Review Of Finance (2011) 15: 475–510
doi: 10.1093/rof/rfq032
Advance Access publication: February 23, 2011

Is There a Relationship Benefit in Credit Ratings?*

THOMAS MÄHLMANN
University of Cologne

Abstract. This paper shows that firms with longer rating agency relationships have better credit ratings, conditional on observables. The paper also finds that (1) controlling for observables, firms with longer relationships, while having higher average ratings, do not have lower default rates, (2) relationship benefits are larger among firms with a greater incentive to game their information supplied to agencies or to pressure agencies into giving higher ratings, and (3) investors demand a (price) discount on bonds sold by relationship firms and the correlation between bond yield spreads and ratings is decreasing with relationship length. In sum, the evidence is inconsistent with first-order credit quality explanations but rather supports a ‘‘learning-to-gaming’’ and an ‘‘adverse incentives’’ story.

JEL Classification: D82, G24, G28

1. Introduction

Understanding the properties of credit ratings is important, given that credit ratings play a significant and increasingly important role in firms' access to capital (e.g., Faulkender and Petersen, 2006), in federal and state legislation, in the Basel II capital adequacy rules issued by bank regulators, and in corporate debt contracts (Beaver, Shakespeare, and Soliman, 2006). This paper starts with the empirical observation that (unconditionally) the average rating is increasing (i.e., closer to AAA) in the firm–agency relationship duration (see Figure 1).[1] One obvious explanation for this pattern is that firms with longer rating histories are better credit risks and hence receive higher ratings. This is what I call the "credit quality" explanation. But there is also a second "learning-to-gaming" explanation that

* I would like to thank an anonymous referee for helpful suggestions and Georg Keienburg and Clara Sukonnik for invaluable research support. Financial support from the Institute of Banking at the University of Cologne is gratefully acknowledged.

[1] Figure 1 shows that the mean rating in my sample declines roughly monotonically with the length of the agency relationship (Relate). Note that a numerical rating of 5.0 corresponds to the BB letter level, 4.0 to BBB, 3.0 to A, etc. Hence, the solid line in Figure 1 documents a positive average impact of relationship length on ratings. For example, the unconditional mean rating of a firm with an 11-year agency relationship (4.16) is approximately one half of a full rating letter better than the mean rating of a firm that has only a 1-year relationship (4.62).

builds on the economic implications of learning-by-doing (Arrow, 1962)[2] or learning-by-observing (DeLong and DeYoung, 2007) and that can also generate better ratings for longer firm–agency relationships.

Suppose rating agencies are of good intent and do their best to produce an accurate rating, based on the information provided by the firm. Moreover, rating agencies are exempt from Regulation Fair Disclosure (FD), which means that firms can share private information with them.[3] Over time, firms learn the effects on their rating of sharing certain information with the rating agencies. Therefore, they will strategically release private information, which will lead to a larger bias as the selection of released information becomes more efficient in influencing the rating due to firm experience. The same holds for objections against a less favorable rating than expected. Over time, the firm will learn better how to appeal against a less favorable rating (of course, the firm would not do so for a rating that is too positive).

It could also be argued that, instead of firms and their managers, rating agencies are indeed the ones of "bad intent." In particular, the prediction that ratings become more favorable and less accurate with relationship time is also consistent with rating agencies simply catering the interests of their best clients (i.e., those with long issuance histories).[4] Since there appears to be no easy way to disentangle this "adverse incentives" argument (fuelled by agencies' issuer-pays business model) from the learning-to-gaming theory, this paper contrasts the credit quality story with all others, viewing the adverse incentives theory about as plausible as the

[2] Arrow does not explicitly consider learning-by-doing through communication and information transmission but analyzes the economic implications of learning-by-doing in general. However, his view that "Learning is the product of experience (. . .) and therefore only takes place during activity" is the same perspective of learning-by-doing than the one I have in mind in this paper.

[3] There is ample evidence indicating that ratings contain private information otherwise not available to investors. For example, Jorion, Liu, and Shi (2005) observe that Regulation FD limited the availability of inside information to investors but not to "nationally recognized statistical rating organizations". The authors find that rating changes cause larger stock price movements after Regulation FD than before.

[4] In contrast to structured finance ratings, it is somewhat harder to motivate the benefit of this behavior for Standard & Poor's (S&P) and Moody's in the US corporate bond market which is highly concentrated and can be characterized by a "two-rating norm"—that is, to access a broad investor pool, bond issuers are implicitly required to obtain ratings from both major rating agencies Moody's and S&P. For example, Bongaerts, Cremers, and Goetzmann (2009) find that virtually all large US bonds in their sample, issued over the period 2002 through 2006, are rated by both Moody's and S&P, whereas between 40% and 60% of the bonds are also rated by Fitch. The fact that the market is dominated by Moody's and S&P and that their ratings are no substitutes but rather complements implies that both agencies should have no incentive to reduce standards but rather to apply strict rating policies since that would help them to keep their monopoly. In empirical work, however, Becker and Milbourn (2009) find that in the market for corporate bonds, increased market share by Fitch decreased the accuracy of ratings from Moody's and S&P and made them pander more to issuers, a finding consistent with the model in Bolton, Freixas, and Shapiro (2009).

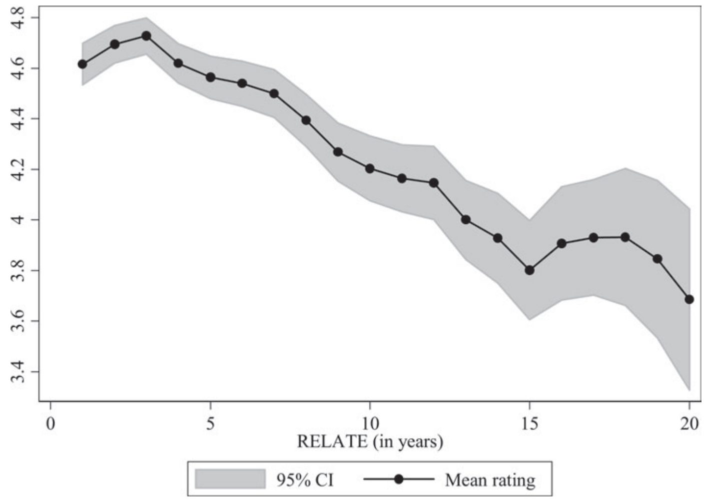

Figure 1. Effect of relationship duration on mean rating: unconditional approach. The figure graphs the mean credit rating as a function of the firm–agency relationship length. The letter ratings are transformed into numerical equivalents using an ordinal scale ranging from 1 for the highest rated firms (AAA) to 7 for the lowest rated firms (below B). Hence, the value 3 corresponds to A and 4 to BBB. The sample is based on 7,626 firm-years from COMPUSTAT over the period 1986–2005.

learning-to-gaming theory. More specifically, by focusing on credit and rating quality, one can tell these stories apart: whereas the credit quality explanation implies that a firms' credit quality is increasing and rating quality is nondecreasing in relationship duration, the learning-to-gaming/adverse incentives argument predicts a negative impact of relationship duration on rating quality.

Based on a large US sample of 1,263 publicly traded firms (and their bonds), being rated by S&P for at least three consecutive years over the period 1986–2005, the paper's empirical tests proceed in three parts. First, I investigate whether the unconditional pattern between relationship duration and average rating levels documented in Figure 1 still holds when I control for observables. By way of preview, I find in (firm and bond) rating-level regressions that relationship benefits are statistically and economically highly significant. To provide an indication of the economic significance of the results, I find that firms with an 11-year agency relationship are predicted to have average ratings 0.61 notches higher than comparable firms that are the same in every observable way except that they only have a 1-year relationship. I find slightly higher relationship benefits in a sample of bond ratings. To be sure that these effects are not driven by omitted variables that are correlated with relationship length, I control for a wide range of firm (or bond)-specific characteristics, including several proxies for the nature of a firm's information environment and industry and time effects.

In a second step, I ask whether credit quality is also increasing in relationship duration, controlling for the same observables contained in the rating-level regressions. If this were indeed the case, then relationship duration would proxy for some unobservable credit quality measure (e.g., favorable "soft" information discovered over the course of the firm–agency relationship), thereby justifying the higher average ratings for firms with longer agency relationships. Controlling for observables, however, I find that older firms with longer relationships, while having higher average ratings, do not possess lower default rates. In addition, I also show that firms with pronounced incentives to game their ratings (i.e., firms with bad prospects as indicated by *ex post* increasing bond yield spreads) face larger relationship benefits. In sum, this is evidence against the credit quality explanation and suggests that, *ceteris paribus*, the upward pattern in rating levels is not accompanied by a corresponding upward pattern in credit quality. In turn, the evidence implies that rating quality is decreasing with relationship duration.

To investigate this point more clearly, I directly estimate the effect of relationship duration on rating quality. In a first exercise, I compare initial bond yield spreads to examine whether the market differentiates the *ex ante* credit quality of bonds sold by relationships firms versus similar bonds sold by nonrelationship firms. If the market believes that relationship bonds have higher credit risk than equally rated nonrelationship bonds, initial yield spreads on the former ought to be higher, conditional on the rating. The evidence presented is consistent with this prediction. Furthermore, I find that price discounts for relationship bonds at issuance are not persistent but decrease over a bond's life, in line with S&P more aggressively downgrading relationship bonds after issuance. In rating informativeness tests, I observe that the correlation between bond yield spreads and ratings falls for bonds from relationship firms. In other words, ratings appear less informative about yield spreads. These results are robust to the inclusion of various types of fixed effects. Overall, the evidence from the rating quality tests supports the learning-to-gaming/adverse incentives argument that predicts lower quality (less informative) ratings when the agency–firm relationship matures.

In a final exercise, I analyze the timing of bond rating changes. In one set of tests, I estimate Cox proportional hazard models for the first bond rating change after issuance. I find that relationship bonds have a significantly higher downgrade (but not upgrade) hazard compared to nonrelationship bonds. I also examine yield spread changes leading up rating changes. The evidence indicates that relationship bonds exhibit smaller yield spread increases prior to a downgrade, meaning that S&P acts more promptly to downgrade relationship bonds compared to equally rated bonds sold by nonrelationship firms. Taken together, these two sets of results are consistent with overly positive initial ratings that subsequently come back to earth.

This study contributes to the credit ratings literature since it is one among only few papers that directly examine the impact of firm-level incentives on the properties of ratings. In related empirical research, Bongaerts, Cremers, and Goetzmann (2009) and Mählmann (2009) investigate whether observed differences between ratings from distinct agencies can be explained by firms strategically seeking the number of ratings that fit their needs best. Faure-Grimaud, Peyrache, and Quesada (2009) and Mählmann (2008) study theoretically the incentives of firms to hide low ratings, assuming that a firms' endogenous decision to get a rating is not observable by the market. This paper points to a possible "dark side" or "hidden cost" of long-term relationships: Opportunism in the form of self-seeking behavior might cause high-reputation (or trust) relationships to be less effective and provide lower quality products. The paper's results also suggest that rating agencies' conflicts of interests are probably not well managed by reputation concerns alone.

The remainder of the paper is organized as follows. Section 2 develops the papers' central set of hypotheses, and Section 3 introduces the data set and sample selection procedure and motivates the variables employed in the analysis. Section 4 presents the empirical models used to investigate the relation between the firm–agency relationship duration and firms' credit ratings along with the main empirical results. Section 5 presents additional robustness checks, and Section 6 concludes.

2. Relationships and Credit Ratings: Hypotheses Development

There has been growing interest in economics and finance in studying the effect of experience on decision making. As outlined in the classical learning-by-doing models (Arrow, 1962; Grossman, Kihlstrom, and Mirman, 1977), economic agents may learn through experience to make better decisions as they acquire better information about the environment or about the quality of their information signals. Recent empirical papers that find some learning in various settings include DeLong and DeYoung (2007), Pástor, Taylor, and Veronesi (2009), and Seru, Shumway, and Stoffman (2010).

Focusing more narrowly on private information acquisition in long-term business relationships through learning and experience, a large banking literature studies the relative benefits of relationship lending compared to transaction lending (see Boot, 2000, for a review). One part of this literature predicts significant relationship benefits for the lender, coming in part from information reusability over time (Greenbaum and Thakor, 2007) or sharing of sensitive information (Bhattacharya and Chiesa, 1995). However, there is at least one major downside due to one party's opportunistic behavior, questioning the efficiency of long-term banking relationships: the hold-up problem (Sharpe, 1990; Rajan, 1992). The information monopoly a bank generates in the course of lending through learning and experience accumulation enables it to make loans at noncompetitive terms in the future to the borrower. More generally, in long-term relationships, characterized by high levels of mutual trust and learning, a risk exists that one party takes advantage of its buildup information stock and exploits the other party in opportunistic ways. This paper employs a closely related idea.

I hypothesize that firms with long agency relationships have learned, by observing the sensitivity of their rating to the information provided to the rating agency in the past, how to better manage their private information release in a way to obtain more favorable ratings. In addition, by being initially very supportive and reporting private information truthfully, firms can build up a reputation for being truthful that likely lowers agencies' auditing and due diligence levels over time. This, in turn, might increase a firm's incentive to engage in opportunistic information disclosure as the relationship matures.

An alternative "adverse incentives" argument for possible inflated corporate bond ratings could be based on the conflicts of interest inherent in the agencies' issuer-pay business model. Since rating fees are transaction based, that is, they are only paid if the issue takes place, it is possible for the issuer to capture the agency, that is, to pressure the agency into giving the company the rating it is looking for the issue to proceed.[5] By assuming that agencies are more willing to stretch ratings for client firms who provide a lot of business (i.e., frequent issuers with long issuance histories), relationship benefits can emerge. However, by inflating its ratings, an agency increases its current revenue but faces the risk of a decrease in future income, due to a loss in reputation. Bolton, Freixas, and Shapiro (2009) model the agencies' conflict of understating credit risk to attract more business and find that rating agencies are more prone to inflate ratings when there is a larger fraction of naive investors in the market or when expected reputational costs are lower. In line with the empirical results of Becker and Milbourn (2009), increased competition may reduce the effectiveness of the reputational mechanism, leading to lower quality ratings. The trade-off between current revenue and reputation is also examined by Mathis *et al.* (2009) in a dynamic model where reputation is endogenous. Their analysis provides an interesting additional insight—the possibility of confidence or reputation cycles.

I formalize the above "learning-to-gaming" and "adverse incentives" arguments into the following empirically testable hypothesis.

[5] There is a large amount of anecdotal evidence suggesting that issuer payments may influence ratings. For example, a survey of investment professionals published by the Chartered Financial Analyst Institute said that 211 (11%) of 1,956 respondents had witnessed a rating agency change its rating in response to pressure from a bond issuer or underwriter (see "Finance Group Questions Bond-Rating Proposals" by A. Lucchetti, *Wall Street Journal*, July 7 2008).

Hypothesis 1 (H1). *The level of ratings is increasing (i.e., ratings become more favorable), ceteris paribus, with the length of the firm–agency relationship.*

I use relationship duration as my main proxy for experience accumulation and issuer pressure. Please note that there appears to be no straightforward test that plausibly distinguishes between the learning-to-gaming and the adverse incentives view. Hence, this paper does not contrast these two arguments for positive relationship effects. As discussed above, there are alternative "first-order" effects expected to drive a positive correlation between rating levels and relationship length. Most obviously, better quality firms survive longer (and, hence, receive higher ratings after long relationships) or firms that survive become safer over time as they mature. I call this the "credit quality" explanation. In addition, when management and agencies interact over multiple periods, managers can build reputations for truthful information disclosures (Stocken, 2000) or for being uninformed such that agencies interpret the absence of disclosure less skeptically (Einhorn and Ziv, 2008). Such a reputation likely reduces the cautionary bias in assigned ratings and the information risk faced by agencies, thereby also producing a positive relationship effect on ratings. I call this the "issuer reputation building" explanation. To disentangle these alternative explanations from the learning-to-gaming/adverse incentives view, I employ two identifying settings: managerial incentives and rating quality.

Concerning the first, managerial motives to engage in strategic information disclosure (or to exert pressure on agencies) are widely studied in the finance and accounting literature. For instance, managers' career concerns can affect their decisions to withhold their private information. Since managers' career concerns are especially heightened when a firm is approaching a state of financial distress (Gilson, 1989; Weisbach, 1988), the link between financial distress and management turnover provides managers with incentives to delay the bad news in the hope of an eventual turnaround (Kothari, Shu, and Wysocki, 2009). Consistent with this prediction, Frost (1997) finds evidence that voluntary disclosure credibility declines for financially distressed firms. Furthermore, studies that have analyzed discretionary accounting choices of managers in financially troubled firms (i.e., firms that have reported debt covenant violations) find evidence consistent with income-increasing accounting policies in years prior to the violation (DeFond and Jiambalvo, 1994; Sweeney, 1994). Finally, a large literature indicates that firms make opportunistic accounting choices and engage in earnings management during periods when they raise external capital. For example, Dietrich, Harris, and Muller (2000) find that firms make accounting method choices regarding fair value estimates of investment properties to boost earnings and time asset sales to help smooth earnings before raising debt.

Based on these findings, I hypothesize that managers with negative private information about their firms' future prospects are more likely to manipulate the information disclosed to rating agencies or to pressure the agency into giving higher ratings prior to a new public debt issue. This leads to the second hypothesis, which I refer to as the "managerial incentives" hypothesis (phrased in favor of the learning-to-gaming/adverse incentives argument).

Hypothesis 2 (H2). *Relationship benefits are larger in situations where managers have increased incentives to game ratings directly or to pressure the agency into giving higher ratings (i.e., prior to a new public debt issue of firms with bad prospects).*

Since the credit quality and the reputation building explanations make no prediction about the effect of incentives on the correlation between relationship duration and ratings, evidence in favor of Hypothesis 2 solely supports the learning-to-gaming/adverse incentives view. In my final hypothesis test, I focus on rating quality. Obviously, if firm managers gain experience over time how to successfully manipulate their credit ratings or are more successful in capturing the agency, rating quality should be a decreasing function of relationship duration. In contrast, the two alternative explanations predict that rating quality is increasing over the course of the relationship (or, at least, nondecreasing) due to, for example, a survival of more high-quality firms that are easier to identify or due to more accurate information disclosure. These arguments lead to my second identifying hypothesis, which I refer to as the "rating quality" hypothesis (phrased in favor of the learning-to-gaming/adverse incentives argument).

Hypothesis 3 (H3). *Rating quality is decreasing with relationship duration (i.e., with accumulated experience and an agencies' willingness to stretch ratings).*

3. Data

3.1 THE FIRM SAMPLE

Sample selection

My measure of credit rating is the S&P Long-Term Domestic Issuer Credit Rating extracted from COMPUSTAT (data item 280). This rating is defined as S&P's "current opinion on an issuer's overall capacity to pay its financial obligations, i.e., its fundamental creditworthiness. It generally indicates the likelihood of default regarding all financial obligations of the company, ... [and does] not take into account recovery prospects" (S&P, 2008). Besides these long-term "corporate credit ratings," COMPUSTAT also reports short-term issuer ratings (i.e., commercial paper (CP) ratings) and subordinated debt (i.e., issue) ratings. However, because these additional ratings have a close correspondence to corporate ratings[6], and “a corporate credit rating is published for all companies that have issue ratings” (S&P, 2008), it seems justified to focus on corporate ratings when defining rating agency–firm relationships. The letter ratings are transformed into numerical equivalents using an ordinal scale ranging from 1 for the highest rated firms (AAA) to 17 for the lowest rated firms (below B−).[7] Central to the analysis in this paper is the notion of a firm–agency relationship spell. A relationship spell defines runs of consecutive rating observations as recorded in COMPUSTAT. A spell starts with the first observed rating in COMPUSTAT and ends whenever for two or more consecutive years ratings are missing. To be included in the sample, a relationship spell has to meet the following conditions: First, the start of the spell must be in 1986 or later. Second, each spell must be at least 3 years of length, that is, a firm must have three or more consecutive rating observations in COMPUSTAT. Third, I require each firm to have at least 2 years of nonmissing sales and total assets in COMPUSTAT prior to the spell start.[8]

I consider 1986 as the first sample year because COMPUSTAT started to collect ratings in 1985, and therefore, all ratings in 1985 would be classified by definition as indicating the start of a relationship spell. This lump goes away by 1986. I discuss the implications of this sampling procedure in the section “Possible limitations of the sampling procedure” below. Since COMPUSTAT does not report rating histories prior to 1985, I cannot be sure whether the first

[6] To rate subordinated bonds, issues are notched down from the corporate credit rating level—typically one notch down for investment-grade and two notches down for speculative-grade companies. A short-term rating is an assessment of an issuer’s default likelihood with respect to an instrument considered short term in the relevant market (in the USA, any obligation with an original maturity of no more than 365 days). Short-term or CP ratings range from “A1+/A1”, for the highest-quality obligations, to “D”, for the lowest. Even if knowing the long-term rating will not fully determine a CP rating, investment-grade short-term ratings, in particular, are highly correlated with long-term corporate ratings (S&P, 2008).

[7] I checked whether I introduce a selection bias by treating a number of the rating declines (downgrades among CCC, CC, and C) as unchanged, thereby possibly producing an upwardly biased average rating change. However, the number of below B− ratings is small (86 for CCC+, 49 for CCC, 10 for CCC−, 17 for CC, and 0 for C) and all results are qualitatively unchanged when I extend the ordinal scale to further differentiate between them.

[8] Given the above definition, each firm can have several distinct relationship spells. For example, a firm might be rated on an ongoing basis between 1988 and 1995, not rated between 1996 and 1999, and rated again from 2000 to 2005. Since learning benefits are presumably most clearly visible at the ultimate start of a long-term relationship, I confine the analysis to the first relationship spell for each firm. Hence, each firm enters the sample with at most one spell.

observed rating in COMPUSTAT really indicates the start of the rating relationship.[9] To identify firms with previous ratings from S&P not captured by COMPUSTAT, I augment the sample with data from the Securities Data Corporation's (SDC) Domestic New Issuances database that records debt and equity issuances in the USA since 1970. In particular, I exclude all firms (spells) with S&P-rated debt issuances (public bonds or syndicated loans) one or more years prior to the start of the spell.[10] Furthermore, I exclude financial firms (SIC codes 6000–6999) but emphasize that all results are robust to their inclusion. Finally, I drop all spells with insufficient data from annual COMPUSTAT files and Center for Research in Security Prices (CRSP) daily stock files to calculate the control variables (discussed below) for at least one firm-year in the spell. The final sample consists of 1,263 relationship spells (firms), yielding 7,626 firm-years over the period from 1986 to 2005. The mean (median) spell length amounts to 7.65 (7) years with a minimum of 3 and a maximum of 20 years.

The variable of key interest (Relate) measures the elapsed time in years between the start of a spell and the end of the actual (fiscal) firm-year; hence, it denotes the current length of the first relationship of a firm with S&P. In unreported robustness checks, I used alternative proxies for firm learning within relationships or for the size of an issuer's rating business with S&P (e.g., the cumulative number or notional volume of past S&P-rated debt issues, obtained from SDC) but found the results quite similar to those obtained with relationship duration. Hence, I focus on Relate for the rest of the paper. As can be seen in the first row of Table I, the average relationship length declines with the rating, at least for the grades A (7.43 years) to B (4.95 years), providing a first indication of a relationship benefit. However, to present more convincing evidence in this respect, I have to control for a variety of other factors shown in previous studies (e.g., Blume, Lim, and

[9] While S&P has provided ratings for more than 80 years, it currently only sells databases (CreditPro) with rating histories starting in 1981, due to various changes in methodology affecting comparisons of ratings across time periods. Even if CreditPro gives four more years of historical rating data compared to COMPUSTAT, the same caveat applies since nothing is known about a firm's rating history before 1981.

[10] Of the 1,404 first-time rating relationship spells identified in COMPUSTAT, 860 spells could be matched with SDC data. Of these, 488 have a spell starting date within 1 year of their first S&P-rated debt issuance date in SDC, 231 have a spell starting date more than 1 year before, and 141 more than 1 year after their first-time S&P-rated debt issuance date. These 141 spells are excluded from the analysis since COMPUSTAT does not adequately define the start of the rating relationship, and I have no data about the course of the rating relationship prior to the COMPUSTAT starting date. Hence, I cannot calculate the exact relationship length for any firm-year in COMPUSTAT. It is important to note that the qualitative results of this paper are robust to inclusion or exclusion of the 544 spells for which SDC data on the first rated debt issuance date were not available.

Table I. Sample means across rating categories and correlations with Relate

The means are calculated using a panel data sample of 7,626 observations from 1986 through 2005. The sample is all nonfinancial COMPUSTAT firms that (i) established a rating relationship with S&P for at least 3 years, (ii) the relationship starts in 1986 or later, (iii) have at least 2 years of nonmissing sales and total assets data in COMPUSTAT prior to the start of the relationship, and (iv) have no S&P-rated debt issue in SDC's Domestic New Issuances database prior to the start of the relationship. The sample excludes observations with missing values for any of the variables. Relate is the number of years the firm has had a relationship with S&P. Interest coverage is the ratio of operating income after depreciation plus interest expense to interest expense. Roa is the ratio of operating income before depreciation to total assets. LT debt is the ratio of long-term debt to total assets. Total debt is the ratio of total debt to total assets. Three-year averages of these ratios are computed for each firm-year. Market value (in $ billion) is the year-end market value of equity deflated by the consumer price index. Market model beta is estimated from the market model using daily stock returns in each fiscal year. The beta estimates are controlled for nonsynchronous trading effects using the Dimson (1979) procedure. Idiosyncratic volatility is the standard error from the market model. The estimates of betas and volatilities for each year are further adjusted for the variation in the means of betas and volatilities over time, respectively (divided by their respective cross-sectional mean of that year). Age denotes the number of years since the firm's first appearance in COMPUSTAT with nonmissing (and positive) total assets. Fraction rated is the percentage of rated firms in the firm's industry. MtB is (total assets − book equity + market equity)/total assets. Tangible is property, plant, and equipment scaled by total assets. The last column displays Spearman's rank correlation coefficients for each variable with Relate.

	Rating							
	AAA − AA (n = 295)	A (n = 1,285)	BBB (n = 2,015)	BB (n = 2,203)	B (n = 1,666)	Below B (n = 162)	All (n = 7,626)	Corr. with Relate
Relate	6.88	7.43	6.63	5.65	4.95	5.23	6.09	1.000
Coverage	14.85	12.81	8.73	5.98	4.03	1.39	7.68	−0.018
Roa	0.18	0.16	0.15	0.13	0.09	0.04	0.13	0.028
LT debt	0.15	0.20	0.26	0.38	0.46	0.52	0.33	0.011
Total debt	0.22	0.26	0.30	0.41	0.50	0.63	0.37	0.010
Market value (in $ billion)	37.18	12.03	4.71	1.77	0.91	0.67	5.43	0.259
Beta	1.01	0.91	0.95	1.10	1.17	0.97	1.04	−0.062
Volatility	0.67	0.75	0.86	1.13	1.59	2.39	1.11	−0.207
Age	25.22	25.06	19.29	14.21	11.97	11.67	17.26	0.462
Fraction rated	0.20	0.29	0.28	0.27	0.24	0.30	0.27	0.149
MtB	2.33	2.09	1.68	1.56	1.54	1.59	1.71	0.027
Tangible	0.33	0.38	0.38	0.38	0.36	0.46	0.37	−0.004

MacKinlay, 1998; Jorion, Shi, and Zhang, 2009) to reflect a typical rating assignment process.

Control variables

The basic control variables include four accounting ratios (interest coverage, return on assets, long-term debt leverage, and total debt leverage) meant to capture a firm's financial risk and three equity market variables (a firm's market value of equity and two equity risk measures: beta and volatility) meant to capture business risk.[11] An extensive motivation for the inclusion of these variables in rating assignment models can be found in Blume, Lim, and MacKinlay (1998) or Amato and Furfine (2004). As shown in Table I, the means of the seven basic control variables are roughly monotonic across rating categories in the expected way, except for the market model beta. In sum, higher rated firms are, on average, larger, have a lower residual volatility in their stock returns, have higher interest coverage, are more profitable, and are less leveraged.

Livingston, Naranjo, and Zhou (2007) and Morgan (2002) find that firms with asset opaqueness problems face a higher probability of receiving split ratings from Moody's and S&P. Adam and Goyal (2008) provide evidence that market-to-book (MtB) is the most informative proxy for growth opportunities, and firms with larger growth opportunities tend to be younger firms in newer industries, making them more opaque and harder to value. To control for cross-sectional differences in asset opaqueness, I include the MtB ratio defined as market value of equity (COMPUSTAT item 25 · item 199) minus book value of equity (item 60) plus total assets (item 6) divided by total assets. As an additional accounting proxy for asset opaqueness, tangible denotes the amount of fixed assets (item 8) as a percentage of total assets.[12] Firm age measures the publicly available component of information. Finally, I control for a firm's uniqueness and the rating agency's industry competence by including a variable (fraction rated) that indicates for each firm-year the percentage of firms in the same industry that have rated debt.

[11] Interest coverage is the ratio of operating income after depreciation (178) plus interest expense (15) to interest expense (15), where the numbers in parentheses are the COMPUSTAT item numbers. Return on assets (Roa) is the ratio of operating income before depreciation (13) to total assets (6). Note that Roa provides a slightly better fit to the data than operating margin (13/12) used, for example, by Blume, Lim, and MacKinlay (1998). Long-term debt leverage (LT debt) is the ratio of long-term debt (9) to total assets (6). Total debt leverage (total debt) is the ratio of long-term debt (9) plus debt in current liabilities (34) to total assets (6). Estimates of beta and residual standard errors (Volatility) are obtained from the market model using at least 200 daily returns from the previous fiscal year. To adjust for nonsynchronous trading effects in the beta estimates, I adopt the Dimson (1979) procedure with one leading and lagging value of the CRSP value-weighted market return. As in Blume, Lim, and MacKinlay (1998), all accounting ratios are reported as 3-year moving averages, and the equity risk measures are scaled by dividing by the cross-sectional mean for each year, which eliminates any trend.

[12] Obviously, the tangibility variable could also capture expected recovery values or the ability of the firm to pledge assets as collateral to prevent defaults. However, since corporate credit ratings focus solely on default likelihood and leave recovery prospects aside (S&P, 2008, p. 11), this channel is partly ruled out. Similarly, because expected recovery values do not impact issuer ratings (in contrast to issue ratings), I can reject the objection that relationship length is just capturing the impact of increasing recovery rates for older firms or firms with longer credit histories.

The remaining rows in Table I contain the distribution of the control variables' means by rating category and overall. We can see that higher rated firms are older and have less tangible assets and more growth options as measured by MtB. However, the unexpected decrease in mean MtB for lower rating levels is probably due to the high positive correlation between MtB and market value of equity ($\rho = 0.35$). The last column in Table I displays Spearman's rank correlation coefficients of the control variables with the relationship measure. As expected, Relate is positively correlated with firm age and market value. Interestingly, firms with higher equity market risk (as measured by beta and volatility) have shorter relationships. This complements findings in Faulkender and Petersen (2006) and Lemmon and Zender (2010) who report that firms with less volatile assets and stock returns are more likely to have a rating (a proxy for public bond market access).

Possible limitations of the sampling procedure

It should be noted that the number of sample observations is declining with Relate, resulting in just 4.1% of observations with relationship durations beyond 15 years compared to 53.7% of the sample with values for Relate equal to or less than 5 years. This relatively low frequency of long-term relationship spells is due to an important limitation in the sample selection process. As discussed earlier, since COMPUSTAT's coverage of S&P annual long-term issuer ratings starts in 1985, I begin to record spells starting in 1986. Hence, all firms with agency relationships established prior to 1986 are not taken into account. In fact, I undersample long-term relationships in the first years, making the sample of relationship spells not a truly random sample. However, this lump should go away as I move further in time from 1986. The solid line in Figure 2 shows that the mean relationship length is steadily increasing over time. Whereas relationship length has a mean of 4.01 years across observations in 1990, this number more than doubles to 9.49 years in 2005.

To account for this undersampling of long-term relationships, I will also present results for a later subsample, spanning the period 1995–2005. Note that the mean relationship duration between firm and agency increases across the two samples, from 6.09 years for the full sample to 6.80 years for the later sample. These differences are statistically significant at 1%. The later subsample also shows a higher standard deviation for Relate, indicating that it covers a more comprehensive range of duration values.[13]

[13] The standard deviations are 4.22 years for the full sample and 4.45 years for the later sample. As a comparative value, the standard deviation of Relate for the period between 1986 and 1990 is 1.33 years (mean 2.69 years).

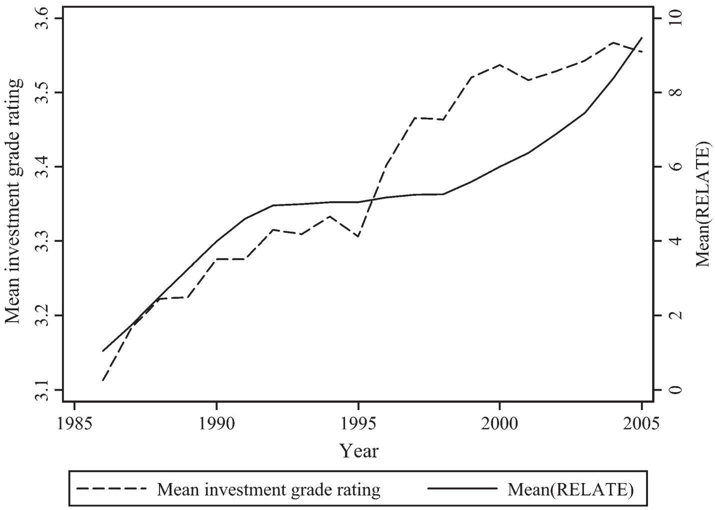

Figure 2. Sample average credit quality and relationship duration over time. The figure graphs the mean investment-grade rating (BBB or above, corresponding to the numerical equivalents 1 for AAA, 2 for AA, 3 for A, and 4 for BBB) and the mean firm–agency relationship length as a function of time. The sample is based on 7,626 firm-years (3,595 with investment-grade ratings) from COMPUSTAT over the period 1986–2005.

Blume, Lim, and MacKinlay (1998) and Jorion, Shi, and Zhang (2009) document a downward trend in the average credit quality of US firms over the 1980s and 1990s.[14] Such a picture also emerges from my data set. In particular, over the 20 years in the sample, the fraction of A firms has dropped from 27% in 1986 to 15% in 2005. The decrease in the fraction of AA and A firms (from 33% to 17%) is offset by an increase in the fraction of BBB and BB firms (from 33% to 63%). The dashed line in Figure 2 shows that for investment-grade firms (rated BBB or better), the average rating has dropped from a value (3.1) close to A to a value (3.6) more in line with BBB+. Unfortunately, this general decrease in credit quality of US firms is positively correlated with mean relationship length—that is, higher mean values of Relate are associated with lower quality average ratings (see Figure 2). Hence, the sample selection procedure together with the secular decline in credit quality

[14] By using ordered probit panel regressions of the determinants of investment-grade ratings including annual dummies and several control variables, Blume, Lim, and MacKinlay (1998) find a downward trend in the annual intercepts, which they interpret as a systematic tightening of rating standards, *ceteris paribus*. In a recent paper, Jorion, Shi, and Zhang (2009) show that this downward trend does not apply to speculative-grade issuers. More importantly, they argue that the annual dummies in the model of Blume *et al.* are more likely picking up the effects of an omitted variable, such as a secular decrease in the value relevance of commonly used accounting ratios or increased earnings management for investment-grade firms.

of US firms (as measured by S&P issuer ratings) can bias my results against finding a positive relationship effect. To disentangle the secular change in credit quality from possible relationship benefits, all regressions include time dummies.

3.2 THE BOND SAMPLE

In addition to a sample of firm-years and issuer ratings, I collect a sample of bond months and issue ratings. In particular, I use the Lehman Brothers (now Barclay's) Bond Database (LBBD) as the source of monthly bond data. The LBBD, which has been extensively used in the finance literature,[15] provides month-end security-specific information on private and publicly traded corporate debt, including bid price, yield to maturity, S&P's and Moody's issue ratings, call and put provisions, and maturity for both investment- and noninvestment-grade debt. I search the LBBD for all straight bonds that were issued by the 1,263 firms in the firm sample over the period 1989–2005. As in Elton *et al.* (2001), I eliminate puttable bonds, bonds with a sinking fund, floating/variable rate bonds and zero-coupon bonds, bonds with a remaining maturity of less than 1 year or above 30 years, as well as bonds that are matrix priced. This results in a final sample of 54,656 monthly bond observations from 1,437 bonds issued by 543 firms. For a subsample of 992 bonds (issued by 415 firms), I obtain pricing and rating information for the end of the issuance month.

Table II summarizes the distribution of Relate in the three different samples: the firm sample, the bond sample at issuance, and the bond sample at month-end. Since the regressions below employ the natural logs of Relate (i.e., ln(Relate)), summary statistics for this variable are reported in parentheses to facilitate interpretation of the regressions. In any case, Relate measures the elapsed time in years between the start of the relationship, as identified in COMPUSTAT, and the end of the actual (fiscal) firm-year or bond month. As can be seen, mean and median relationship durations are higher in both bond samples compared to the firm sample, mainly due to the fact that firms with longer relationships issue bonds more frequently.

4. Empirical Results

4.1 RELATIONSHIP DURATION AND RATING LEVELS

Firm rating levels

The first test of rating quality and how this is affected by increased relationship duration is for the level of firm credit ratings. I regress firm ratings (coded 1

[15] See Hong and Warga (2000) and Elton et al. (2001) for a more detailed description of the LBBD.

Table II. Descriptive statistics for the distribution of Relate (and ln(Relate))

Relate measures the elapsed time in years between the start of the relationship, as identified in COMPUSTAT, and the end of the actual (fiscal) firm-year or bond month. The numbers in parentheses report summary statistics for the natural logs of Relate. The firm sample includes firm-years from 1,263 nonfinancial COMPUSTAT firms that (i) established a rating relationship with S&P for at least 3 years, (ii) the relationship starts in 1986 or later, (iii) have at least 2 years of nonmissing sales and total assets data in COMPUSTAT prior to the start of the relationship, and (iv) have no S&P-rated debt issue in SDC's Domestic New Issuances database prior to the start of the relationship. The bond sample includes bonds issued by these firms between 1989 and 2005. Monthly bond data are obtained from LBBD. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years.

Relate	Firm sample (1986–2005)	Bond sample (1989–2005)	
		At issuance	At month-end
Mean	6.09 (1.54)	6.92 (1.60)	8.31 (1.93)
Median	5.00 (1.61)	6.25 (1.83)	7.67 (2.04)
Standard deviation	4.22 (0.77)	4.68 (0.95)	4.54 (0.67)
25%	3.00 (1.10)	3.16 (1.15)	4.67 (1.54)
75%	8.00 (2.08)	10.37 (2.34)	11.49 (2.44)
Skewness	1.04 (−0.39)	0.58 (−1.03)	0.48 (−0.90)
Kurtosis	3.61 (2.46)	2.59 (3.56)	2.44 (3.87)
No. of observations	7,626	992	54,656
No. of firms	1,263	415	543
No. of bonds		992	1,437

for AAA, 2 for AA+, . . ., 17 for below B−) on Relate, using both ordered probit and OLS specifications. Results for the full sample are presented in the first three columns of Table III and for the later sample in the remaining columns. Standard errors are adjusted for heteroskedasticity and correlation within firm-level clusters since firm-years from the same firm are likely correlated.[16] Table III displays only coefficient estimates for the variable of interest, Relate, because the estimates for the control variables largely agree with the ones reported in previous literature (e.g., Blume, Lim, and MacKinlay, 1998; Jorion, Shi, and Zhang, 2009).[17] To allow for the possibility of diminishing marginal effects of additional years in a relationship, I specify the natural logs of Relate.[18] Consistent with rating relationships being

[16] By using simulations, Petersen (2009) shows that controlling for clustering produces correct standard errors when there is within-cluster correlation in the residuals.

[17] I follow Blume, Lim, and MacKinlay (1998) who argue that interest coverage (COV) should enter the equation in a nonlinear fashion. They set COV $= 0$ for negative entries and COV $= 100$ for entries above 100. Four variables are then set as follows: c1 $=$ COV for $0 \leq$ COV < 5 and c1 $= 5$ otherwise; c2 $=$ COV $- 5$ for $5 \leq$ COV < 10, c2 $= 0$ if COV < 5 and c2 $= 5$ if COV ≥ 10; c3 $=$ COV $- 10$ for $10 \leq$ COV < 20, c3 $= 0$ if COV < 10 and c3 $= 10$ if COV ≥ 20; c4 $=$ COV $- 20$ for $20 \leq$ COV and 0 otherwise. In addition, market value, firm age, and fraction rated are taken in logs.

[18] I also run robustness checks with Relate measured in levels, rather than logs, and with second-order terms in both the logs and the levels. However, the data do not prefer these alternative specifications.

Table III. Firm rating–level regressions

The dependent variable is S&P's long-term issuer credit rating (coded 1 for AAA, 2 for AA+, . . ., 17 for below B−). Firm controls are all variables listed in Table I. Interest coverage is treated in a nonlinear fashion, as defined in the text. Market value, firm age, and fraction rated are taken in logs. Industry fixed effects (FE) are on the one-digit SIC level. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. *** denotes significance at the 1% level.

	Dependent variable: firm credit rating (at fiscal year-end)					
Time period	1986–2005			1995–2005		
Regression model	Ordered probit (1)	OLS (2)	OLS (3)	Ordered probit (4)	OLS (5)	OLS (6)
ln(Relate)	−0.115***	−0.239***	−0.253***	−0.182***	−0.337***	−0.268***
	(0.038)	(0.066)	(0.093)	(0.044)	(0.074)	(0.090)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes		Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE			Yes			Yes
(Pseudo) R^2	0.261	0.728	0.950	0.269	0.737	0.956
n	7,626	7,626	7,626	5,935	5,935	5,935

valuable for firms, I find a significant negative coefficient on ln(Relate) across all specifications in Table III, suggesting that longer relationships are associated with favorable ratings. Importantly, observe that the coefficient on ln(Relate) remains negative and significantly different from zero when I relax the assumption of uniform differences in benefits (costs) between rating categories and employ ordered probit regression instead of OLS, and when I include firm fixed effects, which absorb any cross-sectional variation in the dependent variable (firm ratings are fairly stable). To get some insight into the economic significance of relationship effects, consider specification (3) that includes firm fixed effects (making industry fixed effects redundant).[19] The coefficient estimate of −0.253 indicates that firms with an 11-year agency relationship are predicted to have average ratings 0.61 notches higher (i.e., closer to AAA) than comparable firms that are the same in every observable way except that they only have a 1-year relationship. This corresponds to a one rating notch upgrade (e.g., from A+ to AA−) of approximately three of

[19] I use OLS instead of ordered probit because the statistical problem with nonlinear estimators containing fixed effects is that, as the number of groups (e.g., firms) tends to infinity, the number of estimated parameters increases at the same rate, producing inconsistent estimates unless I have a large number of observations within each group. This is the well-known incidental parameters problem (see, e.g., Greene, 2008, p. 800–806). Since the average number of sample observations per firm is merely 6.0, firm-level fixed effects ordered probit is not appropriate.

every five firms.[20] As expected, coefficient estimates are larger (in absolute terms) in the later sample for all three specifications, supporting the argument that results in the full sample are possibly downward biased due to an undersampling of long relationships in the first few years.

Bond rating levels

Up to this point, all regressions are based on S&P's fiscal year-end firm or issuer ratings. The use of annual (say, versus monthly) observations tries to minimize the inclusion of observations that would effectively lead to "double-counting," keeping in mind that much of the information on each firm is not updated frequently. However, since monitoring is costly and it is unlikely that the rating agencies can devote proper resources to examining all rated firms on a continuous basis, this could lead to staleness in ratings, meaning that the link between the rating and the factors that influence its determination might not truly reflect the decision-making behavior of the agency. To overcome this potential problem, I turn now to ratings of individual bonds at the time of bond issuance. When a firm issues a new rated public bond, we can be relatively certain that the firm has been recently investigated by the rating agency. Studying bond ratings at times of new debt issues should also allow me to more clearly identify the impact, if any, of managerial incentives on the rating agencies' decision process (see Hypothesis H2). A subsample of 992 bond ratings at issuance dates (month-ends) over the period 1989–2005 is extracted from LBBD, and balance sheet data for the fiscal quarter ending immediately before the issue date is obtained from quarterly COMPUSTAT. In addition, I control for the following bond features: offering amount, maturity, coupon, whether the bond is callable, and whether the issue is privately placed or public. In the first two columns of Table IV, I report the estimates of regressions of individual bond ratings at issuance on relationship duration. The results are slightly more pronounced compared to the ones for the firm sample. Focusing on the OLS specification (2), the coefficient indicates that ten more years of relationship (from 1 to 11 years) are expected to increase issuance ratings by an average of 0.74 notches, that is, three of four bonds will see an increased rating of one notch, a larger effect than that found for firm-level ratings. As a further robustness check, I estimate regressions for the full sample of monthly bond observations.

[20] Focusing on the ordered probit specifications (1) and (4), ten more years of relationship (i.e., from 1 year to 11 years) increase the probability of receiving an investment-grade rating by 11%–16%, all else equal. The predicted investment-grade probabilities are 0.348 versus 0.454 for the full sample and 0.296 versus 0.460 for the later sample. As a comparative value, a one standard deviation increase in log market value around its mean raises the predicted investment-grade probability by approximately 37% in both samples.

Table IV. Bond issue rating–level regressions

The dependent variable is S&P's issue credit rating (coded 1 for AAA, 2 for AA+, ..., 17 for below B−). The sample includes bonds issued between 1989 and 2005. Monthly bond data are obtained from LBBD. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years. Callable bonds are included. Firm controls are all variables listed in Table I. Interest coverage is treated in a nonlinear fashion, as defined in the text. Market value, firm age, and fraction rated are taken in logs. All accounting variables are measured at the previous fiscal quarter end just before the bond observation. Bond controls are log of offering amount (and its square), log of bond maturity (and its square), coupon, dummies for callable bonds and for private placements, and—in models (3) and (4) only—log of bond age (and its square). Industry fixed effects (FE) are on the one-digit SIC level. Date FE are on the year level in models (1) and (2) and on the year × month level in models (3) and (4). Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms (in models (1) and (2)) or on bonds (in models (3) and (4)). *** denotes significance at the 1% level.

	Dependent variable: bond issue credit rating			
Time	At issuance		At month-end	
Regression model	Ordered probit (1)	OLS (2)	OLS (3)	OLS (4)
ln(Relate)	−0.215*** (0.063)	−0.309*** (0.094)	−0.320*** (0.086)	−0.563*** (0.126)
Firm controls	Yes	Yes	Yes	Yes
Bond controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	
Date FE	Yes	Yes	Yes	Yes
Issue FE				Yes
(Pseudo) R^2	0.324	0.834	0.817	0.974
n	992	992	54,656	54,656

Since the same bond appears many times in the full sample (there are on average 38 monthly observations for each bond), I can include bond issue fixed effects and estimate the effect of Relate holding the subject of the rating fixed. In this specification, shown in the last column of Table IV, ln(Relate) is again negatively and significantly related to ratings. Moreover, the estimated effect is about 75% larger than without bond fixed effects. Overall, these results rule out the argument that the relationship effect is driven by any time-invariant differences between bonds (or firms).

Does relationship length proxy for firm age and size?

It could be the case that Relate is just proxying for a nonlinearity between the rating and the age of the firm or the rating and firm size, not controlled for in the previous regressions. To test this issue, I define three different size proxies (i.e., market equity, total assets, and sales) and four different age variables. Among the firm age proxies, first appearance counts the number of years since the firm's first appearance in COMPUSTAT with nonmissing (and positive) total assets. COMPUSTAT

IPO age controls for firm age since listing, which is available for 666 firms (3,460 firm years) in COMPUSTAT. The last two age proxies define age using founding dates (founding age) and IPO dates from Jay Ritter's web page, available for 605 firms (3,269 firm-years). Two functional forms are applied to each size and age proxy: logs (i.e., the natural logarithm) and levels (i.e., the first- and second-order term of the proxy measured in levels). For each of the resulting (6×8) 48 possibilities to combine one age and one size proxy at a time, I estimate the standard firm rating–level regression (the ordered probit specification (1) in Table III) and report the obtained coefficients for ln(Relate) in Table V. As can be seen, relationship benefits are robust to the particular age/size specification used. In all 48 regressions, the coefficient on ln(Relate) has the correct negative sign, and for 44 regressions the coefficient is significant at the 10% level or better. In unreported regressions, I also control for two age measures simultaneously (e.g., for COMPUSTAT age and founding age in logs or levels). The results for ln(Relate) are qualitatively unchanged. In sum, the above evidence indicates that a meaningful part of the variation in relationship length is unrelated to firm

Table V. Ordered probit firm rating–level regressions—alternative firm size and age measures

Each cell refers to one ordered probit specification (different rows (columns) represent regressions which differ in the firm size (age) proxy used). The dependent variable is S&P's long-term issuer credit rating (coded 1 for AAA, 2 for AA+, ..., 17 for below B−). Each regression includes year and industry fixed effects (FE) and all firm controls listed in Table I. For each regression, the coefficient estimate for ln(Relate) is reported. First appearance counts the number of years since the firm's first appearance in COMPUSTAT with nonmissing (and positive) total assets. COMPUSTAT IPO age controls for firm age since listing, which is available for 666 firms (3,460 firm-years) in COMPUSTAT. The last two age proxies define age using founding dates (Founding age) and IPO dates from Jay Ritter's web page, available for 605 firms (3,269 firm-years). Two functional forms are applied to each size and age proxy: logs (i.e., the natural logarithm) and levels (i.e., the first- and second-order term of the proxy measured in levels). Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Size \ Age		First appearance (COMPUSTAT)		IPO age (COMPUSTAT)		Founding age (Ritter)		IPO age (Ritter)	
		Logs	Levels	Logs	Levels	Logs	Levels	Logs	Levels
Market equity	Logs	−0.104***	−0.112***	−0.104*	−0.096*	−0.204***	−0.212***	−0.182***	−0.178***
		(0.039)	(0.038)	(0.058)	(0.057)	(0.053)	(0.052)	(0.057)	(0.057)
	Levels	−0.173***	−0.176***	−0.164***	−0.150***	−0.245***	−0.254***	−0.250***	−0.243***
		(0.038)	(0.037)	(0.062)	(0.061)	(0.053)	(0.053)	(0.056)	(0.057)
Total assets	Logs	−0.071*	−0.078**	−0.058	−0.043	−0.144***	−0.153***	−0.116**	−0.106*
		(0.040)	(0.039)	(0.058)	(0.057)	(0.055)	(0.054)	(0.058)	(0.058)
	Levels	−0.181***	−0.185***	−0.174***	−0.151***	−0.237***	−0.246***	−0.233***	−0.227***
		(0.039)	(0.038)	(0.063)	(0.061)	(0.054)	(0.054)	(0.057)	(0.057)
Sales	Logs	−0.077**	−0.079**	−0.036	−0.024	−0.152***	−0.160***	−0.138***	−0.132**
		(0.040)	(0.039)	(0.060)	(0.059)	(0.056)	(0.056)	(0.058)	(0.059)
	Levels	−0.164***	−0.168***	−0.165***	−0.136**	−0.228***	−0.239***	−0.230***	−0.224***
		(0.039)	(0.038)	(0.063)	(0.062)	(0.055)	(0.054)	(0.057)	(0.057)

age or size. Hence, ln(Relate) measures the length of a firm's commercial relationship with S&P and not just firm maturity.

4.2 RELATIONSHIP DURATION AND CUMULATIVE DEFAULT PROBABILITIES

To test whether relationship benefits reflect unobserved credit quality, I estimate default prediction models of various time horizons. Default information for the sample of firm-years over the period 1986–2005 is collected from COMPUSTAT and Moody's and S&P's annual default reports. The dependent binary variable in the probit regressions is equal to 1 for defaulting observations and 0 for surviving observations. For a given time horizon T, surviving observations are observations of firms surviving beyond T years and defaulting observations are observations of firms defaulting within T years. Hence, the probit regressions model T-year cumulative default rates (i.e., the probability of default in the coming T years).

Estimation results for $T = 3$, 5, and 7 years are shown in Table VI.[21] In a univariate context, longer agency relationships decrease cumulative default risk for all horizons, as predicted by the credit quality hypothesis. However, this negative relationship disappears and the coefficient of ln(Relate) becomes insignificantly different from zero when I control for the observable risk factors that are also contained in the rating-level regressions. This means that the positive effect of relationships on ratings established above coincides with precisely no trend in credit quality, a finding that disagrees with the credit quality hypothesis. However, standard errors are wide, and I can also not reject that there is a small negative effect of relationships on default risk (e.g., -0.05 is within the confidence interval for all three prediction horizons), too small to be detectable given the restricted number of defaults. A powerful rejection of the credit quality hypothesis would require that I find (larger) relationship benefits among firms with conditionally declining credit quality. Hence, the next section investigates relationship benefits among firms with decreasing credit quality as measured by subsequent increases in bond yield spreads.

4.3 RELATIONSHIP DURATION AND MANAGERIAL INCENTIVES

In this section, I test whether experience benefits are increasing at times when managers have pronounced incentives to game ratings (see Hypothesis H2).

[21] Since default information is available through the end of 2007, calculating cumulative default probabilities will lose observations at the end of the sample. In particular, the $T = 3$ results are estimated from the 1986–2004 sample, the $T = 5$ results from 1986–2002, and the $T = 7$ results from 1986–2000. The coefficient estimates for ln(Relate) in ordered probit firm rating–level regressions for these three subsamples amount to (p-values in parentheses) -0.109 (0.004), -0.101 (0.009), and -0.112 (0.008), respectively.

Table VI. Default prediction models

The dependent binary variable in the probit regressions is 1 for defaulting observations (firm-years from firms defaulting within the default prediction time horizon) and 0 for all surviving observations (firm-years from firms surviving within the default prediction time horizon). All models contain time and industry dummies. Since default information is available until the end of 2007, the parameters for the 3-year horizon are estimated from the 1986–2004 period, from 1986–2002 for the 5-year horizon, and from 1986–2000 for the 7-year horizon. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Default prediction time horizon	3 years		5 years		7 years	
	Model I	Model II	Model I	Model II	Model I	Model II
ln(Relate)	−0.125** (0.052)	0.056 (0.091)	−0.172*** (0.058)	−0.041 (0.091)	−0.233*** (0.068)	−0.142 (0.094)
Interest coverage						
c1		0.017 (0.054)		0.016 (0.054)		−0.020 (0.059)
c2		0.016 (0.050)		0.038 (0.048)		0.015 (0.046)
c3		−0.088 (0.066)		−0.055 (0.056)		−0.062 (0.047)
c4		0.018* (0.011)		0.008 (0.010)		0.018** (0.007)
Roa		−3.407*** (0.874)		−2.986*** (0.939)		−1.967** (0.975)
Total debt		0.665** (0.311)		0.730** (0.324)		0.655* (0.351)
ln(Market value)		−0.241*** (0.047)		−0.205*** (0.049)		−0.173*** (0.050)
Beta		0.132** (0.063)		0.107 (0.066)		0.143** (0.063)
Volatility		0.306*** (0.095)		0.290*** (0.096)		0.273** (0.111)
MtB		−0.312** (0.132)		−0.301*** (0.120)		−0.234** (0.111)
Tangible		0.031 (0.253)		−0.001 (0.263)		−0.068 (0.277)
ln(Age)		−0.196** (0.089)		−0.150* (0.090)		−0.126 (0.097)
ln(fraction rated)		1.448** (0.710)		1.900** (0.814)		2.459*** (0.915)
No. of defaults	216		271		260	
n	6,501		5,121		3,839	
Pseudo R^2	0.058	0.336	0.056	0.285	0.063	0.249

According to the credit quality hypothesis, positive relationship duration effects are driven by older firms being, on average, better credit risks. Hence, ln(Relate) proxies for some unobservable forward-looking credit quality measure. A powerful way to test this prediction is to look at relationship effects for firms that face increasing default risk *ex post*. According to the credit quality hypothesis, there should be no positive relationship benefits for firms with increasing future default risk. If agencies, moreover, are better able to assess the default risk of firms with whom they have longer relationships, higher values of Relate should indeed be associated with lower (i.e., closer to default) ratings, suggesting a positive coefficient for ln(Relate) in rating-level regressions for firms with decreasing future credit quality. If, on the other hand, successful selective information disclosure and/or an agencies' willingness to stretch ratings for client firms who provide a lot of business (i.e., issue a lot of debt) produces positive relationship benefits, managers of firms with private information signaling increasing default risk should have the strongest incentives to withhold/manipulate this information or to enhance the pressure on agencies, resulting in more positive relationship benefits for these firms.

To identify firms with bad private information (i.e., decreasing future credit quality), I focus on bond yield spread changes after issuance. In particular, I separate the 992 bonds for which I have the yield spread at issuance into a subsample of bonds with cumulative increasing spreads over t months after issuance and a sample of bonds with decreasing spreads over the same period. Over time and assuming sufficiently efficient bond markets, negative private information about a firm's future prospects becomes public and leads to increasing bond yield spreads and vice versa for positive private information. Hence, the direction of future bond yield spread changes should be an appropriate proxy for the content of private firm-specific information at bond issuance.

Table VII displays the results of ordered probit bond rating–level regressions separately for bonds with increasing credit spreads (hence increasing credit risk) over $t = 3, 6, 9, 12$ months after issuance and for bonds with decreasing spreads (hence, decreasing credit risk) over the same periods. The regressions are estimated for bond ratings at issuance since this should be the time at which managers care most about their firm's credit rating and incentives to manipulate information submission are strongest. As can be seen, the relationship effect varies across firm types: the coefficient is negative and highly significant for bonds from firms with decreasing credit quality but much smaller (up to 73% for $t = 6$) and in no case significant at the 1% level for bonds from firms with increasing credit quality. This agrees with the learning-to-gaming/adverse incentives view: firms with negative private information are more successful in hiding this information and/or capturing agencies when they have learned the effects on their rating of sharing certain information with the agency and/or when they are more frequent debt issuers

Table VII. Conditional (on future credit risk) bond rating–level regressions at issuance

The dependent variable in the ordered probit regressions is S&P's issue credit rating (coded 1 for AAA, 2 for AA+, …, 17 for below B−), recorded at issuance. The sample includes bonds issued between 1989 and 2005. Monthly bond data are obtained from LBBD. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years. Callable bonds are included. The final sample comprises 992 issues that are classified into bonds from firms with future increasing/decreasing credit risk according to their yield spread changes over t months after issuance. Firm controls are all variables listed in Table I. Interest coverage is treated in a nonlinear fashion, as defined in the text. Market value, firm age and fraction rated are taken in logs. All accounting variables are measured at the previous fiscal quarter end just before bond issuance. Bond controls are: log of offering amount (and its square), log of bond maturity (and its square), coupon, and dummies for callable bonds and for private placements. Industry fixed effects (FE) are on the one-digit SIC level. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Dependent variable: bond credit rating at issuance							
	Bonds with increasing credit spreads over t months				Bonds with decreasing credit spreads over t months			
	$t=3$	$t=6$	$t=9$	$t=12$	$t=3$	$t=6$	$t=9$	$t=12$
ln(Relate)	−0.326***	−0.319***	−0.323***	−0.226***	−0.169*	−0.086	−0.127	−0.213**
	(0.078)	(0.082)	(0.084)	(0.075)	(0.086)	(0.089)	(0.085)	(0.101)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bond controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.359	0.380	0.368	0.356	0.312	0.308	0.314	0.318
n	500	541	567	579	492	451	425	413

(i.e., long-term customers).[22] In contrast, for firms with positive information about their future default risk, relationships appear to be of no particular relevance.

4.4 RELATIONSHIP DURATION AND BOND YIELD SPREADS

Next, I investigate the correlation between bond yield spreads and relationship duration. If investors suspect that rating agencies intentionally favor frequent debt issuers or these issuers game their ratings upward and this goes unnoticed by the agency, investors should demand a higher yield spread to compensate for the higher

[22] In line with the credit quality hypothesis, it could be argued that among firms with increasing credit spreads, those with longer relationships display smaller increases, thereby justifying better ratings. However, when I regress cumulative spread changes, conditional on being positive, on ln(Relate) and control variables, the coefficient on ln(Relate) is small (negative for $t = 3, 9, 12$ and positive for $t = 6$) and insignificant in each case. Similar insignificant results are obtained for regressions of spread changes, conditional on being negative.

risk. To test whether investors demand a (price) discount on bonds from firms with longer relationships, I report results from yield spread regressions at issuance in the first two columns of Table VIII. As can be seen, the coefficients on ln(Relate) are positive and significant at the 5% level. According to specification (2), controlling for the rating and other firm/bond features, bonds from firms with 11 years of relationship have to pay on average 24 basis points (bp) more at issuance than bonds from firms with only one relationship year. In a second step, I ask whether relationship bonds still trade at a discount compared to equally rated bonds from nonrelationship firms when they age. To test for the persistence of price discounts, I include an interaction of ln(Relate) and log of bond age (in months) into monthly

Table VIII. Bond yield spreads and relationship duration

Each column presents the coefficient estimates from an OLS regression. The dependent variable is the yield spread (i.e., the yield to maturity minus the yield to maturity of the government bond with the closest maturity) of a bond. The spread distribution is truncated at the 1% and 99% level. The sample period is from 1989 until 2005. Monthly bond data are obtained from LBBD. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years. Callable bonds are included. Firm controls include log of firm age, total debt, log of market value, beta, and volatility. Bond controls are log of offering amount (and its square), log of bond maturity (and its square), coupon, and dummies for callable bonds and for private placements. Bond age is measured in months. Rating fixed effects (FE) are based on S&P's issue credit rating (coded 1 for AAA, 2 for AA+, . . ., 17 for below B−). Industry FE are on the one-digit SIC level. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms or on year × bond age clusters (in model (5)). *** and ** denote significance at the 1% and 5% levels, respectively.

	Dependent variable: bond yield spread				
	At issuance		At month-end		
Time	(1)	(2)	(3)	(4)	(5)
ln(Relate)	8.659**	9.932**	22.863***	32.394***	20.215***
	(3.975)	(4.248)	(6.555)	(11.717)	(2.527)
ln(Relate) ×			−7.145***	−5.063**	−5.354***
ln(1 + bond age)			(2.755)	(2.152)	(0.981)
ln(1+bond age)			20.294***	20.259***	
			(5.812)	(4.362)	
Firm controls	Yes	Yes	Yes	Yes	Yes
Bond controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes		Yes
Year FE	Yes	Yes	Yes	Yes	
Rating FE		Yes	Yes	Yes	Yes
Firm FE				Yes	
Year × bond age FE					Yes
R^2	0.823	0.827	0.680	0.760	0.694
n	992	992	53,469	53,469	53,469

bond yield spread regressions. The last three columns of Table VIII show that the coefficient on the interaction is negative and significant, indicating that relative price discounts at issuance are not persistent over the bond's life. For example, the coefficients from specification (3) imply that at issuance, the market demands a yield spread of 55 bp higher for bonds from relationship firms (with 11 years of relationship) compared to equally rated bonds from nonrelationship firms (with only 1 year), but this spread differential decreases to just 14 bp 10 months after issuance. The numbers are larger (78 versus 49 bp) when I include firm fixed effects in specification (4). Since Relate displays a time trend (see Figure 2), it could be argued that the coefficient on the interaction is simply capturing a temporal trend in how bond age relates to bond yields. The inclusion of year $\times$ bond age fixed effects in specification (5), however, shows that this is not the case. Note that the missing persistence of price discounts is consistent with two explanations: either rating agencies are aggressively downgrading misjudged relationship bonds to their spread implied rating levels after issuance or the market adjusts the prices of these bonds upward. Section 5.2 presents evidence in line with the first view.

4.5 RELATIONSHIP DURATION AND RATING INFORMATIVENESS

If a firm's incentive and ability to strategically disclose private information to agencies is an increasing function of relationship duration and/or agencies favor better (i.e., long-term) customers, the informativeness of ratings should decrease with ln(Relate) (see Hypothesis H3). I test this by examining the conditional correlation of bond yield spreads and ratings for varying levels of relationship length. In particular, I include an interaction of a bond's credit rating from S&P (the variable Rating, coded 1 for AAA, 2 for AA+, . . ., 17 for below B−) and a dummy Long, taking on the value 1 for bond months with relationship durations above the median value of 7.67 years in the bond sample, in a regression of bond yield spreads on bond/firm characteristics and time controls. Regression results are reported in Table IX. In all specifications examined, the coefficient on Rating is positive and significant, implying that bonds with better ratings pay lower spreads. The coefficient on the interaction of Rating and Long is negative and significant (in three of four cases), indicating that the correlation of credit ratings and bond spreads is lower when the agency–firm relationship is longer. The magnitude of this effect is economically large. For example, in specification (1), for bonds with relationship durations above 7.67 years, the coefficient on Rating is approximately 62% of the value when the relationship duration is below the median number. This is consistent with the view that issuer learning and selective information disclosure in long-term agency relationships reduce the information content of ratings. Alternatively, agencies may favor long-term customers by lowering rating standards.

Table IX. Bond yield spreads and ratings—the effect of relationship duration

Each column presents the coefficient estimates from an OLS regression. The dependent variable is the yield spread (i.e., the yield to maturity minus the yield to maturity of the government bond with the closest maturity) of a bond. The spread distribution is truncated at the 1% and 99% level. The sample period is from 1989 until 2005. Monthly bond data are obtained from LBBD. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years. Callable bonds are included. Rating denotes S&P's issue credit rating (coded 1 for AAA, 2 for AA+, . . ., 17 for below B−). Long is a dummy that takes the value 1 for bond months with relationship durations above the median value of 7.67 years in the bond sample and 0 otherwise. Firm controls are all variables listed in Table I. Interest coverage is treated in a nonlinear fashion, as defined in the text. Market value, firm age, and fraction rated are taken in logs. All accounting variables are measured at the previous fiscal quarter end just before the bond observation. Bond age is measured in months. Industry fixed effects (FE) are on the one-digit SIC level. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms (in models (1) and (2)), on firm × year clusters (in model (3)), or on rating × year clusters in model (4). *** and ** denote significance at the 1% and 5% levels, respectively.

	Dependent variable: bond yield spread at month-end			
	(1)	(2)	(3)	(4)
Rating	31.919*** (4.714)	32.868*** (8.007)	28.363*** (9.182)	
Rating × Long	−12.174*** (3.881)	−9.358** (4.626)	−4.253 (5.472)	−5.706** (2.589)
Long	114.235*** (34.657)	92.941** (45.180)	29.607 (45.631)	44.750** (21.618)
Firm controls	Yes	Yes	Yes	Yes
Bond controls	Yes	Yes	Yes	Yes
Industry FE	Yes			Yes
Year FE	Yes	Yes		
Bond age FE	Yes	Yes	Yes	Yes
Firm FE		Yes		
Firm × year FE			Yes	
Rating × year FE				Yes
R^2	0.673	0.799	0.882	0.754
n	53,469	53,469	53,469	53,469

The coefficient on the interaction is still negative when I include firm fixed effects or firm × year fixed effects but becomes insignificant in the latter case. This is likely due to the fact that firm × year fixed effects only leave within firm-year spread variation unexplained, and Long does not vary much within firm-years.[23] If ratings show a time trend in their relation to bond yields, this could generate a spurious finding because relationship duration is, on average, increasing over the sample period. Hence, in specification (4), I include Rating × year fixed effects to capture any time variation in the rating–yield spread relation. The estimated

[23] Only for 176 (or 6.5%) of the 2,706 firm-years, I observe a change in the value of Long (from 0 to 1) within the year. Consequently, Rating × Long interaction values are quite stable within firm-years, the average within percentage of monthly observations with equal interaction values amounts to 84.5%.

coefficient on the Rating × Long interaction is very similar to the one found in the previous specifications. In sum, the above results show that bond spreads are less related to credit ratings when the agency–firm relationship gets older. The implication is that credit ratings are less informative for spreads when relationships are longer or contain less spread-relevant information. This supports the learning-to-gaming/adverse incentives argument, predicting an adverse impact of relationship duration on rating quality (see Hypothesis H3), and is not consistent with the credit quality explanation.

Interestingly, since the coefficients for Long are all positive (and significant in three of four cases), Table IX complements the evidence in Table VIII that investors price the lower quality of bond ratings for firms with longer relationships. For instance, according to the estimates from specification (4), AA-rated bonds from long relationship firms trade at a 28 bp higher average spread than equally rated bonds sold by short relationship firms. Consistent with the notion that issuers or agencies game ratings upward, spread differentials decrease for ratings signaling lower credit quality (e.g., the differential is just 11 bp for A-rated bonds).

5. Robustness

5.1 ENDOGENEITY OF RELATIONSHIPS

The empirical strategy of this paper relies on exogenous variation in the duration of a firm's commercial relationship with S&P. However, variation in relationship length is mainly driven by a firm's decision to enter or exit the rated public bond market.[24] Hence, rating relationships are likely endogenous in the sense that a firm's rating history affects and predicts its future capital structure choices and use of rating agencies' services. One approach to address endogeneity of relationships is to find an instrument that is not subject to the same potential problem. I employ a measure of a firm's debt issuance history as an instrument for ln(Relate). In particular, I construct the variable No_issues (lagged by 1 year) that counts for each firm-year the cumulative number of past S&P-rated debt issues (according to SDC), conditional on the firm having at least one rated debt issue in the SDC's Domestic

[24] Hale and Santos (2008) study the timing of a firm's bond IPO decision. They find that more creditworthy firms with higher external funds demand and an established relationship with investment banks (through the issuance of private bonds or syndicated loans) issue their first public bond earlier. In addition, Faulkender and Petersen (2006) and Lemmon and Zender (2010) analyze what differentiates firms with access to the public bond market from firms without (i.e., rated from unrated firms). They find that larger and older firms, firms with more tangible assets and smaller MtB ratios, more leveraged firms with longer debt maturities, and firms with less volatile assets and stock returns are more likely to have a rating.

New Issuances database since the start of its rating relationship spell. The intuition of the instrument is that firms with more previous debt issues are also more likely to have tapped the public bond market for a longer period and, hence, experienced longer agency–firm relationships. Furthermore, since this measure is lagged by 1 year, any concern about endogeneity is much weaker. As expected, No_issues is highly correlated with relationship duration (the unconditional correlation is 0.436, significant at the 0.1% level). I estimate an ordered probit model with instrumental variables (No_issues and its square) in a two-stage process by following Nelson and Olsen (1978). To avoid concerns in the estimation of a nonlinear limited dependent variable model with endogenous variables, I alternatively treat every rating step as equal and estimate the resulting linear IV regression using the generalized method of moments (GMM) estimator.

The instrumental variable estimates are predicted to be smaller (in absolute terms) than those from the reduced-form regressions if there is negative correlation between ratings and relationship duration due to reverse causality (i.e., firms with worse ratings or anticipated downgrades choose to terminate their relationship earlier). In contrast, the IV estimate will be larger than the reduced-form one if reverse causality produces a positive correlation between ratings and relationship duration. For example, nontransparent, high-risk firms might benefit the most from additional monitoring services and information production by rating agencies, increasing their capital market access and debt capacity (Sufi, 2009). Hence, they are less likely to terminate their rating relationship than low-risk firms.

Results of reduced-form and IV rating-level regressions are reported in Table X. Columns 1 and 2 replicate the ordered probit and linear versions of the standard firm rating–level regression on the smaller sample with available information for No_issues. The corresponding IV regressions are presented in Columns 3 and 4, respectively. The coefficients for ln(Relate) are all negative and significant at the 1% level. Furthermore, the instruments are highly significant and have the predicted signs, assuring unbiased estimates for the effect of relationship duration on ratings.[25] Since the IV estimates are consistently larger (in absolute terms) than the reduced-form estimates, it is likely that some form of endogeneity is operating against my finding, suggesting that the true effect of relationships on ratings may in fact be larger than the reduced-form findings imply. Correspondingly, the results of a generalized Hausman test in the last row of Table X provide some evidence for ln(Relate) being truly endogenous.

[25] In the two-stage GMM case, the F test of the joint significance of the two instruments in the first-stage regression is 42.10 (p-value 0.00) and the partial R^2 amounts to 0.105.

Table X. Firm rating–level regressions—instrumental variables approach

The dependent variable is S&P's long-term issuer credit rating (coded 1 for AAA, 2 for AA+, . . ., 17 for below B−). Columns 1 and 2 show results for the reduced-form ordered probit and OLS regressions that take ln(Relate) as an exogenous variable. The linear specification treats every step of the rating system as equal. Columns 3 and 4 show results for the corresponding (nonlinear and linear) regressions that instrument ln(Relate) with No_issues and its square. No_issues measures for each firm-year the cumulative number of past S&P-rated debt issues (according to SDC), conditional on the firm having at least one rated debt issue in SDC's Domestic New Issuances database since the start of its rating relationship spell. The instruments are lagged by 1 year. Results from the first-stage regressions are not reported, except for No_issues and No_issues^2. The last row tests the null hypothesis that ln(Relate) can be treated as exogenous, using a generalized Hausman test (see Baum, Schaffer, and Stillman, 2003). Firm controls are all variables listed in Table I. Interest coverage is treated in a nonlinear fashion, as defined in the text. Market value, firm age, and fraction rated are taken in logs. Industry fixed effects (FE) are on the one-digit SIC level. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. *** denotes significance at the 1% level.

	Reduced-form		Instrumental variable regressions	
Regression model	Ordered probit (nonlinear)	OLS (linear)	Two-stage ordered probit (nonlinear)	Two-stage GMM (linear)
ln(Relate)	−0.285*** (0.067)	−0.465*** (0.101)	−0.803*** (0.293)	−1.525*** (0.444)
Firm controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
First stage				
No_issues			0.051*** (0.007)	0.051*** (0.007)
No_issues^2			−0.001*** (<0.001)	−0.001*** (<0.001)
n	3,901	3,901	3,901	3,901
Endogeneity test (p-value)			3.270 (0.071)	6.174 (0.013)

5.2 THE TIMING OF RATING CHANGES

If relationship firms receive upwardly biased ratings and agencies care about their reputation, we would expect to see these misleading ratings to come down as soon as information concerning a bond's true credit risk becomes public. Recall from Table VIII that relative bond price discounts at issuance are not persistent, consistent with the notion that S&P more aggressively downgrades relationship bonds with (at issue) yield spreads higher than their ratings imply.

To directly test for differences in the rating change behavior of bonds from relationship firms compared to nonrelationship firms, I estimate Cox proportional hazard models, separately for downgrades and upgrades. The dependent variable in these regressions is the time to the first rating change after bond issuance. Dropping 39 bonds that start life in the best (AAA) or worst (below B−) rating class, and, hence, cannot be upgraded or downgraded, leaves 953 bonds (438 downgrades, 299 upgrades, and 216 bonds with no observed rating change until December 2009), for which the regression results are displayed in Table XI. All firm and bond control variables are measured at issuance or immediately prior. Focusing first on downgrades in Column (1), we can see that the downgrade hazard of bonds from relationship firms (Relate = 11 years) is 43% (i.e., exp[0.148 × ln(11)]) higher than the downgrade hazard of bonds from nonrelationship (Relate = 1 year) firms. This difference is significant at the 10% level. However, I do not find any significant impact of relationship duration on a bonds' upgrade hazard. The coefficient in Column (3) is much smaller and insignificant. Whereas this evidence is consistent with more aggressive downgrades (but not upgrades) of relationship bonds, one possible objection concerns the fact that bonds of the same firm are typically simultaneously

Table XI. Cox proportional hazard model coefficient estimates for downgrades/upgrades

The dependent variable is the time to the first rating change after bond issuance. The sample includes bonds issued between 1989 and 2005. The rating history of bonds is observed until December 2009 or until the bond drops out from LBBD, whichever comes first. In models (1) and (2), bonds are censored when they experienced no rating change over the sample period or when the first change after issuance was an upgrade. In models (3) and (4), bonds are censored when they experienced no rating change over the sample period or when the first change after issuance was a downgrade. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years. Callable bonds are included. In addition, I drop all issues that start in the best (AAA) or worst (below B−) rating class because they cannot be upgraded or downgraded, respectively. The final sample includes 953 bonds (from 404 firms) of which 438 (299) were first downgraded (upgraded) after issuance. Models (2) and (4) include a gamma-distributed firm-level frailty to account for unobserved within-firm correlation. Firm controls include log of firm age, total debt, log of market value, beta, and volatility. Bond controls are log of offering amount (and its square), log of bond maturity (and its square), coupon, and dummies for callable bonds and for private placements. Industry fixed effects (FE) are on the one-digit SIC level. In models (1) and (3), standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. *** and * denote significance at the 1% and 10% levels, respectively.

	Dependent variable: time to first bond rating change			
	Downgrade hazard		Upgrade hazard	
Regression model	Cox (1)	Cox with shared frailty (2)	Cox (3)	Cox with shared frailty (4)
ln(Relate)	0.148* (0.080)	0.432*** (0.127)	0.039 (0.104)	0.051 (0.111)
Firm controls	Yes	Yes	Yes	Yes
Bond controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Issue year FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Frailty variance		2.339*** (0.345)		1.158*** (0.231)
n	953	953	953	953

downgraded or upgraded, and this is not adequately controlled for by firm cluster–adjusted standard errors. Hence, I alternatively model the within-firm correlation by assuming that the correlation is the result of a latent firm-level effect or frailty. Results from a Cox model with gamma-distributed frailty are contained in Columns (2) and (4), respectively. The estimates reinforce the previous findings: the downgrade hazard ratio between bonds from relationship and nonrelationship firms is now 2.82 (and significant at 1%), whereas the upgrade hazard ratio is still small (1.13) and insignificant. The significant values found for the frailty variance further support the shared frailty models, meaning that the correlation within firms cannot be ignored.

In a second set of tests, I focus on whether bonds sold by relationship firms are less sensitive to changes in market prices/yields than those sold by nonrelationship firms. In other words, I test for the tendency of downgrades/upgrades to follow rather than lead the market. If relationship benefits really reflect unobserved credit quality, downgrades should occur more slowly, that is, after greater increases in yield spreads, compared to similar bonds sold by nonrelationship firms. Table XII reports regression results of the cumulative (relative) yield spread change from issuance to the time of first downgrade or upgrade. The dependent variable equals the change between the (month-end) yield spread immediately before the first rating change and the yield spread at issuance. In these tests, I only include the bonds that received either a downgrade (Columns (1) and (2)) or an upgrade (Columns (3) and (4)); for bonds that have multiple rating changes over the sample period, I only look at the first. For downgrades, the relationship effect is negative and significant (at 1%), indicating that bonds from relationship firms (with Relate = 11 years) experience smaller (62% smaller in specification (2) including rating fixed effects) yield spread increases prior to first downgrade than those issued by nonrelationship (Relate = 1 year) firms. In other words, S&P acts faster to downgrade bonds sold by relationship firms. By contrast, relationship effects on spread decreases leading up rating upgrades are small and insignificant. In sum, the results from Table XII are consistent with S&P acting more aggressively to downgrade bonds issued by relationship firms to protect its reputation. This, in turn, is inconsistent with the notion that relationship benefits simply reflect credit quality.

Finally, in a recent study, Becker and Milbourn (2009) present empirical evidence that increased competition from Fitch leads to more issuer-friendly and less informative ratings from S&P. Relying on the Mergent Fixed Income Securities Database, the authors' measure of competition is the fraction of all bond ratings in a year–industry cell performed by Fitch. They show that over the sample period 1995–2006, Fitch's market share increased especially fast in the post-2000 period, mainly due to two acquisitions. This time pattern of increasing competitive pressure for S&P coincides with longer average relationships in my sample (see Figure 2)

Table XII. Cumulative yield spread changes up to the first rating change

Each column presents the coefficient estimates from an OLS regression. The dependent variable is the relative change between the yield spread of a bond in the month immediately before the month with the first rating change and the yield spread at issuance. All yield spreads are measured at month-ends. The first two specifications are for bonds with a downgrade as first rating change after issuance and the last two for bonds being upgraded after issuance. The cumulative spread change distribution is truncated at the 1% and 99% levels, separately for first downgrades and upgrades. The sample includes bonds issued between 1989 and 2005. Monthly bond data are obtained from LBBD. The price history of bonds is observed until December 2009 or until the bond drops out from the database, whichever comes first. I discard all bonds that are puttable, have sinking funds, have floating/variable or zero coupons, are matrix priced, or have a remaining maturity below 1 year or above 30 years. Callable bonds are included. Firm controls include log of firm age, total debt, log of market value, beta, and volatility. Bond controls are log of offering amount (and its square), log of bond maturity (and its square), coupon, and dummies for callable bonds and for private placements. Rating fixed effects (FE) are based on S&P's issue credit rating (coded 1 for AAA, 2 for AA+, ..., 17 for below B−). Industry FE are on the one-digit SIC level. Clustered standard errors (in parentheses) are robust standard errors adjusted for clustering on firms. *** denotes significance at the 1% level.

	Dependent variable: cumulative (relative) spread change			
	Downgrades		Upgrades	
	(1)	(2)	(3)	(4)
ln(Relate)	−0.268*** (0.102)	−0.259*** (0.100)	0.019 (0.025)	0.027 (0.024)
Firm controls	Yes	Yes	Yes	Yes
Bond controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Issue year FE	Yes	Yes	Yes	Yes
Rating FE		Yes		Yes
R^2	0.272	0.279	0.343	0.381
n	429	429	309	309

and higher (in absolute terms) coefficient estimates for ln(Relate) in firm rating–level regressions based on more recent data (see Table III). Hence, it might be argued that ln(Relate) is just capturing an assumed negative impact of increased competition on S&P's rating standards. In addition, if raters really favor relationship clients over new customers, as suggested by the adverse incentives view, we should see larger relationship benefits for firms in year–industry cells exposed to more competitive pressure by Fitch. I test this argument by including an interaction of ln(Relate) and Fitch's market share into the standard firm rating–level regressions (results not reported), using the same competition proxy as employed by Becker and Milbourn.[26] I find that the coefficient on the interaction is negative in all specifications, consistent with relationship benefits increasing with

[26] I am indebted to Bo Becker for sharing the variable Fitch's market share.

competitive pressure, but insignificant in any case. In contrast, the coefficient on ln(Relate) stays negative and highly significant. However, since the number of bond ratings issued by Fitch as a fraction of those issued by the three main raters in total is probably a noisy measure of competition, these results on the impact of competition are at best preliminary and require further research.

6. Conclusion

This paper documents the existence of relationship benefits in credit ratings. In particular, the paper shows that firms with longer rating agency relationships have better (i.e., closer to AAA) credit ratings. The paper provides several pieces of evidence supporting a learning-to-gaming and an adverse incentives explanation of the observed pattern but ruling out credit quality explanations. First, controlling for observables, the firms with longer relationships, while having higher ratings, do not have lower default rates. Second, relationship benefits are larger among firms with a higher incentive to game their information supplied to agencies or to pressure the agency into giving higher ratings (i.e., firms with bad future prospects as proxied for by *ex post* increasing bond yield spreads). Third, investors demand a (price) discount on bonds sold by relationship firms and the correlation between bond yield spreads and ratings (a measure of rating informativeness) is decreasing with relationship length. Finally, the higher downgrade hazards and smaller spread increases leading up downgrades of relationship bonds indicate that S&P indeed cares about its reputation and is more aggressively downgrading relationship bonds after issuance. Whereas the evidence presented here is in general inconsistent with the notion that relationship benefits simply reflect credit quality, but rather points to biased relationship ratings, it is beyond the scope of this paper (and an interesting topic for future research) to answer the question of who is the one of bad intent: the firm by consistently fooling S&P or the agency being more willing to stretch ratings for clients who bring a lot of business or both?

References

Adam, T. and Goyal, V. K. (2008) The investment opportunity set and its proxy variables: Theory and evidence, *Journal of Financial Research* **31**, 41–63.

Amato, J. D. and Furfine, C. H. (2004) Are credit ratings procyclical? *Journal of Banking and Finance* **28**, 2641–2677.

Arrow, K. (1962) The economic implications of learning by doing, *Review of Economic Studies* **29**, 155–173.

Baum, C. F., Schaffer, M. E., and Stillman, S. (2003) Instrumental variables and GMM: Estimation and testing, *Stata Journal* **3**, 1–31.

Beaver, W. H., Shakespeare, C., and Soliman, M. T. (2006) Differential properties in the ratings of certified versus non-certified bond-rating agencies, *Journal of Accounting and Economics* **42**, 303–334.

Becker, B. and Milbourn, T. (2009) Reputation and Competition: Evidence from the Credit Rating Industry, Working paper, Harvard Business School.

Bhattacharya, S. and Chiesa, G. (1995) Proprietary information, financial intermediation and research incentives, *Journal of Financial Intermediation* **4**, 328–357.

Blume, M. E., Lim, F., and MacKinlay, A. C. (1998) The declining credit quality of U.S. corporate debt: Myth or reality? *Journal of Finance* **53**, 1389–1413.

Bolton, P., Freixas, X., and Shapiro, J. (2009) The credit ratings game, Working paper, Columbia Business School and Universitat Pompeu Fabra.

Bongaerts, D., Cremers, M., and Goetzmann, W. N. (2009) Tiebreaker: Certification and multiple credit ratings, Working paper, Yale University.

Boot, A. W. A. (2000) Relationship banking: What do we know? *Journal of Financial Intermediation* **9**, 7–25.

DeFond, M. and Jiambalvo, J. (1994) Debt covenant violation and manipulation of accruals, *Journal of Accounting and Economics* **17**, 145–176.

DeLong, G. and DeYoung, R. (2007) Learning by observing: Information spillovers in the execution and valuation of commercial bank M&As, *Journal of Finance* **62**, 181–216.

Dietrich, J., Harris, M., and Muller, K. (2000) The reliability of investment property fair value estimates, *Journal of Accounting and Economics* **30**, 125–158.

Dimson, E. (1979) Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* **7**, 197–226.

Einhorn, E. and Ziv, A. (2008) Intertemporal dynamics of corporate voluntary disclosures, *Journal of Accounting Research* **46**, 567–589.

Elton, E., Gruber, M., Agrawal, D., and Mann, C. (2001) Explaining the rate spread on corporate bonds, *Journal of Finance* **56**, 247–277.

Faulkender, M. and Petersen, M. A. (2006) Does the source of capital affect capital structure, *Review of Financial Studies* **19**, 45–79.

Faure-Grimaud, A., Peyrache, E., and Quesada, L. (2009) The ownership of ratings, *Rand Journal of Economics* **40**, 234–257.

Frost, C. A. (1997) Disclosure policy choices of UK firms receiving modified audit reports, *Journal of Accounting and Economics* **23**, 163–187.

Gilson, S. (1989) Management turnover and financial distress, *Journal of Financial Economics* **25**, 241–262.

Greenbaum, S. I. and Thakor, A. V. (2007) Contemporary Financial Intermediation, Dryden Press, New York.

Greene, W. H. (2008) Econometric Analysis, Prentice Hall, Upper Saddle River, New Jersey.

Grossman, S. J., Kihlstrom, R. E., and Mirman, L. J. (1977) A Bayesian approach to the production of information and learning by doing, *Review of Economic Studies* **44**, 533–547.

Hale, G. and Santos, J. A. C. (2008) The decision to first enter the public bond market: The role of firm reputation, funding choices, and bank relationships, *Journal of Banking and Finance* **32**, 1928–1940.

Hong, G. and Warga, A. (2000) An empirical study of bond market transactions, *Financial Analysts Journal* **56**, 32–46.

Jorion, P., Liu, Z., and Shi, C. (2005) Informational effects of regulation FD: Evidence from rating agencies, *Journal of Financial Economics* **76**, 309–330.

Jorion, P., Shi, C., and Zhang, S. (2009) Tightening credit standards: The role of accounting quality, *Review of Accounting Studies* **14**, 123–160.

Kothari, S. P., Shu, S., and Wysocki, P. D. (2009) Do managers withhold bad news? *Journal of Accounting Research* **47**, 241–276.
Lemmon, M. and Zender, J. (2010) Debt capacity and tests of capital structure theories, *Journal of Financial and Quantitative Analysis* **45**, 1161–1187.
Livingston, M., Naranjo, A., and Zhou, L. (2007) Asset opaqueness and split bond ratings, *Financial Management* **36**, 49–62.
Mählmann, T. (2008) Rating agencies and the role of rating publication rights, *Journal of Banking and Finance* **32**, 2412–2422.
Mählmann, T. (2009) Multiple credit ratings, cost of debt and self-selection, *Journal of Business Finance and Accounting* **36**, 1228–1251.
Mathis, J., McAndrews, J., and Rochet, J.-C. (2009) Rating the raters: Are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics* **56**, 657–674.
Morgan, D. P. (2002) Rating banks: Risk and uncertainty in an opaque industry, *American Economic Review* **92**, 874–888.
Nelson, F. and Olsen, L. (1978) Specification and estimation of a simultaneous equation model with limited dependent variables, *International Economic Review* **19**, 695–705.
Pástor, L., Taylor, L., and Veronesi, P. (2009) Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability, *Review of Financial Studies* **22**, 3005–3046.
Petersen, M. A. (2009) Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* **22**, 435–480.
Rajan, R. G. (1992) Insiders and outsiders: The choice between relationship and arm's length debt, *Journal of Finance* **47**, 1367–1400.
Seru, A., Shumway, T., and Stoffman, N. (2010) Learning by trading, *Review of Financial Studies* **23**, 705–739.
Sharpe, S. (1990) Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships, *Journal of Finance* **45**, 1069–1087.
Standard & Poor's. 2008. *Corporate Ratings Criteria*. Standard & Poor's, New York.
Stocken, P. C. (2000) Credibility of voluntary disclosure, *The RAND Journal of Economics* **31**, 359–374.
Sufi, A. (2009) The real effects of debt certification: Evidence from the introduction of bank loan ratings, *Review of Financial Studies* **22**, 1659–1691.
Sweeney, A. (1994) Debt-covenant violations and managers' accounting responses, *Journal of Accounting and Economics* **17**, 281–308.
Weisbach, M. (1988) Outside directors and CEO turnover, *Journal of Financial Economics* **23**, 431–460.