Do unsolicited ratings contain a strategic rating component? Evidence from S&P*

Christina E. Bannier† Patrick Behr‡ André Güttler§

March 17, 2008

Abstract

This paper examines why, for non-U.S. firms, unsolicited ratings tend to be lower than solicited ratings. Both adverse selection and "strategic rating" arguments such as agency conservatism or blackmailing may be reasonable explanations. Comparing empirical default rates of firms with solicited and unsolicited S&P ratings between January 1996 and December 2006, we cannot reject the adverse selection hypothesis for the total sample. However, focussing on the more opaque sub-sample of banks we find that strategic rating seems to play an important role. Our results are robust to various additional tests, including CreditWatch and outlook information, the use of different default horizons, and of alternative outcome measures.

JEL Classification: G15, G24

Keywords: Unsolicited Ratings, Adverse Selection, Rating Agency Conservatism

---

*We thank Philip Bond, Itay Goldstein, Alistair Milne, Peter Raupach, Michael Roberts, seminar participants at the Deutsche Bundesbank, the Goethe University Frankfurt, the University of Pennsylvania's Wharton School, Standard and Poor's in New York, and participants at the annual meeting of the German Finance Association 2007, and CREDIT 2007 for helpful discussions and comments. We further thank Felix Miebs for excellent research assistance and Standard and Poor's for data support. Financial support by the E-Finance Lab Frankfurt is gratefully acknowledged.

†Commerzbank Professor of Corporate Finance / SME Finance, Department of Banking and Finance, Frankfurt School of Finance and Management, Sonnemannstr. 9-11, 60314 Frankfurt, Germany E-mail: c.bannier@frankfurt-school.de.

‡Department of Finance, Goethe-University Frankfurt, Mertonsr. 17, 60054 Frankfurt, Germany, Phone: +49 69 798-23984, Fax: +49 69 798-28272, E-mail: behr@finance.uni-frankfurt.de. (Corresponding author)

§HCI Endowed Chair of Financial Services, Department of Finance, Accounting and Real Estate, European Business School, International University, Rheingaustr. 1, 65375 Oestrich-Winkel, Germany, E-mail: andre.guettler@ebs.edu.

1 Introduction

Among the most controversial aspects of the credit rating business is the practice of unsolicited ratings. Unsolicited ratings are assessments of credit quality "that credit rating agencies conduct without being formally engaged to do so by the issuer" (IOSCO, 2003). As a consequence, unsolicited ratings do not entail the payment of a rating fee and they are usually not based on private information or other forms of cooperation between rated entity and rating agency.

Credit rating agencies emphasize various benefits of announcing not-requested ratings. Particularly the three biggest credit rating agencies, Moody's, Standard and Poor's (S&P), and Fitch,[1] claim that unsolicited ratings serve investors' interests of a full market coverage, i.e. coverage also with respect to issuers that are unwilling to undergo a time-consuming rating process or to pay a rating fee. The publication of unsolicited ratings also allows - possibly smaller or younger - rating agencies to demonstrate their market knowledge even if they do not (yet) dispose of considerable market shares, thereby increasing competition in the rating market.

While traditionally rating information was requested and paid for by bondholders rather than issuers, such "unsolicited" ratings vanished during the 1970s and 1980s. Due to the low cost of photocopying and the consequential ease in free-riding on single subscribers' rating information, rating agencies changed their business model and requested bond issuers to pay for a rating announcement. In the early 1990s, however, the main credit rating agencies reintroduced the practice of announcing unsolicited ratings. In 2000, the proportion of unsolicited ratings with respect to the total number of oustanding ratings varied between 6% and 27% in industrial countries, depending on rating agency and region (Fight, 2001). In developing countries, unsolicited ratings are even more common, thus corroborating the "business expansion" argument. Interestingly, however, particularly the biggest rating agencies seem to use the concept of unsolicited ratings to penetrate new markets. As such, unsolicited ratings are often criticized as additional barriers to entry into the rating industry by increasing the market power of the already large rating agencies rather than opening up competition.

Recent critique with respect to unsolicited ratings has centered on two main points. First, there is general concern that unsolicited ratings "do not appear to be empirically as favorable as solicited ratings" (SEC, 2005), i.e. credit rating agencies seem to assign lower unsolicited ratings than when asked and paid to do so. Second, the agencies seem somewhat reluctant to announce whether a published rating is commissioned or not (Gasparino, 1996).[2] Given that the information basis should vary significantly between the two types of rating (public and private information for solicited ratings, only public

[1]Moody's, S&P, and Fitch belong to the group of Nationally Recognized Statistical Rating Organizations (NRSROs), whose ratings are permitted by the SEC to be used for regulatory purposes. Currently, there are eight NRSROs. Apart from the above mentioned, biggest ones, these are A.M. Best, Dominion Bond Rating Service, Japan Credit Rating Agency, R&I, and Egan-Jones.

[2]S&P started in 1996 to label unsolicited ratings with "pi" subscripts, at least outside the U.S., referring to "public information". Fitch has announced unsolicited ratings since acquiring Thomson BankWatch's Credit Evaluation in 2000 and refers to these credit gradings as "shadow ratings". Moody's has differentiated between solicited and unsolicited ratings only in the initial assignment announcement, starting this practice in 2001.

information for unsolicited ones),[3] bondholders would reasonably want to know exactly which type of information enters the rating assignment process.

Several incidences have fueled the particular concern that the suspected downward bias of unsolicited ratings may be deliberately generated by rating agencies to blackmail issuers into paying for a (better) solicited rating statement. In 1996, for instance, Moody's was sued by the Jefferson County School District for allegedly having posted negative unsolicited comments on a municipal bond issue that cost the district \$769,000.[4] Similarly, in 1998 Hannover Re, a German reinsurer, was approached by Moody's to subscribe to its rating services. Being rejected,[5] Moody's announced an unsolicited rating that was decreasing strongly over the following years. The company's subordinated debt was even downgraded to junk bond status on March 25, 2003. Despite the fact that S&P and A.M. Best retained their relatively strong assessments of Hannover Re's credit quality, Moody's severe downgrade led to a 16.8% abnormal decline in the company's stock value between March 25 and 26. Similar cases (e.g. Simon Property Group Inc. or Compuware Corp. in the late 1990s) were also purported with respect to Fitch and S&P ratings. Most often, however, the rated issuers eventually agreed to solicit a rating and pay for the agencies' services, in the hope of improving the rating level and thereby reducing borrowing costs (Economist, 2005; Klein, 2004).

Concerns with regard to unsolicited ratings are particularly strongly voiced in Europe and Asia and have led to a vehement debate among market participants and regulators about their use (JCIF, 1999-2001). Generally, the new capital adequacy rules issued by the Basel Committee on Banking Supervision - known as Basel II - allow the consideration of unsolicited ratings for the determination of regulatory capital according to the standardized approach. However, regulators in Japan and Austria currently discuss the exclusion of unsolicited ratings because of the controversial discussions surrounding them. Corroborating these concerns, empirical research has shown that unsolicited ratings seem to influence investment decisions just as strongly as solicited rating information (Behr and Güttler, forthcoming).[6]

This paper conducts a comprehensive study of the presumed downward bias of unsolicited ratings. We differentiate and test between reasons endogenous and exogenous to the rating process for unsolicited ratings' potential downward bias. Endogenous arguments are referred to as "strategic rating" in the following and pertain to the inherent characteristics of the rating assignment process. Apart from blackmailing incentives, also agency conservatism may lead to lower unsolicited ratings according to this line of

[3] Particularly Moody's and Fitch claim that their unsolicited ratings contain also a private information component. Certainly, however, this must be a minor component as otherwise solicited ratings would be obsolete. Also, as the rated entity typically does not provide any "inside" information, the private component should be expected to consist mainly of the agency's historical knowledge about similar firms with respect to business, industry, country, size etc.

[4] The case was dismissed in 1997. The judge ruled that rating statements were opinions protected by the First Amendment.

[5] Hannover Re had already engaged S&P and A.M. Best to rate its credit quality.

[6] Behr and Güttler (forthcoming) analyze stock market reactions to the assignment of an initial unsolicited rating and to subsequent changes of the unsolicited rating. They find significant abnormal returns for both kinds of events, underlining the importance of unsolicited ratings for firms' funding costs.

reasoning. Both endogenous explanations should lead to different solicited and unsolicited ratings for firms with identical credit quality. As a consequence, identically-rated entities with differing solicitation status should show different empirical default rates. According to exogenous arguments, in contrast, the observed downward bias of unsolicited ratings is due to a self-selection of companies into the solicited or unsolicited rating status. It eventually leads to both solicited and unsolicited ratings correctly representing an issuer's credit quality, i.e. observed default rates should not differ with respect to solicitation status.

To the best of our knowledge, our empirical study is the first to use an extensive dataset where information on rating solicitation status - our center of interest - is given directly, which delivers a most uncontaminated database. In order to account for the different lines of critique regarding unsolicited ratings and the fact that they are mainly voiced outside the U.S., we focus on non-U.S. data and split our dataset into industry-specific sub-samples. In this way, we are able to appropriately comment on the different aspects of the current discussion, for instance with respect to the treatment of banks, for whom it is often argued that the rating process is more intricate due to their inherent opaqueness.

In line with an analysis by Gan (2004) on U.S. industrial firms, we find that the empirical default rates of non-U.S. industrial companies do not differ significantly with respect to solicitation status, controlling for rating level, time, and regional effects. This suggests that rating level differences for industrial firms can be explained by factors exogenous to the rating process such as adverse selection and also strengthens Gan's results that were derived from only indirect - and hence rather weak - information on solicitation status. However, for banks we find strong and robust evidence that empirical default rates do differ between firms with unsolicited and those with solicited ratings. Exogenous factors hence cannot (fully) explain the observed rating level differences, which leaves some space for deliberate action on the part of the rating agencies. The inherent opaqueness of banks may thus have forced (excess conservatism) or enabled (blackmailing) rating agencies to announce more conservative, i.e. lower, unsolicited ratings.

The remainder of the paper is organized as follows. Section 2 provides a literature review. Section 3 derives the hypothesis to be tested in the following. Section 4 uses empirical default rates for our sample of non-U.S. firms to examine the theoretical implications. The last section concludes.

2 Related literature

Lately, concerns about unsolicited ratings have sparked a burgeoning literature on this topic. The question whether there is indeed a difference in levels between solicited and unsolicited ratings, is usually answered within a so-called *ex-ante approach*. It compares solicited and unsolicited ratings by regressing the rating level on a vector of control variables such as company and country specific risk estimates and a dummy variable indicating the rating status. Relying on a sample of 595 non-U.S. companies with S&P ratings between 1998 and 2000, Poon (2003) finds that unsolicited ratings are lower (i.e. a significantly negative unsolicited dummy is obtained) for Japanese firms. Poon and Firth (2005) and Van Roy (2006), in contrast, use Fitch rating data. They observe that Fitch's so-called shadow ratings - which are supposedly unsolicited

- are lower than solicited ratings for Asian and international bank ratings, respectively. However, a word of caution is in order regarding Fitch's shadow ratings. As they may be heavily supplemented by company-related private information, it is not entirely clear which type of information actually enters the rating assessment process. The results by Poon and Firth (2005) and Van Roy (2006) should therefore be interpreted with a considerable amount of caution.

While appropriate for detecting a downward bias in unsolicited ratings in the first place, the ex-ante approach is obviously not able to distinguish between the potential causes of this bias. Often, these earlier studies conclude more or less implicitly that the reported downward bias results from deliberate actions by the rating agencies, and, hence, favor the blackmailing hypothesis: by announcing inappropriately low unsolicited ratings, agencies try to pressure firms to solicit their rating, thus succumbing to internal conflicts of interest. Interestingly, this argument is analogous to accusations regarding stock recommendations. Stock analysts in investment banks face a similar problem as rating agencies: on the one hand, their compensation often relies on the analyst's "helpfulness" to the corporate finance division, such that a better recommendation after an IPO might enhance the probability that the company will include this investment bank in the consortium of the next security issuance. On the other hand, an analyst's (external) reputation depends heavily on her forecasting quality. Regarding empirical evidence, Michaely and Womack (1999) find for IPOs that in the month after the quiet period, lead underwriter analysts issue 50% more buy recommendations than do analysts of other investment banks. They also observe significantly inferior short- and long-run performance of lead underwriters' buy recommendations. In the latter case, the difference between the underwriter's and non-underwriter's mean size-adjusted buy-and-hold return accounts for more than 50% for a two year period beginning after the IPO. In a more recent paper, Barber et al. (forthcoming) test the SEC's accusation against several investment banks that analysts' conflicts of interest resulted in a reluctance to downgrade buy-rated stocks during the bear market of the early 2000s. For a sample period from 1996 to mid-2003, they find that independent research firms' buy recommendations outperform those of the investment banks by a daily abnormal return of 3.1 basis points. Partitioning the time frame into the period prior to the market peak in March 2000 and afterwards (the bear market), Barber et al. observe that investment banks' buy recommendations following equity offerings significantly underperform the congruent buy recommendations of independent research firms by 8.7 basis points per day during the bear market.

With respect to credit rating agencies, several theoretical papers have hinted at the particular role that reputation plays. This may level off any blackmailing incentives. Among the earliest and most prominent studies, Ramakrishnan and Thakor (1984) and Millon and Thakor (1985) remarked on the strong dependence of financial intermediaries such as credit rating agencies on their reputation. The fear of a loss of reputation should therefore counterbalance rating agencies' incentives to not act prudently (Cantor and Packer, 1994). In a more recent study, Covitz and Harrison (2003) verify empirically on a U.S. database using credit rating migrations from Moody's and S&P that "rating agencies are motivated primarily by reputation-related incentives". Conflicts of interest, they claim, seem to be well-managed by the agencies. Still, other studies such as by Partnoy (2001, 2006) or Hill (2004) doubt that reputation is the sole dominating objective guiding the agencies' business.

However, given the fundamental role that credit rating agencies play as delegated monitors on capital markets, the industry has voluntarily set up a code of conduct to ensure that rating accuracy is seen to (OICU-IOSCO, 2004, 2007). Following certain guidelines in the rating process appears to be desirable particularly with respect to the rating of opaque borrowers. Due to their complex businesses, it is often argued that especially banks fall into this category (Morgan, 2002; Hirtle, 2006). Yet, the evidence regarding the opaqueness of banks is not unequivocal (Flannery et al., 2004). Still, the credit quality of banks is often seen as an indicator for the financial strength of the whole banking sector and, hence, of one of the backbones of a country's economy. As, furthermore, regulatory requirements of Basel II also apply to interbank lending, an appropriate measurement of banks' credit quality by rating agencies is strongly desired. The question regarding credit rating agencies' motivation is hence particularly pressing with respect to banks' credit ratings.

In contrast to the earlier empirical work that mainly confirmed the suspected level difference between solicited and unsolicited ratings, our study aims at testing explicitly between two different causes of this observation. Additionally to the above mentioned reputation-related arguments, also reasons exogenous to the rating process may lead to the downward bias in unsolicited ratings. Simple adverse selection processes à la Akerlof (1970) may deliver lower unsolicited ratings as borrowers of relatively bad credit quality refrain from soliciting an expensive rating in which both public and - probably unfavourable - private information about the company enter. A simple model delivering a self-selection of borrowers into high-quality debtors with solicited and low-quality debtors with unsolicited ratings has been derived by Bannier and Tyrell (2006). Its contents will be sketched in the following section.

With respect to the methodology employed in this paper, it is obvious that even though the above mentioned ex-ante analysis is helpful in detecting rating level differences, it does not yield any insights on the reasons for the level differences. In order to differentiate between endogenous and exogenous causes, we therefore employ an ex post approach, the so-called *outcome test*. In essence, it compares a prediction to the respective empirical realization.[7] This kind of analysis is also quite common in other areas of empirical research such as the economic analysis of crime.[8] In our study, the outcome test is used to compare default predictions - the ratings - with actual default realizations.

Gan (2004) is the first to apply an outcome test on rating data. Regressing *actual default observations* on a vector of control variables including the rating level and a dummy variable indicating the rating status, she finds no significant results for the unsolicited dummy. This leads her to conclude that exogenous reasons such as adverse selection seem to have driven the rating level differences between solicited and unsolicited ratings of U.S. industrial firms. While being an interesting and intuitive result, her analysis succumbs to serious critique with respect to the data employed - particularly with respect to data on rating solicitation status, which is not observed directly but based

[7]See also the nobel lecture of Becker (1993) who remarks on the outcome test.

[8]See, for instance, the study by Anwar and Fang (2006) who use an outcome test to analyze whether there is racial prejudice in motor vehicle searches.

on two-stage estimations. First, she uses S-3 registration statements of bond issuances. Most of these statements include an item 14 that exhibits estimates of the aggregate rating agency fees and the bond nominal volume for the specific bond issuance. Given the assigned ratings from S&P and Moody's, she then uses approximate rating fee calculations to estimate which of the ratings are unsolicited. This yields a database with very noisy information about the actual rating status.

As we employ S&P rating data for non U.S. firms where the rating status is directly observable and verifiable, our dataset is very clear and uncontaminated in this respect. Furthermore, we focus solely on non-U.S. firms as the concerns about unsolicited ratings are most often voiced outside the U.S. Finally, our dataset allows us to differentiate between different types of firms such as industrials and banks. Our results are therefore not only unique but also allow us to draw much clearer conclusions about the reasons for the observed rating level differences as compared to earlier studies.

3 Derivation of hypothesis

3.1 Strategic rating - Excess conservatism and blackmailing

Explanatory factors that are endogenous in the rating process are subsumed as "strategic rating" arguments in the following. They result from the fact that, contrary to solicited ratings, unsolicited credit gradings are based solely on publicly available information and are not paid for.

The simple lack of private information per se cannot explain any downward bias of unsolicited ratings, unless we could safely assume that private information about an entity's credit quality were always better than public information. As there is no reason why this should be the case, there must be some strategic aspects at work leading to lower unsolicited ratings. One factor may be agencies' general concern with reputation. Unsolicited ratings are generated from a weaker, i.e. more noisy, information base, so that credit rating agencies may fear that unsolicited ratings lead to larger rating "errors" than solicited assessments. Rating errors, however, fall in two categories with different consequences for bondholders as the main users of rating information. A type-I error occurs, if an issuer is assessed as low risk and assigned a high rating, but defaults nonetheless. A type-II error, in contrast, refers to the non-default of a lowly-rated issuer. Given that type-I errors are much more costly to bondholders, rating agencies thriving for a maximum of reputation and reliability would try to minimize the probability of such errors (at the cost of maximizing type-II errors at the same time). Strong concerns for reputation may therefore lead to excess conservatism, i.e. to ratings that are downward-biased. This bias should be the stronger, the noisier the information base of the rating. There should hence be a strong difference in average rating levels between solicited and unsolicited ratings. Furthermore, among the group of unsolicited ratings, the downward bias should be the stronger, the less precise the public information about the rated entity is. As such, we may expect the downward bias of unsolicited ratings to be strongest for very opaque issuers such as, e.g., banks.

Reputation, however, may not be the only element entering a credit rating agency's utility function. Given that rating agencies do not obtain any fee income from announcing unsolicited credit ratings, they may try to use this instrument to "incentivize" issuers into soliciting and, hence, paying for a rating. Issuers will only have an interest in doing so, though, if there is a possibility of rating improvement after solicitation: a firm will request a rating, if the expected decrease in credit costs due to the rating improvement is sufficiently strong to offset the fee payment to the rating agency. However, there is a tradeoff between this financial incentive and concerns of reputation. If solicitation always leads to a huge upwards movement in rating level, bondholders will rationally either disregard the information content of unsolicited ratings completely or take the downward bias into account, thus removing the effect on credit costs. Issuers will then no longer have an interest in requesting and paying for a rating. Consequently, for unsolicited ratings to be incentive compatible in this sense, the downward bias will have to be sufficiently small to uphold the rating agency's reputation, but sufficiently strong to give issuers an incentive to solicit a rating. This mirrors a credit rating agency's tradeoff between reputational and financial objectives and corresponds to a utility function in the sense of Boot et al. (1993). Interestingly, the blackmailing argument may again lead to a stronger downward bias in unsolicited ratings for more opaque issuers. For these, even a relatively strong upgrade upon solicitation may be justified by the agency by alluding to the noisier information base.

Both strategic rating arguments lead to the following consequence: firms of identical credit quality obtain different credit ratings, depending on their solicitation status. I.e. comparing firms with identical credit quality, those who did not solicit a rating will obtain a lower (unsolicited) rating than those who did ask and pay for a (solicited) rating. Stated differently, when comparing issuers with identical rating but different solicitation status, we should find that issuers with unsolicited ratings default less often than issuers with solicited ratings.

3.2 Adverse selection

Apart from the above mentioned endogenous factors, implying "strategic" and, hence, deliberate rating action causing the downward bias of unsolicited ratings, factors exogenous to the rating process may trigger the same result. Given that agency ratings are credible measures of an entity's credit quality, issuers would generally like to improve the rating level in order to save on credit costs. Due to the information asymmetry between debt issuers on the one side and market participants and rating agencies on the other, an issuer may want to disclose private information in a rating solicitation process in order to increase the rating level.

Self-selection will therefore lead issuers, who believe that their unsolicited rating is unfairly low and are certain that private information disclosure will improve the rating level, to commission a rating. The solicitation process then naturally leads to a rating upgrade and a corresponding decrease in credit costs for the issuer. In contrast, issuers who feel that their unsolicited rating correctly reflects their credit quality, will not decide to pay for a solicited rating as this may not improve the rating level, so that the rating fee will not be compensated by a reduction in credit costs. Low-quality issuers will thus remain with their relatively low unsolicited rating.

This general "adverse selection" argument by Akerlof (1970) has also been applied to the context of debt issuers and credit ratings by Bannier and Tyrell (2006). They show in a theoretical model that a credit rating agency's coordinating role on financial markets can strengthen self-selection effects of this type. The underlying reason is that bondholders' investment behavior displays strategic complementarities in that identical investment decisions become favorable.[9] In such a framework, bondholders react strongly (sometimes even overreact) on both solicited and unsolicited rating information and do no longer appropriately account for the differing information base contained in them. In this respect, the rating announcement by an agency "coordinates", i.e. determines almost exclusively, investors' behavior. Adverse selection among debt issuers with respect to solicitation status should therefore be particularly severe if the announcing credit rating agency has a dominant market position, leading to a strong coordination role.

Contrary to the implications of strategic rating, adverse selection should lead to ratings always correctly reflecting an issuer's credit quality, irrespective of the solicitation status. We may hence test between the two explanatory lines (endogenous vs. exogenous) by comparing the ex-post realized default probabilities from companies with unsolicited and solicited ratings within each rating class. We thus phrase the hypothesis that empirical default rates of identically-rated companies do not differ depending on their solicitation status (solicited / unsolicited). This adverse selection hypothesis will be tested in the subsequent section.

4 Empirical test of the adverse selection hypothesis

4.1 Descriptive statistics and univariate analysis of default frequencies

Our data set comprises S&P data for solicited and unsolicited ratings for the time period January 1996 to December 2006. The sample period starts in January 1996 because S&P introduced the distinction between solicited and unsolicited ratings only then. We did not include Moody's or Fitch's rating data as these agencies do not reliably disclose the solicitation status of the announced ratings. Rating data were extracted from S&P's credit ratings database provided by Wharton Research Data Services (WRDS). We translate the original ratings into numerical ratings which are scaled from 1 (AAA) to 18 (CC to C).[10] All data on unsolicited ratings were manually cross-checked with S&P's RatingsDirect database. For the robustness tests we additionally included rating outlooks and CreditWatch entries to consider all available information regarding an issuer's creditworthiness.

For the tests on empirical default rates, we record whether a company with a given rating defaulted in the following one-year period, which we will refer to as realization period. For example, using firm ratings per year-end of 1996, we record whether or not these firms defaulted in 1997. Afterwards, we use all ratings per year-end of 1997 and check whether these companies defaulted in 1998, etc. Our sample thus covers the whole non-U.S. universe of S&P ratings from January 1996 to December 2006 and the corresponding default data. A default is registered upon the occurrence of a D

[9] Strategic complementarities refer to a situation where it is the more attractive to choose a specific strategy the larger the number of other investors who choose the same strategy.

[10] This method follows Fenn (2000).

(regular or full default), SD (selected default[11]) or R rating in the realization period. An R rating indicates cases in which financial companies were regulated by national supervisory bodies. As regulated companies cannot freely decide to continue their debt repayments, we treat them as defaults. We cross-checked all recorded defaults with S&P's annual default reports.

Table 1 provides a descriptive overview of the sectoral and regional distribution of our sample. Panel I shows the sample's sectoral distribution. It is further subdivided into firm-year observations of companies that did not default in the subsequent one-year realization period ("survivors") and companies that defaulted ("defaulters"). Also, the table distinguishes between firms with solicited and unsolicited ratings. Overall, the sample contains 26,642 firm-year observations. Of these, 59.6% come from financial firms[12] and 40.4% from non-financial firms. Among financial firms, 63.1% possess solicited ratings and 36.9% hold unsolicited ratings. Particularly insurance companies seem to be affected from unsolicited ratings, with roughly equal proportions of solicited and unsolicited ratings. Among non-financial firms, in contrast, the proportion of companies with solicited ratings is much higher (89.3% versus 10.7% unsolicited ratings). With respect to defaults, we find that 56 financial firms with a solicited and 23 with an unsolicited rating defaulted in the realization period, representing 0.56% respectively 0.40% of the firm-year observations. Among non-financial firms, a much larger number of defaults is observed: 154 (1.62%) with a solicited and 6 (0.52%) with an unsolicited rating defaulted. In sum, the sample contains 239 defaults of S&P-rated firms in the period January 1996 to December 2006. Of these, 180 were defaults of companies with a solicited rating with 154 coming from non-financial firms.

The regional distribution, which is depicted in Panel II, shows that the majority of our data comes from Europe/Middle East/Africa, followed by Asia Pacific. Interestingly, the Asia Pacific region displays the highest proportion of unsolicited ratings: roughly 50% of all firm-year observations are unsolicited ratings. In Europe/Middle East/Africa, it is only 38.5% unsolicited ratings, while in the other three regions very few unsolicited ratings are observed (proportion between 5% and 12%). With respect to defaults, we find the highest default frequency among firms with solicited ratings in Latin America (3.98%) followed by Europe/Middle East/Africa (0.68%). On the other hand, the highest proportion of defaults of companies with unsolicited ratings is observed in the Asia Pacific region (0.93%).

Table 2 shows the rating distribution per year. We subdivide the sample into financial and non-financial firm-year observations[13] and in a second step into survivors and defaulters with respect to the one-year realization period following the year-end rating observation. We then calculate mean ratings per year by using the numerical rating scale ranging from 1 (AAA) to 18 (CC to C) and by taking the last valid rating of each firm in a given year. As the next step, we compute the mean over all rating observations for each year, separately for survivors and defaulting firms with solicited and unsolicited ratings.

[11]SD ratings were introduced by S&P in 1999 and reflect cases where the respective company stopped its debt service on some but not all of its outstanding obligations.

[12]The category of other financials comprises asset managers, brokerage firms, and the like.

[13]The rating distribution is not split into all sub-samples due to space restrictions. Results are available from the authors upon request.

With respect to the time-series dimension of our sample, we observe that among the non-defaulters, the number of firms with solicited ratings has more than doubled, while the number of firms with unsolicited ratings has developed less steadily. For financial companies, the number of unsolicited ratings has increased strongly from 50 in 1996 to 777 in 2001, but has declined slightly in recent years. For non-financial firms, the number has remained more or less constant at about 150 from 2000 on. Due to these developments, the ratio of unsolicited to solicited ratings for financial firms has been varying over the years until it recently declined (from a maximum of 85.2% in 1999) to roughly one third, while the ratio is much steadier and much lower for non-financial firms at about 12%. We also find that mean solicited ratings of non-defaulting firms slightly deteriorated over the years (by 2-3 rating notches), which may be a sign of a generally decreasing credit quality or of a hardening in S&P's rating standards (Blume et al., 1998). With respect to the mean difference between unsolicited and solicited ratings, we observe a positive spread for all subgroups. The difference is particularly strong (almost 4 rating notches) for financial firms that did not default within the observation period. This first indication with respect to a downward bias in unsolicited ratings will be tested further in the following.

Table 3 presents univariate test results for the observed rating level differences. Panel I provides evidence for the total sample. Overall, the mean solicited rating for a non-defaulting firm is 7.20, while the mean unsolicited rating is 9.53. This rating difference of 2.32 numerical rating notches is highly significant (on the 1%-level) both according to the t-test and the Wilcoxon signed ranks test. For defaulting firms, in contrast, the mean rating difference is only 0.24 notches and not significant. In order to delve deeper into the significant rating difference for non-defaulting companies, Panel II runs the same test on the sub-samples of non-defaulting financial firms and non-financial firms. Both the difference for financial firms (3.68 notches) and for non-financial firms (0.90 notches), that have already been alluded to in Table 2, are highly significant. Obviously, there is a strong difference between average solicited and unsolicited ratings that seems to be driven by financial firms of relatively high quality, i.e. that did not run into default.

We further checked whether these results remain stable after applying historical default rates instead of the rating level. The former measure of default risk may be more appropriate because the relationship between default risk and rating class is highly non-linear.[14] To reduce the effect of annual variations in historical default rates we use S&P's long-year average over the years 1981-2005 (S&P, 2006a) and smooth these averages exponentially. While differences in default frequencies of non-defaulting financial firms are still significantly positive on the 1%-level according to both the t-test and the Wilcoxon signed ranks test, these differences become negative for non-financial firms.[15] For financial firms, however, the results are even more pronounced than when using rating levels: the mean default frequency of surviving financial firms with solicited ratings equals 0.46%, and almost triples (to 1.28%) for respective observations with unsolicited ratings.

Summing up, our univariate results show that the observed rating level difference is

[14] In our robustness tests in Section (4.3.2) we come back to this issue and explore it even further.

[15] Results are not displayed but available upon request.

only significant for non-defaulting firms, but not for defaulting companies. Results are particularly pronounced for financial firms, less so for non-financial companies. Lacking any default event in the realization periods, however, there is no obvious reason for this rating difference. Still, given that defaults are relatively rare events, there may be some other factors lying at the heart of the observed rating level deviation. This may be a first hint that exogenous effects such as adverse selection may not be a sufficient explanation.

4.2 Multivariate analysis of empirical default rates

In this section we test the adverse selection hypothesis by using multivariate regression approaches. We refer to this analysis as ex-post or outcome test, as we relate estimates of default probability, i.e. credit ratings, for companies with solicited and unsolicited ratings to ex-post realizations, i.e. default or survival. The outcome test is hence used to evaluate the reasons for ex-ante rating differences.

Our basic multivariate analysis is similar to the one undertaken by Gan (2004). To test for differences in the empirical default rates we employ a pooled logit regression model. We define $default_i$ as a dummy variable indicating whether company i defaulted in the realization period (one for default, zero otherwise), and $default_i^*$ as the unobserved linking variable, which is continuous and ranges over the set of real numbers. Hence, we estimate

$$default_i^* = \alpha + \beta_1 \cdot rating_i + \beta_2 \cdot unsolicited_i + \gamma \cdot D + \epsilon_i \tag{1}$$

with

$$default_i = \begin{cases} 1 & if \quad default_i^* & > 0, \\ 0 & if \quad default_i^* & \leq 0, \end{cases}$$

where *rating* is the rating level expressed as numerical value for company i between 1 (AAA) and 18 (CC to C) that is followed by the realization period in which the default status is recorded. *Unsolicited* represents a dummy for the rating status (1 for unsolicited, 0 for solicited ratings). In addition, γ is a vector of coefficients for a vector of dummy variables, D, capturing nine realization periods, four business sectors, and four regions. The year-dummies are necessary to control for varying macroeconomic environments that affect firms' default probabilities (Nickell et al., 2000). Since rating agencies claim that their rating assessments are consistent over all covered business sectors and regions, in principal, no further controls should be necessary. However, particularly in light of our univariate analysis that delivers differing results for financial and non-financial firms, we control for business sector differences in the default frequencies by classifying our observations into five different sub-samples: banks, insurance companies, other financial firms, utilities, and corporates. The latter serves as the reference group. Additionally, since our descriptive analysis reveals huge regional differences in the prevalence of unsolicited ratings, we use region dummies for Asia Pacific, Australia/New Zealand, Europe/Middle East/Africa, Latin America and Canada. Again, the latter serves as the reference group. We do not control for any company-specific variables such as size, return on equity, liquidity, and the like. The reason is that we include the company rating which is an aggregate measure for all company-specific variables.

If unsolicited ratings were inadvertently too low we would not expect any difference in the outcome, i.e. the default frequency, between companies with solicited and unsolicited ratings. As a reason, consider that if unsolicited ratings properly reflected the default risk of a firm, the empirical default rates should not differ significantly from the default rates of firms with solicited ratings. There is no obvious reason why two groups of firms with the same rating - one group with solicited the other with unsolicited ratings - should have a significantly different default realization unless one of the two is intentionally rated too low. Rather, adverse selection arguments should lead firms with positive internal information to order a paid rating while firms with weak internal information should stick to the unsolicited rating. Therefore, from an ex-post point of view, there should be no differences between unsolicited and solicited ratings even though the unsolicited ratings show lower levels than the solicited ratings. However, if unsolicited ratings are too low because of endogenous reasons within the rating process then they should be associated with lower empirical default rates. In this case the unsolicited dummy in the regression model would be significantly negative.

Table 4 contains the results of the logit model with one-year realization periods. Regression model I contains the results for the whole sample.[16] As can be seen, firms with a bad rating default significantly more often. The unsolicited dummy is significantly negative, which indicates on first sight that endogenous reasons such as excess conservatism or blackmailing seem to be driving the differences between solicited and unsolicited ratings. Year dummies are significantly positive for the years 2000 and 2001. This indicates that default rates are higher in the years 2001 and 2002 than in the reference year 1997. This is plausible since these two years experienced a relative large number of defaults. Regarding the business sector dummies, other financial companies experienced significantly higher default rates than corporates. Additionally, Europe/Middle East/Africa shows lower default rates than Canada.

Regression model II contains the results for a sub-sample including only financial firms. In this model, we employ banks as the reference group. We are forced to omit the region dummy Australia/New Zealand and the year 1997 from the regression model because both categories do not feature any default observations. We again find a negative unsolicited dummy, which is significant on the 1% level. For non-financial firms, on the other hand, we fail to detect statistical significance of the unsolicited dummy. This finding, which is not shown in the table, is in line with the univariate results. It also serves as a sign of caution with respect to interpreting the results of model I. Probably the cleanest way of stating our results would therefore be to say that for the total sample adverse selection alone is not sufficient to drive the observed downward bias in unsolicited ratings.

As pointed out above, banks arguably are more difficult to rate because of their inherent opacity (e.g., Morgan, 2002). We therefore further subdivide the sample of financial firms into banks and insurance companies and run our multivariate analysis. Regression model III provides the results for the sub-sample of banks only. In addition to the region

[16]We tested for autocorrelation of the residuals, but did not detect any. Hence, we only use a heteroskedasticity consistent covariance matrix estimator (White, 1980).

dummy Australia/New Zealand and the year dummy of 1997, we also exclude the year dummy of 2003. Again, we obtain a negative unsolicited dummy, which is significant on the 1% level. For the insurance companies and the sub-sample of other financial firms we do not find any significant effects of the unsolicited dummy.[17] This suggests that the observed effect is mainly driven by the banks in our sample. In the remainder of the paper we thus primarily run robustness tests for the sub-sample of banks.

Summing up our results, we may state that the existing rating level differences seem not to be fully explainable by adverse selection arguments, but rather that endogenous reasons, i.e. deliberate actions on the side of the rating agency, contribute to their explanation. Endogenous effects seem to be particularly strong for the sub-sample of banks, while strategic rating is obviously the least pronounced for industrial firms.

Note that although we use longitudinal data we do not employ a (logit) panel model, but rather a pooled regression approach. As in the non-linear case fixed effects models include only those observations with a change in the dependent variable over time, this would have reduced our sample to the 239 companies with default events. From this point of view only random effect models seem to be feasible. However, it is not very plausible for our sample to assume the observations to be random draws from a large population because we cover the whole non-U.S. market. We therefore decided against using a panel approach.[18]

4.3 Robustness tests

As our empirical results may depend on the choice of the employed variables or the underlying model assumptions, this section provides a series of tests to diagnose the robustness of our results. We first use rating modifiers (outlook/CreditWatch entries) as a further refinement of the risk prediction expressed by the rating, then we employ default frequencies as an explanatory variable instead of the rating level. Both robustness tests are done for the whole sample, the sample of financial firms, and for banks only. We then concentrate on the sub-sample of banks and extend the realization period from a one-year to a three-year period to better match the long-term through-the-cycle character of credit ratings. Finally, we employ the bank-specific z-score as an alternative outcome measure for banks.

4.3.1 Rating modifiers

Further to rating changes we now also use rating outlooks and CreditWatch entries. S&P assigns a rating outlook to all long-term debt issuers to assess the potential of a future rating change. Outlooks have a rather long-term horizon that typically spans two years. They reflect trends or risks with less certain implications for credit quality. On the other hand, companies appear on the CreditWatch when a certain event has occurred or is expected by the analysts and further information is necessary to assign a

[17] Results are not displayed here but are available upon request.

[18] As an additional robust test we estimated a random effects logit model. The results remain qualitatively the same as in the pooled logit approach. The unsolicited dummy is significantly negative for the sub-sample of banks.

new rating. Usually, these reviews are done within 90 days, unless the outcome of the respective event is still unclear (S&P, 2006b). There are several empirical studies which show that CreditWatch and outlook information provides timely information to market participants about the company's creditworthiness.[19] For 4,849 firm-year observations we observe positive or negative outlooks, and for 1,439 firm-year observations positive or negative CreditWatch entries. We neglect stable outlook information and CreditWatch entries with the developing status because there are only very few of them and because it is unclear what kind of information stable outlooks and developing CreditWatch entries reveal.

S&P seems to employ negative implications more often. For example, whereas solicited firm-year observations are associated in 17.7% of all cases with a negative outlook this ratio is only 6.4% for positive outlooks. On the other hand, outlook and CreditWatch information is rarely used for unsolicited ratings. Only 1.31% (0.34%) of all firm-year observations with unsolicited ratings carry a negative (positive) outlook.

To build the rating modifiers into our regression analysis we include dummy variables for positive and negative outlook and CreditWatch observations. Otherwise, the three previous regression models remain unchanged. Table 5 contains the results. The negative outlook and CreditWatch dummies are significantly positive in regression model I consisting of the whole sample. Thus, observations with negative outlook or a negative CreditWatch entry default more frequently ceteris paribus. On the other hand, through the refinement of the rating level the unsolicited dummy is not different from zero anymore. In regression model II that comprises the financial firms, only the negative CreditWatch entry dummy is significantly positive and the unsolicited dummy remains significantly negative on the 5%-level. In the case of the sub-sample of banks - regression model III - we have to remove both dummies with positive indications due to lacking defaults in these categories. The negative outlook and CreditWatch dummies are not significant anymore, but the unsolicited dummy is still significantly negative on the 5%-level. Hence, the previous results for banks are robust to the inclusion of outlook and CreditWatch information. Again, we checked the unsolicited dummy for the sub-samples of insurance companies and other financial firms and found them not to be significantly different from zero.

4.3.2 Default frequencies

The use of the rating level in our regression analysis implies a linear relationship between the rating and the default probability expressed by it. However, in reality the relationship between the rating and the default probability is highly non-linear. Whereas only small increases in default rates are associated with a rating downgrade of highly rated issuers, say from AAA to AA-, default rates increase significantly for non-investment grade rated companies.[20] To test the robustness of our results we hence substitute the rating level in our regression models with the long-term average default frequency.

[19]For instance, refer to Holthausen and Leftwich (1986) and Hand et al. (1992). Additionally, the work by Boot et al. (2006) provides a theoretical framework pointing out the importance of CreditWatch entries.

[20]To account for this non-linearity Jorion et al. (2005) introduce a dummy variable to distinguish between investment and non-investment grade issuers.

We use S&P's average default rates per rating class over the years 1981 to 2005 (S&P, 2006a). As empirical default frequencies tend to be erratic, we smooth them exponentially in order to obtain strictly monotonically increasing default rates for deteriorating credit ratings. We then run all regressions for the whole sample, the financial firms and the banks again.

The results for all three regression models are displayed in Table 6. As expected, the default frequency itself is highly significant in all regression models, indicating that firms with a higher empirical default frequency default significantly more often. Confirming our earlier results, the unsolicited dummy is not significantly different from zero for the whole sample. Additionally, the dummy for the financial firms also looses significance. For the sub-sample of banks, however, the unsolicited dummy is still significantly negative. This underlines our finding that the assignment of unsolicited ratings to banks is subject to adjustments by the rating agency.

4.3.3 Three-year realization periods

So far we used one-year realization periods and the resulting default observations as dependent variables. It is, however, an established fact that external credit ratings intend to "look through the business cycle" and are thus long-term estimates for default risk (Gordy and Howells, 2006). Although it is not known how rating agencies define the properties and length of this business cycle, we employ overlapping three-year instead of one-year realization periods in order to match S&P's rating approach more closely.[21]

As a consequence we lose the last two years of rating observations, that is, we only include all rating information up to 2003 to have the period 2004-2006 as the last realization period. As the use of overlapping realization periods might cause autocorrelation among residuals we carefully test for this. For all regression models presented in this sub-section we find strong autocorrelation. Thus, we use an heteroskedasticity and autocorrelation consistent covariance matrix estimator (Newey and West, 1987) with maximum lag size of two periods. The number of lags is chosen according to the maximum number of overlapping realization periods.

This robustness test focuses on the sub-sample of banks because the results for the other industry groups are not robust to the change of the underlying model assumptions. The most striking difference to the case of the one-year realization period is that the number of employable defaults increases. While we could only employ 39 defaults for the sub-sample of banks in case of the one-year realization period we can now make use of 79 defaults.

We estimate three different pooled logit models to see whether our results from the basic analysis remain stable. Table 7 contains the results. Regression model I repeats the basic analysis with the rating level as explanatory variable. The unsolicited dummy is significantly negative on the 1%-level. The explanatory power of the model measured in terms of the adjusted McFadden R^2 increases from 0.322 (one-year realization periods) to 0.373. Regression model II also includes the outlook and CreditWatch entry dummies. In contrast to the basic analysis, the higher number of defaults enables the use of the

[21] We have also experimented with two-year and four-year realization periods. The results are not notably different from the results for the three-year default horizon.

positive outlook dummy as well. The main result with respect to the unsolicited dummy remains the same as in model I. By using the default frequency instead of the rating level we find a significantly negative unsolicited dummy on the 1%-level (model III). Thus, for both regression model II and III results are stronger than in the first robustness check in section 4.3.2 where we only found significance on the 5%-level.

These results suggest that the use of longer-term default horizons indeed captures the rating approach used by S&P better and that our results for the sub-sample of banks are robust to fundamental changes of our initial model assumptions.

4.3.4 The bank specific z-score as an alternative outcome measure

As a final robustness test we substitute our dependent variable, i.e. default observations of subsequent realization periods, with an alternative outcome measure. Again we only focus on the bank sub-sample. The alternative outcome measure employed is the z-score for banks. It measures an individual bank's risk and is defined as the return on average assets plus the capital-asset ratio divided by the standard deviation of asset returns. The z-score combines accounting measures of profitability, leverage and volatility. Specifically, z indicates the number of standard deviations that a bank's return on assets has to drop below its expected value before equity is depleted and the bank is insolvent (Roy, 1952; Hannan and Hanwick, 1988).

The data for the computation of the z-score are extracted from Compustat. We calculate the standard deviation of asset returns of each bank over a four-year horizon (Laeven and Levine, 2006). For instance, we compute the z-scores as of year-end 1997 by making use of the volatility of return on assets over the four years 1994-1997. We cover both the one-year and the three-year horizon. In particular, to compare the default risk prediction expressed by the rating with the risk realization expressed by the z-score, we first record the end-of-year rating of a given bank and use the bank's z-score in year t+1 (respectively t+3 for the three-year horizon). Hence, we record the year-end rating in 1996 and then map the z-score of the bank per year-end 1997 (year-end 1999 for the three-year horizon) to it. This z-score then becomes our dependent variable.

As we do not have accounting data for all banks, our sample is considerably reduced to 1,689 (1,306) firm-year observations in case of the one-year (three-year) realization period. For both realization periods we estimate three regression models. Model I includes only the rating level as explanatory variable, model II refines the rating level by making use of outlook and CreditWatch information, and model III uses the empirical default frequency instead of the rating level as explanatory variable. In all models we employ regional and time control variables. Note that we are not estimating pooled logit regression models anymore, but pooled OLS regressions because the dependent variable is now continuous. As in case of the previous robustness check we find for all regression models of this sub-section strong autocorrelation. Thus, we use an heteroskedasticity and autocorrelation consistent covariance matrix estimator (Newey and West, 1987) with maximum lag size of two periods. The number of lags is chosen according to the maximum number of overlapping realization periods for the three-year horizon.[22]

Panel I of Table 8 shows the results for the one-year horizon. We see that the unsolicited

[22] To ensure consistency, we use the same number of lags for the one-year horizon.

dummy in model I has a positive and significant coefficient. This illustrates that banks with unsolicited ratings have significantly higher z-scores than banks with solicited ratings. As higher z-scores reflect less risk, this result underlines our previous findings that unsolicited ratings for banks seem to be intentionally too low. The coefficient is not only statistically significant but its size also suggests a high economic significance. Furthermore, we see that the rating level coefficient is negative and significant. This illustrates that companies with higher numerical ratings (higher risk) have lower z-scores in the realization period. This makes intuitively sense and is in line with the previous findings. The results remain robust if we include rating modifiers to refine the rating level (model II). The unsolicited dummy is still significantly positive and the rating level coefficient significantly negative. Additionally, we find that the negative outlook dummy is significantly negative, implying higher risk for banks with a negative outlook. In model III we use the empirical default frequencies instead of the rating level as explanatory variable. Now we see that the unsolicited dummy becomes insignificant.

The results for the three-year realization periods are shown in panel II. The unsolicited dummy is significantly negative in model I and model II, and even in model III we detect statistical significance on the 10% level. Hence, they confirm the results for the one-year realization period as well as the results we obtained when we used default observations as dependent variable.

5 Conclusion

In this study, we analyze possible reasons for rating level differences between unsolicited and solicited ratings. A potential downward bias of unsolicited ratings may be due to reasons purely exogenous to the rating process such as adverse selection or may stem from strategic effects inherent in the rating process. Based on a data set of non-U.S. firms in the time period January 1996 to December 2006, we test the adverse selection hypothesis by comparing ex-post default rates of firms with unsolicited and those with solicited S&P ratings. For non-financial firms we cannot reject self-selection as a dominating cause, as empirical default rates seem to be adequate for both rating types. This result is in line with the finding of Gan (2004) for a sample of U.S. industrial firms. For these companies, factors exogenous to the rating process hence seem to play a major explanatory role. The critical assessment of unsolicited ratings in general is therefore not necessarily fully warranted.

However, for the sub-sample of banks our results indicate that unsolicited ratings are not only lower than solicited ratings, but that they are too low given the empirical default rates as well as their z-scores. This new result suggests that for banks the observed differences in solicited and unsolicited rating levels seem to be - at least partly - driven by factors endogenous in the rating process. This further seems to be the case not only in Asia, where critical voices about unsolicited ratings are loudest, but generally in other regions outside the U.S. In fact, we did not find any peculiarities associated with Asian firms in our data.

Since we do not find a consistent pattern of too low default rates across all business sectors, we are reluctant to accept the frequently uttered blackmailing accusation associated with unsolicited ratings of U.S. rating agencies as fully plausible. Rather, our results for non-U.S. banks may also be caused by excessive rating agency conservatism, particularly as banks are perceived to be more opaque than other firms. Still, we may state that the different treatment of solicited vis-à-vis unsolicited ratings is to the detriment of the rated banks. The fact that unsolicited and solicited rating levels for banks relate to different empirical default rates and as such are not fully comparable is also not desirable from the agencies' point of view.

The ultimate question is, of course, what are the economic costs of too low unsolicited ratings and who bears them. Naturally, the rated firms themselves suffer as the assigned ratings determine their funding costs. If the rating level is too low given the true creditworthiness, this unnecessarily drives up funding costs. This argument is even strengthened by the regulatory use of unsolicited ratings. It is questionable, though, whether investors are adversely affected by the bias in unsolicited ratings as well. It may even be conceivable that investors do not object to the downward bias if they display the same degree of risk aversion as the rating agencies. Still, any bias between solicited and unsolicited ratings creates a framework in which the two types of ratings are no longer comparable as they refer to different levels of ex-post default probability. In the long run, this should not leave the agencies' reputation unaffected.

The often-voiced critical view on unsolicited ratings of banks hence seems to be justified. Policymakers should therefore carefully evaluate the use of unsolicited ratings for regulatory purposes. As endogenous effects in the rating process seem to play a major role for the reported lower unsolicited rating levels of banks, our results should be taken as a significant step in investigating the rating process more closely.

Table 1: Sectoral (Panel I) and regional (Panel II) sample distribution. We use S&P rating and default data for the time period January 1996 to December 2006. S&P rating data are extracted from the S&P credit ratings database via WRDS. Data on unsolicited ratings are cross-checked with RatingsDirect, default data are cross-checked with S&P's annual default reports.

Panel I: Business sector	Banks	Insurance	Other financials	Corporates	Utilities	Sum
<u>Survivors:</u>						
Solicited rating	4,825	4,372	732	7,561	1,972	19,462
Unsolicited rating	1,295	4,485	14	1,109	38	6,941
<u>Defaulters:</u>						
Solicited rating	31	3	22	131	23	210
Unsolicited rating	8	11	4	6	0	29
Sum	6,159	8,871	772	8,807	2,033	26,642

Panel II: Region	Asia Pacific	Australia/ NZ	Canada	Europe/Mid East/Africa	Latin America	Sum
<u>Survivors:</u>						
Solicited rating	2,634	2,410	2,062	10,398	1,958	19,462
Unsolicited rating	2,374	113	178	4,029	247	6,941
<u>Defaulters:</u>						
Solicited rating	22	5	34	71	78	210
Unsolicited rating	22	0	0	5	2	29
Sum	5,052	2,528	2,274	14,503	2,285	26,642

Table 2: Mean year-end ratings for all surviving and defaulting firms according to the one-year realization period following the rating announcement. We use S&P rating and default data for the time period January 1996 to December 2006. S&P rating data are extracted from the S&P credit ratings database via WRDS. Data on unsolicited ratings are cross-checked with RatingsDirect, default data are cross-checked with S&P's annual default reports. Ratings are translated into a numerical rating scale ranging from 1 (AAA) to 18 (CC to C). Panel I shows the mean numerical rating level at a certain point in time for all surviving (Panel Ia) and defaulting (Panel Ib) financial firms with solicited and unsolicited ratings. Panel II provides the respective information for non-financial firms.

	Rating per year-end of											
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Sum	Mean
Panel I: Financials												
Ia) Survivors:												
Number of solicited ratings	655	811	897	914	975	1,049	1,068	1,111	1,183	1,266	9,929	
Mean solicited rating	4.76	5.19	5.33	5.53	5.54	5.67	6.09	6.43	6.51	6.47		5.85
Number of unsolicited ratings	50	616	702	779	760	777	674	545	478	413	5,794	
Mean unsolicited rating	10.38	9.57	9.55	9.35	9.26	9.32	9.63	9.90	9.80	9.57		9.53
Ib) Defaulters:												
Number of solicited ratings	0	9	5	2	12	17	8	1	0	2	56	
Mean solicited rating	0.00	13.44	16.40	16.00	13.58	14.53	14.88	18.00	0.00	15.00		14.50
Number of unsolicited ratings	0	0	3	7	5	2	4	1	1	0	23	
Mean unsolicited rating	0.00	0.00	15.67	15.57	13.80	14.00	15.75	18.00	17.00	0.00		15.26
Panel II: Non-financials												
IIa) Survivors:												
Number of solicited ratings	495	632	751	824	914	1,030	1,085	1,167	1,285	1,350	9,533	
Mean solicited rating	6.98	7.80	8.40	8.60	8.46	8.39	8.64	8.87	9.08	9.30		8.61
Number of unsolicited ratings	0	45	71	114	150	150	149	140	170	158	1,147	
Mean unsolicited rating	0.00	9.27	9.55	10.25	10.23	10.31	10.32	8.98	8.54	8.33		9.51
IIb) Defaulters:												
Number of solicited ratings	1	1	20	12	30	59	18	3	5	5	154	
Mean solicited rating	17.00	14.00	15.50	15.75	15.57	15.32	15.61	15.67	15.00	14.20		15.42
Number of unsolicited ratings	0	0	0	0	1	2	1	1	1	0	6	
Mean unsolicited rating	0.00	0.00	0.00	0.00	15.00	16.00	18.00	14.00	17.00	0.00		16.00

Table 3: Univariate test results of rating level differences between solicited and unsolicited ratings. We use S&P rating and default data for the time period January 1996 to December 2006. S&P rating data are extracted from the S&P credit ratings database via WRDS. Data on unsolicited ratings are cross-checked with RatingsDirect, default data are cross-checked by using S&P's annual default reports. Ratings are translated into a numerical rating scale ranging from 1 (AAA) to 18 (CC to C). We conduct the analysis for the total sample (Panel I) and for firm-year observations of companies that survived the one-year realization period (Panel II). The second panel is subdivided into financial and non-financial firms. The third column provides the mean rating before the realization period and the forth column the respective difference between firm-year observations of companies with unsolicited and solicited ratings. The fifth and sixth columns show the p values of the t-test and the Wilcoxon signed ranks test.

	Observations	Mean rating	Difference	t-test p value	Wilcoxon p value
Panel I: Total sample					
Survivors:					
Solicited rating	19,462	7.20			
Unsolicited rating	6,941	9.53	2.32	<.001	<.001
Defaulters:					
Solicited rating	210	15.18			
Unsolicited rating	29	15.41	0.24	0.6059	0.8310
Panel II: Survivors					
Financial firms:					
Solicited rating	5,794	5.85			
Unsolicited rating	9,929	9.53	3.68	<.001	<.001
Non-financial firms:					
Solicited rating	1,147	8.61			
Unsolicited rating	9,533	9.51	0.90	<.001	<.001

Table 4: Ex-post analysis of empirical default rates using a pooled logit regression approach. The dependent variable takes the value one in case of a default and zero if the firm has survived the next one-year realization period. The independent variables are: unsolicited signifies a dummy for rating solicitation (1 for unsolicited ratings, 0 otherwise), rating indicates the rating level in the estimation period expressed as numerical value for company i between 1 (AAA) and 18 (CC to C) that is followed by the realization period in which the default status is recorded. We further use sectoral and regional classifications, and year dummies. In regression model I (II) the year 1996, corporates (banks) and Canadian firms serve as reference. In regression models II and III we omitted the respective variable if we did not observe any default in a certain region or year. We use a heteroskedasticity consistent covariance matrix estimator (White, 1980).

	(I)		(II)		(III)	
Independent	Total sample		Financials		Banks	
variable	Coefficient	p value	Coefficient	p value	Coefficient	p value
Intercept	-13.485	<.001	-10.866	<.001	-10.476	<.001
Unsolicited	-1.142	<.001	-1.209	<.001	-1.433	0.002
Rating level	0.647	<.001	0.607	<.001	0.604	<.001
Bank	-0.267	0.223				
Insurance	-0.378	0.262	-0.172	0.610		
Other financial	1.552	<.001	1.704	<.001		
Utilities	0.467	0.096				
Asia Pacific	-0.333	0.208	-0.467	0.338	-1.447	0.074
Australia/New Zealand	-0.248	0.622				
Europe/Mid East/Africa	-0.522	0.027	-0.971	0.049	-1.963	0.013
Latin America	0.308	0.185	-0.199	0.679	-0.538	0.495
Year 1997	1.276	0.206				
Year 1998	1.586	0.103	-0.603	0.269	0.134	0.832
Year 1999	1.234	0.207	-0.399	0.470	-0.436	0.561
Year 2000	2.175	0.024	0.357	0.466	0.807	0.145
Year 2001	2.460	0.010	0.213	0.671	0.416	0.452
Year 2002	1.386	0.155	-0.325	0.556	-0.174	0.768
Year 2003	-0.354	0.733	-2.394	0.002		
Year 2004	-0.089	0.931	-2.792	0.013	-1.712	0.125
Year 2005	-0.162	0.875	-1.981	0.015	-1.747	0.113
McFadden adj. R^2	0.433		0.404		0.322	
Observations	26,642		15,802		6,159	

Table 5: Ex-post analysis of empirical default rates with additional outlook and CreditWatch information. The dependent variable takes the value one in case of a default and zero if the firm has survived the next one-year realization period. The independent variables are: unsolicited signifies a dummy for rating solicitation (1 for unsolicited ratings, 0 otherwise), rating indicates the rating level in the estimation period expressed as numerical value for company i between 1 (AAA) and 18 (CC to C) that is followed by the realization period in which the default status is recorded, negative outlook (1 for negative outlook, 0 otherwise), positive outlook (1 for positive outlook, 0 otherwise), negative CreditWatch (1 for negative CreditWatch entry, 0 otherwise), positive CreditWatch (1 for positive CreditWatch entry, 0 otherwise). We further use sectoral and regional classifications, and year dummies. In regression model I (II) the year 1996, corporates (banks) and Canadian firms serve as reference. In regression models II and III we omitted the respective variable if we did not observe any default in a certain region or year. We use a heteroskedasticity consistent covariance matrix estimator (White, 1980).

	(I)		(II)		(III)	
Independent	Total sample		Financials		Banks	
variable	Coefficient	p value	Coefficient	p value	Coefficient	p value
Intercept	-13.043	<.001	-13.221	<.001	-10.304	<.001
Unsolicited	-0.356	0.236	-0.929	0.016	-1.208	0.014
Rating level	0.593	<.001	0.588	<.001	0.588	<.001
Negative Outlook	0.963	<.001	0.025	0.955	0.067	0.901
Positive Outlook	-0.670	0.192	-0.781	0.381		
Negative CreditWatch	1.859	<.001	1.320	0.001	1.012	0.055
Positive CreditWatch	0.554	0.168	-0.184	0.796		
Bank	-0.132	0.560				
Insurance	-0.549	0.104	-0.335	0.332		
Other financial	1.570	<.001	1.618	<.001		
Utilities	0.404	0.139				
Asia Pacific	-0.577	0.039	-0.543	0.294	-1.612	0.051
Australia/New Zealand	-0.476	0.410				
Europe/Mid East/Africa	-0.418	0.088	-0.874	0.083	-1.961	0.013
Latin America	0.222	0.364	-0.156	0.754	-0.593	0.451
Year 1997	1.050	0.317	2.577	0.001		
Year 1998	1.110	0.275	1.756	0.024	0.041	0.950
Year 1999	0.931	0.361	1.997	0.010	-0.499	0.539
Year 2000	1.899	0.058	2.769	<.001	0.762	0.148
Year 2001	2.148	0.032	2.676	0.000	0.414	0.450
Year 2002	1.014	0.317	2.120	0.003	-0.210	0.719
Year 2003	-0.549	0.605				
Year 2004	-0.084	0.937	-0.218	0.855	-1.659	0.140
Year 2005	-0.173	0.871	0.643	0.483	-1.674	0.128
McFadden adj. R^2	0.462		0.413		0.320	
Observations	26,642		15,802		6,159	

Table 6: Ex-post analysis of empirical default frequencies using a pooled logit regression approach. The dependent variable takes the value one in case of a default and zero if the firm has survived the next one-year realization period. The independent variables are: unsolicited signifies a dummy for rating solicitation (1 for unsolicited ratings, 0 otherwise), default frequency indicates the smoothed, long-term average default frequency of the estimation period that is followed by the realization period in which the default status is recorded. We further use sectoral and regional classifications, and year dummies. In regression model I (II) the year 1996, corporates (banks) and Canadian firms serve as reference. In regression models II and III we omitted the respective variable if we did not observe any default in a certain region or year. We use a heteroskedasticity consistent covariance matrix estimator (White, 1980).

	(I)		(II)		(III)	
Independent	Total sample		Financials		Banks	
variable	Coefficient	p value	Coefficient	p value	Coefficient	p value
Intercept	-6.421	<.001	-8.366	<.001	-5.537	<.001
Unsolicited	-0.366	0.178	-0.536	0.116	-1.237	0.034
Default frequency	21.268	<.001	22.814	<.001	22.318	<.001
Bank	-0.515	0.024				
Insurance	-1.474	<.001	-0.894	0.008		
Other financial	1.021	<.001	1.598	<.001		
Utilities	-0.343	0.213				
Asia Pacific	-0.981	0.001	-0.064	0.895	-0.250	0.747
Australia/New Zealand	-1.639	0.001				
Europe/Mid East/Africa	-0.871	0.000	-0.511	0.234	-0.894	0.184
Latin America	0.479	0.039	0.911	0.036	1.394	0.038
Year 1997	1.761	0.091	3.383	<.001		
Year 1998	1.946	0.057	2.539	0.000	0.295	0.646
Year 1999	1.694	0.098	2.718	<.001	-0.501	0.627
Year 2000	2.705	0.007	3.552	<.001	1.032	0.067
Year 2001	2.873	0.004	3.351	<.001	0.575	0.312
Year 2002	1.634	0.110	2.733	<.001	-0.113	0.846
Year 2003	-0.092	0.933				
Year 2004	0.524	0.622	0.451	0.696	-1.514	0.182
Year 2005	0.500	0.640	1.293	0.146	-1.418	0.196
McFadden adj. R^2	0.360		0.308		0.239	
Observations	26,642		15,802		6,159	

Table 7: Ex-post analysis of empirical default rates using a pooled logit regression approach with three-year realization periods. The sample comprises only banks. The dependent variable takes the value one in case of a default and zero if the firm has survived the next three-year realization period. The independent variables are: unsolicited signifies a dummy for rating solicitation (1 for unsolicited ratings, 0 otherwise), rating indicates the rating level of the estimation period expressed as numerical value for company i between 1 (AAA) and 18 (CC to C) that is followed by the realization period in which the default status is recorded, default frequency indicates the smoothed, long-term average default frequency of the estimation period that is followed by the realization period in which the default status is recorded, negative outlook (1 for negative outlook, 0 otherwise), positive outlook (1 for positive outlook, 0 otherwise), negative CreditWatch (1 for negative CreditWatch entry, 0 otherwise). We further use sectoral and regional classifications, and year dummies. The year 1996 and Canadian firms serve as reference. The dummy variable for Australia/New Zealand was excluded because we did not record any bank default in this region. We use a heteroskedasticity and autocorrelation consistent covariance matrix estimator (Newey and West, 1987).

Independent	(I)		(II)		(III)	
variable	Coefficient	p value	Coefficient	p value	Coefficient	p value
Intercept	-8.240	<.001	-8.209	<.001	-4.392	<.001
Unsolicited	-1.511	<.001	-1.468	<.001	-1.151	0.002
Rating level	0.570	<.001	0.566	<.001		
Default frequency					24.108	<.001
Negative Outlook			0.258	0.490		
Positive Outlook			-1.347	0.208		
Negative CreditWatch			0.102	0.859		
Asia Pacific	-1.623	0.002	-1.622	0.003	-0.494	0.377
Europe/Mid East/Africa	-1.972	<.001	-1.909	<.001	-1.014	0.021
Latin America	-0.289	0.555	-0.282	0.556	1.760	<.001
Year 1997	0.350	0.656	0.314	0.698	0.710	0.292
Year 1998	-0.428	0.591	-0.503	0.536	0.124	0.852
Year 1999	-0.203	0.790	-0.250	0.750	0.424	0.510
Year 2000	-0.181	0.812	-0.146	0.854	0.478	0.456
Year 2001	-0.722	0.352	-0.710	0.379	-0.156	0.811
Year 2002	-1.716	0.038	-1.716	0.044	-1.273	0.073
Year 2003	-3.613	0.005	-3.600	0.005	-3.310	0.010
McFadden adj. R^2	0.373		0.370		0.279	
Observations	4,757		4,757		4,757	

Table 8: Ex-post analysis of empirical default rates using an OLS regression approach. The sample comprises only banks. In Panel I (II) the dependent variable takes the bank-specific z-score of the next (third-next) year. Accounting data for the z-score calculation are extracted from Compustat. The independent variables are: unsolicited signifies a dummy for rating solicitation (1 for unsolicited ratings, 0 otherwise), rating indicates the rating level of the estimation period expressed as numerical value for company i between 1 (AAA) and 18 (CC to C) that is followed by the realization period in which the z-score is computed, default frequency indicates the smoothed, long-term average default frequency of the estimation period that is followed by the realization period in which the z-score is computed, negative outlook (1 for negative outlook, 0 otherwise), positive outlook (1 for positive outlook, 0 otherwise), negative CreditWatch (1 for negative CreditWatch entry, 0 otherwise), positive CreditWatch (1 for positive CreditWatch entry, 0 otherwise). We further use sectoral and regional classifications, and year dummies. Results for these control variables are omitted. The year 1996 and Canadian firms serve as reference. We use a heteroskedasticity and autocorrelation consistent covariance matrix estimator (Newey and West, 1987).

Independent	(I)		(II)		(III)	
variable	Coefficient	p value	Coefficient	p value	Coefficient	p value
Panel I: one-year horizon:						
Intercept	25.619	<.001	27.773	<.001	19.727	<.001
Unsolicited	3.477	0.005	2.870	0.030	0.618	0.632
Rating level	-1.634	<.001	-1.692	<.001		
Default frequency					-146.727	<.001
Negative Outlook			-5.262	<.001		
Positive Outlook			3.920	0.094		
Negative CreditWatch			-3.343	0.123		
Positive CreditWatch			2.379	0.484		
Adjusted R^2	0.154		0.163		0.136	
Observations	1,689		1,689		1,689	
Panel II: three-year horizon:						
Intercept	23.619	<.001	24.413	<.001	17.303	<.001
Unsolicited	5.575	<.001	5.191	0.007	2.712	0.095
Rating level	-1.704	<.001	-1.728	<.001		
Default frequency					-182.907	<.001
Negative Outlook			-1.871	0.347		
Positive Outlook			2.083	0.500		
Negative CreditWatch			-3.388	0.113		
Positive CreditWatch			3.743	0.364		
Adjusted R^2	0.184		0.183		0.178	
Observations	1,306		1,306		1,306	

References

Akerlof, G. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 89, 488–500.

Anwar, S., and H. Fang (2006): "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Review*, 96, 127–151.

Bannier, C., and M. Tyrell (2006): "Modelling the Role of Credit Rating Agencies - Do they Spark off a Virtous Circle?," *Working Paper, Goethe University Frankfurt.*

Barber, B. M., R. Lehavy, and B. Trueman (forthcoming): "Comparing the Stock Recommendation Performance of Investment Banks and Independent Research Firms," *Journal of Financial Economics.*

Becker, G. S. (1993): "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, 101, 385–409.

Behr, P., and A. Güttler (forthcoming): "The Informational Content of Unsolicited Ratings," *Journal of Banking and Finance.*

Blume, M. E., F. Lim, and A. C. MacKinley (1998): "The Declining Credit Quality of U.S. Corporate Debt: Myth or Reality?," *Journal of Finance*, 53, 1389–1413.

Boot, A. W., S. Greenbaum, and A. Thakor (1993): "Reputation and Discretion in Financial Contracting," *American Economic Review*, 83, 1165–1183.

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

Cantor, R., and F. Packer (1994): "The Credit Rating Industry," *Federal Reserve Bank of New York Quarterly Review*, Summer-Fall, 1–26.

Covitz, D. M., and P. Harrison (2003): "Testing Conflicts of Interest at Bond Rating Agencies with Market Anticipation: Evidence that Reputation Incentives Dominate," *Finance and Economics Discussion Series, Federal Reserve Board*, 2003-68.

Economist (2005): "Who Rates the Raters?," *Economist*, March, 67–69.

Fenn, G. W. (2000): "Speed of Issuance and the Adequacy of Disclosure in the 144A High-yield Debt Market," *Journal of Financial Economics*, 56, 383–405.

Fight, A. (2001): "The Ratings Game," John Wiley and Sons.

Flannery, M. J., S. H. Kwan, and M. Nimalendran (2004): "Market Evidence on the Opaqueness of Banking Firms' Assets," *Journal of Financial Economics*, 71, 419–460.

Gan, Y. H. (2004): "Why Do Firms Pay for Bond Ratings When They Can Get Them for Free?," *Working Paper, Wharton School.*

Gasparino, C. (1996): "Bond-rating Firms May Be Required to Disclose When Work is Unsolicited," *The Wall Street Journal.*

GORDY, M., AND B. HOWELLS (2006): "Procyclicality in Basel II: Can We Treat the Disease Without Killing the Patient?," *Journal of Financial Intermediation*, 15, 395–417.

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

HANNAN, T. H., AND G. A. HANWECK (1988): "Bank Insolveny Risk and the Market for Large Certificates of Deposits," *Journal of Money, Credit, and Banking*, 20, 203–211.

HILL, C. (2004): "Regulating the Rating Agencies," *Washington University Law Quarterly*, 82, 42–95.

HIRTLE, B. (2006): "Stock Market Reaction to Financial Statement Certification by Bank Holding Company CEOs," *Journal of Money, Credit, and Banking*, 38, 1263–1291.

HOLTHAUSEN, R., AND R. LEFTWICH (1986): "The Effect of Bond Rating Changes on Common Stock Prices," *Journal of Financial Economics*, 17, 57–89.

INTERNATIONAL ORGANIZATION OF SECURITIES COMMISSION (IOSCO) (2003): "Report on the Activities of Credit Rating Agencies," IOSCO, Madrid.

JAPAN CENTER FOR INTERNATIONAL FINANCE (JCIF) (1999-2001): "Characteristics and Appraisal of Major Rating Companies," JCIF.

JORION, P., Z. LIU, AND S. CHARLES (2005): "Informational Effects of Regulation FD: Evidence from Rating Agencies," *Journal of Financial Economics*, 76, 309–330.

KLEIN (2004): "Credit Raters' Power Leads to Abuses, Some Borrowers Say," Washington Post Staff Writer.

LAEVEN, L., AND R. LEVINE (2006): "Corporate Governance, Regulation, and Bank Risk Taking," *Working Paper, Brown University / IMF*.

MICHAELY, R., AND K. WOMACK (1999): "Conflict of Interest and the Credibility of Underwriter Analyst Recommendations," *Review of Financial Studies*, 12, 653–686.

MILLON, M., AND A. THAKOR (1985): "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies," *Journal of Finance*, 40, 1403–1422.

MORGAN, D. P. (2002): "Rating Banks: Risk and Uncertainty in an Opaque Industry," *American Economic Review*, 92, 874–889.

NEWEY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

NICKELL, P., W. PERRAUDIN, AND S. VAROTTO (2000): "Stability of Rating Transitions," *Journal of Banking and Finance*, 24, 203–227.

OICU-IOSCO (2004): "Code of Conduct Fundamentals for Credit Rating Agencies," mimeo.

——— (2007): "Review of Implementation of the IOSCO Fundamentals of a Code of Conduct for Credit Rating Agencies," mimeo.

Partnoy, F. (2001): "The Paradox of Credit Ratings," *University of San Diego Law and Economics Research Paper No. 20.*

——— (2006): "How and Why Credit Rating Agencies are Not Like Other Gatekeepers," *University of San Diego Legal Studies Research Paper No. 07/46.*

Poon, W. P. H. (2003): "Are Unsolicited Ratings Downward Biased?," *Journal of Banking and Finance*, 27, 593–614.

Poon, W. P. H., and M. Firth (2005): "Are Unsolicited Credit Ratings lower? International Evidence from Bank Ratings," *Journal of Business Finance and Accounting*, 32, 1741–1771.

Ramakrishnan, R., and A. Thakor (1984): "Information Reliability and a Theory of Financial Intermediation," *Review of Economic Studies*, 51, 415–432.

Roy, A. D. (1952): "Safety First and the Holding of Assets," *Econometrica*, 20, 431–449.

Securities and Exchange Commission (SEC) (2005): "Regulatory Issues and Economic Principles," SEC, Washington DC.

Standard and Poor's (2006a): "Annual 2005 Global Corporate Default Study And Rating Transitions," Standard and Poor's, Rating's Direct.

——— (2006b): "Corporate Ratings Criteria," Standard and Poor's.

Van Roy, P. (2006): "Is There a Difference Between Solicited and Unsolicited Bank Ratings and, If So, Why?," *National Bank of Belgium Working Paper No. 79.*

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817–838.