ELSEVIER

Journal of Economic Behavior & Organization
Vol. 56 (2005) 365–381

JOURNAL OF
Economic Behavior
& Organization

www.elsevier.com/locate/econbase

Avoiding the rating bounce: why rating agencies are slow to react to new information

Gunter Löffler*

Department of Finance, University of Ulm, Helmholtzstrasse 18, 89081 Ulm, Germany

Received 17 September 2002; received in revised form 13 March 2003; accepted 18 September 2003
Available online 3 August 2004

Abstract

Rating agencies state that they take a rating action only when it is unlikely to be reversed shortly afterwards. Using a formal representation of the rating process, I show that such a policy provides a good explanation for the empirical evidence: rating changes relatively seldom occur, they exhibit serial dependence, and they lag changes in the issuers' default risk. In terms of informational losses, avoiding rating reversals can be more harmful than monitoring credit quality only twice per year.

JEL classification: G2; G21

Keywords: Credit rating; Rating agencies; Conservatism; Rating migration

1. Introduction

Moody's Investors Service, one of the leading credit rating agencies, takes a rating action only "when it is unlikely to be reversed within a relatively short period of time" (Cantor, 2001, p. 175). As an explanation for this rating policy, Cantor cites the market's "expectation for stable ratings." Intriguingly, rating agencies are often accused of being too slow to adjust their ratings.[1] Could it be that the criticism rating agencies receive is the

* Tel.: +49 731 5023597; fax: +49 731 5023950.
E-mail address: loeffler@mathematik.uni-ulm.de (G. Löffler).

[1] See, for example, the Economist (1997, p. 70) on the Asian crisis: "The raters, firms such as Moody's Investors Service, Standard and Poor's, Duff and Phelps and IBCA, are supposed to be the financial markets' early warning system. Instead, the agencies have spent the past few months belatedly reacting to events."

0167-2681/$ – see front matter
doi:10.1016/j.jebo.2003.09.015

outcome of their desire to meet the market's preferences? This is the question I am going to address.

My analysis is built on a formal representation of the rating process. I model ratings as a mapping of a continuous variable, called credit quality, into discrete categories. Unmanaged, the discreteness produces dependencies in rating changes. The mechanics behind this feature, which is reminiscent of discreteness effects in stock price returns (see Campbell et al., 1997), is as follows: if credit quality follows a probability distribution whose density declines monotonically towards the tails, a threshold triggering a rating change is more likely to be crossed by a small amount than by a large one. The closer the new credit quality is to the rating boundary just crossed, however, the larger is the probability of a subsequent rating reversal relative to the probability of observing another rating change in the same direction. This bias towards rating reversals can be avoided by managing ratings as described above. The reasons for doing so are related to the use of ratings in investment management and capital market transactions.[2] Many pension and mutual funds are subject to investment restrictions concerning risky bonds. A rating reversal can thus induce a reversal of transactions. Since the corporate bond market is notoriously illiquid, this can lead to substantial transaction costs. It is not only reversals of downgrades that are damaging. A fund restricted to invest in investment-grade bonds may have bought a bond upon an upgrade, only to find some time later that the bond has to be sold again. For bond issuers or counterparties in other financial transactions, a rating change can entail irreversible costs even if it is reversed after a short period of time. This holds not only when new debt is issued in between the two rating changes; coupon payments on outstanding debt, for instance, are often related to the current rating.

In the paper, rating management is implemented by setting tolerance regions around rating boundaries. If credit quality surpasses a boundary but lies within the tolerance regions, the rating change is suppressed. Through simulations, I show that a policy of rating bounce avoidance could explain many of the empirical rating characteristics that have been taken to indicate low rating quality. With rating management, ratings are relatively stable, while rating changes are serially correlated and preceded by substantial changes in default probabilities.

Apart from rating bounce avoidance, the agencies' rating systems are characterized by another peculiarity. Most rating agencies employ a through-the-cycle approach that neglects cyclical variations in credit quality when assigning ratings. Though related, the two features are distinct. They both lead to a decrease in rating volatility, but the problem of rating reversals arises even if credit quality is not cyclical.[3] Löffler (2004) shows that the through-the-cycle method, while able to explain important stylized facts like rating stability, fails to account for the predictability of rating changes.

Another possible explanation for the stylized facts is a slow processing of new information. Such an underreaction could be of a psychological nature, reflecting a common human trait (Edwards, 1968). It could also be due to infrequent revisions of ratings. The fact that rating agencies do not monitor ratings continuously is evident from their placing issuers on

[2] See Moody's Investors Service (2002).

[3] Consistent with this view, Cantor states that the avoidance of rating reversals supports the through-the-cycle approach in reducing rating volatility.

watchlists. Agencies would not have to devote special attention to individual issuers if all issuers were under continuous review anyway.

One could suspect that putting an issuer on credit watch indicates a situation in which the credit quality is no longer in line with the current rating but where the rating change has been suppressed in order to avoid its likely reversal. If used in this way, credit watch could mitigate informational losses from rating bounce avoidance because it would signal the true credit quality to outside observers. However, this is not what agencies claim to do.

> "These watchlists list the names of credits whose Moody's ratings have a likelihood of changing. These names are actively under review because of developing trends or events which, in Moody's opinion, warrant a more extensive examination."[4]

> "CreditWatch highlights the potential direction of a short- or long-term rating. It focuses on identifiable events and short-term trends that cause ratings to be placed under special surveillance by Standard & Poor's analytical staff. These may include mergers, recapitalizations, voter referendums, regulatory action, or anticipated operating developments."[5]

Putting a borrower on watch indicates a situation in which, due to imminent events, the probability of a change in credit quality is relatively high, not one where credit quality has already changed.

This paper cannot answer the question whether it is the raters' policy or shortcomings in information processing that underlie the stylized facts. It does show, for example, that the effects of rating bounce avoidance can lead to substantial informational losses similar to those brought about by infrequent monitoring. Any critique of the rating agencies runs the risk of being partial as long as it does not take the official rating policy into account. In consequence, market participants should ask rating agencies to reveal their rating policy in sufficient detail. Otherwise the market will not know what it gets, nor will it be able to evaluate the quality of rating agencies.

The related literature includes papers on empirical characteristics of agency ratings. Carey and Hrycay (2001) and Kealhofer et al. (1998) find that agency ratings are relatively stable compared to alternative rating systems. Altman and Kao (1992) and Lando and Skødeberg (2002) document the existence of serial dependence in rating changes. Delianedis and Geske (1999) show that rating changes lag changes in default probabilities. Their evidence is in line with previous findings that stock prices and stock analyst forecasts predict rating changes (e.g. Holthausen and Leftwich, 1986, and Ederington and Goh, 1998). Extant normative or descriptive papers on rating systems do not address the problem of rating bounces. Krahnen and Weber (2001) propose general standards for good rating practice. A report by the Basel Committee on Banking Supervision (2000) provides a comprehensive overview of rating practices. Crouhy et al. (2001) describe the rating system of Moody's and Standard & Poor's and propose a prototype rating system for bank internal ratings. Kuhner

[4] http://www.moodys.com/moodys/cust/watchlist/watchlist.asp, 23 January 2001.
[5] http://www.standardandpoors.com/ResourceCenter/RatingsDefinitions.html# creditwatch, 23 January 2001.

(2001), finally, presents a signaling game in which rating agencies can have incentives to misrepresent credit quality in times of enhanced systemic risk.

The remainder of the paper is organized as follows. Section 2 presents a formalization of rating processes. Section 3 uses simulations to quantify the effects of rating policies on rating dynamics. Section 4 assesses the costs and benefits of rating management in a fund management context. Section 5 concludes.

2. Formalizing rating policies

Credit ratings can be viewed as a mapping of credit quality into discrete categories $i = 1,\ldots,N$. A borrower receives rating grade i if the credit quality z lies within the boundaries for that grade, b^i_{lower} and b^i_{upper}. I consider a system with 17 rating categories ($N = 17$) excluding default, which is the number of (modified) rating grades for which Moody's and Standard & Poor's publish default rate statistics. As in most rating systems,[6] credit quality is inversely related to the probability of default.

I assume that credit quality z follows a random walk with normally distributed innovations:

$$z_t = z_{t-1} + \varepsilon_t, \quad \varepsilon_t \text{ iid} \quad N(0, \sigma^2) \tag{1}$$

In the simulations, periodicity is one month; the annual variance of credit quality changes is set to unity, which implies $\sigma = 12^{-0.5}$. For the purpose of the paper, the random walk specification is useful because it leaves no role for a through-the-cycle approach. Although it introduces an inconsistency because credit quality diverges to extreme levels as time passes, the robustness checks show that choosing a mean-reverting process does not change conclusions. One could also object that the empirical dynamics of agency ratings are difficult to replicate with structural models that rely on a normally distributed state variable (Gordy and Heitfield, 2001). As demonstrated in Section 3.4, however, such apparent departures from normality can be due to rating management.

For borrowers situated right in the middle between the boundaries of their rating category, the probability that the rating remains stable on a one-year horizon is set to 35 percent.[7] To obtain a rating stability of 35 percent for these median borrowers, the width of a rating class has to be set equal to $-2\Phi^{-1}((1-0.35)/2)$, with $\Phi(\cdot)$ denoting the standard normal cumulative distribution function. Arbitrarily setting the lower boundary of the worst grade (17 $\sim$ CCC) to zero, rating boundaries are as follows:

$$b^i_{\text{lower}} = -2(17-i)\Phi^{-1}((1-0.35)/2), \quad i = 1, \ldots, 17 \tag{2}$$

A rating stability of 35 percent is below the figures reported in Standard & Poor's (2001), where the median stability across grades AAA to CCC is equal to 74.1 percent. However, the

[6] See Basel Committee on Banking Supervision (2000) for a description of rating systems.

[7] The robustness checks will consider different values of rating stability and allow stability to differ across rating categories.

Table 1
One-year default probabilities (percent) within the model

Rating	Default probability at midpoint of rating boundaries	Historical default rates	
		S&P	Moody's
1 ~ AAA	0.04	0	0
2 ~ AA+	0.05	0	0
3 ~ AA	0.06	0	0
4 ~ AA−	0.07	0.03	0.06
5 ~ A+	0.08	0.02	0
6 ~ A	0.09	0.05	0
7 ~ A−	0.10	0.05	0
8 ~ BBB+	0.12	0.12	0.07
9 ~ BBB	0.22	0.22	0.06
10 ~ BBB−	0.35	0.35	0.39
11 ~ BB+	0.44	0.44	0.64
12 ~ BB	0.94	0.94	0.54
13 ~ BB−	1.33	1.33	2.47
14 ~ B+	2.91	2.91	3.48
15 ~ B	8.38	8.38	6.23
16 ~ B−	10.32	10.32	11.88
17 ~ CCC	21.94	21.94	18.85

Default rates are from Standard & Poor's (2001) and Moody's Investors Service (2001).

empirical stability of agency ratings is likely to be affected by the rating policy. Kealhofer et al. (1998) report a transition matrix for ratings that are based on statistical estimates of default probabilities, that is, a transition matrix not affected by active rating management. There, the median stability is 44.4 percent. Since I use 17 rating grades instead of seven as in Kealhofer et al., I regard a value lower than 44.4 percent to be appropriate.

To associate a given credit quality with a default probability, I set the one-year default probabilities of median borrowers equal to the historical default rates of the corresponding S&P rating categories. In particular for ratings better than BBB, historical default rates are imprecise estimates of the underlying default probabilities; sometimes they are zero. I therefore depart from the historical default rates and let the model default probabilities decline linearly from 0.1 percent for category 6 (∼A) to 0.04 percent for category 1 (∼AAA). The model default probabilities as well as the empirical default rates for S&P and Moody's rating grades are listed in Table 1.

Having specified the default probabilities for median borrowers, the default probability for any given z is obtained through linear interpolation.[8] In the simulations, default is modeled as an exogenous event whose probability depends on credit quality. In month t, the probability of default is one twelfth of the default probability associated with the previous credit quality z_{t-1}. Due to the non-linear relationship between credit quality and default probabilities, the resulting one-year default frequencies need not be equal to the specified

[8] The minimum default probability is set to 0.04 percent; for credit qualities at or below zero, the monthly default probability is set to 100 percent.

one-year default probabilities.[9] Differences are negligible; I simulate 100,000 one-year paths to determine default frequencies. They amount to 0.224 percent (0.988 percent) for borrowers with an initial model default probability of 0.22 percent (0.94 percent). Despite the large sample size, these differences are not statistically significant. Note, too, that the conclusions of this paper do not rest on an analysis of realized default rates.

Even though the state variable z follows a random walk, the rating derived from this variable will not. Consider a borrower whose credit quality crosses a rating boundary. Since credit quality changes follow a bell-shaped distribution, the boundary is more likely to be exceeded by a small than by a large amount. Conditional on a rating change, the probability that the rating change is reversed is thus larger than the probability that the rating change is followed by another change in the same direction. In the limit, when the credit quality just hits the boundary, the probability of a reversal is 50 percent, while the probability of observing another change in the same direction is, on a one-year horizon, equal to $\Phi(2\Phi^{-1}((1-0.35)/2))$ = 18.21 percent.

As noted in the introduction, Moody's claims to take a rating action only when it is unlikely to be reversed within a relatively short period of time. The description neither specifies the time horizon nor what is exactly meant by unlikely. Assume that, at each rating review date, the rating agency wants to keep the probability of a reversal within the next m years below P^*. Within the rating model described above, such a policy can be formulated as follows. The probability P that a rating change is reversed in the next m years depends on the difference between the credit quality and the rating boundary just crossed:

$$\text{Prob(reversal)} = P = \begin{cases} \text{Prob}(z_{t+m} \leq b^{\text{crossed in } t}) = \Phi\left((b^{\text{crossed in } t} - z_t)/\sqrt{m}\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{after upgrades} \\ \text{Prob}(z_{t+m} > b^{\text{crossed in } t}) = \Phi\left((z_t - b^{\text{crossed in } t})/\sqrt{m}\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{after downgrades} \end{cases} \tag{3}$$

The rating policy prescribes that the probability of reversal P is smaller than a target value P^*.[10] It can be implemented by requiring credit quality to exceed a rating boundary by at least $\left|\Phi^{-1}(P^*)\sqrt{m}\right|$ in order for a rating change to occur. What happens if credit quality crosses two boundaries, but fails to exceed the second boundary by the critical amount? In this case, the rating will be adjusted by one grade rather than two. Fig. 1 shows the various possibilities for a single-period change in credit quality.

The following example illustrates the conservatism that can be introduced by such a rating policy. Assume P^*= 0.2 and m = 1, that is, the rating agency wants to avoid situations

[9] Consider two one-year paths for a BBB− borrower (default probability = 0.22 percent): (i) in the first month, credit quality rises to BBB+ (default probability = 0.12 percent) and stays there; (ii) in the first month, credit quality falls to BBB− (default probability = 0.35 percent) and stays there. One-year default probabilities are 0.22 percent/12 + 0.12 percent × (11/12) and 0.22 percent/12 + 0.35 percent × (11/12) for cases (i) and (ii), respectively. If the two cases are equally likely, the average default probability is 0.23 percent.

[10] Strictly speaking, the probability of reversal examined here is the probability that credit quality moves to a level consistent with the previous rating. Since rating management influences not only current but also subsequent rating decisions, this is different from the probability of actual rating reversals prevailing under a volatility-reducing rating policy.

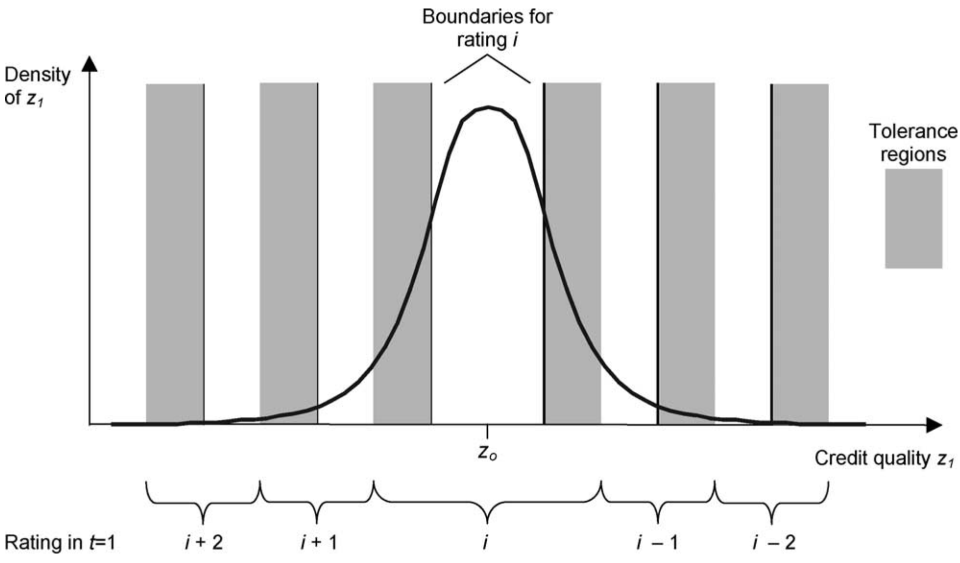

Fig. 1. Schematic representation of the rating policy for a one-period change in credit quality of a borrower rated i in $t = 0$.

where rating changes are reversed with a probability of 20 percent within one year. A rating boundary then has to be exceeded by $|\Phi^{-1}(0.2)| = 0.842$. This tolerance region is almost as wide as the interval pertaining to one rating grade, which has a width of 0.908 for a rating stability of 35 percent. In effect, such a rating policy would blur differences between neighboring rating grades.

3. Rating bounce avoidance as an explanation of stylized facts

In this section, I use the rating model described above to assess whether the desire to avoid frequent rating reversals could underlie the peculiarities of agency ratings that have been documented in the literature. Notably,

- agency ratings appear relatively stable compared to other rating systems;
- ratings exhibit drift. Subsequent rating changes in the same direction are more frequent than subsequent rating changes in opposite directions;
- ratings lag changes in issuers' default probabilities.

These stylized facts will be addressed one after another within the framework laid out in the previous section. I assume that rating agencies pursue a policy of avoiding rating reversals. The tolerated probability for rating reversals P^* is set at 0.1, 0.2 or 0.3; the time period m is chosen to be 0.25, 0.5 and 1, corresponding to time intervals of three, six and twelve months, respectively.

Table 2
Simulated one-year stability of credit ratings for different rating policies

P^* (tolerated reversal probability)	m (time horizon for reversal probability)		
	0.25	0.5	1
0.1	0.58	0.73	0.87
0.2	0.45	0.55	0.70
0.3	0.38	0.43	0.51
0.5	0.35	0.35	0.35

The effects of rating bounce avoidance are assessed through Monte Carlo simulations. Periodicity is one month. In one run of the simulations, I generate a random path for the credit quality z. According to the mapping rules from Section 2, the credit quality z is translated into ratings. Since the assumed credit quality dynamics are independent of the current and past credit qualities and the width of rating categories is uniform across grades, the starting value for the credit quality is not decisive for the results. I choose the initial credit quality to be the one of a median borrower within rating category BB. On a one-year horizon, the associated default probability is 0.94 percent. One run of the Monte Carlo simulations extends over a period of ten years. I perform 10,000 replications for each parameterization.

3.1. *Ratings are relatively stable*

Kealhofer et al. (1998) and Carey and Hrycay (2001) find that agency ratings are significantly less volatile than alternative ratings. Kealhofer et al. estimate default probabilities based on the Merton (1974) model and categorize borrowers according to these probabilities. Carey and Hrycay use a logit model to assign borrowers to rating grades. Typically, 40 percent to 50 percent of these ratings remain stable over a one-year horizon, compared to 80 percent to 90 percent in the case of agency ratings. The ratings constructed in Kealhofer et al. and in Carey and Hrycay are based on seven and five categories, respectively. Since I use 17 grades in this paper, rating stability will generally be lower. However, the simulation results and the empirical evidence can still be compared with respect to relative differences in rating stability.

Table 2 summarizes simulated one-year transition probabilities for various assumptions about the acceptable reversal probability P^* and the time horizon m used to compute this probability. If there is no rating bounce avoidance (that is, P^* is equal to the maximum value of 0.5) rating stability is equal to the 35 percent that was used to calibrate the model. For $P^* < 0.5$, rating stability increases. It ranges from 38 percent to 87 percent. With a tolerated reversal probability of 0.2 and a time horizon of six months, the rating stability is 55 percent. Rating bounce avoidance can thus lead to a considerable increase in the stability of credit ratings. The simulated figures largely mirror the empirical differences between agency ratings and rating systems that are known not to be influenced by ratings management. Since the precise rating policy of the agencies is unknown, it is difficult to judge whether rating bounce avoidance completely explains the empirical evidence. For the intermediate parameter combination $P^* = 0.2$ and $m = 0.5$, for example, one could argue that it does not. Both in Kealhofer et al. (1998) and in Carey and Hrycay (2001), the stability of agency ratings is up to twice that of alternative rating systems; in Standard & Poor's (2001), the maximum empirical stability for modified grades is 90 percent; the median across the 17 grades is 78 percent. However, this would not leave us with a puzzle. The agencies' policy of rating through the cycle, which is not modeled here, can also lead to a significant increase in rating stability (see Löffler). Together, the two peculiarities of the agencies' rating approach could well explain the empirical facts, even if rating bounce avoidance alone did not.

3.2. Rating changes are serially dependent

Empirical studies of rating changes have documented significant positive serial dependence (Altman and Kao, and Lando and Skødeberg). Such a dependence can arise even if ratings are continuous and rating analysts efficiently use available information. An analyst who learns that the default probability of a firm will decrease over time will not completely incorporate this information into the current rating if the rating horizon is shorter than the time span in which the firm's restructuring is completed. Partial responses to new information, however, will create positive serial dependence. Since I model credit quality as a random walk, this explanation can be ruled out for the simulation experiments conducted here.

Altman and Kao (1992) examine the rating dynamics of 1970–1985 new bond issues. They measure serial dependence through a statistic defined as the frequency of subsequent rating changes in one direction divided by the frequency of subsequent rating changes in opposite directions. If ratings exhibit positive drift, the statistic is larger than one. If an upgrade is more likely to be followed by a downgrade, and vice versa, the statistic is smaller than one. I compute the statistic within the simulated samples. In each run, which spreads over ten years, I take only the first two rating changes to compute the statistic; this is appropriate because Altman and Kao examine rating changes of newly rated bond issues. If a simulation run contains less than two rating changes, it does not enter the calculation of the test statistic.

The results are reported in Table 3. If raters do not try to avoid rating reversals (P^* = 0.5), the statistic is 0.39, meaning that the probability of observing rating changes in opposite directions is more than twice the one of observing rating changes in identical

Table 3
Simulated serial dependence statistics for different rating policies

P^* (tolerated reversal probability)	m (time horizon for reversal probability)		
	0.25	0.5	1
0.1	1.86	2.45	4.01
0.2	1.28	1.71	2.20
0.3	0.98	1.15	1.50
0.5	0.39	0.39	0.39

The statistic is defined as the frequency of observing subsequent rating changes in the same direction divided by the frequency of observing subsequent rating changes in opposite directions. It is greater than one for positive serial dependence.

directions. This is the reflection of the rating bounce. If ratings are set to avoid this bounce, the statistics range from 0.98 to 4.01. For most of the parameter values chosen here, rating bounce avoidance thus leads to positive serial dependence in rating changes. The values are in line with the ones reported by Altman and Kao separately for issuer groups, and for up- and downgrades. The mean (median) of their statistics is 1.752 (1.475), with a range of 0.2–3.83.

The rating policy modeled here could thus account for the existing evidence. This is important since another peculiarity of agency ratings, the through-the-cycle approach, cannot (see Löffler). Lando and Skødeberg document that the rating drift is especially pronounced for downgrades. This is sometimes explained by noting that agencies "dole out the bad news in small doses rather than savaging the bond issuer – who is, after all, their customer – all in one go" (Economist, 1997, p. 71), but it could also be explained through avoidance of rating reversals. It seems likely that the costs of rating changes are larger for downgrades. After a downgrade, investment restrictions may force investors to sell bonds, and covenants may restrict the flexibility of the borrower. If rating agencies act in the interest of their clients, they could try to avoid rating reversals particularly for downgrades.[11] In the framework of this paper, such a policy is not the same as "doling out bad news in small doses" because rating changes are suppressed rather than handed out piecemeal. Rating drift arises because rating changes are only made when the credit quality is relatively close to the boundary triggering a further rating change.

In a related experiment, I examine whether serial dependence could be due to infrequent rating reviews. Assume that agencies monitor ratings only in six month intervals. In the simulation, monitoring dates are thus t = 0, 6, 12 and so forth. At a monitoring date, the rating is set according to the credit quality so that there is no rating bounce avoidance. With such an infrequent monitoring, the simulated autocorrelation statistic is 0.80. If the frequency is further reduced to one rating review per year the statistic increases to 0.89. It is thus difficult to explain the empirical evidence on serial correlation with infrequent monitoring.

3.3. Ratings lag changes in default probabilities

Based on the option-theoretic models of Merton (1974) and Geske (1977), Delianedis and Geske (1999) use balance sheet data, equity values and equity volatilities to compute risk-neutral default probabilities[12] for borrowers rated by Standard & Poor's. They examine how these default probabilities evolve before a rating change and find that they rise (fall) several months before a downgrade (upgrade).

[11] "Especially in the case of downgrades, the potentially self-fulfilling nature of ratings requires that Moody's particularly endeavor to avoid 'false' negative predictions" (Moody's Investors Service, 2002, p. 4).

[12] In option-theoretic models, default is triggered if a firm's asset value falls below the value of liabilities. The higher the growth rate of assets, the lower the default probability, ceteris paribus. Risk-neutral default probabilities are obtained under the assumption that all investors are risk-neutral and, in consequence, all assets earn the risk-free rate of return. The change in the rate of return is achieved by a change of measure from actual to risk-neutral probabilities, and not by changing cash flows. In the real world, risky assets command a premium. Actual default probabilities are thus lower than risk-neutral ones. One reason for focusing on the latter is that one does not need to estimate expected asset returns.

Table 4
Simulated median default probabilities (percent) in the month before a downgrade from grade 12 (default probability = 0.94 percent) for different rating policies

P^* (tolerated reversal probability)	m (time horizon for reversal probability)		
	0.25	0.5	1
0.1	1.32	1.76	2.41
0.2	1.23	1.30	1.65
0.3	1.16	1.21	1.27
0.5	1.04	1.04	1.04

Within the simulated samples, I examine the default probabilities one month before the first downgrade, regardless of the magnitude of the downgrade. Recall from Table 1 that the initial default probability of grade 12 is 0.94 percent; the median default probability of the next lower rating class is 1.33 percent. If the rating agency does not aim at avoiding rating reversals, a downgrade occurs as soon as the default probability exceeds 0.5 (0.94 + 1.33 percent) = 1.14 percent. Due to the discrete nature of the rating system, the default probability one month before a downgrade will not be equal to the initial one. Downgrades are more likely to be observed if the credit quality has declined within the range associated with the initial rating. This effect is documented in Table 4. Even if raters are not concerned about reversals ($P^* = 0.5$), the median default probability one month before a downgrade has increased relative to the initial one, from 0.94 percent to 1.04 percent The effect is relatively small, however, which is due to the rating system being relatively fine. With rating bounce avoidance ($P^* < 0.5$), rating changes lag much further behind changes in default probabilities. Depending on the parameters, the median default probability one month before a rating change can be up to 2.41 percent, more than twice the initial default probability of 0.94 percent.

How do these figures compare to the results in Delianedis and Geske? For investment-grade bonds, the median default probability one month before any downgrade is 1.1 percent, while the median default probability of a benchmark sample that does not contain downgrades is 0.7 percent. These figures are very similar to the ones generated through intermediate assumptions on rating bounce avoidance. Setting $P^* = 0.2$ and $m = 0.5$, for example, the comparable figures are 1.30 and 0.94 percent, respectively.[13]

Another study questioning the information content of agency ratings is by Perraudin and Taylor (2004) who show that bond yields often lie above (below) the average yield of bonds with the next lower (higher) letter rating. In the examples presented here, tolerance regions do not spread across more than two ratings; an issuer with credit quality corresponding to AA− can have a rating of A+ or A, but not of A−. Together with pricing errors or uncontrolled factors, however, rating management can help to explain the empirical evidence.

Some empirical studies (e.g. Hand et al., 1992) suggest that bond price reactions to bond rating changes are relatively weak. It is, therefore, interesting to ask whether rating

[13] Delianedis and Geske compute risk-neutral probabilities. To check whether this affects the comparison, I translate them to actual probabilities as described in Crouhy et al. (equation 15). With a Sharpe ratio of 0.5, 0.7 percent and 1.1 percent translate into actual probabilities of 0.16 and 0.26 percent, respectively. Rating management leads to very similar values when the initial rating is 8 (∼BBB+) instead of 12.

management could have such an effect. As modeled in this paper, rating management does not lead to a situation in which rating changes are associated with smaller changes in credit quality. Managed or unmanaged, ratings change when credit quality crosses a threshold; the main difference is that, under rating management, thresholds are path-dependent. However, rating management makes it more difficult for outsiders to infer the underlying credit quality from ratings. In the information aggregation process leading to market prices, managed ratings will receive a weight that is smaller than the one investors would attach to unmanaged ones. As a consequence, price reactions to rating changes do not fully reflect the information produced by rating agencies.

Finally, I compare the effects of the rating policy to the ones that would arise from infrequent rating reviews. With semi-annual monitoring, the simulated median default probability before a downgrade is 1.16 percent, which is lower than some of the values that obtain with rating bounce avoidance. With P^* = 0.2 and m = 0.5, for example, the median default probability is 1.30 percent (see Table 4). This shows that rating bounce avoidance can be more harmful to the timeliness of a rating system than restricting the number of rating reviews to only two per year.

3.4. Sensitivity analyses

To examine the robustness of the results, I re-run the analyses for the parameter combination P^* = 0.2 and m = 0.5, making the following, non-accumulating variations:

(A) The initial credit quality conforms to rating 6 (∼A) rather than 12 (∼BB).
(B) Rating stability without rating management is set to 30 percent instead of 35 percent.
(C) Rating stability without rating management is set to 40 percent instead of 35 percent.
(D) Rating stability increases linearly from 30 percent (grade 17 ∼CCC) to 62 percent (grade 1 ∼AAA) instead of being constant at 35 percent. Stability for grade 12 (∼BB) is 40 percent.
(E) Credit quality follows a mean-reverting process instead of a random walk:

$$z_t - z_{t-1} = 0.015(z_0 - z_{t-1}) + u_t, u_t \sim \text{iid} \quad N(0, 0.090)$$

The process has an annual variance of one. The annual speed of adjustment is $(1 - 0.985^{12}) = 0.166$, which is at the upper end of the estimates that Fama and French (2002) obtain for the speed of adjustment to target leverage ratios. The tolerance regions account for the fact that the expected change in credit quality is non-zero whenever the credit quality differs from the initial one.

Table 5 compares the simulation results with the previous ones. Differences are small, or as expected. Changing the initial credit quality from BB to A does not change rating stability or the autocorrelation statistic. The latter increases when rating stability without rating management is lowered to 30 percent because tolerance regions spread further into the next rating category; increasing stability to 40 percent leads to opposite effects. Making rating stability heterogeneous produces much the same results as a uniform stability of 40 percent; in both cases, the stability of the initial credit quality is 40 percent. With mean reversion the width of the rating categories remains the same but tolerance regions widen because mean

Table 5
Sensitivity analyses (reversal probability P^* = 0.2 and time horizon m = 0.5 for each experiment)

Base case		Variation				
		A	B	C	D	E
One-year stability	0.55	0.55	0.50	0.60	0.61	0.67
Autocorrelation statistic	1.71	1.67	1.98	1.46	1.47	1.20
Default probrobability before downgrade (percent)	1.30	0.10	1.33	1.28	1.29	1.60

A: initial credit quality ∼A; B: rating stability without rating management = 30 percent; C: rating stability without rating management = 40 percent; D: rating stability without rating management decreasing from 62 percent to 30 percent; E: mean-reverting credit quality.

reversion increases the probability of a reversal. In consequence, rating stability and the time-lag in rating actions increase. Since the credit quality is now negatively autocorrelated, the autocorrelation statistic decreases. Nevertheless, previous conclusions can be upheld.

Gordy and Heitfield use a structural model similar to the one presented in Section 2 to replicate empirical rating transition data. They find that choosing a fat-tailed distribution for credit quality provides a better fit than the normal. To check whether this finding can be the result of rating management, I perform an analysis similar to theirs. From the simulated data, I obtain one-year transition frequencies of grade 12 issuers for the case that rating policy tolerates a reversal probability of 0.2 at a six-month horizon. I take the rating thresholds defined in Section 2 as given and calibrate the credit quality distribution that best replicates the simulated transition frequencies. Specifically, I assume that one-year changes in credit quality follow a scaled *t* distribution and numerically search for the variance and the degrees of freedom that minimize the sum of squared differences between simulated transition frequencies and model transition probabilities of median issuers within grade 12.[14] The *t* distribution approaches the normal as the degrees of freedom grow to infinity; the differences between the two distributions are small if the degrees of freedom are above 30. The best fit is obtained with 3 degrees of freedom; repeating the analysis for transition frequencies of grade 6 issuers produces the same result. Thus, empirical evidence of leptokurtosis is consistent with credit quality following a normal distribution and agencies pursuing a policy of rating bounce avoidance.

4. Usefulness of rating management

One reason for reducing the frequency of rating reversals is to reduce transaction costs of pension or mutual funds that are subject to rating-based investment rules. In the following, I present some results which shall help to assess the costs and benefits of rating bounce avoidance in this context. Specifically, I consider a fund that is restricted to invest in investment-grade bonds (i.e. bonds with a rating better than BB+) and examine the effects of rating bounce avoidance in a simulation experiment. I randomly draw a credit quality z from

[14] The calibration does not include the transition to default because default is modeled as an exogenous event in this paper.

the range encompassing rating grades 5 (∼A+) to 10 (∼BBB−).[15] I then simulate credit quality over 12 months and assign ratings in each month: (i) without rating management, and (ii) with rating management ($P^* = 0.2$; $m = 1$). Following a downgrade to a rating worse than 10, the bond is replaced by an investment-grade bond whose credit quality is again randomly drawn. This analysis is repeated 100,000 times.

The costs of rating management are measured through the default frequency. It should be higher with rating management because credit quality could fall below investment grade even though the managed rating does not. Simulated differences are small. Annual default rates are 0.17 and 0.15 percent with and without rating management, respectively. Assuming a typical recovery rate of 50 percent (Altman and Kishore, 1996), rating management thus increases expected losses from default by (0.17 percent − 0.15 percent) × 50 percent = 0.01 percent. Its benefits are lower transaction costs. I examine the number of cases in which an issuer is downgraded to a speculative grade in the course of the year, but regains investment-grade credit quality at the end of year. These cases are instances in which the fund manager may have decided not to sell the bond if she knew in advance that it would again be an admissible investment. The relative frequency of such reversals is 0.05 and 4.83 percent with and without rating management, respectively. If round-trip transaction costs are set to the empirical estimate of 0.59 percent obtained by Chen and Wei (2001), rating management reduces transaction costs by 0.59 percent × (4.83 percent − 0.05 percent) = 0.026 percent. This is more than twice the increase in expected losses from default.

An average advantage of 0.026 percent may not appear to be economically significant, but there are several reasons why it can reach higher levels. First, there is considerable variation in transaction costs: the fifth quintile of estimated round-trip costs in Chen and Wei is 5.49 percent, ten times higher than the median; Schultz (2001) documents that inactive buy-and-hold investors (the traditional clientele of rating agencies) bear above-average trading costs. Second, selling pressure upon a downgrade may depress prices below their fundamental value. Finally, fund managers will be concerned about more than average transaction costs, because short-term performance is an important element of their evaluation. Reversals that occur in a given period can easily exceed the average 4.83 percent.[16]

Of course, the presented measures of costs and benefits are relatively simple. Deriving and incorporating more complex assumptions on investor objectives, investment policies, and transaction costs would be beyond the scope of this paper. In addition, the analysis does not establish that rating management is the optimal solution to the problem. Transaction costs could also be reduced by making investment rules more flexible. Nevertheless, the example indicates that rating management could be a sensible response of rating agencies to the needs of investors.

[15] Most investment-grade issuers have grades below AA− (73 percent according to Standard and Poor's, 2001).

[16] Assume that rating changes are independent across issuers and that the probability of a reversal is 4.83 percent. It follows from the binomial distribution that, in a portfolio with 30 issuers, reversal frequencies of 10 percent or more are observed with a probability of 17.5 percent. Incorporating dependencies would increase the variation in reversal frequencies.

5. Concluding remarks

The paper has shown that the wish to avoid frequent reversals of credit ratings could account for the stylized facts of agency ratings. Empirically, rating changes occur relatively seldom, and they are serially dependent and predictable using borrower fundamentals. Simulations reveal that rating bounce avoidance can explain these peculiarities very well. Moreover, predictability cannot be explained by another characteristic of the agencies' rating system, the through-the-cycle approach (Löffler, 2004). Rating bounce avoidance thus is an important candidate for explaining the stylized facts of agency ratings. Another candidate is biases in the processing of new information. If rating agencies are slow to react to new information, stability will increase, and rating changes will become predictable. Differentiating between these alternative explanations is difficult. The analysis has shown, however, that rating bounce avoidance can reduce the informational content of ratings by more than a rating system that reviews credit quality only twice per year. In addition, infrequent reviews cannot explain the observed serial dependence of rating changes.

Moody's claims that it manages ratings in order to "balance the market's need for timely updates on issuer risk profiles, with its conflicting expectation for stable ratings" (Cantor, 2001, p. 175). It is beyond the scope of this paper to evaluate what the market really wants and whether rating agencies act in response to these preferences or use them to cover any deficiencies of their ratings. It seems obvious, however, that the market's preferences are not homogeneous. Rating management cannot serve all market participants alike. In addition, even if rating management meets an investor's expectation for stability, there may be situations where this particular investor might want to know the precise credit quality, not the one obscured by rating management.

There seem to be two ways of reducing informational losses due to rating bounce avoidance. One is to communicate the precise rating policy; the other is to change the rating system. A move towards greater transparency would be to state how wide the tolerance regions are in terms of rating grades. The analysis has shown that rating management can blur differences between adjacent rating categories. In effect, rating management can offset the increase in accuracy achieved through the rating modification (+ and − in the case of Standard & Poor's) introduced by the rating agencies in the early 1980s. Rating agencies could try to elicit market feedback on whether such an inaccuracy is indeed what the market wants. In the aftermath of the Enron default, Moody's has initiated a dialogue on the quality and timeliness of ratings (cf. Moody's Investors Service, 2002). The market response has confirmed Moody's in its policy of avoiding rating reversals. Even though Moody's aims at greater transparency, however, Moody's has not specified its policy in more detail.

The problem of rating bounces could be reduced by moving from a discrete rating system to a continuous one. This does not imply that the rating is equated with default probabilities; it could still reflect other dimensions of credit risk such as recovery risk, or be based on a combination of default probabilities for various time horizons. There are various possible arguments against continuous ratings. For cognitive reasons, rating analysts might find it easier to aggregate their information into discrete categories, rating agencies might introduce random variation into ratings in order to prove that ratings are reviewed frequently, and market participants might overestimate the accuracy of such a continuous rating. These arguments are appealing, but it has been shown in other contexts that continuously measured expectations can provide better results than qualitative ones. Batchelor (1986), for example, recommends asking for continuous expectations of consumer price inflation rather than for qualitative responses.

The upcoming reform of capital adequacy requirements (Basel Committee on Banking Supervision, 2001) has spurred a discussion on the design of rating systems.[17] Since rating bounce avoidance appears to be a driving factor behind rating dynamics, it should receive more attention in this discussion. The issue is not confined to external rating agencies. Banks might have incentives to manage internal ratings in a way similar to rating agencies.

Acknowledgement

I thank Hergen Frerichs, Jan Krahnen, Lars Norden, Peter Raupach, Mark Wahrenburg, two anonymous referees, the editor, and participants of the 2002 banking workshop in Münster and the 2003 Global Finance Conference in Frankfurt for helpful comments.

References

Altman, E., Kao, D.L., 1992. Rating drift in high-yield bonds. Journal of Fixed Income 2, 15–20.

Altman, E., Kishore, V., 1996. Almost everything you wanted to know about recoveries on defaulted bonds. Financial Analysts Journal 52, 57–64.

Basel Committee on Banking Supervision, 2000. Credit ratings and complementary sources of credit quality information. Bank for International Settlements, Basel.

Basel Committee on Banking Supervision, 2001. Overview of the new Basel capital accord. Bank for International Settlements, Basel.

Batchelor, R.A., 1986. Quantitative v. qualitative measures of inflation expectations. Oxford Bulletin of Economics and Statistics 48, 99–120.

Campbell, J., Lo, A., MacKinlay, C., 1997. The Econometrics of Financial Markets. Princeton: Princeton University Press.

Cantor, R., 2001. Moody's investors service response to the consultative paper issued by the Basel Committee on Banking Supervision and its implications for the rating agency industry. Journal of Banking and Finance 25, 171–186.

Carey, M., Hrycay, M., 2001. Parameterizing credit risk models with rating data. Journal of Banking and Finance 25, 197–270.

Chen, L., Wei, J., 2001. An indirect estimate of transaction costs for corporate bonds. Working paper, University of Michigan, Toronto.

Crouhy, M., Galai, D., Mark, R., 2001. Prototype risk rating system. Journal of Banking and Finance 25, 47–95.

Delianedis, G., Geske, R., 1999. Credit risk and risk neutral default probabilities: information about rating migrations and defaults. Working paper, UCLA.

Economist, 1997. Risks beyond measure. December 13, Economist 375, 70–71.

Ederington, L., Goh, J., 1998. Bond rating agencies and stock analysts: who knows what when? Journal of Financial and Quantitative Analysis 33, 569–585.

Edwards, W., 1968. Conservatism in human information processing. New York: Wiley.

[17] See, for example, Krahnen and Weber.

Fama, E., French, K., 2002. Testing tradeoff and pecking order predictions about dividends and debt. Review of Financial Studies 15, 1–33.

Geske, R., 1977. The valuation of corporate liabilities as compound options. Journal of Financial and Quantitative Analysis 12, 541–552.

Gordy, M., Heitfield, E., 2001. Of Moody's and Merton: a structural model of bond rating transitions. Working Paper, Board of Governors of the U.S. Federal Reserve System.

Hand, J., Holthausen, R., Leftwich, R., 1992. The effect of bond rating agency announcements on bond and stock prices. Journal of Finance 47, 733–752.

Holthausen, R., Leftwich, R., 1986. The effect of bond rating changes on common stock prices. Journal of Financial Economics 17, 57–89.

Kealhofer, S., Kwok, S., Weng, W., 1998. Uses and abuses of default rates. CreditMetrics Monitor, 1st Quarter, 37–55.

Krahnen, J.P., Weber, M., 2001. Generally accepted rating principles: a primer. Journal of Banking and Finance 25, 3–24.

Kuhner, C., 2001. Financial rating agencies: are they credible? Insights into the reporting incentives of rating agencies in times of enhanced systemic risk. Schmalenbach Business Review 53, 2–26.

Lando, D., Skødeberg, T.M., 2002. Analyzing rating transitions and rating drift with continuous observations. Journal of Banking and Finance 26, 423–444.

Löffler, G., 2004. An anatomy of rating through the cycle. Journal of Banking and Finance 28, 695–720.

Merton, R.C., 1974. On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance 29, 449–470.

Moody's Investors Service, 2001. Default and recovery of corporate bond issuers: 2001. Moody's Investors Service, New York.

Moody's Investors Service, 2002. Understanding Moody's corporate bond ratings and rating process. Moody's Investors Service, New York.

Perraudin, W., Taylor, A. P., 2004. Credit risk orderings implicit in ratings and bond market yields. Journal of Banking and Finance, forthcoming.

Schultz, P., 2001. Corporate bond trading costs: a peek behind the curtain. Journal of Finance 56, 677–698.

Standard & Poor's, 2001. Ratings performance 2000. Default, transition, recovery, and spreads. Standard & Poor's, New York.