

Discussion Paper

Deutsche Bundesbank
No 13/2016

The joint dynamics of sovereign ratings and government bond yields

Makram El-Shagi

(Henan University Kaifeng, School of Economics
and Halle Institute for Economic Research (IWH))

Gregor von Schweinitz

(Halle Institute for Economics Research (IWH)
and Deutsche Bundesbank)

Editorial Board:

Daniel Foos
Thomas Kick
Jochen Mankart
Christoph Memmel
Panagiota Tzamourani

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet <http://www.bundesbank.de>

Reproduction permitted only if source is stated.

ISBN 978-3-95729-251-3 (Printversion)

ISBN 978-3-95729-252-0 (Internetversion)

Non-technical summary

Research Question

This paper presents new results on the joint dynamics of sovereign ratings and government bond yields. We are mainly interested in two features of this dynamic relationship. First, we search for a vicious cycle of rating downgrades and increasing yields driving a country into default. Second, we are interested in the short- to medium-run reaction of yields to rating shocks.

Contribution

Our large dataset of 46 advanced and developing countries, starting in January 1980 allows us to account for a large number of empirical features in our estimation. Among others, we are able to capture the asymmetric effects of upgrades and downgrades, the nonlinear relationship between ratings and yields for different default probabilities, and the high persistence of ratings. Including these features in the estimation is necessary in order to gain insights into the short- and long-run relationship between ratings and yields.

Results

Our main finding is that ratings and yields slowly converge to a single equilibrium of high ratings and low yields. There is no empirical evidence at any rating level for a vicious cycle driving countries into default after an initial negative shock. The persistence of ratings is none-theless high enough such that a rating downgrade can increase the interest burden substantially. If ratings are downgraded to highly speculative (a rating of B) or below, interest rates increase sharply, and for a long time. Yet, we show that it is still highly improbable that rating agencies can be made responsible for the most dramatic spikes in interest rates.

Nichttechnische Zusammenfassung

Fragestellung

In diesem Papier werden neue Resultate zur gemeinsamen Dynamik von Bonitätsbewertungen von Staaten und Zinsen auf Staatsanleihen vorgelegt. Von besonderem Interesse sind hierbei zwei Eigenschaften dieser Dynamik. Erstens wird die Existenz eines Teufelskreises von Herabstufungen der Bonität und steigenden Zinsen überprüft. Zweitens wird neben diesem langfristigen Zusammenhang auch die kurz- bis mittelfristige Reaktion von Zinsen auf Bonitätsveränderungen untersucht.

Beitrag

Der außerordentlich große Datensatz von Beobachtungen aus 46 Industrie- und Schwellenländern seit Januar 1980 erlaubt es, in der Schätzung auf zahlreiche empirische Besonderheiten in der untersuchten dynamischen Beziehung Rücksicht zu nehmen. Unter anderem können asymmetrische Effekte von Aufwertungen und Herabstufungen, das hohe Beharrungsvermögen von Ratings sowie die hochgradig nichtlineare Beziehung bei unterschiedlichen Ausfallwahrscheinlichkeiten berücksichtigt werden. Durch die gleichzeitige Berücksichtigung dieser Besonderheiten können tragfähige Aussagen über die Kurz- und Langfristbeziehung von Bonitätsbewertungen und Zinsen getroffen werden.

Ergebnisse

Unser Hauptergebnis ist, dass Bonitätsbewertungen und Zinsen langsam zu einem einzigen Gleichgewicht mit niedrigen Zinsen und guter Bewertung konvergieren. Ein Teufelskreis hin zu einem unausweichlichen Staatsausfall kann für keine einzige Bonitätsstufe nachgewiesen werden. Allerdings ist das Beharrungsvermögen der Bonitätsbewertungen so hoch, dass eine einmalige Herabstufung die Zinslast substantiell erhöhen kann. Wenn die Bonität auf das Niveau "hochspekulativ" (eine Bonität von B) oder darunter herabgestuft wird, steigen die Zinsen für einen längeren Zeitraum deutlich an. Trotzdem ist es höchst unwahrscheinlich, dass extreme Zinssprünge an Kapitalmärkten auf Bonitätsherabstufungen zurückzuführen sind.

The Joint Dynamics of Sovereign Ratings and Government Bond Yields*

Makram El-Shagi

Henan University, Kaifeng, School of Economics
Halle Institute for Economic Research (IWH)

Gregor von Schweinitz

Halle Institute for Economics Research (IWH)
Deutsche Bundesbank

Abstract

Can a negative shock to sovereign ratings invoke a vicious cycle of increasing government bond yields and further downgrades, ultimately pushing a country toward default? The narratives of public and political discussions, as well as of some widely cited papers, suggest this possibility. In this paper, we will investigate the possible existence of such a vicious cycle. We find no evidence of a bad long-run equilibrium and cannot confirm a negative feedback loop leading into default as a transitory state for all but the very worst ratings. We use a bivariate semiparametric dynamic panel model to reproduce the joint dynamics of sovereign ratings and government bond yields. The individual equations resemble Pesaran-type cointegration models, which allow for valid inference regardless of whether the employed variables display unit-root behavior. To incorporate most of the empirical features previously documented (separately) in the literature, we allow for different long-run relationships in both equations, nonlinearities in the level effects of ratings, and asymmetric effects in changes of ratings and yields. Our finding of a single good equilibrium implies the slow convergence of ratings and yields toward this equilibrium. However, the persistence of ratings is sufficiently high that a rating shock can have substantial costs if it occurs at a highly speculative rating or lower. Rating shocks that drive the rating below this threshold can increase the interest rate sharply, and for a long time. Yet, simulation studies based on our estimations show that it is highly improbable that rating agencies can be made responsible for the most dramatic spikes in interest rates.

Keywords: Sovereign Risk, Rating Agencies, Semiparametric Models, Nonlinearities

JEL classification: C14, C25, F34, G24.

*Contact address: Gregor von Schweinitz. E-Mail: gregor.von.schweinitz@bundesbank.de, gregorvon.schweinitz@iwh-halle.de. Parts of this research were conducted while Gregor von Schweinitz was a visiting scholar at the Henan University, Kaifeng. Research of Gregor von Schweinitz was partly funded by the European Regional Development Fund through the programme “Investing in your Future” and by the IWH Speed Project 2014/01. For valuable comments and suggestions, the authors thank William Barnett, Manuel Buchholz, Stefan Eichler, Oliver Holtemöller, and Stephen Miller, as well as participants of the 2nd international workshop on Financial Markets and Nonlinear Dynamics, the WEAI annual conference 2015, and the CES 2015 annual conference. Discussion Papers represent the authors’ personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

1 Introduction

Can a negative shock to sovereign ratings invoke a vicious cycle of increasing government bond yields and further downgrades that might ultimately push a country into default? In particular, the narratives of public and political discussions, supported by some widely cited papers such as [Ferri, Liu, and Stiglitz \(1999\)](#), suggest this possibility. In this paper, we propose a semiparametric bivariate framework to analyze the interaction between sovereign ratings and government bond yields to assess whether this narrative is empirically plausible.

Credit rating agencies – especially the so-called Big Three: Moody’s, Standard and Poor’s (S&P) and Fitch IBCA – took heavy blame for the recent financial crisis and subsequent (and partly still ongoing) great recession. When a crisis happens, it is almost tautological to note that rating agencies failed to predict it. After all, if rating agencies had foreseen that specific assets were highly risky, whether they are senior tranches of asset-backed securities, corporate bonds or sovereign bonds, then over-investment in these specific asset classes would probably have been avoided. However, regarding government bonds in particular, which are the main interest of this paper, the most frequently voiced concern is not the rating agencies’ failure to predict crises, but the possibility that unfavorable rating changes cause capital flight, driving the risk premium up and thereby causing further problems that are sanctioned with another rating downgrade. When S&P downgraded the French rating from triple to double A on August 11, 2013, the first response of the French government (through the minister of finance, Pierre Moscovici) was to criticize the decision.¹ It is obvious why politicians favor the view of a vicious cycle that can befall the best of us. This argument essentially builds on a result that was first proposed by [Ferri et al. \(1999\)](#) in their widely cited paper on the Asian Flu in the late 1990s. Yet, their result is far from uncontroversial. The original study has been harshly criticized by [Mora \(2006\)](#) and [El-Shagi \(2010\)](#). Similarly, other recent papers that address the same question in the context of the European debt crisis reach very different conclusions. While [Baum, Karpava, Schäfer, and Stephan \(2014\)](#) also find evidence of a substantial impact of ratings on capital allocation, [De Vries and De Haan \(2014\)](#) focus on the finding that the increased volatility following a rating downgrade was only temporary.

The seeming contradiction between the latter two contributions highlights a key omission in the literature that the present paper aims to fill. Much of the literature criticizing rating agencies focuses on their short-term impact or aims to show that there is some arbitrariness to ratings (e.g. [Bolton, Freixas, and Shapiro, 2012](#)). Nevertheless, as noted by [El-Shagi and von Schweinitz \(2015\)](#), neither of those effects provides sufficient empirical evidence of a vicious cycle between ratings and the risk premium that can push a country from a good to a bad equilibrium. Even if a rating downgrade does increase the interest rate, the new high interest is merely paid on new and rolled over debt. That is, if the average maturity is not extremely short (which usually occurs only in countries that have low ratings to begin with), the increase of the interest rate has to be sustained for a considerable length of time to actually increase the fiscal burden.² To demonstrate the

¹See, for example, <http://www.bloomberg.com/news/articles/2013-11-08/france-credit-rating-cut-to-aa-by-s-p-on-weak-growth-prospects>.

²Rating downgrades can have a negative impact on government finances beyond their direct effects. Due to the so-called sovereign ceiling (an implicit rule whereby companies only rarely obtain a better

existence of a vicious cycle that inevitably leads to default unless the country is affected by subsequent positive shocks, it is necessary to prove the existence of explosive behavior in ratings and yields under particular conditions (such as a threshold for risk beyond which behavior becomes explosive). A weaker form of a vicious cycle – which is still strong enough to cause default, at least hypothetically – is the combination of the strong self-reinforcing behavior of rating downgrades, combined with high persistence in rating levels in the absence of shocks and/or a lesser degree of self-reinforcing behavior for rating upgrades. Under those conditions, a negative rating shock might multiply before the system stabilizes (for a while) at a far lower rating, thus increasing interest rates and the fiscal burden. While technically not a second equilibrium, this kind of self-reinforcement with transitory stabilization and slow recovery would qualify as a vicious cycle in the sense used by the major critics of rating agencies.

In our paper, we augment a recently suggested approach by [El-Shagi and von Schweinitz \(2015\)](#), who simultaneously model long-run relationship and short-run dynamics in a model that explicitly allows for multiple equilibria. That is, [El-Shagi and von Schweinitz](#) avoid the validity problems associated with estimated long- or short-run dependencies that exist when only one of the two horizons is explicitly modeled. While the original paper pays little attention to the short-run dynamics, mostly modeling them to allow robust inference about the long-run relationship, the current paper focuses on the joint dynamics of government ratings and sovereign bond yields.

Those dynamic aspects have been widely ignored in the previous empirical analysis of sovereign ratings. While there is a huge body of literature on the immediate impact of ratings, the dynamic part of the analysis usually takes the form of a simple event study. Rare exceptions (such as [De Santis, 2012](#)) derive impulse response functions based on VAR models, that treat ratings as a continuous variable. Contrary to a standard VAR approach, we are able to account for the ordinal nature of ratings, include nonlinearities and asymmetries. Yet, this comes at a cost. The model we estimate is computationally demanding even in the bivariate form presented in this paper. Thus, we cannot account for the role of other macroeconomic variables. Since the interest rates react fairly quickly, and institutional variables are well covered by fixed effects, we do however believe that the benefits of our approach outweigh the losses.

We augment the dynamic part of the model of [El-Shagi and von Schweinitz](#) substantially, most importantly by allowing for asymmetric effects of both rating and yield changes. This modification allows for the type of vicious cycle that is driven by short-run dynamics rather than by convergence to a bad equilibrium. We confirm their finding that there is strong evidence of a single good equilibrium. At no point is the typical risk premium associated with a rating sufficient to justify further rating downgrades. Yet, we find that downgrades can come at a substantial cost. Over the short run, rating changes tend to mildly reinforce themselves, slightly increasing the risk of further downgrades. Additionally, we do observe sharply increasing risk premia when ratings fall below the B+ level. Due to the high persistence of ratings, those interest premia can last many years, thus generating substantial macroeconomic costs without being vicious cycles. Yet, simulation studies based on downgrade episodes from the past decades show that unfavorable developments that have occasionally been observed after initial downgrades cannot be

rating than their home country), government downgrades will negatively affect ratings of companies, increasing interest payments and worsening the economic outlook ([Durbin and Ng, 2005](#)).

explained through the common joint dynamics of ratings and yields. They are thus most likely driven by an actual change in the fundamentals (or a correction in their assessment). The remainder of this paper is structured as follows. In Section 2, we briefly review the previous empirical findings, provide some introductory stylized facts on the dynamics of ratings and yields and discuss some of the associated measurement problems. In Section 3, we explain our econometric model and the methods employed. Section 4 presents our results on the long- and short-run relation between ratings and yields, including the scenario simulations reproducing previous downgrade episodes using our model. Section 5 concludes.

2 Previous evidence and stylized facts

In particular, the critics of rating agencies believe them to hold – as [Ferri et al. \(1999\)](#) puts it – “a tremendous power”. They claim that the major rating agencies, Moody’s, S&P and Fitch – frequently dubbed *the Big Three* – strongly shape the behavior of international investors, despite their finding that they add only limited information.

There is a long and extensive literature that aims to identify the factors that influence sovereign credit rating decisions. In general, debt sustainability measures, the degree of economic development and the default history ([Cantor and Packer, 1996](#); [Gärtner, Griesbach, and Jung, 2011](#)) as well as political stability and governance indicators ([Mellios and Paget-Blanc, 2006](#)) are found to be important. One of the most general results regarding ratings is that agencies react to past fundamentals rather than successfully predict future shocks ([Cantor and Packer, 1996](#); [Reisen and von Maltzan, 1999](#)). However, while providing only little new information in normal times, rating agencies aim to reestablish their reputation after missing an emerging crisis, responding with overly restrictive downgrades ([Ferri et al., 1999](#); [White, 2010](#)). This creates capital flight, increasing the interest rate and thus the fiscal burden of the government, which is driven even closer to an actual default. In the case of sovereign ratings, this problem is particularly severe because sovereign ratings also serve as the so-called sovereign ceiling, that is, a best-case rating for all but a few companies with headquarters in that country ([Durbin and Ng, 2005](#)). Therefore, capital flight can also affect the private sector, slowing the economy and eroding the tax base. This problem is exacerbated by the reliance of regulators on ratings. These regulations often indicate that assets with ratings below a certain threshold are not considered as “investment” but speculation; thus, these assets are strongly restricted or penalized ([White, 2010](#)). Of course, this argument of a self-fulfilling prophecy is controversial, e.g., [El-Shagi \(2010\)](#) notes the inconsistency in simultaneously claiming that rating agencies obviously and systematically err and noting that the majority of investors does not recognize this and follows the rating agencies despite their shortcomings.

Yet, the responsibility of rating agencies for financial market turbulence is a well-established narrative, particularly in politics. The ongoing debate in the European Union about the need for a large European rating agency is a textbook example of the magnitude of the influence that is attributed to rating agencies by politicians. The fear of the Big Three – all headquartered in the US and UK – became so severe that even major politicians in the European Union (such as Rainer Brüderle, the then head of the FDP (libertarian) group in the German Bundestag and a former federal minister of economic affairs) became worried that American institutions tried to deliberately exploit their influence to hurt the

Euro (area).³

The original example that fueled the debate about the dangers of rating agencies is the so-called Asian Flu in the late 1990s. [Ferri et al. \(1999\)](#) argue that they played a significant role in accelerating the crisis. Their argument is based on the first downgrades of Thailand (October 1997, from A- to BBB), Malaysia (December 1997, from A+ to A) and Indonesia (December 1997, from BBB to BB+). Their conclusion is challenged from two directions: [Mora \(2006\)](#) finds ratings to be sticky rather than procyclical.⁴ [El-Shagi \(2010\)](#) goes one step further and documents that there were many rating adjustments following these first downgrades, the last and most significant of them occurring shortly before the end of the crisis and sometimes even after. That is, there is at least as much evidence that rating agencies merely follow the market rather than triggering or worsening a crisis by downgrading a country.

The second prime example proposed by critics of rating agencies relates to the ongoing European debt crisis. It was argued that the actions of (US-based) agencies unduly increased market pressure on European periphery countries, increasing their government bond yields to unsustainable levels, thus triggering a public debt crisis with severe long-run macroeconomic costs. [Arezki, Candelon, and Sy \(2011\)](#) find that some downgrades in the Euro area, such as the one of Greece from A- to BBB+ by Fitch on December 8, 2009, had systematic spillover effects to other European countries (see also [Beirne and Fratzscher, 2013](#)). That is, the downgrading of Greece is found to have increased not only the CDS spreads of government bonds (a measure of credit default risk) in Greece but also in a number of other European countries. The authors claim that these spillovers alone may trigger further financial instability. However, their results for spillover effects are quite heterogeneous and thus may not be strong enough to support their claim in a more general setting. The findings of [Afonso, Furceri, and Gomes \(2012\)](#) point to a much more balanced view of the question at hand. First, they find a strong and asymmetric effect of rating changes. They also find that these changes are mostly unanticipated one to two months ahead (with evidence of limited bidirectional Granger causality afterwards), concluding that there may be limited evidence of the market power of rating agencies. However, markets react quickly, which partly refutes the argument that ratings can actually drive a country into default. A similarly balanced position is taken by [Gärtner et al. \(2011\)](#), who use yearly data and a fundamental estimation of ratings as in [Cantor and Packer \(1996\)](#) to determine non-fundamentally justified aspects of ratings. They argue that these aspects of rating decisions also affect yield spreads. That is, an erroneous (arbitrary) downgrading decision might trigger yield increases, which would open the possibility of further downgrades in the future. This result is partially challenged by the finding of [De Vries and De Haan \(2014\)](#) that credit ratings and yields have recently become disentangled: after the summer of 2012, the yield levels of European periphery countries decreased quickly while ratings stayed at very low levels. The authors attribute this to either unconventional monetary policy or increasing conservativeness among credit rating agencies. However, their econometric model does not include short-

³See for example <http://www.welt.de/politik/deutschland/article13756954/Bruederle-fordert-europaeische-Ratingagentur.html>.

⁴The stickiness of ratings may have two causes: first, they could be due to shortcomings in information processing; second, to a tendency of rating agencies to avoid rating reversals if default probabilities fluctuate near the boundary of two discrete rating classes ([Löffler, 2005](#)).

run effects and allows for varying effects of different rating levels only to a very limited extent. Therefore, the econometric model of [De Vries and De Haan](#) may be misspecified, and the slow adjustment of ratings may simply be due to their general stickiness.

After briefly introducing our dataset, we will present a few stylized facts regarding the joint dynamics of ratings and yields to motivate our own econometric approach in the following subsections. In particular, we will argue that it is necessary to consider (a) both short- and long-run effects in the model, (b) nonlinearities in the long-run relation, and (c) asymmetries in the short-run relation.

2.1 Measurement and sample selection

Ratings: To maximize data coverage, we use the ratings of foreign currency denominated government bonds as provided by Moody's, S&P and Fitch.⁵ Due to higher international demand, foreign currency bonds have often been rated long before the agencies considered domestic currency denominated bonds. Yet, the ratings for different bonds issued by the same country are generally highly correlated.⁶

The three agencies use grades to assess the probability of a default over the medium- to long-term, where better grades correspond to lower default probabilities. The names of grades differ across agencies. However, this difference is merely notational ([Cantor and Packer, 1996](#)). Therefore, grades can be easily compared and transformed into the ordinal scale given in [Table 1](#). Throughout the remainder of this paper, we will use the S&P notation.

As market movements are often found to be strong around rating announcements (which provide new signals), the most important signal is probably provided by the first agency to adjust its rating. However, there seems to be some evidence for specialization and leadership of the agencies in specific markets ([Hill and Faff, 2010](#)), which is why we should include information from all agencies rather than concentrate on a single agency. In addition to accounting for the timeliness of new information, average ratings provide an implicit safeguard against random judgment errors.

Where a rational representation of the rating is required, we use the mean rating of all three agencies (see also [De Vries and De Haan, 2014](#)). However, for most of our analysis, we aim to maintain the ordinal nature of the ratings, contrary to the majority of the literature considering rating levels. The reason for this is the non-linear relationship between ratings and default probabilities. Our rating class dummies are generated by rounding the mean rating to the next integer and considering the joint rating as belonging to the corresponding rating class defined in [Table 1](#). It is fairly well documented that the ratings of different agencies seldom differ by much, even during times of higher uncertainty when rating agencies adjust their assessments more frequently ([Ferri et al., 1999](#)).⁷ Therefore, this averaging generally corresponds to the majority rating of the three agencies.

As a robustness check, we also employ the median of the three agencies. However, the differences in results (see [Figures A4 and A5](#) in the appendix) are marginal.

⁵The data are collected from <http://countryeconomy.com/>

⁶The correlation is 94% for ratings, and 74% of domestic rating changes occur at the same time as foreign rating changes.

⁷Empirically, the standard deviation across agencies is 0.77 notches, the mean absolute difference to the average is 0.5 notches. That is, rating agencies are on average far less than one notch apart.

Table 1: Rating grades and transformation

Grade	Moody's	S&P	Fitch	Assigned Value
Prime	Aaa	AAA	AAA	24
High grade	Aa1	AA+	AA+	23
	Aa2	AA	AA	22
	Aa3	AA-	AA-	21
Upper medium grade	A1	A+	A+	20
	A2	A	A	19
	A3	A-	A-	18
Lower medium grade	Baa1	BBB+	BBB+	17
	Baa2	BBB	BBB	16
	Baa3	BBB-	BBB-	15
Non-investment grade speculative	Ba1	BB+	BB+	14
	Ba2	BB	BB	13
	Ba3	BB-	BB-	12
Highly speculative	B1	B+	B+	11
	B2	B	B	10
	B3	B-	B-	9
Substantial risks	Caa1	CCC+	CCC+	8
Extremely speculative	Caa2	CCC	CCC	7
	Caa3	CCC-	CCC-	6
In default with little prospect for recovery	Ca	CC	CC	5
		C	C	4
In default	C	D	DDD	3
			DD	2
			D	1

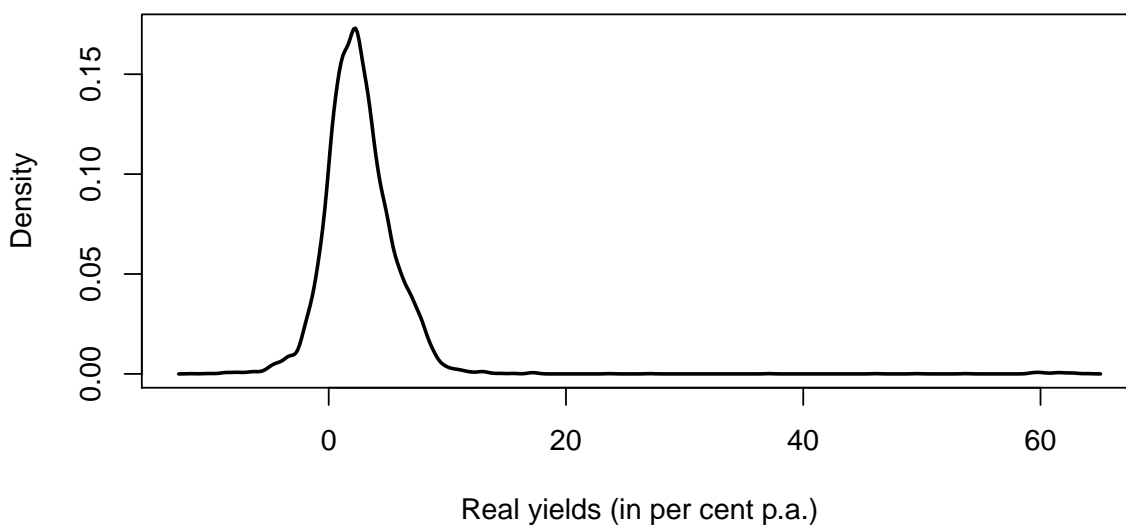


Figure 1: Density of yields

Yields: We measure the risk premium through the real government bond yield on sovereign bonds with a maturity of 5 to 10 years, denominated in the domestic currency.⁸ Again, this choice is mostly enforced by data availability and feasibility. While real yields are strongly driven by default risk, they simultaneously capture other factors, such as the degree of market liquidity and global risk aversion (von Hagen, Schuknecht, and Wolswijk, 2011). Thus, they go slightly beyond the claim of rating agencies to consider solely the probability of default. The pure credit default risk component could also be captured by the prices of CDS, which are essentially insurance contracts against the event of default. However, using CDS prices would reduce our sample significantly, especially reducing the number of observations with low ratings.

Conversely, yields on domestic currency bonds are widely available for a broad range of countries for extended periods. The 5 to 10 year maturity mirrors the risk horizon of rating agencies. To produce comparable yields for different countries, we deflate the yields using the year-on-year inflation of the previous 12 months. While past inflation is not a perfect measure of inflation expectations, which would be required to compute expected real returns, inflation time series are available for far more countries and periods than survey data or other more direct measures of expectations.

Figure 1 shows the distribution of those real yields. In approximately 85% of the periods, yields are positive, and they have a mean and median near 2.5%. This is plausible, as investors would only be willing to accept yields below current inflation rates if they can

⁸Thus, there exists a currency mismatch between ratings and yields. However, the correlation between domestic and foreign ratings is very high. The results are broadly robust to using domestic instead of foreign ratings or to using the emerging market bond index (EMBI, produced by JP Morgan) in US dollars. Both alternative measures for ratings or yields come at the cost of a significantly reduced dataset: domestic ratings are only available much later, and the EMBI only covers emerging markets.

be nearly certain that the latter will decrease in the near future. Table A2 suggests that the distribution of yields does not differ much among countries. Only in Greece between October 2011 and January 2014, real yields have been higher than 20%.⁹ In all other countries, high nominal yields were usually accompanied by high inflation rates, leading to a narrow distribution of real yields.

On average, we expect developing countries to have higher credit risk and inflation volatility, both contributing to higher average real yields. However, these expectations are not fulfilled by the summary statistics in Table 2.¹⁰ Advanced economies have, on average, higher real yields. This counterintuitive result can largely be attributed to Greece. If we exclude it, average yields are comparable between groups, with a much lower standard deviation (i.e., volatility) in advanced economies.

Table 2: Summary statistics of ratings and yields

Variable	IMF classification	mean	sd	min	max
ratings	total sample	20.75	3.93	4.50	24.00
	Advanced	22.51	2.36	4.50	24.00
	Developing	14.83	2.76	7.67	20.67
	Transition	18.21	1.97	13.67	22.00
yields	total sample	2.74	4.06	-11.59	64.00
	Advanced	2.83	4.35	-8.63	64.00
	Developing	2.53	3.27	-11.59	17.34
	Transition	2.28	2.07	-3.10	9.78

Sample: The indicators defined above are available for an unbalanced monthly panel of 46 countries from January 1980 to January 2014. It covers 27 advanced economies as well as four Eastern European transition economies and 15 developing economies. Transition and developing economies often have low data availability: there are several countries for which yield data are only available after 2001. However, we still obtain approximately 9,100 observations in total.

2.2 Stylized Facts

Nonlinearities: As ratings measure default probabilities in a nonlinear way (Löffler, 2005), it is not very likely that ratings (or their assigned values) can be used linearly. In the present context, this limitation especially holds when the potential effect of an investment grade threshold and the possibly nonlinear relationship between default probabilities and sovereign yields is taken into account. Therefore, different authors used various transfor-

⁹In a second exception, the nominal yield in Sri Lanka exceeded 450% in July 2010. This value is so unreasonably high that we exclude it from our estimation as an outlier. All other available observations are included in our sample.

¹⁰The classification of countries follows IMF (1997). A detailed list of countries with data availability can be found in Table A1 of the appendix.

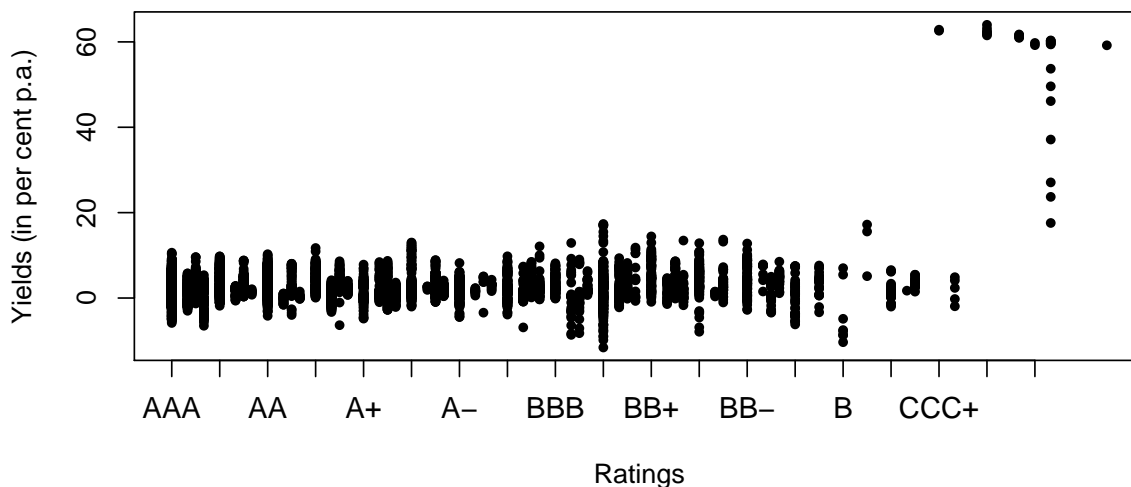


Figure 2: Scatterplot of ratings and yields

Note: The rating scale (notation of S&P) is inverted in this and all following graphs. That is, the x-axis displays increasing ratings (i.e., lower risk) when going from right to left.

mations of ratings when they seek to explain (medium-run) yield movements.¹¹ Larraín, Reisen, and von Maltzan (1997) test both a linear and a logistic transformation of ratings; Ferri et al. (1999) provide evidence for an exponential conversion, which is also used (along with linear and cubic conversions) by Gärtner and Griesbach (2012).

Given that there is no agreement on how to transform ratings such that a linear relationship between transformed ratings and yields can be expected, we think that a more flexible transformation should be employed. This transformation will be presented in the following section along with our method of addressing the short- and long-run interactions of ratings and yields.

In Figure 2, it can be seen that the level relation between ratings and yields is basically flat for all but the lowest rating classes, even well below the investment grade threshold between BBB- and BB+. How can this “non-relation” be reconciled with the anecdotal evidence and the reasonable assumption that (inverted) ratings and yields should be positively correlated, as they are both measuring the same thing? First, ratings are constant for long periods of time during which yields may slowly adjust to new risk levels. Second, real yields are affected by many more factors in addition to sovereign risk. Rather than interpreting the flat slope as the absence of a risk premium, it should be interpreted as a risk premium of an order of magnitude that is overshadowed by the general variance of interest rates. Only if risk becomes substantial, does the risk premium begins to quickly increase. This is not purely related to the nonlinearity of risk measurement; it is actually in line with theory, which predicts that the risk premium goes to infinity when the default probability approaches one.

¹¹In event studies with windows of only a few days, the nonlinearity of ratings does not play such a large role.

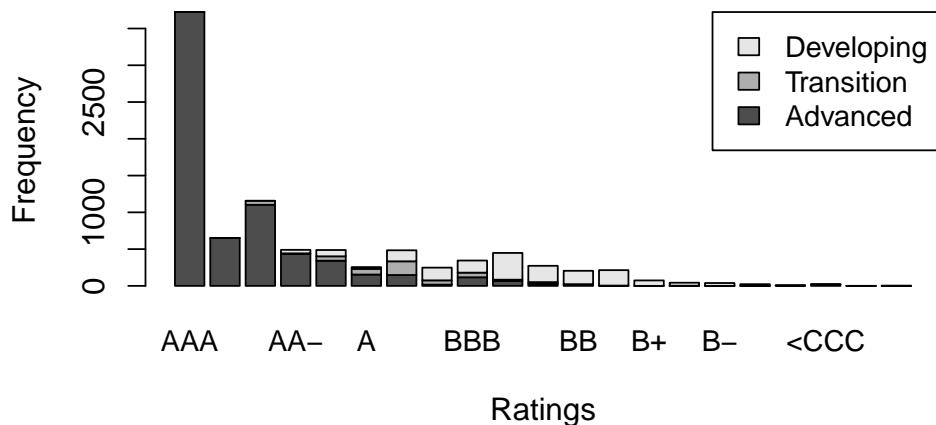


Figure 3: Histogram of ratings

Persistence: Figure 3 shows the histogram of average ratings. Nearly 60% of our ratings are high-grade, which can partly be explained by the greater data availability of advanced economies (Table A1 in the appendix), which tend to have higher ratings. In our dataset, ten of these countries have never received a rating below AA, while only two industrialized countries (Greece and Israel) have never achieved a rating above A+; see Table A2 in the appendix. Transition and developing countries, however, tend to have much lower ratings. Among that group, only Slovenia achieved a “high grade” rating until they were downgraded during the European debt crisis in January 2012. Yet, the distribution features no bimodality that would suggest that countries below a certain rating face consecutive downgrades until they eventually default.

While this is at least some indication of convergence to a good equilibrium, ratings are characterized by enormous persistence. The share of observations with a rating change is a mere 3.5%. In our entire sample, we observe 187 upgrades and 133 downgrades (which are on average slightly larger than upgrades). This low share of periods with rating changes is not driven by countries that have already achieved peak ratings. Even when excluding observations with ratings of AAA and AA+, the probability of a rating adjustment barely exceeds 5%.

This degree of stickiness makes conducting an analysis in a traditional AR framework difficult, even if the variables of interest are technically stationary in the sense that they slowly return to a unique equilibrium (rather than an equilibrium curve as in a cointegration setting). Yet, ignoring the short-run dynamics would imply ignoring the shocks that drive ratings away from this long-run equilibrium, which is why it is crucial to use a model that combines short- and long-run effects.

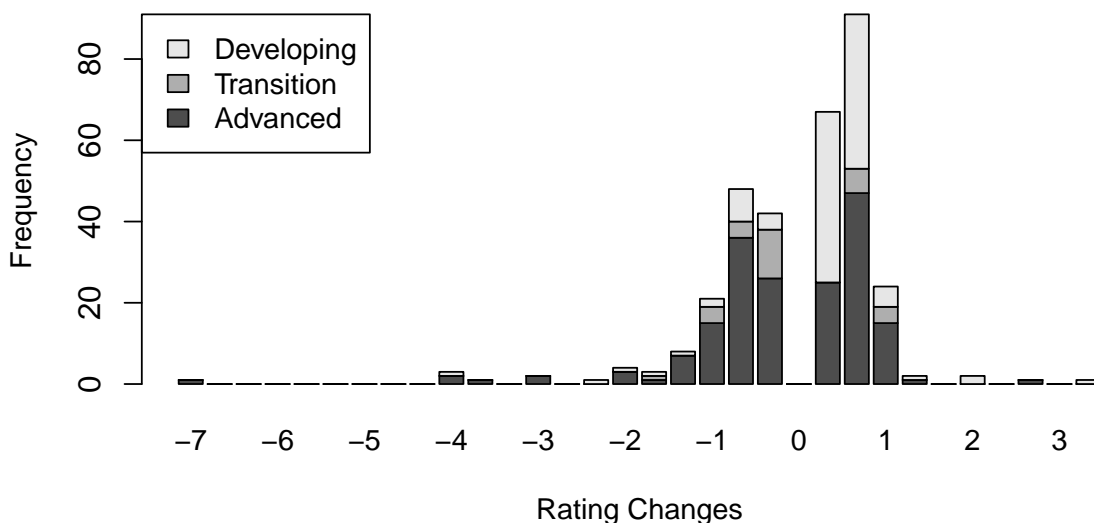


Figure 4: Histogram of rating changes

Note: Differences between bars are 1/3, i.e., a rating change by one notch by one of the three agencies. Periods with no rating changes excluded

Asymmetries: Even when considering the magnitude of rating changes in a single event, ratings move very slowly; see Figure 4. If one or more agencies adjust their assessments, the average moves by at most one notch in more than 80% of downgrades and 95% of upgrades. Yet, another reading of those numbers is that rating movements of more than one step are four times more likely for downgrades than they are for upgrades. That is, while we often observe a staggering of rating adjustments, with one or two agencies moving first and the third following the next month, downwards dynamics seem much more intense than the recovery. Such strong downgrades occur most often in advanced economies (where there is admittedly more room for downgrades, on average). This impression is confirmed if we look at longer horizons. There, we can see that an initial downgrade may trigger several more in advanced economies, which amounts to a large total downgrade: the six largest cumulative rating downgrades over one year, between 4.5 and 9.5 notches, occurred in advanced economies. However, cumulative rating improvements are more or less equally distributed over advanced, transition and developing economies.

Asymmetry is not limited to the magnitude of a change, but more importantly, to the dynamics of change. Figure 5 shows the development of yields (normalized to 100 in the month of the rating change) at longer horizons of 12 months before and after a rating change. While this does not provide conclusive evidence, the figure roughly identifies two stories. During periods of downgrades, ratings changes, more often than not, seem to occur jointly with a peak in the interest rate. That is, the data confirm the finding of [El-Shagi \(2010\)](#) on a much broader level that rating downgrades occur late in an adjustment process, immediately before yields decrease again. Yet, during times of upgrades, rating

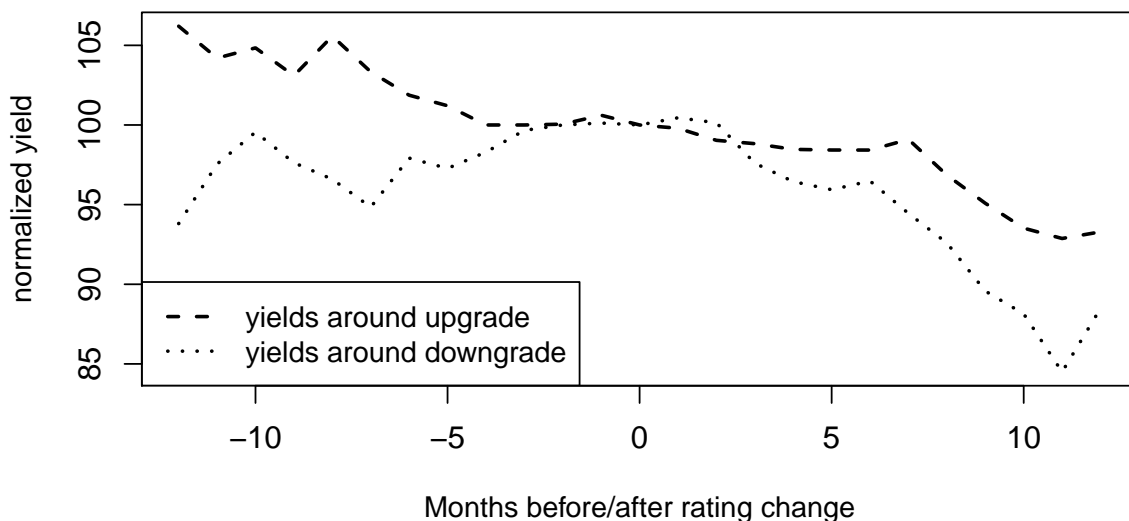


Figure 5: Development of yields in a two year window around rating changes

changes are adjustments during an ongoing decline in risk premia.

While the general finding of asymmetry is shared by most of the rich literature on event studies of rating changes, many of those studies imply much stronger interest rate dynamics. However, most of those studies (e.g., [Ferri et al. \(1999\)](#), [Kiff, Nowak, and Schumacher \(2012\)](#), [Afonso et al. \(2012\)](#)) are limited by the fact that they use comparably short windows of approximately 14 days before and after a rating announcement. While the short-run fluctuations they analyze may be highly relevant for speculation purposes, they seem negligible given a longer perspective.

Event studies usually find strong effects of rating announcements on yields on the days before and after the event. It is sometimes argued that this is a sign of an anticipation effect ([Hill and Faff, 2010](#)). However, the argument that ratings are sticky ([Mora, 2006](#)) or lagging ([Reisen and von Maltzan, 1999](#); [El-Shagi, 2010](#)) seems more convincing. That is, we should expect Granger causality in both directions in general rather than only for short windows around events.

3 Model and estimation technique

In this section, we present the econometric model we use to answer the questions regarding (a) the possible existence of multiple long-run equilibria in the relation between ratings and yields and (b) the short-run dynamics after shocks. Our model is a simultaneously estimated bivariate two-equation model consisting of a continuous yield equation and an ordered probit ratings equation, which provides several extensions of the simpler model in [El-Shagi and von Schweinitz \(2015\)](#). Each equation is inspired by the structure proposed by [Pesaran, Shin, and Smith \(2001\)](#) in his seminal paper on the bounds cointegration test.

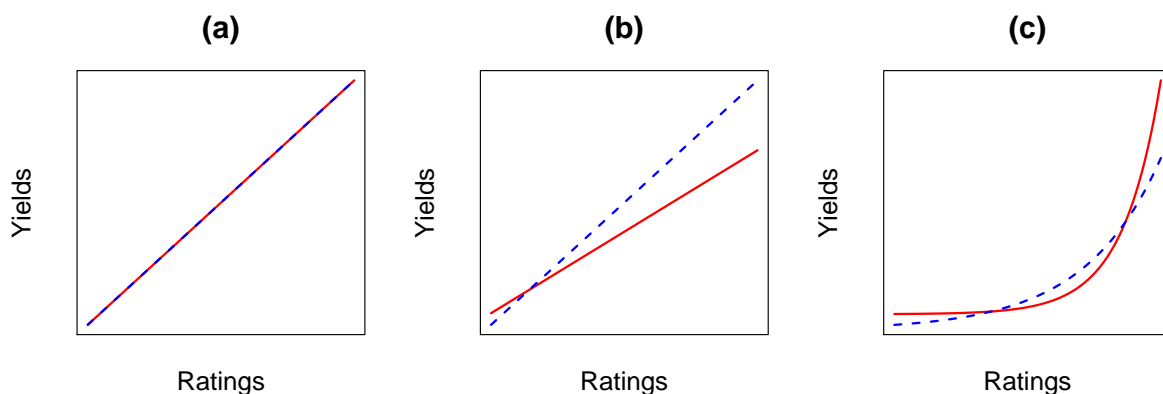


Figure 6: Exemplary long-run relations of ratings (depending on yields, blue) and yields (depending on ratings, red)

A priori, there is no reason for a unique long-run equilibrium or a unique long-run relationship (implying an infinite number of equilibria along the relationship). Figure 6 displays three pairs of stylized long-run relationships. The last case (c), inspired by the theoretical model of [Gärtner and Griesbach \(2012\)](#) shows a relationship with two equilibria: a good and stable equilibrium of low yields and high ratings, and a bad and unstable equilibrium of high yields and low ratings. If a country receives a rating below the bad equilibrium, this sets off the vicious cycle of rising yields and downgrads already described above.

As indicated by previous results and stylized facts, our model needs to account for the features described in section 2. In subsection 3.1, we explain how we perform a simultaneous estimation of the short- and long-run relations and how we perform the identification of (possibly many) long-run equilibria in principle. Subsection 3.2 presents the specific model. In subsection 3.3, we explain how we incorporate nonlinearity in rating levels into the model while ensuring a certain degree of smoothness. Subsection 3.4 explains the bootstrap procedure needed to account for the time and cross-sectional heterogeneity of shocks, while subsection 3.5 describes how we adapt the identification of long-run equilibria (and testing for multiple equilibria) due to the existence of asymmetric short-run effects.

3.1 Overview: the model structure

Our model is an extension of the model by [El-Shagi and von Schweinitz \(2015\)](#), which in turn relies on an equation structure proposed by [Pesaran et al.](#):

$$\Delta x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 y_{t-1} + \sum_{i=1}^p \gamma_i \Delta x_{t-i} + \sum_{j=0}^p \eta_j \Delta y_{t-j} + \varepsilon_t \quad (1)$$

The reason for adopting this model setup is that the equation structure estimated by [Pesaran et al.](#) allows valid inference (i.e., consistent and unbiased estimates) regardless of whether the two included variables x and y show unit root behavior.

The estimator only provides a single long-run relationship identified through the ratio of β_1 and β_2 . However, we would like to estimate the long-run relationship implied

by changes in ratings and yields separately, thereby identifying the equilibrium yield as a function of the current rating and the equilibrium rating as function of the current yield. We require some assumptions to allow for this kind of semistructural identification. To achieve identification, we borrow from the literature on structural VAR and impose restrictions on the contemporaneous effects.¹²

We assume that ratings have a contemporaneous effect on yields, but not vice versa. This assumption is reasonable for two reasons. First, ratings are characterized by high persistence at monthly frequencies. Second, and more importantly, we measure ratings on the first day of the month and average yields over the month. That is, “contemporaneous” yields contain the information of up to a full month after the measurement of ratings. The identification of the contemporaneous effect of ratings on yields is achieved by including the contemporaneous rating change in the yield equation (following Pesaran et al. quite closely). This provides causal identification equivalent to the standard SVAR, which uses Cholesky decomposition of the covariance matrix. However, our approach comes at a considerable advantage. The ordered probit model used to explain (discrete) rating changes does not provide residuals. Therefore, a Cholesky decomposition could only be achieved by simultaneous estimation of the complete system of both equations and the covariance matrix. Our identification strategy, on the contrary, orthogonalizes residuals and thus (following the seminal argument by Sims, 1980) allows for block-wise estimation. Through the approach outlined above, rather than estimating *the* long-run relation, we estimate the long-run relation *implied by each* of the change variables (i.e., the first difference of interest rates and the presence of up- or downgrades). Finding two individually significant yet different long-run relations in both equations implies that – while having enough persistence to differentiate between long- and short-run effects – the variables are not technically cointegrated in the traditional sense. Rather, they are either stationary or exhibit some type of regime-switching behavior.

Consider again the exemplary long-run relations featured in Figure 6 (a) and (b). Only in the first case (a), where *both* long-run relations are the same, is there an infinite number of equilibria in the sense that none of the points on the curve features a model intrinsic tendency for deviation (or no equilibrium in the sense that no single point acts as a permanent center of gravity). That is, a shock can potentially have permanent impact, as in a nonstationary system. If the lines intersect once as in case (b), only this combination of yields and ratings is stable over the long run, and both yields and ratings will move toward that combination over the very long run. That is, despite high persistence, as indicated by the very existence of a long-run relationship, the system is essentially stationary. Because our model should accommodate both the traditional cointegration case (a) and the more complex case (b), it is essential that the model remains well identified regardless of whether there is stationarity or not.

In order to allow our model to identify the frequently feared situation of a *good* and a *default* equilibrium (see Figure 6 (c)), we need to allow for nonlinearities in the long-run relations. To this end, rather than including the rating as a continuous variable, we use

¹²In the original framework of Pesaran et al., the contemporaneous correlation is captured through η_0 , i.e., the contemporaneous first difference of y without making any assumption as to whether this actually reflects an impact of y on x or vice versa. To identify a unique long-run relationship, this is not consequential. Equation 1 can easily be solved for Δy_t as dependent variable without changing the relation of the coefficients of the level variables β_1 and β_2 .

a semiparametric approach to estimate a functional form over a set of rating dummies.

3.2 The model

The yields model: Our interest rate equation takes the form:

$$\begin{aligned} \Delta i_t = & \beta_0 + \beta_1 i_{t-1} + \sum_{c=7}^{24} \beta_c r_{c,t-1} \\ & + \sum_{l=1}^{p_i} (\alpha_{l,p} 1_{\Delta i_{t-l} \geq 0} + \alpha_{l,n} 1_{\Delta i_{t-l} < 0}) \Delta i_{t-l} \\ & + \sum_{l=0}^{p_i} (\gamma_{l,p} 1_{\Delta r_{t-l} \geq 0} + \gamma_{l,n} 1_{\Delta r_{t-l} < 0}) \Delta r_{t-l} + \varepsilon_t, \end{aligned} \quad (2)$$

where i is the interest rate, $\Delta r \in \{-1, 0, 1\}$ the change of the rating, r_c a rating dummy, t the time index, and $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ an error term. We use up to p_i lags and the contemporaneous rating changes. The possibly asymmetric effect of past changes is captured by the different coefficients on positive ($\alpha_{l,p}, \gamma_{l,p}$) and negative ($\alpha_{l,n}, \gamma_{l,n}$) rating and yield changes. The rating dummies are defined such that r_c indicates that the rating is lower than or equal to c . That is, for a rating of A, the first 13 dummies (7-19) would be set to 1. For estimation purposes, we combine all default ratings (CCC- and below) in one reference class (captured in the constant) because very few ratings at or below “Extremely Speculative” (CCC) are observed. This leaves us with 18 different rating dummies. Defining the rating dummies cumulatively (i.e., a dummy equals 1 not only for countries with the respective rating but also for countries with better ratings) simplifies the assessment of whether the difference in effects between adjacent rating classes is statistically meaningful. The modeling of rating levels with dummies introduces a potentially strong degree of nonlinearity. Thus, we are also able to detect possible structural breaks, which could for example exist at or around the investment grade threshold (a rating of B or below). In particular, our model allows to identify long-run relationships as depicted in Figure 6 (c), where the behavior of the system can become explosive beyond a certain threshold and the country is driven automatically into default.

While rating dummies refer to the average rating class of the three rating agencies, rating changes Δr capture every rating change (both smaller and larger than one notch). We do this because the rating decision of a single agency may have a short-run effect on yields even if the average rating class does not change.

The asymmetric effects of negative and positive changes of both ratings and yields is an extension of the model in [El-Shagi and von Schweinitz \(2015\)](#). We test models with asymmetric effects against its symmetric counterpart at different numbers of lags and find that the symmetric model is always significantly outperformed by the asymmetric one.

The rating model: Exploiting the categorical nature of the ratings, we estimate an ordered probit model for the change of the rating, again allowing for asymmetries:

$$\begin{aligned}
r_t^* = & \psi_1 i_{t-1} + \sum_{c=7}^{24} \psi_c r_{c,t-1} \\
& + \sum_{l=1}^{p_r+1} (\rho_{l,p} 1_{\Delta i_{t-l} \geq 0} + \rho_{l,n} 1_{\Delta i_{t-l} < 0}) \Delta i_{t-l} \\
& + \sum_{l=1}^{p_r} (\omega_{l,p} 1_{\Delta r_{t-l} \geq 0} + \omega_{l,n} 1_{\Delta r_{t-l} < 0}) \Delta r_{t-l} + \eta_t.
\end{aligned} \tag{3}$$

$$\begin{aligned}
\text{rating downgrade if: } & r_t^* < \mu_1, \\
\text{rating upgrade if: } & r_t^* > \mu_2
\end{aligned}$$

where r^* is the latent variable that is linearly dependent on the explanatory variables, and $\eta_t \sim \mathcal{N}(0, 1)$ is again an error term. We include $p_r + 1$ lags of yields to remain consistent with the framework of [Pesaran et al. \(2001\)](#) (which uses an additional lag of the “exogenous” variable) while adhering to our identification strategy.

Specification: Lags are selected in each equation separately by the Bayesian information criterion with a maximum of six lags. We adjust the dataset to estimate both equations with exactly the same observations.

Inclusion of fixed effects: Fixed effects are usually avoided in ordinal models because they are no longer identified when a cross-sectional unit is constantly in one of the extreme groups. While this is true for the level of the rating (with several advanced economies always being rated AAA), this does not occur when using rating changes as the dependent variable (because no country is permanently down- or upgraded). Thus, fixed effects for N countries could be easily included in the form of $N - 1$ dummies in both equations.¹³ In our view, the inclusion of country fixed effects has both advantages and disadvantages. On the one hand, they account for possible differences between countries that may affect yields and ratings. Due to the structure of our econometric model, short- and long-run effects from other sources on our two variables of interest are simultaneously captured by fixed effects (if they are not already captured by the lag structures). On the other hand, many rating classes are only observed in a small number of countries, and (conversely) only few countries experience a large range of ratings. Therefore, fixed effects could also blur the relation between yields and ratings by capturing the effects of individual rating classes instead of the effects for individual countries. As an econometric distinction between these two very different channels is hardly possible, we present the results without fixed effects as a baseline and use fixed effects estimation as a robustness check.

¹³Incidental parameters are sometimes a problem in ordered-choice models with panel data. However, this is of no concern as we have a large T dimension and relatively few countries N .

3.3 Exploiting the ordinal nature of ratings to model nonlinearity

The semiparametric estimation: When estimating equations (2) and (3) through standard estimators, the representation of the ratings through a series of dummies is problematic. Because some rating classes are observed in very few situations, time and country idiosyncrasies would drive the estimated coefficients rather than the actual impact of a rating of the corresponding class. While technically allowing “nonlinearities” in the impact of ratings (compared to treating ratings as pseudo-continuous), this creates considerable and economically unwarranted differences in the effects of very similar ratings. Essentially, despite modeling ratings through class dummies, we would like to obtain a well-behaved smooth function over rating classes, unless there is strong evidence (as in many data points) suggesting otherwise.

To achieve this objective, we borrow from an approach suggested by [Breitung and Roling \(2015\)](#) for mixed frequency data sampling (MIDAS). In MIDAS approaches, low-frequency data (e.g., monthly inflation) are explained through high-frequency data (e.g., daily oil price movements). Merely estimating many coefficients for the high-frequency lags usually yields substantial identification problems. The existing literature has addressed this problem by restricting the coefficients on high-frequency lags to follow a specific functional form that can be described by few parameters. However, [Breitung and Roling \(2015\)](#) argue that this might be overly restrictive and suggests a more flexible nonparametric approach. Instead of enforcing a specific functional form, [Breitung and Roling](#) augment an objective function (such as the likelihood function) by a term that penalizes second differences between coefficients for lags of adjacent periods.

We employ the same strategy to enforce smooth behavior of the impact of ratings. Because our ratings dummies are defined cumulatively, we restrict not second but first differences between coefficient estimates. This is equivalent to minimizing second differences between mutually exclusive rating dummies, where the dummy equals 1 if and only if a country has the corresponding rating.

That is, instead of using the traditional likelihood function for our models (denoted by LL_{model}), we augment the likelihood function as follows:

$$LL_{smooth} = \sum_{c=8}^{24} \ln(\phi(\sqrt{\lambda}(\zeta_c - \zeta_{c-1}), 0, \sigma_{model}^2)) \quad (4)$$

$$LL = LL_{model} + LL_{smooth},$$

where, depending on the model, ζ can be either β or ψ , and $\phi(\zeta, 0, \sigma_{model}^2)$ is the density of a normal distribution with mean zero and variance drawn from the errors in the respective model (σ_{ε}^2 or σ_{η}^2), measured at point ζ .

By increasing the weight λ of the penalty LL_{smooth} , it is possible to enforce the smooth behavior of adjacent coefficients.¹⁴ In the limit, when the penalty weight goes to infinity, the coefficients are forced to be identical, i.e., we would only estimate a single coefficient. If the weight of the penalty goes to zero, the results approach those of the unrestricted model, i.e., we would estimate 18 coefficients. [Breitung and Roling](#) accordingly show that

¹⁴The likelihood contains $\sqrt{\lambda}$, as equation (2) can be estimated by simple OLS with a quadratic penalty term ([Breitung and Roling, 2015](#)).

λ can generally be mapped onto the effective loss of degrees of freedom. That is, we can use standard information criteria to select the degree of smoothing. On the one hand, if high differences between the ratings are actually needed to explain the behavior of interest rates or ratings, we will choose a low degree of smoothing in the respective equation. On the other hand, if volatility in those coefficients is merely driven by very few observations in specific classes, we will opt for smoothness.

In contrast to [El-Shagi and von Schweinitz \(2015\)](#), we estimate a single λ for both equations. This allows a more clear-cut comparison of the model, where the long-run relation is restricted to be identical in both equations, which implicitly includes a single rather than individual λ s. That is, in our setup, a rejection of a single long-run relationship cannot be attributed to the influence of using different λ s.

3.4 Addressing heteroscedasticity

As mentioned in subsection 2.1, government bond yields are affected by more than one risk channel. Especially during times of turmoil in financial markets, market participants may shift their portfolios toward government bonds that they deem safe ([Vayanos, 2004](#); [Beber, Brandt, and Kavajecz, 2009](#)). This safe-haven effect (which is essentially herding behavior) may lead to self-fulfilling crises in other countries when creditors with a risk that is slightly worse than “safe” (but far from being in default) are shunned by financial markets ([De Grauwe and Ji, 2013](#)). Herding behavior also reinforces a tendency to treat superficially comparable countries similarly without performing in-depth analyses of the individual debtors. This, in turn, may lead to spillovers of risk from one country to the next ([Beirne and Fratzscher, 2013](#)), that is, shocks to yields may be correlated across countries.

Shocks to the interest rate especially deviate from the usual i.i.d. assumptions in our estimation. Indeed, we find substantial heteroscedasticity, cross-sectional correlation and heavy tails (as could, for example, be inferred from the results of [Arezki et al., 2011](#)). We use a bootstrap method to account for heteroscedasticity (both over time and across countries) and cross-sectional correlation. It is based on the wild bootstrap originally proposed by [Wu \(1986\)](#). In the wild bootstrap, rather than resampling the original residuals, the simulations are generated using error terms that are obtained by multiplying the original residual ε for the respective observation with a random multiplier v . The distribution of v ensures that the expected value of εv is 0 and that the first few moments of the distribution of ε and εv are identical or at least very similar. In this paper, the random multiplier v is drawn from a 6-point distribution proposed by [Webb \(2013\)](#), which has been shown to have even more desirable properties than the traditionally used distributions suggested by [Mammen \(1993\)](#) and [Davidson and Flachaire \(2008\)](#). To reproduce the cross-country correlations found in the original sample, we use the same multiplier for all countries at a given point in time, i.e., $v_{i,t} = v_{j,t}$ for all pairs of i and j . This follows an approach suggested by [Davidson and MacKinnon \(2010\)](#), who used the same technique to reproduce the correlations among the residuals of several equations. For the lack of a better alternative, the probit equation is simulated using i.i.d. errors drawn from a standard normal distribution. The starting values for the ratings and yields as well as the lagged differences are drawn from the empirical joint distribution. Because we need low ratings in the simulated sample to have identification of all coefficients, we

combine the bootstrap with an acceptance-rejection algorithm that discards simulations where extremely low ratings do not occur. The reported confidence bounds are based on 1,000 accepted simulations of the entire two equation system.

3.5 Assessing the long-run relation

Identifying the long-run relation: Including asymmetries in the model slightly changes how we have to treat the long-run relation implied by each equation. Usually, we simply set the first differences of all variables to zero on both sides of the equation, solving for the corresponding relationship of the level variables.¹⁵ Of course, this approach relies on the idea that the expected value of the impact of a right-hand side variable is zero, if the expected value of said variable is zero. However, this logic no longer holds in the case of asymmetric effects. Indeed, the impact of the lagged first differences of interest rates and the lagged rating changes is no longer zero, on average, even in the event that we are in equilibrium and those deviations from zero are purely random. To partially compensate for the asymmetric effects of positive and negative deviations, we use typical past changes in the interest rate. That is, rather than setting the past lagged changes to zero, we replace all lagged changes by the standard deviation of the changes observed in the data. Because rating changes are extremely rare events, we retain a value of lagged rating changes of 0.

Testing for the number of long-run relationships: To assess whether the model allowing for individual long-run relationships in both equations outperforms the more traditional cointegration approach with identical long-run relationships in both equations, we use a standard likelihood ratio test. First, we estimate a restricted cointegration model where the long-run coefficients in both equations are equal ($\beta = \psi$). Second, we compare the restricted and the unrestricted models using model-specific optimal smoothing parameters λ . Third, to guarantee comparability, the likelihood ratio test is performed twice more for pairwise identical λ . That is, we compare both models using the optimum λ selected for the cointegration model, and we compare both models using the λ selected for the model with individual long-run relationships. The model comparison is based on the ML results for computational reasons.¹⁶ This decision comes with a caveat: Because the residuals are not normally distributed – which is the very reason for our bootstrapping procedure – the test results have to be interpreted with caution. However, the three tests all return similar results. In all cases, the restricted cointegration model is rejected at the 1% level.

¹⁵As outlined in [El-Shagi and von Schweinitz \(2015\)](#), this corresponds to setting the left-hand side of the ordered probit model equal to the mean of the thresholds, implying equal (and low) probabilities of up- and downgrades, i.e., an expected value of no adjustment at all.

¹⁶The clearly outperformed cointegration model takes an extremely long time to bootstrap. Because evidence from the model with multiple long-run relations indicates that the coefficients change only slightly, we carry over the ML results to avoid conducting a full-fledged bootstrap of the cointegration system.

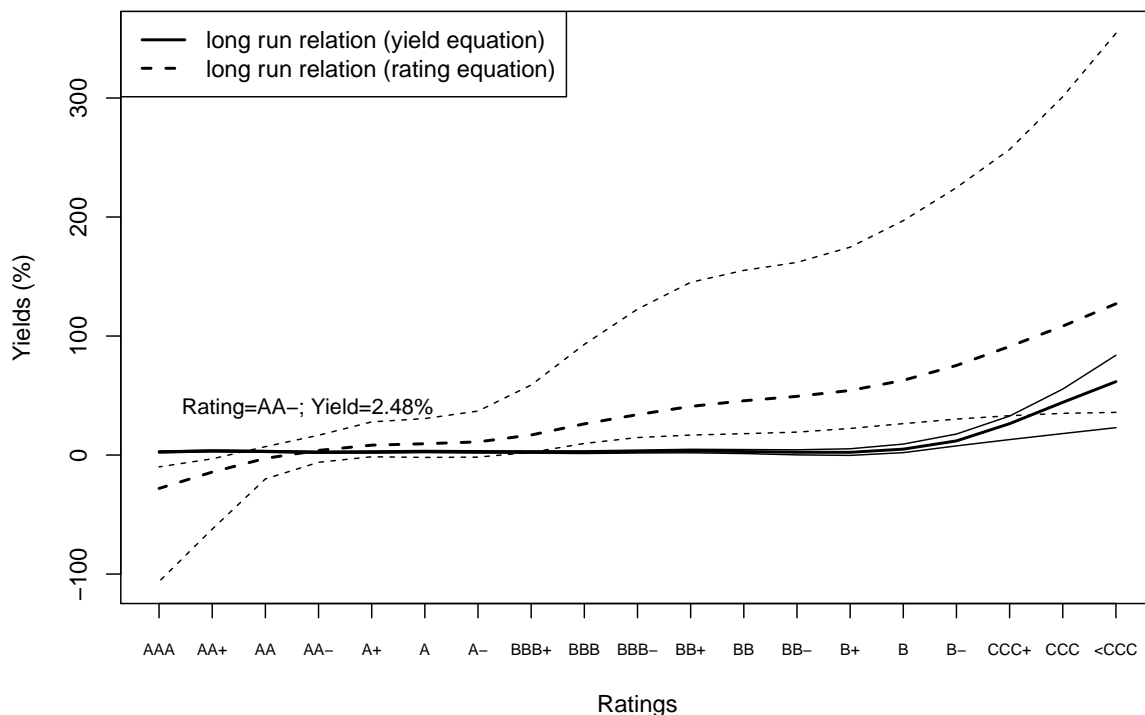


Figure 7: long-run relation of ratings and yields, accounting for asymmetric shocks
Note: In addition to median long-run relationship, we show the 10%- and 90% confidence bands from the bootstrap.

4 Results

4.1 The long-run relation of ratings and yields

Our preferred model – allowing for different long-run relations in both equations – strongly outperforms the basic cointegration approach. There is further evidence for the superiority of our benchmark model. Figure 7 plots the long-run relationships implied by our two equations, including their confidence bounds. It is evident that the long-run relations are not merely different for most rating levels, they are economically and statistically significantly different.

As theory predicts, both curves are significantly negatively sloped, implying that (a) ratings tend to deteriorate if the interest rate is high (dashed line, from equation (3)) and that (b) the risk premium (measured through the interest rate) increases for countries with low ratings (solid line, from equation (2)). However, as indicated by the stylized facts presented in Section 2, we only find a risk premium of noteworthy magnitude at extremely low ratings. Only far below the investment grade threshold of B does the equilibrium interest rate, as indicated by equation (2), start to increase, quickly reaching extreme heights.

This interest rate behavior is one of the main reasons why our point estimates indicate

a single intersection of the long-run relations, which occurs at the highest ratings and at very low interest rates. When the rating is below AA-, the interest rate implied by the current rating is so low that it implies pressure for a rating upgrade. This is particularly true for low ratings, which are thus highly transitory.

The confidence bands of the rating equation grow extremely wide for low ratings. Yet, the impression of a possible second equilibrium is partly driven by the reporting of independent confidence bounds for both equations. When assessing the existence of a second equilibrium for each bootstrapped set of long-run relationships individually, we see that we can still significantly reject the existence of a bad equilibrium. The high significance of rejection in the presence of only a limited number of rating changes and – correspondingly – wide confidence bands gives further credence to our results. For 98% of the bootstrapped sets of long-run relations, we find that the equilibrium interest rate (according to the yield curve) creates upward pressure in the rating equation at all ratings below A-. Even in the unlikely case that a second equilibrium exists, that second equilibrium is unstable in the sense that there is no process driving the rating toward it. Essentially, a second intersection of the rating and yield curves below a rating of CCC would not imply that a country can be driven into default by a vicious cycle of ratings and yields.¹⁷ This merely indicates that a country requires a positive impulse to escape default ratings. Because negative rating shocks are impossible at a D rating, a positive shock initiating recovery would be inevitable.

As a robustness test, we also estimate the model that includes country-specific effects. Because this implies different intercepts in the long-run relationships, each country now has a separate equilibrium. For all countries except Pakistan, this equilibrium corresponds to a yield of less than 7 percent, and just as in the baseline estimation, there is no evidence of a second intersection for any country. Regarding the dynamics, as outlined in the following subsection, the results are even more similar; thus, they are not discussed separately.

4.2 Short-run dynamics and impulse response functions

Like the analysis of the long-term relationship, which found no gravitational pull toward default, the short-run dynamics do not provide evidence of a vicious cycle. While there is some autocorrelation in changes of ratings (especially for downgrades), it is far too little to imply the economically meaningful self-reinforcement of a negative shock. Only when the initial ratings are very close to default, the interest rate volatility generated by a negative rating causes an economically meaningful increase in the probability of further downgrades. Yet, this does not completely eliminate the role of rating shocks in creating macroeconomic distress. If a rating downgrade drives the rating below B, the impact is considerable. Due to the steep increase of the risk premium below this rating level, the interest rate quickly rises to extreme levels and remains high over an extended period due to the high persistence of ratings.

To illustrate this problem, we provide impulse response functions (IRF) based on the bootstrapped coefficients after a relatively large rating shock, a downgrade of 2 notches.

¹⁷This finding is robust with regard to the number of bootstrap draws. Tests with 5,000 simulations showed virtually no difference. Due to the high computational burden of more simulations, not all settings could be estimated with 5,000 simulations.

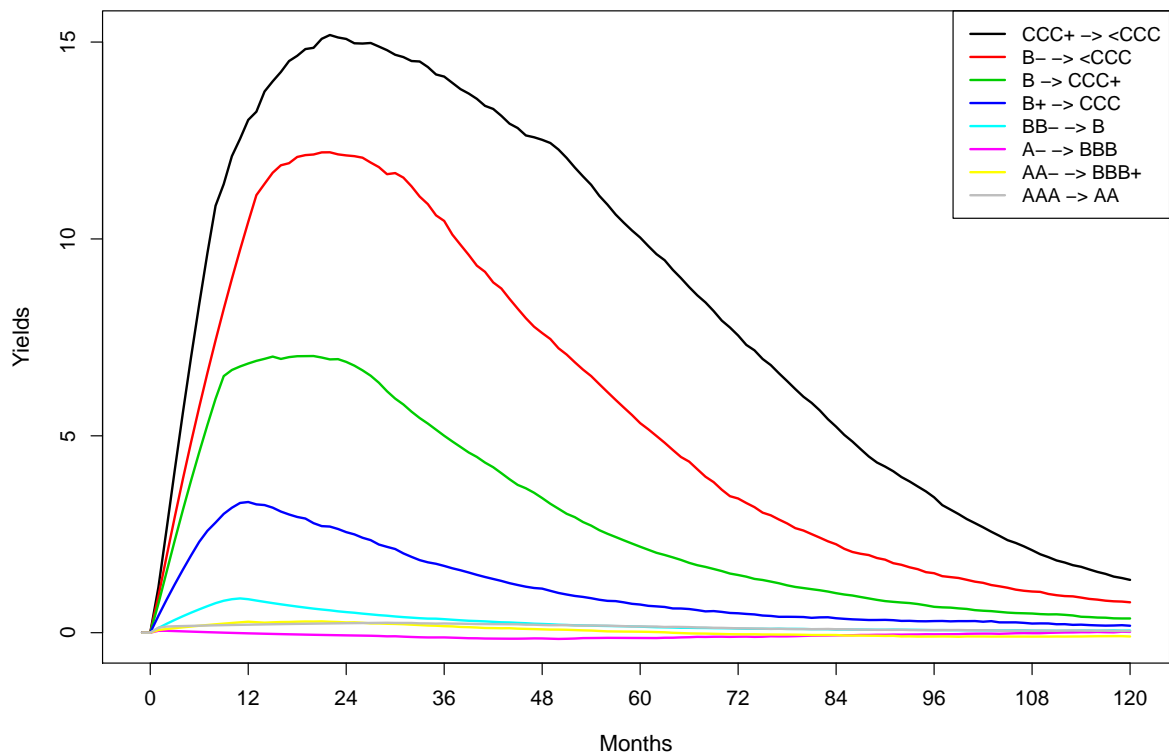


Figure 8: Median impulse-response function of yields after a rating shock of two notches, selected ratings

Note: Impulse response functions are significant at the 5% level for approximately 60 periods for all rating shocks occurring from an initial rating of BB- or below. For higher ratings, the entire IRF becomes insignificant. Selection of impulse response functions including confidence bands in Figure A1 in the appendix.

The standard method of calculating an IRF is to simulate the reaction after a shock leading away from equilibrium. In our context, however, this is inadequate. Recovery from default ratings to the long-run equilibrium takes more than eighty years in the median (see Table 3). Because the nonlinearity in the risk premium causes shocks to have markedly different effects based on the original rating, it is essential to consider shocks to countries with different risk assessments. Thus, rather than reporting a single impulse response function starting at the long-run equilibrium, we report a range of IRFs for a negative 2-notch rating shock. Each IRF uses a different original rating level and the corresponding yield as given by the equilibrium yield curve. We report the difference to the recovery path without the initial rating shock. The resulting IRFs of yields are presented in Figure 8, while the IRFs of ratings are given in Figure 9.¹⁸

Table 3: Median years until first occurrence of a rating after a D rating

Rating	D	CCC	CCC+B-	B	B+	BB-	BB	BB+BBB-	BBB	BBB+A-	A	A+	AA-	AA	AA	AAA			
Years	0	2	4	7	11	14	17	19	22	26	30	34	42	53	66	81	105	148	231

Note: The number of years until a rating is reached for the first time is calculated as the median of 5,000 simulations beginning from a rating of 6 and a yield of 60% (as indicated by the yield equation). This estimation thus includes the (in this case beneficial) positive autocorrelation of rating changes. While the table includes all ratings, only the movement back to a level of AA is a recovery in the sense of a return to equilibrium. Further improvement in ratings is driven purely by chance, which explains the apparent break in the recovery times.

The impact of the rating shock on the interest rate is only significant if the rating shock drives the rating below a rating of B; see Figure A1 in the appendix. That is, only when the rating of a country is already (far) below the investment grade threshold, further downgrades have a significant increasing effect on yields. Yet, as soon as this happens, the effects are considerable. A rating shock driving the rating from B+ to B- causes a significant increase in the interest rate for more than six years, peaking at approximately 3% after one year. The impact of a shock driving the rating from B- to CCC remains significant for more than a decade, peaking at more than 12% after two years. For nearly five years, the interest rate is increased by 5% or more compared to the benchmark recovery path.

The duration of the interest premium suggests that there will be a substantial increase in the interest burden, even given moderate debt rollover during this time. For countries with a low average maturity – which might be the most likely to be hit by this type of

¹⁸While impulse response functions are usually computed deterministically, i.e., without further shocks, this is unfeasible in our case due to the (ordered probit) rating equation. At each point in time, the most likely outcome is no rating adjustment. Yet, over time, the cumulative probability of a rating adjustment is increasing. Thus, the impulse responses are computed as the median of a range of simulations where both ratings and yields are subject to disturbances. By taking the median over many observations, we remove the idiosyncratic effect of shocks. Rather than computing the difference between the median development with the initial rating shock and the median development without said shock, we always conduct the simulations with and without the 2-notch shock with identical simulated disturbances. For each pair, we compute the difference individually and then report the median of those differences. For the computation of confidence bounds, which should include parameter uncertainty but not the idiosyncrasies of further disturbances, we repeat the aforementioned procedure for every bootstrapped set of coefficients. That is, our confidence bounds are quantiles of those median IRFs for different possible coefficients.

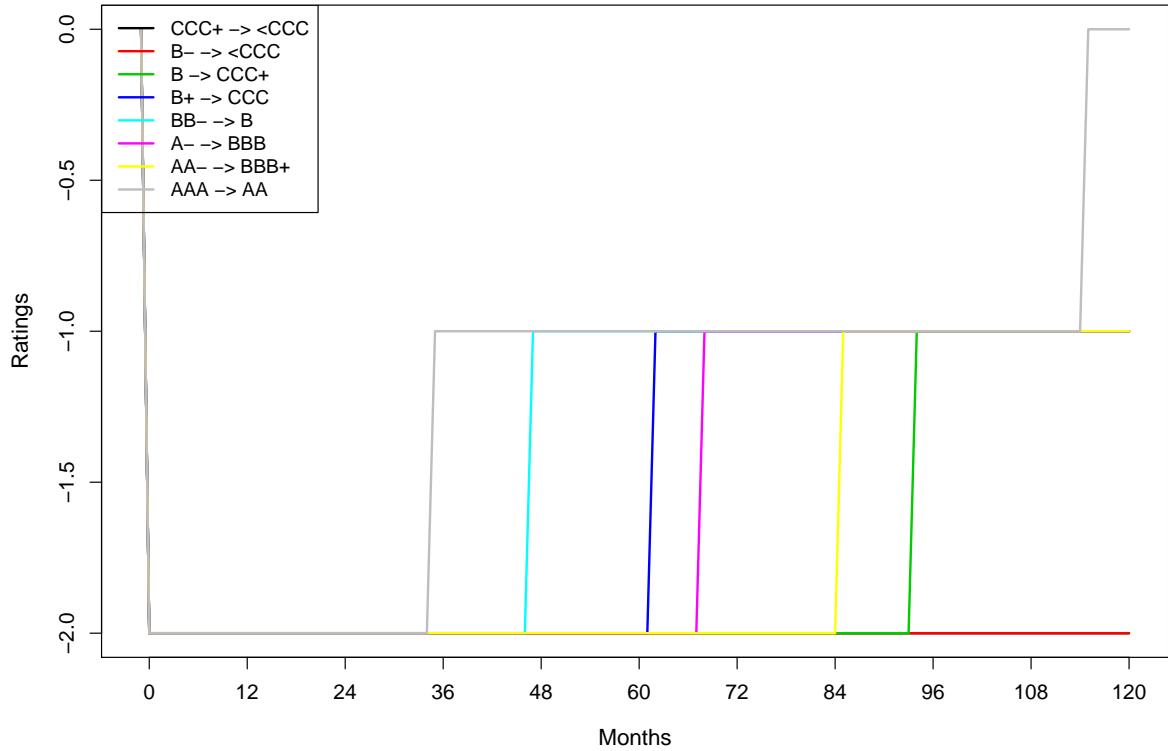


Figure 9: Median impulse-response function of ratings after a rating shock of two notches, selected ratings

Note: Impulse response functions are significant at the 5% level for approximately 40 periods for the lowest initial rating. For higher starting ratings, more periods are significant, and from BBB+ onwards, ratings remain significantly below the benchmark for the entire simulated period. Selection of impulse response functions including confidence bands in Figure A1 in the appendix.

shock (Arellano and Ramanarayanan, 2012) – the cost is even higher. This will strain the government budget, thereby limiting public good provision substantially. At the same time, the sovereign ceiling function of the sovereign rating will cause similar rating changes in the private sector, most likely corresponding to similar interest premia, which hamper investment. Given the high path dependence of economic development, the consequences of such periods might be noticeable long after the interest rate has normalized.

Concerning the dynamics of ratings, our results are mixed. On the one hand, the already discussed persistence of ratings makes it nearly impossible to catch up to the equilibrium recovery path after a substantial shock. Even after eight years, ratings remain significantly below the original undisturbed recovery path for all starting ratings. This is partly due to the effect of past yield changes on rating changes. Both increasing and decreasing yields make rating downgrades more likely; see table A3 in the appendix. This seemingly contradictory result may be explained by the inability of rating agencies to differentiate between fundamentally justified yield movements and increased market volatility. As the second is a sign of financial distress, which may by itself negatively affect the sustain-

ability of government debt, the probability of a downgrade increases (slightly). However, ratings do not deteriorate further after an initial shock. Thus, we find no evidence of self-reinforcing rating dynamics.

The response to a corresponding upgrade is fairly similar (although in the opposite direction), as shown in Figure A2 in the appendix. While the dynamics are slightly different due to the strongly asymmetric impact of interest rate changes on ratings, the order of magnitude and the shape are comparable.

As a robustness test of the results provided above, we test an alternative downgrade scenario, wherein we distribute the two-notch shock of the previous scenario over two consecutive months. As our rating equation employs two lags, this scenario tests the potential self-reinforcing behavior of downgrades. Without taking yield changes and level effects into account, two successive downgrades would increase the probability of a third downgrade from 1% to 44%. Therefore, these results should be interpreted as a worst-case response to a large rating shock. Both IRFs (yields and ratings) show larger differences to the benchmark adjustment cases, as shown in Figure A3 in the appendix. We can confirm that even transitory default after a negative rating shock is highly unlikely if the initial rating is not already close to default. Even for a starting rating of B, the probability of a default (as a transitory state) is only 2%, and it is almost zero for better ratings.

The results are also robust with regard to the case where different rating agencies have strongly differing opinions. Taking the median of the different rating decisions instead of its mean does not affect our results, as can be seen in Figures A4 and A5 in the appendix.

4.3 Scenario analysis

To provide more conclusive evidence for the claim that the downward spirals of countries such as Greece were not caused by their initial rating downgrade, this section presents scenario analyses in which the impulse response developments are initialized with data from some episodes of major financial distress and compared to the observed development. The scenarios (i.e., the shock and the subsequent development) are depicted in Figures 10 and 11 along with the impulse response functions and the two sets of confidence bands. In addition to the confidence bounds reported for the IRFs, which only include parameter uncertainty, we provide confidence bounds including uncertainty concerning future shocks (not unlike the confidence bounds used in forecasting). This gives a more reasonable benchmark to assess the probability of an observed development, which does, of course, include an entire sequence of disturbances rather than merely the initial event. We report the results for two events: the downgrade of Italy after the EMS crisis in the early 1990s and the first sequence of downgrades of Greece during the ongoing debt crisis.

Italy in 1992: First, we examine Italy. As one of the founding members of the European exchange rate mechanism (ERM), Italy enjoyed a AAA rating until July 1991 when Moody's first downgraded it by one notch to AA+ (Aa1 in Moody's notation), making Italy the only G7 country with a rating below AAA at that time. A second downgrade, this time by two notches to AA-, followed on August 13, 1992. The reasons for these downgrades were the exceptionally high debt levels the Italian government had amassed in previous years (exceeding 100% of GDP in 1992) together with large external imbalances. Harsh austerity measures, privatization and laws aimed at reducing labor costs by the

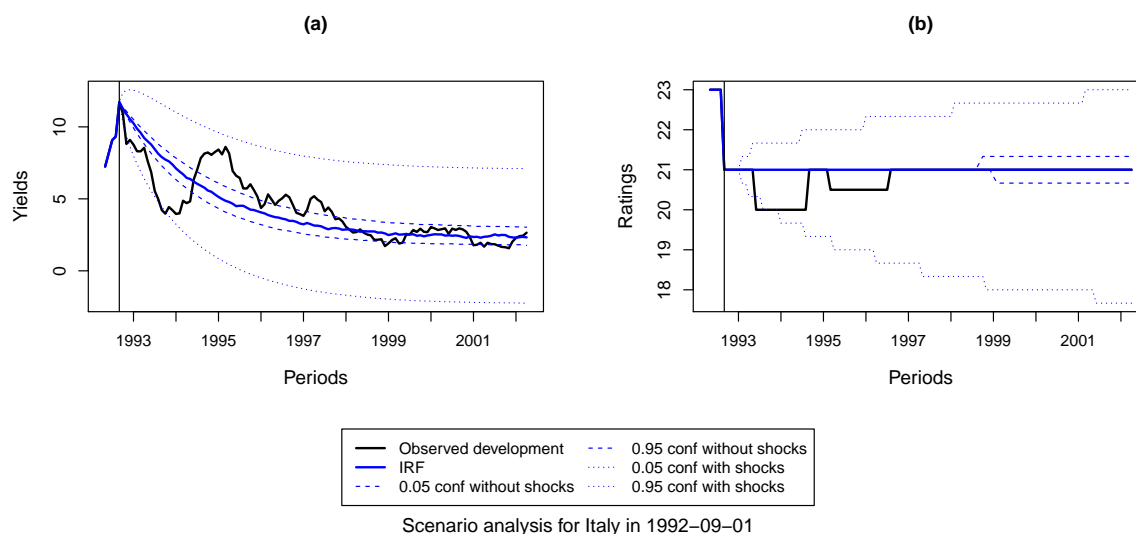


Figure 10: Observed development and impulse reaction functions in Italy after a downgrade in August 1992

newly elected Italian government were considered positive by Moody's. However, the rating agency held that these measures had come too late. Therefore, the agency predicted that government debt would grow further during the 1990s (as it in fact did). Another fact that was thought to aggravate debt problems was the overvalued Italian currency, which was restricted by the rules of the ERM. Again, the rating agencies' assessment was largely sound. Italy left the ERM one month later in September 1992 and devalued strongly (see also [El-Shagi, Lindner, and von Schweinitz, 2016](#)).

It seems that the decision to downgrade the rating was largely a consequence of previous developments. Judging the reforms as coming too late, Moody's itself acknowledged that there had been enough to announce an earlier downgrade. Similarly, the European community had previously voiced its concerns, citing the Maastricht criteria (among them, a maximum government debt level of 60% of GDP) that were introduced in February 1992. Even the Italian minister of the treasury was not excessively concerned by the downgrade. Therefore, it is not surprising that the reaction of the markets was normal; see Figure 10. The observed development of yields is quite close to the development of the simulated IRF, with all differences (except for the first month) being within the wider set of confidence bands. The observed ratings are identical to the simulated series for most of the time. Overall, this is a typical case in which the discrete rating change reflected an assessment of the sustainability of government debt that was shared by the markets, which holds for the large majority of rating changes.

Greece in 2010: A completely different case is observed Greece in 2010. Membership in the Euro area had led to lower interest rates and improved ratings until October 20th, 2009, when the newly elected government opened its books and announced that the previous government had provided false low estimates of the expected government

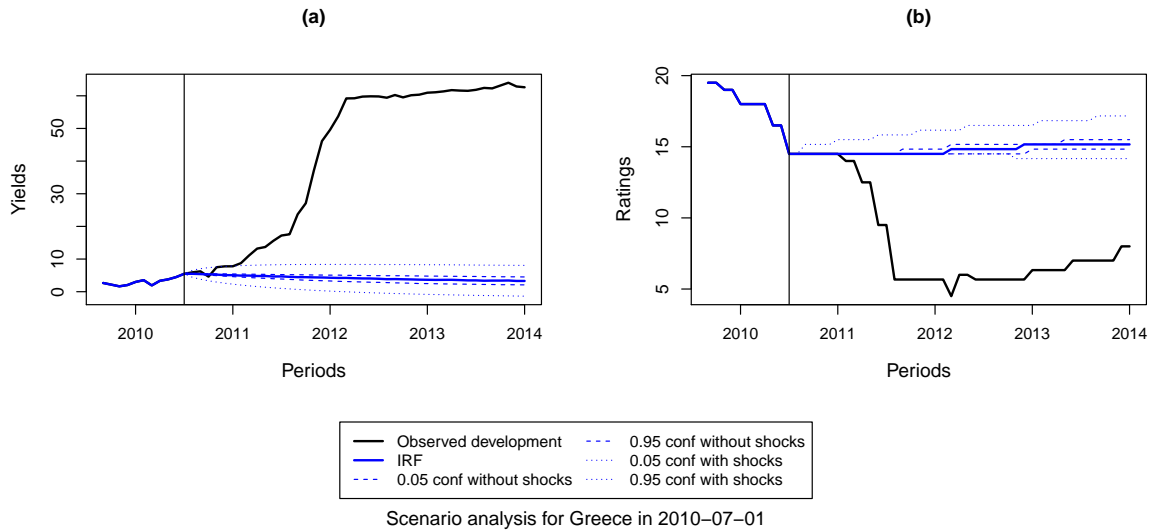


Figure 11: Observed development and impulse reaction functions in Greece after a series of downgrades until June 2010

deficit. Only two days later, Fitch downgraded Greek sovereign bonds to A-. While yields increased and other members of the Euro area (together with the IMF, assisted by the ECB) issued a first rescue package in March 2010 and introduced the European Financial Stability Facility, ratings deteriorated further. The starting point of our simulation is June 2010, when rating agencies downgraded Greece to the investment grade threshold (Fitch just above, Moody's just below). Approximately until the end of 2010, the observed ratings and yields are in line with our simulated results. However, at that time, both recession and strong political opposition to reform led to renewed doubts regarding the ability and willingness of Greece to repay its debt. This in turn led to strongly increasing yields and further downgrades and, when private investors became part of a second rescue package in July 2011, a final downgrade to CCC- (in default with little prospect for recovery). Taken together, it should be clear that not ratings but political developments likely fueled the collapse toward default in 2012. That is, this development cannot be attributed to the initial (weak) downgrades of the rating agencies.

5 Conclusion

In conclusion, our evidence reconciles the two prominent conflicting views featured in the existing literature, which partly stresses the dangers of low ratings and partly denies the importance of ratings.

On the one hand, we find rather strong evidence against the theory of a vicious cycle. This is true for both the strong form of this theory that speculates the existence of a second – bad – equilibrium that might emerge after a rating is driven below moderate risk levels and the weaker form of the theory that focuses on self reinforcing short-run dynamics. Neither cycle is found in our data, at least not at a meaningful level. The usual

interaction between ratings and yields fails to explain the downward spirals observed in a few cases. The vicious cycle theory seems driven by misinterpreting a few individual observations as typical.

On the other hand, there can be substantial costs to rating downgrades. If a rating shock drives a country below a B rating, the risk premium can virtually explode. While the impact of ratings is negligible for better ratings, the increase is considerable for countries that begin at problematic levels before the downgrade and can easily reach a 2-digit magnitude. Due to the persistence of ratings, it is more than likely that the country will pay this cost for an extended period.

References

- Afonso, A., D. Furceri, and P. Gomes (2012). Sovereign Credit Ratings and Financial Markets Linkages: Application to European Data. *Journal of International Money and Finance* 31(3), 606–638.
- Arellano, C. and A. Ramanarayanan (2012). Default and the Maturity Structure in Sovereign Bonds. *Journal of Political Economy* 120(2), 187–232.
- 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 Papers 11/68.
- Baum, C. F., M. Karpava, D. Schäfer, and A. Stephan (2014). Credit Rating Agency Downgrades and the Eurozone Sovereign Debt Crises. NBP Working Paper 177.
- Beber, A., M. W. Brandt, and K. A. Kavajecz (2009). Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market. *Review of Financial Studies* 22(3), 925–957.
- Beirne, J. and M. Fratzscher (2013). The Pricing of Sovereign Risk and Contagion During the European Sovereign Debt Crisis. *Journal of International Money and Finance* 34, 60–82.
- Bolton, P., X. Freixas, and J. Shapiro (2012). The Credit Ratings Game. *The Journal of Finance* 67(1), 85–111.
- Breitung, J. and C. Roling (2015). Forecasting Inflation Rates Using Daily Data: A Nonparametric Midas Approach. *Journal of Forecasting* 34(7), 588–603.
- Cantor, R. and F. Packer (1996). Determinants and Impact of Sovereign Credit Ratings. *The Journal of Fixed Income* 6(3), 76–91.
- Davidson, R. and E. Flachaire (2008). The Wild Bootstrap, Tamed At Last. *Journal of Econometrics* 146(1), 162–169.
- Davidson, R. and J. G. MacKinnon (2010). Wild Bootstrap Tests for IV Regression. *Journal of Business & Economic Statistics* 28(1).

- De Grauwe, P. and Y. Ji (2013). Self-fulfilling Crises in the Eurozone: An Empirical Test. *Journal of International Money and Finance* 34, 15–36.
- 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. ECB Working Paper 1419.
- De Vries, T. and J. De Haan (2014). Credit Ratings and Bond Spreads of the GIIPS. DNB Working Paper 432.
- Durbin, E. and D. Ng (2005). The Sovereign Ceiling and Emerging Market Corporate Bond Spreads. *Journal of International Money and Finance* 24(4), 631–649.
- El-Shagi, M. (2010). The Role of Rating Agencies in Financial Crises: Event Studies from the Asian Flu. *Cambridge Journal of Economics* 34(4), 671–685.
- El-Shagi, M., A. Lindner, and G. von Schweinitz (2016). Real Effective Exchange Rate Misalignment in the Euro Area: A Counterfactual Analysis. *Review of International Economics* 24(1), 37–66.
- El-Shagi, M. and G. von Schweinitz (2015). Risk and Return: Is there an Unholy Cycle of Ratings and Yields? *Economics Letters* 129, 49–51.
- Ferri, G., L.-G. Liu, and J. E. Stiglitz (1999). The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis. *Economic Notes* 28(3), 335–355.
- Gärtner, M. and B. Griesbach (2012). Rating Agencies, Self-fulfilling Prophecy and Multiple Equilibria?: An Empirical Model of the European Sovereign Debt Crisis 2009-2011. mimeo, University of St. Gallen.
- Gärtner, M., B. Griesbach, and F. Jung (2011). PIGS or Lambs? The European Sovereign Debt Crisis and the Role of Rating Agencies. *International Advances in Economic Research* 17(3), 288–299.
- Hill, P. and R. Faff (2010). The Market Impact of Relative Agency Activity in the Sovereign Ratings Market. *Journal of Business Finance & Accounting* 37(9-10), 1309–1347.
- IMF (1997, May). World Economic Outlook May 1997 – Globalization: Opportunity and Challenges.
- Kiff, J., S. B. Nowak, and L. Schumacher (2012). Are Rating Agencies Powerful? An Investigation Into the Impact and Accuracy of Sovereign Ratings. IMF Working Paper 12/23.
- Larraín, G., H. Reisen, and J. von Maltzan (1997). Emerging Market Risk and Sovereign Credit Ratings. OECD Development Centre Working Paper 124.
- Löffler, G. (2005). Avoiding the rating bounce: Why rating agencies are slow to react to new information. *Journal of Economic Behavior & Organization* 56(3), 365–381.

- Mammen, E. (1993). Bootstrap and Wild Bootstrap for High Dimensional Linear Models. *The Annals of Statistics*, 255–285.
- Mellios, C. and E. Paget-Blanc (2006). Which Factors Determine Sovereign Credit Ratings? *The European Journal of Finance* 12(4), 361–377.
- Mora, N. (2006). Sovereign Credit Ratings: Guilty Beyond Reasonable Doubt? *Journal of Banking & Finance* 30(7), 2041–2062.
- Pesaran, M. H., Y. Shin, and R. J. Smith (2001). Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics* 16(3), 289–326.
- Reisen, H. and J. von Maltzan (1999). Boom and Bust and Sovereign Ratings. *International Finance* 2(2), 273–293.
- Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica* 48(1), 1–48.
- Vayanos, D. (2004). Flight to Quality, Flight to Liquidity, and the Pricing of Risk. NBER Working Paper 10327.
- von Hagen, J., L. Schuknecht, and G. Wolswijk (2011). Government Bond Risk Premiums in the EU Revisited: The Impact of the Financial Crisis. *European Journal of Political Economy* 27(1), 36–43.
- Webb, M. D. (2013). Reworking Wild Bootstrap Based Inference for Clustered Errors. Working Papers 1315, Queen’s University, Department of Economics.
- White, L. J. (2010). Markets: The Credit Rating Agencies. *The Journal of Economic Perspectives* 24(2), 211–226.
- Wu, C.-F. J. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. *The Annals of Statistics* 14(4), 1261–1295.

Appendix

Table A1: List of countries and data availability

Country	IMF classification	Data availability
Argentina	Developing	2009-10-01 – 2013-12-01
Australia	Advanced	1980-01-01 – 2014-01-01
Austria	Advanced	2003-02-01 – 2014-01-01
Belgium	Advanced	2005-01-01 – 2014-01-01
Brazil	Developing	2003-02-01 – 2014-01-01
Canada	Advanced	1986-05-01 – 2014-01-01
Chile	Developing	2009-01-01 – 2014-01-01
China	Developing	2002-06-01 – 2014-01-01
Colombia	Developing	2002-09-01 – 2014-01-01
Denmark	Advanced	1986-09-01 – 2014-01-01
Finland	Advanced	1992-01-01 – 2014-01-01
France	Advanced	1989-01-01 – 2014-01-01
Germany	Advanced	1986-01-01 – 2014-01-01
Greece	Advanced	1997-06-01 – 2014-01-01
Hong.Kong	Advanced	1994-09-01 – 2014-01-01
Hungary	Transition	1997-02-01 – 2014-01-01
Iceland	Advanced	2003-08-01 – 2014-01-01
India	Developing	1993-10-01 – 2014-01-01
Indonesia	Developing	2003-05-01 – 2014-01-01
Ireland	Advanced	1992-01-01 – 2014-01-01
Israel	Advanced	2002-04-01 – 2014-01-01
Italy	Advanced	1988-11-01 – 2014-01-01
Japan	Advanced	1985-12-01 – 2014-01-01
Malaysia	Developing	2001-10-01 – 2014-01-01
Mexico	Developing	2001-08-01 – 2014-01-01
Netherlands	Advanced	1994-02-01 – 2014-01-01
New.Zealand	Advanced	1994-04-01 – 2014-01-01
Norway	Advanced	1992-11-01 – 2014-01-01
Pakistan	Developing	2009-10-01 – 2014-01-01
Peru	Developing	2009-10-01 – 2014-01-01
Philippines	Developing	2001-02-01 – 2014-01-01
Poland	Transition	1999-03-01 – 2014-01-01
Portugal	Advanced	1994-12-01 – 2014-01-01
Singapore	Advanced	1990-06-01 – 2014-01-01
Slovakia	Transition	2007-04-01 – 2014-01-01
Slovenia	Transition	2007-04-01 – 2014-01-01
South.Korea	Advanced	1997-12-01 – 2014-01-01
Spain	Advanced	1988-06-01 – 2014-01-01
Sri.Lanka	Developing	2006-10-01 – 2014-01-01
Sweden	Advanced	1985-01-01 – 2014-01-01
Switzerland	Advanced	1994-01-01 – 2014-01-01
Taiwan	Advanced	1999-04-01 – 2014-01-01
Thailand	Developing	1999-09-01 – 2014-01-01
Turkey	Developing	2005-08-01 – 2014-01-01
United.Kingdom	Advanced	1989-01-01 – 2014-01-01
United.States	Advanced	1994-09-01 – 2014-01-01

Table A2: Summary statistics of ratings and yields

Country	Real Yields				Average Ratings			
	meany	sdy	miny	maxy	meanr	sdr	minr	maxr
Argentina	2.58	4.32	-6.18	7.88	10.86	1.46	7.67	12.50
Australia	3.99	2.29	-1.21	10.62	23.05	0.87	22.00	24.00
Austria	0.76	1.39	-2.19	3.43	23.94	0.13	23.67	24.00
Belgium	0.66	1.46	-1.59	4.46	22.64	0.54	21.67	23.00
Brazil	3.68	4.25	-10.34	9.35	13.67	2.04	10.00	16.00
Canada	3.22	2.23	-1.73	8.81	23.47	0.80	22.00	24.00
Chile	3.31	1.87	-0.95	8.43	20.15	0.41	19.33	20.67
China	0.42	1.90	-4.81	4.46	19.52	0.96	18.00	20.67
Colombia	4.63	1.58	2.38	11.73	13.79	0.83	13.00	15.67
Denmark	3.17	2.67	-2.42	8.65	23.47	0.46	23.00	24.00
Finland	2.83	2.81	-2.35	10.27	23.44	0.91	21.50	24.00
France	2.95	2.09	-1.29	7.41	23.95	0.20	22.67	24.00
Germany	2.76	2.04	-1.76	8.64	24.00	0.00	24.00	24.00
Greece	10.09	19.42	-1.28	64.00	16.49	4.68	4.50	20.00
Hong Kong	2.30	4.50	-6.49	13.00	20.62	1.59	19.00	23.33
Hungary	2.07	2.26	-3.10	9.78	16.93	1.77	13.67	19.00
Iceland	1.45	3.33	-8.63	8.30	18.70	3.63	14.67	22.50
India	1.45	3.56	-8.90	10.86	14.34	0.83	13.00	15.00
Indonesia	2.30	2.80	-5.43	7.61	12.45	1.58	9.50	14.67
Ireland	2.73	3.07	-2.20	12.10	22.16	2.65	16.00	24.00
Israel	2.79	2.52	-2.43	8.90	19.03	0.54	18.50	19.67
Italy	3.46	2.53	0.09	11.72	21.26	1.64	16.33	24.00
Japan	1.74	1.48	-1.41	5.66	23.15	0.98	20.67	24.00
Malaysia	1.31	1.63	-4.45	6.15	17.69	0.59	16.00	18.00
Mexico	3.22	1.52	-0.31	6.75	16.07	0.77	14.50	17.00
Netherlands	1.58	1.58	-2.39	4.88	24.00	0.03	23.67	24.00
New Zealand	3.69	1.72	-1.28	7.17	22.86	0.63	22.00	23.50
Norway	2.76	1.92	-1.38	7.27	23.90	0.18	23.50	24.00
Pakistan	2.11	1.72	-1.57	5.47	8.83	0.24	8.50	9.00
Peru	1.84	1.46	-0.13	6.32	15.48	0.57	14.50	16.67
Philippines	3.95	3.52	-2.78	11.42	12.98	0.90	12.00	15.00
Poland	3.36	1.71	0.42	8.55	17.87	0.64	16.00	18.50
Portugal	3.02	2.98	-0.94	14.43	20.44	2.95	13.00	22.00
Singapore	0.69	2.50	-5.84	6.37	23.21	1.09	21.00	24.00
Slovakia	0.78	1.45	-2.10	3.25	19.71	0.31	19.33	20.00
Slovenia	1.96	1.60	-2.17	4.53	20.68	2.08	16.33	22.00
South Korea	2.60	2.36	-0.97	9.02	18.23	1.88	11.50	20.67
Spain	2.94	2.58	-1.13	8.59	22.22	2.01	15.33	24.00
Sri Lanka	3.28	5.80	-11.59	17.34	13.24	1.95	11.00	15.50
Sweden	3.70	2.41	-1.81	9.39	23.18	1.08	21.00	24.00
Switzerland	1.92	1.41	-0.53	5.40	24.00	0.00	24.00	24.00
Taiwan	1.26	2.16	-3.27	6.83	20.72	0.14	20.50	21.00
Thailand	1.23	1.97	-3.88	7.37	16.25	0.89	14.50	17.00
Turkey	4.26	3.77	-1.69	12.77	12.77	0.93	11.50	14.67
United Kingdom	2.84	2.52	-3.93	7.09	23.98	0.11	23.33	24.00
United States	1.52	1.89	-2.98	5.11	23.96	0.11	23.67	24.00

*: excluding an extreme outlier in July 2010, when real yields shot up to more than 400%. Note: Real yields are the average monthly rating minus yoy-inflation in that month. Ratings are taken as the average rating at the end of the month.

Table A3: Median estimated coefficients from the yield and rating equation

coefficient	dyields0	dratings0
lyields	-0.033 ***	-0.01 *
facratings7	-0.542 ***	-0.187 **
facratings8	-0.533 ***	-0.179 **
facratings9	-0.427 ***	-0.16 *
facratings10	-0.239 ***	-0.124 *
facratings11	-0.09 **	-0.084 *
facratings12	-0.003	-0.054
facratings13	0.022	-0.041
facratings14	0.016	-0.052
facratings15	-0.015	-0.076 **
facratings16	-0.02	-0.097 **
facratings17	0.003	-0.106 **
facratings18	0.005	-0.066 *
facratings19	0.01	-0.017
facratings20	-0.014	-0.016
facratings21	-0.005	-0.056 *
facratings22	0.023	-0.084 **
facratings23	0.013	-0.133 ***
facratings24	-0.028 *	-0.163 ***
dyields1 _p	0.234 ***	-0.151 **
dyields2 _p		-0.037
dyields3 _p		-0.189 ***
dyields1 _n	0.109 ***	0.153 ***
dyields2 _n		0.116 **
dyields3 _n		0.118 **
dratings0 _p	-0.034	
dratings1 _p	0.002	0.586 ***
dratings2 _p		0.12
dratings0 _n	0.06	
dratings1 _n	-0.06	1.038 ***
dratings2 _n		1.09 ***
const	1.949 ***	
-1—0		-3.914 ***
0—1		0.658
λ	55.142	55.142
# Effective Coefficients	18.305	23.264
Country Fixed Effects	No	No
LL(Data)	3938.304	-1481.072
LL(Smoothing)	-18.191	-16.036
R^2	0.062	0.063
R^2_{adj}	0.061	0.05
BIC	-7709.655	3174.08
AIC	-7839.997	3008.671

Note: *lyields* is the coefficient of lagged yields; *facratings* the dummy for ratings that are at most as good as this level; *dyields* and *dratings* denote lagged differences, with the asymmetric separation captured by the subindices *p* and *n*; -1—0 and 0—1 are the thresholds of the ordered probit ratings model, the equivalent to the constant in the yields model. The smoothing coefficient λ is restricted to be equal in both equations. ***, **, * denote significance of a one sided test at the 1%, 5% and 10%-confidence levels.

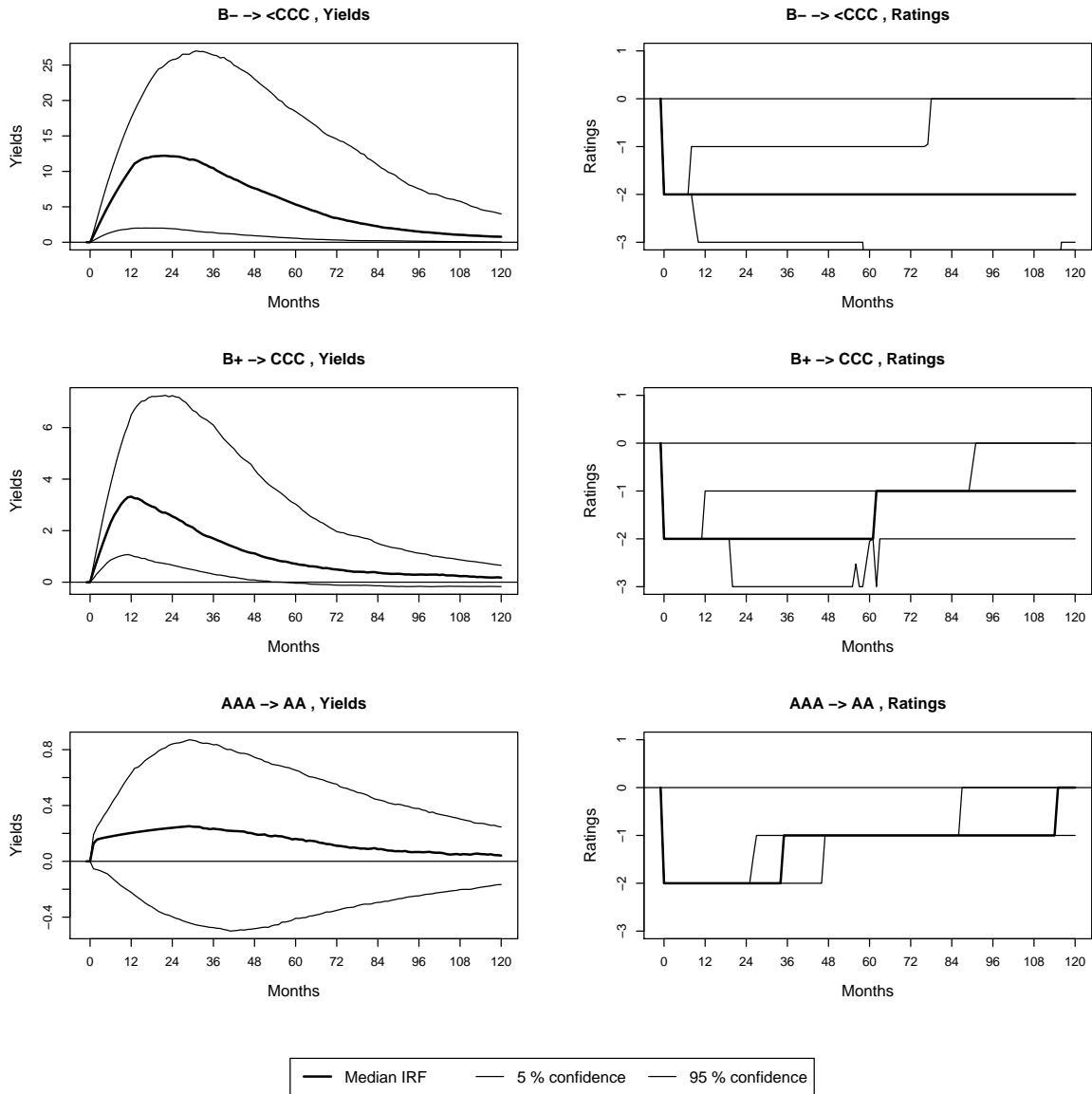


Figure A1: Selected impulse-response functions after a rating shock with confidence bounds.

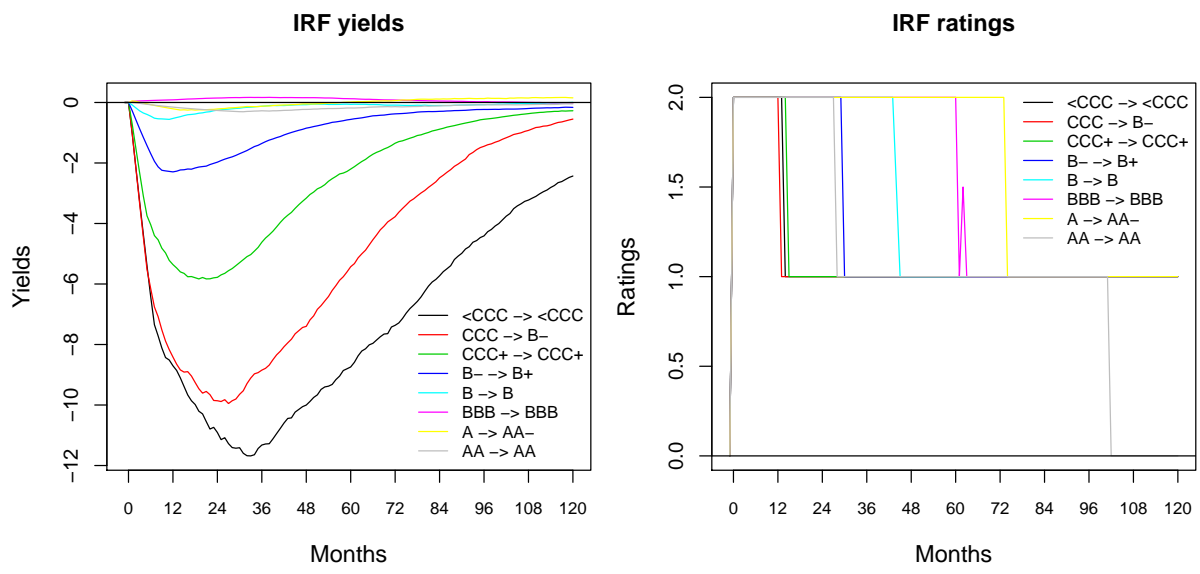


Figure A2: Selected impulse-response functions after a positive rating shock of two notches.

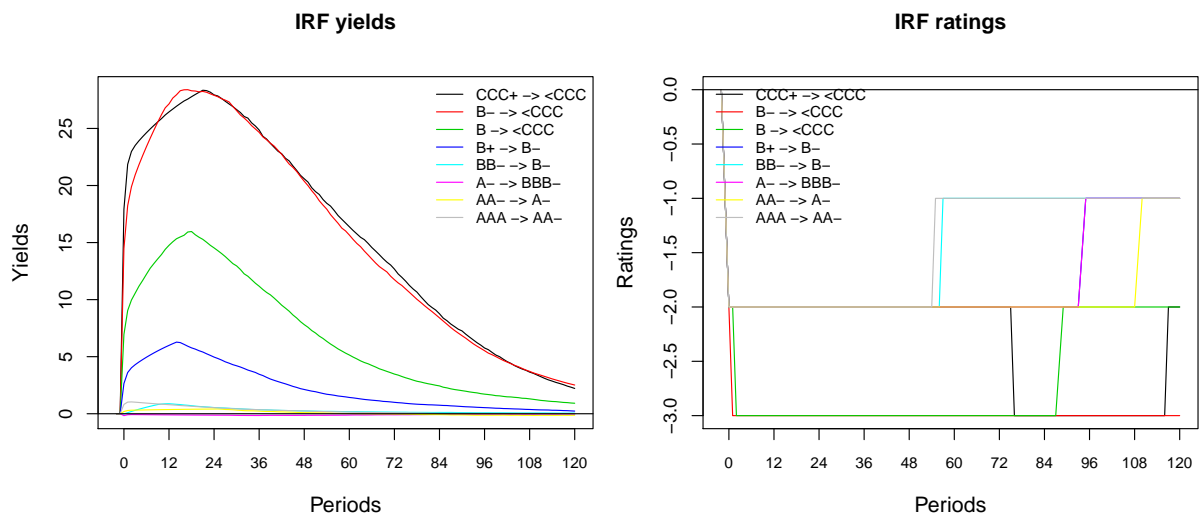


Figure A3: Selected impulse-response functions after a staggered rating shock.

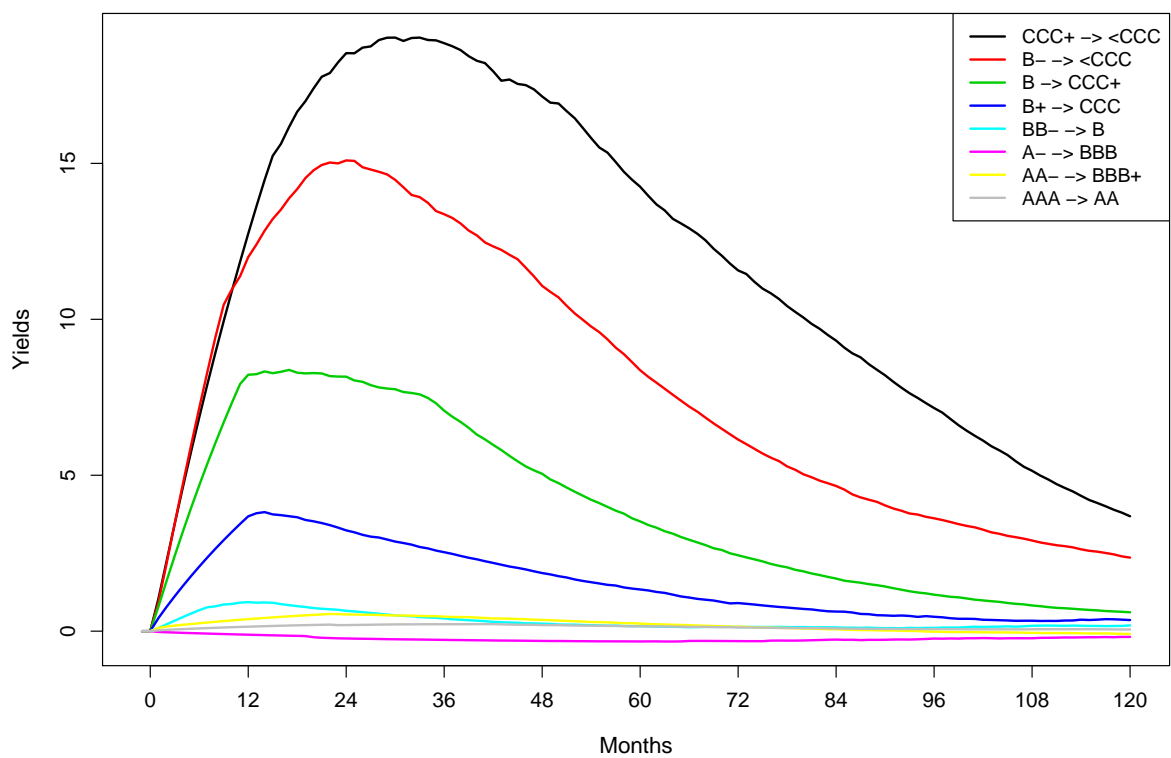


Figure A4: Robustness test with median ratings: Median impulse-response function of yields after a rating shock of two notches, selected ratings.

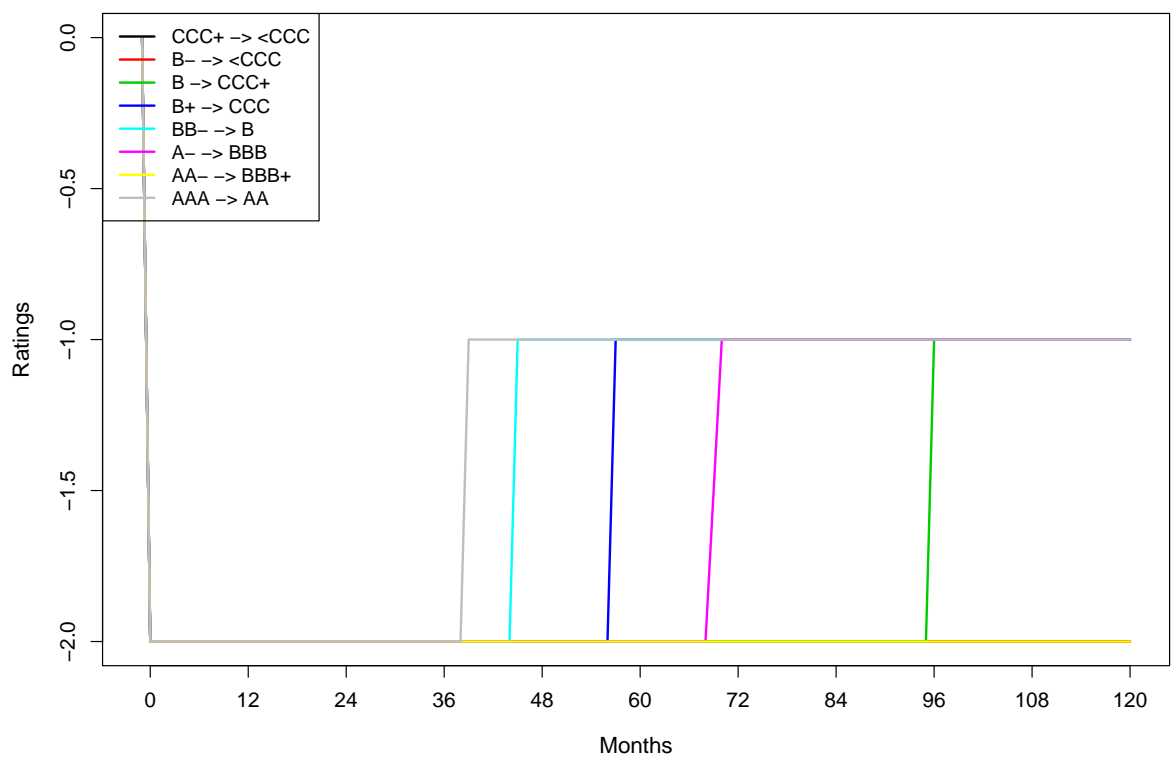


Figure A5: Robustness test with median ratings: Median impulse-response function of ratings after a rating shock of two notches, selected ratings.