# Why they keep missing: An empirical investigation of rational inattention of rating agencies*

Makram El-Shagi[†1,2] and Gregor von Schweinitz[2,3]

[1]School of Economics, Henan University, China
[2]Halle Institute for Economic Research, Germany
[3]Deutsche Bundesbank, Germany

This version: January 10, 2017

**Abstract**

Although there is a wide consensus that rating agencies have frequently failed to predict major crises, the literature on sovereign ratings has so far mostly focused on explaining the rating level rather than explaining the timing of the rating decision. In this paper we aim to fill this gap in the literature. Moreover, we go beyond the previous literature by explicitly differentiating between a decision to assess a country and the actual rating decision. Thereby, we take rational inattention of rating agencies into account that should exist due to the cognitive and informational costs of a reassessment. Exploiting information of rating announcements, we can show that (i) the differentiation between the two decision processes significantly improves the model explaining rating decisions; (ii) rating agencies take many nonfundamental factors in their decision to reassess a country into account; (iii) markets only react to ratings if these ratings supply genuinely new information; and (iv) that developed country get preferential treatment.

**Keywords:** Rating agencies; sovereign risk; rational inattention
**JEL-Classification:** C14; C25; F34; G24

*This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff.

†Corresponding author. Email: elshagi@henu.edu.cn

# 1 Introduction

If there is one thing the literature on sovereign ratings agrees on, it is the fact that rating agencies usually act considerably after signs of a changed risk perception (such as capital flows or changing yields) are clearly visible in the market. Originally this was seen as an anticipation effect (Cantor & Packer 1996). However, recent research focusing on the dynamics of ratings (Hu, Kiesel & Perraudin 2002, El-Shagi & von Schweinitz 2015) shows the great difficulties in precisely predicting rating events due to generally low transition probabilities. This casts some doubt on the anticipation hypothesis.

Thus, the interpretation that currently dominates the literature is far less favorable for rating agencies. Critics of rating agencies suspect that rating agencies – rather than predicting risk – respond to the market. Since investors take the rating as news anyways, market based risk indicators such as the government bond yield increase as response to ratings. Thus, this can generate a vicious cycle of downgrades and increasing risk premia or capital flight (Ferri, Liu & Stiglitz 1999, White 2010). Contrarily, the defendants of rating agencies believe that rating agencies respond to the market or at least later than the market. Since this is well understood by investors, they have only limited impact on the market. In short, both concurrent views on rating agencies share the view that rating agencies - despite their role as risk assessors, are rarely if ever the first to notice upcoming problems that increase the risk of default (Mora 2006, El-Shagi 2010, El-Shagi & von Schweinitz 2015).

Yet, it seems premature to judge that rating agencies are unable to provide information. In many instances there is evidence that interest rates respond to rating changes (e.g. Ferri et al. 1999, De Santis 2012). This response is heterogeneous in the degree of surprise of these changes, see Goh & Ederington (1993) (for the corporate bond market) and El-Shagi (2016). Given the evidence that rating agencies are generally able to provide new information, it is hard to believe that the frequent delay in ratings is caused by incompetence or a lack of understanding of the rated markets. Rather, it seems that the rating agencies are often fairly accurate if they evaluate a country, but they often fail to do necessary evaluations in time. In other words, it seems as if rating agencies did a reasonable job once they decide to conduct a thorough analysis and provide a rating update, but very often they fail to get active in the first place.

So far, this aspect of sovereign creditor ratings has been underappreciated by the literature, partly because most of the literature aims to explain rating levels rather than rating decisions, and is thus unable to distinguish between deviations from the appropriate rating that are caused by misjudgment and those that are driven by lack of rating activity in general and the corresponding stickiness of ratings. Only few papers account for this stickiness. El-Shagi & von Schweinitz (2015) estimate an error correction type ordered probit model, explaining rating changes through lagged levels of ratings and interest rates, rather then estimating the long run relationship between yields and ratings directly. Even for fairly large deviations between the long-run equilibrium and the current situation, they find fairly low monthly adjustment probabilities. Dimitrakopoulos & Kolossiatis (2015) estimate a transition matrix between (clustered) rating levels allowing for different levels of stickiness at different rating levels. Yet, while those approaches account for the stickiness of ratings per se, they do not identify the sources of said stickiness.

In the present paper, the persistence of ratings, and specifically variations in the persistence, are explicitly modeled. In the spirit of the rational inattention (Sims 2003), we propose a simultaneous equation approach that separates (i) the rating agencies' decision whether or not to gather new information and update a rating from (ii) the decision how to update the rating.[1]

This approach does not only allow to explain, under which circumstances rating agencies fail to adjust a rating in time. It also allows uncovering the latent rating function of agencies, that is often obfuscated by mistaking the afore mentioned rational inattention for a deliberate decision to not adjust the rating based on available information.

We are able to identify several periods, where inattention prevented rating downgrades in countries where the reaction function would predict strong pressure to act. We find that rational inattention is not only driven by the time that passed since the last rating, but also by the rating class itself (i.e. that are ratings that can always be considered 'on probation' while others are very persistent), and income. Yet, despite controlling for those variables in the probit equation describing the decision to update a rating, we still find non-fundamental variables (such as lagged rating changes) in the ordered probit equation. This indicates, that the rating decision itself is subject to considerations other than economic conditions. In particular rating agencies seem to avoid large rating updates, often preferring a sequence of downgrades over a major single downgrade.

In a separate analysis, we look at the reactions of markets to rating changes. We find that markets consider deviations from the afore mentioned latent rating function as particularly surprising. Market uncertainty increases, and yields increase (decrease) strongly in reaction to downgrades (upgrades). Deviations from a rating function that is estimated without taking rational inattention into account lead to much weaker market reactions. This provides further evidence of the validity of our approach.

The remainder of the paper is structured as follows. Section 2 discusses the previous empirical findings and outlines our contribution over the previous literature. Section 3 introduces our method in more detail. In Section 4, we describe our data and some important stylized facts. Section 5 presents our results and Section 6 concludes.

# 2 Literature review

**Determinants of ratings** Over the past three decades, there has been an abundance of literature on ratings. Roughly 60 papers – to our knowledge – essentially deal with the same question, i.e. what is driving sovereign ratings. Over the time, plenty of potential indicators have been discussed, some being standard in the literature now, some still obscure footnotes in the literature. The literature is often traced back to the seminal paper by Cantor & Packer (1996), who were indeed the first to look at the ratings provided by major credit rating agencies. However, research on the creditworthiness (perception) actually traces back much further. Starting with Feder & Uy (1985), there have been a range of papers assessing the Institutional Investor ratings and Euromoney ratings (see e.g. Brewer & Rivoli (1990), Cosset & Roy (1991) and Lee (1993)).

The core set of variables used in the current literature is still the one that has been established in the seminal paper by Cantor & Packer (1996), that essentially looks at

[1] To our knowledge, we are the first to employ this approach. There is, however, a project that is simultaneously developed by Hantzsche (2015) using a midpoint inflated ordered probit method method. This approach assumes we cannot observes whether or not keeping a rating fixed is driven by the rating agencies decides to not rate a country, or if the rating agencies consider updating but intentionally decide to keep the original rating. Given the frequent updates of ratings with no change, we believe that they are indeed informative, and indicate that usually the decision to update is observable.

a combination of debt, the fiscal balance and a range of macroeconomic fundamentals, such as income per capita, inflation etc. A large number of additional indicators has been tested in later contributions for their potential impact on ratings. These extensions can be broadly grouped in two strands.

First, a fairly large range of papers has rediscovered the role of political and institutional factors for ratings. Those had already been covered in the early literature, and indeed been the focus of Brewer & Rivoli (1990), but was omitted in Cantor & Packer (1996). Depken, LaFountain, Butters et al. (2007) introduce corruption into the baseline model, which was found to be fairly successful and has been a staple variable in the later literature either as part of a wider index (Depken et al. 2007) or as a separate indicator (Amstad & Packer 2015). In a similar vein, Butler & Fauver (2006) institutional quality and legal origin as indicators of the soundness of institutions. Haque, Mark & Mathieson (1998) look at a wide range of indicators of political stability, such as coup d'etats, strikes, demonstrations, and Block & Vaaler (2004) consider the impact of elections.

Second, a lot of authors investigate split samples testing the assumption that different country groups are treated structurally different by rating agencies. Gültekin-Karakaş, Hisarcıklılar & Öztürk (2011) split their sample in emerging markets and developed economies, presenting a natural extension of previous papers that include a developed country dummy on top of income per capita measures. Butler & Fauver (2006) split their sample by the level of debt. Their findings are particularly interesting, since they suggest more than a mere club effect that might explain the different treatment of, say, OECD countries, but rather indicates actual nonlinearities in the rating process.

Our paper takes a fairly wide approach, including – where available – all drivers that have been identified robustly in the previous literature. We also account for different treatment of different country groups. However, rather than doing sample splits as most of the previous literature, we include a battery of interactions in our models.

The key difference between our paper and the majority of the previous literature is that we explicitly account for the dynamics and persistence in rating decisions.

**Ratings, persistence and timing** The key criticism concerning rating agencies is the timeliness of their ratings and their dynamic interaction of rating changes with the macroeconomy. Yet, the vast majority of papers digging deeper into determinants of ratings study rating levels (rather than changes) in a cross section of countries. To name just a few important contributions, Cantor & Packer (1996), Afonso (2003) and Amstad & Packer (2015) follow this approach. Even where panel data is utilized, empirical strategies often aim to explain long-run rating levels in a nondynamic framework, (see for example Ferri et al. 1999, Depken et al. 2007). However, neither of these strands of literature account for the strong dynamic aspect of ratings, i.e., persistence and speed of adjustment (as mentioned by Mora 2006, El-Shagi 2010, El-Shagi & von Schweinitz 2015).

Yet, there are notable exceptions. While not standard in the literature, there is a range of papers controlling for lagged rating levels. This implicitly accounts for persistence, although by no means explaining it (see e.g. Haque, Kumar, Mark & Mathieson 1996, Haque et al. 1998, Mulder & Monfort 2000). Some, such as Al-Sakka & ap Gwilym (2009), and estimate models in first differences (accounting for persistence by construction), where they explicitly account for momentum in changes. Hu et al. (2002) estimate transition matrices, augmenting the simple autoregressive models of ratings, and thus accounting for heterogeneity in rating persistence in some more details. Schumacher (2014) and El-Shagi & von Schweinitz (2015) estimate VAR models that jointly consider the macroeconomy,

thereby also shedding more light on the dynamic aspects of rating decisions.
However, very few papers explicitly discuss persistence in depths, namely Dimitrakopoulos & Kolossiatis (2015) and Hantzsche (2015). While the former estimates higher order AR models, the latter is probably the one closest to us as it is to our knowledge the only other paper that explicitly tries to explain persistence, rather than just taking persistence as something that exists and is constant over time. Contrary to the estimation by Hantzsche (2015), our paper exploits information on rating announcements whether or not the rating is actually changed, uses a much wider sample, and accounts for more indicators and potential nonlinearities. In particular, we differentiate between persistence (i) due to fundamentals (ii) from smoothing and staggered adjustment, and (iii) from rational inattention.

# 3 Method

Rating agencies – as argued in the literature review above – do not necessarily reevaluate rating decisions continuously. In the majority of periods, the probability of coming to a new rating conclusion is insufficient to justify the cognitive and informational costs of a full reassessment (Sims 2003). Instead, there may be long periods of time where agencies do not even consider a reevaluation. That is, rating agencies face two decision problems at every point in time. First, they need to decide if a rating should be reevaluated or not. Second and only in case of reevaluation, a new rating level needs to be determined and announced. Ideally, periods of no reevaluation should coincide with periods of comparable stability in fundamental variables such that the last rating decision proves to be correct even though it is not confirmed. As long as there is no reason to change a rating, a costly reevaluation would generate no benefit. However, it may well be that there are several other determinants driving the first decision (to reevaluate or not) that are unrelated to the development of fundamentals.
Statistically, our approach is very similar to a Heckman selection model, where we assess the direction of change in a limited dataset of observations where the rating has been assessed, and a "selection" equation determining when a country will be evaluated. Theoretically, those equations can be substituted in one another, to compute the total effect of various indicators on the probability for upgrades and downgrades. Yet, contrary to selection models we are not only after estimating in the joint effect, but actually interested in the individual equations, because both have an interesting story to tell about how rating dynamics work. Economically, we are thus much closer related to questions that have been assessed by midpoint inflated ordered probit (MIOP) models, where an ordered variable (such as rating changes) is modeled using two equations (Brooks, Harris, Spencer et al. 2007, Bagozzi & Mukherjee 2012).[2] Again, one equation describes whether change is even considered while a second one tracks the direction of change. Those models do, however, assume that the data does not allow to distinguish whether there is the deliberate decision to not change the variable of interest, or if change has not even been considered. The most prominent macroeconomic application of this method has probably been interest rate setting, but it has also been applied to sovereign bond ratings. The key difference of our case to the MIOP scenario is that rating agencies frequently publish rating announcements without changing the rating: 17% of our observations contain an-

[2]The MIOP was developed based on the zero-inflated ordered probit of Harris & Zhao (2007), which inflates the lowest instead of the middle category.

nouncements, of which only 20% (i.e., 3.5% of total observations) are rating changes. Due to the large difference between announcements and rating changes, and the missing incentive for rating agencies to perform costly reevaluation exercises without sending a public signal about its actions, we can safely assume that we do indeed have the information on rating assessments.

In order to determine the drivers of the two decision problems of rating agencies, we model rating changes $y$ as a combination of two processes $y^d$ and $\tilde{y}$. The first process $y^d$ describes the decision reevaluate a rating. We assume that every reevalation is followed by an announcement of the rating agency, such that our announcement variable gives us full knowledge about the reevaluation decision. The second process is the direction of rating changes $\tilde{y}$ in case of reevaluation. There are three categories of rating changes, *downgrade* (-1), *no change* (0), or *upgrade* (+1), which can only be observed in periods where an actual reevaluation takes place. That is, only reevaluation periods ($y^d = 1$) are informative on the influence of explanatory variables on the direction of rating changes.

We model the reevaluation decision $y^d$ by a probit model with explanatory variables $X$ (including an intercept):

$$\begin{aligned} P(y^d = 0|X) &= 1 - \Phi(X\beta) \\ P(y^d = 1|X) &= \Phi(X\beta) \end{aligned}$$

The directional decision $\tilde{y}$ in case of a reevaluation is given by an ordered probit model with explanatory variables $Z$ (including an intercept) and a positive threshold $\mu$. We further account for the bounded nature of rating levels. For the highest (lowest) rating classes, further upgrades (downgrades) are impossible and should therefore have a probability of zero. Consider a country with a AAA rating. An upgrade beyond the current rating would not be possible even if there would be overwhelming reason to upgrade this country further. Following (Hantzsche 2015), the necessary adjustment can be introduced by adding upgrade (downgrade) probabilities for boundary observations to the probability of no change. Modelling the actual outcome of no change for these countries more precisely avoids an estimation bias as documented by (Hantzsche 2015). Denoting with $r$ the rating level, we introduce two dummy variables that are one for observations with rating levels at the boundary, $D^{AAA} = 1_{r=AAA}$ and $D^D = 1_{r=D}$. Using $1-\Phi(x) = \Phi(-x)$, we thus model the directional decision $\tilde{y}$ as follows:

$$\begin{aligned} P(\tilde{y} = -1|Z) &= (1 - D^D)\Phi(-Z\gamma) \\ P(\tilde{y} = 0|Z) &= (\Phi(Z\gamma) - \Phi(Z\gamma - \mu)) + \left(D^D\Phi(-Z\gamma) + D^{AAA}\Phi(Z\gamma - \mu)\right) \\ P(\tilde{y} = 1|Z) &= (1 - D^{AAA})\Phi(Z\gamma - \mu) \end{aligned}$$

These two models are combined to model the observed rating decision $y$, which now depends both on $X$ (the determinants of rating reevaluations) and $Z$ (the determinants of rating decisions in case of reevaluation). In case of uncorrelated errors, the model is fairly simple:

$$\begin{aligned} P(y = -1|X, Z) &= P(y^d = 1|X)P(\tilde{y} = -1|Z) \\ P(y = 0|X, Z) &= P(y^d = 0|X) + P(y^d = 1|X)P(\tilde{y} = 0|Z) \\ P(y = 1|X, Z) &= P(y^d = 1|X)P(\tilde{y} = 1|Z) \end{aligned} \tag{1}$$

The probability of a rating *downgrade* is the joint probability of a rating reevaluation ($X\beta > 0$) and a downgrade decision in case of reevaluation ($Z\gamma \leq 0$). Similarly, the

probability of a rating *upgrade* is the joint probability of a rating reevaluation and an upgrade decision in case of reevaluation ($Z\gamma > \mu$). The probability of *no change* is then the remaining probability. It needs to be noted, that the probability of *no change* combines two distinct observable cases, i.e. the case of no reevaluation ($Pr(y^d = 0|X)$) and the case of reevaluation with no rating change ($Pr(y^d = 1|X)Pr(\tilde{y} = 0|Z)$).
That is, we have de facto four different states jointly determined by $y^d$ and $\tilde{y}$. The coefficients of our model ($\hat{\beta}', \hat{\gamma}', \hat{\mu}, \hat{\rho}$) can be easily determined by maximization of the log-likelihood over those four states:

$$\begin{aligned}
\max_{(\hat{\beta}', \hat{\gamma}', \hat{\mu})} LL(y|X,Z) &= L(y^d, \tilde{y}|X,Z) \\
&= \sum_{n=1}^{N} 1_{y^d=0} \ln(P(y^d = 0|X)) \\
&\quad + 1_{y^d=1,\tilde{y}=-1} \ln(P(y^d = 1|X)P(\tilde{y} = -1|Z)) \\
&\quad + 1_{y^d=1,\tilde{y}=0} \ln(P(y^d = 1|X)P(\tilde{y} = 0|Z)) \\
&\quad + 1_{y^d=1,\tilde{y}=1} \ln(P(y^d = 1|X)P(\tilde{y} = 1|Z)).
\end{aligned}$$

The threshold $\mu$ needs to be positive. We enforce this by using the transformation $\tilde{\mu} = \ln(\mu)$ in the maximum-likelihood estimation.
As a robustness check, we will also allow for correlated errors. In this case, the probability of an observed rating decision $y$ can be derived from a bivariate normal distribution, which we denote by $\Phi_2$ with correlation parameter $\rho$:[3]

$$\begin{aligned}
Pr(y = -1|X,Z) &= (1 - D^D)\Phi_2(X\beta, -Z\gamma, -\rho) \\
Pr(y = 0|X,Z) &= 1 - \Phi(X\beta) + (\Phi_2(X\beta, Z\gamma, \rho) - \Phi_2(X\beta, Z\gamma - \mu, \rho)) \\
&\quad + \left(D^D\Phi_2(X\beta, -Z\gamma, -\rho) + D^{AAA}\Phi_2(X\beta, Z\gamma - \mu, \rho)\right) \\
Pr(y = 1|X,Z) &= (1 - D^{AAA})\Phi_2(X\beta, Z\gamma - \mu, \rho).
\end{aligned}$$

Similar to the above, we use Fishers z-transformation $\tilde{\rho} = 0.5 \ln\left(\frac{1+\rho}{1-\rho}\right)$ to incorporate the condition $|\rho| \leq 1$. The transformed $\tilde{\rho}$ is again normally distributed with standard confidence bands.

# 4 Ratings and their determinants

Our analysis encompasses three different groups of variables: (a) announcements by rating agencies and non-fundamental variables derived from this information (most importantly, rating changes $\tilde{y}$ and reevaluation decisions $y^d$); (b) fundamental variables related to government default probabilities; (c) economic development and political variables (the latter used only in a robustness check). In order to be able to describe the rating decision process at a more granular level, we will work with monthly data. This, however, forces us to interpolate some of the fundamental variables, especially for developing economies where the availability of good data on a monthly frequency is scarce, as also explained in Subsection 4.4.

[3]Please note that the correlation parameter $\rho$ changes sign if one of the two variables in the bivariate normal distribution (i.e., $Z\gamma$) is multiplied by $-1$.

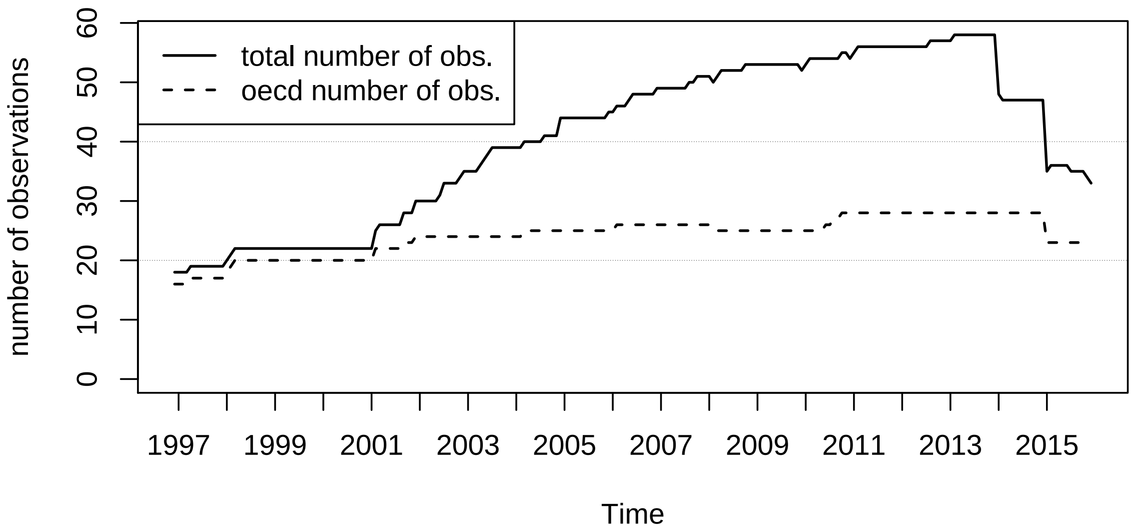


Figure 1: Number of observations over time

### 4.1 Rating and default data

**Ratings** Rating levels and announcement dates are drawn from the website http://www.countryeconomy.com. The website collects rating data for the three big rating agencies: Moody's, Standard & Poors and Fitch. Data on foreign-currency denominated loans currently span 138 countries partly going back as far as 1974.[4] As a rating stays constant from one announcement to the next, availability of rating data per country is solely determined by the first (reported) announcement. Figure 1 shows the number of countries for which rating data are available at each date. For the first part of the sample, our rating information is concentrated on OECD countries. Later, more and more countries are rated and data availability increases. In the last part of our data, certain explanatory variables are unavailable in some countries, reducing the scope of our analysis slightly.

As in the previous literature, we code rating levels on a discrete scale where 24 is a AAA rating and where 0-3 denotes different default ratings. We then average over the three large rating agencies (El-Shagi & von Schweinitz 2015). In contrast to the wide literature on ratings that empirically assesses rating levels in the cross section, our panel approach requires to look at rating changes. While there have been incidences when ratings were changed by several agencies at once, or even by several notches by an individual agencies, those instances do not provide a clear enough picture to be empirically exploited. We therefore opt for a simple ternary indicator of change as our dependent variable, only distinguishing upgrades (1), downgrades (-1) or unchanged ratings (0) within the current month. Yet, the rating level itself (*rating*) and its square (*rating.sq*) are used as explanatory variable, giving some error correction interpretation to the model, thus allowing to draw level conclusions from a first difference model. Figure 2 shows the share of

[4] However, due to lower availability of explanatory variables, we are only able to work with data from 56 countries, starting in December 1996.

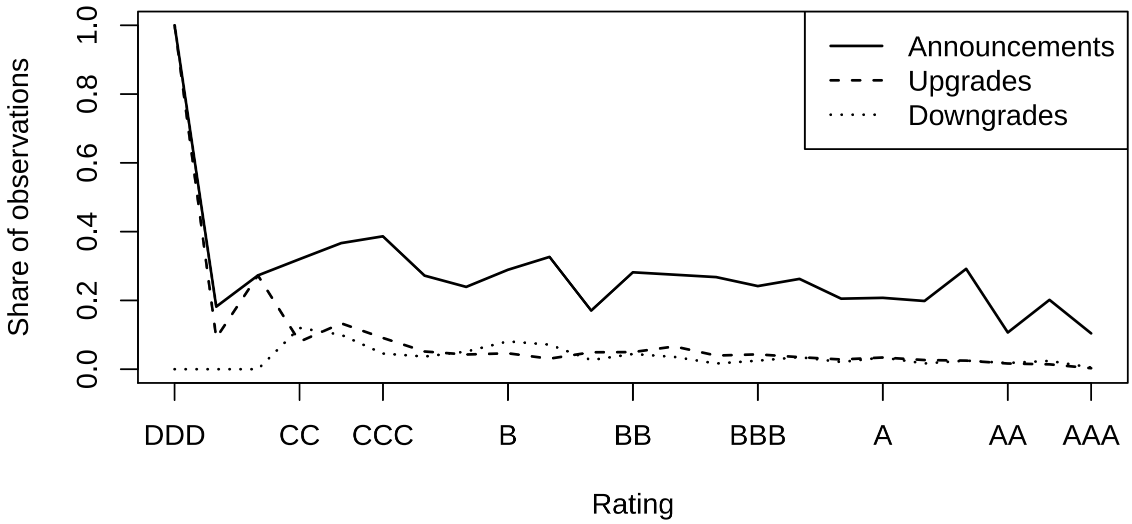


Figure 2: Rating announcements and changes for different rating classes

observations where there was an announcement or a rating change. While the frequency of announcements decreases only slightly for higher ratings, the share of rating changes drops towards zero very quickly.

**Time and (rational) inattention** The key innovation of this paper is to model ratings in a form that explicitly accounts for the often criticized – but possibly rational – inattention rating agencies seem to show. The most obvious reason for rational inattention is that the fundamental reasons behind a rating change slowly. That is, unless there are specific reasons to look at a country at a certain time, a country will be only screened occasionally. The most obvious way to implement this in a model is looking at the passing of time itself. In our model we include both time since the last rating (*years*) and its square (*years.sq*) to model this rough regularity. The reason for this nonlinear specification is evidence for staggering adjustment of ratings, both within and between the three rating agencies. In the absence of staggering adjustments, one would expect to see very low readjustment probability right after a rating has been set to its new, in the eyes of the rater "correct" rating, that then increases over time. However, due to the documented persistence of ratings an adjustment to the new level does not happen in one step. We therefore expect clustered rating adjustments, i.e. high adjustment probabilities right after an initial change, followed by a drop when the new appropriate level is reached. From this trough we would then expect to see the aforementioned rise in rating adjustment probabilities as time goes by, and the information underlying the last rating becomes more outdated. Yet, for a richer modeling of the dynamics during a staggering adjustment, we also include the total number of rating upgrades and rating downgrades in a country within the past 12 months (*Up12* and *Down12*).

**Other indicators of rational inattention based on rating dates** Outdatedness is probably not the only factor that triggers attention. We expect rating agencies to be more careful when they receive news that indicate the necessity of deeper investigation. In our paper this possibility is captured in two ways. First, analogue to the rating changes in a country in the past 12 months, we use the shares of countries with rating upgrades and rating downgrades in other countries within the past 12 month (*UpAll12* and *DownAll12*).[5] This captures both the possibility of spillovers in the sense that downgrades cause downgrades in other countries where the fundamentals might not justify those (as discussed in the contagion literature, see e.g. Forbes & Rigobon 2001), and the possibility of raised awareness after having to rerate several countries (if the general probability of change is affected by those indicators). Second, we generate an indicator of fundamental change since the last rating announcement at a country level (*changefund*). This indicator is based on all fundamental variables except GNI per capita, as discussed in the next subsection. It is the squared average windsorized change since the last rating announcement, where change for each variable is defined in relation to its average monthly change, and is given by:

$$fc(c,t) = \left[ k^{-1} w \left( \sum_{j=1}^{k} \frac{(I_{j,c,t} - I_{j,c,s})/I_{j,c,s}}{N^{-1} \sum_{c=1}^{C} \sum_{\tau=1}^{T} \Delta I_{j,c,\tau}/I_{j,c,\tau-1}} \right) \right]^2, \tag{2}$$

where $j \in \{1, \dots, k\}$ is an index of fundamental indicators, $I_{j,t,c}$ correspondingly is the $j^{th}$ indicator in country $c$ at time $t$, $N$ the number of observations and $w$ a windsorizing function that windsorizes at the $5^{th}$ and $95^{th}$ percentiles respectively. $s$ is the time of the last rating change. Thus, $(I_{j,c,t} - I_{j,c,s})/I_{j,c,s}$ denotes the percentage change of indicator $j$ in country $c$ since the last rating announcement.

**Default history** In the cross sectional literature it has been established that the default history of a country has a major impact on ratings. Rather than using the default history as implied by ratings (i.e. a rating of D), we use the Bank of Canada Database on Sovereign Defaults (see Beers & Nadeau (2014)). This allows to track defaults further back than the available ratings.

We tested two different indicators. First, motivated by the cross sectional literature, we use a dummy indicating whether a country has ever defaulted in the past (*default*). In an alternative specification, we use an indicator exploiting the time dimension, indicating whether or not a country had debt in default during the past 10 years (*recentdefault*). While our core results remain robust, we lose explanatory power compared to the baseline, indicating that rating agencies are surprisingly unforgiving when it comes to default. While it has often been mentioned that countries can return to the capital markets surprisingly quick after a default (see e.g. the survey article by Panizza, Sturzenegger & Zettelmeyer (2009)), it seems indeed as if something always "sticks", stigmatizing defaulting countries over extremely long periods.

## 4.2 Fundamental variables

Our paper includes a range of fundamentals, mostly suggested by the previous literature, that can be roughly divided into the subsets macroeconomic risk, development and macroeconomic performance and policy.

[5] We use the share rather than the total number to account for changes in the cross-sectional dimension over time.

**Macroeonomic risk** Following the bulk of the literature we consider both the debt to GDP ratio (*debt*) and the fiscal balance (*fiscbalance*) (again relative to GDP) as indicators of fiscal risk. While theoretically appealing to cover fiscal sustainability, the ratio of interest rate payments to revenues is only available for a very limited subset of countries. In particular, there are many gaps in the data, rendering a dynamic approach like ours almost impossible. We do, however, include real government bond yields (*yield*) to have a proxy of the market assessment of debt sustainability. To cover external risks, we focus on the current account balance (*current*). Finally, we include the growth rate of central bank reserves (*reserves*), to capture the possibilities that the reserves of a country can counteract potential debt or crisis problems.

**Development** We capture development in a range of ways. First, we use GNI per capita (*gnipc*). Also adding squared GNI per capita (*gnipc.sq*) allows us to pick up possible effects of a middle-income trap, where higher income comes with risk of substantial growth slowdowns warranting rating changes (Eichengreen, Park & Shin 2013). To overcome stationarity issues, we measure GNI per capita relative to the US which has been considered the primary economic force on the globe throughout our sample. Second, we measure institutional development using the corruption perception index (*corrupt*).

**Macroeconomic performance and policy** Our dataset covers industrial production growth (*ip*), inflation (*inf*), real effective exchange rate growth (*reer*) as indicator of competitiveness, and real money growth (*money*). These variables are more short-run in nature than the macroeconomic risk variables. However, large and persistent deviations from the level that is perceived as consistent with their rating should still be taken into account by rating agencies.

**Nonlinearities based on economic development** The previous literature has shown some evidence that rich countries are treated differently. We address this possibility in three different ways by testing interactions of our fundamentals with (a) a dummy variable indicating OECD membership (*dumoecd*), (b) a dummy variable indicating European Union membership (*dumeu*), and (c) *gnipc*.

## 4.3 Political variables

In a robustness test, we also include variables from the database of political institutions provided by the World Bank. The data we include is mostly meant to capture two aspects of the political system that matter for risk and (strategic) risk assessment. First, we want to measure political stability, second, we want to address strategical issues in the timing of rating announcements around election. Due to the availability of election dates, the database automatically allows to construct monthly data.

The first indicator we use is whether or not the government has a parliamentary majority (*majority*), thereby accounting for situations as recently seen in the US where the government had to shut down for some time in October 2013, greatly increasing the perception of risk concerning American debt.

Second, we include years in office (*yroffice*). While a long time in office might represent stability, extremely long times in office might rather indicate authoritarian regimes. To avoid mixing these two effects, we also include the square of this indicator (*yroffice.sq*).

Third, we include dummies indicating the 12 months before and (in a separate dummy) the 12 months after an election for the legislative (*legelecpre, legelecpost*) or executive (*exelecpre, exelecpost*) branch of government.

### 4.4 Data treatment

Quite a few of our fundamentals are not available at a monthly frequency, but only quarterly or even annually, see Table 6 in the Appendix. Yet, we want to avoid losing too many indicators that have been identified as important in the cross-sectional literature. Rather than dropping low frequency variables, we thus perform cubic interpolation of quarterly and yearly data to monthly frequency. Although not perfect, this approach can be justified with two arguments. Low-frequency fundamental data display a high persistence. The slow changes should affect the decisions of rating agencies, which claim to "rate through the cycle". Moreover, interpolation tries to mimic the fact that news on the fundamental development occur continuously.
Additionally, we collect data from many different sources and merge them afterward. Mostly, this just means that data for different countries is obtained from different sources with comparable data definitions. In some cases, however, we have overlapping data from several sources for individual countries. In this case we try to extend the data of the longest or most suitable available series and adjust data from other sources based on a simple bivariate regression on the overlapping sample. In order to assure that series merged by regression are indeed consistent, we add the two following steps: (a) regression is restricted to series with a correlation of at least 90% in the overlap window; (b) the intercept of the regression is adjusted such that there is no break in the merged series at the date of merging.
All data is deseasonalized using X13-ARIMA-SEATS. Growth rates are always computed as year over year growth rates. To avoid that the interpolation plays too big a role, we always lag the variables by one unit of their original frequency. Details on data treatment are provided in Table 6, and summary statistcs are given in Table 8 in the Appendix.

## 5 Results

### 5.1 The baseline model

The following discussion will be separated into a discussion of the selection part of the model, and the rating setting part of the model, before we move to joint (marginal) effects in the next subsection. Where we feel that a result can only be fully appreciated in the joint effect, we will omit the discussion in the following paragraphs on selection and rating adjustment.

**Is modeling inattention important?** Our econometric model, combining a probit model for the decision to reevaluate and an ordered probit model for the evaluation decision, allows for much richer dynamics than a simple ordered probit model that only takes observed rating changes into account. Our baseline model clearly and significantly outperforms a simple ordered probit model (see Table 1 in row *baseline oprob*). Even just on its own, this lends some support to our hypothesis of – possibly rational – inattention of rating agencies. Moreover, we find that the indicators of rational inattention jointly play a significant role in explaining rating behavior. Not only does a two equation model with those indicators outperform a model that only includes fundamentals (see row *fundamentals*), but even in the simple ordered probit they add some explanatory power.[6] That is, we find substantial evidence that rather than providing a continuous flow of new information, rating agencies become active, when they feel the need to, which is not necessarily reflecting a change in actual risk.

Table 1: Model comparison

| | # eq | LL | LL(oprob) | # coeffs | p(LLR-test) | against |
|---|---|---|---|---|---|---|
| *baseline* | 2 | -4813.06 | -1760.47 | 45 | | |
| *baseline oprob* | 1 | | -1798.51 | 23 | $<$0.0001 | baseline |
| *fundamentals* | 2 | -5473.80 | -1854.00 | 23 | $<$0.0001 | baseline |
| *fund. oprob* | 1 | | -1903.36 | 10 | $<$0.0001 | fundamentals |

*Note:* To make the models comparable, we compute an ordered probit equivalent log likelihood for our two equation model in the column *LL(oprob)*, where we only consider the implied three total probabilities from equation (1). That is, we use only the three cases entering the simple ordered probit model (upgrades, downgrades and no adjustment), adding up the probabilities for no announcement and an announcement without rating change.

**Times of inattention** In particular the coefficients on *years* and on its square *years.sq* given in Table 2 indicate the expected u-shape.[7] Rating announcements do indeed occur clustered, but once things have stabilized they tend to remain constant for some time, before the probability for a new rating evaluation is increasing again (as the square term starts to dominate the marginal effect of time passed). This is fully in-line with the idea that enough new information needs to accumulate until a costly evaluation of this information is sensible. *Up12* partly compensates the initial clustering effect when upgrades are concerned, i.e. a country will be subject to more scrutiny after downgrades. This is consistent with the evidence for stronger staggering of downgrades than upgrades we describe later. Like the passing of time after the initial months, the total structural change since the last rating change – that is of course correlated with *years* – increases the probability of reassessment, much as expected. As countries with a high rating are reevaluated less often (as seen in the negative coefficients of *rating* and *rating.sq*), more time can pass between ratings and more structural change can accumulate. Yet, the magnitude of the coefficients of *years*, *years.sq*, and *changefund* indicates that it takes a long time and substantial structural change since the last rating to compensate the strong negative impact of *rating* on reevaluation probability.
This still holds despite the fairly surprising result that high income (*gnipc*) makes new rating evaluations more likely.[8] This positive coefficient is not completely implausible. Given the same conditions, a richer and more developed country is typically more relevant

[6] This is implied by the significant rejection of the LLR-test of *fund. oprob* against *fundamentals*, combined with the fact that the likelihood of *baseline oprob* is larger than that of *fundamentals* at the same number of coefficients.

[7] We demeaned *years* such that a value of zero on *years* and *years.sq* is at roughly 21 months.

[8] Technically, we find a positive effect of *gnipc* and a negative effect of *gnipc.sq*, i.e. a hump shape. Yet, the hump peaks for extremely rich and developed countries at a GNI per capita of roughly 30% more than the US level. The probability of reassessment only decreases marginally for even higher levels of *gnipc*.

for investors, and thus more relevant to rating agencies. In total, this confirms our original hypotheses.
A similar argument for rational inattention can be made concerning the level of debt. Higher debt levels mean that ratings cover significantly more assets. This, in turn, should lead to increased monitoring and more frequent reevaluations.

**Setting ratings** In the ordered probit equation, it is interesting to see that the inattention variables often also drive the direction of the rating decision.[9]
Unsurprisingly countries with high ratings rather tend to be downgraded, while countries with low ratings tend to be upgraded. Notably, this is not a consequence of the fact that an AAA-rating cannot be further upgraded, as the model explicitly corrects for this bias. Interestingly, waves of up- and downgrades in other countries do not influence the direction of reevaluation (much), while own past rating changes have a very strong self-reinforcing effect, which is slightly higher for downgrades than for upgrades (although the difference is not statistically significant). The coefficients on *Up12* and *Down12* indeed indicate very strong staggering of rating adjustments towards the new level, consistent with both asynchronous adaptation across different rating agencies and smoothed adjustment within individual agencies. There is a slight tendency that the time since the last reevaluation makes upgrades more likely than downgrades, however, the difference is rather small. The negative effect on *default* confirms the stigma of past defaults that has previously been found in the literature. The same can be said of most fundamental variables: stronger growth, stronger fiscal and external balance and lower corruption induce agencies to upgrades (see e.g. Ferri et al. (1999), Mora (2006), Amstad & Packer (2015)). The strongest effects by far come from the fiscal and external balance, as these two indicators have immediate effects on medium-run indebtedness and are good indicators of fiscal and financial sustainability. Growing real effective exchange rates point to strong development (Balassa 1964, Samuelson 1964) and thus increase the probability of upgrades.

## 5.2 Looking at the equations simultaneously: Marginal and joint effects

Some indicators only reveal their full economic implications when simultaneously looking at both equations. We do this both by looking at the coefficients of both equations simultaneously, and by evaluating the total marginal effect of a change on upgrade and downgrades probabilities evaluated at the median for all variables, see Table 3. While reporting marginal results at the mean is more common, the mean is quite misleading for a few of our indicators, in particular dummies such as *default*. The nature of other indicators makes the actual marginal effects, i.e. the slope of the probability at the evaluation point, hard to interpret. This is particularly true for variables that move in discrete steps. Therefore, for those variables we technically do not report marginal effects, but probability differences when changing the variable by one step. For simplicity we will still refer to those as marginal effects for the remainder of this section. Variables concerned are *rating*, *Up12*, *Down12*, and *default* where we assess steps of one full rating level, and one rating change in the past 12 months respectively.

[9] This observation directly relates to the question if estimation errors in the two equations might be correlated. However, estimating a model that explicitly allows for correlated errors does not change estimation results. Indeed, the correlation coefficient is found to be insignificant.

Table 2: Estimation coefficients, baseline model

| | Reevaluation | Rating decision |
|---|---|---|
| rating | -1.357 *** | -2.418 *** |
| rating.sq | -2.353 *** | 1.019 |
| default | -0.097 * | -0.377 *** |
| UpAll12 | 0.540 *** | 0.356 |
| DownAll12 | 1.226 *** | -0.165 |
| Up12 | -0.435 *** | 1.167 *** |
| Down12 | 0.226 | -1.446 *** |
| years | -0.359 *** | 0.283 *** |
| years.sq | 0.066 *** | -0.062 *** |
| changefund | 0.106 *** | -0.168 ** |
| gnipc | 0.921 *** | -1.643 *** |
| gnipc.sq | -0.356 *** | 0.776 ** |
| ip | 0.014 | 0.198 *** |
| reserves | 0.021 | 0.095 |
| inf | -0.149 | -0.675 * |
| reer | -0.013 | 0.160 *** |
| yield | -0.119 *** | -0.042 |
| debt | 0.105 * | -0.065 |
| fiscbal | 0.139 | 6.292 *** |
| current | -0.079 | 3.088 *** |
| corrupt | -0.067 | 0.559 *** |
| Constant | -2.331 *** | 2.564 *** |
| Thresh 0.1 | - | 2.666 *** |
| LL | -4813.058 | |
| N | 9296 | |

**Ratings and default history** Most interestingly, changes of the rating itself have only little impact in the neighborhood of the median country which has a fairly high rating. Our model finds a trade-off between announcements and directions of rating evaluations. As mentioned above, rich and highly rated countries are in general reevaluated less often. However, in the rare cases they are reevaluated, they face a negative pressure on the rating, as indicated by the negative coefficients on *rating* and *gnipc*. At the median, a higher rating makes further upgrades a bit more unlikely as "the air gets thin" at the top. Yet, the increase of downgrade probabilities is inconsequentially small.
Also, despite the general evidence that we find for staggered downward adjustments, this does not happen close to the median. *Up12* and *Down12* merely reduce the probability for opposite movements, rather than increasing the probability for another step in the same direction significantly. I.e., there is neither bounceback nor staggering at median levels.
Looking at *UpAll12* and *DownAll12* in both equations strongly challenges any arguments on spillover effects of negative ratings, that have been suspected in the previous literature. Upgrades and in particular downgrades in other countries do not force a countries' rating in the same direction. Rather, countries all across the globe are subject to more attention by rating agencies when many ratings have been adjusted in either direction. For the median country, upgrades and downgrades in other countries increase the probability for both upgrades and downgrades considerably, with the impact on upgrades being more substantial. This indicates, that the waves of downgrades in the past have not been pure contagion, but have indeed been driven by a correlation of other (structural) indicators between the affected countries. For example, if there is some kind of business cycle correlation between countries, rating decisions based on fundamental values will go in the same direction. This alone might create situations that look like spillovers, when not separately modeling the decision to rate and the decision how to rate.
Default history increases the probability of downgrades in the ordered probit equation, while simultaneously reducing the probability of changes. This can contribute to a lock-in effect, making bad ratings very persistent, once markets had a truly negative experience (i.e., a default) with this country. When looking at marginal effects at the median, this is reflected in a significant negative impact on upgrade probabilities, and an insignificant increase in downgrade probabilities. Yet, is has to be considered that a country with a past default but a current rating of 20 (i.e. a country that is created by changing *default* away from its median) is fairly unrealistic in first place.

**Structural indicators** For most indicators the impact evaluated at the median is what theory suggests. There are, however, two fairly surprising results that stand out.
First, real yields (*yield*) are only found to be significant in the probit equation, where they enter with a negative coefficient. That is, yields are not necessarily needed to explain rating changes. Higher real yields (which are correlated to high volatilities, see Ball 1992) instead discourage rating agencies to reevaluate sovereign ratings, rather than leading to downgrades due to perceived increases in credit risk. This result is supported by previous findings that rating agencies – due to their tendency to rate "through the cycle" – usually do not react immediately to market movements (Ferri et al. 1999, El-Shagi 2010). Also, a large part of the effect of high yields might actually be an indicator of high profitability rather than risk premium after controlling for many of the structural risk factors behind high yields, such as debt and deficit.[10] The total marginal effect on

[10]See e.g. (Bernoth, Von Hagen & Schuknecht 2012) for a discussion of risk factors driving sovereign

both upgrades and downgrades is therefore significantly negative. It should, however, be kept in mind that the effect is economically small at least when evaluated at the median. Only the most extreme yield fluctuations would have a meaningful impact on up- and downgrade probabilities.

Second, we find that high *gnipc* creates some downward pressure. Since we control for ratings (both in levels and squares) it is unlikely that this effect merely reflects the fact that highly rated, usually rich countries have only one way to go. To some extent this might be explained by a middle income trap.[11] The median observation has a *gnipc* of about 50% of the US. For countries at this level of income, higher income might be perceived as being bought at the risk of an ongoing stagnation once the catching up comes to a sudden halt in the middle income region. But again, most importantly, while statistically significant, the effect is economically fairly irrelevant. After all, a change of one unit in our indicator (equal to changing GNI per capita by the entire level of US GNI per capita) would increase downgrade probabilities by a mere 2.8%.

All other fundamentals point in the expected direction. Growth of industrial production, as well as that of reserves and real effective exchange rates decreases downgrade probabilities. The same is true for improving fiscal and external balance. Also lower corruption (measured by an increase in the corruption index) lowers downgrade probabilities. Rising inflation, on the other hand, mainly works against upgrade probabilities, but is by itself apparently insufficient to create downgrade pressure.

Table 3: Marginal effects, baseline model

| | value | Downgrades | Upgrades | Decision down | Decision up | Reevaluation |
|---|---|---|---|---|---|---|
| rating | 19.500 | 0.002 * | -0.026 *** | 0.043 | -0.081 *** | -0.090 *** |
| UpAll12 | 0.020 | 0.015 *** | 0.999 ** | -1.051 *** | 2.264 * | 3.951 *** |
| DownAll12 | 0.016 | 0.435 ** | 0.873 ** | 0.377 ** | -0.806 | 6.932 *** |
| Up12 | 0.000 | -0.090 *** | 0.128 | -0.414 *** | 1.207 *** | -0.442 *** |
| Down12 | 0.000 | 0.036 | -0.049 *** | 0.175 | -0.308 *** | 0.060 * |
| years | 5.290 | -0.134 *** | -0.016 *** | -0.398 *** | 0.853 *** | -1.162 *** |
| changefund | -0.516 | 0.477 | -0.294 *** | 1.805 | -3.869 *** | 2.795 *** |
| gnipc | 0.423 | 0.028 ** | -0.017 *** | 0.106 | -0.224 *** | 0.164 *** |
| ip | 2.129 | -0.043 *** | 0.103 | -0.246 *** | 0.529 *** | 0.043 |
| reserves | 5.964 | -0.003 *** | 0.010 | -0.021 *** | 0.046 * | 0.012 |
| inf | 2.605 | 0.018 | -0.056 *** | 0.120 | -0.257 ** | -0.065 |
| reer | 0.289 | -0.033 *** | 0.062 | -0.172 *** | 0.370 *** | -0.034 |
| yield | 2.126 | -0.015 *** | -0.116 *** | 0.082 | -0.176 | -0.569 *** |
| debt | 48.894 | 0.003 | 0.001 | 0.007 | -0.015 | 0.028 ** |
| fiscbal | -2.411 | -0.121 *** | 0.271 | -0.673 *** | 1.456 *** | 0.037 |
| current | -0.408 | -0.062 *** | 0.127 | -0.331 *** | 0.713 *** | -0.021 |
| corrupt | 54.391 | -0.053 *** | 0.093 | -0.268 *** | 0.576 *** | -0.079 * |
| default | 0.000 | 0.730 | -1.472 *** | 5.438 | -7.033 *** | -2.447 ** |

**Nonlinearities** In particular for variables that enter our estimation in a nonlinear fashion, the impact on upgrade and downgrade probabilities can change drastically as the variable is changing. In this section, we present the probabilities for those variables across the full spectrum of possible values. In all the figures, we take an observation

bond yields.

[11] See e.g. Eichengreen et al. (2013) for a deeper discussion of this phenomenon.

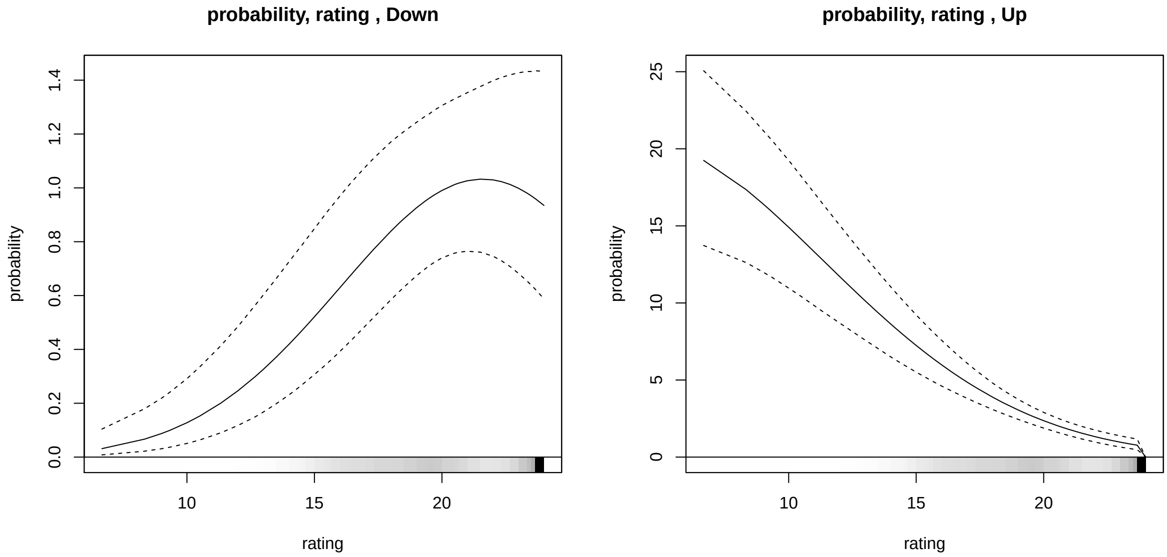


Figure 3: Downgrade and upgrade probabilities for different ratings

*Note:* The plot shows probability estimates (solid line) with 90% confidence bands (dashed lines). Darker colors in the gray bar at the bottom of the plot indicate a higher density of observations in the direct neighbourhood of median observations.

with median values of all variables and vary one variable, taking squares (and later interactions) into account.[12] These figures provide some additional insight into the basic results described with the help of coefficients and marginal effects.
The marginal effects in Table 3 indicate that rating increases at the median rating mostly reduce further upgrade probabilities. This is fully consistent with Figure 3, which shows the probability of downgrades and upgrades for all levels of *rating*. Downgrade and upgrade probabilities display a certain degree of mean reversion. However, the difference in scaling has to be strongly emphasized: Only at the very highest rating levels become upgrades less likely than downgrades. If a median observation would have extremely low rating levels, upgrade probabilities would be close to 20% (per month)! Figure 3 also displays the density of different ratings in our dataset in the form of a gray bar at the bottom of the figure, showing the familiar sight that around 25% of our observations have a rating of AAA, while we have very few observations with low rating levels.
For *gnipc*, Figure 4 shows that the surprising positive impact on downgrade probability seems to hold over most of the feasible values of GNI per capita. However, when looking closer, it becomes obvious that the confidence bounds explode once we approach the US level. That is, the effect seems to be mostly driven by the afore mentioned middle income group.
The hump-shape of *years* in the ordered probit equation implies higher downgrade probabilities directly and long after the last announcement, while upgrade probabilities are highest around 40 months after the last announcement. Figure 5 shows how this effect interacts with the u-shaped influence on the probability of a reevaluation of ratings. For downgrades, the u-shape is extremely pronounced, while it is much weakened for upgrade probabilities.

[12]In order to put the figures into perspective, it is helpful to remember that average upgrade and downgrade probabilities in our sample are 2.6% and 2.1%, respectively.

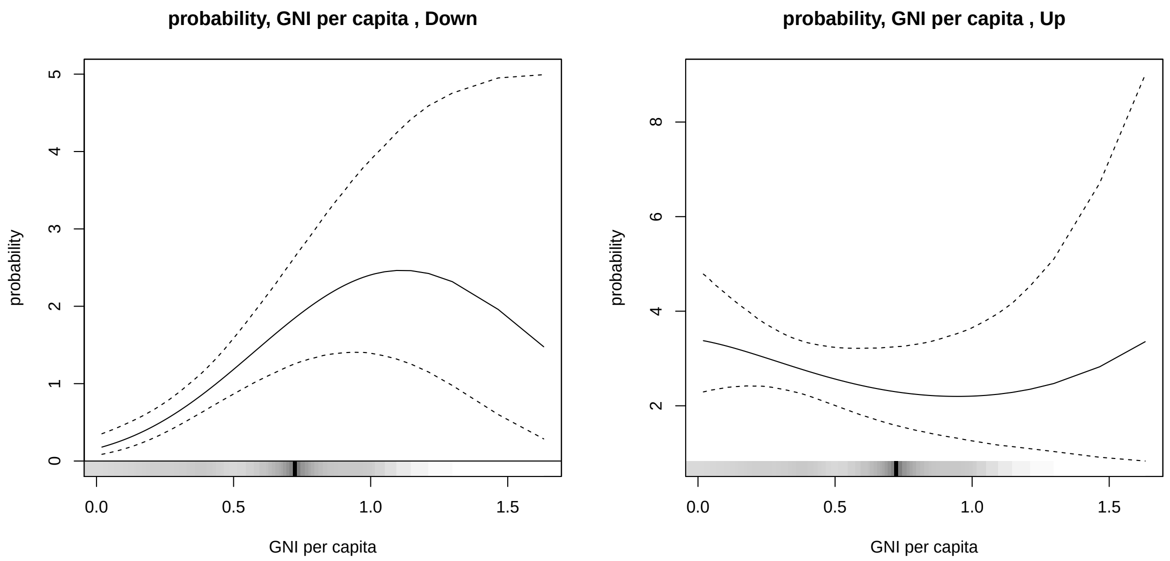


Figure 4: Downgrade and upgrade probabilities for different gnipc

*Note:* The plot shows probability estimates (solid line) with 90% confidence bands (dashed lines). Darker colors in the gray bar at the bottom of the plot indicate a higher density of observations in the direct neighbourhood of median observations.

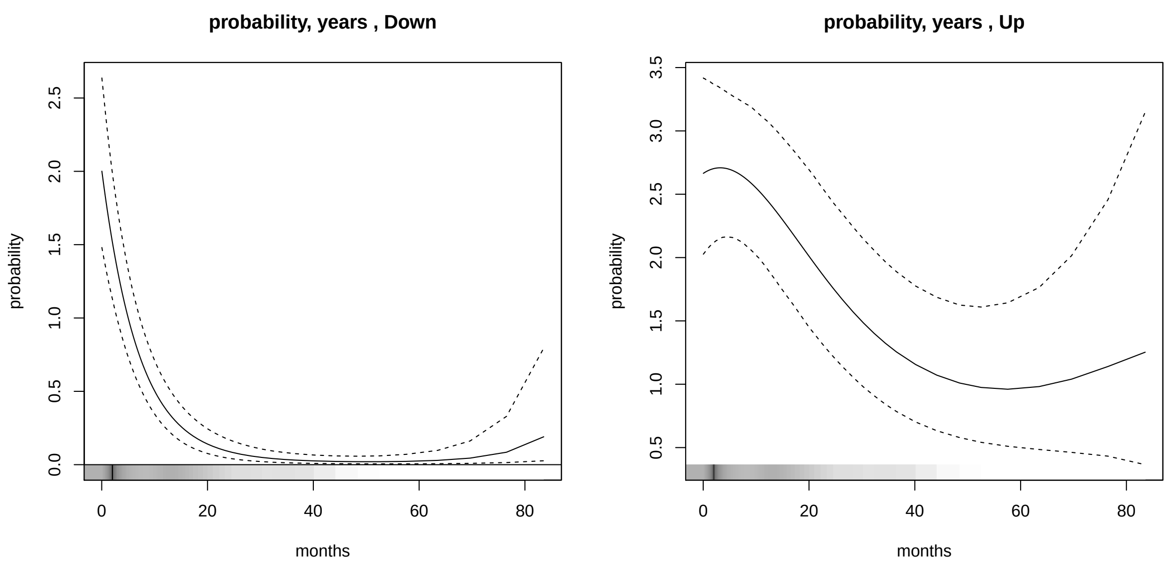


Figure 5: Downgrade and upgrade probabilities for time (in years) since last announcement

*Note:* The plot shows probability estimates (solid line) with 90% confidence bands (dashed lines). Darker colors in the gray bar at the bottom of the plot indicate a higher density of observations in the direct neighbourhood of median observations.

## 5.3 Rational inattention and market surprises

In this subsection, we will look at unusual periods in the sense that there was either an unusual absence of rating agency activity, or a surprise rating change. Both types of periods are interesting for us, because they show that our empirical strategy does not merely add explanatory power in a statistical sense, but also allows to deepen our economic understanding of rating mechanisms.

### 5.3.1 Periods of rational inattention

What makes our approach so attractive is that we can – to some extent – identify why countries were not reevaluated despite problems that seem obvious at least with the benefit of hindsight. Consider that our ordered probit equation implicitly contains all the information necessary to derive the current equilibrium rating, i.e. whether there is up- or downward pressure. Our probit equation explains persistence, very similar to the role of a (time-varying) autoregressive coefficient in monetary policy rules. In other words, we can identify episodes, where the pressure to upgrade or downgrade (indicated by the ordered probit equation) is ignored, because the current persistence (indicated by a low probability to reevaluate) is extremely high.
In this subsection, we look at long periods without rating changes, where there has been an unusually high downgrade or upgrade probability according to the ordered probit equation, paired with an unusually low reassessment probability.
We identify these periods in two steps. First, we regress the downgrade (upgrade) probability obtained from the ordered probit equation of our model, on the corresponding joint probability from the full model. I.e., we run the following two regressions:

$$\begin{aligned} P(\tilde{y} = -1|Z) &= \alpha + \beta P(y = -1|X, Z) + \varepsilon^{down} \\ P(\tilde{y} = 1|Z) &= \alpha + \beta P(y = 1|X, Z) + \varepsilon^{up} \end{aligned}$$

Unusually high residuals in this equation imply that the (downgrade or upgrade) probability estimated by the ordered probit component of the model is "dampened" unusually strongly by the probit component of the model that describes reassessment probabilities, and vice versa.
Second, we apply a simple filter to identify blocks of periods where the regression errors have been unusually large for a long time. More specifically, we focus on those periods where $\varepsilon^{down}$ ($\varepsilon^{up}$) has been consistently above its median for at least four years without announcement. In order to account for short-run fluctuations in probabilities, we smooth over single periods with smaller regression errors.
This process leaves us with six cases, where high change probabilities as indicated by the ordered probit model have been counteracted by unusually low announcement probabilities as indicated by the probit model. Four of these cases indicate a high downgrade probability with no action by rating agencies: Italy (Nov 2006–May 2011); Spain (Jan 2004–April 2010); the UK (June 2004–Sep 2010) and the US (Dec 2003–June 2011). The two remaining cases indicate high upgrade probabilities: Australia (Dec 2003–Jan 2008); Tunisia (Dec 2003–Nov 2008).
The cases of Italy and the US are special among these six cases as regression errors are (on average) among the 10% largest errors in the whole sample. Moreover, they are robust to other estimations described in later subsections. Thus, we consistently find that long

before the crisis hit, macroeconomic conditions in the US and Italy would have suggested a downgrade.
For the US our finding is very much in line with the frequently voiced fear that the US – where all of the big three rating agencies are based – are unduly favored by the agencies. This has even led to a political debate of the necessity of a European rating agency. Yet, our results indicate that the European Union might have received similarly favorable treatment. With Spain and Italy, both being periphery countries in the Euro area that turned out to be in worrying economic conditions during the crisis, and the UK included in our list, it is hard to argue for a purely American problem. Instead, it seems as if large highly developed countries generally receive the benefit of doubt.
While both periods of ignored upgrade pressure in Tunisia and Australia were ended by the Great Recession, which worsened economic conditions to an extent that the upgrade was no longer adequate, the periods where downgrade pressure was ignored were all ended by a rating announcement. However, only Spain and Italy[13] had to actually face a downgrade, while the ratings in the UK and the US were confirmed. This gives some indication, that Western countries do not merely benefit from less frequent ratings (that make it easier to go through rough times without facing downgrades), but are generally reviewed less critically.

### 5.3.2 Market reactions to surprise rating changes

The flipside of periods of rational inattention are periods with unforeseen rating changes. In the spirit of event studies, we investigate yield changes (measured in differences) during windows of different length around rating changes and compare them to yield changes in periods without rating announcements. We interpret unusually large market reactions as evidence of periods when rating changes provide news to the market.
We calculate yield differences during windows of different length around downgrades and upgrades (2 to 20 days, centered around the day of the rating change) and compare the distribution of these yield differences to the corresponding distribution of differences during windows around days without rating announcements (i.e., excluding also windows around announcements without changing ratings). Our estimations allow us to differentiate rating changes, and thus also market reactions around those changes, along two different dimensions: (i) upgrades and downgrades, and (ii) surprising and unsurprising changes. The results of these splits are shown graphically in Figures 8 and 9 in the Appendix.
In a first step, we look at upgrades and downgrades separately. Our findings confirm the results from the previous event study literature that on average yields decrease around upgrades and increase around downgrades. Furthermore, market reaction is on average stronger around downgrades than around upgrades. However, differences are in general not significant, which may be well explained by the fact that most rating changes can indeed by expected by markets.
Differentiating explicitly between surprising and predictable changes supports this hypothesis. Our definition of surprise essentially boils down to the question of predictability, which we approach from three different angles in terms of the underlying probability for a down- or upgrade: (i) the probability of our full model ($P(y = \cdot|X, Z)$), (ii) the proba-

[13] Technically, in Italy the period is ended by a rating announcements by Moody's stating that Italy is "under review" followed by downgrades by Moody's and Fitch a few months later, and S&P in January 2012.

bility from the ordered probit part of our model ($P(\tilde{y} = \cdot|Z)$), and (iii), for comparison, the probability from a simple ordered probit model that does not employ the information from announcements ($P(y = \cdot|Z)$). For all three, we interpret a change to be surprising if its probability is below 5%. A change is considered unsurprising if the probability is larger than 10%.

As would be expected, uncertainty around yield changes increases with window size. We can directly see from the Figures 8 and 9 in the Appendix that standard deviations for yield changes during windows of all rating changes are substantially larger than in normal periods. However, it seems to make a difference which probability is used to identify surprise changes. Our full model (i) and a simple ordered probit (iii) give very similar results, while changes unforeseen by the ordered probit part of our model (ii) seem to be particularly disrupting. That is, the true surprise for markets is a change unforeseen by case (ii), when there is no rating pressure in that direction.

This finding is particularly interesting, because it suggests that not all rating announcements provide the same level of information. Indeed, the part of the rating that is based on observables that are available to the market – as they are to us in estimating our econometric model – is not relevant for market participants per se. Markets seem to respond mostly to the change in fundamentals, whether or not the rating agencies adjust their rating. If the rating agencies eventually confirm this movement, it triggers only limited further reactions.[14] On the contrary, rating changes that are not in line with the usual rating procedures seem to be considered as new, previously unavailable information, which is reflected by a severe market response. Assuming efficient markets for government bond yields, where market participants exploit freely available information whether or not rating agencies have already responded, this is further evidence that our model provides a much better description of the actions of rating agencies than a simple ordered probit model.[15] Further evidence for this conclusion comes from the comparison of mean changes around surprising and unsurprising rating changes. Unsurprising rating changes only lead to slight increases (reductions) of yields around downgrades (upgrades). However, t-tests indicate that the means in both groups are mostly not significantly different. On the contrary, yields increase strongly and very significantly after surprise downgrades. The decrease after surprise upgrades is slightly smaller, but equally significant. Again, the most drastic changes occur in surprise periods identified by case (ii).

## 5.4 Rich vs. poor countries

The results of our baseline are based on pooled data for all countries and evaluated at the full sample median. Yet, even with this simplified model, we find some evidence that rating agencies might treat richer countries differently from poorer countries. To assess whether development is just one risk factor entering rating decisions, or whether rich countries are assessed in a different way compared to less developed coutries, we look at models where we interact the ratings and fundamentals with an indicator of eco-

[14]This might also be the underlying reason for the finding of Altdörfer, De las Salas, Guettler & Löffler (2016) that announcements by Fitch triggered no significant market reactions during the European sovereign debt crisis. One reason the authors provide is that Fitch acted during that time as a follower on the European rating market, i.e., it reacted slower to new information than the other two rating agencies.

[15]This result is not at all driven by the fact that in our identification scheme surprises under case (iii) are actually a subset of surprises under case (i). Results similarly hold if we look at probabilities below the 33% quantile (surprising changes) and above the 66% quantile (unsurprising changes).

nomic development. As indicators of development we alternatively consider the dummies indicating OECD and EU membership and *gnipc*. Generally, those interaction models outperform the simple baseline significantly according to a likelihood ratio test. We find that the model using *dumoecd* performs best, which is why the results presented in this chapter are based on this model. This gives some indication that the difference in treatment comes from a club effect and not development alone. However, the results with all three indicators are qualitatively very similar.[16]

There may be good reasons to differentiate between different levels of development, as higher developed countries usually experience more moderate, but also more stable development. Evidence of this can be found in the summary statistics in Table 8 in the Appendix, which also differentiates between OECD and non-OECD countries. The former not only have a higher *gnipc*, but also experience (on average) lower growth of industrial production, reserves and real effective exchange rates. They are able to sustain higher public debt levels as well as higher fiscal and external deficits. Despite this, OECD countries have on average lower inflation and lower real yields and less corruption. The transition from low to high levels of development might in turn lead to different assessments (and different assessment processes) by rating agencies.

**Interpretation of the interactions** The coefficients of the interacted variables are summarized in Table 4. Being an OECD country (as indicated by the coefficients on *dumoecd*) comes with fewer evaluation periods and higher upgrade probabilities.

One of the most interesting differences is seen in the impact of *changefund*. While we find an increase in the probability to be reassessed for both OECD and non OECD countries, the negative impact on the direction of change seems to come from non-OECD members only. That is, (fast) structural change seems to be considered as a risk for developing countries, but neutral when fully developed economies are concerned.

One thing that stands out is the positive coefficient on the growth rate of industrial production in the probit equation for OECD countries. Essentially good macroeconomic conditions reduce the general rigidity that ratings exhibit in well rated countries, and particularly so for OECD countries as shown by the interaction of the rating level and OECD membership. Since growth is pushing the rating up this implies that the upward rigidity is lower than the downward rigidity for OECD countries.

Inflation now has a clearly negative impact on ratings for less developed nations, but a positive one for OECD members. Yet, this last finding might be driven by decreasing inflation rates and even deflation during the Great Recession, where several OECD members faced downgrades. The same reason might be behind the negative coefficient on central bank reserves in OECD countries. Reserves usually compensate exchange rate risk in countries with pegged currencies. OECD countries, however, usually have very stable, yet free-floating, currencies, that often also serve as a safe haven. Thus, external stability does not depend strongly on central bank reserves. Instead, reserve growth in the industrial world was predominantly observed as a side-effect of monetary expansion during the Great Recession.

**Evaluating (marginal) effects for different income groups** Another way to look at different treatments is to assess the marginal effects and probabilities predicted by the

[16]Results from interaction models with *dumeu* and *gnipc* are shown in the appendix in Tables 9 and 10.

Table 4: Estimation coefficients, interaction model with dumoecd

| | Basic effects | | Interaction with dumoecd | |
|---|---|---|---|---|
| | Reevaluation | Rating decision | Reevaluation | Rating decision |
| rating | -0.634 * | -2.899 *** | -1.986 *** | 0.055 |
| rating.sq | -1.277 | 0.773 | 1.804 | 2.468 |
| default | -0.088 | -0.403 *** | | |
| UpAll12 | 0.482 *** | 0.328 | | |
| DownAll12 | 1.135 *** | -0.059 | | |
| Up12 | -0.418 *** | 0.924 *** | | |
| Down12 | 0.229 | -1.448 *** | | |
| years | -0.329 *** | 0.222 *** | | |
| years.sq | 0.060 *** | -0.062 *** | | |
| changefund | 0.119 *** | -0.219 *** | -0.051 | 0.254 ** |
| gnipc | 0.807 | -1.729 * | 0.231 | -0.419 |
| gnipc.sq | -0.213 | -0.077 | -0.183 | 1.177 |
| ip | -0.030 | 0.146 ** | 0.121 *** | 0.066 |
| reserves | -0.003 | 0.375 *** | 0.046 | -0.548 *** |
| inf | -0.129 | -0.867 ** | -2.058 ** | 3.971 ** |
| reer | -0.023 | 0.100 * | 0.028 | 0.145 |
| yield | -0.070 ** | -0.054 | -0.404 *** | 0.030 |
| debt | 0.114 | -0.072 | -0.095 | 0.030 |
| fiscbal | -0.385 | 7.711 *** | -0.454 | 0.219 |
| current | -0.383 | 4.185 *** | 0.041 | -1.014 |
| corrupt | -0.109 | 0.694 *** | 0.034 | -0.176 |
| dumoecd | -0.592 ** | 1.228 ** | | |
| Constant | -2.205 *** | 2.558 *** | | |
| Thresh 0.1 | - | 2.727 *** | | |
| LL | -4763.962 | | | |
| N | 9296 | | | |

*Note:* Coefficients on interaction terms given in the last two columns.

model when being evaluated under different conditions, i.e. computing them at subsample medians for OECD members and non members rather than the full sample median. Table 5 gives the marginal effects at median values conditional on *dumoecd*. The second and the fifth column now contain the values at which marginal effects are evaluated, while the other columns give the marginal effects on downgrade and upgrade probabilities for OECD countries and non-members.

Table 5: Marginal effects from the interaction model, separately for OECD and non-OECD members

| | dumoecd=0 | | | dumoecd=1 | | |
|---|---|---|---|---|---|---|
| | value | Downgrades | Upgrades | value | Downgrades | Upgrades |
| rating | 15.667 | 0.012 | -0.021 *** | 23.000 | -0.002 *** | -0.013 *** |
| default | 1.000 | -0.673 *** | 2.533 | 0.000 | 0.436 | -0.839 *** |
| UpAll12 | 0.020 | 0.009 *** | 0.793 ** | 0.020 | 0.023 *** | 0.546 ** |
| DownAll12 | 0.017 | 0.465 ** | 0.738 ** | 0.015 | 0.235 ** | 0.580 ** |
| Up12 | 0.000 | -0.101 *** | 0.092 | 0.000 | -0.044 *** | 0.041 |
| Down12 | 0.000 | 0.046 | -0.046 *** | 0.000 | 0.020 | -0.026 *** |
| years | 3.614 | -0.158 *** | -0.004 *** | 7.589 | -0.067 *** | -0.027 *** |
| changefund | -0.537 | 0.765 | -0.474 *** | -0.487 | 0.024 ** | 0.254 |
| gnipc | 0.109 | 0.059 * | -0.044 *** | 0.728 | 0.010 ** | -0.001 *** |
| ip | 2.431 | -0.049 *** | 0.055 | 1.954 | -0.017 *** | 0.083 |
| reserves | 8.692 | -0.020 *** | 0.031 | 4.423 | 0.005 | -0.006 *** |
| inf | 4.390 | 0.032 | -0.063 *** | 2.060 | -0.088 *** | 0.027 |
| reer | 0.281 | -0.030 *** | 0.032 | 0.290 | -0.026 *** | 0.058 |
| yield | 2.343 | 0.002 ** | -0.080 *** | 2.006 | -0.074 *** | -0.226 *** |
| debt | 40.107 | 0.004 | 0.001 * | 56.467 | 0.001 * | -0.001 *** |
| fiscbal | -2.046 | -0.200 *** | 0.293 | -2.638 | -0.093 *** | 0.163 |
| current | -0.553 | -0.112 *** | 0.153 | -0.309 | -0.037 *** | 0.065 |
| corrupt | 37.243 | -0.087 *** | 0.106 | 72.319 | -0.028 *** | 0.045 |
| dumoecd | 0.000 | -0.622 *** | 1.701 | 1.000 | 1.552 | -0.531 *** |

Some variables have considerably different marginal effects when evaluated for the typical OECD member and the typical (i.e. median) non-member respectively. Yet, we find only two examples where a coefficient is statistically significant with different signs at different points. We find this for *yield*, where increasing real yields increase downgrade (and decrease upgrade) probabilities for OECD countries, while it decreases both probabilities for other countries. Thus, the finding of the baseline model that higher real yields might reduce rating activity due to increased uncertainty is most likely driven by non-OECD countries. For OECD countries, the traditional credit risk channel is much more important. Another example is inflation. In non-OECD countries higher inflation seems to be a sign of instability, decreasing upgrade probabilities. In OECD countries, however, higher inflation (if close to the target of the majority of OECD-central banks) is a sign of strong growth, reducing downgrade probabilities instead.

In most cases the difference between OECD members and non-members is more subtle. What we find occasionally are situations where one group of countries might experience lower downgrade probabilities, while the other group would enjoy higher upgrade probabilities. This channel is (for example) at work for *changefund*. The observed increase in

downgrade probabilities for OECD countries in the case of increasing changes in fundamentals (rather than decreasing upgrade probabilities) is consistent with the high rating levels in OECD levels that induce low upgrade probabilities in general.
The clearest benefits of OECD membership in terms of ratings can be seen when looking at the impact of ratings over the entire spectrum of potential ratings, see also Figure 6. The downgrade probability of the typical OECD countries is negligible at all rating levels, while upgrade probabilities become extremely high for low rating levels. That is, rating agencies tend to assign low ratings to OECD countries only on a rather temporary basis, while the same rating levels can be very persistent for non-OECD countries. This insight is further strengthened if we take a look at announcement probabilities. While they vary only slightly (with nearly insignificant differences) over rating classes for non-OECD countries, OECD countries get more and more announcements as they have lower ratings. That is, they are under more or less constant supervision if ratings drop below BBB-. However, there is still a large degree of uncertainty on the exact effect, as only 5% of observations from OECD countries have such a rating.
As far as *gnipc* is concerned, the middle income trap hypothesis still seems valid (Figure 7). While the downgrade probability seems to be increasing over all income levels even for non members, the estimation has actually little to say about the very odd construct of a rich non-OECD country. Indeed only Hong Kong and Singapore in our sample have a GNI per capita relative to the US exceeding 0.7. Correspondingly, we literally find an explosion of the confidence bounds once we approach the US level of GNI per capita.[17]
While the shape of the function looks similar at a first glance for OECD members, the effect of income is comparably very low, and economically inconsequential.

## 5.5 The influence of politics

We estimate two new models both including years in office (and its square) and the government majority dummy, with the first model accounting for the time just before and after election of the legislative branch, and another one accounting for the time just before and after elections of the executive branch. We separate those models because joint elections – the usual practice in many countries – would induce multicollinearity problems. Lower data availability for these variables, however, reduces our sample size by about 1'000 observations.
Our baseline results are fairly robust to including political variables in our model. The only variable added with a highly statistically significant coefficient is the government majority which enters both equations significantly positively. The marginal effect at the median, however, is only statistically significant in decreasing downgrade probability (though quite sizable in increasing upgrade probabilities). This result is highly plausible. A clear majority does not indicate stability, which is positive in itself, but also reduces the probability of logrolling, where progress is only made by expensive compromises satisfying the clientele of political interest groups.
Although the coefficients of *yroffice* and *yroffice.sq* are insignificant or barely significant, the marginal joint effect is not. With about four years in office, we find a small reduction of downgrade probabilities that is mostly driven by a reduction in reassessment prob-

[17]The conditions in those countries might actually drive the point estimator for rich non OECD countries. The median of fiscal (5.6%) and external (17.2%) surplusses of these two countries is slightly above the best values for OECD countries. Thus, they are able to counteract higher downgrade probabilities. Last, the plot extrapolates for *gnipc*>1.02.

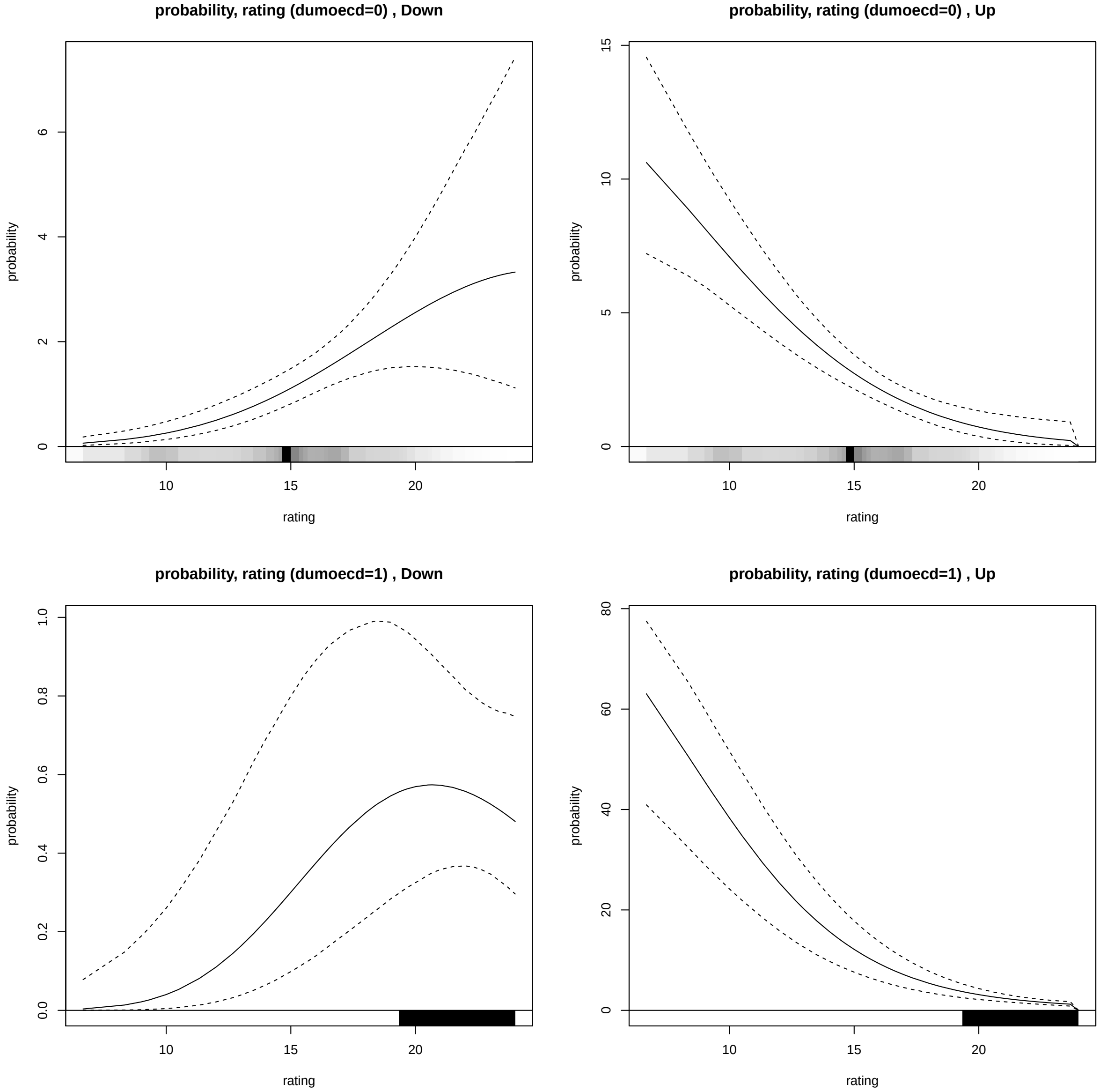


Figure 6: Downgrade and upgrade probabilities for different ratings, split by dumoecd

*Note:* The plot shows probability estimates (solid line) with 90% confidence bands (dashed lines). Darker colors in the gray bar at the bottom of the plot indicate a higher density of observations in the direct neighbourhood of median observations. For OECD countries, the range of rating values is shown instead of their density.

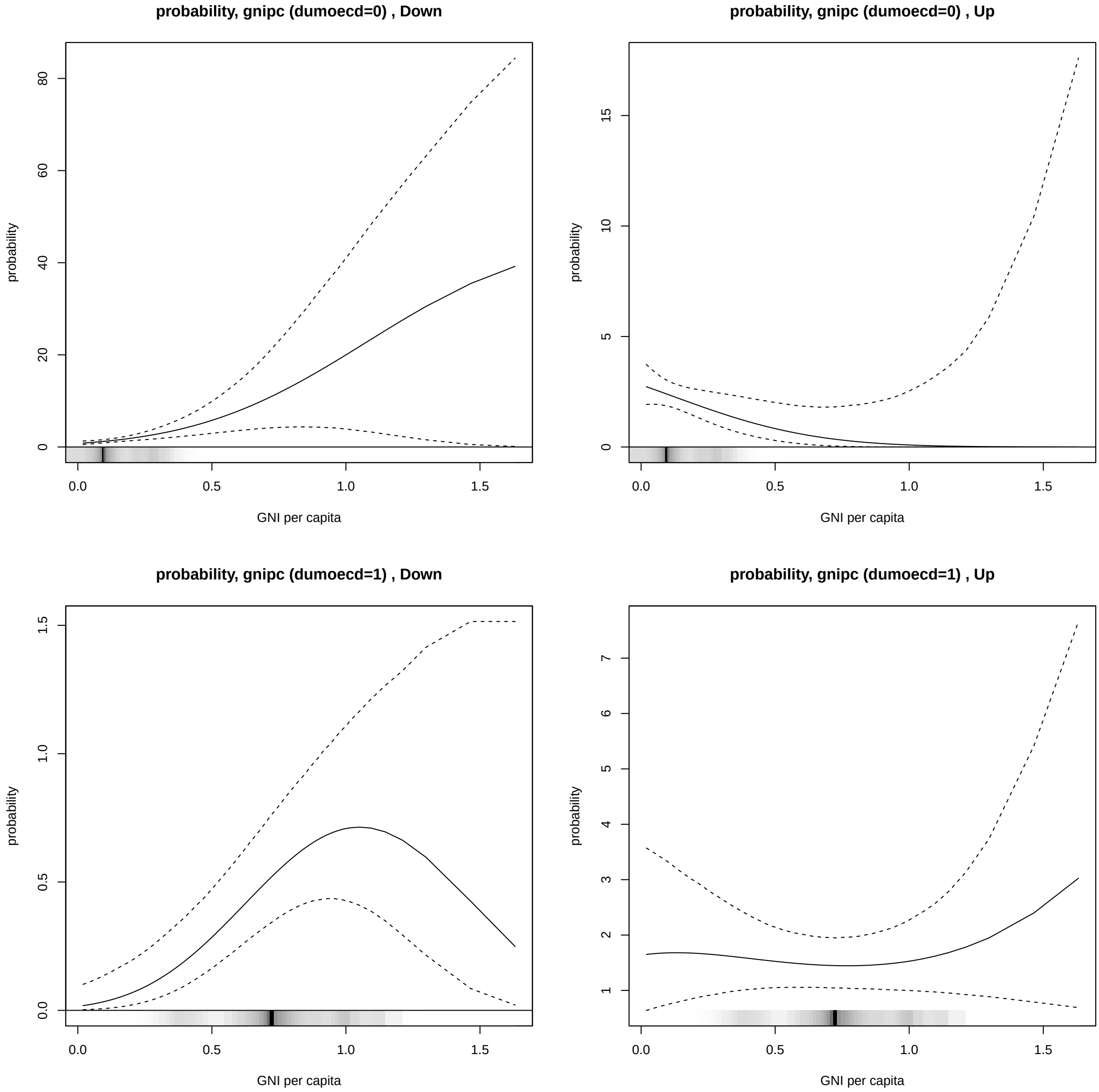


Figure 7: Downgrade and upgrade probabilities for different gnipc, split by dumoecd

*Note:* The plot shows probability estimates (solid line) with 90% confidence bands (dashed lines). Darker colors in the gray bar at the bottom of the plot indicate a higher density of observations in the direct neighbourhood of median observations.

ability. This also, matches our expectations, as four years in office indicates a level of stability where the government serves a typical election cycle, but is clearly far away from authoritarian regimes.

# 6 Conclusion

Our results provide strong evidence that the decisions whether or not to rate and which rating to assign have to be considered separately in order to truly understand the dynamics of rating decisions. Previously, a lot of delays in rating decisions have been blamed on the rating agencies' inability to correctly assess country risk. Yet, we find that the delays are often driven by rational inattention and strategical considerations (such as not alienating major OECD countries unnecessarily). Rating agencies only update if they have reason to believe that there has been enough change since the last assessment, and that the potential reputation loss by not adjusting the rating in time outweighs the cognitive cost of reassessment. Once the rating agencies decide to assess, their assessment of fundamentals that are observable is in line with economic theory and moreover seems to be widely shared by the market. Indeed, markets trust rating agencies enough such that rating actions strongly deviating from ratings implied by current fundamentals provide news and impact government bond yields strongly. Instead, deviations driven by inactivity are not perceived in the same way. In other words, the markets already adjust to easily observable macroeconomic and institutional changes before the rating agencies do. Thus, once the rating is adjusted to the level implied by those fundamentals, the impact on interest rates is inconsequentially small. This ambiguity may explain the seeming contradiction in the previous literature, that partly claims that ratings drive markets, and partly claims that rating agencies usually follow the market because they fail to recognize risk. In a way, both explanations seem to be true. Rating agencies sometimes follow the market in the sense that they respond to the same events later (due to their original inattention without necessarily actually following the market assessment itself). Yet, their detailed reports can truly contain new information that goes beyond usually considered fundamentals, thus driving market participants.

# References


Afonso, A. (2003). Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies, *Journal of Economics and Finance* **27**(1): 56–74.

Al-Sakka, R. & ap Gwilym, O. (2009). Heterogeneity of sovereign rating migrations in emerging countries, *Emerging Markets Review* **10**(2): 151–165.

Altdörfer, M., De las Salas, C. A., Guettler, A. & Löffler, G. (2016). European versus anglo-saxon credit view: Evidence from the eurozone sovereign debt crisis, *IWH Discussionpaper 34/2016*.

Amstad, M. & Packer, F. (2015). Sovereign Ratings of Advanced and Emerging Economies after the Crisis, *BIS Quarterly Review December* .

Bagozzi, B. E. & Mukherjee, B. (2012). A Mixture Model for Middle Category Inflation in Ordered Survey Responses, *Political Analysis* **20**(3): 369–386.

Balassa, B. (1964). The Purchasing-Power Parity Doctrine: a Reappraisal, *Journal of Political Economy* **72**(6): 584–596.

Ball, L. (1992). Why does High Inflation Raise Inflation Uncertainty?, *Journal of Monetary Economics* **29**(3): 371–388.

Beers, D. T. & Nadeau, J.-S. (2014). Introducing a new database of sovereign defaults, *Available at SSRN 2609162* .

Bernoth, K., Von Hagen, J. & Schuknecht, L. (2012). Sovereign Risk Premiums in the European Government Bond Market, *Journal of International Money and Finance* **31**(5): 975–995.

Block, S. A. & Vaaler, P. M. (2004). The price of democracy: sovereign risk ratings, bond spreads and political business cycles in developing countries, *Journal of International Money and Finance* **23**(6): 917–946.

Brewer, T. L. & Rivoli, P. (1990). Politics and perceived country creditworthiness in international banking, *Journal of Money, Credit and Banking* **22**(3): 357–369.

Brooks, R., Harris, M., Spencer, C. et al. (2007). An Inflated Ordered Probit Model of Monetary Policy: Evidence From MPC Voting Data, *MPRA Discussion Paper 8509*.

Butler, A. W. & Fauver, L. (2006). Institutional environment and sovereign credit ratings, *Financial Management* **35**(3): 53–79.

Cantor, R. & Packer, F. (1996). Determinants and Impact of Sovereign Credit Ratings, *The Journal of Fixed Income* **6**(3): 76–91.

Cosset, J.-C. & Roy, J. (1991). The determinants of country risk ratings, *Journal of International Business Studies* **22**(1): 135–142.

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*.

Depken, C., LaFountain, C., Butters, R. et al. (2007). Corruption and creditworthiness: Evidence from sovereign credit ratings.

Dimitrakopoulos, S. & Kolossiatis, M. (2015). State Dependence and Stickiness of Sovereign Credit Ratings: Evidence from a Panel of Countries, *Journal of Applied Econometrics* .

Eichengreen, B., Park, D. & Shin, K. (2013). Growth slowdowns redux: New evidence on the middle-income trap, *NBER Working Paper 18673*.

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. (2016). Much Ado About Nothing: Sovereign Bond Ratings in the OECD, *Technical report*.

El-Shagi, M. & von Schweinitz, G. (2015). Risk and Return: Is there an Unholy Cycle of Ratings and Yields?, *Economics Letters* **129**: 49–51.

Feder, G. & Uy, L. V. (1985). The Determinants of International Creditworthiness and their Policy Implications, *Journal of Policy Modeling* **7**(1): 133–156.

Ferri, G., Liu, L.-G. & Stiglitz, J. E. (1999). The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis, *Economic Notes* **28**(3): 335–355.

Forbes, K. & Rigobon, R. (2001). Measuring contagion: conceptual and empirical issues, *International financial contagion*, Springer, pp. 43–66.

Goh, J. C. & Ederington, L. H. (1993). Is a Bond Rating Downgrade Bad News, Good News, or No News for Stockholders?, *The Journal of Finance* **48**(5): 2001–2008.

Gültekin-Karakaş, D., Hisarcıklılar, M. & Öztürk, H. (2011). Sovereign risk ratings: biased toward developed countries?, *Emerging Markets Finance and Trade* **47**(sup2): 69–87.

Hantzsche, A. (2015). An Estimation Strategy for Analysing the Determinants of Sovereign Credit Ratings. mimeo.

Haque, N. U., Kumar, M. S., Mark, N. & Mathieson, D. J. (1996). The economic content of indicators of developing country creditworthiness, *Staff Papers* **43**(4): 688–724.

Haque, N. U., Mark, N. C. & Mathieson, D. J. (1998). The relative importance of political and economic variables in creditworthiness ratings.

Harris, M. N. & Zhao, X. (2007). A Zero-Inflated Ordered Probit Model, With an Application to Modelling Tobacco Consumption, *Journal of Econometrics* **141**(2): 1073–1099.

Hu, Y.-T., Kiesel, R. & Perraudin, W. (2002). The estimation of transition matrices for sovereign credit ratings, *Journal of Banking & Finance* **26**(7): 1383–1406.

Lee, S. H. (1993). Are the credit ratings assigned by bankers based on the willingness of ldc borrowers to repay?, *Journal of Development Economics* **40**(2): 349–359.

Mora, N. (2006). Sovereign Credit Ratings: Guilty Beyond Reasonable Doubt?, *Journal of Banking & Finance* **30**(7): 2041–2062.

Mulder, C. B. & Monfort, B. (2000). Using credit ratings for capital requirementson lending to emerging market economies: Possible impact of a new basel accord, *Technical Report 69*, IMF Working Papers.

Panizza, U., Sturzenegger, F. & Zettelmeyer, J. (2009). The economics and law of sovereign debt and default, *Journal of Economic Literature* **47**(3): 651–698.

Samuelson, P. A. (1964). Theoretical Notes on Trade Problems, *Review of Economics and Statistics* **46**(2): 145–154.

Schumacher, I. (2014). On the self-fulfilling prophecy of changes in sovereign ratings, *Economic Modelling* **38**: 351–356.

Sims, C. A. (2003). Implications of rational inattention, *Journal of monetary Economics* **50**(3): 665–690.

White, L. J. (2010). Markets: The Credit Rating Agencies, *The Journal of Economic Perspectives* **24**(2): 211–226.

# Appendix

Table 6: Description of variables

| Variable | Full name | Group | Source[a] | Freq. | Transformation | Combination Rule[b] | Normalization[c] |
|---|---|---|---|---|---|---|---|
| $y^d$ | Rating announcement (Dummy) | Explained variable | CE | Daily | Sum | | |
| $\tilde{y}$ | Rating change (Ternary) | Explained variable | CE | Daily | Sum | | |
| rating | Rating level | Ratings | CE | Daily | first day/month | | demean, norm (24) |
| years | Years since last announcement | Ratings | CE | Monthly | | | demean, norm (12) |
| Up12 | # upgrades in prev. year | Ratings | CE | Monthly | | | norm |
| Down12 | # downgrades in prev. year | Ratings | CE | Monthly | | | norm |
| UpAll12 | % foreign upgrades in prev. year | Ratings | CE | Monthly | | | norm |
| DownAll12 | % foreign downgrades in prev. year | Ratings | CE | Monthly | | | norm |
| default | Dummy: default since 1970 | Ratings | DSD | Yearly | | | |
| recentdefault | Dummy: default in last 10 years | Ratings | DSD | Yearly | | | |

[a]Sources: Bank of Canada Database on Sovereign Defaults (*DSD*); Bank for International Settlement (*BIS*); countryeconomy.com (*CE*); Interamerican Development Bank: Database of Political Institutions 2015 (*DPI*); International Monetary Fund: International Financial Statistics (*IMF-IFS*); national statistical offices (*NSO*); national central banks (*NCB*); Transparency International (*TI*); Thomson Reuters (*TR*); World Bank: World Development Indicators (*WDI*)

[b]Combination Rules: *LS* uses the longest available series in every country; *OC* uses series, if available, in the order given in the table; *REG* performs a regression of all series on a benchmark (country-specific or on the full sample), and employs the estimated coefficients for combination; the regression constant is adjusted in order to remove breakpoints.

[c]Normalization Rules: *win* indicates windsorizing at the 99% (and 1%) level; *demean* indicates demeaned series; *sd* standardizes series to have mean zero and standard deviation one; *norm* indicates normalization to the range $[0, 1]$ or by the given factor.

Table 7: Description of variables, continued

| Variable | Full name | Group | Source | Freq. | Transformation | Combination Rule | Normalization |
|---|---|---|---|---|---|---|---|
| ip | Growth of industrial production | Fundamentals | NSO | Monthly | growth rates | *LS*, first on monthly, then incl. quarterly | sd |
| | | | NCB | Monthly | | | |
| | | | NSO | Quarterly | | | |
| | | | IMF-IFS | Quarterly | | | |
| money | M2 | Fundamentals | NCB | Monthly | growth rates | Country-specific *REG* on longest series | win (upper tail), sd |
| | M3 | | NCB | Monthly | | | |
| | M2 | | IMF-IFS | Monthly | | | |
| reserves | Central bank reserves | Fundamentals | NCB | Monthly | growth rates | | win (upper tail), sd |
| inf | Yearly change of consumer prices | Fundamentals | IMF-IFS | Monthly | | | win (upper tail), sd |
| reer[a] | Real effective exchange rates | Fundamentals | IMF-IFS | Monthly | growth rates | *OC* | win (both tails), sd |
| | | | BIS | Monthly | | | |
| | | | NCB | Monthly | | | |
| yield | Benchmark bond yields (5-10 years) | Fundamentals | TR | Monthly | defl. by curr. inf. | *LS* of benchmark bonds, *OC* with EMBI | win (upper tail), sd |
| | Emerging Markets Bond Index | Fundamentals | JPM | Monthly | | | |
| debt | General government debt | Fundamentals | IMF-IFS | Yearly | | Country-specific *REG* on general government debt | norm (100) |
| | Central government debt | | WDI | Yearly | | | |
| fiscbal | Fiscal balance | Fundamentals | IMF-IFS | Yearly | | *LS* | norm (100) |
| | Fiscal balance | | WDI | Yearly | | | |
| current | Current account balance | Fundamentals | IMF-IFS | Yearly | | Full *REG* on IMF-IFS series | norm (100) |
| | | | WDI | Yearly | | | |
| corrupt | Corruption perception index | Fundamentals | TI | Yearly | | | sd |
| gnipc | Gross national income per capita | Fundamentals | WDI | Yearly | relative to US GNI | | |
| dumoecd | OECD membership dummy | Institutional | | Monthly | | | |
| dumeu | EU membership dummy | Institutional | | Monthly | | | |
| yrsoffc | Exec. years in office | Institutional | DPI | Monthly | | | sd |
| maj | Majority of govt in parliament | Institutional | DPI | Monthly | | | sd |
| exelecpost | # exec. elec. in next year | Institutional | DPI | Monthly | | | |
| exelecpre | # exec. elec. in prev. year | Institutional | DPI | Monthly | | | |
| legelecpost | # leg. elec. in next year | Institutional | DPI | Monthly | | | |
| legelecpre | # leg. elec. in prev. year | Institutional | DPI | Monthly | | | |

[a]Series inverted for Chile, Costa Rica, Croatia, Hungary, Mexico, Mongolia, Sri Lanka , Sweden, Zambia

Table 8: Summary statistics

| | Full sample | | | | | OECD | | | | | other | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | sd | median | min | max | mean | sd | median | min | max | mean | sd | median | min | max | n |
| rating | 19.06 | 4.71 | 19.50 | 2.00 | 24.00 | 4.00 | 23.00 | 21.43 | 3.36 | 24.00 | 2.00 | 15.67 | 15.59 | 4.22 | 24.00 | 9296 |
| default | 0.39 | 0.49 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.15 | 0.36 | 1.00 | 0.00 | 1.00 | 0.75 | 0.43 | 1.00 | 9296 |
| UpAll12 | 0.02 | 0.01 | 0.02 | 0.00 | 0.04 | 0.00 | 0.02 | 0.02 | 0.01 | 0.04 | 0.00 | 0.02 | 0.02 | 0.01 | 0.04 | 9296 |
| DownAll12 | 0.02 | 0.01 | 0.02 | 0.00 | 0.05 | 0.00 | 0.02 | 0.02 | 0.01 | 0.05 | 0.00 | 0.02 | 0.02 | 0.01 | 0.05 | 9296 |
| Up12 | 0.30 | 0.64 | 0.00 | 0.00 | 5.00 | 0.00 | 0.00 | 0.21 | 0.56 | 4.00 | 0.00 | 0.00 | 0.43 | 0.74 | 5.00 | 9296 |
| Down12 | 0.29 | 1.00 | 0.00 | 0.00 | 10.00 | 0.00 | 0.00 | 0.25 | 0.95 | 10.00 | 0.00 | 0.00 | 0.35 | 1.06 | 8.00 | 9296 |
| years | 12.92 | 18.41 | 5.29 | 0.00 | 94.52 | 0.00 | 7.59 | 16.59 | 20.94 | 93.54 | 0.00 | 3.61 | 7.53 | 12.00 | 94.52 | 9296 |
| changefund | 0.00 | 1.00 | -0.52 | -0.60 | 3.24 | -0.60 | -0.49 | 0.11 | 1.09 | 3.24 | -0.60 | -0.54 | -0.16 | 0.82 | 3.10 | 9296 |
| gnipc | 0.50 | 0.40 | 0.42 | 0.01 | 1.94 | 0.12 | 0.73 | 0.71 | 0.36 | 1.94 | 0.01 | 0.11 | 0.20 | 0.21 | 1.02 | 9296 |
| ip | 2.11 | 7.33 | 2.13 | -34.70 | 58.39 | -34.70 | 1.95 | 1.98 | 6.55 | 38.92 | -32.56 | 2.43 | 2.31 | 8.34 | 58.39 | 9296 |
| reserves | 9.32 | 28.92 | 5.96 | -99.05 | 261.82 | -99.05 | 4.42 | 6.77 | 27.41 | 261.82 | -91.22 | 8.69 | 13.07 | 30.62 | 261.82 | 9296 |
| inf | 3.68 | 5.11 | 2.61 | -6.56 | 109.73 | -6.56 | 2.06 | 2.32 | 2.14 | 18.76 | -4.38 | 4.39 | 5.66 | 7.15 | 109.73 | 9296 |
| reer | 0.33 | 6.73 | 0.29 | -35.99 | 37.63 | -35.99 | 0.29 | 0.26 | 5.93 | 37.63 | -35.99 | 0.28 | 0.45 | 7.77 | 36.81 | 9296 |
| yield | 2.41 | 3.26 | 2.13 | -12.10 | 23.13 | -3.49 | 2.01 | 2.23 | 2.18 | 23.13 | -12.10 | 2.34 | 2.69 | 4.37 | 23.13 | 9296 |
| debt | 56.16 | 33.99 | 48.89 | 0.07 | 249.08 | 5.29 | 56.47 | 64.77 | 36.87 | 249.08 | 0.07 | 40.11 | 43.51 | 24.21 | 106.03 | 9296 |
| fiscbal | -2.30 | 3.54 | -2.41 | -32.30 | 12.21 | -32.30 | -2.64 | -2.74 | 3.38 | 5.50 | -11.26 | -2.05 | -1.66 | 3.66 | 12.21 | 9296 |
| current | 0.34 | 6.32 | -0.41 | -24.85 | 27.02 | -14.62 | -0.31 | 0.14 | 5.39 | 17.10 | -24.85 | -0.55 | 0.63 | 7.48 | 27.02 | 9296 |
| corrupt | 58.19 | 22.46 | 54.39 | 18.88 | 100.21 | 27.60 | 72.32 | 68.71 | 18.39 | 100.21 | 18.88 | 37.24 | 42.72 | 18.64 | 94.21 | 9296 |
| dumoecd | 0.60 | 0.49 | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 9296 |
| dumeu | 0.44 | 0.50 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.63 | 0.48 | 1.00 | 0.00 | 0.00 | 0.16 | 0.37 | 1.00 | 9296 |
| yroffice | 4.78 | 3.99 | 3.83 | 0.00 | 31.08 | 0.00 | 3.83 | 4.93 | 4.19 | 31.08 | 0.00 | 3.75 | 4.57 | 3.65 | 22.83 | 8280 |
| maj | 0.57 | 0.16 | 0.55 | 0.05 | 1.00 | 0.05 | 0.55 | 0.58 | 0.16 | 1.00 | 0.05 | 0.55 | 0.56 | 0.15 | 1.00 | 8280 |
| exelecpre | 0.09 | 0.29 | 0.00 | 0.00 | 2.00 | 0.00 | 0.00 | 0.09 | 0.29 | 1.00 | 0.00 | 0.00 | 0.10 | 0.31 | 2.00 | 8280 |
| exelecpost | 0.10 | 0.31 | 0.00 | 0.00 | 2.00 | 0.00 | 0.00 | 0.10 | 0.31 | 2.00 | 0.00 | 0.00 | 0.10 | 0.31 | 2.00 | 8280 |
| legelecpre | 0.26 | 0.44 | 0.00 | 0.00 | 2.00 | 0.00 | 0.00 | 0.26 | 0.44 | 2.00 | 0.00 | 0.00 | 0.25 | 0.44 | 2.00 | 8280 |
| legelecpost | 0.27 | 0.45 | 0.00 | 0.00 | 2.00 | 0.00 | 0.00 | 0.28 | 0.45 | 2.00 | 0.00 | 0.00 | 0.26 | 0.44 | 2.00 | 8280 |

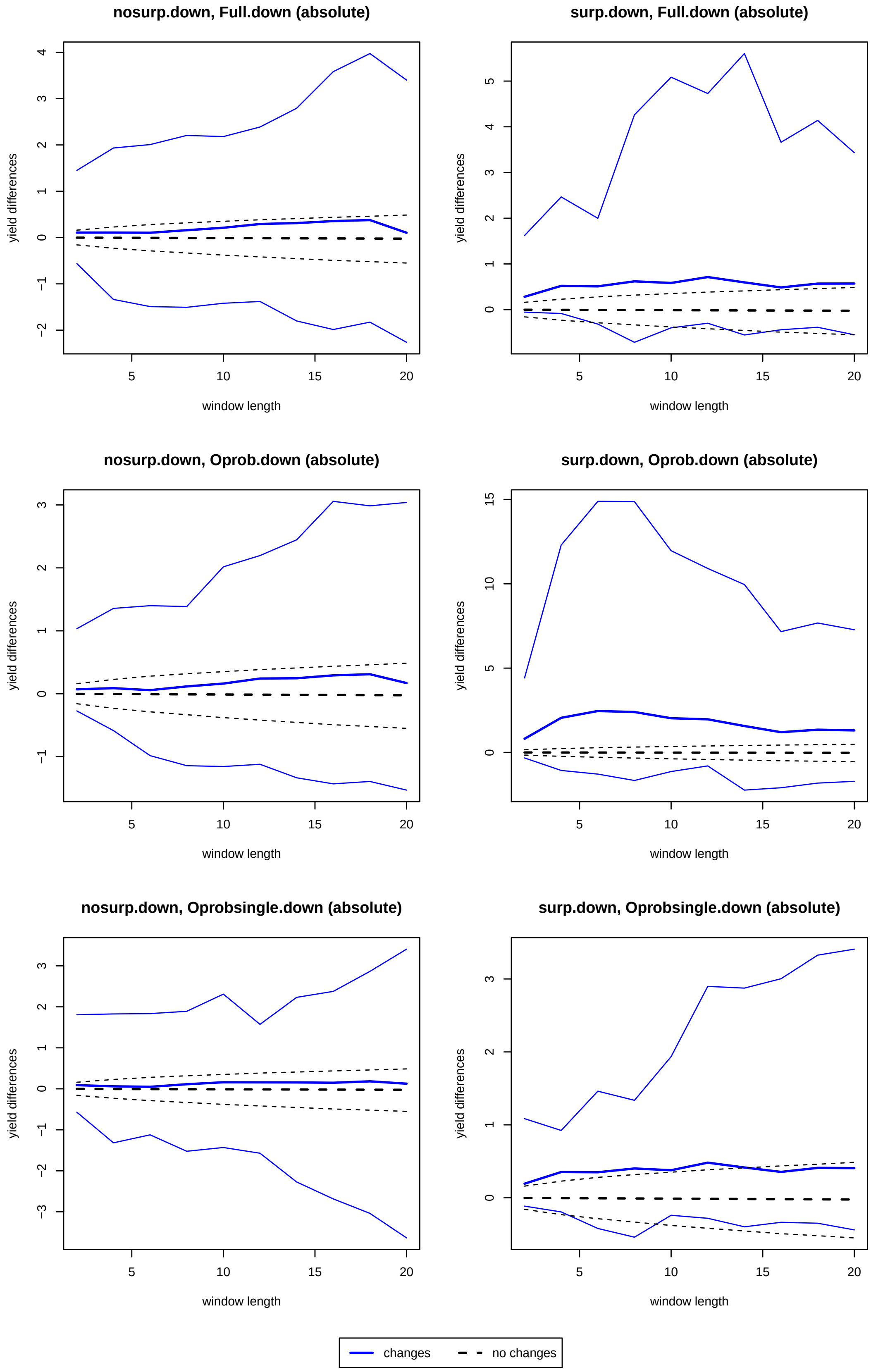


Figure 8: Yield changes around downgrades

*Note:* Yield changes in windows of given length around downgrades and in tranquil times. The left column contains results for unsurprising downgrades, the right column those for surprises. The rows contain different probabilities employed to identify surprises.

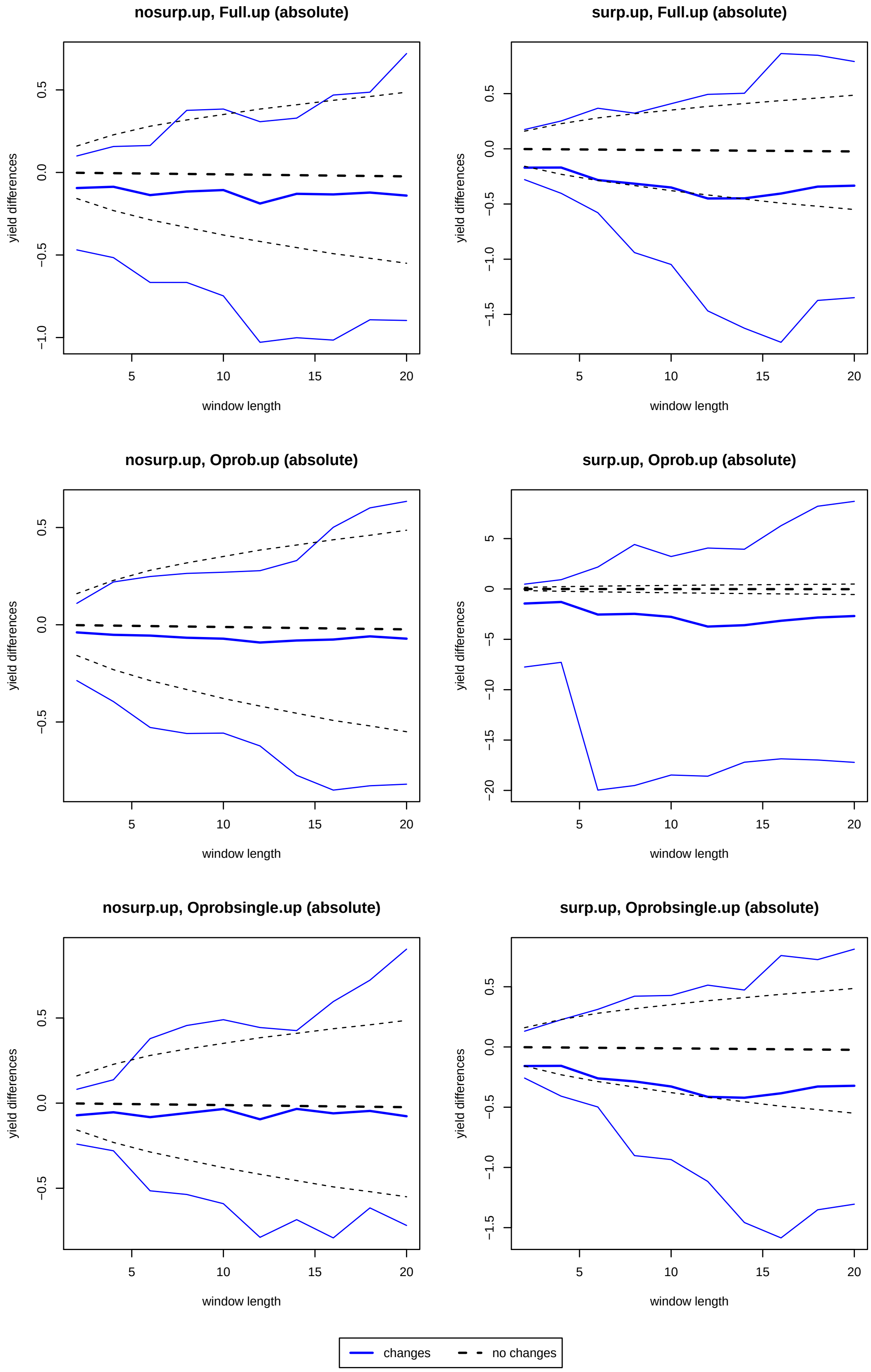


Figure 9: Yield changes around upgrades

*Note:* Yield changes in windows of given length around upgrades and in tranquil times. The left column contains results for unsurprising upgrades, the right column those for surprises. The rows contain different probabilities employed to identify surprises.

Table 9: Estimation coefficients, interaction model with dumeu

| | Basic effects | | Interaction with dumeu | |
|---|---|---|---|---|
| | Reevaluation | Rating decision | Reevaluation | Rating decision |
| rating | -0.886 *** | -2.493 *** | -1.330 *** | 0.073 |
| rating.sq | -1.869 ** | -0.564 | 1.100 | 6.307 *** |
| default | -0.041 | -0.456 *** | | |
| UpAll12 | 0.462 *** | 0.043 | | |
| DownAll12 | 1.188 *** | -0.226 | | |
| Up12 | -0.470 *** | 0.922 *** | | |
| Down12 | 0.217 | -1.212 *** | | |
| years | -0.344 *** | 0.228 *** | | |
| years.sq | 0.063 *** | -0.061 *** | | |
| changefund | 0.103 *** | -0.191 ** | -0.031 | 0.209 * |
| gnipc | 0.887 *** | -1.772 *** | 0.086 | -0.595 |
| gnipc.sq | -0.381 *** | 1.024 ** | 0.106 | -0.321 |
| ip | 0.015 | 0.165 *** | 0.041 | 0.107 |
| reserves | 0.001 | 0.427 *** | 0.035 | -0.670 *** |
| inf | 0.053 | -0.747 * | -3.296 *** | -2.444 |
| reer | -0.043 | 0.125 ** | 0.221 *** | 0.292 * |
| yield | -0.086 ** | -0.092 | -0.240 *** | -0.116 |
| debt | 0.074 | -0.194 | -0.050 | 0.392 |
| fiscbal | -1.209 | 6.550 *** | 1.750 | -1.173 |
| current | 0.901 * | 0.792 | -2.031 ** | 4.251 ** |
| corrupt | -0.120 ** | 0.588 *** | 0.052 | -0.083 |
| dumeu | -0.787 ** | -0.867 | | |
| Constant | -2.277 *** | 2.966 *** | | |
| Thresh 0.1 | - | 2.724 *** | | |
| LL | -4769.522 | | | |
| N | 9296 | | | |

*Note:* Coefficients on interaction terms given in the last two columns.

Table 10: Estimation coefficients, interaction model with gnipc

| | Basic effects | | | | Interaction with gnipc | | | |
|---|---|---|---|---|---|---|---|---|
| | Reevaluation | | Rating decision | | Reevaluation | | Rating decision | |
| rating | -1.203 | *** | -0.610 | | -1.829 | ** | -4.060 | ** |
| rating.sq | -3.502 | *** | 3.027 | | 4.926 | ** | 0.742 | |
| default | -0.099 | * | -0.297 | *** | | | | |
| UpAll12 | 0.468 | *** | 0.320 | | | | | |
| DownAll12 | 1.107 | *** | -0.106 | | | | | |
| Up12 | -0.376 | *** | 1.006 | *** | | | | |
| Down12 | 0.284 | | -1.494 | *** | | | | |
| years | -0.316 | *** | 0.244 | *** | | | | |
| years.sq | 0.062 | *** | -0.061 | *** | | | | |
| changefund | 0.174 | *** | -0.193 | ** | -0.164 | *** | 0.172 | |
| gnipc | -0.460 | | -0.815 | | | | | |
| gnipc.sq | -0.410 | * | 1.766 | ** | | | | |
| ip | -0.032 | | 0.131 | ** | 0.145 | ** | 0.218 | |
| reserves | 0.005 | | 0.399 | *** | 0.046 | | -0.835 | *** |
| inf | 0.479 | | -1.247 | ** | -5.499 | *** | 5.821 | |
| reer | -0.037 | | 0.171 | ** | 0.073 | | -0.051 | |
| yield | 0.021 | | -0.075 | | -0.784 | *** | 0.176 | |
| debt | 0.066 | | -0.081 | | -0.101 | | 0.104 | |
| fiscbal | -0.910 | | 4.945 | ** | 0.789 | | 3.446 | |
| current | 0.248 | | 3.029 | *** | -0.921 | | -0.173 | |
| corrupt | -0.040 | | 0.435 | ** | -0.070 | | 0.231 | |
| Constant | -1.988 | *** | 2.419 | *** | | | | |
| Thresh 0.1 | - | | 2.702 | *** | | | | |
| LL | -4766.310 | | | | | | | |
| N | 9296 | | | | | | | |

*Note:* Coefficients on interaction terms given in the last two columns.