico	Mexico Bolsa Index
Malaysia	Kuala Lumpur Composite Index
Nigeria	Nigeria Stock Exchange
Pakistan	Karachi All Share Index
Panama	Panama Stock Exchange General
Peru	Peru Lima General Index
Philippines	Philippine Stock Exchange Index
Poland	Wse Wig Index
Russia	Russia Stock Market Index
El Salvador	Not available in Bloomberg
Thailand	Stock Exchange of Thailand Index
Tunisia	Tunise Stock Exchange Tunindex
Turkey	Ise Industrials
Ukraine	Ukraine Pfts Index
Uruguay	Not available in Bloomberg
Venezuela	Venezuela Stock Market Index
South Africa	Africa All Share Index

DOI: 10.1002/ijfe

REFERENCES

Afonso A, Gomes P, Rother P. 2007. What 'Hides' behind Sovereign Debt Ratings? Working Paper 711. Frankfurt, Germany: European Central Bank.

Ashcraft A, Schuermann T. 2008. Understanding the Securitization of Subprime Mortgage Credit, mimeo, Federal Reserve Bank of New York.

Becker B, Milbourn T. 2008. Reputation and competition: evidence from the credit rating industry. Mimeo, Havard Business School.

Behr P, Guttler A. 2008. The informational content of unsolicited ratings. *Journal of Banking & Finance* **32**(4): 587–599, April.

Bolton P, Freixas X, Shapiro JD. 2009. The Credit Ratings Game. NBER Working Paper No. w14712.

Brooks R, Faff R, Hillier D, Hillier J. 2004. The national market impact of sovereign rating changes. *Journal of Banking & Finance* **28**(1): 233–250, January.

Campbell JY, Lo AW, Mckinlay AC. 1997. *The Econometrics of Financial Markets*. Princeton University Press: Princeton, United States.

Cantor R. 2004. An introduction to recent research on credit ratings, Recent Research on Credit Ratings. *Journal of Banking & Finance* **28**(11): 2565–2573, November.

Cantor R, Packer F. 1996. Determinants and Impact of Sovereign Credit Ratings. *Economic Policy Review* **2**(2): 37–53. New York, United States: Federal Reserve Bank of New York.

Dell'Ariccia G, Schabel I, Zettelmeyer J. 2006. How Do Official Bailouts Affect the Risk of Investing in Emerging Markets? *Journal of Money, Credit, and Banking* **38**(7): 1689–1714.

Eichengreen B, Mody A. 1998. What Explains Changing Spreads on Emerging-Market Debt: Fundamentals Or Market Sentiment? NBER Working Paper 6408. Cambridge, United States: National Bureau of Economic Research.

Ferreira M, Gama P. 2007. Does sovereign debt ratings news spill over to international stock markets? *Journal of Banking & Finance* **31**(10):3162–3182, October.

Ferri G, Liu L-G, Majnoni G. 2001. The role of rating agency assessments in less developed countries: Impact of the proposed Basel guidelines. *Journal of Banking & Finance* **25**: 115–148.

Ferri G, Liu L-G, Stiglitz J. 2001. The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis. *Economic Notes* **28**(3): 335–355.

González Rozada M, Levy Yeyati E. 2008. Global Factors and Emerging Market Spreads. *Economic Journal, Royal Economic Society* **118**(533): 1917–1936, November.

Hausman J. 1978. Specification Tests in Econometrics. *Econometrica* **46**(6): 1251–1271.

Kaminsky G, Schmukler S. 2002. Emerging Market Instability: Do Sovereign Ratings Affect Country Risk and Stock Returns? *The World Bank Economic Review* **16**(2): 171–195.

Levich R, Majnoni G, Reinhart C (eds). 2002. Ratings, Rating Agencies and the Global Financial System, Kluwer.

Portes R. 2008. Ratings Agency Reform. Voxeu.org, 22 January 2008.

Powell A, Martínez J. 2007. On Emerging Economy Sovereign Spreads and Ratings. Research Department Working Paper 629. Washington, DC, United States: Inter-American Development Bank.

Reinhart CM.2002. Default, Currency Crises, and Sovereign Credit Ratings, *World Bank Economic Review* **16**(2): 151–170.

DOI: 10.1002/ijfe