



Universität St.Gallen

Standards are Poor: On Competence and Professional Integrity at the Leading Rating Agency

Manfred Gärtner

July 2014 Discussion Paper no. 2014-18

Editor: Martina Flockerzi
University of St.Gallen
School of Economics and Political Science
Department of Economics
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25
Fax +41 71 224 31 35
Email seps@unisg.ch

Publisher: School of Economics and Political Science
Department of Economics
University of St.Gallen
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25
Fax +41 71 224 31 35

Electronic Publication: <http://www.seps.unisg.ch>

Standards are Poor:
On Competence and Professional Integrity at the Leading Rating Agency

Manfred Gärtner

Author's address:

Professor Dr. Manfred Gärtner
Institute of Economics (FGN-HSG)
Bodanstrasse 1
CH-9000 St. Gallen
Phone +41 71 224 2311
Fax +41 71 224 2874
Email manfred.gaertner@unisg.ch
Website www.fgn.unisg.ch/en/professoren/manfred+gaertner

Abstract

Competence behind the rating of sovereign countries is crucial, given that the market for government bonds may be vulnerable to multiple equilibria and self-fulfilling prophecies. With this in mind, this paper reviews and scrutinizes an official Standard and Poor's (S&P) publication that addresses key issues surrounding the market for government bonds and the role of sovereign ratings. It encounters: a low level of competence, revealed in an inability to engage in logical discourse and an inadequate grasp of crucial concepts such as multiple equilibria and self-fulfilling prophecy; obliviousness to S&P's own rating methodology; and a nonchalant treatment of facts that casts a poor light on its professional integrity.

Keywords

Credit rating agencies, competence, integrity, sovereign ratings, Standard and Poor's.

JEL Classification

E65, G12, G24, H63.

The mistakes of a major credit rating agency (CRA) may have serious consequences for the financial sector and, as the financial and sovereign debt crises demonstrate, for the economy and society at large. This is virtually mandated by the regulations that are in place for the financial industry, but also due to and boosted by financial markets' reliance on the competence of rating agencies. Mistakes cannot be avoided, as our understanding of a continuously evolving world remains imperfect and economies and political processes appear to be subject to a constantly changing barrage of shocks, large and small, that may be impossible or very difficult to anticipate. However, an appropriately designed incentive structure that eliminates conflicts of interest and a high level of expertise behind the rating process would appear to constitute necessary conditions for keeping rating errors and the fallout that these may generate to a minimum. With this in mind, this paper examines the expertise that drives sovereign ratings at the world's leading rating agency.

1. Assessing the sovereign rating competence of CRAs

Given the power of CRAs and the decisive role that rating labels play in today's financial markets, the quality of these ratings becomes essential. When assessing this quality, or better, these qualities, it is mandatory to consider each area in which CRAs engage in business separately. Crudely speaking, CRAs evaluate

- (i) the creditworthiness of businesses
- (ii) the riskiness of financial products
- (iii) the solvency of sovereign nations

While the first area has a century-old tradition and is at the very origin of the rating industry, the other two are rather new. It is important to realize that, while the end result is the same in each of these areas – a label picked from a rating scale using the first four letters of the alphabet – each of these three fields needs its own, idiosyncratic kind of expertise. Gauging the risks of highly complex derivatives, which have flooded the financial markets in recent decades, is completely different from judging the business prospects of commercial companies or the creditworthiness of a sovereign country.

This paper looks at the professional competence behind *sovereign credit ratings*. While the rating of sovereigns had already been popular for a spell between the First World War

and the Second World War, as documented by Gaillard (2012), it had largely vanished after the Second World War. As Figure 1 shows, the big three CRAs, Fitch Ratings, Moody's Investors Service and Standard & Poor's Ratings Services, rated no more than about a dozen sovereigns, if any, before the collapse of the communist Eastern bloc in the late 1980s. After that, this business segment took off and by now each of the world's biggest CRAs rates more than 100 countries.

[Figure 1 near here]

In principle, there appear to be three basic ways to evaluate the quality of sovereign credit ratings: (i) evaluate the *forecasting success* of sovereign ratings; (ii) gauge the *consistency* in sovereign ratings over time and between countries; and (iii) appraise the *qualifications and expertise* of the professionals who rate sovereigns.

(i) *Look at the forecasting quality of sovereign ratings*

While this appears to be the natural thing to do, it does not really work – for two reasons.

First and foremost, there are no data, or no sufficient data available that would permit this kind of statistical exercise. No industrial country with a credit rating by Fitch, Moody's or Standard and Poor's defaulted in modern times, up until the Greek haircut in 2012.¹ This problem was even conceded by rating agencies well before the financial crisis:

It is important, though, that investors realise the limitations of this [sovereign debt rating] exercise, which is necessarily far less certain than our ability to analyse either bank or corporate risks of default. The essential problem is that the world of sovereign borrowers is far smaller than the world of large banks or corporations, and that the number of instances of default in the modern period when we have reasonable national accounts is tinier still (See Fitch, 2002, pp. 3f.).

Second, even if we had a sufficiently large number of sovereign defaults, which we do not, we would still have to deal with the problem of self-fulfilling prophecy. As generations of graduate students have learned from Romer's (2011) leading advanced macroeconomics textbook, there is a distinct possibility that more than one equilibrium exists

¹ While there is a limited number of developing countries with a credit rating that had to default or restructure their debt in recent years and decades, this number remains insufficient for serious statistical analysis. Additionally, the question of comparability remains; whether studying the defaults of developing countries generates useful lessons for the rating and solvency issues of high income countries that operate in completely different institutional settings is highly questionable.

in the market for a country's government bonds. Once perceived default risk crosses a given threshold, even initially unfounded surges may generate their own justification in terms of increasing interest rates, rising debt or receding economic growth. The same holds for shocks to the interest rate. In such a situation even sovereign ratings derived from the roll of a die would be warranted by subsequent events and would appear to be competently generated forecasts.

If there is no way to gauge the quality of sovereign ratings by studying their predictive success, we might at least

(b) *Look at the consistency of sovereign ratings procedures over time and between countries*

This is a much softer criterion, of course. Even if consistency was confirmed, this would reveal nothing about the quality of this consistently applied procedure. Since option (a) is not available, however, option (b) has indeed been considered in a number of empirical papers, with results that ask serious questions. This line of research has suggested that

■ **Sovereign ratings are pro-cyclical** and, thus, may well aggravate financial and debt crises (Ferri, Liu and Stiglitz, 1999).

■ **Ratings are biased against debt titles issued by governmental entities** (Cornachia, Cornachia and Hund, 2012). Debt titles issued by the public sector receive ratings that are several notches worse than those of private-sector instruments in the same risk category.

■ **Downgrades of the eurozone's peripheral countries appear to be excessive** (Gärtner, Griesbach and Jung, 2011; Gärtner and Griesbach, 2012; Gärtner, Griesbach and Mennillo, 2013; Vernazza, Nielsen and Gkionakis, 2014). The rating changes during the recent debt crisis cannot not be justified on the basis of econometric rating equations that 'explain' the ratings of other countries or ratings in general before 2009.

■ **Sovereign ratings suffer from a cultural and home bias.** They favour the country where the CRA is domiciled, and English speaking countries in general (Fuchs and Gehring, 2013). These biases have grown during the financial and debt crises.

CRA's have taken issue with this approach that looks for consistency, and thus ducked the questions asked by the results reported in these papers. Accepting their criticism, if only for argument's sake, would force us to retreat to a still weaker criterion, namely:

(c) *Look at the qualifications and expertise of the people who rate sovereigns*

This is a very weak test indeed. Even if sovereign ratings departments of CRAs were exclusively manned by excellently educated ‘Einsteins’, conflicts of interest would still loom and might lead to rating biases or even encourage abuse. On the other hand, without the qualifications needed to understand the idiosyncrasies of the market for government bonds, and how these are related to the economy and the political arena at large, even the most benevolent CRA would fail to carry out the task properly. Therefore, we are looking at a *necessary condition* for high-quality sovereign debt ratings, albeit not at a sufficient one.

It has already been questioned whether CRAs possess the qualifications needed for competent analysis of sovereign finances. Gaillard (2011), reporting on information obtained while researching for his book (Gaillard, 2012), noted in an interview that “(t)here is a lack of macroeconomic expertise at Fitch, Moody’s and Standard & Poor’s”. More recently, after scrutinizing sovereign rating procedures in a larger group of CRAs, the European Securities and Markets Authority (ESMA) gave the following warning and demanded remedial action from CRAs:

ESMA also observed the increasing reliance that lead analysts place on very junior support staff. ESMA is concerned that ... this could pose a risk to the quality of ratings (ESMA, 2013, p. 14).

The current paper takes these concerns, which derive from a perceived lack of credentials in the CRAs sovereign rating departments, seriously. It looks at this issue from a slightly different angle, however, by scrutinizing an official contribution by Standard and Poor’s (henceforth referred to as S&P) to the ongoing discussion about the CRAs’ role in the so-called European sovereign debt crisis. This contribution (Krämer, 2012) was written and disseminated in reaction to Gärtner and Griesbach (2012) (henceforth referred to as GG) and the questions raised in this paper. A brief summary of GG (2012) should prove helpful when evaluating the arguments advanced by S&P.

2. The GG paper in a nutshell

The GG paper starts with a sketch of David Romer’s (2011) model of the market for government bonds. This fairly uncontroversial little model, featuring in the world’s leading advanced macroeconomics textbook for many years, conveys the essence of Calvo’s (1988) optimizing model of sovereign debt crises by means of two structural equations. The first

equation suggests that rising sovereign default risk pushes up the interest rate on government debt. The second equation proposes that a rising interest rate, along with negative developments in economic and political fundamentals, increases a country's propensity to default. Depending on the parameters and the state of a country's economy and public finances, this model may generate two equilibria that are locally stable: a good one, in which both the interest rate and the default risk are low; and a bad one, in which the interest rates rise out of bounds and the government defaults. The neighbourhoods within which these two equilibria are stable are separated by an insolvency threshold (i.e., a third, *unstable* equilibrium). Once the interest rate or perceived default risk crosses its respective threshold value, a self-propelling process is initiated which drives the country closer and closer to insolvency.

Next the paper explores whether this purely theoretical model fits empirical data and may shed light on the sovereign debt crisis that has haunted the euro area since 2010. For this purpose, Romer's model is modified by adding two assumptions:

(i) *Financial investors rely on sovereign debt ratings as an indicator of default risk.*

This permits the use of sovereign ratings data as an independent stand-in for perceived default risk. It also brings CRAs into the picture, which is interesting in itself in the light of concerns about their role during this crisis.²

(ii) *Changes in the interest rate may affect default risk with a lag.*

This assumption derives from the empirical observation that most countries' debt titles have long maturities that extend well beyond the one-period horizon proposed, for simplicity, in the Romer model. Furthermore, it acknowledges that financial markets may not always behave rationally, particularly during times of crisis.³

Having set up this augmented model, displayed graphically in Figure 2, the paper attempts to derive the slopes and curvatures of the two relationships from empirical data for

² For academic contributions which scrutinize the role of credit rating agencies during the recent sovereign debt crisis see Fuchs and Gehring (2013), Gärtner, Griesbach and Jung (2011), Gärtner, Griesbach and Mennillo (2013), Vernazza, Nielsen and Gkionakis (2014) and the papers cited therein. There have also been interventions by politicians (see Reuters, 2010) and by supervisory bodies (see ESMA, 2013).

³ On this latter point, see IMF chief economist Olivier Blanchard (2011), who included "financial investors are schizophrenic", "post the 2008-09 crisis, the world economy is pregnant with multiple equilibria—self-fulfilling outcomes of pessimism or optimism" and "perception moulds reality" among the four hard truths he learned from the year 2011. The assumption is also in line with results from behavioural economics, suggesting that financial investors may form expectations adaptively. See, e.g., Haruvy, Lahav and Noussair (2007).

twenty-five OECD countries. Two versions of the empirical model are provided and discussed.

[Figure 2 near here]

The first version uses coefficient estimates from an earlier paper [Gärtner, Griesbach and Jung (2011)]. This paper explores and estimates *linear* relationships between ratings and interest rates only. Estimates may be interpreted as linear approximations of the curves in the neighbourhood of the good equilibrium. This equilibrium indeed turns out to be locally stable in the sense that erroneous rating downgrades, or increases in the interest rate, have only temporary effects. They do *not* trigger a destabilising, self-propelling process of continuously increasing interest rates and sovereign downgrades that would drive the country towards default.

The second version moves beyond the results presented by Gärtner, Griesbach and Jung (2011) by exploring the possibility of nonlinear relationships between interest rates and ratings as the key contribution of the paper. Moreover, it expands the data range to the year 2011, when the debt crisis approached its peak.⁴ The main findings of this core section are:

(1) *Ratings respond linearly to changes in the interest rate.*

Graphically, this results in the straight rating line shown in Figure 2. As mentioned, interest rates are allowed to affect ratings with a lag in order to accommodate the reality of long-maturity debt titles. The line moves to the right when a country's financial or macroeconomic situation deteriorates, as measured by variables such as debt or deficit ratios or income per capita.

(2) *There is pronounced nonlinearity in how ratings affect interest rates*

This is shown by the convex interest rate line in Figure 2. Successive downgrades trigger bigger and bigger increases in the interest rate. This curve moves upwards when the risk-free rate increases, for which Germany's interest rate is used as a proxy.

(3) *The model may give rise to multiple equilibria and self-fulfilling prophecy.*

The situation displayed in Figure 2 is synthetic in the sense that sample averages for the exogenous variables are used to position the two lines. There are two equilibria, identified by the points of intersection. In line with the Romer model, the good equilibrium is locally

⁴ The period covered by the panel regressions is 1999-2011.

stable. The second equilibrium is unstable and constitutes a threshold. Once this threshold is crossed, the interest rate increases without limits, the country's debt is downgraded to 'junk', and the government is unable to refinance expiring bond titles and defaults. This situation constitutes a third, stable 'equilibrium', even though it is not identified by a point of intersection. On average, for the countries and years included in the data set, the insolvency threshold or abyss appears to lie between the A and the B segment of the rating scale. For individual countries, the threshold values may be different, depending on the state of their economy and their public finances.

A fourth result suggests that the risks that derive from the possibility of multiple equilibria and self-fulfilling prophecy in the market for government bonds may be real. It states:

(4) Several eurozone countries were downgraded 'excessively' during 2009-2011

when a rating equation is used for reference that explains the pre-crisis ratings of all the countries in the sample and the crisis ratings of other countries. For example, deteriorating fundamentals would have justified a downgrade of Ireland's creditworthiness by one to two rating categories during those two years. In reality, Ireland was downgraded by seven to nine notches by the three leading CRAs.⁵

After a discussion of dynamic extensions of the model, the paper concludes by stating:

(A)t least for countries with sovereign debt ratings outside the A range even erroneous, arbitrary or abusive rating downgrades may easily generate the very conditions that do actually justify the rating. Combined with earlier evidence that many of the rating downgrades of the eurozone's peripheral countries ... could not be justified on the basis of rating algorithms that explain the ratings of other countries or ratings before 2009, this result is highly discomfoting. It urges governments to take a long overdue close look at financial markets in general, and at sovereign bond markets in particular, and at the motivations, dependencies and conflicts of interest of key players in these markets. (p. 28)

⁵ See Gärtner, Griesbach and Mennillo (2013) for a more detailed account of this episode.

3. The reaction of Standard and Poor's

Standard & Poor's Rating Services, the world's leading credit rating agency in terms of revenue and employment, published a sweeping defence to GG (2012), succinctly summarized by the publication's heading: *S&P's Ratings Are Not "Self-Fulfilling Prophecies"*.⁶

One might timidly object that the GG paper does not really say anything about S&P because it used Fitch's data in the reported econometric work. However, this would be splitting hairs. One might also insist that the GG paper does not conclude that ratings "are" self-fulfilling, but only that ratings *may* become self-fulfilling once they cross a certain threshold – a threshold that varies across countries and over time.

Putting such subtleties aside, the key question is: How does S&P substantiate its criticism and denial? We start by looking at the reception of the GG paper by S&P. This is followed by a discussion of the specific points of criticism and arguments, in an attempt to check them for validity.

3.1. Standard and Poor's reception of the GG paper

S&P starts by alleging that the GG paper "accuses rating agencies of 'erroneous, arbitrary or abusive rating downgrades' in the context of the European debt crisis". This is as if climate scientists had concluded from their computer-based model that global warming might accelerate if cars' powertrains do not become more efficient and then Toyota interpreted this as having been accused of not producing more efficient cars. The citation that S&P registered as an *accusation* is the implication of a thought experiment, as is apparent from reading the quote in context:

A more detailed look at the dynamics of the effect of debt rating downgrades on interest rates revealed that at least for countries with sovereign debt ratings outside the A range even erroneous, arbitrary or abusive rating downgrades may easily generate the very conditions that do actually justify the rating (GG, 2012, p. 28).

S&P appears to have interpreted a thought experiment, conducted in the context of an empirical model, as an allusion to reality.

⁶ The critique is signed by Moritz Krämer, Managing Director of Sovereign Ratings for Europe, the Middle East and Africa. Therefore, it appears fair to read it as Standard and Poor's position on the addressed issues. A shorter version of this critique was published in a German economic weekly (Krämer, 2012a).

Another, serious S&P misperception of what the GG paper expressed and was trying to achieve appears near the end of the S&P critique. There S&P claims that the GG paper “does offer an alternative rating approach with its quantitative algorithm” and generously “welcome(s) the competition of ideas and diverse methodologies to assess sovereign credit risk. Whether users of ratings will accept the algorithm remains to be seen, however.”

Contrary to S&P’s perception, the rating equation presented by GG (2012) does not propose an alternative or purportedly superior rating methodology. It does not tell CRAs how they *should* conduct their rating of sovereigns. It offers an interpretation of how they *did* conduct their rating. It is an attempt, employing established methods from empirical economic research, to retrace what appears to drive actual ratings. Looking through the glasses of mainstream macroeconomics, it explores whether and how those variables that are considered crucial for a country’s solvency, the debt ratio, the deficit ratio, economic growth etc., affect a country’s credit rating. The resulting numerical equation quantifies how those key variables made their way through the complex rating procedures of CRAs. This equation cannot explain or trace every nuance of rating differences and changes. However, it may be seen as a core process behind the pertinent work of CRAs, a core process that explains a major part of the rating gaps between countries and of rating changes over time. No claim was made that CRAs are consciously using this empirical model.⁷

The sobering conclusion here is that S&P unfortunately failed to grasp the direction of GG (2012) and, hence, misses the questions posed by the paper’s results.

3.2. Standard and Poor’s points of criticism

Expanding on the claim made in its comment’s title, S&P maintains that GG’s “confusion of correlation and causality is the flawed foundation of their claim that rating actions are self-fulfilling prophecies”.

Let us structure this argument. Self-fulfilling prophecies would be triggered and propelled by what S&P calls a “negative-feedback loop”.⁸ The two segments of such a loop,

⁷ If applied economists estimate that German exports fall by two percent when the Euro appreciates by one percent, this would not require UK car buyers to be aware of this and consciously decide to purchase a two percent cheaper German car or switch to a non-German brand with a probability of five percent.

⁸ This terminology is a little unconventional, because *negative* feedback is usually associated with self-correcting processes. Self-reinforcing processes, which is what self-fulfilling prophecies amount to, are the result of *positive* feedback.

the existence of which S&P denied, would be the interest rate equation and the rating equation estimated by GG (2012). These imply that:

- (1) deteriorating credit ratings lead to higher interest rates
- (2) rising interest rates have a negative effect on a country's credit rating

The interesting – and in the context of the market for governments bonds discomfoting – property of such an interactive process is that any unjustified increase in the interest rate would lead to a downgrade, which would, ex post, provide justification for the initially unjustified interest rate hike. By the same token, an unjustified downgrade would drive up the interest rate which, ex post, provides justification for the initially erroneous or arbitrary downgrade.

If the quantitative responses are small enough, this process is *stable* (or self-correcting), meaning that after a shock to the interest rate or to the credit rating the market returns to the initial equilibrium. There is still an element of self-fulfilling prophecy, but this is not large enough to provide full justification for the initial error (or shock). Therefore, the effect remains temporary and vanishes over time.

Large responses, however, render the system *unstable* (or self-reinforcing). Any errors committed by CRAs, or by the market in general, become self-fulfilling in full guise, triggering an avalanche of further downgrades and interest rate increases that ultimately ends in default.

S&P denies the possibility of a self-fulfilling prophecy and, hence, the existence of, in its terminology, a negative-feedback loop in which its ratings may 'mould reality'. This is tantamount to denying link (1) described above, or link (2), or both. Since the statistical correlation between the interest rate and the rating is undisputed, this boils down to the issue of causality.

Let us consider what we can say about links (1) and (2):

Link (1): Do sovereign ratings affect the interest rate?

In search for an answer we may (i) consider statistical evidence (ii) take the effects of financial market regulations into account, (iii) learn from real-world experiments and (iv) discusses the consequences of asymmetries, reputation spill-overs and insider information.

(i) *Statistical evidence*

While the issue of causality between interest rates and sovereign ratings is a tricky matter indeed, as is the issue of causality in economics in general, there are statistical methods at our disposal that offer help. One example is the test for Granger causality. Pertinent causality tests have been conducted by a number of authors. Results were reported by Gärtner, Griesbach and Jung (2011), along with references to other studies reaching similar results. Using weekly data, their tests “can never reject the hypothesis that the credit spread [the key component of which is the interest rate] is caused by the rating” (Gärtner, Griesbach and Jung, 2011, p. 295).⁹

Of course, such tests have well-known limitations. They never give final proof, one way or the other, but only serve as pieces of evidence – no more, no less. That said, there is an important difference between abstract statistical reasoning, and econometrics. Econometrics combines the information that we may crunch from the numbers alone with information about the institutions within which the data were generated, with an understanding of the incentive structures underlying the decisions of key participants, and more. So, let us look at the issue of causality from such a more general angle.

(ii) *Regulations*

Financial market regulations such as the Basel II accord oftentimes force financial institutions to respond to the downgrading of a country by adjusting the composition of their portfolio, regardless of whether these institutions consider the downgrade to be justified or not. A downgrade of Portugal may trigger the sale of Portuguese government bonds and, hence, the interest rate on new debt issues, on purely technical terms. Of course, the responsibility for this technical effect lies with the governments that implemented pertinent regulations. However, the effect exists.

(iii) *Real-world experiments*

On occasion reality provides us with experiments from which we can learn more than we can learn from theoretical and statistical discourses. One such experiment played out on 10 November 2011:

⁹ High-frequency data cannot be used in the estimation of the rating equation because a number of key variables, including the public debt ratio or measures of government deficits, are only available on an annual basis.

Standard & Poor's mistakenly announced the downgrade of France's top credit rating on Thursday ... The erroneous alert, which S&P said was sent to some of its subscribers, ... contributed to the worst day for France's government bonds since before the euro was launched in 1999. ... In a statement issued nearly two hours after the fact, S&P said the message resulted from a technical error and not from any action it intended to take against France. (Reuters, 2011)

As word spread Thursday about the message, the euro weakened against the dollar, U.S. stocks slipped and French bond prices fell, pushing yields higher. Cash reversed course, flooding into U.S. Treasuries and German government bonds. (Wall Street Journal, 2011)

(iv) *Asymmetries, reputation transfers, and insider information*

There is a host of other channels through which sovereign ratings may make their way into the market for governments bonds:

Asymmetries in the incentives and reward structure of account managers at banks: Suppose that Portugal suffers a downgrade that my account manager considers unjustified. She may, therefore, recommend that I hold on to my Portuguese government bonds. In case Portugal defaults nevertheless, she will find it hard to justify this advice to her clients and her boss. If she advises me to sell my Portuguese government bonds and Portugal does not default, she has the convenient excuse that she only heeded the "opinion" of some well-reputed rating agency.

Actual or imagined insider knowledge: Many sovereign ratings are unsolicited, which means that they are based on publicly available information. Some are solicited, however, which might mean that these in part reflect information that is not publicly available. Investors oftentimes cannot or do not discriminate between these two categories and thus suspect some private information behind any sovereign rating.

Reputation spill-overs: Ratings of companies have been in existence for a century. As a rule, when CRAs rate commercial companies they draw on information that the typical investor does not have. In this context, ratings thus serve an important purpose, and this is how rating agencies built a reputation. Many market participants transport this reputation to the rather new sovereign rating activities, even though this business area requires an entirely different kind of expertise.

Rationality: Even the most rational investor who does not believe that sovereign ratings have any relevant informational content is forced to respond to them if he believes that other market participants will (be forced to) respond to these ratings.

With all these observations and arguments, it appears difficult to doubt that sovereign ratings exert a causal effect on the interest rate.

Let us now turn to the second segment of the feedback loop.

Link (2): Do interest rates affect sovereign ratings?

This is the loop's segment that was questioned most vehemently by S&P. In their own words:

Standard & Poor's sovereign methodology does not take credit spreads into consideration. We believe that spread movements provide neither diagnostic capabilities nor do they systematically signal insights into the fundamental factors that differentiate one credit from another (Krämer, 2012, p. 3)

This is a strong statement. It implies that it is irrelevant whether Japan pays an interest rate of 2 percent on its more than 200 percent of public debt relative to GDP, or 10 percent or 20 percent. A look at Standard & Poor's (2011) sovereign rating methodology gives comfort and a scare at the same time, however:

Among the "five key factors underlying our sovereign rating analysis" is the "fiscal score" (p. 4). One of two factors determining the fiscal score is a country's "debt burden" (p. 23). And the debt burden is determined by the "General government's interest expenditures as a percentage of government revenues" (p. 27).

The good news is that S&P does indeed take credit spreads into account, which is nothing but the difference between a country's interest rate and the risk-free rate. The bad news is that S&P does not seem to be aware of its own rating methodology. For all its disturbing implications, this borders on the comical, given that S&P "wonder(s) whether the authors [i.e., GG] have even read the [CRAs] methodologies" (p. 3).

Ignorance of its own rating methodology and hastily brushing aside uncomfortable questions plagues other S&P contributions to the ongoing discussion of sovereign debt ratings as well. One recent example is the response to a widely discussed paper by a team of UniCredit economists (Vernazza, Nielsen and Gkionakis, 2014). This paper reported unexplained downgrades of countries from the eurozone's periphery similar to those given by GG (2012). Taking issue with UniCredit's claim that "there is little dispute about which variables should be included in an analysis of a country's creditworthiness", S&P's untitled response retorted:

We dispute UniCredit's analysis and selection of variables. In fact, **none of the ten variables chosen by them is referenced in our own methodology** [my emphasis; M.G.]. (Standard and Poor's, 2014).

Investigating this, Table 1 lists the ten variables used in UniCredit's regression analysis on the left. The second column reports results of a search in the S&P methodology for corresponding items and shows pertinent quotes. After comparing the two columns the verdict is clear: The only variable included in the UniCredit regressions that has no counterpart in the S&P rating manual is nominal GDP.¹⁰ All other nine variables feature in the S&P methodology – either directly, with minor differences in the fine print, or are at least mentioned and discussed as adjustment factors. One example is GDP per capita which both sources measure in U.S. dollars. UniCredit adjusts nominal numbers for purchasing power. S&P makes adjustment for currency over- and undervaluation, which is essentially the same thing. Another example is public debt, for which UniCredit uses gross debt while S&P looks at net debt. A final example, at the bottom of the table, is law. UniCredit uses data from the World Bank Rule of Law index. S&P looks at the *rule of law*.

The devastating verdict is that S&P's statement is grossly misleading, a fabrication. Rather than denying that any of these variables is referenced in S&P's methodology, a truthful statement would have been "yes, we also look at almost all of the variables that enter UniCredit's regression equation". Therefore, beyond splitting hairs because of minor differences in the fine print, the impression arises that a key weapon of S&P in defending their sovereign rating decisions against academic and professional critique is to mislead readers and the public.

[Table 1 near here]

This brings us to the factual claims advanced by S&P. As mentioned and discussed above, a key concern of S&P's critique of the GG paper is the possibility of a feedback loop between the interest rate and the debt rating. Justifying its denial of such a loop and, thus, ruling out the risk of a self-fulfilling prophecy, S&P claimed: "(T)he [GG] paper offers no support for its claim of a negative-feedback loop" (Krämer, 2012, p. 2). Given that GG (2012)

¹⁰ Nominal GDP, as opposed to GDP *per capita*, is a variable that is indeed hard to justify. It implies that the sheer size of an economy matters for the rating of sovereign debt. Unsurprisingly, this variable is insignificant in three of the four regressions shown in the body of the text. In the fourth equation it is significant at the 90 percent level.

distils exactly such a feedback loop from the data, S&P's criticism appears odd. One may take issue with the result reported in GG (2012), one may call for different empirical methods, insist on the inclusion of other variables and of interaction terms, question the time range, the selection of countries, and more. But when S&P asserts that "the authors' claims remain unsupported by evidence" (Krämer, 2012, p. 2) this casts a strange light on their professional integrity.

S&P also maintained:

In our view, the algorithm [reported in GG] ... would have entirely missed the Greek default in early 2012, the largest sovereign restructuring in financial history. By contrast, far from having acted in an "arbitrary or abusive" manner, Standard & Poor's anticipated Greece's default well before it occurred (Krämer, 2012, p. 4).

This only reconfirms S&P's misunderstanding of the questions being posed. When the question asked by GG (2012) and others, such as Vernazza, Nielsen and Gkionakis (2014), is whether Greece defaulted, and other countries ran into similar trouble, *because of* untraceable, excessive-looking downgrades that triggered a run of self-fulfilling prophecy, how can the response be to crow that one "predicted" Greece's 2012 default?

Moreover, apart from the issue and possibility of self-fulfilling prophecy, S&P's understanding of anticipating a default "well before it occurred" is quite unusual. Actually, S&P had bestowed upon Greece an A rating a few months after the country joined the eurozone on 1 January 2001. Greece never slipped below this mark until 14 January 2009, when it was downgraded one notch to A-. This happened four months *after* the collapse of Lehman Brothers, when it had become common knowledge that the financial crisis would have serious consequences for government budgets. Despite this downgrade, Greece's rating was still in the investment grade range which is characterized by "strong capacity to meet financial commitments, but somewhat susceptible to adverse economic conditions and changes in circumstances" (Standard & Poor's, 2014a). To cite this very late and trifling initial downgrade as evidence of S&P's farsightedness may classify as quite an overstatement.

In the same vein, in its response to Vernazza, Nielsen and Gkionakis (2014), S&P boasted: "S&P was the first rating agency to raise concerns about brewing euro area problems when we started to lower peripheral sovereign ratings, beginning with Italy in mid-2004." What had actually happened was that during the five years between 2004 and 2008, S&P issued four downgrades and two upgrades of Eurozone countries, which hardly

classifies as a wake-up call for a currency zone with fifteen member countries (as of 2008). It took until 2009, well after the collapse of Lehman Brothers, when all hell broke loose, one crisis summit chased another and a blind mole saw budget problems coming, for S&P and the other CRAs to start a virtual orgy of procyclical downgrades, during which the downgrades for the well-known eurozone victims massively exceeded what was called for by deteriorating fundamentals. The culmination points were the years 2011 and 2012, with 44 and 43 downgrades by the Big Three, respectively. During those years it felt like no week, sometimes even not a day, passed without another downgrade, outlook deterioration, negative watch issue or other bad news about public finances hitting the media.^{11,12} So, again, S&P is adorning itself with unearned laurels and with deceptive factual claims.

4. Conclusion

Competence behind sovereign ratings is crucial, given that the market for government bonds may be vulnerable to multiple equilibria and self-fulfilling prophecies. Several authors have questioned the expertise that drives sovereign rating decisions, pointing to a lack of advanced formal education in macroeconomics (Gaillard, 2012) in the sovereign ratings departments of CRAs, and to a tendency to shoulder junior staff with crucial tasks and decisions they may not be up to (yet) (ESMA, 2013). The current paper augments such earlier criticism by taking a close look at an official S&P publication that addresses the issues surrounding the market for government bonds and the role of sovereign ratings. The sobering result is that this contribution displays (i) inability to engage in logical discourse, (ii) lack of understanding crucial concepts such as multiple equilibria and self-fulfilling prophecy, (iii) obliviousness of S&P's own rating methodology, and (iv) nonchalance in making factual claims that casts poor light on the professional integrity and the credibility of S&P. This should be worrisome for anybody whose well-being and future depends on sovereign rating verdicts – which includes virtually all of us.

¹¹ With the big three rating agencies very much marching in step, these downgrades were shared more or less equally between Fitch, Moody's and S&P.

¹² The shift looks even more striking when we look at the magnitude of the rating changes published by S&P. Between 2004 and 2008 eurozone member countries received net downgrades – the difference between the sum of the notches (or steps) by which countries were downgraded and the sum of the notches by which they were upgraded – of 4 notches. By comparison, between 2009 and 2013 eurozone members were dealt net downgrades totalling 59 notches!

References

Blanchard, Olivier (2011), In Review: Four Hard Truths. iMFdirect, 21 December. <http://blog-imfdirect.imf.org/2011/12/21/2011-in-review-four-hard-truths/>. Retrieved 11 June 2014.

Calvo, Guillermo A. (1988), Servicing the public debt: The role of expectations. *American Economic Review* 78(4), 647-61.

Cornaggia, Jess, Kimberly Rogers Cornaggia and John Hund (2012), Credit Ratings across Asset Classes: A ≠ A?'. SSRN eLibrary .

ESMA (2013). Credit Rating Agencies. Sovereign ratings investigation. ESMA's assessment of governance, conflicts of interest, resourcing adequacy and confidentiality controls. http://www.esma.europa.eu/system/files/2013-1780_esma_identifies_deficiencies_in_cras_sovereign_ratings_processes.pdf. Retrieved 11 June 2014.

Ferri, Giovanni. Li-Gang Liu and Joseph E. Stiglitz (1999). The procyclical role of rating agencies: evidence from the East Asian crisis. *Economic Notes* 28(3): 335-355.

Fitch Ratings: Sovereign Ratings, Rating Methodology (2002). <http://www.fitchratings.com.bo/Upload/methodology.pdf> . Retrieved 31 January 2012.

Fuchs, Andreas and Kai Gehring (2013). The Home Bias in Sovereign Ratings. University of Heidelberg, Department of Economics Discussion Paper Series No. 552, December.

Gaillard, Norbert (2011). Interview. France 24. <http://www.france24.com/en/20110901-interview-norbert-gaillard-economist--world-bank-credit-ratings-agencies-sovereign-debt-moodys-standard-poors-fitch-us-/>. Retrieved 3 July 2014.

Gaillard, Norbert (2012). *A Century of Sovereign Ratings*. Springer.

Gärtner, Manfred, Björn Griesbach and Florian 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.

Gärtner, Manfred and Björn Griesbach (2012). Rating Agencies, Self-Fulfilling Prophecy and Multiple Equilibria? An Empirical Model of the European Sovereign Debt Crisis 2009-2011. University of St. Gallen, School of Economics and Political Science, Discussion paper no. 2012-15, June 2013. <http://www1.vwa.unisg.ch/RePEc/usg/econwp/EWP-1215.pdf>

Gärtner, Manfred, Björn Griesbach and Giulia Mennillo (2013). The near-death experience of the Celtic Tiger: A model-driven narrative from the European sovereign debt crisis. *Intereconomics - Review of European Economic Policy* 48(6): 358-365.

Haruvy, Ernan, Yaron Lahav, and Charles N. Noussair (2007). Traders' Expectations in Asset Markets: Experimental Evidence. *American Economic Review*, 97(5): 1901-1920.

Krämer, Moritz (2012). S&P's Ratings Are Not "Self-Fulfilling Prophecy". *Ratings Direct*. Standard & Poor's Ratings Services. August. Found 27 March, 2013
http://www.standardandpoors.com/spf/upload/Ratings_US/RatingsDirect_Commentary_1003411_08_31_2012_09_23_36.pdf

Krämer, Moritz (2012a). Ratings sind keine „sich selbst erfüllenden Prophezeiungen“. *Wirtschaftswoche*, 28 September.

Reuters (2010). Rating agencies warned to watch their step.
<http://blogs.reuters.com/global/2010/04/29/rating-agencies-warned-to-watch-their-step/>. Retrieved 11 June 2014.

Reuters (2011). France shocked by S&P downgrade error.
<http://uk.reuters.com/article/2011/11/11/uk-france-ratings-sandp-error-idUKTRE7AA12820111111>. Retrieved 28 March 2014.

Romer, David (2011). *Advanced Macroeconomics*. 4th edition. McGraw-Hill.

Standard & Poor's (2011). Sovereign Government Rating Methodology And Assumptions. 30 June. <http://www.standardandpoors.com/servlet/BlobServer?blobheadername3=MDT-Type&blobcol=urldata&blobtable=MungoBlobs&blobheadervalue2=inline%3B+filename%3DCriteriaGovernmentsSovereignsSovereignGovernmentRatingMethodologyAndAssumptions.pdf&blobheadername2=Content-Disposition&blobheadervalue1=application%2Fpdf&blobkey=id&blobheadername1=content-type&blobwhere=1243943646614&blobheadervalue3=UTF-8>. Retrieved 3 July 2014.

Standard & Poor's (2014). Untitled manuscript. <http://ftalphaville.ft.com/files/2014/03/SP-UniCredit-response-27-Mar-14-1.docx>. Retrieved 29 May 2014.

Standard & Poor's (2014a). Credit Rating definitions & FAQs . <http://www.standardandpoors.com/ratings/definitions-and-faqs/en/us>. Retrieved 6 June 2014.

Vernazza, Daniel, Erik F. Nielsen and Vasileios Gkionakis (2014). The Damaging Bias of Sovereign Ratings. UniCredit Global Themes Series 21, 26 March. http://www.dt.tesoro.it/export/sites/sitodt/modules/documenti_en/analisi_progammazione/rown_bag/GlobalThemesSeries_26Mar14.pdf. Retrieved 2 June 2014.

Wall Street Journal (2011). S&P 'Oops' on Rating of France Is Probed. <http://online.wsj.com/news/articles/SB10001424052970204224604577030083804142906>. Retrieved 28 March 2014.

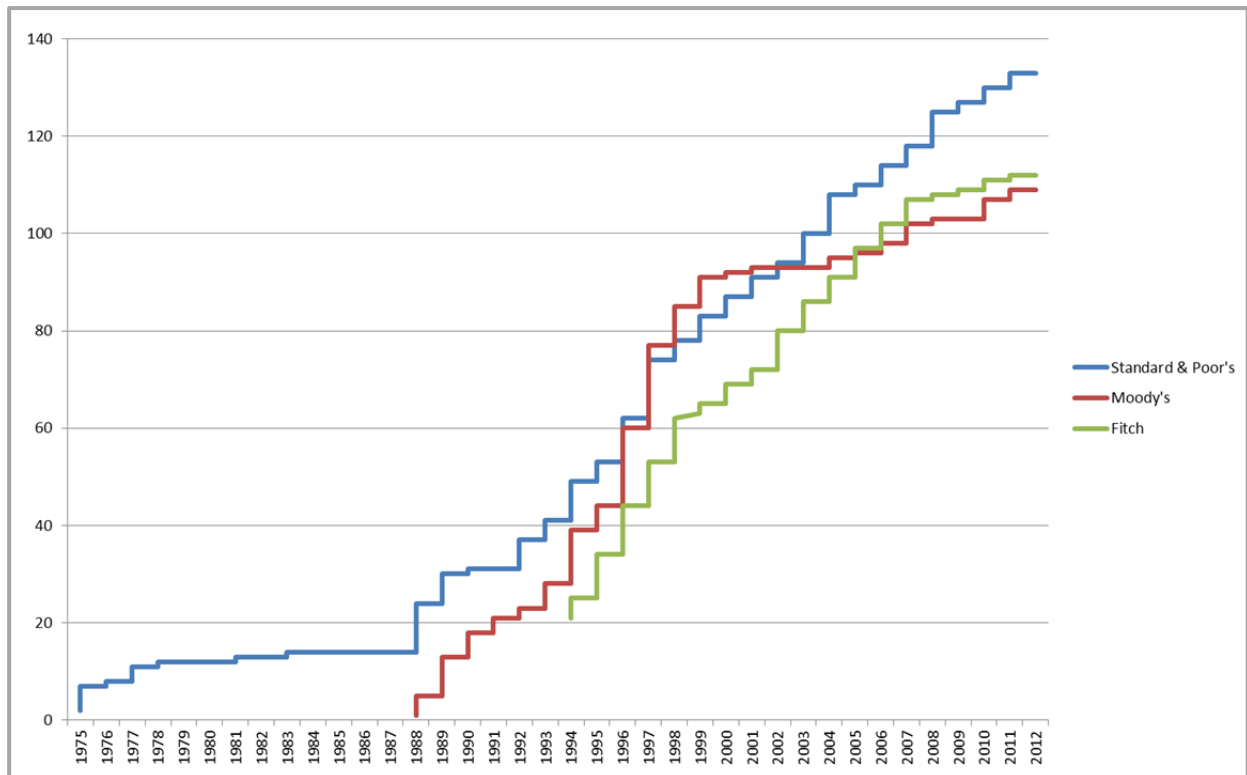


Figure 1.
The number of countries with a sovereign debt rating by the big three rating agencies
Source: Author's calculations from the web pages of Fitch, Moody's and Standard and Poor's.

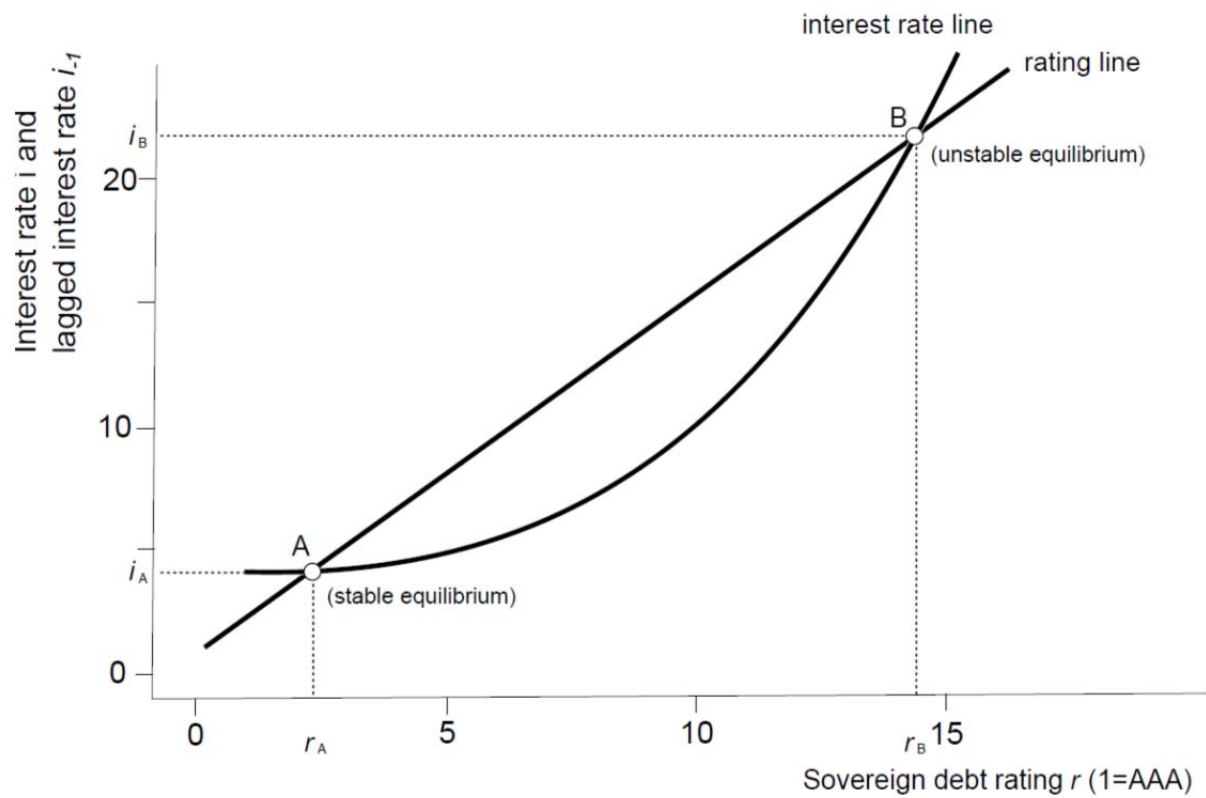


Figure 2.
An empirical model of the market for government bonds

Source: Gärtner, Griesbach and Mennillo (2013).

Table 1.

Comparing regressors used in UniCredit paper with S&P methodology

Claim: “(N)one of the ten variables chosen by [UniCredit] is referenced in our own methodology” (Standard & Poor’s, 2014).

Variables used by UniCredit	Standard & Poor’s methodology
1 Nominal GDP GDP in current prices and exchange rates USD	<i>Not mentioned</i>
2 GDP per capita GDP per person, PPP-adjusted USD	GDP per capita In USD is Standard & Poor’s most prominent measure of income levels [16] <i>Adjustment for over- and undervaluation</i>
3 GDP growth Average annual real GDP growth	Real per capita GDP trend growth [17]
4 Public debt General government gross debt, end-of-year, % of GDP	Debt level Net general government debt as percentage of GDP [27]
5 Current account Annual current account balance, % of GDP	Current account on average over the last historical year, the current year, and the next two forecast years [20]
6 External debt Gross external debt, % of GDP	External indebtedness ratio of “narrow net external debt” to current account receipts [19]
7 Past default Indicator variable that takes the value one in all years following a default event since 1960; zero otherwise	<i>No direct mention, but indirect references</i> (A) government’s debt payment culture represents a credit risk. [14] (W)ith each successive default, serial defaulters have less of a reputation to lose. [14]
8 Advanced country An indicator variable that takes the value one if the country is deemed advanced by the IMF; zero otherwise	<i>No direct mention</i> <i>Many factors in discussed in methodology reflect economic and institutional development.</i> Strength and stability of a country’s institutions [7]
9 Government World Bank Government Effectiveness index	Political score The World Bank’s “Worldwide Governance Indicators,” which measure six broad dimensions of governance ... including ... governance effectiveness ... [12-13] Governance and political risk are among the main drivers ... that lead to default. [9] strength and stability of the government’s institutions, and the effectiveness of its policy-making [7]
10 Law World Bank Rule of Law index	Rule of law [14]

Note: All terms and test strings in regular typeface are quotes from Vernazza, Nielsen and Gkionakis (2014) or Standard & Poor’s (2011) (page number for quotes in brackets). My own comments or qualifications are given in italics.