Credit Ratings and Cross-Border Bond Market Spillovers

Böninghausen, Benjamin and Zabel, Michael

University of Munich (LMU), University of Munich (LMU)

June 2013

Credit Ratings and Cross-Border Bond Market Spillovers*

Benjamin Böninghausen† Michael Zabel‡

June 4, 2013

Abstract

This paper studies spillovers across sovereign debt markets in the wake of sovereign rating changes. To this end, we use an extensive dataset covering all announcements by the three major agencies (Standard & Poor's, Moody's, Fitch) and daily sovereign bond market movements of up to 73 developed and emerging countries between 1994 and 2011. On the basis of an explicit counterfactual and controlling for important dimensions of the announcement environment, we find asymmetric reactions to upgrades and downgrades. While there is strong evidence of negative spillover effects in response to sovereign downgrades, positive spillovers from upgrades are much more limited. Our results also suggest that negative spillover effects are more pronounced for countries within the same region. This does not appear to be due to (measurable) fundamental linkages and similarities, such as trade, which turn out to be strikingly insignificant.

JEL classification: G15, F36

Keywords: Sovereign debt market, credit rating agencies, cross-border spillover effects, international financial integration

*We are grateful for valuable comments and suggestions by seminar participants at LMU Munich. Benjamin Böninghausen gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft through GRK 801.

†Munich Graduate School of Economics, Kaulbachstraße 45, 80539 Munich, Germany. *E-mail*: benjamin.boeninghausen@lrz.uni-muenchen.de

‡University of Munich, Seminar for Macroeconomics, Ludwigstraße 28 (rear building), 80539 Munich, Germany. *E-mail*: michael.zabel@lrz.uni-muenchen.de

1 Introduction

Ever since tensions began to surface in the eurozone in late 2009, the announcements by credit rating agencies (CRAs) on the creditworthiness of member states have continuously made the headlines and rattled financial markets. In particular, while not specific to the ongoing crisis, the notion that rating actions pertaining to one country might have a major impact on the yields of other countries' sovereign bonds, too, has regained the attention of policymakers. In fact, concerns over so-called negative spillover effects seem to be running deep, which is probably best exemplified by reports that the European Commission was at one stage considering a temporary restriction on the issuance of ratings under exceptional circumstances (Financial Times, 2011). This begs two interesting, wider research questions. First, when a rating announcement is made for a given country, do we generally observe significant spillover effects on other countries' bond markets? Second, if so, under which conditions are those effects strongest, and which countries are affected most?

In this paper, we address these questions based on an extensive dataset comprising both a complete history of sovereign rating actions by the "Big Three" (Standard & Poor's, Moody's and Fitch) and daily sovereign bond market movements for up to 73 countries between 1994 and 2011. Not only do we cover both crisis and non-crisis periods and a broad set of developed and emerging countries across all continents. Our dataset also contains sufficient variation to run an explicit counterfactual analysis which pits small revisions in an agency's assessment of a country's creditworthiness against all other changes. Moreover, it characterises along several important dimensions the environment in which a rating change is made. For instance, we are crucially able to account for the fact that an announcement is often followed by a similar one from a different agency soon after, which may influence the reception of the later announcements.[1]

Our main findings can be summarised as follows. Whereas we find strong evidence for the existence of significant cross-border spillover effects of sovereign rating downgrades, reactions to upgrades appear to be, if anything, much more muted. This points to an important asymmetry in the sovereign debt market's treatment of ratings. Regarding the influence of country characteristics, we find that negative spillover effects tend to be more pronounced for countries within the same region,

[1] To the best of our knowledge, our investigation is the first to consider such interactions between the major CRAs in identifying spillover effects.

a finding that persists even after controlling for measurable fundamental links and similarities between countries, which are strikingly insignificant.

Our paper is related to a broad strand of literature that investigates the effects of sovereign rating announcements on different segments of the financial markets. The most common exercise is to conduct an event study gauging the *direct* impact of rating changes on the bonds issued by the country concerned. However, there is also a substantial body of research analysing the reaction of the country's stock and, more recently, of its CDS market. As a general result, this literature finds a strong and significant impact of sovereign rating downgrades, while upgrades have an insignificant or more limited impact (see, eg, Cantor and Packer, 1996; Larraín et al., 1997; Reisen and von Maltzan, 1999; Brooks et al., 2004; Hooper et al., 2008; Hill and Faff, 2010).

Moreover, in recent years a growing body of research has specifically studied whether sovereign rating changes also lead to *spillover* effects on other countries' sovereign bonds. Generally speaking, the literature affirms the existence of such spillovers, meaning that a rating action on one country is found to significantly affect the sovereign bond prices of other countries (eg, Arezki et al., 2011; De Santis, 2012; Ismailescu and Kazemi, 2010). Some studies also point out that spillovers are not limited to sovereign debt markets but that rating changes also affect foreign stock and exchange markets (Kaminsky and Schmukler, 2002; Arezki et al., 2011; Alsakka and ap Gwilym, 2012). Regarding a potential asymmetry in the spillover effects of negative and positive rating events, the results of the literature so far remain inconclusive. Whereas Afonso et al. (2011) find spillovers to matter most for downgrades, with little or no effects of sovereign upgrades, Ismailescu and Kazemi (2010) find positive rating events to have a greater spillover effect on foreign CDS prices than negative ones.

Most of these studies focus either on spillover effects during specific regional crisis episodes (see Afonso et al. (2011), Arezki et al. (2011) and De Santis (2012) for the eurozone crisis; Kaminsky and Schmukler (1999) for the 1997/98 Asian crisis) or on an otherwise homogeneous sample of countries only, such as emerging countries (Ismailescu and Kazemi, 2010; Kaminsky and Schmukler, 2002). This leaves open the question to what extent their findings are specific to the episode analysed.

The study most closely related to our work is Gande and Parsley (2005). Based on a sample of 34 developed and emerging economies from 1991 to 2000, they find evidence for the existence of spillover effects. In line with Afonso et al. (2011), they report asymmetries in the effects of upgrades and downgrades, with the latter triggering large and significant spillovers but no discernible impact of upgrades. Analysing the transmission mechanisms of these spillovers, they find trade and financial linkages to be the most relevant, and physical proximity and cultural or institutional linkages to be of little importance.

Our study differs from theirs in several ways. Firstly, it is based on a larger and more up-to-date sample of up to 73 countries for the 1994–2011 period. Secondly, their dataset only includes information of ratings by Standard & Poor's (S&P), which risks biasing estimation results since, as we demonstrate, a rating action by one agency is often foreshadowed or closely followed by a similar action from another rating agency. Ignoring the informational content of rating actions by other agencies may therefore prove to be problematic. Thirdly, their identification of spillovers uses a "comprehensive credit rating", which combines actual rating changes and credit watch, or review, changes into a single 17-notch scale. Hence, they make additional assumptions on the relative informational content of reviews and ratings. In this study, we exploit the variation in our sample to focus solely on the class of actual rating changes and their relative strength as the only observable difference between them. At the same time, we do allow for differences in the informational content of sovereign upgrades and downgrades by controlling for watchlistings that may build anticipation by foreshadowing actual future rating changes. Finally, due to lack of data, Gande and Parsley (2005) only use an indirect measure of trade integration. Relying on bilateral data, we find trade to be remarkably unrelated to the strength of spillovers.

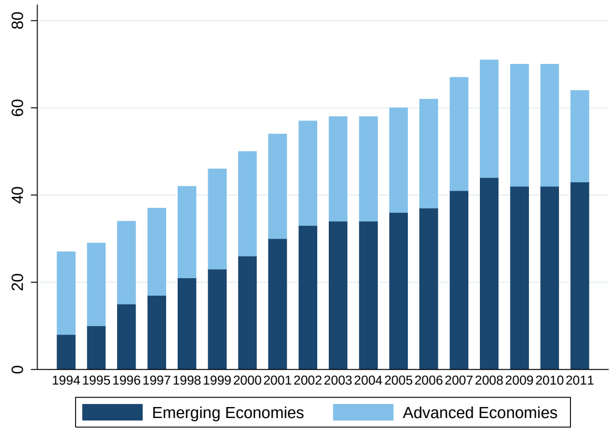
The paper is organised as follows. In the next section, we describe the dataset and highlight some important characteristics of rating announcements. Section 3 discusses the estimation strategy for identifying cross-border spillovers. Section 4 presents our empirical results and discusses their interpretation. We end with a brief conclusion.

2 Data

2.1 The dataset

For our study, we compile a broad dataset of the yields of publicly traded sovereign bonds at daily frequency. The dataset starts in January 1994 and ends in December

Figure 1: **Number of sovereign bonds in dataset**

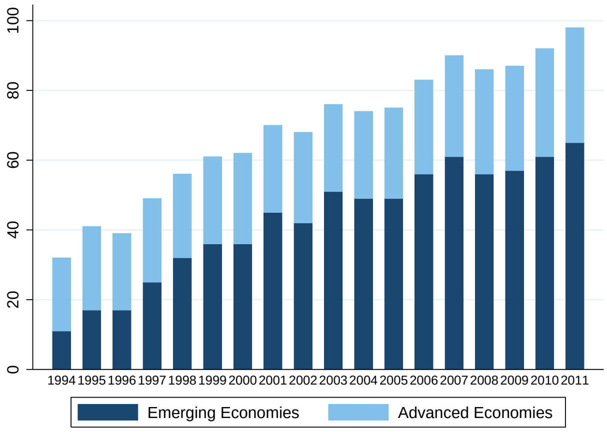


Notes — This figure shows the scope and composition, by economic development, of the sovereign bond sample between 1994 and 2011, highlighting a notable increase in the coverage of emerging economies over time. Countries are classified according to the IMF World Economic Outlook.

2011. Since for many countries data are only available after 1994, we add those countries' sovereign bonds as soon as reliable information becomes available. Whereas our dataset only comprises sovereign bonds issued by 27 countries in 1994, this number increases to 74 countries towards the end of our sample period. This reflects both the increased financing needs of sovereigns and the growing prevalence of bond issuance, as opposed to bank financing, during the last 20 years. While for 1994 sovereign bond yields are mostly available for developed countries, the availability of emerging market bond yields picks up heavily over our sample period. Towards the end of the period, emerging markets even account for the bulk of sovereign bonds in the sample. Figure 1 illustrates the increasing scope of our dataset over time.

In order to consider a broad spectrum of sovereign bonds, our sample draws on data from different sources. Our preferred data source is Bloomberg, from which we use generic 10-year yields for up to 33 countries. If data are not available on Bloomberg, we supplement them with yields from Datastream's 10-year Government Bond Benchmark Index, ensuring that this does not induce structural breaks in the series. Since sovereign bond availability for emerging markets is quite limited both on Bloomberg and on Datastream, we also use data from the JP Morgan Emerging Markets Bond Index Global (henceforth EMBI Global, see JP Morgan, 1999). While bonds included in the EMBI Global have to fulfil strict requirements regarding the availability of reliable daily prices, the average maturity of a country's bond index can vary remarkably from that of the other two sources. We therefore control for

Figure 2: **Number of rated countries**



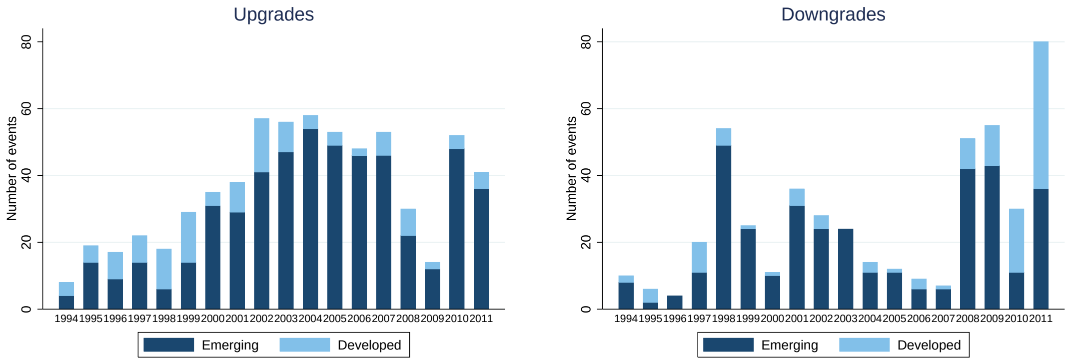
Notes — This figure shows the scope and composition, by economic development, of the sample of countries rated by at least one of the major rating agencies (S&P, Moody's, Fitch) between 1994 and 2011, with a notable increase in the coverage of emerging economies over time. Countries are classified according to the IMF World Economic Outlook.

maturity in all regressions. Table A.1 in the Appendix gives a detailed overview of the sovereign bond market data included in our sample.

For the purpose of our later analysis, we compute sovereign bond *spreads*. The spread is the differential of the country's sovereign bond yield over that of a US Treasury bond of comparable maturity. We use 10-year maturities where possible, which is the case for the developed economies and some emerging markets. For the other emerging economies, we rely on the EMBI Global data. As those correspond to different maturities (depending on the average maturity of eligible instruments a country has issued), we obtain the relevant US Treasury yields by interpolating from the closest published yield curve rates.

Information on sovereign ratings comes from the rating agencies' websites and includes daily information both on rating changes and on sovereign watchlistings by any of the "Big Three" (S&P, Moody's, Fitch) from 1994 to 2011. We choose the year 1994 as a natural starting point for our sample period since Fitch only started to assign sovereign ratings in that year. Like the number of publicly traded sovereign bonds, the scope and composition of countries rated by the "Big Three" changes quite substantially during our sample period. While in 1994 only 34 sovereigns were rated by at least one of the agencies, this number increased to 98 countries in 2011 (see Figure 2).

Figure 3: **Rating actions over time**



Notes — This figure shows upgrades and downgrades of developed and emerging economies made by S&P, Moody's and Fitch between 1994 and 2011. Countries are classified according to the IMF World Economic Outlook.

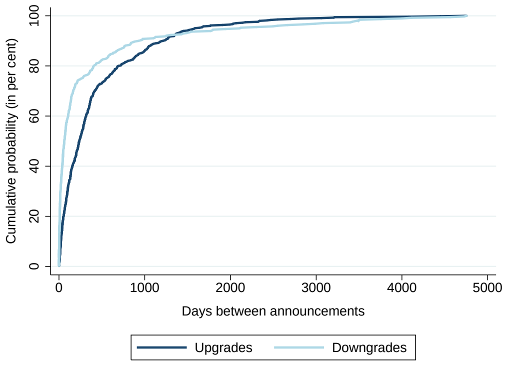
2.2 Characteristics of rating announcements

Over the whole sample period, we are able to consider a total of 1,097 rating changes, of which 635 were upgrades and 462 downgrades. In general, one can observe a significant increase in the number of sovereign credit ratings during our sample period, particularly in emerging market countries.

As Figure 3 illustrates, rating activity is not evenly distributed over time but, especially for downgrades, shows some hefty peaks during specific episodes of crisis. Whereas in "normal times", downgrades tend to be relatively scarce, a severe increase can be observed in the context of the 1997/98 Asian crisis (affecting mostly emerging countries plus South Korea and Hong Kong) and following the 2008–2011 financial and European debt crises (where for the first time advanced economies have been exposed to downgrades at a large scale). This means that similar announcements tend to cluster around certain time periods.

In addition, it is an important stylised fact that the downgrading of a country is frequently followed by yet another downgrade announcement for that same country soon after. This is all the more probable because there is a strong overlap in country coverage by the "Big Three". Almost all countries in our sample are rated by more than one agency only and most are even rated by all three (70 out of 98 countries at the end of 2011). Hence, in what we term *within*-clustering, different agencies may make the same announcement for a *given country* in short succession or even on the same day. Figure 4 illustrates this issue by plotting the cumulative distribution function and summary statistics of the number of days between simi-

Figure 4: **Clustering of rating announcements**



	Upgrades	Downgrades
Mean	453	364
Median	238	63
5th pct	13	1
10th pct	23	3
25th pct	79	12

Notes — This figure shows the cumulative distribution functions and summary statistics of the number of calendar days between an upgrade (downgrade) announcement for a given country and a subsequent upgrade (downgrade) of the same country by any agency. Information is based on the sample of 1,097 rating announcements (635 upgrades, 462 downgrades) made by S&P, Moody's and Fitch between 1994 and 2011.

lar rating actions on the same country. As can be seen, clustering is particularly pronounced for downgrades. In around five per cent of all cases, a downgrade on a country is followed by another downgrade on that country within just one day. For example, in the course of the Asian crisis, S&P, Fitch and Moody's all downgraded South Korea's credit rating on successive days between 25 and 27 November 1997. Similarly, during the ongoing European debt crisis, Fitch issued a downgrade for Greece on 8 December 2009. One week later, S&P downgraded the country as well, as did Moody's yet another six days later.

The presence of clustering might be of crucial importance when examining the spillover effects from a rating announcement since its informational content is likely to vary depending on whether it has been announced in isolation or just a few days after (or even on the same day as) a similar announcement by another agency. Not to control for these cases could seriously bias estimation results for the impart of rating announcements on sovereign bond markets.

Clustering *across* countries may matter, too. When CRAs change the rating of a number of *different countries* in the same direction simultaneously, one needs to control for the fact that some countries will then be both "non-event" and event countries. Otherwise, one might erroneously detect spillovers across sovereign bond markets when, in fact, one is looking at a spillover in ratings. This is all the more important if the countries concerned share a common trait of some form which leads CRAs to make simultaneous announcements for the countries concerned in the first place, as appears to have happened on 3 October 2008 when Fitch downgraded Estonia, Latvia and Lithuania.[2] Is is therefore a major advantage of our dataset that it enables us to explicitly take into account prior and parallel rating actions by other CRAs and on other countries. This should provide for a particularly clean identification of relevant effects, which compares favourably with Gande and Parsley (2005) and other work in the sovereign spillovers literature (eg, Christopher et al., 2012; Ismailescu and Kazemi, 2010).

Similarly, the informational content of a rating change might be conditional on whether it has been preceded by the respective country being put on a watch-list. As the literature on the effects of rating announcements on the refinancing conditions of the very same country shows (eg, Afonso et al., 2011; Ismailescu and Kazemi, 2010), rating changes are often preceded by a similar change in the market's assessment of sovereign risk, especially when countries have been put "on watch", or "review", before.[3] Ignoring these anticipation effects risks underestimating bond market reactions to a sovereign rating action. Since our dataset includes all sovereign watchlistings by the "Big Three", we can directly control for a country's watchlist status and mitigate potential problems with anticipation.

3 Identifying sovereign spillovers

3.1 Counterfactual choice and estimation strategy

The existence of rating spillover effects in the sovereign debt market requires, by definition, that the announcement by a CRA on the creditworthiness of one country (*event country*) impacts significantly on the bond yields of another (*non-event country*). Yet, the mere observation of a change in non-event country yields when an event-country announcement is made does not suffice to establish a causal relation because non-event country yields might have changed regardless. Hence, the key issue in identifying potential spillover effects is to find a suitable counterfactual.

We cannot apply the procedure traditionally used in event studies on *direct* announcement effects, however. This strand of literature focuses on, for instance, the

[2] Other examples may be seen in S&P's downgrade announcements for South Korea and Taiwan during the Asian crisis on 24 October 1997, or in Fitch lowering the ratings of Estonia, Ireland, Latvia and Lithuania on 8 April 2009.

[3] In the following, we use the two terms interchangeably. While S&P and Fitch issue watchlistings, in the Moody's terminology those are called "reviews".

bond yield response of a sovereign that has been downgraded. In this framework, effects are identified by the existence of abnormal returns, meaning that around the announcement (event window), returns are significantly different from normal, as estimated over a longer time frame before the announcement (estimation window). In order to be a reasonable guide to normal returns, the estimation window has to be chosen such that other events with a potentially significant impact on returns are excluded (see, eg, MacKinlay, 1997). In other words, the counterfactual for gauging the impact of rating announcements is "no rating change". While this represents a challenge in direct announcement studies already, which focus on countries in isolation, the identification of *spillover* effects based on this counterfactual is essentially impossible.

The reason is that, in a spillover context, we would require that there are no announcements on *any* rated country within the estimation window.[4] There is obviously a trade-off between the length of that window and the number of announcements eligible for inclusion in the estimation. However, even at a 30-day length commonly used in sovereign event studies, which is towards the shorter end of the event-study literature more generally, only 23 upgrades would be eligible, and 36 downgrades.

We therefore pursue an identification strategy that does not rely on "no rating change at all" as its counterfactual, but which discriminates between rating changes according to their severity. More precisely, rating changes of a single notch serve as the counterfactual for more severe changes of two notches or more.[5] This approach is implemented in the following estimation equation, which we run on upgrades and downgrades separately:[6]

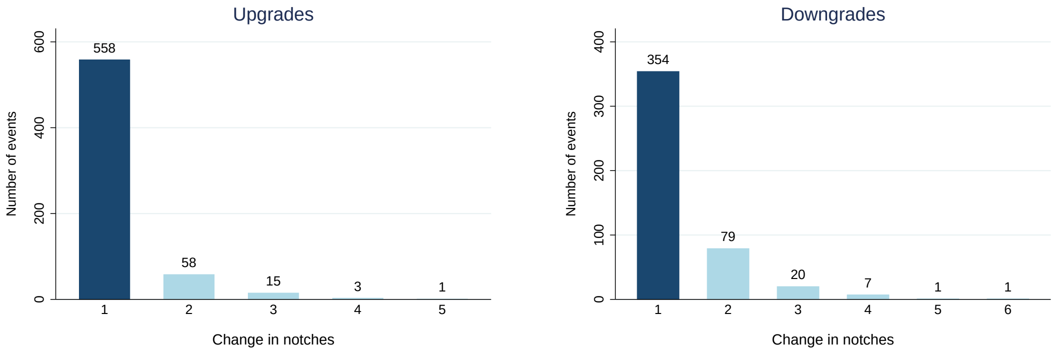
$$\Delta Spread_{n,t} = \alpha + \beta \cdot LARGE_{e,t} + RatEnv_{e,n,t} \cdot \gamma + Other_{e,n,t} \cdot \delta + \omega_{e,n,t} \,.$$

[4]The universe of all rated countries is the relevant benchmark when analysing potential spillover effects in this framework. Of course, if we only required the estimation window to be free of announcements pertaining to the non-event country, the number of events eligible for inclusion would increase substantially. However, this would amount to assuming from the outset that only direct effects, as opposed to spillover effects, could possibly matter, which would defy the purpose of the investigation.

[5]See Table A.2 in the Appendix on the mapping of CRAs' letter ratings into a linear 17-notch scale.

[6]This is in keeping with the methodology employed by the bulk of the spillover and event-study literature. The asymmetries in reactions to positive and negative announcements consistently found therein further reinforce the notion of treating upgrades and downgrades distinctly.

Figure 5: **Distribution of rating changes**



Notes — This figure shows the distribution of the severity of rating changes, measured on a 17-notch scale (see Table A.2 in the Appendix). Numbers are based on the sample of 1,097 rating announcements (635 upgrades, 462 downgrades) made by S&P, Moody's and Fitch between 1994 and 2011.

The dependent variable $\Delta Spread_{n,t}$ is the change in non-event country n's bond spread vis-à-vis the United States over the two-trading-day window $[-1, +1]$ around the announcement on day 0 of a change in the rating of event country e ($\neq n$). The event window length accounts for the fact that by the time a CRA announces a rating change on day 0, markets in some parts of the world may have already closed. Hence, any impact on those would not materialise before day +1, and would go undetected using a shorter $[-1, 0]$ window. The same argument applies to rating announcements made after the exchange has closed in the country concerned, which we cannot distinguish from those made during trading.[7]

The key regressor in identifying possible spillover effects is $LARGE_{e,t}$, a dummy that takes on a value of one if e's rating is changed by two notches or more, and zero otherwise. We thereby treat rating changes of two notches or more as one single group. This is due to the distribution of the severity of upgrades and downgrades in our sample, which is shown in Figure 5.

The vast majority of rating announcements result in a one-notch change in a country's rating. Beyond that, we observe a significant amount of events only for changes of two notches, while changes of three notches or more occur only very rarely. There-

[7]CRAs have made post-trading announcements during the eurozone crisis, for instance (Financial Times, 2010; Wall Street Journal, 2012). In financial markets more generally, information which is deemed highly relevant is frequently released when exchanges are closed in order to limit or smooth the impact on prices.

fore, we do not include separate dummy variables for the latter categories but group all rating changes of two notches or more into a single bin.[8]

In this framework, positive (negative) spillover effects are equivalent to a drop (rise) in the spreads of country n which is significantly more pronounced in response to a two-or-more-notches upgrade (downgrade) of country e than to a single-notch one. We would then expect β to be significantly negative (positive) in the upgrade (downgrade) regressions.

This counterfactual choice also has implications for the estimation technique. Since we do not use "no change" as the counterfactual (due to the estimation window problem outlined above), we identify spillover effects in pooled cross sections of upgrades and downgrades rather than in a true panel setup.[9] We estimate the model by OLS.

At this point, it seems important to address a concern about potential endogeneity of the large-change dummy. The implicit assumption in the above design is that the rating announcement constitutes the dominant event in the event window, and that its severity is not systematically related to other spread-relevant information being released simultaneously. Otherwise, $LARGE$ and the error term ω would be correlated, so that β would be biased. Restricting the event window to two days already goes a long way towards alleviating the problem by limiting the amount of information that might potentially correlate with the large-change dummy. However, there might still remain a concern if, for instance, CRAs downgraded a country instantaneously in reaction to "bad news" and did so by more notches for "particularly bad news".

While one cannot rule this out for every single announcement in the sample — and because soft information is notoriously hard to measure —, the proclaimed practice and a corresponding body of empirical literature suggest that such a bias might not be a major issue. The agencies state a preference for stable ratings (see, eg, Cantor, 2001; Cantor and Mann, 2003, 2007; Standard & Poor's, 2010), intending to announce a change only if it is unlikely to be reversed in the near future. This "through the cycle" approach contrasts with a "point in time" approach in that cyclical phenomena should not, in themselves, trigger rating changes. If CRAs actually pursued a stable rating policy, the fact that cyclical and permanent

[8]For the same reason, Alsakka and ap Gwilym (2012) also apply this grouping in their analysis of the direct effect of sovereign rating changes on exchange rate movements.

[9]Thus, t denotes *generic* rather than actual time and can be thought of as indexing the different rating events.

factors are difficult to disentangle (International Monetary Fund, 2010) should imply some delay between new information becoming available and an ensuing change in the credit rating. Empirical evidence for corporate bond rating indicates that this practice might indeed be followed, thus reducing the timeliness of rating changes (Altman and Rijken, 2004; Liu et al., 2011), and that the CRAs may be "slow" in processing new information (Löffler, 2005). This perception has also been expressed in investor surveys (Association for Financial Professionals, 2002; Baker and Mansi, 2002). Finally, Sy (2004) notes for the sovereign sector that it may simply be concerns about rating changes precipitating significant increases in borrowing costs or outright crises which make CRAs opt for somewhat less timely announcements.

3.2 The rating environment

The rating environment may play an important role for the bond market reaction to an upgrade or downgrade announcement. Our regressions therefore control for a number of different rating variables contained in $RatEnv_{e,n,t}$. For example, the spillover potential of a rating action might depend on the creditworthiness of the event country, which we proxy by the rating it held with the announcing CRA on the day before ($InitRat_{e,t}$). We also include the absolute difference between the event country's initial rating and that of the non-event country ($\Delta InitRat_{e,n,t}$). This is because one might expect bilateral effects to differ depending on how similar countries are in terms of creditworthiness.

In addition, it is well established in the literature that the impact of rating announcements may vary according to whether they have been anticipated by the market (eg, Ismailescu and Kazemi, 2010; Gande and Parsley, 2005; Reisen and von Maltzan, 1999). One potentially important and convenient measure of such anticipation is whether the actual rating action has been foreshadowed by a CRA putting the respective country on watch, or review (Afonso et al., 2011; Kaminsky and Schmukler, 2002; Hand et al., 1992). Hence, we add a dummy that takes on a value of one if a review in the indicated direction has been ongoing at the time of the upgrade or downgrade, and zero otherwise ($OnWatch_{e,t}$).

Introducing an explicit control variable differs from Gande and Parsley (2005), who amalgamate a country's watch status into a "comprehensive credit rating". More precisely, for any given day their measure is defined as the country's actual letter rating on a 17-notch scale, raised (lowered) if the country is on review for an upgrade (downgrade). Presumably due to the counterfactual issue discussed in 3.1, Gande and Parsley (2005) then focus on those days as events on which there is a non-zero change in the comprehensive credit rating. However, this identification crucially involves additional assumptions on how changes in review status and actual rating changes relate to one another quantitatively. Furthermore, one might argue that, despite the potential anticipation effects of watchlistings, the latter are not qualitatively the same as actual rating changes. In any case, our much larger sample allows us to avoid those assumptions. We focus instead on the class of actual rating changes and their relative strengths only while controlling for anticipation through watchlistings. This should provide for a cleaner identification of spillover effects.

Moreover, we have shown in 2.2 that similar announcements by different CRAs tend to cluster around certain dates, and that this is particularly true for rating downgrades. We account for potential clustering *within* countries by a variable which captures the number of similar announcements made for a particular country by other agencies over a 14-day window before the respective event ($SimActsWdwEvt_{e,t}$). For clustering *across* countries, ie one or more CRAs changing the rating of more than one country in the same direction simultaneously, we include the number of similar announcements made on the same day for the"non-event" country ($SimActsDayNonEvt_{e,t}$).

Finally, we add the volatility measure for the S&P 500 Index in the United States (VIX_t) to control for the "global market sentiment" in which the rating announcement is made. One might, for instance, imagine that in more turbulent times (ie, in which volatility is high) borrowing conditions deteriorate across the board, so that spreads over the event window would be more likely to increase in any case. In that sense, VIX_t can be regarded as a technical control, which also adds a genuine time component to the pooled cross sections.

All regressions include the vector $Other_{e,n,t}$ which contains a fixed set of controls, such as event and non-event country dummies. Importantly, we also account for common time effects in the pooled cross sections through the inclusion of year dummies. These capture global macroeconomic trends which might be reflected in the yields of US Treasuries and, hence, spread changes. For instance, there may be a stronger tendency for investments to flow into the US in some years due to a (perceived) "safe haven" status, or a "global savings glut" that has been discussed for the early 2000s. Moreover, each regression includes the following technical controls: the maturity of non-event country bonds in levels and squares to account for positions on the yield curve, a dummy for EMBI Global bond yields, and a dummy for spread changes that need to be measured over weekends as those correspond to longer intervals in terms of calendar days.

4 Results

4.1 Existence of cross-border spillover effects

Table 1 shows baseline estimation results on the existence of cross-border effects for upgrades and downgrades, respectively. We start with a parsimonious specification in Model 1, which only contains our main variable of interest, the large-change dummy $LARGE$ and initial ratings. We then control for potential anticipation effects from watchlistings as well as clustering within and across countries in Model 2. Finally, Model 3 also accounts for global market turbulence, or risk aversion.

The key result is that the large-change dummy has the expected sign for both upgrades (ie, negative) and downgrades (ie, positive), and that it is highly significant in both cases. Moreover, this finding appears to be remarkably robust as the coefficient on $LARGE$ is very stable and retains its significance across specifications. Comparison of the absolute coefficients, however, indicates an asymmetry in the spillover effects induced by upgrades and downgrades, respectively. Downgrades of two notches or more are associated with an average spread change over the event window which exceeds that of one-notch downgrades by about 2 basis points. In contrast, large upgrades are associated with spread changes that are roughly 1.2 basis points below those of one-notch upgrades. The asymmetry is also reflected in the lower significance levels for upgrades despite a larger number of rating events and observations. To further corroborate this, we confirm in a separate (unreported) regression that the absolute coefficients for upgrades and downgrades are statistically different from each other.[10]

Asymmetries in the reactions to positive and negative events have frequently been documented in the literature. For instance, Gande and Parsley (2005) find for a

[10] To this end, we pool *all* rating changes and replace the event-window spread changes for upgrades with their negative values for the sake of comparison. We then add a downgrade dummy (taking on a value of one for downgrades, and zero for upgrades) to all specifications both in levels and as interactions with the other explanatory variables. The interaction term of $LARGE$ with the downgrade dummy is positive and highly significant throughout, pointing to statistically significant differences in the absolute coefficients for upgrades and downgrades.

Table 1: **Baseline regressions**

	Panel A: Upgrades			Panel B: Downgrades		
	(1)	(2)	(3)	(1)	(2)	(3)
LARGE	-0.0121**	-0.0124*	-0.0128*	0.0187***	0.0224***	0.0207***
	(0.0060)	(0.0064)	(0.0067)	(0.0061)	(0.0065)	(0.0066)
InitRat	0.0001	-0.0005	0.0000	-0.0013	-0.0013	-0.0008
	(0.0008)	(0.0009)	(0.0010)	(0.0014)	(0.0017)	(0.0017)
Δ*InitRat*	0.0010	0.0008	0.0009	0.0006	0.0008	0.0008
	(0.0006)	(0.0006)	(0.0007)	(0.0008)	(0.0009)	(0.0009)
OnWatch		0.0057	0.0070		-0.0100*	-0.0046
		(0.0055)	(0.0058)		(0.0054)	(0.0054)
SimActsWdwEvt		-0.0020	-0.0013		0.0170***	0.0141**
		(0.0057)	(0.0057)		(0.0064)	(0.0065)
SimActsDayNonEvt		-0.0863*	-0.0877		0.1210**	0.1477**
		(0.0512)	(0.0546)		(0.0558)	(0.0635)
VIX			0.0017***			0.0006*
			(0.0004)			(0.0004)
N	31,986	30,564	29,950	23,734	22,413	21,931
Event countries	104	92	92	95	84	84
Non-event countries	73	73	73	73	73	73
Rating actions	635	606	595	462	436	427
R^2	0.0230	0.0216	0.0223	0.0397	0.0400	0.0423

Notes — This table shows baseline regressions explaining the percentage point change Δ*Spread* in non-event country spreads around the rating announcement for up to 635 upgrades and 462 downgrades made by S&P, Moody's and Fitch between 1994 and 2011. For variable definitions, see Table A.3 in the Appendix. All specifications include a constant, dummies for event and non-event countries, years, spread reactions over weekends and JP Morgan EMBI Global data, as well as levels and squares of non-event country bond maturities. Robust standard errors in parentheses. ***, **, and * denote significance at the 1, 5, and 10 per cent levels, respectively.

1990s sample of developed and emerging countries that negative rating events in one country affect sovereign bond spreads in others whereas there is no discernible impact for positive events. Similar results have been obtained regarding the direct effects in sovereign bond and CDS markets (Afonso et al., 2011; Larraín et al., 1997), mirroring a well-established finding from event studies on bond, stock, and CDS returns in the corporate sector (eg, Norden and Weber, 2004; Steiner and Heinke, 2001; Goh and Ederington, 1993; Hand et al., 1992). Recently, however, there has also been evidence of symmetric spillover reactions to sovereign rating announcements in the foreign exchange market (Alsakka and ap Gwilym, 2012), or even that positive announcements in emerging countries have both stronger direct and spillover effects in sovereign CDS markets (Ismailescu and Kazemi, 2010).

Turning to the rating-environment controls, neither the initial rating of the event country just before the rating announcement nor the difference in initial ratings between event and non-event country seem to play a role in terms of spillover effects. Both coefficients are far from significant across specifications. Previous evidence on this has been inconclusive. While Alsakka and ap Gwilym (2012) and Ferreira and Gama (2007) detect stronger spillover effects in the foreign exchange and stock markets, respectively, for event countries with lower initial ratings, Gande and Parsley (2005) find the opposite for bond market reactions (to sovereign downgrades).

We do find some evidence, though, that the impact of an actual rating change on spreads depends on whether it has been foreshadowed by a watchlisting. The corresponding dummy, *OnWatch*, is signed as expected for both upgrades and downgrades, yet there is again an asymmetry: the control variable turns out insignificant in all upgrade specifications but significant at almost the five per cent level for downgrades (Model 2 in Panel B). A possible explanation for this is given by Altman and Rijken (2006). They point out that watchlistings partially ease the tension between the market's expectation of rating stability and the demand for rating timeliness. This suggests that watchlistings contribute to the anticipation of actual rating changes. Given that investors tend to be more concerned about negative news, watchlistings should be more important in building anticipation for downgrades than for upgrades. Figures from our dataset support this notion. While about a third of all downgrades are preceded by a watchlisting, so are only 15 per cent of all upgrades. Finally, it has often been noted that there is an incentive to leak good news (eg, Alsakka and ap Gwilym, 2012; Christopher et al., 2012; Gande and Parsley, 2005; Goh and Ederington, 1993; Holthausen and Leftwich, 1986), so the relevance of watchlistings in building anticipation is conceivably much lower in the case of upgrades. We interpret the fact that our results are consistent with this literature as reassuring in terms of the validity of the regression specifications.

Our results also point to the importance of the clustering of rating announcements, especially for downgrades. While the controls for both clustering within (*SimActsWdwEvt*) and across countries (*SimActsDayNonEvt*) are highly significant in the downgrade regressions, the effect of across-clustering is only marginally significant once for upgrades. This appears plausible in light of the stylised facts presented in 2.2 because simultaneous announcements on several countries by one or more agencies occur much less frequently for upgrades than for downgrades. Moreover, the coefficients are correctly signed for both upgrades and downgrades, suggesting that the spread-decreasing (spread-increasing) spillover effects of an upgrade (downgrade) are all the more pronounced when one or more upgrades (downgrades) are announced for the "non-event" country at the same time.

A similar statement regarding the signs cannot be made with the same degree of confidence for *SimActsWdwEvt*, which measures the number of upgrades (downgrades) announced by other agencies over a 14-day window before the respective upgrade (downgrade).[11] While we again find strong differences in significance between upgrades and downgrades as well as opposing signs, one need not necessarily expect within-clustering to have an additional spread-increasing effect over the event window for downgrades. Instead, the variable might subsume two opposing effects. On the one hand, the clustering of downgrades over a short interval could imply that any announcement is less relevant individually. In that case, one would expect a negative coefficient. On the other hand, clustering is much more prevalent in crisis times (see 2.2). Thus, *SimActsWdwEvt* tends to be higher in times of market turbulence or global risk aversion when spreads against a "safe-haven" investment like US Treasuries are upward-trending, too (eg, González-Rozada and Levy Yeyati, 2008; García-Herrero and Ortíz, 2006; International Monetary Fund, 2004, 2006).

[11] In choosing the window length, we follow Gande and Parsley (2005) who employ a two-week duration for a comparable control variable. However, using a one-week or three-week window instead does not alter the conclusions. Moreover, the reader may note that we do not report a variable capturing similar rating announcements made *on the same day* by other agencies. This is due to the unattractive property that this variable drops out in the upgrade regressions since there is not a single event of multiple upgrades of a country on the same day in our sample. Therefore, in the interest of comparability, we choose not to report downgrade regressions with that control either. These regressions show, however, that the measure is always insignificant for downgrades, regardless of whether it is included in addition to, or as a stand-in for, *SimActsWdwEvt*. All results are available on request.

As this is consistent with a positive sign, the significantly positive coefficients for downgrades suggest that we may be picking up a substantial turbulence component.

Since the literature provides little guidance on whether this is what is driving our results, we include the S&P 500 Volatility Index (*VIX*), a commonly used proxy for global risk aversion (De Santis, 2012). As expected, its coefficient is positive and significant for both upgrades and downgrades, given the relation between market turbulence and yield spread drift. Interestingly, the coefficient on *SimActsWdwEvt* is still positive but slightly lower than before. This may be due to *VIX* picking up some of the turbulence effect previously captured by *SimActsWdwEvt*. Hence, there is indeed evidence that clustering may also reduce the spillover relevance of individual rating events that take place in a period of many similar announcements by other CRAs.

Finally, we subject our baseline regressions for downgrades to some robustness checks regarding extreme rating events and two specific crisis episodes. One might be concerned, for instance, that grouping all downgrades of two notches or more into a single bin could obscure the impact of a very few severe rating changes that might be driving our results (see Figure 5). However, this is not the case as dropping downgrades of four notches or more and three notches or more, respectively, leaves the findings unchanged. Furthermore, we ensure that the results on negative spillovers are not merely the product of the eurozone and Asian crises. These robustness checks are reported in Table A.4 in the Appendix.

4.2 Spillover channels

We now turn to potential channels of spillover effects in the sovereign bond market. While the regressions presented so far control for a multiplicity of factors pertaining to event and non-event countries *on their own*, they do not — with the exception of Δ*InitRat* — account for *bilateral* characteristics of event and non-event countries. However, bond market reactions in the wake of rating announcements in other countries might differ depending on similarities and bilateral linkages, which may be highly relevant from the perspective of policymakers.

We therefore augment our final baseline specification (Model 3 in Table 1) by whether the event and non-event country belong to the same geographical region (*Region*), whether they are members of a common major trade bloc (*TradeBloc*), and the importance of the event country as an export destination for the non-event country (*ExpImpEvt*). We also consider the size of the event country's GDP (*SizeEvt*) as well as differences between event and non-event countries in terms of GDP ($\Delta Size$) and trend growth ($\Delta TrendGrowth$). Definitions and sources for all control variables are reported in Table A.3 in the Appendix. The results are shown in Tables 2 and 3.

There is again a notable asymmetry between the findings on upgrades and those on downgrades. This applies to both the results on the potential channels themselves and to the impact that the inclusion of additional controls has on the robustness of our baseline findings. Whereas the results for downgrades are highly stable and intuitive, they paint a more nuanced picture for upgrades.

In more detail, we find consistently that spillover effects are significantly stronger with the same region in the case of downgrade announcements. The coefficient on *Region* has the correct sign, indicating that borrowing costs increase by up to almost four basis points more for non-event countries in the same region as the event country than for those outside it. Our findings appear plausible since countries in the same geographical region are more likely to share institutional or cultural characteristics and to have important real and financial links to one another. Apart from the fundamental factors, a more mundane explanation might posit that financial markets simply find non-event countries from the same region "guilty by association". The results are also in line with a number of studies which focus on one or more particular regions from the start (eg, Alsakka and ap Gwilym, 2012; Arezki et al., 2011; De Santis, 2012). Surprisingly, we obtain positive coefficients for upgrades as well, which would suggest that those are less likely to induce spillovers within than across regions. While one could imagine that belonging to a particular region does not matter for upgrade announcements due to an asymmetric perception by investors, the fact that the coefficients are often significant is not easily rationalised. On a positive note, though, the magnitude for upgrades is only about a third of that for downgrades. Therefore, in the interest of comparability and as an important economic control, we retain *Region* in all specifications.

The two trade controls, ie common membership in a major trade bloc (*TradeBloc*) and the non-event country's ratio of exports to the event country to domestic GDP (*ExpImpEvt*), are signed as expected throughout, pointing to more pronounced spillover effects for both upgrades and downgrades when such linkages exist, or when they are stronger. However, they are only mildly significant once for upgrades (see Model 7 in Table 2). Moreover, the stability in magnitude and significance of

Table 2: **Spillover channels, upgrades**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LARGE	-0.0128*	-0.0128*	-0.0111	-0.0117*	-0.0142**	-0.0112	-0.0137**
	(0.0067)	(0.0067)	(0.0071)	(0.0068)	(0.0066)	(0.0070)	(0.0069)
InitRat	0.0000	0.0001	-0.0005	0.0027**	0.0031***	0.0008	0.0013
	(0.0010)	(0.0010)	(0.0010)	(0.0013)	(0.0012)	(0.0014)	(0.0013)
Δ*InitRat*	0.0009	0.0010	0.0006	0.0012*	0.0011	0.0009	0.0007
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0008)	(0.0008)
OnWatch	0.0070	0.0070	0.0066	0.0080	0.0085	0.0068	0.0075
	(0.0058)	(0.0058)	(0.0060)	(0.0059)	(0.0061)	(0.0061)	(0.0062)
SimActsWdwEvt	-0.0013	-0.0013	-0.0058	-0.0026	-0.0032	-0.0067	-0.0073
	(0.0057)	(0.0057)	(0.0059)	(0.0058)	(0.0059)	(0.0060)	(0.0061)
SimActsDayNonEvt	-0.0877	-0.0903	-0.1024	-0.0883	-0.0950	-0.1002	-0.1091*
	(0.0546)	(0.0549)	(0.0625)	(0.0546)	(0.0578)	(0.0621)	(0.0663)
VIX	0.0017***	0.0017***	0.0019***	0.0017***	0.0018***	0.0019***	0.0020***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Region		0.0109	0.0146*	0.0128*	0.0125*	0.0162**	0.0169**
		(0.0071)	(0.0080)	(0.0073)	(0.0075)	(0.0080)	(0.0083)
TradeBloc			-0.0100			-0.0107	-0.0133*
			(0.0065)			(0.0066)	(0.0069)

(continued on next page)

Spillover channels, upgrades (continued)

ExpImpEvt			-0.1080			-0.0767	-0.0875
			(0.2149)			(0.2134)	(0.02145)
SizeEvt				0.0279	0.0257	0.0425**	0.0409**
				(0.0190)	(0.0196)	(0.0202)	(0.0208)
Δ*Size*				-0.0399**	-0.0404**	-0.0430**	-0.0442**
				(0.0187)	(0.0194)	(0.0197)	(0.0205)
Δ*TrendGrowth*					-0.0001		-0.0001
					(0.0001)		(0.0001)
N	29,950	29,950	27,962	29,329	28,904	27,798	27,380
Event countries	92	92	90	92	91	90	89
Non-event countries	73	73	71	72	72	71	71
Upgrades	595	595	582	592	584	579	571
R^2	0.0223	0.0223	0.0221	0.0235	0.0271	0.0231	0.0267

Notes — This table shows regressions investigating potential spillover channels for up to 595 upgrade announcements made by S&P, Moody's and Fitch between 1994 and 2011. For variable definitions, see Table A.3 in the Appendix. All specifications include a constant, dummies for event and non-event countries, years, spread reactions over weekends and JP Morgan EMBI Global data, as well as levels and squares of non-event country bond maturities. Robust standard errors in parentheses. ***, **, and * denote significance at the 1, 5, and 10 per cent levels, respectively.

Table 3: **Spillover channels, downgrades**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LARGE	0.0207***	0.0206***	0.0217***	0.0222***	0.0224***	0.0227***	0.0229***
	(0.0066)	(0.0066)	(0.0069)	(0.0070)	(0.0070)	(0.0072)	(0.0072)
InitRat	-0.0008	-0.0006	-0.0010	-0.0017	-0.0017	-0.0023	-0.0024
	(0.0017)	(0.0017)	(0.0018)	(0.0019)	(0.0019)	(0.0020)	(0.0020)
Δ*InitRat*	0.0008	0.0012	0.0017*	0.0008	0.0008	0.0013	0.0013
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0010)	(0.0011)	(0.0011)
OnWatch	-0.0046	-0.0046	-0.0031	-0.0009	-0.0008	0.0009	0.0011
	(0.0054)	(0.0054)	(0.0058)	(0.0056)	(0.0057)	(0.0058)	(0.0059)
SimActsWdwEvt	0.0141**	0.0141**	0.0135**	0.0146**	0.0146**	0.0139**	0.0139**
	(0.0065)	(0.0065)	(0.0066)	(0.0067)	(0.0067)	(0.0068)	(0.0068)
SimActsDayNonEvt	0.1477**	0.1451**	0.1426**	0.1160*	0.1161*	0.1136*	0.1137*
	(0.0648)	(0.0643)	(0.0653)	(0.0623)	(0.0623)	(0.0622)	(0.0622)
VIX	0.0006*	0.0006*	0.0006	0.0006*	0.0006*	0.0006	0.0006
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Region		0.0376**	0.0329**	0.0379**	0.0380**	0.0331**	0.0332**
		(0.0153)	(0.0164)	(0.0157)	(0.0157)	(0.0168)	(0.0168)
TradeBloc			0.0159			0.0151	0.0153
			(0.0111)			(0.0115)	(0.0116)

(continued on next page)

Spillover channels, downgrades (continued)

ExpImpEvt			0.0687			0.0635	0.0594
			(0.2200)			(0.2265)	(0.2259)
SizeEvt				0.0222	0.0221	0.0263	0.0260
				(0.0290)	(0.0294)	(0.0309)	(0.0313)
Δ*Size*				-0.0169	-0.0170	-0.0190	-0.0189
				(0.0218)	(0.0223)	(0.0234)	(0.0240)
Δ*TrendGrowth*					0.0000		0.0000
					(0.0000)		(0.0000)
N	21,931	21,931	20,633	21,031	20,885	20,035	19,896
Event countries	84	84	81	82	82	79	79
Non-event countries	73	73	71	72	72	71	71
Downgrades	427	427	416	416	416	405	405
R^2	0.0423	0.0428	0.0423	0.0441	0.0442	0.0435	0.0436

Notes — This table shows regressions investigating potential spillover channels for up to 427 downgrade announcements made by S&P, Moody's and Fitch between 1994 and 2011. For variable definitions, see Table A.3 in the Appendix. All specifications include a constant, dummies for event and non-event countries, years, spread reactions over weekends and JP Morgan EMBI Global data, as well as levels and squares of non-event country bond maturities. Robust standard errors in parentheses. ***, **, and * denote significance at the 1, 5, and 10 per cent levels, respectively.

Region upon inclusion of the trade variables, in particular for downgrades, seems to indicate that stronger spillover effects within regions cannot easily be explained by real linkages.[12] To the extent that trade also captures a notable portion of variation in bilateral asset holdings, the same applies to financial linkages.[13] Hence, it appears more likely that our findings on belonging to the same region are driven by shared institutional and cultural features or, less sophisticated, that this has an effect by itself.

The evidence on the remaining potential channels is succinctly summarised for downgrades. In no specification do the size of the event country's GDP (*SizeEvt*), its increment over that of the non-event country ($\Delta Size$), or differences in trend growth between event and non-event countries ($\Delta TrendGrowth$) turn out to be significant determinants of the strength of bond market spillovers. At the same time, all results from the baseline and augmented baseline regressions (Models 1 and 2 in Table 3) prove remarkably stable in terms of both magnitude and significance.

This contrasts with the corresponding findings for upgrades. On the one hand, we obtain a number of interesting results for the size and growth controls. On the other hand, the augmented regressions raise some doubts on our our main variable of interest, *LARGE*, in terms of statistical significance. The latter alternates between specifications and vanishes in some, yet in view of the considerably stronger baseline results for downgrades, this is not entirely surprising. It merely serves to underscore the asymmetry that exists between positive and negative rating changes. However, this also means that the evidence on the potential channels for upgrades should be taken with a grain of salt.

[12]The fact that the correlation of the two trade variables with the region control is low does not support multicollinearity as a technical explanation for this result. Moreover, replacing *ExpImpEvt* by other proxies for bilateral trade does not change the picture either (see Table A.5 in the Appendix).

[13]We would like to control directly for financial linkages of event and non-event countries, eg the exposure of non-event country banks to event-country sovereign bonds. Unfortunately, even use of the most comprehensive data from the IMF's Coordinated Portfolio Investment Survey leads to a massive reduction in the number of observations and major selection effects along the time series and country dimensions, which renders virtually impossible any comparison with the baseline results. This also applies to data on foreign direct investment (FDI). However, bilateral trade is likely to pick up some of the financial linkages we intend to capture as there is evidence that trade is a powerful determinant of bilateral (bank) asset holdings (Aviat and Coeurdacier, 2007). In addition, through its correlation with FDI, trade may proxy for cross-country bank exposure since bank lending may follow domestic companies when those set up operations abroad (eg, Goldberg and Saunders, 1980, 1981; Brealey and Kaplanis, 1996; Yamori, 1998). Aiming to maintain broad coverage and composition through time, we rely on trade data.

In this regard, the most interesting result is probably the observation that, given the event country's size and initial rating, positive spillovers are larger the smaller the non-event country relative to the event country ($\Delta Size$). The magnitude of the coefficient suggests that non-event countries which are half (two-thirds) the size of the event country experience an additional positive spillover effect of about four (two) basis points, as compared to non-event countries as large as the event country.[14] While the effect appears to be relatively small, its direction is still interesting, in particular when viewed in conjunction with the fact that, across the whole sample, larger and more highly rated countries induce smaller spillovers (Models 4 to 7 in Table 2).[15] This would be consistent with a world in which positive spillover effects matter primarily within a group of small developed and emerging countries but less so within a group of large, developed countries, and in which the latter have little impact on the former. The insignificance of the absolute difference in trend GDP growth rates between event and non-event countries ($\Delta TrendGrowth$) as a further measure of differences in economic development does nothing to contradict this interpretation. In view of the generally more ambiguous results for upgrades, however, we do not wish to overemphasise this point.

4.3 Discussion

Our results can be condensed into the following stylised facts. First, there is strong evidence of statistically significant, negative spillover effects of downgrade announcements. This result proves highly robust to controlling for anticipation through watchlistings and the clustering of rating announcements. Second, negative spillover effects are more pronounced among countries in a common region, which cannot be explained by measurable fundamental links and similarities between countries. Third, reactions to upgrades are, if anything, much more muted than for downgrades, suggesting important asymmetries in the sovereign bond market's treatment of the two types of announcements. Fourth, evidence on the channels behind pos-

[14]$\Delta Size$ is defined as the difference between the event and non-event country's log GDPs or, equivalently, the log of the ratio of the two GDP levels. Therefore, a decrease in relative non-event country size by half (two-thirds) amounts to an increase in $\Delta Size$ of about one hundred (fifty) per cent. With an absolute coefficient of roughly 0.04, the (semi-elasticity) marginal effects therefore obtain as four and two basis points, respectively.

[15]Larger countries tend to hold more favourable credit ratings. This correlation may explain why, in Models 4 and 5, we find *InitRat* significant and *SizeEvt* insignificant, and vice versa for Models 6 and 7.

itive spillover effects, if any, offers a more complex picture and appears relatively inconclusive.

So, which conclusion to draw from this? To begin with, there is a strong case for the notion that negative sovereign rating announcements, ie those of most concern to policymakers, do matter in inducing spillovers across markets. Such is the outcome of the explicit identification strategy used in this paper, which demonstrates that, all other things equal, "large" downgrades of two notches or more cause larger hikes in spreads than "small" one-notch downgrades. This suggests a role for CRAs and their actions in sovereign bond markets, be it through the revelation of new information on creditworthiness which acts as a "wake-up call" for investors to reassess fundamentals in other countries (Goldstein, 1998), or simply by providing a coordinating signal that shifts expectations from a good to a bad equilibrium (Boot et al., 2006; Masson, 1998).

However, a major regulatory focus on the activities of CRAs would also require negative spillover effects of substantial *economic* magnitude. In this paper, we find the incremental impact of "large" downgrades to be a little over two basis points, which may appear limited at first glance. Yet, it is important to note that this does not represent the total effect that policymakers would be concerned about. This can be thought of as consisting of a "base effect" that "small" downgrades have, compared to a benchmark scenario of no downgrades anywhere, plus an additional impact for "large" downgrades — which is what we measure. Of course, the reason we focus on the latter lies in the impossibility of cleanly identifying the "base effects" of rating changes unless one rules out the existence of rating-induced spillovers from the beginning (see the discussion in 3.1). Nonetheless, the total effect is conceivably a multiple of the one we estimate. At factors of 2 and 5, for instance, the implied total effects amount to approximately 4 and 10 basis points, respectively. To put this into perspective, the average sovereign bond spread vis-à-vis US Treasuries at the time of the downgrade announcements in our sample is 3.25 per cent, or 325 basis points. While the total effect of downgrades is relatively small in comparison, one has to bear in mind that governments often need to refinance large amounts of debt, which magnifies the impact of even small spread differences. Moreover, there is still a regional effect of up to 4 basis points on top of that, suggesting that concerns about negative spillovers in the sovereign debt market should not be lightly dismissed.

Finally, from a policymaker's point of view, the finding that the increased strength of negative spillovers within regions cannot be explained away by measurable linkages and similarities between countries might also be a cause for concern. Even though limited data availability precludes an all-encompassing analysis of potential channels, there is little to suggest that one can comfortably rule out that some countries are found "guilty by association" with the event country. Moreover, such behaviour on the part of investors would likely extend to their reactions to news other than rating announcements. While it is hard to see an obvious remedy, the potential problem would seem to be much more general and, above all, rooted in investor behaviour. Hence, it is not clear that putting the primary emphasis on CRAs would prove effective in this regard.

5 Conclusion

Concerns about negative spillovers across sovereign debt markets in the wake of sovereign rating changes have recently resurfaced on the agenda of policymakers. In this paper, we study the existence and potential channels of such spillover effects. More specifically, we avail of an extensive dataset which covers all sovereign rating announcements made by the three major agencies and daily sovereign bond market movements of up to 73 developed and emerging countries between 1994 and 2011. Based on this, we propose an explicit counterfactual identification strategy which compares the bond market reactions to small changes in an agency's assessment of a country's creditworthiness to those induced by all other, more major revisions. In doing so, we account for a number of factors that might impact on the reception of individual announcements.

We find strong evidence in favour of negative cross-border spillovers in the wake of sovereign downgrades. At the same time, there is no similarly robust indication as to positive spillovers since reactions to upgrades are much more muted at best, which points to an important asymmetry in the sovereign debt market's treatment of positive and negative information. Regarding the channels of negative spillover effects, our results suggest that those are more pronounced for countries within the same region. Strikingly, however, this cannot be explained by (measurable) fundamental linkages and similarities, such as trade, which turn out to be insignificant.

Therefore, there is reason to believe that policymakers' concerns about negative spillover effects are not unfounded. In fact, failure to reject the irrelevance of a set of fundamentals in explaining the added regional component may reinforce, or give rise to, concerns about the ability of investors to discriminate accurately between sovereigns. This could also be of more general interest because such behaviour is likely to carry over to reactions to various kinds of non-CRA news in other markets and sectors, too. Hence, important though they are, a sole focus on CRAs and their actions might be missing a bigger picture.

References

Afonso, A., D. Furceri, and P. Gomes (2011): "Sovereign credit ratings and financial markets linkages: applications to European data," Working Paper No. 1347, European Central Bank.

Alsakka, R. and O. ap Gwilym (2012): "Foreign exchange market reactions to sovereign credit news," *Journal of International Money and Finance*, 31, 845–864.

Altman, E. I. and H. A. Rijken (2004): "How Rating Agencies Achieve Rating Stability," *Journal of Banking & Finance*, 28, 2679–2714.

——— (2006): "The Added Value of Rating Outlooks and Rating Reviews to Corporate Bond Ratings," Working Paper, NYU Salomon Center.

Arezki, R., B. Candelon, and A. N. R. Sy (2011): "Sovereign Rating News and Financial Markets Spillovers: Evidence from the European Debt Crisis," IMF Working Paper Series 11/68.

Association for Financial Professionals (2002): "Rating Agencies Survey: Accuracy, Timeliness, and Regulation," November.

Aviat, A. and N. Coeurdacier (2007): "The geography of trade in goods and asset holdings," *Journal of International Economics*, 71, 22–51.

Baker, H. K. and S. A. Mansi (2002): "Assessing Credit Rating Agencies," *Journal of Business Finance & Accounting*, 29, 1367–1398.

Boot, A. W. A., T. A. Milbourn, and A. Schmeits (2006): "Credit Ratings as Coordination Mechanisms," *Review of Financial Studies*, 19, 81–118.

Brealey, R. A. and E. C. Kaplanis (1996): "The determination of foreign banking location," *Journal of International Money and Finance*, 15, 577–597.

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

Cantor, R. (2001): "Moody's investors service response to the consultative paper issued by the Basel Committee on Bank Supervision "A new capital adequacy framework"," *Journal of Banking & Finance*, 25, 171–185.

Cantor, R. and C. Mann (2003): “Measuring the Performance of Corporate Bond Ratings,” Special comment, Moody’s Investors Service, New York.

——— (2007): “Analyzing the Tradeoff between Ratings Accuracy and Stability,” *Journal of Fixed Income*, 16, 60–68.

Cantor, R. and F. Packer (1996): “Determinants and impact of sovereign credit ratings,” *Economic Policy Review*, 37–53.

Christopher, R., S.-J. Kim, and E. Wu (2012): “Do sovereign credit ratings influence regional stock and bond market interdependencies in emerging countries?” *Journal of International Financial Markets, Institutions and Money*, 22, 1070–1089.

De Santis, R. A. (2012): “The euro area sovereign debt crisis: safe haven, credit rating agencies and the spread of the fever from Greece, Ireland and Portugal,” Working Paper No. 1419, European Central Bank.

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

Financial Times (2010): “German MPs claim Greece needs €120bn,” 28 April 2010.

——— (2011): “Rating agencies face shake-up,” 21 October 2011.

Gande, A. and D. C. Parsley (2005): “News spillovers in the sovereign debt market,” *Journal of Financial Economics*, 75, 691–734.

García-Herrero, A. and A. Ortíz (2006): “The Role of Global Risk Aversion in Explaining Sovereign Spreads,” *Economía*, 7, 125–155.

Goh, J. C. and L. H. Ederington (1993): “Is a Bond Rating Downgrade Bad News, Good News, or No News for Stockholders?” *Journal of Finance*, 48, 2001–2008.

Goldberg, L. G. and A. Saunders (1980): “The Causes of U.S. Bank Expansion Overseas: The Case of Great Britain,” *Journal of Money, Credit and Banking*, 12, 630–643.

——— (1981): "The determinants of foreign banking activity in the United States," *Journal of Banking & Finance*, 5, 17–32.

GOLDSTEIN, M. (1998): *The Asian Crisis: Causes, Cures, and Systemic Implications*, Washington, D.C.: Institute for International Economics.

GONZÁLEZ-ROZADA, M. AND E. LEVY YEYATI (2008): "Global Factors and Emerging Market Spreads," *The Economic Journal*, 118, 1917–1936.

HAND, J. R. M., R. W. HOLTHAUSEN, AND R. W. LEFTWICH (1992): "The Effect of Bond Rating Agency Announcements on Bond and Stock Prices," *Journal of Finance*, 47, 733–752.

HILL, P. AND R. FAFF (2010): "The Market Impact of Relative Agency Activity in the Sovereign Ratings Market," *Journal of Business Finance & Accounting*, 37, 1309–1347.

HOLTHAUSEN, R. W. AND R. W. LEFTWICH (1986): "The Effect of Bond Rating Changes on Common Stock Prices," *Journal of Financial Economics*, 62, 57–89.

HOOPER, V., T. HUME, AND S.-J. KIM (2008): "Sovereign rating changes–Do they provide new information for stock markets?" *Economic Systems*, 32, 142–166.

INTERNATIONAL MONETARY FUND (2004): *Global Financial Stability Report, September 2004*, chap. Global Financial Market Developments, 8–80.

——— (2006): *Global Financial Stability Report, September 2006*, chap. Assessing Global Financial Risks, 1–45.

——— (2010): *Global Financial Stability Report, October 2010*, chap. The Uses and Abuses of Sovereign Credit Ratings, 85–122.

ISMAILESCU, I. AND H. KAZEMI (2010): "The reaction of emerging market credit default swap spreads to sovereign credit rating changes," *Journal of Banking & Finance*, 34, 2861–2873.

JP MORGAN (1999): "Introducing the J.P. Morgan Emerging Markets Bond Index Global (EMBI Global)," August.

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

Kaminsky, G. L. and S. L. Schmukler (1999): "What Triggers Market Jitters? - A Chronicle of the Asian Crisis," *Journal of International Money and Finance*, 18, 537–60.

Larraín, G., H. Reisen, and J. von Maltzan (1997): "Emerging Market Risk and Sovereign Credit Ratings," OECD Development Centre Working Papers 124, OECD Publishing.

Liu, P., J. S. Jones, and J. Y. Gu (2011): "Do Credit Rating Agencies Sacrifice Timeliness by Pursuing Rating Stability? Evidence from Equity Market Reactions to CreditWatch Events," Paper presented at the 2011 EFMA Annual Meeting, Braga (Portugal).

Löffler, G. (2005): "Avoiding the Rating Bounce: Why Rating Agencies are Slow to React to New Information," *Journal of Economic Behavior and Organization*, 56, 365–381.

MacKinlay, A. C. (1997): "Event Studies in Economics and Finance," *Journal of Economic Literature*, 35, 13–39.

Masson, P. (1998): "Contagion: Monsoonal Effects, Spillovers, and Jumps Between Multiple Equilibria," IMF Working Paper Series 98/142.

Norden, L. and M. Weber (2004): "Informational Efficiency of Credit Default Swap and Stock Markets: The Impact of Credit Rating Announcements," *Journal of Banking & Finance*, 28, 2813–2843.

Reisen, H. and J. von Maltzan (1999): "Boom and Bust in Sovereign Ratings," *International Finance*, 2, 273–293.

Standard & Poor's (2010): "Methodology: Credit Stability Criteria," Standard & Poor's RatingsDirect.

Steiner, M. and V. G. Heinke (2001): "Event Study Concerning International Bond Price Effects of Credit Rating Actions," *International Journal of Finance and Economics*, 6, 139–157.

Sy, A. (2004): "Rating the rating agencies: Anticipating currency crises or debt crises?" *Journal of Banking & Finance*, 28, 2845–2867.

Wall Street Journal (2012): "Moody's Poised to Decide on Spain," 28 September 2012.

Yamori, N. (1998): "A note on the location choice of multinational banks: The case of Japanese financial institutions," *Journal of Banking & Finance*, 22, 109–120.

Appendix

Table A.1: **Sovereign bond yield data sources and availability**

Bloomberg (33 countries)	
1994	Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Spain, Sweden, United Kingdom, United States (January), Switzerland (February)
1997	Portugal (February), Greece (July)
1998	Hong Kong (March), Singapore (June), India (November)
1999	Taiwan (April)
2000	Thailand (January), Czech Republic (April), South Korea (December)
2002	Slovakia (June), Romania (August)
2006	Israel (February)
2007	Slovenia (March)
2008	Iceland (April)
JP Morgan EMBI Global (41 countries)	
1994	Argentina, Mexico, Nigeria, Venezuela (January), China (March), Brazil (April), Bulgaria (July), Poland (October), South Africa (December)
1995	Ecuador (February)
1996	Turkey (June), Panama (July), Croatia (August), Malaysia (October)
1997	Colombia (February), Peru (March), Philippines, Russia (December)
1998	Lebanon (April)
1999	Hungary (January), Chile (May)
2000	Ukraine (May)
2001	Pakistan (January), Uruguay (May), Egypt (July), Dominican Republic (November)
2002	El Salvador (April)
2004	Indonesia (May)
2005	Serbia (July), Vietnam (November)
2007	Belize (March), Kazakhstan (June), Ghana, Jamaica (October), Sri Lanka (November), Gabon (December)
2008	Georgia (June)
2011	Jordan (January), Senegal (May), Lithuania, Namibia (November)

Notes — This table lists the sources of the sovereign bond yield data in the sample and the years in which the respective time series are first observed (months in parentheses). If there are gaps in the Bloomberg 10-year generic yield series, we add observations of 10-year generic yields from Datastream, ensuring that this does not induce structural breaks. Moreover, for some emerging countries we include 10-year generic yields until the EMBI Global series become available.

Table A.2: **Rating scales and transformation**

Characterisation of debt and issuer		Letter rating			Linear transformation
		S&P	Moody's	Fitch	
Highest quality	Investment grade	AAA	Aaa	AAA	17
High quality		AA+	Aa1	AA+	16
		AAA	Aa2	AA	15
		AA–	Aa3	AA–	14
Strong payment capacity		A+	A1	A+	13
		A	A2	A	12
		A–	A3	A–	11
Adequate payment capacity		BBB+	Baa1	BBB+	10
		BBB	Baa2	BBB	9
		BBB–	Baa3	BBB–	8
Likely to fulfil obligations, ongoing uncertainty	Speculative grade	BB+	Ba1	BB+	7
		BB	Ba2	BB	6
		BB–	Ba3	BB–	5
High credit risk		B+	B1	B+	4
		B	B2	B	3
		B–	B3	B–	2
Very high credit risk		CCC+	Caa1	CCC+	1
		CCC	Caa2	CCC	
		CCC–	Caa3	CCC–	
Near default with possibility of recovery		CC	Ca	CC	
				C	
Default		SD	C	DDD	
		D		DD	
				D	

Notes — This table shows how the letter ratings used by S&P, Moody's and Fitch correspond to one another and to different degrees of credit risk, and how they are mapped into the linear 17-notch scale used in the investigation. The transformation is the same as in Afonso et al. (2011), from which this table is adapted.

Table A.3: **Variable definitions**

Variable	Definition	Sources
$\Delta Spread$	Change in the non-event country spread vis-à-vis US Treasuries of comparable maturity over the two-trading-day window $[-1, +1]$ around the rating announcement (day 0), measured in percentage points.	Bloomberg, Datastream, JP Morgan, US Department of the Treasury
LARGE	Dummy variable taking on a value of one for "large" rating changes of two notches or more; zero otherwise. Notches are measured according to the linear transformation in Table A.2.	S&P, Moody's, Fitch
InitRat	Credit rating held by the event country with the announcing CRA prior to the event, measured on the 17-notch scale (see Table A.2).	S&P, Moody's, Fitch
$\Delta InitRat$	Absolute difference between *InitRat* and the average of all credit ratings held by the non-event country with the three CRAs, measured on the 17-notch scale (see Table A.2).	S&P, Moody's, Fitch
OnWatch	Dummy variable taking on a value of one if the event country was on watch, or review, by the announcing CRA at the time of the event; zero otherwise.	S&P, Moody's, Fitch
SimActsWdwEvt	Number of upgrade (downgrade) announcements made on the event country by respective other CRAs over the two-week interval $[-14, -1]$ (calendar days) before the upgrade (downgrade) event.	S&P, Moody's, Fitch
SimActsDayNonEvt	Number of upgrade (downgrade) announcements made on the non-event country by any CRA on the same day as the upgrade (downgrade) of the event country.	S&P, Moody's, Fitch

VIX	Volatility measure for the S&P 500 stock market index in the United States.	Bloomberg
Region	Dummy variable taking on a value of one if the event and non-event country belong to the same geographical region; zero otherwise.	CIA World Factbook
TradeBloc	Dummy variable taking on a value of one if the event and non-event country are members of a common major trade bloc; zero otherwise. The trade blocs are: EU, NAFTA, ASEAN, Mercosur, CARICOM, Andean Community, Gulf Cooperation Council, Southern African Customs Union, Economic Community of Central African States, Economic Community of West African States, Organisation of Eastern Caribbean States.	Authors' definition
ExpImpEvt	Importance of the event to the non-event country in terms of exports, measured as the non-event country's ratio of exports to the event country to domestic GDP.	World Bank
SizeEvt	Size of the event country, measured in logs of US dollar GDP.	World Bank
$\Delta Size$	Size differential of the event over the non-event country, measured in logs of US dollar GDP.	World Bank
$\Delta TrendGrowth$	Absolute difference between the event and non-event country's GDP trend growth, calculated for the sample period 1994–2011 on the basis of annual data using a Hodrick-Prescott filter with smoothing parameter 6.25.	World Bank

Table A.4: **Baseline regressions, downgrades — Robustness checks**

	Baseline	Ex notches ≥ 4	Ex notches ≥ 3	Crises I	Crises II
LARGE	0.0207***	0.0206***	0.0263***	0.0216***	0.0231***
	(0.0066)	(0.0068)	(0.0077)	(0.0067)	(0.0064)
InitRat	-0.0008	-0.0020	-0.0019	-0.0012	-0.0013
	(0.0017)	(0.0018)	(0.0019)	(0.0017)	(0.0017)
Δ*InitRat*	0.0008	0.0007	-0.0001	0.0009	0.0009
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
OnWatch	-0.0046	-0.0026	0.0023	-0.0050	-0.0050
	(0.0054)	(0.0056)	(0.0059)	(0.0055)	(0.0055)
SimActsWdwEvt	0.0141**	0.0173***	0.0192***	0.0122*	0.0122*
	(0.0065)	(0.0066)	(0.0074)	(0.0065)	(0.0065)
SimActsDayNonEvt	0.1477**	0.1540**	0.1538**	0.1464**	0.1466**
	(0.0648)	(0.0658)	(0.0674)	(0.0649)	(0.0649)
VIX	0.0006*	0.0008**	0.0008**	0.0006*	0.0006*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Euro				0.0186	0.0264
				(0.0115)	(0.0168)
Asian				0.1355***	0.1412***
				(0.0329)	(0.0293)

(continued on next page)

Baseline regressions, downgrades — Robustness checks (continued)

Euro × *LARGE*					-0.0150
					(0.0178)
Asian × *LARGE*					-0.0118
					(0.0396)
N	21,931	21,519	20,510	21,931	21,931
Event countries	84	84	84	84	84
Non-event countries	73	73	73	73	73
Downgrades	427	418	399	427	427
R^2	0.0423	0.0434	0.0437	0.0432	0.0432

Notes — This table shows the robustness of our baseline results on the main variable of interest, *LARGE*, to extreme downgrades and to two specific crisis episodes. For purposes of comparison, the first column reports the results from the full baseline specification for downgrades (see Panel B, Model 3 in Table 1). Since we group all rating downgrades of two notches or more into a single bin, we ensure that our findings are not driven by downgrades of four and three notches or more, respectively, by dropping those rating events from the sample. Moreover, to check that the results are not solely due to the main crisis episodes over the sample period, namely the eurozone and Asian crises, we add two dummy variables, *Euro* and *Asian*, in levels (Crises I) and as interactions with the large-change dummy (Crises II). *Euro* takes on a value of one if the downgrade was announced in 2010 or 2011 and if both the event and non-event country were members of the eurozone at that time, and zero otherwise. Similarly, *Asian* takes on a value of one for all downgrades between July 1997 and December 1998 in which both the event and the non-event country are from either of the following countries: Indonesia, Malaysia, Philippines, Singapore, South Korea, Thailand.

Table A.5: **Spillover channels, downgrades — Different trade measures**

	Trade measure			
	ExpImpEvt	*TradeImpEvt*	*ExpShEvt*	*TradeShEvt*
LARGE	0.0229***	0.0231***	0.0229***	0.0231***
	(0.0072)	(0.0073)	(0.0072)	(0.0073)
InitRat	-0.0024	-0.0023	-0.0024	-0.0023
	(0.0020)	(0.0020)	(0.0020)	(0.0020)
Δ*InitRat*	0.0013	0.0013	0.0013	0.0013
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
OnWatch	0.0011	0.0010	0.0011	0.0011
	(0.0059)	(0.0059)	(0.0059)	(0.0059)
SimActsWdwEvt	0.0139**	0.0143**	0.0139**	0.0143**
	(0.0068)	(0.0068)	(0.0068)	(0.0068)
SimActsDayNonEvt	0.1137*	0.1131*	0.1138*	0.1131*
	(0.0622)	(0.0623)	(0.0622)	(0.0623)
VIX	0.0006	0.0006	0.0006	0.0006
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Region	0.0332**	0.0308*	0.0330**	0.0309*
	(0.0168)	(0.0167)	(0.0168)	(0.0167)
TradeBloc	0.0153	0.0173	0.0151	0.0173
	(0.0116)	(0.0116)	(0.0115)	(0.0116)
Trade measure	**0.0594**	**0.0506**	**0.0259**	**0.0234**
	(0.2259)	**(0.1137)**	**(0.0657)**	**(0.0535)**
SizeEvt	0.0260	0.0275	0.0258	0.0275
	(0.0313)	(0.0316)	(0.0312)	(0.0315)
Δ*Size*	-0.0189	-0.0233	-0.0188	-0.0232
	(0.0240)	(0.0242)	(0.0240)	(0.0242)
Δ*TrendGrowth*	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
N	19,896	19,683	19,887	19,674
Event countries	79	79	79	79
Non-event countries	71	71	71	71
Downgrades	405	405	405	405
R^2	0.0436	0.0436	0.0436	0.0436

Notes — This table shows the robustness of our results on the spillover channels of downgrade announcements to different measures of bilateral trade linkages. For purposes of comparison, in the first column we report the results from the most comprehensive specification using as trade measure *ExpImpEvt*, the non-event country's exports to the event country relative to non-event country GDP (see Model 7 in Table 3). Alternatively, we use *TradeImpEvt*, which is bilateral trade (imports + exports) with the event country relative to non-event country GDP. Finally, *ExpShEvt* and *TradeShEvt* measure the event country's share in the non-event country's total exports and total bilateral trade, respectively.