ARE RATING AGENCIES STILL CREDIBLE AFTER THE SUBPRIME CRISIS?

Eleonora Isaia*, Marina Damilano*

*University of Turin, Department of Management, Turin, Italy

Abstract

Reputational concerns should discipline credit rating agencies (CRAs), eliminate any conflicts of interest, and motivate them to provide unbiased ratings. However, the recent financial crisis confirms models of CRAs' behavior that predict inflated ratings for complex products and during booms. We test whether CRAs suffered a reputational damage for this behavior. We find strong support in the data for our hypothesis. The stock price reaction to rating revisions is significantly lower after the financial crisis, particularly in the financial sector. In multivariate tests, we find that the stock price reaction is lower, on average, in the post-crisis period by 2.3%.

Keywords: Credit Rating Agencies, Subprime Crisis, Information Asymmetry, Reputational Damage
JEL Classification: G12, G14, G24

1. INTRODUCTION

The Financial Crisis Inquiry Report": "We conclude the failures of credit rating agencies were essential cogs in the wheel of financial destruction. The three credit rating agencies were key enablers of the financial meltdown."

The Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States published in 2011, concluded that the three major credit rating agencies (CRAs) had a key responsibility in the development of the recent financial crisis. The Report blamed the credit rating agencies for failing to properly evaluate the credit risk of collateralized securities, and even claimed that "This crisis could not have happened without the rating agencies." We ask the important question of whether the credit rating agencies suffered a reputational damage for their behavior.

Investors rely on CRAs to reduce information asymmetries in financial markets (Brealey, Leland, and Pyle (1977), Diamond (1984). Reputational concerns should discipline CRAs, eliminate any conflicts of interest, and motivate them to provide unbiased ratings (Schwarcz (2002). Upward biased ratings would in fact ex post hurt their reputation and lead to the loss of future business.

However, due to different incentives, CRAs could, ex-ante, optimally deviate from minimizing information asymmetries. There could be equilibrium outcomes where the credit rating agencies optimally inflate their ratings. Mathis, McAndrews, and Rochet (2009) provide a theoretical framework where the reputation mechanism disciplines CRAs actions only if rating complex products is not a major source of income. But in the years leading to the financial crisis, the business of rating complex products boomed, and this likely weakened the CRAs' reputational concerns. According to the predictions of the model, Cornaggia and Cornaggia (2013) empirically document that issuer-paid agencies tend to cater to issuers' interests and understate credit risk. CRAs did inflate their ratings and this, as underlined by the Financial Crisis Inquiry Report, critically contributed to the recent financial meltdown.

Furthermore, in the model of Bolton, Freixas, and Shapiro (2012) CRAs are more likely to inflate ratings during economic booms. This happens because during booms a larger clientele of investors (naive investors) take ratings at face value and because during booms there is also a lower risk of issuer failure that could damage the raters' reputation. This results in profits from rating inflation that are larger than the expected reputation costs. However, the collapse of the housing market triggered an extremely high (and unexpected) rate of failure among issuers.

In this paper, we empirically test whether the rating agencies suffered, ex-post, a reputational damage. To identify the effects of the reputational damage, we investigate the informational content of credit ratings by looking at the stock price reactions to rating announcements before and after the crisis. If rating agencies suffered a reputational damage, we should observe a weaker stock price reaction to rating announcements.

Using a sample of downgrades and upgrades relative to the biggest fifty stocks of the American, European, and Asian markets over the period 2003-2013, we find results that provide strong support for our hypothesis. We find that the stock price reaction is significantly lower after the financial crisis. The effect is both statistically and economically significant. In multivariate tests, we find that the stock price reaction in the post-crisis period is, on average, lower by 2.3%. These results provide evidence consistent with a significant reputational damage suffered by the rating agencies. Rating agencies had weaker incentives to protect their reputational capital in the wake of the financial crisis due to the increasing importance of rating complex products. This led to inflated ratings and to a reputational damage following the financial meltdown.

The second hypothesis we test is that the CRAs' reputational damage has been particularly severe in the financial sector where the rating

agencies made the biggest mistakes. Our results strongly support this hypothesis. In univariate results and in multivariate tests, the reduction of the stock price reaction is always stronger in the subsample of financial institutions.

Our results are important to better understand the incentives of CRAs and to complement the reputational mechanism with more regulatory oversight. Rating agencies play a fundamental role in the financial market. However, the reputational concerns could not be enough to discipline their actions.

2. LITERATURE REVIEW

This paper relates to two main strands of the literature. On the one hand, it is related to the literature about CRAs' reputation as a disciplining mechanism as in Schwarcz (2002). It is motivated by the theoretical models of Bolton, Freixas, and Shapiro (2012), Mathis, McAndrews, and Rochet (2009), and Bar-Isaac and Shapiro (2013) that predict inflated ratings during booms. The role of reputation is critical for financial intermediaries. Gopalan, Nanda, and Yerramilli (2011) claim that the poor performance of borrowers damages the lead arranger's reputation. In the context of rating agencies, we show that the poor performance of the rated collateralized securities, led to a severe reputational damage.

On the other hand, it contributes to the literature on the information content of credit ratings. Jorion, Liu, and Shi (2005) claim that the information content of downgrades and upgrades is greater in the post-Regulation Fair Disclosure period. Xia (2014) documents a deterioration in the incumbents' ratings quality following the entry of a third investor-paid credit rating agency. Grothe (2013) claim that downgrades have, on average, a greater impact than upgrades have. Afonso, Furceri, and Gomes (2012) show a significant response of government bond yield spreads to rating changes, particularly for negative announcements.

3. DATA AND SUMMARY STATISTICS

To assess the impact of the reputational shock on the information content of ratings, we analyze the stock market reaction to rating revisions before, during, and after the financial crisis. We define the start of the financial crisis with the bankruptcy of Lehman Brothers (September, 15th, 2008). The post-crisis period starts from October 16th, 2009 when the CBOE volatility index (VIX) came back to the pre-crisis mean level. The sample consists of 1,153 rating revisions from November 1st 2003 to November 1st 2013.

We consider the rating revisions of the three major rating agencies, i.e. Standard & Poor's (S&P), Moody's Investors Service, and Fitch Inc. The rating revisions are related to the 150 most representative companies traded on the American, European, and Asian/Pacific markets. In particular, we look at the components of the STOXX USA 50, EURO STOXX 50, and STOXX Asia/Pacific 50.

Afonso, Furceri, and Gomes (2012) perform a numerical transformation of the alphanumerical rating codes. Following a similar procedure, we standardize the conventional alpha-numerical scale that goes from excellent to poor: AAA (or similar) is equal to 17, while CCC+ (or similar) is equal to 1. The positive and negative watches are equal to -0.25 and + 0.25 respectively.

To perform multivariate tests, we need to control for variables that previous research claimed to be related to the level of abnormal stock market reactions. We proxy the issuer's stock risk using the standard deviation of the daily stock returns (DEVST) calculated over the 50 days leading to the rating action. We use the level of the VIX index to proxy for expected risk in the market (Bollerslev, Tauchen, and Zhou (2009)). We obtain the rating revisions and other market data from the Bloomberg database.

The summary statistics are reported in Table 1. All data are winsorized at the 1% and 99% levels. It is interesting to note that 22% of the rating actions are anticipated by a previous watch (Steiner and Heinke (2001)). In robustness, we control for this effect. On average, the starting rating is equal to 12.10 (equivalent to the A rating). Finally, the absolute size of the rating change is slightly bigger for downgrades (0.74 versus 0.67).

Table 1. Summary statistics

	Mean	Median	P25	P75
DEVST (%)	1.87	1.55	1.13	2.17
VIX	18.58	17.57	10.34	25.26
DUMMY_ANTICIP	0.22	0	0	0
LAST_RATING	12.10	12	10.75	14
CHANGE_NOTCH (+)	0.67	0.75	0.25	1
CHANGE_NOTCH (-)	-0.74	-0.75	-1	-0.25

Note: This table reports descriptive statistics for the sample of rating revisions over the period 2003-2013. The table reports the mean, the median, the 25th and the 75th percentile of the control variables used in the multivariate analysis. DEVST is the standard deviation of the daily stock returns calculated over the 50 days leading to the rating action. VIX is the value of the VIX index on the announcement day of the rating revision. DUMMY_ANTICIP is a dummy variable that is equal to one if the rating action is the realization of the previous forecasted watch. LAST_RATING is the value of the rating before the rating action. CHANGE_NOTCH (+) and CHANGE_NOTCH (-) is the positive and negative rating change, respectively.

4. EMPIRICAL ANALYSIS

The objective of this study is to test whether the major CRAs suffered a reputational damage after the subprime crisis. Ratings are signals sent by the CRAs to the market. The magnitude of these signals is observable and is given by the value of the ratings. However, the precision and accuracy of these signals cannot be observed directly. The CRAs' reputation is critical for these signals to be credible. A CRAs' better reputation reinforces the market's beliefs of the precision and accuracy of the CRAs' ratings.

To test whether the CRAs suffered a reputational damage, we empirically estimate the information content of rating revisions before and after the financial crisis. Rating changes are signals of new information about the issuers quality. If CRAs suffered a reputational damage, we should see lower stock price reactions for rating changes after the financial crisis. To capture the effect of the reputational channel on the market's beliefs about the precision of the signals, in univariate and multivariate tests, we employ a standard event study methodology, and we compare the stock price reaction to rating changes before and after the financial crisis. Cumulative abnormal returns (CARs) are calculated over a three-day window centered on the announcement date. We consider the absolute values of the Cumulative Abnormal Returns, CARs (Grothe (2013). This allows to isolate the beliefs of the precision of the signals independently from the information that is already incorporated in the prices before the rating actions.

4.1. Univariate analysis

Table 2 reports the mean absolute value of the CARs around the rating actions for investment grade firms. Panel A reports CARs before and after the crisis. Panel B repeats the same analysis excluding the crisis period.

The univariate evidence supports our hypothesis. Even without controlling for other factors, after the crisis the market reaction is statistically and economically significantly lower. Panel A shows that the average difference is equal to 1.2% that, compared to the average value of the stock price reaction before the end of the crisis of 3.5%, represents a decrease of 34%.

In line with previous research (Holthausen and Leftwich (1986)) the stock price reaction is bigger for downgrades compared to upgrades (4.1% vs. 2.2%). This confirms that the information content of downgrades is higher compared to upgrades. Consistent with the reputational damage hypothesis, it is exactly for downgrades that the reduction is bigger. The decrease is equal to 1.7%, a reduction of 41% compared to the mean value of 4.1% before the end of the subprime crisis.

Panel B shows qualitatively similar results excluding the crisis period. The overall difference before and after the crisis is equal to 0.7%, that compared to the average value pre-crisis of 2.9% represents a decrease of 24%. The decrease for downgrades only is equal to 1.2%, a reduction of 33% compared to the pre-crisis mean value of 3.6%.

Table 2. Univariate results

		Downgrades	Upgrades
	(1)	(2)	(3)
Panel A. Entire sample			
Pre-Crisis + Crisis	0.035	0.041	0.022
	N=551	N=362	N=157
Post-Crisis	0.022	0.024	0.019
	N=261	N=130	N=107
Difference (Post - Pre)	-0.012***	-0.017***	-0.003
Panel B. Entire sample			
Pre-Crisis	0.029	0.036	0.018
	N=485	N=305	N=152
Post-Crisis	0.022	0.024	0.019
	N=261	N=130	N=107
Difference	-0.007***	-0.012***	-0.001
(Post - Pre)			
Panel C. Subsample financials			
Pre-Crisis	0.035	0.045	0.017
	N=204	N=127	N=72
Post-Crisis	0.020	0.022	0.016
	N=81	N=45	N=31
Difference (Post Pre)	-0.015***	-0.023***	-0.001
Panel D. Subsample financials			
Pre-Crisis + Crisis	0.045	0.054	0.025
	N=252	N=170	N=76
Post-Crisis	0.020	0.022	0.017
	N=81	N=45	N=31
Difference (Post – Pre)	-0.025***	-0.032***	-0.007

*Note: Panel A reports the means of ABS_CAR for rating revisions of investment grade firms (no border) before and after the financial crisis. Column (1) considers the entire sample. Column (2) only downgrades, Column (3) only upgrades. The number of observations are reported under the mean values. The last row reports the difference of the mean value post- and the mean value pre-crisis. Panel B includes the crisis together with the pre-crisis period. ***,**,* indicate statistical significance at the 1%, 5% and 10% level, respectively.*

Finally, Panel C and Panel D repeats the same analysis of Panel A and B, respectively, on the subsample of financial institutions. The results are consistent with the hypothesis that the CRAs suffered a bigger reputational damage in the financial sector. Both pooling downgrades and upgrades together and looking at downgrades alone, the stock price reaction is negative and always bigger (in absolute terms) compared to the entire sample.

These preliminary results suggest that the CRAs' reputational damage had a significant impact on how the market interprets the signals sent by CRAs. However, there could be determinants of the CARs correlated with the timing of the crisis that could drive the results. In the following section we examine the impact of the reputational damage in a multivariate framework.

4.2. Multivariate analysis and robustness

In order to evaluate the impact of the reputational damage, we need to control for variables that previous studies have identified as determinants of the CARs. We estimate the following model for the entire sample, separately for downgrades and upgrades, and for the subsample of financial institutions to test our hypotheses:

$$ABS\_CAR_i=\alpha+\beta_1 POST\_CRISIS_i+\beta_2 NO\_BORDER_i+\beta_3 POST\_CRISIS_{ix} NO\_BORDER_i+\beta_4 VIX_i+ \beta_5 DEVST+ \delta_i+\varepsilon_i \quad (1)$$

where, for the rating change $ABS\_CAR_{i,t}$ is the absolute value of the cumulative abnormal return around the rating change of firm *i* at time t; $POST\_CRISIS_{i,t}$ is a dummy variable that equals one if the rating change of firm *i* at time *t* happens after the financial crisis, and zero otherwise; $NO\_BORDER_{i,t}$ is a dummy variable that equals one if the rating change of firm *i* at time *t* concerns a company whose last or current ratings are investment-grade; VIX_t is the CBOE volatility index at time t; $DEVST_i$ is the standard deviation of the issuer's daily stock returns calculated over the 50 days leading to the rating action of firm *i* at time t. To control for unobservable characteristics specific to the American, European, and Asian/Pacific regions we include region fixed effects (δ_i).

The reputation hypothesis predicts a negative coefficient on the variable POST_CRISIS, which suggests that after the financial crisis the information content of the rating changes is lower. When the variable NO_BORDER and the interaction term NO_BORDER x POST_CRISIS are included, our hypothesis predicts a negative value for the sum of the coefficients β_1 and β_3. This sum represents the differential mean value of the absolute CAR before and after the financial crisis for investment-grade companies.

Table 3 reports the results. Panel A considers the entire sample. The coefficient on the variable POST_CRISIS is negative and both statistically and economically significant. After the financial crisis the mean absolute value of the CAR is 2.3% lower (Model (1). Controlling for stock-specific risk and market-expected risk (Model (2), the mean absolute value of the CAR is 2.4% lower. Model (3) allows to isolate the effect on investment-grade companies looking at the sum of the coefficient on POST_CRISIS and the interaction term NO_BORDER x POST_CRISIS. The estimated value (-2.3%) does not change the previous results.

Panel B considers the downgrades and the upgrades separately. Consistent with univariate results, the reduction of the information content of rating changes is greater for downgrades compared to upgrades. This further supports the reputation hypothesis. It is exactly where the CRAs' reputation is more valuable, i.e. for downgrades as documented by previous research (Holthausen and Leftwich (1986)), that the market's beliefs about the precision of the CRAs' signals have been significantly revised after the financial crisis.

Table 3. Main Results

	Dep. Variable: ABS_CAR			Dep. Variable: ABS_CAR			
				Downgrades	Upgrades	Downgrades	Upgrades
	(1)	(2)	(3)	(1)	(2)	(3)	(4)
Panel A: Entire sample				**Panel B: Downgrades and Upgrades**			
Intercept	0.0534***	0.042***	0.036***	**0.063*** **	**0.021*** **	0.057***	**0.021*** **
	(0.002)	(0.005)	(0.005)	**(0.007)**	**(0.006)**	(0.007)	**(0.006)**
POST_CRISIS	-0.023***	-0.024***	-0.012**	**-0.035*** **	**-0.011*** **	**-0.007**	**-0.010***
	(0.004)	(0.001)	(0.006)	**(0.005)**	**(0.005)**	**(0.013)**	**(0.006)**
NO_BORDER			0.002			**0.002**	**-0.001**
			(0.002)			**(0.004)**	**(0.003)**
NO_BORDERx			-0.011*			**-0.030** **	**-0.001**
POST_CRISIS			(0.006)			**(0.013)**	**(0.005)**
VIX		0.001***	0.001	**0.001*** **	**0.001*** **	0.001***	**0.001*** **
		(0.001)	(0.001)	**(0.001)**	**(0.001)**	(0.001)	**(0.001)**
DEVST		0.446***	0.526***	**0.560*** **	**0.399*** **	0.456)**	**0.395**
		(0.172)	(0.168)	**(0.210)**	**(0.255)**	(0.210)	**(0.259)**
Region FE	Yes	Yes	Yes	**Yes**	**Yes**	**Yes**	**Yes**
N	1,153	1,105	1,071	**587**	**414**	587	**414**
R^2	0.02	0.16	0.16	**0.14**	**0.18**	0.16	**0.18**

*Note: This table reports the results of following OLS model (1). The dependent variable is the absolute value of issuers' CAR around the rating revision. POST_CRISIS is a dummy variable that is equal to one after October 15th, 2009 and zero otherwise. The Region Fixed Effects refer to the American, European, and Asia/Pacific regions. Panel A considers the entire sample. Panel B analyses the downgrades and the upgrades separately. For each model the robust standard errors are reported in parenthesis under the coefficients. The final rows of each column report the number of observations and R2. ***,**,* indicate statistical significance at the 1%, 5% and 10% level, respectively.*

Table 4 reports the results on the subsample of financial institutions. The main takeaway of the table is that the reduction of the information content of rating changes is bigger for the financial institutions. After the financial crisis the mean absolute value of the CAR is 3.3% lower (Model (1). The reduction for downgrades (4.5%) is also bigger if compared to the entire sample. This, together with the univariate results, supports the hypothesis that the reputational damage has been more severe for financial institutions, where the CRAs made the biggest mistakes.

Table 4. Subsample of financial institutions

	Dep. Variable: ABS CAR		
	(1)	Downgrades (2)	Upgrades (3)
Intercept	0.053***	0.073***	-0.007
	(0.011)	(0.015)	(0.010)
POST_CRISIS	-0.051***	-0.066***	0.008
	(0.012)	(0.016)	(0.011)
NO_BORDER	-0.011	-0.014	0.021
	(0.009)	(0.010)	(0.008)
NO_BORDERx	0.018*	-0.021	-0.013
POST_CRISIS	(0.011)	(0.015)	(0.009)
VIX	0.002***	0.001***	0.005****
	(0.001)	(0.001)	(0.001)
DEVST	0.512*	0.778***	-0.404
	(0.220)	(0.272)	(0.481)
Region FE	Yes	Yes	Yes
N	347	223	115
R^2	0.29	0.22	0.58

*Note: This table reports the results of the subsample of the financial institutions. Model (1) considers both downgrades and upgrades. Models (2) and (3) analyse the downgrades and the upgrades separately. For each model the robust standard errors are reported in parenthesis under the coefficients. The final rows of each column report the number of observations and R2. ***,**,* indicate statistical significance at the 1%, 5% and 10% level, respectively.*

Table 5 reports some robustness tests. Model (1) explicitly controls for the rating actions that are simply the realization of previous watches. We include a dummy variable (DUMMY_ANTICIP) that is equal to one if the rating action is the realization of the forecasted watch. Finally, we also control for the size of the rating change (Model (2). The results are qualitatively unchanged.

Table 5. Robustness

	Dep. Variable: ABS_CAR	
	(1)	(2)
Intercept	0.040***	0.040***
	(0.005)	(0.004)
POST_CRISIS	-0.015**	-0.013**
	(0.006)	(0.006)
NO_BORDER	0.002	0.001
	(0.003)	(0.003)
NO_BORDERx	-0.012**	-0.013**
POST_CRISIS	(0.006)	(0.006)
VIX	0.001***	0.001***
	(0.001)	(0.000)
DEVST	0.464**	0.422**
	(0.164)	(0.164)
DUMMY_ANTICIP	-0.003	
	(0.003)	
CHANGE_NOTCH		-0.040***
		(0.001)
Region FE	Yes	Yes
N	1,004	1,004
R^2	0.16	0.18

*Note: This table reports the results of robustness tests. Model (1) includes a dummy variable (DUMMY_ANTICIP) that is equal to one if the rating action is the realization of the forecasted watch. Models (2) controls for the size of the rating change (CHANGE_NOTCH). For each model the robust standard errors are reported in parenthesis under the coefficients. The final rows of each column report the number of observations and R2. ***,**,* indicate statistical significance at the 1%, 5% and 10% level, respectively.*

5. CONCLUSION

We investigate whether the credit rating agencies are still credible after the financial crisis. In particular, we investigate whether the rating agencies suffered a reputational damage after the crisis. Our results strongly support the hypothesis that they suffered a reputational damage. The empirical evidence also supports the hypothesis that the damage has been particularly severe for financial institutions, where they made the biggest mistakes.

Our results are important to better understand the behavior of CRAs. Rating agencies play a fundamental role in the financial market. However, the reputational concerns could not be enough to discipline their actions.

REFERENCES

1. Afonso, A., D. Furceri and P.Gomes. (2012), "Sovereign Credit Ratings and Financial Markets Linkages: Application to European Data." Journal of International Money and Finance, No.31(3), pp. 606-638.
2. Heski, B.-I. and J. Shapiro. (2013), "Ratings Quality over the Business Cycle." Journal of Financial Economics, No. 108(1), pp. 62-78.
3. Bollerslev, T., G. Tauchen and H. Zhou. (2009), "Expected Stock Returns and Variance Risk Premia." Review of Financial Studies, No.22(11), 4463-4492.
4. Bolton, P., Xavier F. and J. Shapiro. (2012), "The Credit Ratings Game." Journal of Finance, No.67(1), pp. 85-111.
5. Brealey, R., H.E. Leland and D.H. Pyle. (1977), "Informational Asymmetries, Financial Structure, and Financial Intermediation." The Journal of Finance, 32(2), pp. 371-87.
6. Cornaggia, J. and K. J. Cornaggia. (2013), "Estimating the Costs of Issuer-Paid Credit Ratings." Review of Financial Studies, 26(9), pp. 2229-2269.
7. Diamond, D., 1984, "Financial Intermediation and Delegated Monitoring." Review of Economic Studies, No.51(3), pp. 393-414.
8. Gopalan, R., N.Vikram and Y.Vijay. (2011), "Does Poor Performance Damage the Reputation of Financial Intermediaries? Evidence from the Loan Syndication Market." Journal of Finance, No.66(6), pp. 2083-2120.
9. Grothe, M. (2013), "Market Pricing of Credit Rating Signals.", ECB Working Paper
10. Holthausen, R. W. and R. W. Leftwich. (1986), "The Effect of Bond Rating Changes on Common Stock Prices." Journal of Financial Economics, No.17(1), pp. 57-89.
11. Jorion, P., Z. Liu and C. Shi. (2005), "Informational Effects of Regulation FD: Evidence from Rating Agencies." Journal of Financial Economics,

No.76(2), pp.309-330.
12. Mathis, J., J. McAndrews and J. C. Rochet. (2009), "Rating the Raters: Are Reputation Concerns Powerful Enough to Discipline Rating Agencies?" Journal of Monetary Economics, No.56(5), pp.657-674.
13. Schwarcz, S. (2002), "Private Ordering of Public Markets: The Rating Agency Paradox." University of Illinois Law Review, 1-27.
14. Steiner, M. and V. G. Heinke. (2001), "Event Study Concerning International Bond Price Effects of Credit Rating Actions." International Journal of Finance and Economics, No6(2), pp.139-157.
15. Xia, H. (2014), "Can Investor-Paid Credit Rating Agencies Improve the Information Quality of Issuer-Paid Rating Agenciesa." Journal of Financial Economics, No. 111(2), pp.450-68.