ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Rieber, Alexander; Schechinger, Steffen

Conference Paper

Herding Behavior between Rating Agencies

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Empirical Finance, No. G05-V1

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Rieber, Alexander; Schechinger, Steffen (2019) : Herding Behavior between Rating Agencies, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Empirical Finance, No. G05-V1, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:
https://hdl.handle.net/10419/203580

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Herding Behavior between Rating Agencies

Alexander Rieber Steffen Schechinger

February 21, 2019

Abstract

We investigate whether credit rating agencies systematically follow each otherâ€™s rating decisions. Therefore we rely on the rotation of rating analysts within credit rating agencies and their impact on the rating. Using this institutional setup we can disentangle causal herding behavior from simple co-movement between credit rating agencies due to changes in firm fundamentals. Rating analysts have substantial influence on ratings and we use their individual optimism/pessimism as instrumental variable to estimate causal effects of a rating change induced by an analyst on rating changes by other credit rating agencies. For our comprehensive sample of U.S. and European firms, rated by S&P, Moody's and Fitch between 1995 - 2016, we find significant herding behavior among credit rating agencies. This average herding behavior amounts to 0.4 notches for a one notch change at another credit rating agency, which is roughly half the size of the simple co-movement between credit rating agencies.

1 Introduction

The market for credit rating agencies (CRAs) is highly concentrated, with three CRAs controlling 95% of its share.[1] This oligopoly could be problematic, especially when ratings of different CRAs are not independent of each other, i.e. if one CRA follows another CRA in its rating decisions. The economic literature dealing with the independency of CRAs documents such a rating convergence between different CRAs, see e.g. Güttler (2011) or Norden and Weber (2009). However, CRAs themselves proclaim not to include the rating information from other CRAs.

[1] S&P and Moody's form a quasi-duopoly, controlling 80% of the international market and Fitch controls another 15%, see e.g. https://www.reuters.com/article/uscorpbonds-ratings/big-three-in-credit-ratings-still-dominate-business-idUSL2N17U1L4

They argue to face serious reputation concerns, keeping them away from reacting to outside rating actions (see e.g. Cantor and Packer (1995) or Becker and Milbourn (2011)). To establish a causal interpretation of the convergence between CRAs is tough because all CRAs use the same fundamental data for their rating decisions and (sequential) changes of ratings could just reflect a (sequential) re-evaluation of a given firm by each CRA, arriving at the same rating decision. This is why our study explicitly deals with herding between CRAs. In our definition, herding is the part of co-movement between CRAs that is not based on changes in fundamental information other than the rating itself.

There are mainly two reasons herding is likely to occur between CRAs, especially with reputation concerns, see e.g. Mathis et al. (2009) or Mariano (2012) for theoretical studies. First, CRAs should take all available information into account when providing a rating and the rating of another agency is a valuable information which probably increases the rating accuracy. Second, the rating of another CRA can have a direct effect on the credibility of the rated firm, because it might influence its capital access (see e.g. Cantor and Packer (1995), Devenow and Welch (1996) and Lugo et al. (2014)). However, herding among CRAs is problematic because it undermines the independence of ratings. If ratings are independent and their correlation is only based on fundamental information, their value is higher compared to dependent ratings, where the follow up rating already includes the information from another CRA. This is why herding is socially inefficient[2].

However, previous empirical studies focused mainly on the convergence between CRAs because of the causality problem inherent in the estimation of herding,, see e.g. Güttler (2011), Norden and Weber (2009), Alsakka and Ap Gwilym (2010), Lugo et al. (2014). These studies employ a leader-follower analysis, which does not allow for a causal interpretation of the dependence structure between CRAs. If changes in the evaluation of fundamental data happen sequentially, we could not differentiate a reaction to fundamental changes from herding using a leader-follower analysis. As a result, these studies document differences in the evaluation time of CRAs rather than herding.

In contrast to the previously mentioned empirical studies in the credit rating literature, we employ an instrumental variable approach to test if CRAs herd. We use credit rating analysts as instruments and exploit their influence on corporate ratings. Therefore we directly build on the work by Fracassi et al. (2016) which highlight the influence of individual analysts on corporate ratings. Rating analysts might interpret the same fundamental information about a firm differently, some

[2]Scharfstein and Stein (1990) make similar arguments for individual investors.

rating analysts are generally more optimistic, others more pessimistic, than their peers, as shown in Fracassi et al. (2016). Analysts are quasi randomly assigned to (new) firms within a sector, i.e. CRAs allocate (new) firms to analysts depending on an analysts actual workload and not based on the current rating of a firm or the revenue a CRA can gain from a rated issuer.[3] In our empirical setting we take advantage of the regular rotation of analysts and their quasi-random assignment to new firms. Given, for example, a more pessimistic analyst rotates to a firm where an optimistic analyst previously issued an optimistic rating. After some time the more pessimistic analyst might revise the assigned rating, not because firm fundamentals changed, but because he judges the same fundamentals differently. Such a rating change should not lead to a rating change by another CRA. However, if another CRA reacts to such an analyst induced rating change we have evidence for herding. Our analysis is inspired by recent studies using fixed effects of individuals as instruments. An exemplary study closely related to ours is Aizer and Doyle Jr (2015), who use the influence of judges to estimate causal effects of juvenile incarceration on high school completion and adult recidivism. Judge fixed effects are frequently used as instruments (see e.g. Chang and Schoar (2013), Dobbie and Song (2015), Frandsen et al. (2019) and for a resent overview . Another closely related example is Silver (2016) (physician fixed effects) or more distant related Angrist and Krueger (1991) (month of birth).

We employ a comprehensive sample of more than 100 000 rating actions of Moody's, S&P and Fitch in the period from 1995 until 2016 to analyze the rating assessment of CRAs. We include US-American and EU-European companies in our sample, which have at least one rating action between 1995 and 2016 and are rated by two or more CRAs.

In our empirical analysis, all three major CRAs are affected by the rating of another CRA, beyond a simple co-movement based on fundamental information. These herding effects are sizeable and relevant; they range from 0.3 to 0.5 notches. This means a one notch rating difference induced by an analyst (absent from changes in firm fundamentals) leads to an average rating reaction by other CRAs in the range of roughly 0.3 to 0.5 notches in the same direction. This result supports the existing literature which also implies that the amount of convergence is too high to be based

[3]Moody's and Fitch confirmed to us via e-mail that they assign new firms to analysts with free capacity within a sector, which is in line with statements presented by Fracassi et al. (2016). Further the International Organization of Securities Commissions (IOSCO (2015)) issued a Code of Conduct which explicitly state that analysts are not compensated or evaluated by the amount of revenue that the CRA derives from that obligor(IOSCO (International Organization of Securities Commissions (2015), point 2.12 a). Moody's (2019), S&P (2018) and Fitch (2017) include comparable statements into their Code of Conduct.

solely on fundamental information, see e.g. Güttler and Wahrenburg (2007).
First, this result is relevant for market participants as the informational value of a follow up rating of a different CRA is significantly lower when CRAs herd. Second, this result is relevant for firms rated by more than one CRA because the issuer pays model is prevalent in the rating industry and the value of such a second rating is lower, compared to the initial rating. Third, this result is relevant for lawmakers, because they determine the requirements for insurance or bank contracts often relying on ratings. Fourth, this result is relevant for CRAs, because they might want to prevent herding behavior[4].

2 Data

The core of our database is information on credit ratings from the three major CRAs Moody's, S&P and Fitch. We obtain senior issuer ratings from Moody's Default and Recovery Database, Fitch Connect and S&P Capital-IQ. Our main objective is to study the connection between different CRAs and therefore we opt to use senior issuer ratings instead of bond ratings, because senior issuer ratings provide an overall assessment of a firmâ€™s credit quality. By using rating information from each CRA directly, we get the exact date of a rating change for all firms covered by each CRA. We transform the alphanumeric rating provided by each CRA into a numeric scale following Fracassi et al. (2016).[5] To avoid further confusion note, that the best possible rating is 1 and the worst 21. So a lower rating number corresponds to a better rating.
Alongside each rating action CRAs issue a rating report in which they explain the objective behind their action. In these reports the name(s) and job title(s) of the responsible rating analyst(s) are announced. Each of the three big CRAs, Moody's, S&P and Fitch, provide a comprehensive list of rating reports for each firm they currently cover or have covered in the past. We consider all publicly available rating reports and the inherent analyst information for Moody's and Fitch, including reports on rating, watchlist and outlook changes. Unfortunately, historic rating reports by S&P are not publicly available and we could not obtain analyst information for S&P.
We consider all reports between 1995 and 2016 labeled "Rating Action" by Moody's[6]

[4]There are rating committees ultimately responsible for the rating of firms which should prevent individual analysts from herding. For more details see S&P (2018), Moody's (2019) and Fitch (2017).
[5]In the appendix we provide table 4 with this conversion
[6]We download the reports from www.moodys.com.

and all reports between 1999 and 2016 labeled "Rating action commentary" and "Non-rating action commentary" by Fitch[7]. We include only English reports into our analysis. There are some days with more than one report for the same firm by the same CRA. If this is the case and there is conflicting analyst information, we take a four step procedure to filter out the main report of the responsible analysts: (i) We delete reports with less than 10 words, because they typically refer to another main report; (ii) we keep only the last report written on this day, if headlines are identical; (iii) we delete methodology changes, because these reports refer to more than one firm and are often not written by the responsible analysts for the respective firm. We identify methodology changes by keywords in their headlines.[8]; (iv) for the days where we still observe more than one report we exclude the whole day, because we cannot perfectly define if there was an analyst change or not.

After deleting these methodology changes, we extract and adjust analyst names for our analysis. We correct analyst names : (i) misspelled names, e.g. Willaim Norris instead of William Norris, after making sure that this analyst is indeed the one we are looking for. To this end, we check the period the probably misspelled analyst has been working for the CRA and which firms he had covered. (ii) Some analysts use their nickname from time to time, e.g. Bill Norris instead of William Norris, or (iii) they include their middle name in some reports, e.g. William J. Norris instead of William Norris. We exclude these middle names and make sure to capture the same analyst with a nickname and his full name. We adjust the nickname if we are sure that the analyst is indeed the one we are looking for. The unique number of analysts and number of firms covered by each CRA is shown in Table 1.

Sample selection

First, we match analyst information to rating, outlook and watchlist information provided by each CRA. Second, we combine the information from Moody's, Fitch and S&P, using their Legal Entity Identifier (LEI) Information, CUSIP and Ticker. The LEI is a global 20-digit alphanumeric code uniquely identifying each company.We obtain a list with all firms in Compustat North-America and Compustat Global, their name, CUSIP, LEI, ticker and GVKEY information from S&P Capital-IQ In Moody's Default and Recovery Database we obtain 6-digit CUSIP and Ticker, Fitch Connect includes LEI and 6-digit CUSIP. In a first step, we match Moody's with S&P information based on their CUSIP. In a second step, we include previously unmatched companies from Moody's using their ticker. However, ticker match is

[7]We download the reports from www.fitchratings.com.

[8]The keywords we use are in Appendix table 5.

error-prone, which is why we double check each ticker match manually and keep it only if the name and business description of both companies match, otherwise we drop the company. We include information from Fitch using company LEI and, for companies without LEI, their CUSIP. Finally, we restrict our sample to US-American and EU-European firms and exclude sovereigns. Our final sample consists of 104 339 event observations on 7 787 firms.

Out of this event data we construct a quarterly panel based on GVKEY and quarter, using the last available senior issuer rating in the firm-quarter as the quarterly rating for the firm by CRA. We do not consider the information on outlook or watchlistings, however, we use the reports dealing with such actions. By using the report information we can assess if there was an analyst change in a given quarter in absence of a rating change. However, when building the panel we do not have continuous information about analysts between two reports. Comparable to senior issuer ratings, we use the last available analyst information in a given firm-quarter as the information on the analyst in this quarter. Yet, we opt to consider an analyst only up to the point where he has written his overall last report at the CRA. Employing this approach we set analyst information for all firm-quarters to missing if an analyst issued his last report at the CRA and no new rating report is released yet, i.e. we have no new analyst information. Because we cannot be sure which analyst is responsible for a firmâ€™s rating during the described firm-quarters we opt to exclude them.

Our identification of herding is based on whether the impact of analysts on the rating of their own CRA continues further and impacts the rating of other CRAs. The first necessary condition to be fulfilled is that analysts from a CRA have an impact on the rating of their own CRA. Fracassi et al. (2016) already addressed this issue. They show, that analysts have a tendency to be generally more optimistic or pessimistic with respect to the firms they rate. This general perception significantly influence a firmâ€™s rating.[9]. Different to their study we use separated firm and year-quarter fixed effects instead of combined firm×year-quarter fixed effects, further we include the second analyst in our analysis. When we assess the joint significance for first and second analysts within our setup, we obtain F-statistics of 25.22 for Moody's and 14.47 for Fitch.[10] Due to regulatory standards in Europe, set by EU (2009), analysts must rotate the firms they rate. For example the lead ana-

[9]We replicate the analysis of Fracassi et al. (2016) with our data and obtain similar results with respect to analyst influence on ratings. Using their approach to identify the joint significance of the analyst effect on ratings of Moody's or Fitch we obtain F-Statistics of 9.48, compared to Fracassi et al. (2016) with an F-statistic of 10.56.

[10]Moody's and Fitch) confirmed to us via e-mail that they assign new firms to analysts with free capacity within a sector, which is in line with statements presented by Fracassi et al. (2016).

lysts for a firm must change after a maximum of four years, committee chairpersons after a maximum of seven years. S&P included such a rotation system worldwide, Moody's and Fitch established it only where regulatory requirements forced them to do so.[11]

Table 1: Number of observations and statistics for main variables

	N of firms	Avg. Rating	Analyst 1	Analyst 2	Rating Obs.
For each CRA separately					
S&P	6952	10.24			225247
Moody's	4438	10.94	579	541	147715
Fitch	1122	7.95	645	672	47132
Rated by Moody's and S&P					
Moody's	3512	11.23	502	472	101092
S&P	3512	10.66			101092
Rated by Moody's and Fitch					
Moodys	808	8.52	356	315	29986
Fitch	808	8.13	503	530	29986
Rated by Fitch and S&P					
S&P	856	8.31			31287
Fitch	856	8.10	537	588	31287
Rated by all three CRAs					
Moody's	675	8.66	293	274	22883
Fitch	675	8.34	446	477	22883
S&P	675	8.54			22883

The Table presents descriptive statistics for all CRAs in our sample between 1995 and 2016. It is separated into five parts, where the first part presents results for each CRA separately and in the other four segments dependent for firms rated by two or three CRAs together. *Moody's/Fitch/S&P* denote the observations obtained for the respective CRA. *N of firms* presents the number of firms covered by the CRAs. *Avg. Rating* represents the average rating of all firm-quarters, *Analyst 1* denotes the number of distinct first analysts in our sample period and *Analyst 2* the distinct number of second analysts. *Rating Obs.* represent the number of firm-quarter rating observations.

In Table 1 we present descriptive statistics of our data. We split the table into five parts. In every part we present the data for the number of firms rated by each CRA, the average rating given by each CRA, the number of distinct first analysts and the

[11]For information on the rotation requirements see EU (2009) and the Code of Conduct by each CRA.

number of distinct second analysts. As we study the first and second analysts together we also present the distinct analyst teams observable in our data for each CRA. The last column represents the number of firm×year-quarter observations for which a rating from the CRA is available.

Most firms are rated by S&P, least by Fitch. Hence, observations decrease when we include Fitch into our analysis. S&P and Moody's have a comparable rating distribution, where Moody's is slightly more conservative in their ratings. Remember that a â€žhigherâ€œrating number corresponds to a more negative rating. At first sight, it seems that Fitch has the most lenient rating of all CRAs, which is far better than the rating by S&P or Moody's. But this impression occurs only, because (a) Fitch rates a lower number of firms and (b) they rate predominantly firms that have a higher creditworthiness, i.e. their portfolio of rated firms is less diverse. Appendix Figure 1 provides the rating distribution for all CRAs in our sample. For all subgroups, depending on a Fitch rating being available, the rating distributions align (see also Figure 2 in the Appendix). Fitch has more analysts than Moody's in our sample, despite the fact that (a) we have information about Moody's analysts since 1995 and for Fitch analysts since 1999 and (b) that Moody's rates more firms than Fitch. These statistics suggest that there is a higher fluctuation of analysts at Fitch than at Moody's. With 1 120 first and second analyst combined for Moody's, rating 4 438 different firms and 1 317 first and second analysts combined at Fitch, rating 1 122 different firms, there are many possible shifts in analyst coverage, i.e. many instruments for our regression[12].

3 Empirical Framework and Results

From a theoretical point of view, there are two main reasons why herding among CRAs could be rational. First, CRAs want to provide the best possible rating to investors. This is only possible if they take all available information into account. Until 2010, CRAs had an exemption from Regulation FD (which was reversed in 2010 due to the Dodd-Frank Wall Street Reform and Consumer Protection Act)[13]. Due to this exemption they had access to private information from the firm they rate, a privilege e.g. stock analysts lost due to Regulation FD. With private information they had an advantage over other market participants judging the creditworthiness of a firm. Nevertheless, for an optimal rating decision CRAs should take valuable

[12]Note that the actual number of instruments in our regressions is lower. By design we exclude analyst from our analysis which only rate one firm. The next chapter covers this in more detail

[13]For details on the Dodd-Frank Wall Street Reform and Consumer Protection Act see Dodd-Frank (2010).

publicly available information also into account. One piece of publicly available information is the rating of another CRA, which is an assessment of a firm's creditworthiness and therefore an information worthwhile to consider. Second, ratings themselves present a fundamental information for a rated firm. A rating change directly effects the credibility of a firm, which has to pay higher interest rates for new bond issues. Devenow and Welch (1996) and Lugo et al. (2014) addressed the point that ratings influence a firmâ€™s access to capital. Under this circumstance it seems to be rational for CRAs to include ratings of other agencies into their own assessments. Yet we argue that herding among CRAs is socially inefficient because it distorts the informational value of a subsequent rating from a different CRA.

In our empirical framework we use an instrumental variable approach to find herding between CRAs. While other approaches, like the leader-follower analysis, e.g. from Güttler (2011), Alsakka and Ap Gwilym (2010) or Lugo et al. (2014), are very useful to study convergence and dynamics between rating agencies, we doubt that they are sufficient to detect a causal dependence between CRAs. As changes in firm fundamentals should lead to rating changes we doubt that a leader-follower analysis warrants a causal identification of herding because these rating changes might happen sequentially by each CRA. Hence, a time lag is arguably not enough to make a causal claim for herding.

This is why we focus on the causal part of ratings convergence by exploiting analysts' substantial influence on corporate ratings, outlined in Fracassi et al. (2016). Fracassi et al. (2016) show that some rating analysts are generally more optimistic or pessimistic in their rating decisions than their peers. We enhance the analysis of Fracassi et al. (2016) and estimate the effect an analyst has on a CRA he does not work for. Rating differences based purely on analystsâ€™ perception should not influence the rating of another CRA. Yet, if other CRAs react to a pure analyst induced rating change, we have evidence for herding. Hence, these analysts' fixed effects are suitable instruments because they influence their own agencies' rating but are not correlated with firm fundamentals.

Model

In our firm-year-quarter panel we have rating data on Fitch, Moody's and S&P and analyst data on Moody's and Fitch. We exploit the analyst variation within firm j and year-quarter t in a rating agency i. In the second stage we estimate the rating $ratg_{ijt}$ of firm i on the estimated rating $\widehat{ratg}_{-ijt}$ of agencies $-i$ for firm j at quarter

t including firm and year-quarter fixed effects [14]:

$$ratg_{ijt} = \beta_0 + \beta_{-i} \cdot \widehat{ratg}_{-ijt} + \alpha_j + \alpha_t + \epsilon_{ijt} \tag{1}$$

The typical first stage of a two stage least squares would be $\widehat{ratg}_{-ijt} = \hat{\beta}_0 + \hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\alpha}_j + \hat{\alpha}_t$ with the analyst fixed effects γ_1 and γ_2 for the first and second analyst if the firm is rated by these analysts and zero otherwise. This approach would lead to an unbiased estimation for β_{-i} in the second stage if we possess the true analyst effects. Unfortunately, there is no way of knowing the true effects. If we attempt to estimate them, high ratings which are correlated with the ratings of the other CRAs might lead to potentially higher estimated analyst effects. In this case we would obtain results where the fitted values in the first stage and the error of the second stage ϵ_{ijt} are correlated.

To circumvent the probability that results from the first and second stage are correlated, we use a jackknife instrumental variable estimation, short JIVE[15]. This approach is suitable in cases with many instruments where the number of instruments increases with the number of observations[16].

In a standard setup JIVE is used by leaving one observation out and estimating the coefficients without this observation. In our specific setting this might not be enough due to the structure of our panel, where ratings and analysts stay the same for (many) consecutive quarters. Therefore, we decided to use a "leave one firm out" approach, i.e. we estimate the individual analyst effects in the first stage without the observations of firm j. The estimated rating $\widehat{ratg}_{-ijt}$ of agency $-i$ for firm j at time t is than estimated by

$$\widehat{ratg}_{-ijt} = \hat{\beta}_0 + \hat{\gamma}_1^{-j} + \hat{\gamma}_2^{-j} + \hat{\alpha}_j + \hat{\alpha}_t \tag{2}$$

with the analyst fixed effects $\hat{\gamma}_1^{-j}$ and $\hat{\gamma}_2^{-j}$ if the firm j is rated by these analysts at time t and zero otherwise. These analyst fixed effects are estimated with

$$\widehat{ratg}_{-i-jt} = \hat{\alpha}_{-i} + \hat{\gamma}_1^{-j} + \hat{\gamma}_2^{-j} + \hat{\alpha}_{-j} + \hat{\alpha}_t \tag{3}$$

Since$\hat{\gamma}^{-j}$ is a consistent estimator for γ, the coefficient in the second stage $\hat{\beta}_{-i}$ is a consistent estimator for the herding effect β_{-i}.

Note that our first stage regression is different from the one used by Fracassi et al. (2016), even without jackknife. They estimate analyst optimism or pessimism always

[14] Either for one or both other agencies

[15] See e.g. Quenouille (1956), Blomquist and Dahlberg (1999) and Angrist et al. (1999)

[16] See also Stock et al. (2002) and Kolesár et al. (2015)

relative to another CRA which rates the same firm at the same time. They control for fundamental information using firm × year-quarter fixed effects and exploit the variation of ratings between different CRAs. Unfortunately, we cannot use the same approach if we want to estimate the effect one CRA has on another in the second stage. Hence, our first stage, without the jackknife, is similar to Bertrand and Schoar (2003) who estimate manager fixed effects.

Due to our conservative approach when setting up our panel, where we include an analyst only up to the quarter he wrote his last report within the CRA, we have some firm-year-quarters with missing analysts. If this is the case the analyst probably left the CRA, but there was no new report for that firm, i.e. no new analyst name, and therefore these firm-year-quarters drop out of our estimation. This is why the number of observations might be lower for our JIVE estimation than for the OLS estimation.

Results

Table 2 reports the second stage results. The first row reports the herding effect of the Fitch rating ($\widehat{ratg}$ Fitch), the second row the herding effect of the Moody's rating ($\widehat{ratg}$ Moody's). Because we don't have analyst data for S&P we cannot estimate its effect on the other agencies. All Regressions include firm and year-quarter fixed effects and report bootstrapped standard errors.

Table 2: Estimated herding effect: JIVE estimation

	Rating				
	Moody's	S&P	Fitch	S&P	S&P
$\widehat{ratg}$ Fitch	0.478^{***}	0.444^{***}			0.343^{***}
	(0.0611)	(0.0496)			(0.0616)
$\widehat{ratg}$ Moody's			0.407^{***}	0.467^{***}	0.484^{***}
			(0.0455)	(0.0314)	(0.0552)
Fime FE	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	25713	26078	20681	73441	13429
Adjusted R^2	0.081	0.112	0.088	0.112	0.165

All regressions include firm and year-quarter fixed effects. Bootstrapped standard errors in parentheses with 2000 draws. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

For example, consider the first column of Table 2; it displays the weighted average herding effect Fitch has on Moody's ratings. If an analyst effect changes the rating

at Fitch by one notch, Moody's changes its rating on average by 0.478 notches. Remember that this change is not induced by a change in firm fundamentals but rather due to a different assessment of the firm by a new Fitch analyst. In general, we find significant herding effects, roughly between 0.4 and 0.5 notches, for Fitch and Moody's. These sizeable effects indicate that analysts do not only have a significant influence on their own rating, but on ratings of other CRAs as well. The herding coefficients are approximately between half and two-third the size of the standard OLS coefficients, see Table 3.

The OLS coefficients in Table 3 represent the general co-movement between CRA ratings in our panel. It subsumes all fundamental information, herding and other influences. While the first part (I) of Table 3 shows the overall co-movement, the second part (II) reports the co-movement accounted for firm and year-quarter fixed effects. If we compare the herding effect from Table 2 to the general co-movement it is roughly half its size, i.e. there is a stronger relationship between firm ratings when firm fundamentals change compared to a situation where only the analyst changes. Given the size of the herding effect, especially compared to the simple co-movement, we argue that herding among CRAs is sizeable and economically meaningful.

Interestingly, all CRAs herd. Moody's and Fitch affect the other two CRAs and each CRA is affected by another. Unfortunately we cannot make any statement on the herding in response to a S&P rating because we could not obtain analyst data. However, given the size of the herding effects we argue they are relevant, yet we do not claim that these herding reactions are irrational. There are reputational or informational concerns which can cause rational herding. A downgrade by Moody's might weaken the credibility of a firm even if this downgrade is not based on changes in firm fundamentals. Therefore, it is not irrational if other CRAs adjust their ratings as well. It is however important to understand the amount of independent information contained in a rating. These independence is questioned by our results. Certainly, herding does not mean the agencies just copy each other. There are ways imaginable which are much more subtle or even unconscious. For example, it might be enough to trigger another assessment of the firm to find these effects, as Güttler and Wahrenburg (2007) point out.

Robustness

In this robustness section we split our database according to region or firms which are only rated by two CRAs. Due to this splits the number of observations decreases, which might be problematic when interpreting the estimated coefficients. To assess the stability of our estimates we show the coefficients $\hat{\beta}_{-i}$ of the regressions depending

Table 3: Estimated co-movement effect: Standard OLS regressions

	Rating				
	Moody's	S&P	Fitch	S&P	S&P
	(I) without fixed effects				
ratg Fitch	0.954***	0.883***			0.609***
	(0.0148)	(0.0110)			(0.0370)
ratg Moody's			0.843***	0.815***	0.270***
			(0.0155)	(0.00434)	(0.0320)
Firm FE	No	No	No	No	No
Year-Quarter FE	No	No	No	No	No
Adjusted R^2	0.804	0.872	0.804	0.875	0.887
	(I) with year-quarter and firm fixed effects				
ratg Fitch	0.716***	0.742***			0.520***
	(0.0341)	(0.0255)			(0.0340)
ratg Moody's			0.567***	0.656***	0.291***
			(0.0301)	(0.0123)	(0.0260)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.441	0.595	0.453	0.540	0.650
Observations	29986	31287	29986	101092	22883

(I) Robust standard errors and (II) Firm-cluster robust standard errors in parentheses: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

on the number of observations in the second stage in Figure 3 in the Appendix.. We drew different sized random subsample out of our firms and estimated the coefficients for the respective subsample. We find that below 10 000 observations the variation of coefficients becomes relatively large. Hence, we should interpret coefficients based on a lower number of observations with care.

For the sake of brevity and clarity the results discussed in this section are provided in the Appendix. In our main results we use the lead analyst and the second analyst as instruments. It is not perfectly clear if we should use both analysts or just the lead analyst as instrument because the lead analyst writes the review report for a firm, which is subsequently presented to the rating committee. So arguably the influence of a lead analyst on a firm rating is higher than the influence of a second analyst. This is why we use the lead analyst alone as instrument and present the

results in Table 6 in the Appendix. Our main results hold and are not dependent on the choice of instrumenting the analysts.
Further our results hold if we include additional fixed effects as control variables into our main regression. Specifically we include industry fixed effects and whether the firms are listed in the US and/or Europe into our regression (see Table 7 in the Appendix). The US and Europe are two different jurisdictions with different regulatory requirements, e.g. the need to rotate lead analysts every four years in Europe. Therefore we conduct an additional robustness test where we split our sample according to the region a firm is listed. Results separated by region are weaker for European firms, especially between Moody's and Fitch (see Table 9 in the Appendix)[17]. Yet, we have to consider the low number of observation which, as previously mentioned, might cause the weaker effect in Europe. We run into the same problem when we estimate separate regressions for firms which are rated only by two CRAs within the whole sample period (see Table 8 in the Appendix). While the effects stay significant if S&P is one of the CRAs, the estimated effects between Moody's and Fitch are smaller and partly insignificant. Besides the small number of observations, a possible explanation why we observe these insignificant results for Moody's and Fitch is a selection effect. In our descriptive analysis we show that most firms are rated by S&P and least by Fitch. Becker and Milbourn (2011) argues that rating quality deteriorates for firms where Fitch is present as a third competitor. One possible channel for this deterioration is the review frequency for these firms. If these firms are less frequently reviewed and that rating changes are less likely to trigger a reaction by another agency. On the other hand, we would assume that information is on average more ambiguous for these firm. Therefore, if we assume there is a selection effect, it casts doubts on a pure information based reasoning for herding. This might be an indication that herding is stronger for firms where either reputation or market pressure is higher or where reports are published more frequently.

4 Conclusion

The market for credit ratings is very narrow and signals (ratings) from CRAs are a crucial pillar in the financial world to get insights into firms. CRAs claim to assign ratings independently from each other, but the economic literature documents a

[17]Note, the data of the first stage prediction of $\widehat{ratg}_{-i-jt}$ is the same as in the basic model for this and the following checks. I.e. we assume that effects for specific analysts stay constant over different groups.

strong convergence of ratings between different CRAs, e.g. Güttler (2011) or Norden and Weber (2009)[18]. We contribute to this literature by providing separation of a causal herding connection between the different CRA ratings.

We are the first to show that the co-movement between CRAs is not purely based on fundamental information. Our analysis builds on a comprehensive sample of US-American and EU-European firms, rated by S&P, Moody's or Fitch between 1995 and 2016. We exploit the influence of credit rating analysts on their responsible firm ratings, especially the tendency of some analysts to be more optimistic or pessimistic in their rating decisions. We can use an analystsâ€™ influence to estimate the causal effect of non-fundamentally induced rating decisions on the rating assessment of other CRAs. Our results suggest that CRAs follow each otherâ€™s rating decisions, even though these decisions are not based on changes in firm fundamentals; consequently we call this effect herding. The herding among CRAs is sizeable and relevant for investors and regulators alike. A one notch higher rating induced by an analyst leads to ratings that are on average 0.3 to 0.5 notches higher at other CRAs.

The herding effect constitutes roughly half the size of the general co-movement between CRAs. In other words, we expect ratings of CRAs to move together when firm fundamentals change.[19].

We are aware that our approach, with a clear focus on causality, limits our ability to detect dynamics between CRAs. However, with respect to the strong literature on these dynamics, we believe our results can complement the literature by providing a way to distinguish convergence based on fundamentals to herding.

[18]They suggest that the convergence is too strong to be fully explained by fundamental data

[19]The general co-movement between CRAs is 0.8 and 0.95 notches for a one notch change at another CRA, if we don't account for firm and year-quarter fixed effects, and between 0.55 and 0.75 if we account for firm and year-quarter fixed effects.

5 Appendix

Table 4: Alphanumeric credit ratings

Conversion	Standard and Poor's	Moody's	Fitch
1	AAA	Aaa	AAA
2	AA+	Aa1	AA+
3	AA	Aa2	AA
4	AA-	Aa3	AA-
5	A+	A1	A+
6	A	A2	A
7	A-	A3	A-
8	BBB+	Baa1	BBB+
9	BBB	Baa2	BBB
10	BBB-	Baa3	BBB-
11	BB+	Ba1	BB+
12	BB	Ba2	BB
13	BB-	Ba3	BB-
14	B+	B1	B+
15	B	B2	B
16	B-	B3	B-
17	CCC+	Caa1	CCC+
18	CCC	Caa2	CCC
19	CCC-	Caa3	CCC-
20	CC,C	Ca	CC,C
21	D	C	D,DD,DDD

Table 5: Keywords to identifying methodology changes and specific reports

Group	Keywords	Rating agency
Companies	Companies	Moody's, Fitch
Merger	Merger; Joint Venture; Acquisition	Moody's, Fitch
Synthetic	CDO; RMBS; LOC-Backed; LOC-Supported; CDS; Mortgage; Collateralized; Sup bonds; Synthetic Transactions; Joint-Supported; Jointly-Supported; USPF; LOC Issues	Moody's, Fitch
Hybrid	Hybrid Securities	Moody's, Fitch
Methodology	Methodology; Methodologies; Banks; Institutions; Issuers; Corporates; Criteria; Sectors; Stats Reports; Industry Outlook	Moody's, Fitch

Table 6: Estimated herding effect: only first analyst as instrument

	Rating				
	Moody's	S&P	Fitch	S&P	S&P
$\widehat{ratg}$ Fitch	0.534^{***}	0.483^{***}			0.470^{***}
	(0.0753)	(0.0596)			(0.0706)
$\widehat{ratg}$ Moody's			0.399^{***}	0.511^{***}	0.488^{***}
			(0.0571)	(0.0414)	(0.0689)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	27626	28409	24101	85820	16943
Adjusted R^2	0.071	0.101	0.075	0.098	0.140

Bootstrapped standard errors in parentheses: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 7: Estimated herding effect with additional controls

	Rating				
	Moody's	S&P	Fitch	S&P	S&P
$\widehat{ratg}$ Fitch	0.415***	0.304***			0.275***
	(0.0651)	(0.0487)			(0.0629)
$\widehat{ratg}$ Moody's			0.365***	0.425***	0.419***
			(0.0550)	(0.0334)	(0.0635)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes
Additional FE	Yes	Yes	Yes	Yes	Yes
Observations	22572	21270	18467	65711	12210
Adjusted R^2	0.100	0.101	0.108	0.150	0.169

Firm-cluster robust standard errors in parentheses:
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.
Additional FE include industry (based on sic-code) and regional listing fixed effects.

Table 8: Estimated herding effect if firm is rated by only two agencies

	Rating			
	Moody's	S&P	Fitch	S&P
$\widehat{ratg}$ Fitch	0.263	0.357***		
	(0.166)	(0.0595)		
$\widehat{ratg}$ Moody's			0.151*	0.425***
			(0.0681)	(0.0350)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	6329	6625	4883	57631
Adjusted R^2	0.046	0.087	0.014	0.089

Bootstrapped standard errors in parentheses:
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.
Note: The first stage data is identical to the first stage used in the basic model.

Table 9: Estimated herding effect separated by region

	Rating				
	Moody's	S&P	Fitch	S&P	S&P
			US		
$\widehat{ratg}$ Fitch	0.460***	0.324***			0.271***
	(0.0762)	(0.0585)			(0.0740)
$\widehat{ratg}$ Moody's			0.433***	0.477***	0.504***
			(0.0693)	(0.0376)	(0.0798)
Observations	20154	18569	16541	64215	10880
Adjusted R^2	0.067	0.058	0.065	0.104	0.131
			Europe		
$\widehat{ratg}$ Fitch	0.301***	0.530***			0.516***
	(0.0896)	(0.0656)			(0.0994)
$\widehat{ratg}$ Moody's			0.262***	0.329***	0.286***
			(0.0537)	(0.0542)	(0.0667)
Observations	5733	7710	4311	9669	2635
Adjusted R^2	0.208	0.314	0.256	0.188	0.364
Firm FE	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes

Bootstrapped standard errors in parentheses: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

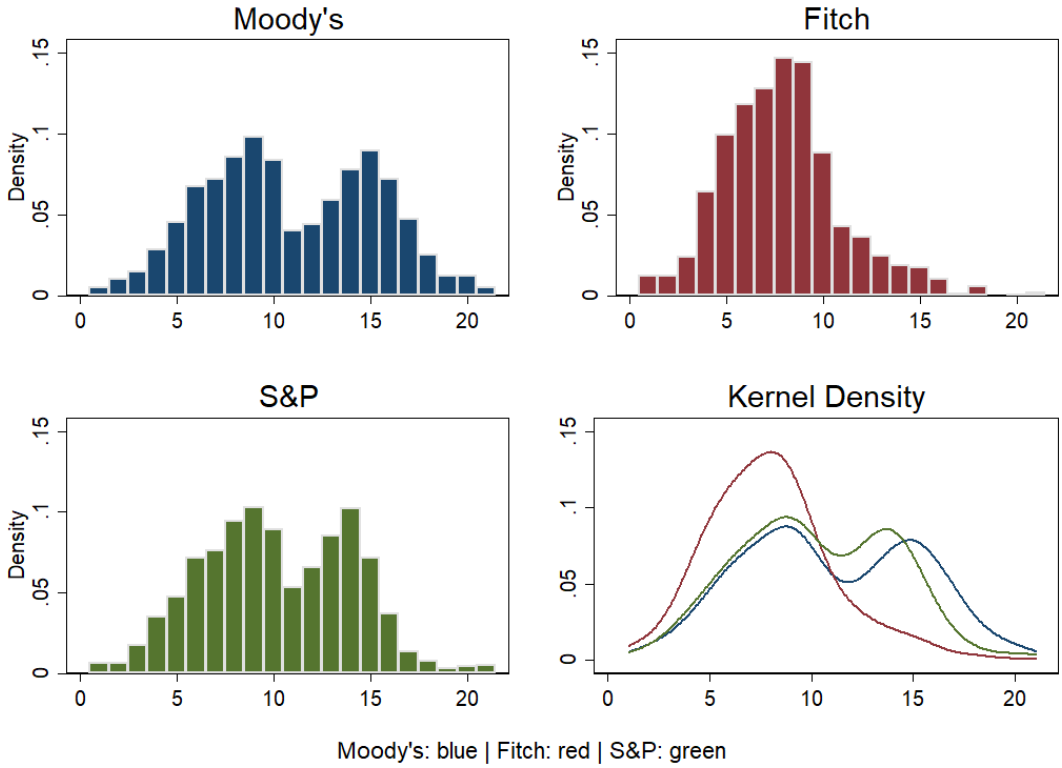

Figure 1: Rating distributions

Histogram and kernel density estimation for the ratings of the agencies. The pictures show the distributions of the firm-quarter ratings for all firms rated by the particular agency.

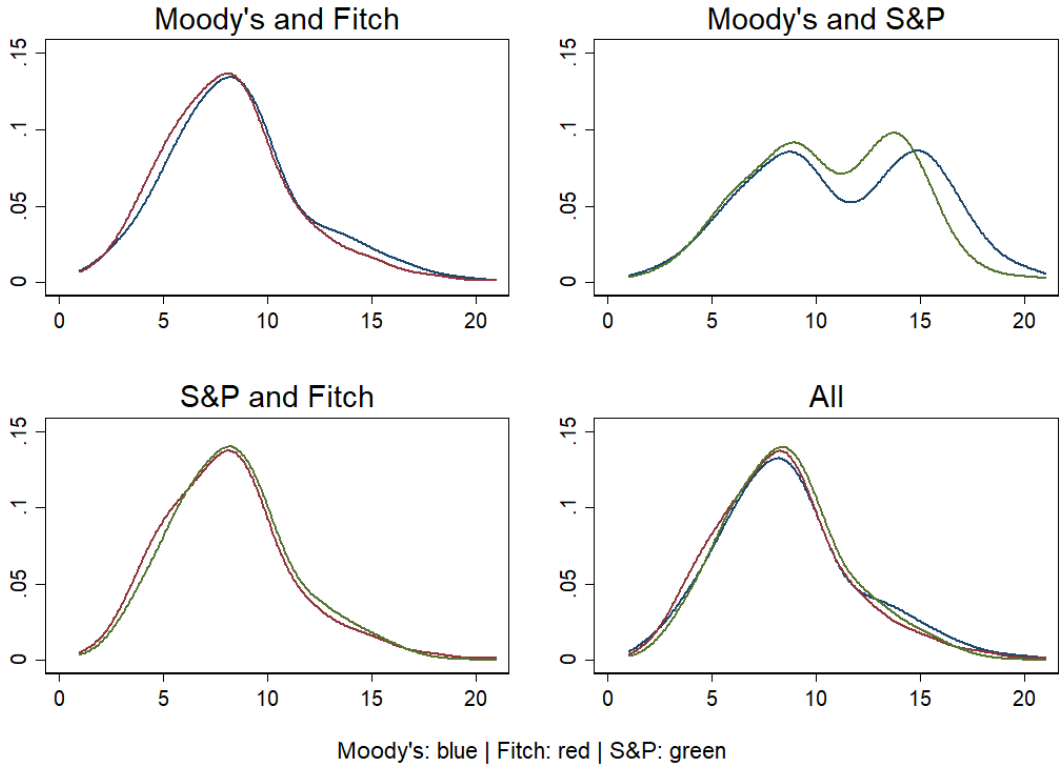

Figure 2: Rating distributions for firms rated by two/three CRAs

Kernel density estimation for the ratings of the agencies. The pictures show the distributions of the firm-quarter ratings for firms rated by both/all three agencies.

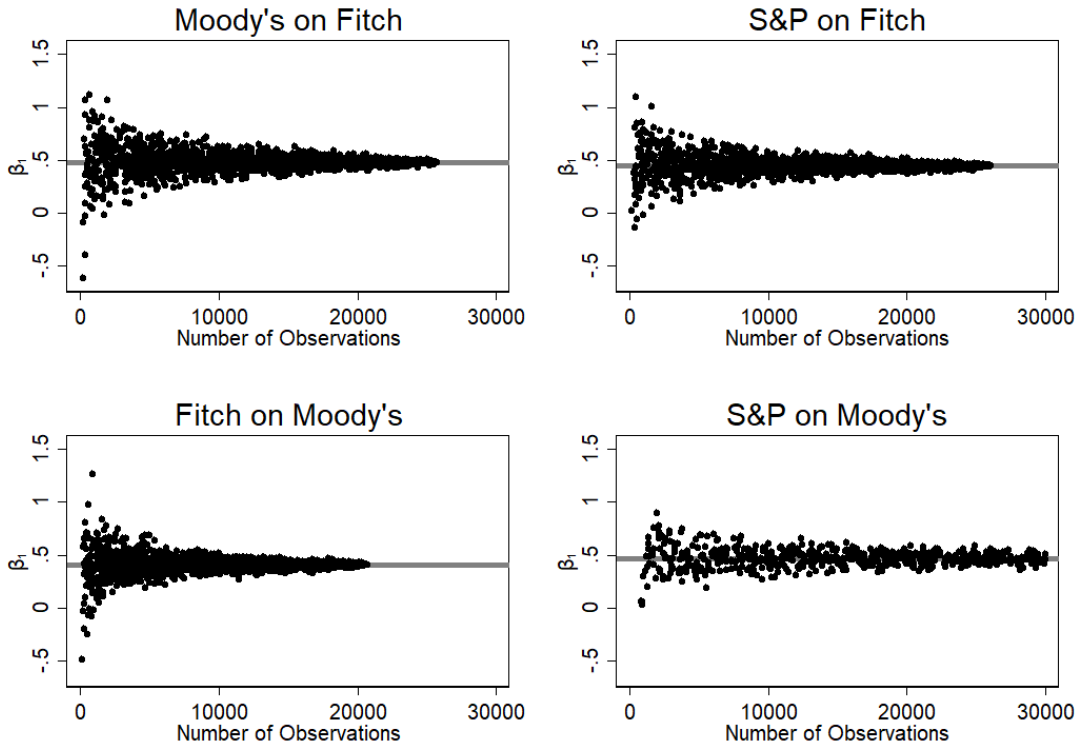

Figure 3: Herding Coefficient depending on the Number of Observations

Coefficients of the regressions depending on the number of observations in the second stage. We drew different sized subsample from our firms and estimated the second stage coefficients for the respective subsample. The horizontal line displays the coefficient from table 2.

References

Aizer, A., Doyle Jr, J.J., 2015. Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. The Quarterly Journal of Economics 130, 759–803.

Alsakka, R., Ap Gwilym, O., 2010. Leads and lags in sovereign credit ratings. Journal of Banking & Finance 34, 2614–2626.

Angrist, J.D., Imbens, G.W., Krueger, A.B., 1999. Jackknife instrumental variables estimation. Journal of Applied Econometrics 14, 57–67.

Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? The Quarterly Journal of Economics 106, 979–1014.

Becker, B., Milbourn, T., 2011. How did increased competition affect credit ratings? Journal of Financial Economics 101, 493–514.

Bertrand, M., Schoar, A., 2003. Managing with style: The effect of managers on firm policies. The Quarterly Journal of Economics 118, 1169–1208.

Blomquist, S., Dahlberg, M., 1999. Small sample properties of liml and jackknife iv estimators: experiments with weak instruments. Journal of Applied Econometrics 14, 69–88.

Cantor, R., Packer, F., 1995. The credit rating industry. Journal of fixed income 5, 10–34.

Chang, T., Schoar, A., 2013. Judge specific differences in chapter 11 and firm outcomes. Revise and Resubmit at Journal of Finance .

Devenow, A., Welch, I., 1996. Rational herding in financial economics. European Economic Review 40, 603–615.

Dobbie, W., Song, J., 2015. Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. American Economic Review 105, 1272–1311.

Dodd-Frank, 2010. Dodd-frank wall street reform and consumer protection act. One hundred and eleventh Congress of the United States .

EU, 2009. Regulation (EU) No 1060/2009 Of the European Parliament and of the Council.

Fitch, 2017. Code of conduct & ethics.

Fracassi, C., Petry, S., Tate, G., 2016. Does rating analyst subjectivity affect corporate debt pricing? Journal of Financial Economics 120, 514–538.

Frandsen, B.R., Lefgren, L.J., Leslie, E.C., 2019. Judging Judge Fixed Effects. Working Paper 25528. National Bureau of Economic Research. URL: http://www.nber.org/papers/w25528, doi:10.3386/w25528.

Güttler, A., 2011. Lead-lag relationships and rating convergence among credit rating agencies. Journal of Credit Risk 7, 95–119.

Güttler, A., Wahrenburg, M., 2007. The adjustment of credit ratings in advance of defaults. Journal of Banking & Finance 31, 751–767.

IOSCO (International Organization of Securities Commissions, 2015. Code of conduct fundamentals for credit rating agencies.

Kolesár, M., Chetty, R., Friedman, J., Glaeser, E., Imbens, G.W., 2015. Identification and inference with many invalid instruments. Journal of Business & Economic Statistics 33, 474–484.

Lugo, S., Croce, A., Faff, R., 2014. Herding behavior and rating convergence among credit rating agencies: evidence from the subprime crisis. Review of Finance 19, 1703–1731.

Mariano, B., 2012. Market power and reputational concerns in the ratings industry. Journal of banking & Finance 36, 1616–1626.

Mathis, J., McAndrews, J., Rochet, J.C., 2009. Rating the raters: are reputation concerns powerful enough to discipline rating agencies? Journal of monetary economics 56, 657–674.

Moody's, 2019. Code of professional conduct.

Norden, L., Weber, M., 2009. The co-movement of credit default swap, bond and stock markets: An empirical analysis. European financial management 15, 529–562.

Quenouille, M.H., 1956. Notes on bias in estimation. Biometrika 43, 353–360.

Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. The American Economic Review , 465–479.

Silver, D., 2016. Haste or waste? Peer pressure and the distribution of marginal returns to health care. Technical Report. Mimeo: UC Berkeley.

S&P, 2018. S&P Global Ratings Code of Conduct.

Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics 20, 518–529.