Sovereign Debt Rating Changes, Corruption and the Stock Market*

Alexander Michaelides

University of Cyprus, CEPR and NETSPAR†

Andreas Milidonis‡ George Nishiotis§

University of Cyprus University of Cyprus

Panayiotis Papakyriacou¶

University of Cyprus

August 2012

*We thank the University of Cyprus for a research grant in pursuing this research paper. We also thank George Constantinides, Bernardo Guimaraes, Ravi Jagannathan, Peter Hoffmann (discussant), Robert Korajczyk, Nadia Massoud, Lenos Trigeorgis and seminar participants at Manchester Business School, SED and EFMA meetings for useful comments.

†Department of Public and Business Administration, University of Cyprus, Lefkosia, 1678, Cyprus. Email: a.michaelides@ucy.ac.cy.

‡Department of Public and Business Administration, University of Cyprus, Lefkosia, 1678, Cyprus. Email: milidonis.andreas@ucy.ac.cy.

§Department of Public and Business Administration, University of Cyprus, Lefkosia, 1678, Cyprus. Email: nishioti@ucy.ac.cy.

¶Department of Public and Business Administration, University of Cyprus, Lefkosia, 1678,

Abstract

We use an event-study methodology to analyze the effect of sovereign debt rating changes on daily stock market returns around the world. We find evidence that the stock market moves before the public announcement of a sovereign rating downgrade, resulting in a statistically and economically significant abnormal market reaction prior to the event. Using instrumental variable techniques we argue that these findings are more pronounced in non-developed markets, in countries with civil (relative to common) legal systems, with lower measures of law and order institutional quality, and with higher measures of corruption.

JEL Classification: G14, G15, G24.

Key Words: sovereign ratings, event studies, international finance, institutional quality.

Cyprus. Email: papakyriacou.panayiotis@ucy.ac.cy.

1 Introduction

Rating agencies, their actions and the effects these actions have on yields, returns and government policies have become an important topic of discussion among market participants and regulators in the last twenty years. The global financial crisis of 2008 and the Euro area sovereign debt crisis of 2011 have only served to heighten the intensity of the discussion. As an illustration of the potential effects that rating changes might have, Figure 1 plots the Dow Jones cumulative raw daily returns twenty trading days before, and twenty trading days after, the August 5th 2011 Standard and Poor's (S&P) U.S. government debt downgrade. The Figure provides motivation for two hypotheses. First, the drop in the stock market starts 10 days before the actual downgrade. Second, the change in stock market returns is economically significant.

Our purpose is to empirically investigate whether the case for the U.S. is an exception, or whether it is a more widespread phenomenon around the world. It should be noted that anecdotal evidence in the popular press (Wall Street Journal, September 20th, 2011) indicates that information about the imminent U.S. downgrade leaked to the market before the actual announcement. Moreover, the U.S. Securities and Exchange Commission (SEC) launched an investigation regarding the potential leakage of information before the downgrade.[1] If rating announcements

[1] In the SEC's September 30th 2011 summary report after examining ten rating agencies under its oversight, the SEC identified a number of concerns: "These concerns included apparent failures in some instances to follow ratings methodologies and procedures, to make timely and accurate disclosures, to establish effective internal control structures for the rating process and to adequately manage conflicts of interest."

in a tightly regulated/monitored capital market such as the American one generate "concerns" (according to the SEC), then rating announcements might generate even more "concerns" in other less-regulated/monitored capital markets.[2] Motivated by these events we empirically investigate abnormal stock market reactions around sovereign debt rating announcements throughout the world. We then use institutional quality variables to better understand the cross-sectional variation in the documented abnormal returns.

To achieve this goal, we employ an event-study methodology to examine local, daily, stock market reactions to changes in sovereign debt ratings, outlooks, and watch list inclusions. We are interested in possible abnormal local stock market returns before, at and after these public announcements. To test our hypothesis we collect the universe of these changes by Fitch, Moody's and Standard & Poor's (S&P) from February 1988 to December 2011. We focus on the three largest credit-rating agencies, both because these agencies hold a substantial fraction of market power in the industry, but also because they have recently come under intense

[2]The experience of Cyprus (downgraded by Fitch on Wednesday August 10th 2011) serves as additional motivation from a less developed capital market. On Thursday August 4th 2011, the Cyprus stock market fell 3.6% and the next trading day an additional 4.1%. On Saturday the Head of the Cyprus Securities and Exchange Commission made a public plea on national television that anyone having information that should be in the public domain must immediately disclose it. The following trading day in Cyprus (Monday, August 8th) coincided with a public holiday in London (Fitch covers Cyprus out of its London office), and the Cypriot stock market fell another 5.69%. Fitch downgraded Cypriot sovereign debt by two notches on Wednesday August 10th 2011, which was associated with a stock market increase of 0.6% (and a 0% change on August 9th). Anecdotal evidence in Cyprus indicates that the imminent downgrade was leaked during the consultation process between the rating agency and local auhorities and before its public announcement.

scrutiny in the market for corporate bond ratings (SEC, 2003; Beaver et al., 2006; SEC, 2008; Cheng and Neamtiu, 2009).

For each rated country, we collect daily local stock market index return data. Using a short-horizon event study analysis, which is "relatively straightforward and trouble-free" according to the recent excellent survey by Kothari and Warner (2007) on the econometrics of event studies, we examine the behavior of local stock market returns twenty days before, and twenty days after, each announcement.

To mitigate the problem arising from simultaneous rating actions across agencies, we construct our preferred definition of an event that takes into account the rating agency that moves first in making a public announcement. Intuitively, we view the first change as more important for the stock market than changes that might occur soon after the initial move. We therefore use a "first mover" filter to construct our baseline case: the event stays in the sample if it is not preceded by a change in rating by the same, or another, rating agency in the twenty trading days prior to the event. To address the criticism that even with the high-frequency (daily) data we use, the rating change might be anticipated and therefore may not be a proper shock in the market, we create two additional samples that include changes in outlooks and watchlist inclusions.

For all three samples we find statistically and economically significant movements in local stock market returns for the periods before, at, and after the actual announcements of sovereign debt rating downgrades[3]. The pre-announcement neg-

[3] The economic significance of the market reaction to upgrades appears to be significantly muted relative to the market reaction to downgrades consistent with findings in the corporate bond ratings literature (Holthausen and Leftwich, 1986; Hand et al., 1992; Ederington and Goh, 1998).

ative abnormal returns are sizable and strongly statistically significant, followed by significant announcement effects. The negative abnormal returns appear to be partially reversed in the post-announcement period, generating a cumulative abnormal return graph with a near "V" shape around the event. The results are robust to estimators that take into account event-induced variance as in Boehmer (1991), to changing the event-window size and to controlling for cross-sectional correlation of abnormal returns as outlined in Kolari and Pynnonen (2010).[4]

We then identify country-level institutional characteristics that are more likely to be associated with these abnormal stock returns. We proceed in two steps. In the first step we ignore endogeneity or error-in-variables issues and try to determine whether there are observable characteristics across countries that correlate strongly with these results. In the second step we build a causal story linking country characteristics and abnormal stock return behavior. Starting with the first step, we conduct event studies separately for developed versus non-developed economies. We find that the results are largely driven by events in non-developed countries. Moreover, countries with civil (relative to common) law are more likely to generate this abnormal stock return pattern, consistent with the conclusions in the survey by La Porta et al. (2008). We also sort countries according to law and order quality, different measures of corruption and an investor protection index. Our

This is also consistent with evidence in the accounting literature of asymmetric market reaction to surprise negative earnings relative to positive earnings announcements (Skinner, 1994; Soffer, Thiagarajan, and Walther, 2000; Hutton, Miller, and Skinner, 2003; Anilowski, Feng, and Skinner, 2007; Kothari, Shu, and Wysocki, 2009).

[4] Excluding periods of high volatility in financial markets (for instance, the post-2008 global financial crisis period) does not change our results either.

results illustrate that the quality of the institutional framework correlates with stock market abnormal returns around sovereign debt downgrades.

In the second step we move away from correlations to build a causal story. A positive correlation between stock returns and institutional quality by no means implies causation. Error-in-variables problems can also affect our conclusions. For instance, using a proxy variable to measure institutional quality probably means that this proxy measures underlying quality with error. This is a classic errors-in-variables problem, generating biased estimates. To resolve endogeneity or error-in-variables problems, we identify appropriate instrumental variables for the variables proxying for corruption and/or institutional quality.

The instrumental variables we use are combinations of recently used variables in the literature proxying for institutional quality. Specifically, we use the origin of the local legal system (La Porta et al., 2008), ethnic and religious fractionalization (Alesina et. al., 2003) and a zero-one indicator for the country being landlocked (Easterly and Levine, 2003). The chosen instruments pass weak instrumental variable tests and the final regressions pass the over-identification Sargan/Hansen test statistic. Our results provide evidence for a causal relation between institutional quality and stock market reaction before a downgrade announcement: the coefficients in the regression are statistically significant and have the expected sign.

These causal estimates are also economically significant. Less developed countries generate cumulative average abnormal returns ($CAARs$) of about 2.7% (p-value of 0.010) lower than those in developed countries in the pre-announcement period (from five days before to three days before the announcement; $CAAR[-5,-3]$).

Moreover, a one-standard deviation decrease in the Transparency International corruption index score gives a 1.3% (p-value of 0.001) decrease in $CAAR[-5,-3]$. Similarly, the Law & Order variable indicates an overall decrease in the $CAAR$ of 1.0% (p-value of 0.013) when the score decreases by one standard deviation, a result which is also obtained for the a one-standard deviation decrease in the Investor Protection Index score (a 1.0% decrease in the $CAAR$ with a p-value of 0.070).

One possible explanation for these findings is that information leaks to the market before the public announcement. This might be coming either from the rating agency[5] or local government officials[6] during the consultation process. Another explanation is that the market anticipates the event through other public information sources. Our cross-sectional results seem more consistent with the leakage of information hypothesis because it is harder to justify that markets with lower institutional quality are better at anticipating credit rating actions.

We want to emphasize, however, that other possible explanations exist. For instance, another possible interpretation might arise from investor difficulty in actually assessing the consequences of imminent sovereign debt downgrades, even if they are expected. This difficulty might be stronger in a country with non-transparent institutions or with smaller stock markets that are less widely followed by analysts. This interpretation could also justify a mild post-event correction as less-informed investors have more time analyzing the news after the event.

[5] In July 2011 and January 2012 Italian prosecutors raided Moody's and Fitch offices respectively and accused the two rating agencies that there was leakage to the market before the sovereign debt downgrades were made publicly available (Rossi, 2012).

[6] Even in well regulated markets like the U.S. the idea of "soft corruption" is prevalent in the press as illustrated by the November 2011 CBS 60 Minutes "Insiders" Report (Kroft, 2012).

Why should we care about these empirical findings? There has been increasing regulatory activity related to rating agencies (2002 Sarbanes-Oxley Act section 702 (b); 2006 Credit Rating Agency Duopoly Relief Act). The abnormal stock return pattern and the characteristics of the countries where this pattern is more pronounced raise concerns about capital market regulation around the public announcement of downgrades. To the extent that the leakage interpretation is correct, our results indicate that capital market regulators need to worry about insider activity around rating announcements. Regulators in countries with lower indicators of institutional quality seem to be the ones that should be worrying the most about information leakage. Moreover, in the emerging field of household finance, a recent explanation of limited stock market participation is the low level of trust in the stock market (Guiso, Sapienza and Zingales (2008)). To the extent that our results are rationalized by informational advantages from insiders, trust in the stock market should be lower in countries with weaker institutions, rationalizing limited stock market participation in such environments.

The literature on the effects of sovereign debt downgrades on stock markets is relatively nascent and recent. Kaminsky and Schmukler (2002) analyze the issue in a similar fashion but we differ by having a more extended data set (both in terms of country and time coverage) and explicitly making the connection between the potential for leakage of information ahead of a rating announcement and the quality of institutions. Brooks et al. (2004) also find a negative effect of rating downgrades on stock returns, but we differ by emphasizing that in our empirical results the effect seems to show up earlier than the actual announcement. Martell

(2005) and Hill and Faff (2010) also find evidence for movements in stock returns before ratings announcements. We differ from both papers primarily because we document a causal link between sovereign institutional quality and stock market reaction before ratings downgrades.

The remainder of the paper is organized as follows. In Section 2, we present descriptive statistics on the assembled data set. In section 3 we present our empirical results and perform robustness checks. In Section 4 we examine how our results differ across institutional regimes. Section 5 concludes.

2 Data and Descriptive Statistics

We use historical sovereign ratings data from the websites of Fitch, Moody's and S&P. S&P and Fitch publish letter ratings corresponding to the same scale. Moody's uses letter grades that are slightly different. Following prior articles in the bond rating literature (Johnson, 2004 and Beaver et al. 2006 among others), we transform letter grades by S&P and Fitch (Moody's) as follows: "AAA" (Aaa) = 1; "AA+" (Aa1) = 2; "AA" (Aa2) = 3; "AA-" (Aa3) = 4; "A+" (A1) = 5; "A" (A2) = 6; "A-" (A3) = 7; "BBB+" (Baa1) = 8; "BBB" (Baa2) = 9; "BBB-" (Baa3) = 10; "BB+" (Ba1) = 11; "BB" (Ba2) = 12; "BB-" (Ba3) = 13; "B+" (B1) = 14; "B" (B2) = 15; "B-" (B3) = 16; "CCC+" (Caa1) = 17; "CCC" (Caa2) = 18; "CCC-" (Caa3) = 19; "CC" (Ca) = 20; "C" (C) = 21. In the case of default, restricted default or other action associated with a sovereign in financial distress (i.e. ratings of D, RD, SD e.t.c.) we assign the number 22.

We identify changes in (local and foreign currency) ratings and outlooks by comparing successive letter grades for each country. The samples for Fitch, Moody's and S&P begin in 1994, 1986 and 1983, with 318 (201), 336 (185) and 434 (350) changes in ratings (outlooks and watchlist inclusions)[7], respectively. To test market reactions around the announcement of ratings changes, we match the union of these ratings changes with the panel of daily prices for each country's local currency stock market index and also the World MSCI index from Datastream. Our analysis begins with the earliest date of the world MSCI index (01/01/1988) and ends on 31/12/2011. After removing duplicate observations (i.e. changes in ratings happening on the same day) and observations with no index return data, the sample of changes in ratings comprises 874 observations (456 upgrades and 418 downgrades) for 65 countries, and the sample for outlook changes is made up of 600 observations (334 positive and 266 negative).

Figure 2, Panel A reports the total number of changes in sovereign debt ratings for the three largest agencies (Fitch, Moody's and S&P) from February 1989 to December 2011. Downgrades seem to be more concentrated than upgrades and tend to happen in periods of recession or global financial turmoil. The 1997 East Asian crisis, the 1998 Russian crisis, the short 2001 U.S. recession, and the ongoing world financial crisis since 2008 figure prominently in the number of downgrades in Panel A. Our analysis is done on events for each agency separately but also after considering all events together from all agencies (Figure 2 Panel A).

Multiple ratings for the same sovereign around the same time are very unlikely to have the same impact on the local stock market returns because they are not

[7] For the remainder of the paper references to outlooks imply also the use of watchlist inclusions in a sample.

independent of each other as they are based, to a large degree, on analyzing the same information about the sovereign. To mitigate the problem arising from such cross-correlation across rating actions, we construct our preferred definition of an event that takes into account the rating agency that moves first in making a public announcement. Intuitively, we expect the second change for the same sovereign to have a less pronounced effect than the original change. To prevent this synchronization from contaminating our analysis, we use a "first mover" filter to construct our baseline case.[8]

We implement this idea as follows. We keep ratings for a sovereign that are not preceded by other changes in ratings of the same, or other, rating agency in the previous twenty trading days. We call this, the “ratings FMR” sample. This generates 293 downgrade events (from an initial of 418) and 400 upgrade events (from an initial of 456) from 65 countries, which represent the core events that make up the baseline specification (“ratings FMR” sample). Figure 2, Panel B plots the time series distribution of the “ratings FMR” sample. Comparing Panel A and B we observe that the definition does matter since a substantial number of individual rating changes from Panel A are removed, especially during periods of heightened financial turmoil and/or recession (1997, 1998, 2001, 2008 and 2011). In our analysis we conduct a sensitivity analysis based on the number of days used in our definition of “first-mover”, by altering the number of days required to have no other change in rating before the observation chosen.[9]

[8] This methodology is also followed by Martell (2005).

[9] Specifically, we change the twenty-day window to thirty, ten and zero and our results are unchanged.

Figure 3 plots the grade-level distribution of the sovereign debt "ratings FMR" sample after the rating change (downgrade or upgrade). We report the frequency distribution separately for each rating agency and also for our "ratings FMR" definition. The graphs illustrate that rating changes are not concentrated on one specific grade-level. Instead, changes generate a spectrum of resulting grades, both for upgrades and downgrades.

The next step before preparing the data for the event study analysis is requiring each rating change to have at least 60 daily observations in the estimation period (from day -270 to day -21), consistent with Low (2009). This filter removes two additional observations from our dataset. The final data step deals with potential problems of illiquidity due to including frontier markets in the analysis. We follow Bekaert et al. (2007) in using the percentage of zero returns as a proxy for illiquidity. Specifically, we examine the number of days of zero returns that exist in each country's testing period (day -21 to +21) and exclude observations that have more than ten days of zero returns. The "ratings FMR" sample has 376 upgrades and 271 downgrades. We confirm that our results are not affected by this filter, by re-running our analysis without it.

Outlook changes might also affect the way information is transmitted to stock market valuations. For robustness purposes we therefore create two additional samples by also considering changes in outlooks. The first robustness sample treats changes in ratings and outlooks symmetrically. To construct it we create the union of changes in ratings and outlooks and apply the "first mover" filter. The event stays in the sample if it is not preceded by a change in rating or outlook by the same, or another, rating agency in the twenty trading days prior to the event (FMRO filter).[10] We call this sample "ratings and outlooks FMRO". Our second robustness sample comprises changes in ratings that are not preceded by changes in ratings and outlooks in the previous twenty days ("ratings FMRO" sample).

In section 4, our analysis focuses on the institutional characteristics in rated countries. These are: (a) the legal system (common vs. civil); (b) the country's classification by the World Bank (developed vs. emerging and frontier); (c) the level of corruption from the Political Risk Services Group (high vs. low relative to the median score); (d) the Law & Order level also from the Political Risk Services Group (high vs. low relative to the median score); (e) the Corruption Perceptions Index (CPI) from Transparency International (high vs. low relative to the median score); and (f) the strength of investor protection from Djankov et al. (2008) (high vs. low relative to the median score). Definitions and the range of values for each institutional characteristic are given in the appendix.

For each of these institutional characteristics, we identify appropriate instrumental variables to address potential endogeneity or errors in variables problems. We select from the following list: the type of legal system (common vs. civil law, (La Porta et al., 1998)), the ethnicity and religion fractionalization measures developed by Alesina et al. (2003), and a landlocked indicator (Easterly and Levine, 2003). In section 4.2 we describe the methodology and in the data appendix we provide further details about these variables.

[10] FMRO stands for First Mover using Ratings and Outlooks.

3 Empirical Results

3.1 Econometric Analysis

We use a short-horizon event-study methodology using daily return data on the stock market index of all countries in our sample to capture the dynamic effects of ratings changes on stock returns. The estimated abnormal returns around a rating change (downgrade or upgrade) can provide evidence on the effect of the rating agency's action on the local stock market. They can also help us assess whether the impact of rating changes on the stock market is merely transient or more sustained.

We use the world CAPM to calculate abnormal returns as follows.[11] For every country, the following time series regression is estimated using data in the window $[-270, -21]$ trading days relative to the event date:

$$R_{it} = \alpha_i + \beta_i R_{Wt} + \varepsilon_{it} \tag{1}$$

where R_{it} is the country i's MSCI index return, and R_{Wt} is the world MSCI index return . We then calculate abnormal returns (AR) from the residuals for the window $[t_1, t_2] = [-20, +20]$ around the event:

$$AR_{it} = R_{it} - \widehat{\alpha}_i - \widehat{\beta}_i R_{Wt} \tag{2}$$

Finally, we obtain cumulative abnormal returns (CARs) for different sub-periods

[11] We also conduct our analysis using raw returns and these are reported later in the paper.

$[t_1, t_2]$ by adding up the corresponding AR's over the event study window

$$CAR_i[t_1, t_2] = AR_{it_1} + ... + AR_{it2} \tag{3}$$

We use different estimators to test for the statistical significance of average abnormal returns and average cumulative abnormal returns (and we do this separately for upgrades and downgrades). We first form a test using the cross-sectional variation of abnormal returns in the event window under the assumption that AR_{it} are independently and identically distributed following a normal distribution with mean zero (under the null) and variance σ^2 (see Charest (1978) and Penman (1982)). Using s_t as an estimator for σ, we can then define the test statistic based on the average abnormal return (AAR_t)

$$Z = \sqrt{N}\ \frac{AAR_t}{s_t} \sim t_{N-1} \tag{4}$$

where N is the number of events and

$$AAR_t = \frac{1}{N}\sum_{i=1}^{N} AR_{it} \tag{5}$$

$$s_t = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(AR_{it} - AAR_t)^2} \tag{6}$$

In a similar fashion, for the CARs we define the following test statistic

$$Z = \sqrt{N}\frac{CAAR_i[t_1, t_2]}{s} \sim N(0, 1) \tag{7}$$

where the Cumulative Average Abnormal Return ($CAAR$) is

$$CAAR[t_1, t_2] = \frac{1}{N} \sum_{i=1}^{N} CAR_i[t_1, t_2] \tag{8}$$

and the standard deviation is

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (CAR_i[t_1, t_2] - CAAR[t_1, t_2])^2} \tag{9}$$

This test statistic accounts for event-induced variance as it uses an estimate of the cross-sectional variation of abnormal returns in the event window (testing period). An alternative way to account for event-induced variance is proposed by Boehmer et al. (1991) and is based on standardized abnormal returns as in Patell (1976). Abnormal returns AR_i in the event window are standardized by the time series standard deviation of AR_{it} in the estimation period $[-270, -21]$. We define

$$\overline{AR_i} = \frac{1}{250} \sum_{t=1}^{250} AR_{it} \tag{10}$$

and

$$\overline{s_i} = \sqrt{\frac{1}{249} \sum_{t=1}^{250} (AR_{it} - \overline{AR_i})^2} \tag{11}$$

The standardized abnormal returns are then defined as

$$SAR_{it} = \frac{AR_{it}}{\overline{s_i}} \tag{12}$$

The Boehmer et al. (1991) t-test is constructed by dividing the average SAR_{it} by their cross-sectional standard deviation:

$$T_{BMP} = \sqrt{N} \, \frac{ASAR_t}{S} \tag{13}$$

where

$$ASAR_t = \frac{1}{N} \sum_{i=1}^{N} SAR_{it} \tag{14}$$

and

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(SAR_{it} - ASAR_t)^2} \tag{15}$$

Kolari and Pynnonen (2010) provide a further adjustment to the Boehmer et al. (1991) test that also accounts for cross-sectional correlation of abnormal returns:

$$T_{KP} = T_{BMP} * \sqrt{\frac{1-\overline{r}}{1+(N-1)\overline{r}}} \tag{16}$$

where $\overline{r}$ is the average of the sample cross-correlations of the estimation period residuals.

We also use the more traditional method proposed by Brown and Warner (1980). This method estimates the standard deviation of average abnormal returns from the time series of average abnormal returns in the estimation period $[-270, -21]$:

$$\overline{s} = \sqrt{\frac{1}{249}\sum_{t=1}^{250}(AAR_t - \overline{AAR})^2} \tag{17}$$

where AAR_t is defined in (5) and

$$\overline{AAR} = \frac{1}{250}\sum_{t=1}^{250} AAR_t \tag{18}$$

The corresponding estimation of the standard deviation for the CAARs for a window $[t_1, t_2]$ is given by:

$$s^* = \sqrt{(t_2 - t_1 + 1)}\,\overline{s} \tag{19}$$

We use the cross-sectional variation of abnormal returns as defined in (4) for the baseline case. In the robustness section we also report results using the other three estimators.

3.2 Results

Table 1 reports daily average abnormal returns (AAR_t) for the event window [-10, +10], along with their statistical significance based on the test-statistic defined in equation (4). Results are reported for both upgrades and downgrades.[12] The corresponding raw return results are also reported for completeness.

An examination of the daily average abnormal returns reveals several important conclusions regarding the pre-announcement, announcement, and post announcement periods. First, the economic impact of downgrades appears to be significantly higher than that of upgrades as revealed by the absolute magnitude of the corresponding AAR around the announcement of the rating change. For example in the window $[-5, +1]$ statistically significant $AARs$ range from -0.366% to -0.257% for downgrades and from 0.117% to 0.189% for upgrades. Second, for downgrades, there is a statistically significant abnormal market reaction prior to the announcement of the ratings change, with AAR_{-5}, AAR_{-4}, AAR_{-3}, and AAR_{-1} being negative and statistically significant at least at the 5% level. For upgrades, only AAR_{-3} is statistically significant at the 5% level and AAR_{-4} at the 10% level.

[12] We have also experimented with separating changes in ratings beyond one grade, the idea being that changes above one grade might have a larger effect on the stock market than the single grade changes. We did not find any statistically different results relative to our baseline and we therefore report the results for all changes without distinguishing between the number of grades being changed.

Third, both upgrades and downgrades exhibit a further market reaction in the expected direction at the announcement window of [0, 1]. For downgrades both AAR_0 and AAR_{+1} are negative (-0.237% and -0.366%, respectively), and statistically significant at the 10% and 5% level, respectively. Results for upgrades are weaker, however, with the corresponding average abnormal returns not being statistically significant at $t = 0$ but statistically significant at the 5% level for $t = 1$ with $AAR_{+1} = 0.141\%$.

Finally, for both upgrades and downgrades we observe a statistically significant market reaction after the announcement of the ratings change, but in the opposite direction to the one found in the pre-announcement and announcement periods. More specifically, for downgrades we find positive and statistically significant average abnormal returns on days +3 and +4 relative to the announcement day (0.504% and 0.297%, respectively) and for upgrades we document negative and significant average abnormal returns on days +2 and +3 (-0.203% and -0.162%, respectively).

These observations about abnormal return behavior around the announcement of sovereign rating changes are confirmed in Figure 4A and in Table 2. Figure 4A graphs the Cumulative Average Abnormal Returns ($CAARs$) for both upgrades and downgrades. Figure 4B also reports the cumulative average raw returns for completeness. The stark difference in abnormal returns around rating downgrade announcements relative to upgrade announcements is immediately obvious in Figure 4A. Figure 4A shows that the stock market reacts more strongly throughout the pre-announcement period for downgrades rather than for upgrades. Moreover, the post-announcement effect is also larger after downgrades than upgrades and it goes in the opposite direction relative to the pre-announcement period.

Table 2 presents the information in Figure $4A$ differently by cumulating returns over different windows in the pre-announcement, announcement, and post-announcement periods. Table 2 also offers statistical tests for the respective $CAARs$[13] and also performs this analysis including information for the robustness samples: "ratings and outlooks FMRO" in panel B; "ratings FMRO" in panel C. For downgrades, in Panel A, we document an economically and statistically significant negative market reaction in the pre-announcement period ($CAAR[-5,-1] = -1.0\%$ with a p-value of 0.003), accompanied by a weaker, but still significant announcement effect ($CAAR[0,+1] = -0.6\%$ with a p-value of 0.01), and a significantly positive reaction in the post announcement period ($CAAR[+2,+5] = 1.1\%$ with a p-value of 0.005). The corresponding results in Panels B and C, are very similar, albeit with somewhat lower absolute values on the $CAAR$ estimates. The $CAAR[-5,+5]$ is negative across all three panels but not statistically significant, and is consistent with the near V-shape in Figure $4A$.

For upgrades, Panel A reports weaker evidence of positive abnormal returns in the pre-announcement period ($CAAR[-5,-1] = 0.1\%$ with a p-value of 0.286; $CAAR[-5,-3] = 0.3\%$ with a p-value of 0.030), accompanied by a statistically significant announcement effect ($CAAR[0,+1] = 0.2\%$ with a p-value of 0.015), and a statistically significant negative reaction in the post announcement period ($CAAR[+2,+5] = -0.4\%$ with a p-value of 0.007). We conclude that the market reaction for upgrades follows a similar pattern as the one for downgrades, but the

[13] Statistical significance is based on the test statistic in equation 7.

absolute magnitudes of $CAARs$ in all three phases (pre-, at and post- announcement) are much weaker. Examining the results in Panels B and C indicates that our conclusions are strongly robust in the manner we incorporate information from outlook changes to our baseline approach of ignoring outlooks (Panel A). We interpret these findings as evidence that stock market reactions after downgrades tend to be stronger than after upgrades.

Overall for our sample of downgrades, the pre-announcement and announcement evidence looks suspicious or "raises concerns". The pattern can be consistent with either a leakage of information in the days prior to the announcement of the rating downgrade or an anticipation of not only the downgrade, but its approximate timing as well. The magnitudes of the CAAR effects could also be driven by differences in liquidity, an issue we will revisit later. The post-announcement positive market reaction points to an over-reaction in the pre-announcement period and a correction after the announcement. For upgrades there appears to be weaker evidence of information leakage or anticipation of the announcement and a stronger, statistically significant announcement effect in the predicted direction. However, the post-announcement period exhibits a significant reversal of the documented announcement effect.

3.3 Robustness Checks

3.3.1 Incorporating Outlooks

We repeat our analysis for the "ratings FMR" sample for the two additional, robustness samples. For the "ratings and outlooks FMRO" the results are shown in

Figure 5A, and for the "ratings FMRO" in Figure 5B. In these two panels we show the market-adjusted, cumulative abnormal returns for each sample before and after the event. We observe similar results to the "ratings FMR" sample.

3.3.2 Estimator Choice

We also investigate how robust our results are to estimator choice. We repeat the same analysis as in Table 1, but use three additional ways to test for statistical significance for abnormal returns. In addition to the cross-sectional method in the baseline case, we also report results from the following methods. BW80 is the Brown and Warner (1980) method that estimates the standard deviation outside the event window (see equation 17); BMP91 is the Boehmer et al. (1991) method that allows for event-induced variance (see equation 13); and KP10 is the Kolari and Pynnonen (2010) method that allows for both event-induced variance and cross-correlation across rating changes simultaneously (see equation 16).

Table 3 shows the statistical significance of average abnormal returns using the four different estimators, for both upgrades and downgrades. For downgrades, Table 3 shows that our baseline results are robust to all estimators. We document significant pre-announcement and announcement effects, as well as significant post-announcement effects in the opposite direction. In fact, the results from the additional estimators are typically at least as statistically significant as the baseline case. Using the Brown-Warner (1980) test statistic, the only estimator that does not account for event-induced variance, generates higher significance relative to the baseline estimator consistent with the presence of event-induced variance in our sample. Due to space considerations we do not report CAAR results, but statistical significance carries over from the average abnormal returns to the CAARs. We conclude that the downgrade results are robust to the choice of estimator.

The evidence in Table 3 regarding upgrades is also typically robust across the four estimators. We observe weaker evidence of positive pre-announcement effects only on days -4 and -3, a positive and consistently significant announcement effect only on day $+1$ and a significant negative abnormal return on days $+2$ and $+3$. Taken together, the evidence from using all estimators in Table 3, shows that upgrades seem to be less important than downgrades in generating statistically significant abnormal returns.[14]

3.3.3 Without "First Mover"

A skeptic might also wonder whether our definition of "first mover" might also be responsible for our findings. We repeat the same methodology by including all rating agency changes (that is, without the "first mover" filter) and the results are plotted in Figure 5C. Our conclusions remain unchanged as the graph remains very similar to our baseline one. Downgrades show substantial pre-announcement stock market effects, with partial reversal after the announcement, while that is not the case for upgrades.[15]

[14] Results in Table 3 are robust across all samples: "Ratings FMR", "Ratings and Outlooks FMRO" and "Ratings FMRO".

[15] Results in Table 3 are robust across all samples: "Ratings (FMR)", "Ratings and Outlooks (FMRO)" and "Ratings (FMRO)".

3.3.4 Without "The Great Recession"

A skeptic might wonder whether the higher volatility during recessions is not well accounted for in the estimators of the previous subsection. Another worry might be that downgrades are more numerous in recessionary periods (as figure 2 suggests) when the stock market is more likely to fall. A particularly volatile time is the period of "The Great Recession" after 2008 (figure 2A shows that downgrades were the highest in October 2008 in our 1989 to 2011 sample period). We therefore repeat the same analysis but exclude all events after 2008, focussing solely on the "ratings FMR" sample of rating changes during the period of "The Great Moderation" (up to 2007).[16] Our results are depicted in Figure 5D and they are very similar to our baseline results (figure 4A), confirming that our conclusions are not driven by what happened by events after 2007.

4 What Drives the Results?

4.1 Correlations

In this subsection we correlate the calculated CAARs with the quality of institutions around the world. To do so we use various measures of institutional quality that have been used in the literature. Some of these measures can be thought of as exogenous variables because they were determined many decades before the actual rating change. The differential experience of common law versus civil law

[16] Results are robust across all samples: Ratings (FMR), Ratings and Outlooks (FMRO) and Ratings (FMRO).

countries is one such dichotomy. Any differential results found there can, given the exogeneity of legal origins to trading mechanisms post-1991, be interpreted as causal. Other categorizations might suffer more from endogeneity issues. The division of the countries in the sample between non-developed and developed offers one such example. Nevertheless, we think that uncovering such correlations is still informative for future research on the topic, while in the next subsection we are more careful in trying to disentangle cause and effect.

Given our results to date (that downgrades have a larger stock market impact than upgrades), we focus on downgrades. To better understand the behavior of abnormal returns around the announcement of sovereign downgrades, we repeat our econometric analysis conditioning on country level characteristics aimed to proxy for the quality of a country's institutional framework or government. We first condition on emerging/frontier (non-developed) and developed countries based on the World Bank classification. We also condition on the origin of the legal system (civil vs common law), the law and order and corruption indices from The Political Risk Services Group (PRS-law and order and PRS-corruption, respectively), the corruption perception index from Transparency International (TI-corruption), and the investor protection index from the World Bank's Doing Business website. The data appendix contains further details about variable definitions and data sources.

Figure 6 plots the cumulative abnormal returns around downgrades after sorting the countries in our dataset according to the variables described above. For countries that have a continuous index, we take a very conservative approach and separate countries above and below the "ratings FMR" sample's median value of that measure (results are naturally stronger if we compare the top to the bottom percentiles). The main message from the results that follow is that the identified patterns in abnormal returns are more pronounced in countries with "lower quality" institutions/government.

Figure 6, Panel A plots civil versus common law systems, and there is evidence that downgrades have a bigger impact on abnormal returns before the event in civil law countries. Table 4 shows that the results for civil law countries are statistically significant for all $CAARs$ from ($t = -5$ *to* $t = -3, -2, -1$) and with a magnitude ranging from -1.14% to -1.4%, whereas the results for common law countries and for the same time windows are statistically insignificant from zero. At the same time, for civil law countries $CAAR[+2, +5]$ is positive (1.28%) and statistically significant (p-value 0.008), whereas the common law coefficient is not statistically significant. We conclude that downgrades are more likely to have an impact in civil law, rather than common law, countries in the pre- and post- announcement periods.

Figure 6, Panel B plots the $CAARs$ for non-developed relative to developed (advanced) economies. The graph illustrates that non-developed countries tend to exhibit larger CAAR effects than developed countries. Table 4 illustrates this statistically. For the $CAAR$ from ($t = -5$ *to* $t = -3, -2, -1$) the non-developed market effect is economically larger and statistically significant relative to the developed market effect, which is statistically insignificant. The positive effect after the event is also statistically significant for non-developed, but not for developed economies. We conclude that our empirical results are more likely to appear in non-developed rather than developed economies.

The TI-corruption results for downgrades are reported in the next column of Table 4, and graphically shown in Figure 6, Panel C. The results are striking: countries with a higher corruption perception index react much more strongly prior, at and after the downgrade. For all $CAAR$ windows for high-corruption countries from ($t = -5$ *to* $t = -3, -2, -1$) there is a pre-event effect ranging between -1.51% and -1.80%, which is statistically significant at the 1% level. During the announcement window, the reaction is -1.10%, which is also statistically significant at the 1% level. Most of these negative returns are reversed in the post-announcement window $CAAR[+2, +5]$ where there is a positive abnormal return of 1.54% with a p-value of 0.001. On the contrary, the pre- , at-, and post- announcement effects for countries with a low corruption perception index are nowhere statistically significant.

The next column (and graphically in Panel D, Figure 6) reports the results from splitting countries according to the PRS-law and order index. The results match very closely the results from the TI-corruption index. Table 4 reports that for all $CAAR$ windows from ($t = -5$ *to* $t = -3, -2, -1$) the effect is statistically significant at the 1% level and ranges between -1.38% and -1.62% for countries with a low PRS-law and order ranking. For the same group of countries Table 4 also documents statistically significant $CAARs$ for the announcement ($CAAR[0, 1]$) and post announcement ($CAAR[+2, +5]$) windows of -0.84% and 1.73%, respectively. The results are statistically insignificant for the countries with a high ranking in the law and order index, except for one case after the event ($CAAR[+2, +5]$) that has a p-value of 0.074.

The final two columns report results from splitting countries according to the PRS-corruption index and the investor protection index, respectively. The graphical results in Panel E and F in Figure 6 illustrate that the differences across countries grouped according to these two variables might not be statistically different from each other. The last two columns of Table 4 confirm this impression. We observe statistically significant reactions before, at and after the downgrade using both measures, but the differences across the groupings are not as striking as when using the previous four variables. Nevertheless, there is still some evidence for a significant pre-announcement effect for the low investor protection countries for $CAAR$ windows from ($t = -5$ *to* $t = -3, -2$), while the equivalent results for the high investor protection countries are not statistically significant.

One possible concern might be that the CAAR magnitudes are exacerbated by lower capital market liquidity, a feature more likely to appear in countries with lower institutional quality. To address this concern we use local market turnover as a measure of liquidity[17] and sort our sample across both corruption and turnover characteristics (we focus on TI Corruption for brevity considerations, since results are similar for civil vs. common law, developed vs. emerging and frontier, and PRS Law & Order). Figure 7 shows that differences in liquidity do not affect our conclusions. Specifically, Figure 7 shows that the observed pattern of abnormal return behavior is still present only in the low transparency sample, regardless of

[17] We measure liquidty using local market turnover at an annual frequency. We define liquidity as total market volume traded divided by market capitalization at the end of the year. Data are taken from the World Bank Development Indicators Database.

the liquidity measure.

We conclude that even with a very conservative split of country characteristics (above and below the median for continuous, imperfect measures of institutional quality), there is evidence for a statistically significant reaction in the stock market before the rating downgrade. Moreover, this effect is mostly present in countries that tend to be associated with "lower quality institutional frameworks". Furthermore, our results are also robust to using the other two samples (that is, incorporating information from changes in outlook and watchlist inclusions).

4.2 Causal Evidence

Most, if not all, independent variables in the previous subsection (transparency index, law and order, corruption and investor protection indices) are likely to suffer from either endogeneity or errors-in-variables bias. A positive correlation between abnormal stock returns and institutional quality by no means implies that institutional quality causally affects the documented abnormal stock return pattern. Error-in-variables problems can also affect our conclusions. Consider, for example, using a proxy variable to measure government corruption or institutional quality and this proxy being an imperfect measure of the true corruption/quality. Regressing cumulative abnormal returns on the proxy variable will suffer from a classic errors-in-variables problem, generating biased estimates depending on the degree of measurement error.

The classic solution to both problems is to search for suitable instrumental variables. In this section we therefore use instrumental variables techniques to give a causal interpretation to the uncovered correlations. In so doing, we provide evidence supporting the idea that the mechanism through which rating announcements reach the capital market needs to concern capital market regulators around the world. Specifically, we conduct two stage least squares (2SLS) regressions of $CAARs$ before and after the event on each of the potentially endogenous variables that can proxy for institutional quality: Emerging/Frontier vs. Developed; TI Corruption; PRS Law and Order; PRS Corruption; Investor Protection.

What are appropriate instrumental variables for these regressors? First, we consider the separation based on the legal system: common vs. civil law (La Porta et al., 1998), where an indicator takes the value of 1 for common law system and 0 otherwise. The next two instruments are the ethnicity and religion fractionalization measures developed by Alesina et al. (2003) as explained in the data appendix. The fourth candidate instrument is the landlocked indicator (Easterly and Levine, 2003). The four instrumental variables are arguably exogenous because they have been determined many decades before the ratings events we study. Moreover, legal origin, fractionalization and geography are good candidates for random variation that might be correlated with the endogenous variable of interest (different measures of institutional quality) but not directly affect the dependent variable (cumulative stock returns), the two conditions needed for a valid instrument.

We view our five endogenous variables as proxying the quality of capital market institutions in a country. We therefore do five separate 2SLS regressions, one for each of the five endogenous variables. For each of these variables we identify the best instruments using the method of Baum, Schaffer and Stillman (2010). Specifically we compute and report tests for model under-, weak-, over-identification as well as for the redundancy of instrumental variables. For all endogenous variables we start by assuming that the four instruments are valid. The null hypothesis under the redundancy test is that the specified instrument is redundant. As shown in Table 5, most of the chosen instruments are valid (p-values<0.01; panel A). The procedure is repeated for each endogenous variable until no more redundant instruments appear (panel B).[18] For instance, for the endogenous variable TI Corruption, all but one IV (landlocked) is statistically significant at a p-value of 1% or lower, therefore in the second step, the redundant IV (landlocked) is not included. Therefore, in the 2SLS estimation, TI Corruption is proxied by the three non-redundant IVs (round 2 in Table 5): common/civil law, ethnic fractionalization and religion fractionalization.

Results of the 2SLS regressions are shown in Tables 6 and 7. Our dependent variables are $CAAR[-5,-3]$[19] (Table 6) and $CAAR[+2,+5]$ (Table 7). We expect that abnormal market reactions before the event will have a positive relation with measures of institutional quality (i.e. since $CAARs$ are negative, we expect that when institutional quality is better, $CAARs$ will be less negative). We find both statistically and economically significant results for four out of the five endogenous variables we use.

Panel A of Table 6, reports the results for our baseline sample. Emerging/Frontier countries (using common law, ethnic fractionalization and landlocked

[18] As a sensitivity test in the 2SLS regressions we re-run our analysis with no restriction on redundancy of instrumental variables (i.e. we use the same four instruments for all endogenous variables) and obtain similar results with higher statistical signfinance, except for investor protection.

[19] We also conduct our regressions on $CAARs$ ($t = -5$ *to* $t = -1$) and we obtain the same results.

as instruments) generate $CAARs$ of about 2.7% (p-value of 0.010) lower than those of developed countries. Similarly, a positive (and statistically significant at the 1% level) coefficient is obtained for the TI Corruption score (using common law, ethnic fractionalization and religion fractionalization as IVs). A one-standard deviation (2.14) decrease in the TI corruption score gives a 1.28% (p-value of 0.001) decrease in the $CAAR$, where all identification tests are again satisfied. The coefficient for PRS Law & Order (using common law and ethnic fractionalization as instruments) is also positive, indicating an overall decrease in $CAAR$ of 1.01% (p-value of 0.013) when the score decreases by one standard deviation (1.26).[20] Similarly, a one-standard deviation (1.46) decrease in the Investor Protection Index (using common law and landlocked as instruments) gives a 1.02% (p-value of 0.070) decrease in the $CAAR$, where all identification tests are again satisfied. Lastly, the coefficients for PRS Corruption (using landlocked as instrument) are not statistically significant.

The statistical tests strongly reject under-identification (UID) for all regressions, while in most models we reject the weak instrumental variables (WID) hypothesis. Both reported test statistics (Cragg-Donald for i.i.d. error disturbances and Kleibergen-Paap for non-i.i.d. errors) exceed the Staiger and Stock (1997) rule of thumb of ten to reject the hypothesis of weak IVs. The Cragg-Donald F-test rejects the hypothesis in four out of five cases in our baseline specification, when compared to the Stock and Yogo relative bias and relative size tests. For the PRS Corruption index where the hypothesis is not rejected, we do not find statistically significant results. The Kleibergen-Paap test also passes the Stock and Yogo (2005) critical

[20] Countries with the highest (lowest) score in TI Corruption have the lowest (highest) corruption. The same rationale applies for PRS Law & Order.

values for 10% maximal IV relative bias and IV size, even though strictly speaking these values should be compared to the i.i.d. case. On balance, we interpret these results as rejecting the weak IV hypothesis. Also, the Hansen J statistic, does not reject the over-identification (OID) hypothesis.

Panels B and C repeat the analysis in Panel A using the IV selection process in Table 5 on the "ratings and outlooks FMRO" and the "ratings FMRO" samples, respectively. The robustness of the results is quite striking since statistical significance does not change for any of the variables, while most coefficients either remain unchanged or only change at the third decimal place (except for PRS Law and Order in panel B).

Mostly for completeness, we also report in Table 7 the 2SLS regressions on abnormal stock market returns after the event using $CAAR[+2, +5]$. The table serves to illustrate that there is only very weak to non-robust evidence for reversion after the event in the opposite direction. From all the variables being considered, only TI corruption is statistically significant. We conclude that the pre-event response is the most robust of our empirical findings.

We also repeat the 2SLS regressions controlling for liquidity, using our turnover measure as an independent variable. The turnover measure turns out to be statistically insignificant and our conclusions remain unchanged across all samples and all variables. For brevity we do not report these findings.

From these results we can conclude that pre-announcement CAAR reactions, tend to take place mostly in countries with lower institutional quality. There are different possible explanations for these findings. These results are more consistent with a leakage of information about the content and timing of the pending announcement hypothesis, rather than the market anticipation story. We take this view because the presence of significant negative pre-event abnormal returns predominantly in low institutional quality markets points to actions that raise “concerns”, since it is hard to justify that markets with low institutional quality are better at anticipating credit rating actions. The leakage of information could be coming from the rating agency itself, but might also be coming from local government bodies that the rating agency is obliged to inform and ask for their feedback after the rating is close to completion, but before the rating’s public announcement. Further evidence is necessary before reaching firm conclusions, but we view our results as supportive of the idea that the mechanism through which rating announcements reach the capital market needs to be a possible concern for capital market regulators around the world.

5 Conclusion

We find evidence that the stock market moves before the public announcement of a sovereign rating downgrade, resulting in a significant market reaction prior to the event. Including information from outlook changes and watchlist inclusions to control for potential anticipation effects does not alter our findings. The results are robust over periods with lower volatility in the stock market ("The Great Moderation"), across different estimators and without our more involved definition of a “first mover” when defining the event.

More importantly, we document a link between different measures of corruption and/or institutional quality and the pre-downgrade negative abnormal return. Specifically, we find empirical evidence that this result is mostly present in countries that are less developed, tend to be more corrupt, have weaker law enforcement, and are under the civil (rather than common) law system. Using instrumental variables techniques we also build a causal argument that the negative abnormal stock market reaction before sovereign debt downgrade announcements tend to occur in countries with lower institutional quality in terms of law enforcement, corruption and development.

Data Appendix: Country Level Characteristics

Common vs. civil law countries: We use the country's legal origin classification from Djankov et al. (2008). La Porta et al. (1998) find that common law countries provide stronger legal protections for investors relative to civil law countries.

Emerging and frontier (non-developed) versus developed country classification: In our analysis we differentiate between developed and emerging/frontier markets following the World Bank classification system into emerging, frontier and developed countries. We account for countries that moved from Emerging to developed market status during our sample period.

The measure of corruption is from the Political Risk Services Group (PRS-corruption) and is defined as follows: "A measure of corruption within the political system that is a threat to foreign investment by distorting the economic and financial environment, reducing the efficiency of government and business by enabling people to assume positions of power through patronage rather than ability, and introducing inherent instability into the political process". This variable takes values from 0 to 6 with 6 denoting the lowest level of corruption. The variable is available on a monthly frequency.

Law and order is another measure taken from the Political Risk Services Group (PRS-law and order): Two measures comprise this risk component. Each sub-component equals half of the total. The "law" sub-component assesses the strength and impartiality of the legal system, and the "order" sub-component assesses popular observance of the law. This variable takes values from 0 to 6 with 6 denoting countries scoring the highest on law and order quality. The variable is available on a monthly frequency.

We also use an alternative measure of corruption, namely the Corruption Perceptions Index from Transparency International (TI-corruption). First launched in 1995, it has been widely credited with putting the issue of corruption on the international policy agenda. The CPI ranks almost 200 countries by their perceived levels of corruption, as determined by expert assessments and opinion surveys. The variable is available on an annual frequency. We use the 1995 values for events before 1995.

Strength of investor protection index: The strength of investor protection index is the average of the extent of disclosure index, the extent of director liability index and the ease of shareholder suits index. The index ranges from 0 to 10, with higher values indicating more investor protection. This methodology was developed in Djankov, La Porta, Lopez-de-Silanes, Schleifer (2008).

Ethnic and religious fractionalization measures the ethnic and religious heterogeneity in a country, respectively. These measures are developed by Alesina et al. (2003). In both cases, fractionalization takes values from 0 to 1, where 1 shows no fractionalization and 0 shows the existence of multiple ethnic and religious groups (fractions) in each country.

The landlocked indicator takes the value of 1 if the country has no outlet to the sea and 0 otherwise. Easterly and Levine (2003) show that not having access to the sea is negatively correlated with institutional quality.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, 2003, Fractionalization, *Journal of Economic Growth* 8, 155-194.

Anilowski, Carol, Mei Feng, and Douglas J. Skinner, 2007, Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance, *Journal of Accounting and Economics* 44, 36-63.

Baum, Christopher, F. , Mark Schaffer, E. , and Steven Stillman, 2010, Ivreg2: Stata module for extended instrumental variables/2sls and gmm estimation, (Boston College Department of Economics).

Beaver, William H., Catherine Shakespeare, and Mark T. Soliman, 2006, Differential properties in the ratings of certified versus non-certified bond-rating agencies, *Journal of Accounting and Economics* 42, 303-334.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad, 2007, Liquidity and expected returns: Lessons from emerging markets, *Review of Financial Studies* 20, 1783-1831.

Boehmer, Ekkehart, Jim Masumeci, and Annette B. Poulsen, 1991, Event-study methodology under conditions of event-induced variance, *Journal of Financial Economics* 30, 253-272.

Brooks, Robert, Robert W. Faff, David Hillier, and Joseph Hillier, 2004, The national market impact of sovereign rating changes, *Journal of Banking &Finance* 28, 233-250.

Brown, Stephen J., and Jerold B. Warner, 1980, Measuring security price performance, *Journal of Financial Economics* 8, 205-258.

Charest, Guy, 1978, Dividend information, stock returns and market efficiency-ii, *Journal of Financial Economics* 6, 297-330.

Cheng, Mei, and Monica Neamtiu, 2009, An empirical analysis of changes in credit rating properties: Timeliness, accuracy and volatility, *Journal of Accounting and Economics* 47, 108-130.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2008, The law and economics of self-dealing, *Journal of Financial Economics* 88, 430-465.

Easterly, William, and Ross Levine, 2003, Tropics, germs, and crops: How endowments influence economic development, *Journal of Monetary Economics* 50, 3-39.

Ederington, Louis H., and Jeremy C. Goh, 1998, Bond rating agencies and stock analysts: Who knows what when?, *Journal of Financial and Quantitative Analysis* 33, 569-585.

Guiso, Luigi, Paola Sapienza and Luigi Zingales, 2008, Trusting the Stock Market, *The Journal of Finance*, 63 (6), 2557-2600.

Hand, John R. M., Robert W. Holthausen, and Richard W. Leftwich, 1992, The effect of bond rating agency announcements on bond and stock prices, *Journal of Finance* 47, 733-752.

Hill, Paula, and Robert Faff, 2010, The market impact of relative agency activity in the sovereign ratings market, *Journal of Business Finance & Accounting* 37, 1309-1347.

Holthausen, Robert W., and Richard W. Leftwich, 1986, The effect of bond ratings changes on common stock prices, *Journal of Financial Economics* 17, 57-89.

Hutton, Amy P., Gregory S. Miller, and Douglas J. Skinner, 2003, The role of supplementary statements with management earnings forecasts, *Journal of Accounting Research* 41, 867-890.

Johnson, Richard, 2004. Rating Agency Actions Around the Investment-Grade Boundary. *Journal of Fixed Income* 13, 25-37.

Kaminsky, Graciela, and Sergio L. Schmukler, 2002, Emerging market instability: Do sovereign ratings affect country risk and stock returns?, *The World Bank Economic Review* 16, 171-195.

Kolari, James W., and Seppo Pynnonen, 2010, Event study testing with cross-sectional correlation of abnormal returns, *Review of Financial Studies* 23, 3996-4025.

Kothari, S. P., Susan Shu, and Peter D. Wysocki, 2009, Do managers withhold bad news?, *Journal of Accounting Research* 47, 241-276.

Kothari, S.P, and Jerold B. Warner, 2007, Econometrics of event studies, in B. Espen Eckbo, ed.: The handbook of corporate finance: Empirical corporate finance (Elsevier, Oxford, UK).

Kroft, Steve, 2012, Congress: Trading stock on inside information? accessed August 27, 2012. Available from http://www.cbsnews.com/8301-18560_162-57323527/congress-trading-stock-on-inside-information/?tag=contentMain;contentBody

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2008, The economic consequences of legal origins, *Journal of Economic Literature* 46, 285-

332.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny, 1998, Law and finance, *Journal of Political Economy* 106, 1113-1155.

Low, Angie, 2009, Managerial risk-taking behavior and equity-based compensation, *Journal of Financial Economics* 92, 470-490.

Martell, Rodolfo, 2005, The effect of sovereign credit rating changes on emerging stock markets, *Purdue University working paper.*

Patell, James, 1976, Corporate Forecasts of Earnings Per Share and Stock Price Behavior: Empirical Test. *Journal of Accounting Research* 14, 246-76.

Penman, Stephen H., 1982, Insider trading and the dissemination of firms' forecast information, *The Journal of Business* 55, 479-503.

Rossi, Sara, 2012, Italy extends rating agencies probe to Fitch, accessed August 27, 2012. Available from http://www.reuters.com/article/2012/01/24/us-italy-fitch-idUSTRE80N1II20120124?feedType=RSS&feedName=businessNews&utm_source=dlvr.it&utm_medium=twitter&dlvrit=56943.

SEC, 2003, Report on the role and function of credit rating agencies in the operation of the securities markets as required by section 702(b) of the sarbanes-oxley act of 2002., (U.S. Securities and Exchange Commission).

SEC, 2008, Summary report of issues identified in the commission staff's examinations of select credit rating agencies, (U.S. Securities and Exchange Commission).

Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news, *Journal of Accounting Research* 32, 38-60.

Soffer, Leonard C., S. Ramu Thiagarajan, and Beverly R. Walther, 2000, Earnings preannouncement strategies, *Review of Accounting Studies* 5, 5-26.

Staiger, Douglas, and James H. Stock, 1997, Instrumental Variables Regression with Weak Instruments. *Econometrica* 65, 557-86.

Stock, James H., and Motohiro Yogo, 2005, Testing for Weak Instruments in Linear IV Regression, in Donald W. K. Andrews and James H. Stock, eds., Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, Cambridge: Cambridge University Press, 80–108.

Figure 1: U.S. Downgrade by Standard & Poor's: Cumulative Raw Returns in the twenty days before and after the downgrade of the U.S. sovereign debt rating by Standard & Poor's on August 5th, 2011.

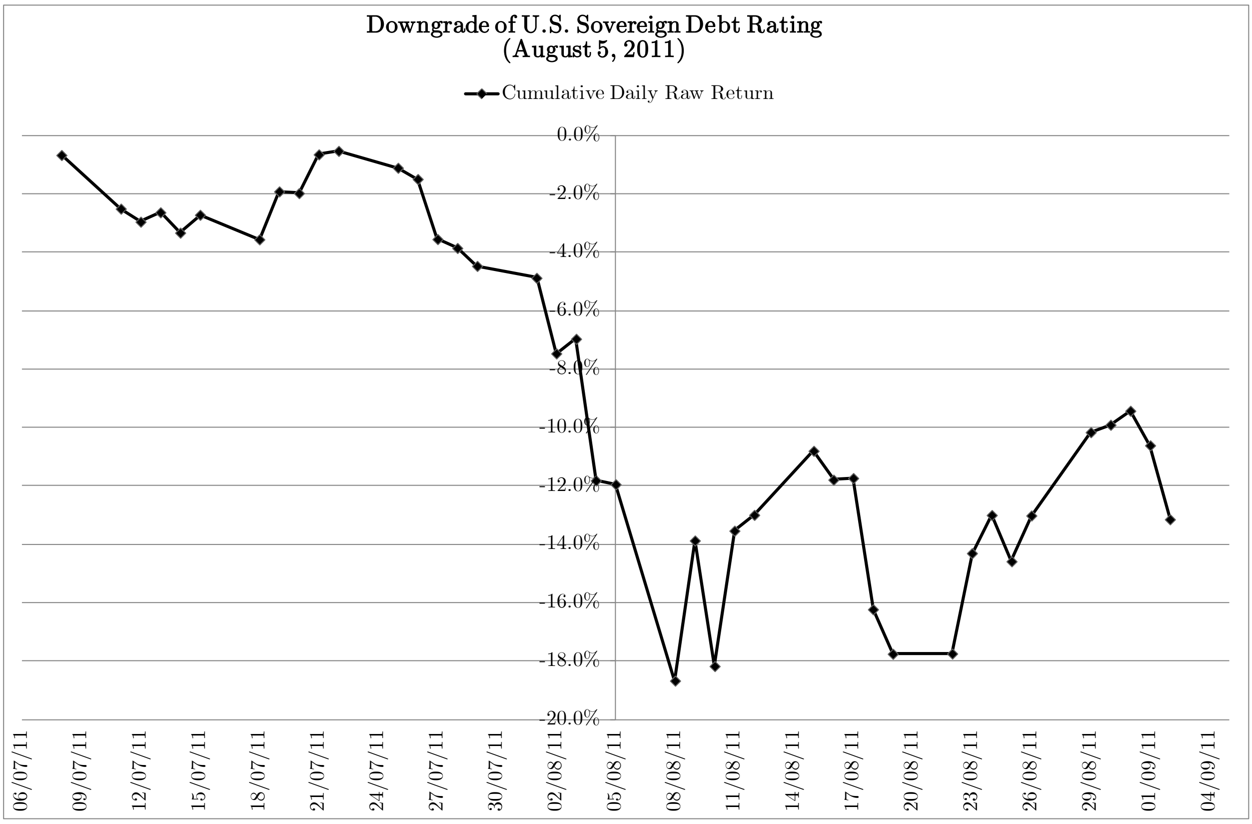

Figure 2: Time series distribution of Changes in Sovereign Debt Ratings. In Figure 2A we show all changes in ratings by Fitch, Moody's and Standard & Poor's, over time. The sample comprises 456 upgrades and 418 downgrades for 65 countries. In Figure 2B, we show the changes in ratings that are free from other changes in ratings in the previous twenty days from the same or another rating agency ("First Mover" Rating agency). These comprise 400 upgrades and 293 downgrades from 65 countries.

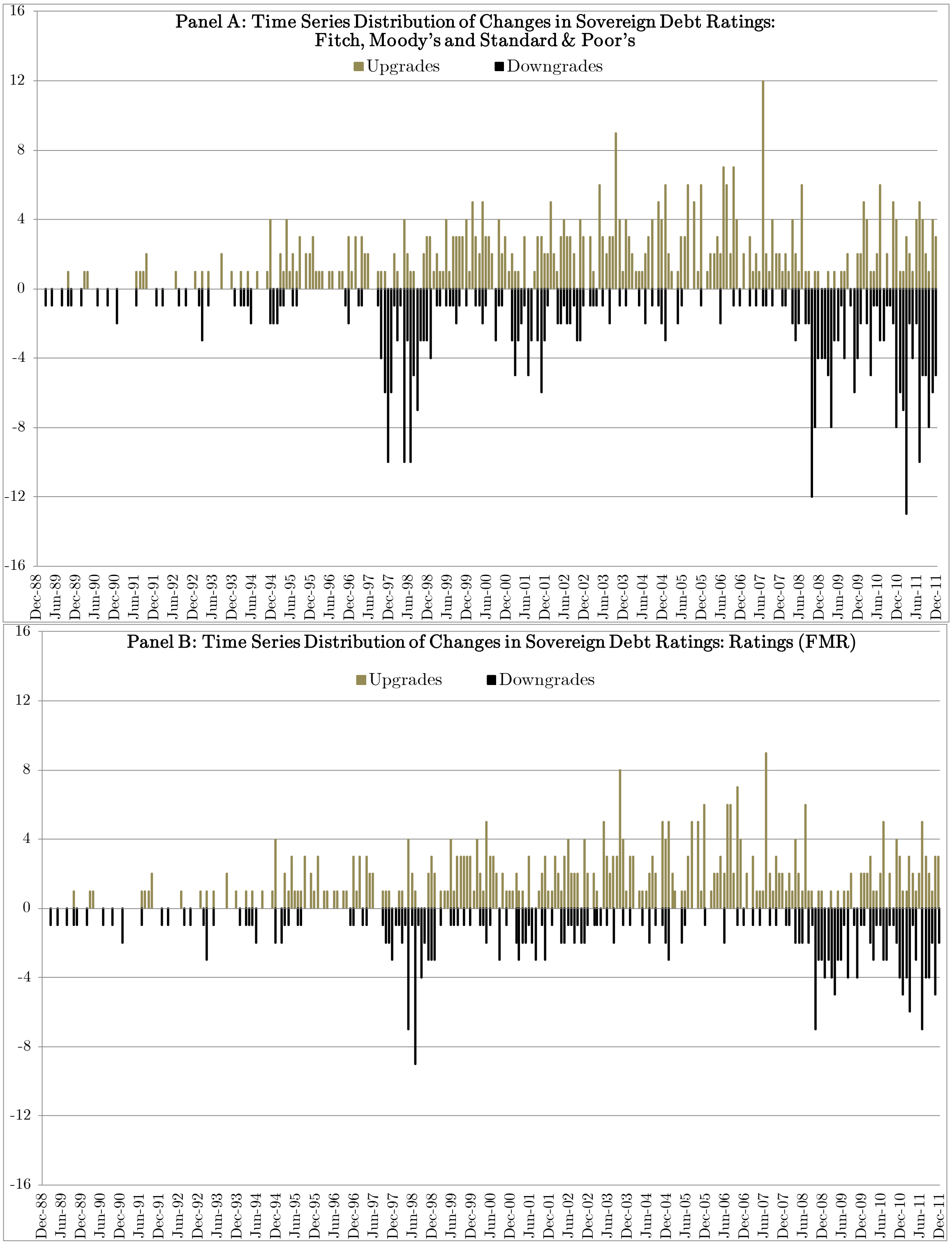

Figure 3: Grade-Level Distribution of Changes in Sovereign Debt Ratings by Rating Agency. A change in rating is defined by either a change in the Local or Foreign Currency Rating. The sample shown is called "Ratings FMR" and it comprises changes in ratings by all rating agencies (changes in outlooks are excluded), which are not preceded by other changes in ratings by the same or other rating agencies in the previous twenty trading days (FMR stands for First Mover using Ratings). In Figure 3A we plot the 117 upgrades and 86 downgrades by Fitch Ratings; In Figure 3B we plot the 133 upgrades and 70 downgrades by Moody's Ratings; In Figure 3C we plot the 150 upgrades and 137 downgrades by Standard & Poor's Ratings. In Figure 3D we plot the 400 upgrades and 293 downgrades for the Ratings (FMR) sample. The horizontal axis shows the categories of ratings (higher numbers indicate lower debt quality).

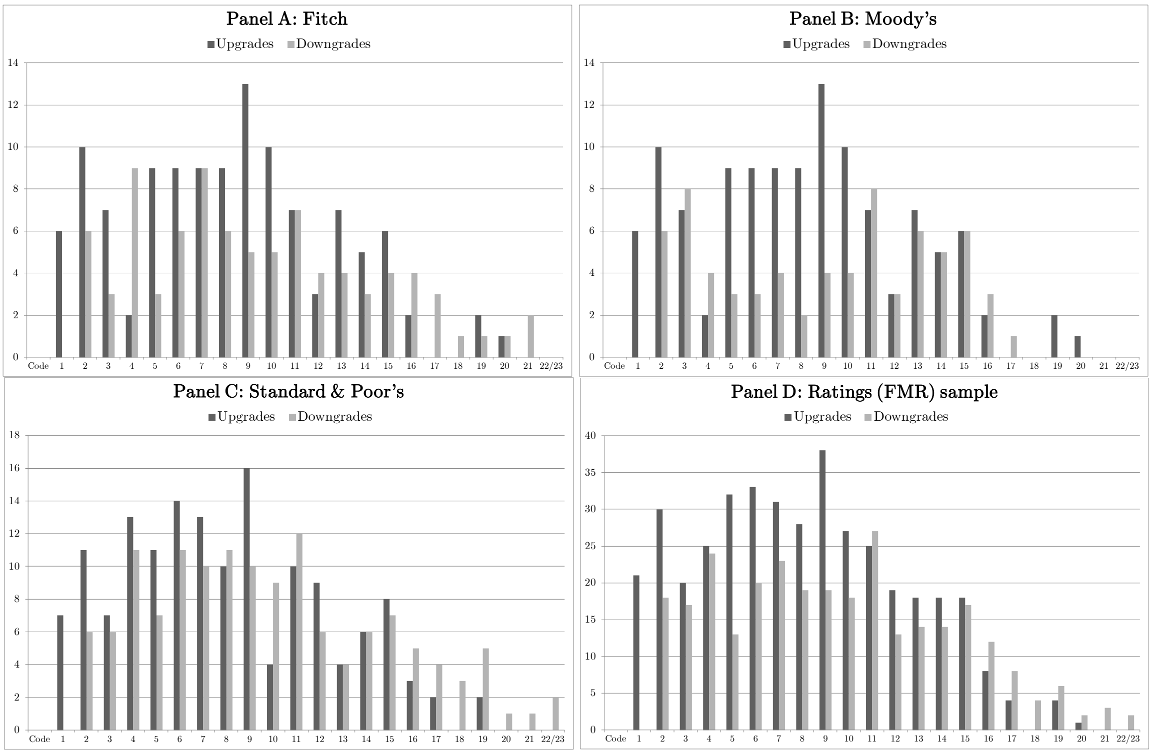

Figure 4: Cumulative Returns around Changes in Sovereign Ratings. Panel 4A shows market-adjusted cumulative average abnormal returns (CAARs) for the "ratings FMR" sample: 376 upgrades and 271 downgrades. Panel 4B shows cumulative raw returns around the time of upgrades and downgrades respectively. Changes in ratings are free from noise from other rating changes ("First Mover" using ratings; FMR), as all rating changes in the preceding twenty days from the same or other rating agencies are removed ("Ratings FMR" sample).

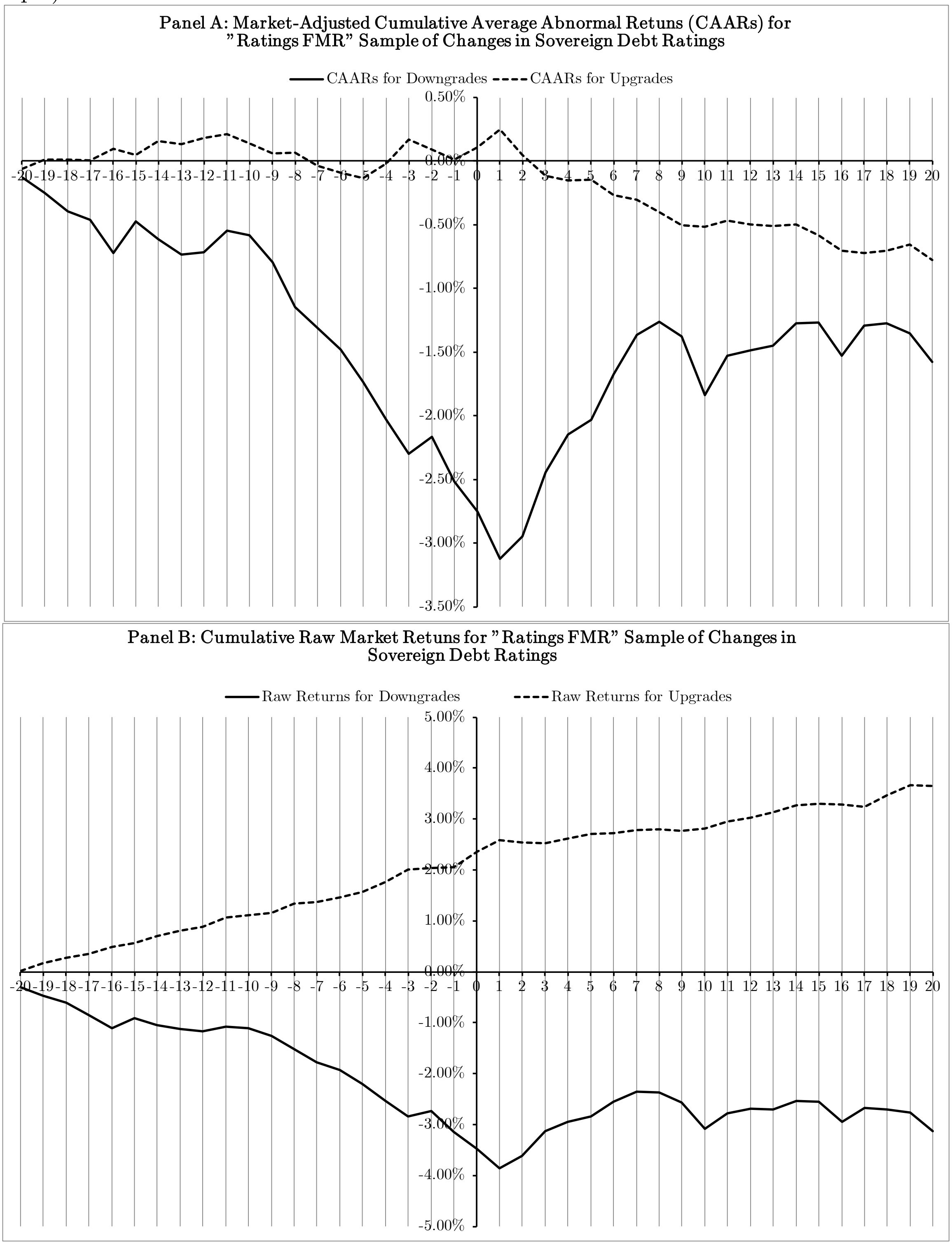

Figure 5: Robustness of Market-Adjusted Cumulative Average Abnormal Returns (CAARs), around Changes in Sovereign Ratings. The analysis in Figure 4A is repeated for four robustness checks. Panel 5A shows CAARs for the "Ratings and Outlooks FMRO" sample (638 upgrades and 456 downgrades). Panel 5B shows the "Ratings FMRO" sample (358 upgrades and 240 downgrades). Panel 5C shows CARs for the sample in 4A without the "First Mover" filter (i.e. union of changes in ratings; 430 upgrades and 381 downgrades). Panel 5D shows the same sample as 5C, but excluding the recent financial crisis (303 upgrades and 157 downgrades). The description of the sample construction process is described in Table 2.

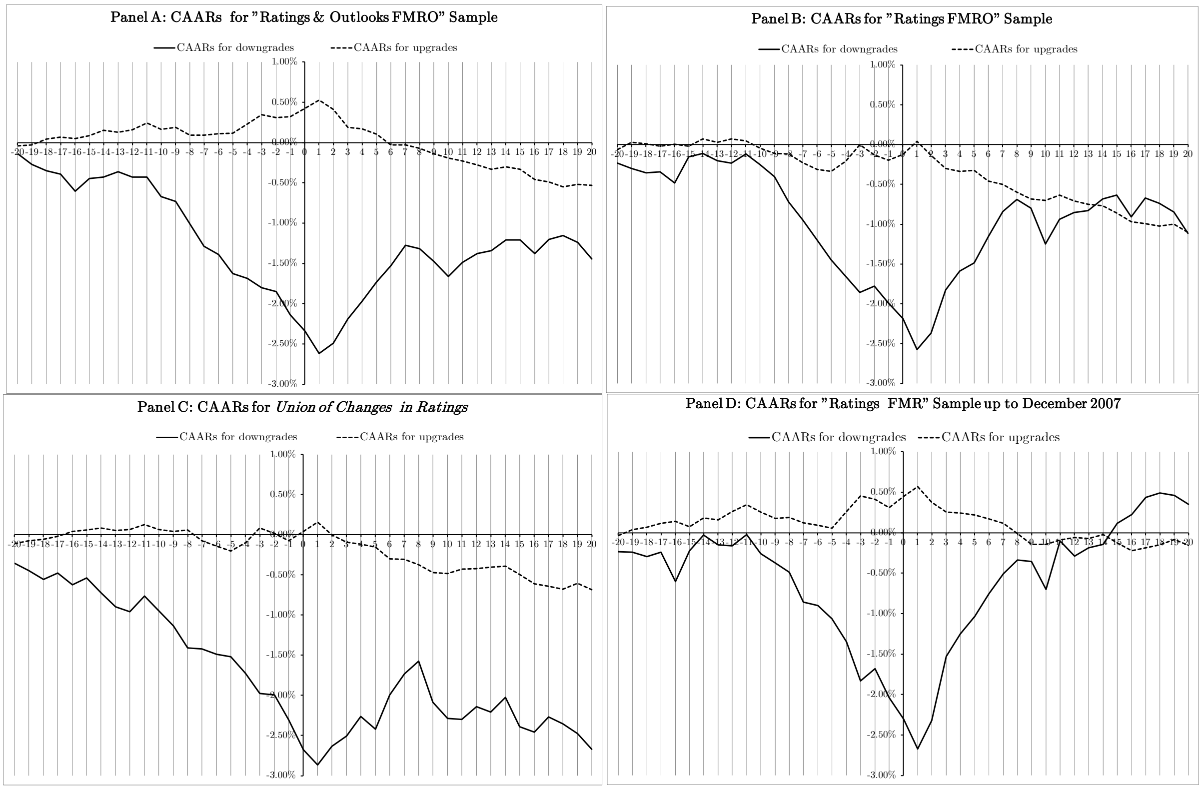

Figure 6: Breakdown of Cumulative Average Abnormal Returns Around Ratings' Downgrades. All Panels show the cumulative average abnormal returns (CAARs) for downgrades in sovereign ratings, according to individual country characteristics ("Ratings FMR" sample). Panel 6A shows the breakdown according to the legal system law. There are a total of 203 observations with Civil law and 68 with Common Legal System. Panel 6B shows the breakdown based on the World Bank's classification of Developed vs. Emerging and Frontier countries. There are a total of 180 observations classified as Emerging & Frontier, and 91 as developed countries. Panel 6C shows the breakdown according to the Transparency Index (low vs. high has 133 vs. 130 observations, respectively). Panel 6D shows the breakdown according to the Law and Order Index (low vs. high has 127 vs. 125 observations, respectively). Panel 6E shows the breakdown according to the Corruption Index (low corruption score vs. high corruption score has 112 vs. 139 observations respectively; Low score means low institutional quality). Panel 6F shows the breakdown according to the Investor Protection Index (low vs. high has 153 vs. 118 observations, respectively). The separation of each category is made at the median value of the "Ratings FMR" sample.

Figure 7: Double Sorting Cumulative Average Abnormal Returns Around Ratings' Downgrades, on Institutional Quality and Liquidity.
The graph shows the cumulative average abnormal returns (CAARs) for downgrades in sovereign ratings, according to the level of Institutional Quality and Liquidity. We proxy institutional quality using the Transparency Index (TI). High (low) TI comprises downgrades with above (below) median transparency scores. We proxy liquidity (liq.) using the turnover in the year of the event. Turnover is the ratio of total market volume traded divided by market capitalization at the end of the year. High (low) liquidity comprises downgrades with above (below) median liquidity levels. The "Ratings FMR" sample is used to construct 4 graphs with the following characteristics: High TI and High Liq. (n=54); High TI and Low Liq. (n=87); Low TI and High Liq. (n=75); Low TI and Low Liq. (n=57).

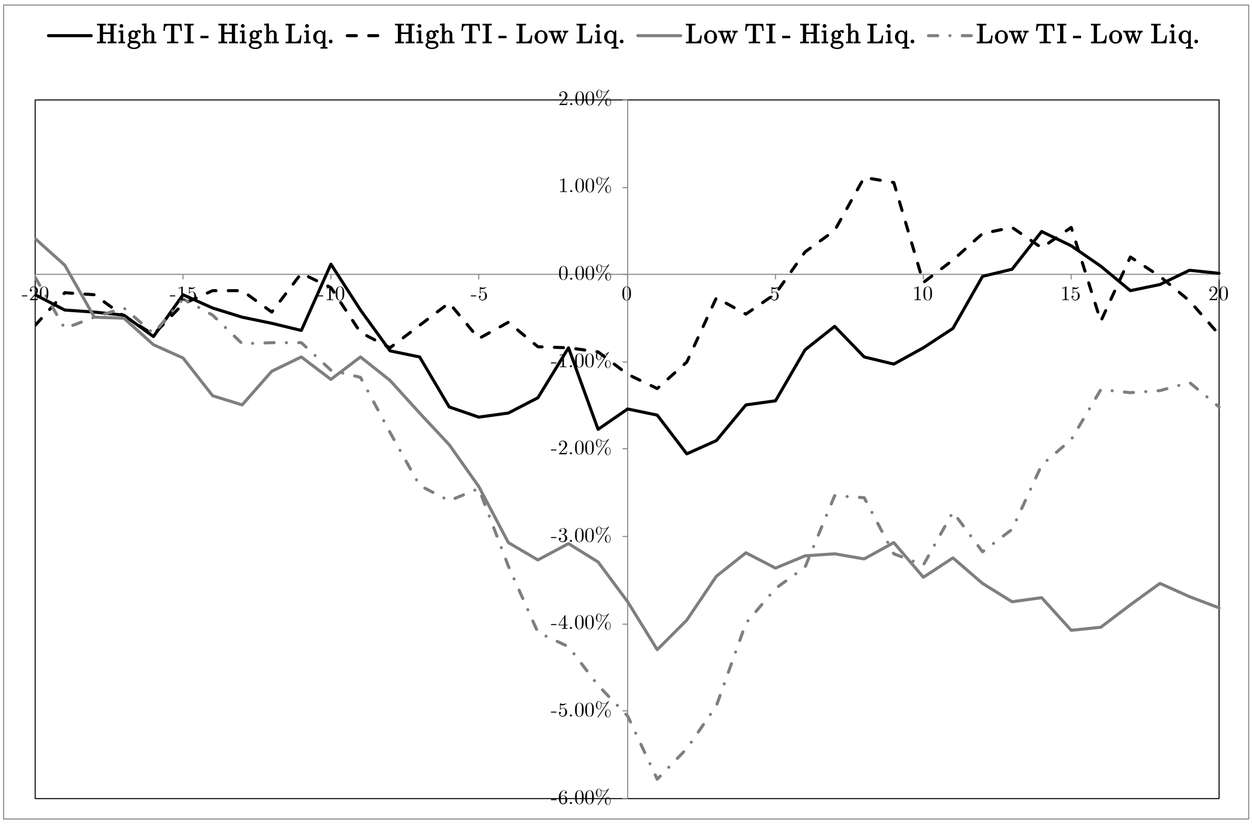

Table 1

Event Study of Changes in Sovereign Ratings on Local Stock Market Indices

This table presents the event-study results of how changes in sovereign debt ratings affect the respective sovereign daily, stock market return. Results are reported separately for upgrades and downgrades. The sample comprises the union of changes in ratings, from Fitch, Moody's and Standard & Poor's, filtered by first mover filter using "Ratings FMR". FMR means that all observations preceded by changes in ratings by the same or other rating agency in the previous twenty trading days are deleted. Rel. Day is the trading day relative to the event day (day 0). $AAR(i,t)$ is the average abnormal return for all observations i for each day t,using a world-CAPM model. Mean *R(i,t)* is the average raw return for all observations *i* for each day *t*. The sample includes 376 upgrades and 271 downgrades. P-values are based on the cross-sectional approach (Equation 4). ***,**, and * denote statistical significance at the 1, 5, and 10 percent level.

	Upgrades						Downgrades					
Rel. Day	AAR (i,t) (%)	P-value	SS	Mean R(i,t) (%)	P-value	SS	AAR(i,t) (%)	P-value	SS	Mean R(i,t) (%)	P-value	SS
-10	-0.075	0.164		0.046	0.293		-0.038	0.397		-0.031	0.427	
-9	-0.076	0.158		0.050	0.266		-0.211	0.049	**	-0.160	0.120	
-8	0.004	0.481		0.190	0.014	**	-0.354	0.005	***	-0.256	0.040	**
-7	-0.106	0.087	*	0.031	0.352		-0.161	0.125		-0.258	0.051	*
-6	-0.053	0.283		0.088	0.187		-0.168	0.155		-0.149	0.193	
-5	-0.043	0.291		0.107	0.111		-0.257	0.031	**	-0.281	0.028	**
-4	0.117	0.065	*	0.191	0.009	***	-0.294	0.021	**	-0.322	0.015	**
-3	0.189	0.021	**	0.251	0.005	***	-0.268	0.043	**	-0.304	0.035	**
-2	-0.083	0.138		0.021	0.397		0.129	0.243		0.096	0.304	
-1	-0.079	0.148		0.017	0.413		-0.352	0.015	**	-0.404	0.008	***
0	0.099	0.123		0.304	0.001	***	-0.237	0.059	*	-0.340	0.019	**
1	0.141	0.035	**	0.237	0.003	***	-0.366	0.026	**	-0.383	0.025	**
2	-0.203	0.007	***	-0.059	0.246		0.173	0.179		0.254	0.100	*
3	-0.162	0.027	**	-0.012	0.446		0.504	0.011	**	0.477	0.018	**
4	-0.034	0.320		0.090	0.114		0.297	0.026	**	0.181	0.115	
5	0.002	0.489		0.097	0.121		0.117	0.207		0.116	0.226	
6	-0.120	0.065	*	0.019	0.418		0.358	0.016	**	0.294	0.040	**
7	-0.037	0.325		0.060	0.241		0.311	0.022	**	0.191	0.118	
8	-0.095	0.100	*	0.017	0.413		0.104	0.225		-0.024	0.440	
9	-0.103	0.084	*	-0.032	0.350		-0.121	0.197		-0.186	0.101	
10	-0.014	0.431		0.034	0.359		-0.460	0.110		-0.515	0.073	*

Table 2

Cumulative Average Abnormal Returns Around Changes in Sovereign Debt Ratings

This table presents cumulative average abnormal returns in the sovereign stock index, in the period of twenty days before and after changes in sovereign debt ratings. Results are reported separately for upgrades and downgrades. In all panels we apply a first mover (FM) filter: FMR means that all observations preceded by changes in ratings by the same or other rating agency (Fitch, Moody's and Standard & Poor's) in the previous twenty trading days are deleted. FMRO is a first mover filter using ratings and outlooks in the previous twenty days. In panel A we show the "Ratings FMR" sample: the union of all changes in ratings, filtered using a ratings' FM filter. In panel B we show the "Ratings and Outlooks FMRO" sample: the union of changes in ratings and outlooks filtered by changes in ratings and outlooks. In panel C we show the "Ratings FMRO" sample: the union of all changes in ratings, filtered using a ratings and outlooks' FM filter. CAAR is the cumulative average abnormal return for the specified event window. P-values using the cross sectional method are reported.***,**, and * denote statistical significance (SS) at the 1, 5, and 10 percent level, respectively.

Panel A:	**"Ratings FMR"**							
	Upgrades				Downgrades			
Event Window	N	CAAR (%)	P-Value	SS	N	CAAR (%)	P-Value	SS
(-5,-3)	376	0.263	0.030	**	271	-0.818	0.002	***
(-5,-2)	376	0.180	0.127		271	-0.690	0.044	**
(-5,-1)	376	0.101	0.286		271	-1.041	0.003	***
(0,+1)	376	0.240	0.015	**	271	-0.602	0.010	***
(+2,+5)	376	-0.396	0.007	***	271	1.090	0.005	***
(-5,+5)	376	-0.055	0.425		271	-0.553	0.162	
Panel B:	**"Ratings & Outlooks FMRO"**							
	Upgrades				Downgrades			
Event Window	N	CAAR (%)	P-Value	SS	N	CAAR (%)	P-Value	SS
(-5,-3)	638	0.237	0.012	**	456	-0.411	0.026	**
(-5,-2)	638	0.199	0.056	*	456	-0.456	0.052	*
(-5,-1)	638	0.215	0.067	*	456	-0.749	0.003	***
(0,+1)	638	0.202	0.014	**	456	-0.478	0.009	***
(+2,+5)	638	-0.424	0.000	***	456	0.892	0.001	***
(-5,+5)	638	-0.007	0.487		456	-0.336	0.196	
Panel C:	**"Ratings FMRO"**							
	Upgrades				Downgrades			
Event Window	N	CAAR (%)	P-Value	SS	N	CAAR (%)	P-Value	SS
(-5,-3)	358	0.312	0.014	**	240	-0.658	0.017	**
(-5,-2)	358	0.184	0.126		240	-0.577	0.096	*
(-5,-1)	358	0.115	0.265		240	-0.791	0.025	**
(0,+1)	358	0.237	0.018	**	240	-0.583	0.016	**
(+2,+5)	358	-0.364	0.011	**	240	1.089	0.010	***
(-5,+5)	358	-0.012	0.484		240	-0.285	0.319	

Table 3

Event Study of Changes in Sovereign Ratings on Local Stock Indices (Robustness)

This table presents the event-study results under different specifications to show how changes in sovereign debt ratings affect the respective sovereign daily stock market returns. Results are reported separately for upgrades and downgrades. We show results for the "Ratings FMR" sample: the union of all changes in ratings by Fitch, Moody's and Standard & Poor's, filtered using other changes in ratings by the same or different rating agency, in the previous twenty trading days. Rel day is the trading day relative to the event day (day 0). AAR(i,t) is the average abnormal return for all observations i for each day t,using a world-CAPM model. The sample includes 376 upgrades and 271 downgrades. P-values (P-val) using four methods are reported: Cross-sect. is the cross sectional method; BW80 is the Brown and Warner (1980) method; BMP91 is the Boehmer et al. (1991) method; KP10 is the Kolari and Pynnonen (2010) method. ***,**, and * denote statistical significance (SS) at the 1, 5, and 10 percent level, respectively.

	Upgrades									Downgrades								
Rel Day	AAR (i,t) (%)	P-val (Cross-sect.)	SS	P-val (BW80)	SS	P-val (BMP91)	SS	P-val (KP10)	SS	AAR (i,t) (%)	P-val (Cross-sect.)	SS	P-val (BW80)	SS	P-val (BMP91)	SS	P-val (KP10)	SS
-10	-0.075	0.164		0.197		0.144		0.142		-0.038	0.397		0.383		0.386		0.391	
-9	-0.076	0.158		0.192		0.364		0.363		-0.211	0.049	**	0.049	**	0.087	*	0.097	*
-8	0.004	0.481		0.484		0.372		0.371		-0.354	0.005	***	0.003	***	0.042	**	0.049	**
-7	-0.106	0.087	*	0.115		0.181		0.180		-0.161	0.125		0.103		0.037	**	0.044	**
-6	-0.053	0.283		0.273		0.227		0.225		-0.168	0.155		0.093	*	0.069	*	0.079	*
-5	-0.043	0.291		0.313		0.221		0.219		-0.257	0.031	**	0.022	**	0.017	**	0.021	**
-4	0.117	0.065	*	0.092	*	0.098	*	0.097	*	-0.294	0.021	**	0.010	**	0.009	***	0.012	**
-3	0.189	0.021	**	0.016	**	0.042	**	0.041	**	-0.268	0.043	**	0.018	**	0.058	*	0.066	*
-2	-0.083	0.138		0.174		0.083	*	0.081	*	0.129	0.243		0.156		0.029	**	0.035	**
-1	-0.079	0.148		0.185		0.122		0.121		-0.352	0.015	**	0.003	***	0.004	***	0.006	***
0	0.099	0.123		0.130		0.062	*	0.061	*	-0.237	0.059	*	0.031	**	0.020	**	0.025	**
1	0.141	0.035	**	0.054	*	0.021	**	0.020	**	-0.366	0.026	**	0.002	***	0.006	***	0.008	***
2	-0.203	0.007	***	0.011	**	0.028	**	0.027	**	0.173	0.179		0.087	*	0.210		0.220	
3	-0.162	0.027	**	0.033	**	0.087	*	0.086	*	0.504	0.011	**	0.000	***	0.051	*	0.059	*
4	-0.034	0.320		0.350		0.373		0.373		0.297	0.026	**	0.010	***	0.030	**	0.036	**
5	0.002	0.489		0.491		0.387		0.387		0.117	0.207		0.178		0.474		0.476	
6	-0.120	0.065	*	0.086	*	0.026	**	0.025	**	0.358	0.016	**	0.002	***	0.005	***	0.007	***
7	-0.037	0.325		0.338		0.335		0.334		0.311	0.022	**	0.007	***	0.043	**	0.050	*
8	-0.095	0.100	*	0.139		0.119		0.118		0.104	0.225		0.207		0.211		0.222	
9	-0.103	0.084	*	0.120		0.330		0.328		-0.121	0.197		0.171		0.096	*	0.106	
10	-0.014	0.431		0.437		0.401		0.401		-0.460	0.110		0.000	***	0.084	*	0.094	*

Table 4

Cumulative Average Abnormal Returns Around Downgrades in Sovereign Debt Ratings by Country Characteristics

This table presents cumulative abnormal returns in the sovereign stock market index, in the period of twenty days before and after downgrades in sovereign debt ratings. We show results for the downgrades in the "Ratings FMR" sample: the union of all changes in ratings by Fitch, Moody's and Standard & Poor's, filtered using other changes in ratings by the same or different rating agency, in the previous twenty trading days. CAAR[t_1,t_2] is the cumulative average abnormal return for the period staring on t_1 and ending at t_2 relative to event day (day 0). We examine CAARs separately for each of the six categories: Civil Law (vs. Common Law) is shown on the first (second) row; Emerging & Frontier (vs. Developed) is shown on the first (second) row; TI Corruption Index (Low vs. High; low score - first row - implies low institutional quality); PRS Law & Order (Low vs. High; low score - first row - implies low institutional quality); PRS Corruption (Low vs. High; low score - first row - implies low institutional quality); Investor Protection (Low vs. High; low score - first row - implies low institutional quality). The separation of each category is made at the median value of the "Ratings FMR" sample. N is the number of observations in each subcategory. P-values using the cross sectional method (equation 7) are reported.***,**, and * denote statistical significance at the 1, 5, and 10 percent level, respectively.

	Civil/Common Law				Emerging & Frontier / Developed				TI Corruption (Low/High)				PRS Law & Order (Low/High)				PRS Corruption (Low/High)				Investor Protection (Low/High)			
Event Window	N	CAAR (%)	P-Val	SS	N	CAAR (%)	P-Val	SS	N	CAAR (%)	P-Val	SS	N	CAAR (%)	P-Val	SS	N	CAAR (%)	P-Val	SS	N	CAAR (%)	P-Val	SS
(-5,-3)	203	-1.18%	0.000	***	180	-1.04%	0.004	***	133	-1.61%	0.000	***	125	-1.52%	0.000	***	112	-0.61%	0.027	**	153	-1.03%	0.004	***
	68	0.27%	0.344		91	-0.38%	0.167		130	-0.10%	0.404		127	-0.22%	0.306		139	-1.08%	0.016	**	118	-0.54%	0.104	
(-5,-2)	203	-1.14%	0.002	***	180	-0.80%	0.061	*	133	-1.51%	0.001	***	125	-1.38%	0.004	***	112	-0.38%	0.177		153	-0.95%	0.035	**
	68	0.65%	0.267		91	-0.47%	0.228		130	0.08%	0.452		127	-0.14%	0.417		139	-1.06%	0.067	*	118	-0.36%	0.288	
(-5,-1)	203	-1.40%	0.001	***	180	-1.31%	0.004	***	133	-1.80%	0.001	***	125	-1.62%	0.002	***	112	-0.83%	0.041	**	153	-1.29%	0.010	**
	68	0.03%	0.484		91	-0.50%	0.203		130	-0.39%	0.239		127	-0.58%	0.160		139	-1.31%	0.020	**	118	-0.71%	0.079	*
(0,+1)	203	-0.61%	0.025	**	180	-0.87%	0.004	***	133	-1.10%	0.003	***	125	-0.84%	0.018	**	112	-0.66%	0.012	**	153	-0.78%	0.026	**
	68	-0.58%	0.091	*	91	-0.08%	0.422		130	-0.10%	0.385		127	-0.39%	0.148		139	-0.58%	0.091	*	118	-0.37%	0.091	*
(+2,+5)	203	1.28%	0.008	***	180	1.54%	0.006	***	133	1.54%	0.001	***	125	1.73%	0.020	**	112	0.77%	0.048	**	153	1.21%	0.042	**
	68	0.51%	0.178		91	0.20%	0.276		130	0.61%	0.199		127	0.51%	0.074	*	139	1.38%	0.031	**	118	0.93%	0.006	***
(-5,+5)	203	-0.73%	0.129		180	-0.64%	0.209		133	-1.35%	0.029	**	125	-0.73%	0.225		112	-0.72%	0.157		153	-0.86%	0.150	
	68	-0.03%	0.490		91	-0.38%	0.262		130	0.12%	0.445		127	-0.46%	0.257		139	-0.51%	0.290		118	-0.15%	0.414	

Table 5

Selection of Instrumental Variables

Results of the selection of most appropriate instrumental variables for the endogenous regressors approximating institutional quality. The four instrumental variables tested for each of the five endogenous variables are: Common vs. Civil Law (La Porta et al., 1998); Ethnicity fractionalization (Alesina et al., 2003); Religion fractionalization (Alesina et al., 2003); a landlocked indicator (1 if landlocked; 0 otherwise). The Null Hypothesis tested is "Instruments are redundant". We report robust test statistics estimated using (Baum et al., 2010), which are distributed according to a chi-squared distribution with degrees of freedom equal to the product of the number of endogenous regressors (1) and the numbers of instruments tested (total number of observations: 273). The procedure begins with the four instruments listed below, and is repeated successively until all redundant instruments are eliminated. The final list of instrumental variables for each endogenous regressor is determined in Round 2.

Endogenous	Common/Civil Law			Ethnicity			Religion			Landlocked		
(down)	Test Stat	P-Val	SS	Test Stat	P-Val	SS	Test Stat	P-Val	SS	Test Stat	P-Val	SS
Round 1												
Emerging/Frontier	14.853	0.000	***	35.776	0.000	***	4.091	0.043	**	17.500	0.000	***
TI Corruption	20.302	0.000	***	61.256	0.000	***	11.127	0.001	***	6.142	0.013	**
PRS Law & Order	20.095	0.000	***	59.970	0.000	***	1.233	0.267		0.510	0.475	
PRS Corruption	2.533	0.115		0.121	0.728		0.050	0.822		12.188	0.001	***
Investor Protection	36.803	0.000	***	5.851	0.016	**	0.249	0.618		7.393	0.007	***
Round 2												
Emerging/Frontier	19.428	0.000	***	34.903	0.000	***				15.001	0.000	***
TI Corruption	24.023	0.000	***	61.347	0.000	***	10.374	0.001	***			
PRS Law & Order	22.968	0.000	***	60.019	0.000	***						
PRS Corruption										a		
Investor Protection	68.449	0.000	***							4.605	0.032	**

a: This variable only has one IV.

Table 6

Two Stage Least Squares Regression of *Pre-Event* Stock Market Reaction on Institutional Quality

This table presents two-stage least square (2SLS) regressions on the cumulative abnormal returns in the local stock market index, in the period starting five days and ending three days (t = -5 to t =-3) *before* downgrades in sovereign debt ratings. In Panel A we show downgrades from the "Ratings FMR" sample (described in Table 2). Instruments (description in Table 5) used for "Emerging & Frontier" are Common/Civil Law, Ethnic fractionalization and landlocked. Instruments used for "TI Corruption" are Common/Civil Law, Ethnic fractionalization and Religion fractionalization. Instruments used for "PRS Law and Order" are Common/Civil Law and Ethnic fractionalization. The instrument used for "PRS Corruption" is landlocked. Instruments used for Investor Protection" are Common/Civil Law and landlocked. "Exp Sign" is the expected sign of the regression coefficient ("Coeff."). "Z" and "P-val" are the robust z-value and p-value of the coefficient. UID is the under-identification test, which reports the Kleibergen-Paap rk LM statistic and associated chi-square p-value. OID is the over-identification test, which reports the Hansen J Statistic and associated chi-square p-value. WID is the weak-identification test reports both the Cragg-Donald Wald F-statistic and also the Kleibergen-Paap rk Wald F statistic. In the last two rows we report the Stock-Yogo weak ID test critical values (10% maximal) for IV relative bias (Rel. Bias) and Size, respectively. In panel B ("Ratings and Outlooks FMRO" sample) and C ("Ratings FMRO" sample), we repeat the IV selection process in Table 5 and re-run the analysis.

Panel A: "Ratings FMR"	**N**	**Exp Sign**	**Coeff.**	**Z**	**P-val**	**SS**	**Coeff.**	**Z**	**P-val**	**SS**	**Coeff.**	**Z**	**P-val**	**SS**	**Coeff.**	**Z**	**P-val**	**SS**	**Coeff.**	**Z**	**P-val**	**SS**
Intercept			0.010	1.300	0.195		-0.038	-4.130	0.000	***	-0.043	-2.970	0.003	***	-0.019	-0.580	0.563		-0.051	-2.190	0.029	**
Emerging/Frontier	271	-	-0.027	-2.580	0.010	**																
TI Corruption	263	+					0.006	3.430	0.001	***												
PRS Law & Order	252	+									0.008	2.490	0.013	**								
PRS Corruption	251	+													0.003	0.310	0.753					
Inv. Protection	271	+																	0.007	1.810	0.070	*
				Stat	**P-val**			**Stat**	**P-val**			**Stat**	**P-val**			**Stat**	**P-val**			**Stat**	**P-val**	
UID (Kleibergen-Paap rk LM)				48.651	0.000	***		67.620	0.000	***		60.100	0.000	***		11.079	0.001	***		50.842	0.000	***
OID (Hansen J)				4.224	0.121	a		1.275	0.529			2.223	0.136			b				1.882	0.170	
WID (Kleibergen-Paap rk Wald F)				54.464				71.555				76.497				27.978				53.676		
WID (Cragg-Donald Wald F)				22.895				41.430				69.679				9.287				51.828		
Stock-Yogo WID 10% Rel. Bias				9.080				9.080				19.930				16.380				19.930		
Stock-Yogo WID 10% Size				22.300				22.300				n/a				n/a				n/a		

Panel B: "Ratings & Outlooks FMRO"		Exp Sign	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS
Intercept			0.012	1.900	0.058	*	-0.027	-3.350	0.004	***	-0.175	-2.040	0.047	**	0.003	0.110	0.908		-0.038	-2.420	0.016	**
Emerging/Frontier	449	-	-0.024	-2.750	0.006	***																
TI Corruption	437	+					0.005	2.870	0.001	***												
PRS Law & Order	413	+									0.041	1.990	0.041	**								
PRS Corruption	409	-													-0.003	-0.300	0.766					
Inv. Protection	449	+																	0.006	2.100	0.036	**
				Stat	P-val			Stat	P-val			Stat	P-val			Stat	P-val			Stat	P-val	
UID (Kleibergen-Paap rk LM)				59.864	0.000			97.503	0.000			3.873	0.000			14.398	0.006			110.020	0.000	
OID (Hansen J)				4.776	0.189			3.750	0.290					c						4.298	0.117	
WID (Kleibergen-Paap rk Wald F)				16.850				62.083				7.086				37.418				113.604		
WID (Cragg-Donald Wald F)				26.366				42.647				7.364				12.310				78.093		
Stock-Yogo WID 10% Rel. Bias				10.270				10.270				19.930				16.380				9.080		
Stock-Yogo WID 10% Size				24.580				24.580				n/a				n/a				22.300		

Panel C: "Ratings FMRO"		Exp Sign	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS
Intercept			0.009	1.230	0.219		-0.031	-3.370	0.000	***	-0.021	-2.190	0.028	**	-0.021	-0.690	0.493		-0.054	-2.170	0.030	**
Emerging/Frontier	240	-	-0.023	-2.510	0.012	**																
TI Corruption	232	+					0.005	2.770	0.006	***												
PRS Law & Order	222	+									0.006	1.790	0.073	*								
PRS Corruption	222	-													0.005	0.480	0.633					
Inv. Protection	240	+																	0.008	1.870	0.062	*
				Stat	P-val			Stat	P-val			Stat	P-val			Stat	P-val			Stat	P-val	
UID (Kleibergen-Paap rk LM)				45.748	0.000			61.560	0.000			52.874	0.000			7.523	0.006			46.178	0.000	
OID (Hansen J)				5.387	0.146			3.378	0.337			2.689	0.101			0.000	0.000			3.250	0.071	
WID (Kleibergen-Paap rk Wald F)				57.135				51.764				69.194				19.508				66.025		
WID (Cragg-Donald Wald F)				16.892				27.699				60.967				6.768				44.670		
Stock-Yogo WID 10% Rel. Bias				10.270				10.270				19.930				16.380				9.080		
Stock-Yogo WID 10% Size				24.580				24.580				n/a				n/a				22.300		

a: Hansen J statistic was originally significant at 5%, hence Ethnic was dropped.
b: PRS Corruption is "just identified" hence the Hansen J statistic is not defined.
c: Hansen J statistic was originally significant at 5%, hence Ethnic was dropped.

Table 7

Two Stage Least Squares Regression of *Post-Event* Stock Market Reaction on Institutional Quality

This table presents two-stage least square (2SLS) regressions on the cumulative abnormal returns in the local stock market index, in the period starting two days and ending five days (t = +2 to t =+5) *after* downgrades in sovereign debt ratings. We show downgrades from the "Ratings FMR" sample as described in Table 2. Instruments (description in Table 5) used for "Emerging & Frontier" are Common/Civil Law, Ethnic fractionalization and landlocked. Instruments used for "TI Corruption" are Common/Civil Law, Ethnic fractionalization and Religion fractionalization. Instruments used for "PRS Law and Order" are Common/Civil Law and Ethnic fractionalization. The instrument used for "PRS Corruption" is landlocked. Instruments used for Investor Protection" are Common/Civil Law and landlocked. "Z" and "P-val" are the robust z-value and p-value of the coefficient. UID is the under-identification test, which reports the Kleibergen-Paap rk LM statistic and associated chi-square p-value. OID is the over-identification test, which reports the Hansen J Statistic and associated chi-square p-value. WID is the weak-identification test reports both the Cragg-Donald Wald F-statistic and also the Kleibergen-Paap rk Wald F statistic. In the last two rows we report the Stock-Yogo weak ID test critical values (10% maximal) for IV relative bias (Rel. Bias) and Size, respectively. There are 271 observations in each regression.

"Ratings FMR"	N	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS	Coeff.	Z	P-val	SS
Intercept		0.000	0.030	0.977		0.032	2.350	0.019	**	0.033	1.790	0.074	*	0.003	0.070	0.943		0.035	1.340	0.179	
Emerging/Frontier	271	0.016	1.170	0.244																	
TI Corruption	263					-0.005	-1.780	0.075	*												
PRS Law & Order	252									-0.005	-1.240	0.215									
PRS Corruption	251													0.003	0.200	0.838					
Inv. Protection	271																	-0.004	-0.990	0.321	
			Stat	**P-val**			**Stat**	**P-val**			**Stat**	**P-val**			**Stat**	**P-val**			**Stat**	**P-val**	
UID (Kleibergen-Paap rk LM)			48.65	0.000	***		67.62	0.000	***		60.10	0.000	***		11.08	0.001	***		50.84	0.000	***
OID (Hansen J)			0.41	0.815			2.58	0.275			0.48	0.487			-	-			0.10	0.755	
WID (Kleibergen-Paap rk Wald F)			54.46				71.56				76.50				27.98				53.68		
WID (Cragg-Donald Wald F)			22.90				41.43				69.98				9.29				51.83		
Stock-Yogo WID 10% Rel. Bias			9.08				9.08				19.93				16.38				19.93		
Stock-Yogo WID 10% Size			22.30				22.30				n/a				n/a				n/a		