Stability and accuracy of credit ratings
Examining credit assessments from two Norwegian banks

Eric Guangcheng Hua
Department of Industrial Economics and Technology Management
Faculty of Economics and Management, NTNU
guangcheng.hua@gmail.com

Jesper Thuestad Jacobsen
Department of Industrial Economics and Technology Management
Faculty of Economics and Management, NTNU
jesper.t.jac@gmail.com

Petter Eilif de Lange
Associate Professor
Department of International Business
Faculty of Economics and Management NTNU
Petter.e.delange@ntnu.no

Lars Ole Hjelkrem
Department of International Business
Faculty of Economics and Management NTNU
lars.o.hjelkrem@ntnu.no

Abstract

This paper examines the stability and accuracy of credit ratings from two Norwegian savings and loans banks, labeled Bank A and Bank B. Credit Rating Agencies (CRAs) often claim that ratings are relative rankings of firms and largely independent of the business cycle. We find that the intensity of banks' rating changes - both upgrades and downgrades - vary over time depending on the business cycle. This is inconsistent with characterizing their methodology as through-the-cycle. We further find that the volatility and accuracy of Bank B - the bank with more customers exposed to the petroleum industry - seems to be higher than that of Bank A. The accuracy of Bank B's ratings also appears to be less affected by economic slowdowns than those of Bank A. Whereas Bank A's accuracy drops significantly following the oil price shock in 2014-2015, the accuracy of Bank B remains more stable.

1 Introduction

Credit rating agencies (CRAs) specialize in the task of evaluating the creditworthiness of an obligor, thereby helping investors and banks in assessing the riskiness of issuers and their securities (Schroeter (2013)). These agencies are confronted with a difficult trade-off dilemma when assigning credit ratings. On the one hand, they are expected by relevant stakeholders to deliver as accurate estimates of default risks as possible, at a particular point in time. On the other hand, certain stakeholders expect stable ratings that do not change in the short term to match the stakeholders' own decision-making horizons.

Originally, credit ratings were designed for long-term investors. These buy-and-hold type investors were less concerned with short-run and temporary changes in risk profiles that did not have a considerable impact on the probability of default of a company. Therefore, credit ratings were assigned "through-the-cycle" based on fundamental data. Today, the approach most CRAs use is still based on this principle, and the majority of agencies claim that their ratings are through-the-cycle and thus should be immune to short-run changes in the business cycle, as noted by Amato and Furfine (Amato & Furfine (2004)). However, some studies claim that this might not be the case, particularly for U.S. firms (J. Lobo et al. (2017)). It is, therefore, of interest to investigate these contradicting results using credit rating data from Europe. By analyzing whether there is evidence of a trade-off for higher accuracy in exchange for lower stability over time, we can examine the claim that ratings are procyclical - i.e. that short-lived economic changes, such as high or low GDP growth, affect credit ratings in a particular direction.

This study utilizes several statistical methods, some of which were developed by Paulo Carvalho, Paul Laux, and João Pereira (Carvalho et al. (2014)) for testing the characteristics of credit rating processes. We apply these methods to new ratings data. Whereas Carvalho et al. (2014) uses data sets from CRAs based in the U.S., we utilize data sets from two Norwegian savings and loan banks, referred to as Bank A and Bank B. Both data sets span the period 2009-2018.

Our literature study covers mostly solicited ratings to which rating adjustments are made when a CRA determines that a change in the creditworthiness of its rated entities has actually occurred. However, the data from the two banks are snapshots of the year-end credit ratings of their customers, irrespective of changes in their customers' creditworthiness occurring during each year. Our conclusions, therefore, differ slightly from that of previous studies and the same conclusions cannot always be drawn even if the result from a particular analysis is identical. Throughout this paper, we occasionally use the term "CRA" to refer to both traditional credit rating agencies *and* banks responsible for credit assessments.

The aim of this paper is to: (i) quantify and test the stability and accuracy of credit ratings, (ii) investigate whether the state of the business cycle influences rating adjustments, and (iii) analyze the trade-off between accuracy and stability. We examine whether the fact that the two banks have different exposure to a number of industry-specific risks, affects their credit rating methodology. At the core of our analysis is a measure for ratings volatility and instability developed by Carvalho et al. (Carvalho et al. (2014)). It condenses the information contained in a two-dimensional transition matrix into a single number for each time period and thus captures both the number of changes as well as the magnitude of rating changes.

First, we calculate unconditional transition matrices in order to offer insight into probabilities of rating changes of firms (obligors). Next, we calculate five different measures of volatility - two traditional

measures of credit rating volatility (*Large Ratings Changes* (LRC) and *Rating reversals* (RR)) and the three measures developed by Carvalho et al. (*RatVol*, *RatVolU*, and *RatVolD*).
We observe that the *trend* of the volatility of ratings for the two banks differ in a way that appears to be independent of Norwegian mainland GDP, i.e., independent of the business cycle. Finally, we assess the quality of the ratings by calculating the statistical measure: *Accuracy ratio (AR)*.

Recognizing that the state of the business cycle is likely to affect the rate of default, this paper studies the impact of the business cycle on CRAs' credit rating methodologies by performing several linear regressions. Our results from performing multivariate regressions on rating volatility and the business cycle, seem to be partially consistent with that of J. Lobo et al. (2017). For Bank A, we find evidence suggesting that credit ratings are indeed dependent on the business cycle and hence *not* through-the-cycle. Similarly, we do not find conclusive evidence that Bank B adheres to a through-the-cycle methodology.

Examining the relationship between accuracy and the state of the business cycle, we find contradicting evidence for the two banks. Whereas Bank B achieves higher accuracy in times of low GDP growth, in line with previous studies such as Bar-Isaac & Shapiro (2013), we find evidence for the opposite for Bank A. A possible explanation for the conflicting results is that banks do not face the same potential conflict of interest as CRAs, whose ratings are often paid for by the rated firms. Therefore, banks do not have the same financial incentive as CRAs to be overly optimistic in good times, as this will not generate higher income for these entities.

Cantor & Mann (2006) conclude that investors want stable credit ratings, even though this leads to trade-offs in terms of poorer rating accuracy. Therefore, we jointly analyze these two measures - rating stability and rating accuracy. Our results show that rating accuracy is positively correlated with rating stability - i.e. an inverse relationship between accuracy and volatility - for both banks, although only Bank A yielding statistically significant results. This contradicts our original hypothesis. Our conclusion of a lack of trade-off between stability and accuracy is in agreement with the conclusions of Carvalho et al. (2014).

This paper aims at applying new rating volatility measures to Nordic credit ratings. Our work contributes to existing research in several ways. We demonstrate how to determine the way different business cycle variables affect rating stability and accuracy. Furthermore, we apply our framework to analyzing credit ratings from banks, which differ from CRAs with respect to some aspects of their credit rating processes. Banks have different incentives than CRAs and we thus contribute with new results not seen in previous credit rating studies. Lastly, we implement multivariate regression methods, some of which to our knowledge, are previously not employed in credit rating research.

2 Literature Review

Through conversations with investors, issuers and regulators, Cantor & Mann (2006) find that many market participants have a preference for stable and accurate credit ratings. Therefore, credit ratings are expected not to be point-in-time measures of credit risk, but to be through-the-cycle reflecting credit risk over the long term. Consequently, one would expect credit ratings not to be significantly correlated with the business cycle. It is also reasonable to expect that frequent credit rating adjustments would result in more accurate ratings. However, several studies claim to have found evidence contradicting both of these claims.

The financial system is mostly procyclical. Measures of financial activity - e.g. new bond issuance, bank lending, and equity offering - tend to be more prevalent during booms than downturns (Bernanke et al. (1999)). At the same time, CRAs claim only to adjust credit ratings when permanent changes in the risk profiles of companies occur. Several studies investigate the causes of credit rating changes and whether CRAs have a motive for frequent changes. Carvalho et al. (2014) analyze the motivations for CRAs to modify their ratings by examining the stability and accuracy of credit ratings. They conclude that CRAs have more volatile ratings during bad economic times, which is inconsistent with the claim that ratings are simply a relative measure of obligors' riskiness and thus should be independent of the business cycle. This implies that credit ratings are point-in-time measures rather than through-the-cycle, consistent with the findings of Amato & Furfine (2004), but at least partially contradicting the work of Altman & Rijken (2006). Amato and Furfine find that cyclical changes to individual businesses and financial risk attributes play a significant role for rating changes, contrary to a through-the-cycle methodology. Furthermore, they find little evidence of procyclicality in U.S. firms en masse. By contrast, they find evidence for procyclicality in *initial ratings* and in *rating changes*. They reason that CRAs rarely change the rating of a particular firm and generally do not adjust ratings based on small movements in the risk profile of firms. However, when they *do* adjust the ratings, they tend to overreact by being excessively optimistic in booms and pessimistic in downturns. While Amato and Furfine use ratings from 1984 to 2000, J. Lobo et al. (2017) use a larger data set with ratings from 1984 to 2012. They do find procyclical tendencies in credit ratings, particularly in the latest period. Lobo et al. attribute their diverging results to differences in the data sets. They argue that Amato and Furfine look at credit ratings during a period with lower economic and market fluctuations and that evidence for procyclicality is only found when including the additional 12 years of credit ratings. As noted by Cantor & Mann (2003), if the evidence provided by Carvalho et al. and Amato and Furfine is correct, the riskiness of issuers and bonds today cannot be directly compared to the riskiness of issuers and debt instruments, which have been rated similarly in the past.

The variation in the accuracy of credit ratings over time suggests a dependency on the business cycle. In an influential theoretical paper, Bar-Isaac & Shapiro (2013) find that credit rating quality is countercyclical, i.e., moving in the opposite direction to that of the overall state of the economy. They reason that CRAs have incentives to improve their reputation, i.e., accuracy, in bad times when analyst labor is cheap and rating mistakes are less likely to be noticed, in order to increase their income in better times when labor is scarcer and fewer firms default. The same conclusion is reached by Bolton et al. (2012). They suggest that, due to the conflict of interest for CRAs, they have a tendency to understate risk in order to attract new business during economic booms, leading to a rating bias. This, in addition to potentially deteriorating due diligence in such periods, is a possible reason for decreased ratings accuracy during booms. When examining the relationship between credit ratings, the business cycle and the raising of capital, Isil et al. (2012) find evidence that appears to substantiate the conclusions of Bolton et al. Their results suggest that a borrower's credit quality is a significant factor in its ability to raise capital during macroeconomic downturns. Specifically, they find that sub-investment grade firms appear to be shut out of the public capital markets during poor economic conditions. As a result, it is likely that companies perform ratings shopping, choosing the CRA that gives them the most favorable credit rating. CRAs thus have an incentive to assign too high ratings during good times in order to attract new business, thus reducing their rating accuracy.

Carvalho et al. (2014) also find indications of this phenomenon. They conclude that higher rating volatility, i.e. more frequent changes, does not lead to higher ratings accuracy. Instead, their results suggest that CRAs modify ratings not to achieve higher accuracy, but to increase revenue. The reasoning behind this claim is the observation of more intense rating adjustments shortly before new issuance in the primary bond market for seemingly no apparent reason. In their view, more frequent changes may lead to a more favorable view of a CRA among investors, which in turn may cause new issuers to choose this CRA when purchasing a credit rating.

Although credit ratings primarily are relative risk measures distinguishing the credit risk of a company from peers in other rating categories, they can also be used to estimate the probability of default. By analyzing the frequency of rating changes from a given rating to another, an estimate of the risk associated with different ratings can be obtained. The distribution of such rating changes plays a crucial role in many risk models. By generating both unconditional and conditional transition matrices, Nickell et al. (2000) quantify how rating transition probabilities depend on the industry that the obligors operate in and the state of the business cycle. They find significant differences between the transition probabilities of banks and industrials, and in good and bad economic times - referred to as peaks and troughs. The latter result implies that credit ratings are dependent on the business cycle and thus not a through-the-cycle measure of risk, consistent with the findings of Carvalho et al. (2014), Amato and Furfine (2004) and Lobo et al. (2017).

3 Data

As noted above, this paper examines the stability and accuracy of credit ratings and attempts to determine how the state of the business cycle influences the frequency and intensity of rating adjustments. In order to perform such an analysis, two types of data are necessary: (1) historical time series data for the credit ratings and (2) several proxies for the business cycle that can be used for measuring the effect that the business cycle has on credit ratings. We also look at financial indicators, which are often thought to be forward looking indicators of the real economy and the state of the business cycle. This gives us an additional comparative measure for interpreting our results. We also consider the different economic variables' exposure to the petroleum industry.

3.1 Credit Ratings

The methods presented in the following section are implemented on two data sets from one medium sized and one large Norwegian savings and loans banks. In this paper, we will refer to them as Bank A and Bank B.

3.1.1 Bank A

Our first data set is provided by a Norwegian savings and loans bank that we refer to as Bank A. The data set includes the bank's own estimated probabilities of default and the corresponding letter ratings for

8,724 Norwegian companies. All rated companies are the bank's own customers - i.e. companies that have loans at the bank. The first ratings were assigned in December 2009 and the final ratings were assigned in late December 2018 and early January 2019, with yearly adjustments for as long as the firms remained solvent and remained a customer. Letter ratings range from *A* to *K*, where *A* denotes the lowest probability of default and *J* and *K* denote companies currently in default. The outstanding debt of companies with rating *K* has been registered as written off, while the debt of companies with rating *J* has not been written-off. We also assume that previously rated firms that are not rated in a particular year, have decided to retire their status as customers of the bank.

Except for a slight decrease in the number of ratings from 2010 to 2011 in the aftermath of the financial crisis, the total number of ratings has increased every year in the period.

Table 1 shows the default frequency and default rates per year for Bank A. Before calculating the statistics, we adjusted the data by registering companies that remained in one of the default states - i.e. are assigned rating *J* or *K* - for two or more consecutive years as defaulting the *first* time they were assigned the rating. However, companies that leave the bankruptcy state by being assigned a higher rating and then once more enter bankruptcy will receive a second bankruptcy count. In other words, we don't remove bankrupt companies from the data set altogether, but we adjust the data set to take such occurrences into consideration.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Number of new ratings J	40	34	24	19	32	28	22	19	21	31
Number of new ratings K	39	16	15	15	20	23	13	24	14	14
Total new defaults	79	50	39	34	52	51	35	43	35	45
Default rate (J) [%]	1.16	0.95	0.70	0.54	0.89	0.76	0.57	0.47	0.51	0.73
Default rate (K) [%]	1.13	0.45	0.44	0.43	0.56	0.63	0.34	0.60	0.34	0.33
Total (J & K) [%]	2.29	1.40	1.14	0.97	1.45	1.39	0.91	1.07	0.85	1.06

Table 1: Annual data from Bank A for (1) the number of new default ratings and (2) the default rate as a percentage of total ratings that year.

3.1.2 Bank B

The second data set is provided by another Norwegian savings and loans bank, hereafter referred to as Bank B. Its customers are located along the west coast of Norway. A larger part of households work within petroleum related industries. The bank is therefore, more invested in the petroleum sector and exposed to the oil price, compared to its counterpart. Their credit ratings are updated at a monthly frequency. However, the ratings used in this paper were annualized to maintain anonymity. The data set includes the bank's own estimated probabilities of default and their corresponding letter ratings for 5,615 Norwegian companies. All rated companies are the bank's own customers - i.e. companies that have loans at the bank. The first ratings were assigned in 2009 and the final ratings were assigned in 2018, with yearly adjustments for as long as the firms remained solvent and remained a customer. Letter ratings range from *A* to *N*, where *A* denotes the lowest probability of default and *M* and *N* denote companies currently in default. The outstanding debt of companies with rating *N* has been registered as written off, while the debt of companies with rating *M* has not been written-off. We also assume that previously rated firms that are not rated in a particular year, have decided to retire their status as customers of the bank.

Most companies are assigned ratings *B, C, D, E,* and *F*. Relatively few companies are assigned the highest and lowest ratings, i.e. rating *A*, *J*, and *K*. The total number of ratings has steadily decreased over time, from 2,870 in 2009 to 2,243 in 2018. This decrease is due to customers suspending their relationship with the bank after the exogenous oil price shock in 2014-2015. It is thus likely that some of these companies actually defaulted. However, since they discontinued being customers at the bank, these default incidences do not show up in the default rating categories. The percentage distribution among different ratings classes remains approximately the same for the higher rating classes *A* to *F* at the beginning of the period as compared to the end. There is a slight increase in companies of higher rating classes in the years 2010 to 2014 - years with extraordinary high oil prices. The share of firms distributed to the lower rating classes *G* to *N*, however, decreases steadily, making up about 38% of all ratings in 2008 and 21% in 2018.

Table 2 shows the default frequency and default rates per year for Bank B. We repeat the same procedure for Bank B as we did for Bank A. We adjust the data by registering companies that remained in one of the default states - i.e. rating classes *M* or *N* - for two or more consecutive years as defaulting the *first* time they were assigned the rating. Furthermore, we let companies that leave the bankruptcy state by being assigned a higher rating and then once more defaults, receive a second bankruptcy count. Again, we note that the true default rates are probably higher in the years following 2014 than table 2 indicates. Due to defaulting firms discontinuing their customer relationship with the bank before being assigned a default rating, this does not appear in the data from the bank - the companies are simply removed from the bank's customer list.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Number of new ratings M	45	36	16	17	18	11	7	11	10	12
Number of new ratings N	93	32	24	22	12	12	11	8	7	6
Total new defaults	138	68	40	39	30	23	18	19	17	18
Default rate (M) [%]	1.62	1.28	0.57	0.62	0.67	0.43	0.31	0.50	0.45	0.54
Default rate (N) [%]	3.35	1.14	0.85	0.80	0.45	0.47	0.49	0.37	0.32	0.27
Total (M & N) [%]	4.97	2.42	1.42	1.42	1.12	0.90	0.80	0.87	0.77	0.81

Table 2: Annual data from a Bank B for (1) the number of new default ratings and (2) the default rate as a percentage of total ratings that year.

3.2 Measures of the Business Cycle

Financial markets are usually a leading indicator of the business cycle. The reasoning behind this is simple. The overall pattern of the current financial market is well-known to all investors. The future, however, is for obvious reasons uncertain. Consequently, investors attempt to act just before the business cycle turns; they increase their risky positions when they *believe* that the business cycle is at the end of a trough, and they reduce risk prior to *believing* that the business cycle is at a peak, as noted by Calverley (2002). Real GDP is a macroeconomic measure of economic output - a good indicator for the state of the economy. The change in real GDP is lower (or negative) during troughs and higher during peaks. Empirical studies have shown that financial variables can be leading indicators of recessions (see, e.g., Estrella & Mishkin (1998)). Such variables include, but are not limited to, the yield curve spread and the swap rate. The CBOE Volatility Index (VIX) can also be seen as an indicator for the state of the economy. Since this paper employs credit ratings from banks and not traditional CRAs, we also include the change in the monetary value of new bond issuance and loans as a potential measure of the business cycle. Applying the above reasoning, we wish to examine how the banks (in their capacity as CRAs) are affected by financial market cycles.

3.2.1 Real GDP

As a proxy for the state of the business cycle, many studies use real GDP, as noted by Wong et al. (2016) and Carvalho et al. (2014). Real GDP provides a relatively good measure because it contains data covering a broad range of economic activity, thereby reflecting the real economic situation in a country. With the purpose of investigating the effect that the state of the business cycle has on credit ratings, we collect data from Statistics Norway (SSB (2020*b*)). However, GDP data consists of two separate components: 1) a long-run trend component and 2) a business cycle component. As our analysis focuses on the state of the business cycle, we are more interested in the business cycle component. Hence, we isolate this component using a Hodrick-Prescott (HP) filter as described in appendix B, thereby removing the long-run trend component.

3.2.2 Swap rates

As suggested above, the yield curve slope is another potential indication for the condition of the business cycle. In normal times with inflation, the yield curve is positive. This indicates a positive expectation of financial performance in the future and thus increased risk premiums for long-term investments. If the slope is negative - i.e. an inverted yield curve - this could indicate an impending recession. The same rationale applies to the swap rate curve. In other words, prior to recessions, long-term rates can become lower than short-term rates. To model this relationship between long-term and short-term rates, similar studies employing credit data from the United States have considered the difference between 10-year and 2-year U.S Treasury bond yields and analyzed its effect on credit rating adjustments. Our data, however, is collected mostly from Norway. The low demand for Norwegian treasuries leads to poorer liquidity in these securities. Therefore, for our purposes, Norwegian government bond yields are probably not a good proxy for the ”true” yield curve as they do not reflect the true risk of government debt. Instead, we utilize the difference between swap rates of the same maturities - 10-year maturity minus 2-year maturity, which we collect from Macrobond. Next, in order to match the frequency of our credit rating data, we proceed to annualize the swap rate data and calculate the difference between 10-year maturity and 2-year maturity swap rates. There is a decreasing trend of swap rate differences, the curve is flattening and actually inverting as 2020 approaches. This could reflect declining policy rates set by the central bank of Norway during the energy crisis of 2016.

3.2.3 Volatility Index

As a third indicator for the state of the economy, we need a proxy for economic uncertainty. In times of high or low levels of uncertainty, banks and CRAs could potentially decide to make changes in credit ratings. A possible such proxy could be historical volatility for the Oslo Stock Exchange Index. However, this risk measure does not reflect future expectations. Since it is expectations that are indicative of uncertainty - and not necessarily real volatility, and since historical volatility will not be available to CRAs when assessing credit risks at a particular point in time, the Volatility Index (VIX) is a better proxy than historical volatility. The Volatility Index (VIX), created by the Chicago Board Options Exchange (CBOE), represents the market's expectation of the 30-day forward-looking volatility by calculating the implied volatility based on S&P 500 index options. To our knowledge, there are no good Nordic or European alternatives to the VIX index. Since previous studies have found a significant correlation between Norwegian and US stock indices (Kruge & Tysnes (2011)), we conclude that the VIX can be used as a proxy for forward-looking volatility in the Norwegian equity market. We obtain daily VIX data from Macrobond. To match the frequency of our rating data, we proceed to annualize the VIX data by calculating the average yearly VIX. The annualized VIX follows a clear downward trend, falling steadily since the global financial crisis.

3.2.4 New Bond Issuance and New Loans

New corporate debt could potentially be related to the propensity of credit rating adjustments and their accuracy (see, e.g., Carvalho et al. (2014)). If the interests of bond issuers and CRAs align when corporations want to take on new debt, CRAs should perform more upgrades and less downgrades on average leading up to periods of increased new debt financing. The reasoning behind this is that a positive rating change by a CRA will lead to better terms when raising capital for corporations. Banks assessing credit risks and providing loans to the respective clients, will however have an incentive to reflect the *true* credit risk. We want to examine how the appetite for new corporate debt relates to the rating stability and accuracy of credit ratings. Generally speaking, corporations have two means of debt financing; they can issue corporate bonds, or they can apply for bank loans. Corporate bonds are perceived as more risky, but are more liquid and often offer higher yields than bank loans.

We collect data from two different sources to use as a proxy for investors' willingness to take on new debt. Then, we use these proxies to assess whether obtaining debt capital influences rating volatility and accuracy.

The first proxy is the total annual value of new bond issuance in Norway. The total market value of new bond issuance in Norway from 2009 to 2018 is collected from Nordic Trustee's Norwegian Bond Market Report (Trustee (2018)). Nordic Trustee is the leading provider of trustee and agency services for bonds and direct lending in the Nordic region.

The second data set is the balance sheet of all Norwegian banks from 2009 to 2018 collected from the Norway Statistics Bureau (SSB (2020*a*)). Due to missing data from the first quarter of 2009, we will be using the average monthly change in the balance sheet of "loans to and claims of customers" - both companies and private individuals - for each year. The point-in-time measure of loans to customers is not a perfect measure for gauging the value of new loans to banks' clients. However, the net change in the banks' assets will reflect the lending behavior of banks to an acceptable degree. The proxy for new debt in terms of bank loans will be the average monthly net change in asset values of bank loans for each year.

4 Methodology

We wish to analyze ratings stability and ratings quality or accuracy using data sets containing ratings data from two Norwegian savings and loans banks. In order to do so, we employ several statistical methods. First, we construct unconditional transition matrices. Then, we employ a measure developed by Carvalho et al. (2014) that condenses the information contained in two-dimensional transition matrices into a single scalar representing the volatility of ratings for each time period. We further analyze ratings stability by calculating two different alternative measures: 1) *Large Rating Changes* (*LRC*) and 2) *Rating Reversals* (*RR*). Explanations of these alternative measurements are addressed in the Appendix. This is followed by an analysis of the accuracy of ratings, evaluated using the measure *Accuracy Ratio* (*AR*), representing rating quality. In order to understand the effect that the state of the business cycle has on the volatility and quality of ratings, we employ several linear regressions. We examine the relationship between the state of the business cycle and the accuracy ratio and ratings volatility. In addition, we examine the relationship between accuracy ratio and rating volatility, with linear regressions.

4.1 Transition Matrix

Transition matrices provide an approximation of the probability of a transition from one rating class to another in the course of a predefined time period. Given the nature of our data sets and the frequency of revisions of credit ratings, it is more reasonable for us to look at a discrete Markov chain model. The probability estimates are calculated by first collecting historical credit rating changes over a given time period. Next, the frequency with which obligors move from the initial rating *i* to the next rating *j* is collected, denoted N_{ij}. Finally, this measure is transformed into a transition probability by dividing the frequency by the total number of firms in a given rating category *i* in the beginning of a period, denoted N_i . The formula for calculating the probability of migration from a specific rating class to another specific rating class during a single period thus is

$$\hat{p}_{ij} = \frac{N_{ij}}{N_i} \quad \forall\, i, j \tag{1}$$

By repeating this calculation for all rating migration possibilities and for all time periods, the result is a transition matrix containing the *average* empirical probability of rating transitions for a given sample.

4.2 Measure of Rating Volatility

Carvalho et al. (2014) construct a new measure for the stability of credit ratings denoted *Ratings Volatility* (*RatVol*). It is an estimate of the volatility of credit ratings and is very similar to a standard deviation. The measure condenses all information that is contained in a two-dimensional ratings transition matrix, into a single scalar for each time period. Hence, the measure can be utilized in time-series tests. It is based on the same information required to compute a standard ratings transition matrix.

4.3 Definition of RatVol

To calculate *RatVol*, let *t* denote the time in years so that *t* = 1, 2, …, *T* represents the end of each year. Let the weights for all possible rating transitions from rating *s* to rating *f* from period *t* − 1 to period *t* equal

$$w_t(s,f) := \frac{n_t(s,f)}{\sum_{s=1}^{N}\sum_{f=1}^{N} n_t(s,f)} \tag{2}$$

where $n_t(s,f)$ is the number of firms that ended the last year (time *t*−1) with rating s and ended the current year (time *t*) with rating *f*. In this context, *s* and *f* are assigned a numerical value, e.g. 1 for ratings *A* and 11 for ratings *K*. *N* denotes the total number of possible ratings classifications in the data set. The term in the denominator represents the total number of movements in the transition matrix. The ratings volatility is then defined as

$$RatVol_t := \sqrt{\sum_{s=1}^{N}\sum_{f=1}^{N} w_t(s,f) \times (f-s)^2} \quad \forall \ \ t \in T \tag{3}$$

Unlike normal transition matrices, the measure above gives more weight to transition paths with more observations and less weight to paths with fewer observations. It also penalizes larger ratings movements more than smaller ratings adjustments. Therefore, Carvalho et al. (2014) argue that it more correctly depicts the true volatility of ratings compared to transition matrices.

4.3.1 Decomposition of RatVol into Upgrades and Downgrades

RatVol includes the total volatility effects of both upgrades and downgrades. In order to analyze these effects separately, we split them into the volatility due to upgrades, $RatVolU_t$, and the volatility due to downgrades, $RatVolD_t$. These measures are defined as

$$RatVolU_t := \sqrt{\sum_{s=1}^{N}\sum_{f=1}^{N} w_t(s,f) \times (f-s)^2 I_{\{f<s\}}} \quad \forall \ \ t \in T \tag{4}$$

$$RatVolD_t := \sqrt{\sum_{s=1}^{N}\sum_{f=1}^{N} w_t(s,f) \times (f-s)^2 I_{\{f>s\}}} \quad \forall \ \ t \in T \tag{5}$$

where the indicator function $I_{\{f<s\}}$ is equal to 1 when the final rating *f* is lower than the initial rating *s*, i.e. when an upgrade occurs. Conversely, $I_{\{f>s\}}$ is equal to 1 when a downgrade occurs.

4.4 Accuracy of Ratings

The Cumulative Accuracy Profile (CAP) can be used to evaluate the accuracy of a rating system through time. A CAP curve is constructed by plotting the share of defaulters for each of the rating categories, as seen in Figure 1. This curve can then be used to derive a measure known as the *Accuracy Ratio* (*AR*) (Cantor & Mann (2003)).

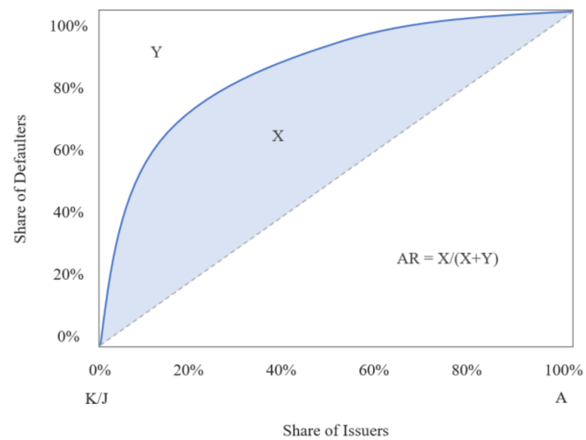

Figure 1: Example of a Cumulative Accuracy Profile (CAP) plot and the derivation of the Accuracy Ratio (AR). Note that the percentages on the y-axis are not necessarily equivalent to the rating stated below due to the ordinal data structure of credit ratings.

The Accuracy ratio (AR) is the summary index of the Cumulative Accuracy Profile (CAP). It condenses all the information depicted in the CAP curve into a single statistic and is equal to the area between the CAP curve and the 45-degree line, divided by the total area above the 45-degree line - i.e. $AR = X/(X+Y)$ as shown in Figure 1. The measure is commonly used to compare the relative accuracy and quality of credit ratings by measuring the correlation coefficients between rating categories and defaults. Furthermore, it is a measure of the discriminatory power of credit score models - i.e. the ability to distinguish *ex ante* between defaulting and solvent firms. Thus, the AR measures how accurate a credit model predicts the probability of default. If only firms in the lowest rating category default (*J* and *K* in our data set), the AR will approach 1. If firms in all rating categories default - i.e. defaults are unrelated to rating categories - the AR will be 0. If only firms in the highest rating category default (A in our data set), the AR will approach −1.

Building on the work of Cantor & Mann (2003) and Carvalho et al. (2014), the AR at time t can be computed using the equation

$$AR_t = \frac{\sum_{\underline{r}}^{\overline{r}}[n(i)-n(i-1)][d(i)-n(i)+d(i-1)-n(i-1)]}{1-\frac{D}{N}} \quad (6)$$

Where

$\underline{r}$ = minimum r in the sample at time t.

$\bar{r}$ = maximum r in the sample at time t.

$n(r) = \sum_{\underline{r}}^{r} N_i/N$, for $r \geq 1$, and $n(0) = 0$

$d(r) = \sum_{\underline{r}}^{r} D_i/D$, for $r \geq 1$, and $d(0) = 0$

N_r = number of issuers with rating r at time t.

D_r = number of issuers with rating r at time t that defaulted over the following year.

N = total number of issuers at time t.

D = total number of defaults over the following year.

4.5 Adjusting the Business Cycle Variable

As previously mentioned, a common proxy for the state of the business cycle is the quarterly or annual change in real GDP. However, GDP data consists of two separate components: 1) a long-run trend component and 2) a short-run business cycle component. As our analysis focuses on the state of the business cycle, we are only interested in the business cycle component. In order to isolate this component from the trend component, we apply a statistical technique developed by Hodrick & Prescott (1997) known as the Hodrick-Prescott (HP) filter. This method is not the focus of our research, and a more detailed explanation is consequently addressed in the Appendix.

4.6 Analysis of Volatility of Ratings

It is natural that the credit ratings of companies change to some extent over a period of time. With data sets containing ratings for several thousand companies, it is of interest to investigate whether there is a fundamental explanatory reason underlying rating upgrades and downgrades. In this paper, we ask if the state of the business cycle, represented by several macroeconomic variables, might be a cause of ratings changes.

We investigate business cycle effects by performing multivariate regressions. To represent the business cycle, we employ $GDP.Dev_t$, $SwapRate_t$, VIX_t, $NewBonds_t$, $NewLoans_t$, and $RateDef_t$ as explanatory variables. The following multivariate regressions are run for each data set:

$$RatVol_t = \alpha + \beta_1 GDP.Dev_t + \beta_2 SwapRate_t + \beta_3 VIX_t + \beta_4 NewBonds_t + \beta_5 NewLoans_t + \beta_6 RateDef_t + \varepsilon_t \quad (7)$$

$$RatVolU_t = \alpha + \beta_1 GDP.Dev_t + \beta_2 SwapRate_t + \beta_3 VIX_t + \beta_4 NewBonds_t + \beta_5 NewLoans_t + \beta_6 RateDef_t + \varepsilon_t \quad (8)$$

$$RatVolD_t = \alpha + \beta_1 GDP.Dev_t + \beta_2 SwapRate_t + \beta_3 VIX_t + \beta_4 NewBonds_t + \beta_5 NewLoans_t + \beta_6 RateDef_t + \varepsilon_t \quad (9)$$

Higher values of *GDP.Dev*$_t$, *SwapRate*, *NewBonds*, *NewLoans* are indicative of better economic times. By contrast, better times are associated with lower values of *VIX* and *RateDef*. As a result, a positive correlation between the business cycle measures *GDP.Dev*$_t$, *SwapRate*, *NewBonds*, *NewLoans* (a negative correlation between the business cycle measures *VIX* and *RateDef*) and the volatility measures *RatVol*$_t$, *RatVolU*$_t$, and *RatVolD*$_t$ indicate that better economic times are associated with higher total volatility, more frequent upgrades, and more frequent downgrades, respectively. Likewise, a positive relation between *GDP.Dev*$_t$, *SwapRate*, *NewBonds*, *NewLoans* (a negative relation between *VIX* and *RateDef*) and *LRC*$_t$ and *RR*$_t$, corresponds to a higher frequency of rating adjustments of three or more notches and a higher frequency of rating reversals, respectively.

4.7 Analysis of Quality of Ratings

CRAs' credit ratings are not perfect assessments of default probabilities. Occasionally, even companies with high credit ratings default. As previously described, the discriminatory measure known as the Accuracy Ratio (AR) can measure how well a CRA performs at assigning "correct" ratings to companies that actually do default. In other words, the AR is a measure of the quality of a CRA's credit ratings. By analyzing the relation between AR and the state of the business cycle, it is possible to determine whether the business cycle has an effect on rating quality.

In order to gain deeper insight into which macroeconomic variables are associated with higher quality of ratings, we employ the following explanatory variables in our multivariate regression: *GDP.Dev*$_t$, *SwapRate*$_t$, *VIX*$_t$, *NewBonds*$_t$, *NewLoans*$_t$, and *RateDef*$_t$. We also include *RatVol*$_t$ to test the effect of our measure of ratings' volatility. Hence, the following multivariate regression is run for each individual bank:

$$AR_t = \alpha + \beta_1 GDP.Dev_t + \beta_2 SwapRate_t + \beta_3 VIX_t + \beta_4 NewBonds_t + \beta_5 NewLoans_t + \beta_6 RateDef_t + \gamma_1 RatVol_t + \varepsilon_t \quad (10)$$

SwapRate$_t$ is the difference between swap rates of 10-year maturity and 2-year maturity. *NewBonds*$_t$ is the total annual value of new bond issuance in Norway. *NewLoans*$_t$ is the average monthly change in loans to Norwegian bank customers. *VIX*$_t$ is the average annual market expectation of the 30-day forward-looking volatility of the S&P 500 index. *RateDef* is the annual default rate among each bank's customers. A positive coefficient for *GDP.Dev*$_t$, *SwapRate*$_t$, *NewBonds*$_t$, and *NewLoans*$_t$ indicates that AR is positively related with business cycle peaks and inversely related with business cycle troughs. In other words, the quality of ratings is higher when the economy is performing well and lower when it is performing poorly. Conversely, a positive coefficient for *VIX*$_t$ and *RateDef* indicates that the rating accuracy is higher during economic troughs and lower during peaks.

4.8 Analysis of the Relation between Accuracy and Stability

Cantor & Mann (2006) claim that CRAs trade off accuracy for higher stability. If this is the case, the relationship between the accuracy ratio (AR) and rating volatility (*RatVol*), rating reversals (RR) and Large Rating Changes (LRC), should be positive. The reason for an expected positive relation is because these three measures - *RatVol*, RR, and LRC - in different ways express the degree of the volatility of ratings. Therefore, if CRAs do indeed trade off accuracy in order to attain more stable changes, a lower accuracy should be accompanied by higher stability, i.e. less volatility, or vice versa.

4.8.1 Relationship between AR and RatVol, LRC, and RR

To examine the relationship between AR and *RatVol*, LRC, and RR, we run the following simple linear OLS regressions, with AR_t as the dependent variable and $RatVol_t$, LRC_t, and RR_t as the independent variables:

$$AR_t = \alpha + \beta RatVol_t + \epsilon_t \tag{11}$$

$$AR_t = \alpha + \beta LRC_t + \epsilon_t \tag{12}$$

$$AR_t = \alpha + \beta RR_t + \epsilon_t \tag{13}$$

A negative β for any of the regressions indicates a negative correlation between the two variables. Such a relation *could* imply a trade-off between accuracy and stability, meaning that CRAs intentionally accept less accurate ratings in order to achieve more stable ratings. However, such a correlation by itself, even if statistically significant, is not enough to conclude that CRAs *actively* pursue such a trade-off. It would, however, strengthen the argument that the particular CRA's ratings are through-the-cycle rather than point-in-time measures, as noted by Altman & Rijken (2006).

4.9 Multivariate Regression Methods

When explanatory variables in a linear regression model exhibit tendencies of multicollinearity, the standard errors of the estimated coefficients can be falsely inflated resulting in inaccurate, non-significant p-values when using the Ordinary Least Squares (OLS) method. Therefore, the potential problem of multicollinearity has to be solved. Correlation can be identified by analyzing the correlation matrix and scatter plots for the independent variables in question, as well as using the variance inflation factor (VIF). Variables with high values of VIF can be removed from the OLS regression. A more detailed explanation of the methodology to test for multicollinearity is found in the Appendix.

5 Results and Discussion

In this section, we present the results from implementing the methods outlined in section 4 on the two data sets presented in section 3. First, we present the unconditional transition matrices summarizing the rating migration probabilities. Next, we demonstrate the stability of the credit ratings as represented by rating volatility (*RatVol*, *RatVolU*, and *RatVolD*), Large Rating Changes (LRC), and Rating Reversals (RR).

Next, we evaluate the quality of the ratings as measured by the accuracy ratio (AR). In an attempt to relate credit ratings to the business cycle, we regress our three volatility measures against several macroeconomic variables. We also regress AR against rating volatility and macroeconomic variables. Finally, we sketch some possibilities for future extension of our work.

5.1 Transition Matrices

A transition matrix provides an approximation for the probability of a transition from one rating category to another. The unconditional transition matrices are presented in Table 3 for Bank A and Table 4 for Bank B. We also calculate the standard deviations of the transition rates for the banks, shown in parentheses below the probabilities.

The diagonal probabilities in the matrices can be interpreted as the probability of retaining a particular rating for two consecutive years.

5.1.1 Unconditional Transition Matrices

In the unconditional transition matrix for Bank A shown in Table 3, the highest probabilities are located along the diagonal. There is one exception for the row representing rating migrations originating from rating category *D*, where the probability of moving from *D* to *C* is higher than that of remaining in *D*. We observe that for companies with ratings *A, B, C, I, J,* and *K*, it is more likely to remain in their current rating category than to migrate to another rating category - i.e. the probability is higher than 50%. These categories represent the upper and lower range of rating categories. We, therefore, conclude that it is more likely for companies with high ratings to remain in the upper range, and companies with low ratings to remain in the lower range.

Meanwhile, companies with ratings in the center of the rating spectrum - ratings *D, E, F, G,* and *H* - are more likely to be upgraded or downgraded than to remain in the same state. From a purely probabilistic standpoint, this conforms with a higher possible range of rating choices for CRAs to select from when assigning new ratings. In the center of the matrix, CRAs can assign both upgrades and downgrades to all companies. At the edges, however, there are fewer possible assignment choices in the immediate vicinity. Also, at the edges the CRAs' rating options are mostly tied to one direction: For high ratings, most possibilities for rating adjustments are downgrades, whereas for low ratings most possibilities for rating adjustments are upgrades. As a result, the probabilities are generally lower in the middle of the diagonal than at the edges. This pattern *could* indicate that CRAs are less certain about the credit worthiness of firms that are located at the center of the rating categories, although we do not have any conclusive evidence for such a claim. Note, for instance, that the ratings with the highest probabilities of remaining in the current rating category are *A* (96.63%), and the two default categories *J* (79.05%) and *K* (79.69%).

In the unconditional transition matrix for Bank B shown in Table 4, probabilities along the diagonal of the matrix are lower than those of Bank A. For all rating categories besides *D*, *G*, *J*, firms are more likely to retain current ratings than to migrate. We note that rating changes are more likely to occur with Bank B than with Bank A. Probabilities along the diagonal of the matrix are still high, but not as high as those of Bank A. For the rating categories *A, B, C, E, F, H,* and *J*, the diagonal entry contains the highest value for all rows in the matrix. However, firms with initial ratings *D, G,* and *I* are more likely to migrate. For the period covered by our data set, companies with rating *D* are more likely to be upgraded to rating *C* than to retain rating *D*. Likewise, for companies with ratings *G* and *I*, the probability of being upgraded to ratings *F* and *H*, respectively, is higher than remaining in their current rating categories.

For companies assigned ratings *A, B, M,* and *N*, the probability of remaining in their current rating category is higher than migrating to any other rating category - i.e. the probability is higher than 50%. These categories represent the upper and lower range of rating categories. Just as was the case for Bank A, it is more likely for companies with the highest ratings *A* and *B* or the lowest ratings *M* and *N* to retain their current ratings.

Meanwhile, companies with ratings in the center of the rating spectrum, namely, ratings *C, D, E, F, G, H, I,* and *J*, are more likely to be upgraded or downgraded than remaining in their current rating categories. As noted above for Bank A, this could be due to a higher possible range of rating choices to select from, when starting from these rating categories. For Bank B, however, the diagonal values at the edges are smaller than for Bank A. For instance, the values for Bank A are 96.63% and 64.40% for ratings *A* and *B*, respectively. The equivalent values for Bank B are 54.19% and 52.98% - significantly lower than Bank A's values. For the two default states - *J* and *K* for Bank A and *M* and *N* for Bank B - the values are 79.05% and 79.69%; 62.50% and 71.32%, respectively. The difference between the two banks are smaller in the default states, but still noticeable. Purely judging by the unconditional transition matrices for the two banks, it seems as if rating changes are more common in Bank B than in Bank A. Later in this section, our measures of volatility (*RatVol*, *RatVolU*, and *RatVolD*) will further strengthen this assumption.

From/to	A	B	C	D	E	F	G	H	I	J	K
A	**96.63**	1.20	0.47	0.60	0.45	0.11	0.19	0.00	0.14	0.21	0.00
	(1.63)	(0.80)	(0.27)	(0.52)	(0.48)	(0.26)	(0.23)	(0.05)	(0.16)	(0.23)	(0.05)
B	0.85	**64.40**	22.13	7.37	3.48	1.26	0.32	0.03	0.06	0.00	0.10
	(0.44)	**(4.88)**	(4.36)	(2.87)	(1.37)	(0.37)	(0.17)	(0.03)	(0.60)	(0.61)	(2.73)
C	0.42	14.97	**54.19**	13.97	9.52	5.13	1.24	0.33	0.09	0.03	0.10
	(0.29)	(3.70)	**(9.74)**	(5.27)	(3.64)	(1.95)	(0.20)	(0.87)	(0.76)	(0.58)	(2.53)
D	0.32	4.51	30.84	**26.87**	18.61	12.11	4.95	1.23	0.36	0.18	0.03
	(0.31)	(1.75)	(11.57)	**(13.93)**	(1.74)	(1.87)	(0.79)	(0.21)	(0.13)	(0.10)	(0.03)
E	0.39	2.29	16.94	21.30	**27.86**	19.12	8.05	2.76	0.90	0.18	0.22
	(0.22)	(0.91)	(6.38)	(2.53)	**(9.86)**	(1.98)	(0.67)	(0.97)	(0.21)	(0.09)	(0.12)
F	0.20	0.76	8.18	10.08	19.01	**34.91**	15.95	7.30	2.85	0.52	0.24
	(0.13)	(0.30)	(3.08)	(2.10)	(4.45)	**(3.05)**	(1.04)	(1.84)	(0.83)	(0.30)	(0.07)
G	0.11	0.32	3.12	5.63	10.53	27.05	**30.23**	14.48	6.85	0.90	0.76
	(0.15)	(0.12)	(1.14)	(1.31)	(2.43)	(2.28)	**(1.31)**	(1.92)	(0.71)	(0.72)	(0.16)
H	0.15	0.10	1.34	2.74	5.17	14.93	24.89	**33.18**	14.72	1.51	1.27
	(0.08)	(0.19)	(0.50)	(0.86)	(1.27)	(2.12)	(2.66)	**(1.38)**	(1.72)	(2.19)	(0.63)
I	0.00	0.00	0.16	0.45	1.23	4.05	12.44	21.90	**50.71**	4.98	4.07
	(0.04)	(0.34)	(0.25)	(0.22)	(0.82)	(0.90)	(1.05)	(2.13)	**(7.72)**	(3.63)	(1.14)
J	0.74	0.00	0.00	0.74	0.00	0.00	1.45	1.53	6.64	**79.05**	9.84
	(0.82)	(0.56)	(0.48)	(0.82)	(0.50)	(1.03)	(1.74)	(5.15)	(7.16)	**(17.77)**	(2.72)
K	0.00	0.00	0.00	0.00	1.31	2.58	1.87	1.61	6.98	5.95	**79.69**
	(0.00)	(1.78)	(1.55)	(0.00)	(1.02)	(1.23)	(0.87)	(1.24)	(3.08)	(6.38)	**(11.08)**

Table 3: Unconditional Transition Matrix for Bank A (2009-2018) with annual frequency. Data is in percentages where values in parenthesis are standard deviations.

From/to	A	B	C	D	E	F	G	H	I	J	M	N
A	**54.19**	31.31	7.57	3.70	2.35	0.22	0.48	0.07	0.00	0.00	0.00	0.10
	(11.76)	(7.59)	(3.45)	(1.73)	(1.47)	(0.23)	(0.51)	(0.08)	(0.00)	(0.00)	(0.00)	(0.11)
B	6.89	**52.98**	28.53	6.52	3.34	1.00	0.34	0.19	0.06	0.14	0.00	0.00
	(2.44)	**(5.94)**	(4.26)	(1.85)	(0.82)	(0.42)	(0.18)	(0.10)	(0.05)	(0.08)	(0.00)	(0.00)
C	1.03	24.46	**45.97**	14.36	7.91	3.52	1.24	0.70	0.25	0.39	0.04	0.12
	(0.37)	(2.89)	**(2.30)**	(2.35)	(1.07)	(0.46)	(0.17)	(0.17)	(0.15)	(0.17)	(0.03)	(0.05)
D	0.68	6.96	29.70	**27.83**	20.02	8.07	2.94	2.10	0.85	0.65	0.12	0.08
	(0.36)	(0.75)	(1.66)	**(2.27)**	(2.97)	(0.88)	(0.35)	(0.35)	(0.28)	(0.30)	(0.09)	(0.06)
E	0.47	3.26	15.23	21.89	**31.20**	15.07	6.64	3.28	1.56	1.11	0.03	0.26
	(0.33)	(0.65)	(1.01)	(3.11)	**(1.34)**	(1.56)	(0.64)	(0.31)	(0.29)	(0.30)	(0.03)	(0.11)
F	0.05	1.69	8.16	11.38	23.97	**26.48**	14.11	8.07	3.30	2.03	0.32	0.46
	(0.05)	(0.39)	(1.13)	(0.98)	(1.12)	**(1.47)**	(1.37)	(0.83)	(0.29)	(0.51)	(0.17)	(0.13)
G	0.45	1.19	4.71	6.57	15.78	22.68	**19.78**	15.64	6.75	4.63	0.69	1.13
	(0.48)	(0.29)	(0.89)	(1.00)	(0.99)	(0.97)	**(1.43)**	(1.60)	(0.97)	(0.66)	(0.18)	(0.23)
H	0.20	0.76	3.25	4.60	9.93	16.21	20.33	**23.74**	9.48	8.93	1.09	1.48
	(0.21)	(0.49)	(1.15)	(1.02)	(0.89)	(1.38)	(1.56)	**(2.24)**	(1.07)	(0.93)	(0.19)	(0.32)
I	0.00	0.21	1.70	2.69	7.67	11.90	15.29	24.64	**16.90**	15.26	1.95	1.79
	(0.00)	(0.15)	(0.57)	(0.68)	(1.04)	(1.46)	(1.57)	(1.61)	**(1.30)**	(1.78)	(0.42)	(0.56)
J	0.05	0.14	0.57	1.35	2.86	7.55	8.51	12.05	13.23	**45.58**	4.54	3.59
	(0.05)	(0.11)	(0.26)	(0.32)	(0.96)	(1.30)	(1.72)	(1.20)	(1.63)	**(2.02)**	(0.75)	(0.91)
M	0.00	0.00	0.00	0.00	0.41	0.00	2.44	1.27	0.41	26.03	**62.50**	6.94
	(0.00)	(0.00)	(0.00)	(0.00)	(0.44)	(0.00)	(1.81)	(0.96)	(0.44)	(10.95)	**(9.59)**	(2.39)
N	0.00	0.62	0.47	1.53	3.37	4.14	4.09	2.82	1.53	5.05	5.05	**71.32**
	(0.00)	(0.46)	(0.34)	(0.82)	(0.99)	(1.57)	(2.11)	(1.71)	(0.94)	(1.87)	(1.28)	**(2.93)**

Table 4: Unconditional Transition Matrix for Bank B (2009-2018) with annual frequency. Data is in percentages where values in parenthesis are standard deviations.

5.2 Rating Stability

Since credit ratings are based on fundamental data and most CRAs claim they are through-the-cycle measures of credit quality, ratings should not change frequently. Occasionally, however, macroeconomic or firm-specific changes lead to adjustments of credit ratings. If a CRA is successful in assessing the creditworthiness of its obligors, rating adjustments should not occur frequently, be large, or change direction regularly. To investigate if this is the case, we calculate and analyze five different measures of volatility for the two banks. These are Rating Volatility (*RatVol*), Rating Volatility due to downgrades (*RatVolD*), Rating Volatility due to upgrades (*RatVolU*), Large Rating Changes (LRC) and Rating Reversals (RR).

5.2.1 Rating Volatility

RatVol is a measure for volatility that condenses volatility data from the two-dimensional ratings transition matrix into a single scalar for each time period. The measure can be further split into volatility caused by downgrades (*RatVolD*) and upgrades (*RatVolU*) as shown in Figures 2. All three measures represent data over the previous year - e.g. $RatVol_t$ describes the volatility from year $t-1$ to t. When performing a Dickey-Fuller test on RatVol for both banks separately, we find that the measures are stationary at a 1% significance level, fulfilling the condition for regression analysis. For Bank A, there has been an upward trend both in total volatility and volatility due to upgrades and downgrades. Because all three measures represent data over the previous year, Bank A achieves a low volatility ”score” in the year leading up to 2011. Likewise, its high is reached in the year leading up to 2016.

From the second graph in Figure 2, we conclude that Bank B's ratings are generally more volatile, in absolute terms, than Bank A. In other words, its ratings are less stable over time. The graph also exhibits a negative trend in both total volatility and volatility due to upgrades and downgrades for Bank B, contrary to the positive trend in Bank A.

The figure indicates that the volatility started to pick up between 2014 and 2015 for Bank A. This is also the period that the oil price hit its ten-year low. Consequently, the period 2015-2017 clearly stands out as a period of strong instability of ratings for this bank. However, there is only insignificant differences between volatility due to upgrades and downgrades. In other words, both upgrades and downgrades contribute with approximately the same proportion to the total volatility throughout the whole time period.

For Bank B on the other hand, *RatVol* and *RatVolU* continue to decrease following the year 2014, despite this bank's strong dependency on the oil and gas sector. One possible reason is that since Bank B is more exposed to the oil and gas sector, it might be more reluctant to provide loans *and assign credit ratings* to firms that pose high credit risk. Bank B's higher exposure to the more cyclic oil and maritime sectors could also explain the higher ratings volatility as compared to Bank A. An even more likely reason is that the total number of firms that remain customers of Bank B has declined since 2014. The data set from Bank B only contains current customers, so if a firm defaults before being assigned a default rating, it does not show up in the default statistics. Thus, these defaulting companies do not affect *RatVol* and other volatility measures the way they should, so that the actual measures could be somewhat different than the results imply.

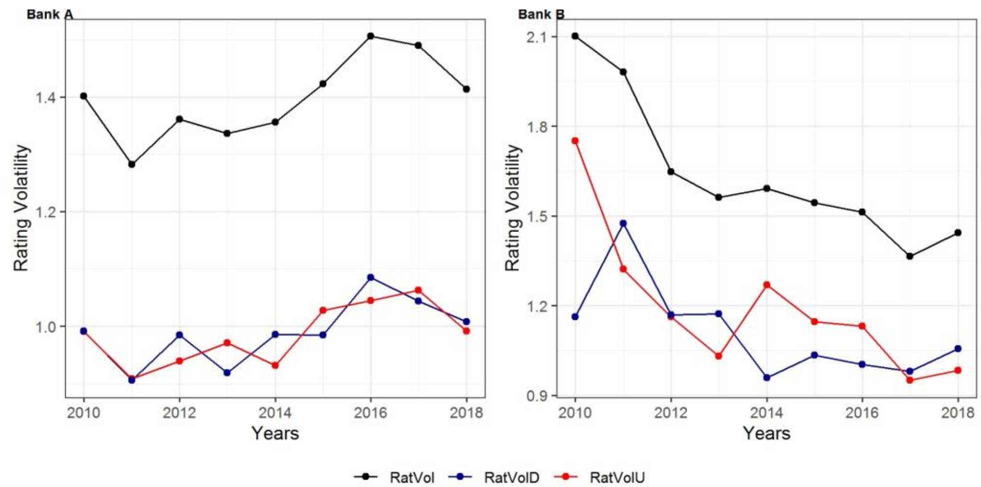

Figure 2: Total rating volatility (*RatVol*), and volatility caused by downgrades (*RatVolD*) and upgrades (*RatVolU*) for Bank A and B during the period 2009-2018. Note that the spacing for the values on the vertical axes on the two subplots are not of equal magnitude.

5.2.2 Large Rating Changes

Large Rating Changes (LRC), i.e. adjustments of three or more notches over two consecutive years during the ten-year period, is shown in Figure 3 for both banks. It is shown in blue on the left axis.

For Bank A, we can see an increase in LRC from 0.7% of annual rating changes at its lowest in 2011 to 1.3% of all rating changes in 2017. Although representing a limited amount of all ratings, LRC almost doubles from 2011 to 2017. Such a drastic increase could be justified if, for instance, there was a change in the business cycle from a peak to a trough. We investigate this possibility and present the results later in this section. In general, however, a rise in the value for LRC may suggest that a CRA has been too slow at incorporating the changes of credit risk in its obligors. For our data, this claim has to be somewhat modified. Credit ratings in our data set are only updated at an annual frequency, not at the bank's own desired frequency. Therefore, one would expect a higher value for LRC at a bank than for a regular CRA, and we cannot directly conclude that the bank is slow to incorporate the change in risk for its obligors.

For Bank B, LRC reaches its highest level at about 4% of annual rating changes in 2011 and its lowest at about 1.2% in 2017. Almost all companies in rating class *A* are downgraded between 2010 and 2011, which could partly explain the sharp increase in LRC in 2011.

Compared to Bank A's high of 1.3% and low of 0.7%, LRC for Bank B, with its high of 4% and low of 1.2%, is much higher than for Bank A throughout the whole duration of our data, except for in 2017. Also, as can be seen in Figure 3, the trend for LRC is *increasing* for Bank A and *decreasing* for Bank B.

Later in this section, we examine whether the banks trade off accuracy for increased volatility, as measured by AR and LRC respectively.

5.2.3 Rating Reversals

The ratio of Rating Reversals (RR) - i.e. a rating adjustment in one direction followed by a rating change in the opposite direction - to the total number of ratings during the ten-year period for Bank A and B, can be seen in Figure 3, in red measured on the right hand axis.

For Bank A, we see that the occurrence of rating reversals is stable throughout the whole time period. Nevertheless, rating reversals overall occur at a relatively high frequency, with between 8% and 9% of all rating changes being rating reversals. Again, part of the reason for this could be the annual frequency of credit rating updates. However, RR should not be affected by the update frequency to the same degree as LRC. In fact, it can be argued that a lower credit update frequency should result in a lower RR since the lower frequency acts as an averaging process. As a result, the RR should be lower than if updates occurred more regularly. Since we do not have any direct comparable data for this, we only mention it but cannot verify if this is the case.

Rating Reversals (RR) for Bank B vary between 8% and 14% of all rating changes. RR spikes in 2014, the same year that the oil price fell from over 100 USD/bbl to less than 50 USD/bbl. Table 2 indicates that the total default rate - i.e. new companies assigned ratings *M* and *N* - appears to fall from 1.12% in 2013 to 0.89% in 2014. Due to this bank's high exposure to the oil industry, we expected the default rate to increase during this period. However, as previously explained, the seemingly lower rate of default is due to the data set only containing current customers, so it does not capture the higher number of defaults in this period. This results in volatility measures that have to be interpreted with care.

When comparing RR for both banks simultaneously in Figure 3, it becomes clear that Rating Reversals are higher and more volatile for Bank B than for Bank A through most of the period. This is in line with the observations from the transition matrices in the previous chapter. Whereas RR for Bank A varied between 8% and 9%, between 8% and 14% of all rating changes were rating reversals for Bank B. The rating volatility for Bank B declines significantly after 2014. We once again point to the fact that the customers of Bank B are generally more dependent on the business cycle and particularly sensitive to oil price shocks. Consequently, large changes in either the business cycle and/or oil price shocks, has a greater effect on this bank than Bank A. In the period our data sets span, this could help explain why the default rates of Bank B remain stable even through oil price shocks, because defaulting companies cease to be customers of the bank in such times and are thus removed from the data set.

In all measurements of ratings volatility, we observe that Bank B has a higher ratings volatility compared to Bank A, in line with observations from transition matrices from the previous chapter. However, ratings volatility of Bank B declines significantly after 2014. Possible explanations might be two-fold. Ratings from Bank B are more dependent on the business cycle, with a larger share of loans to cyclical sectors. Secondly, a large oil induced cyclical shock can quickly wipe out distressed companies and cause them to cease being customers at the bank. Later in this section, we investigate whether the banks trade off accuracy for increased volatility, as measured by AR and RR respectively.

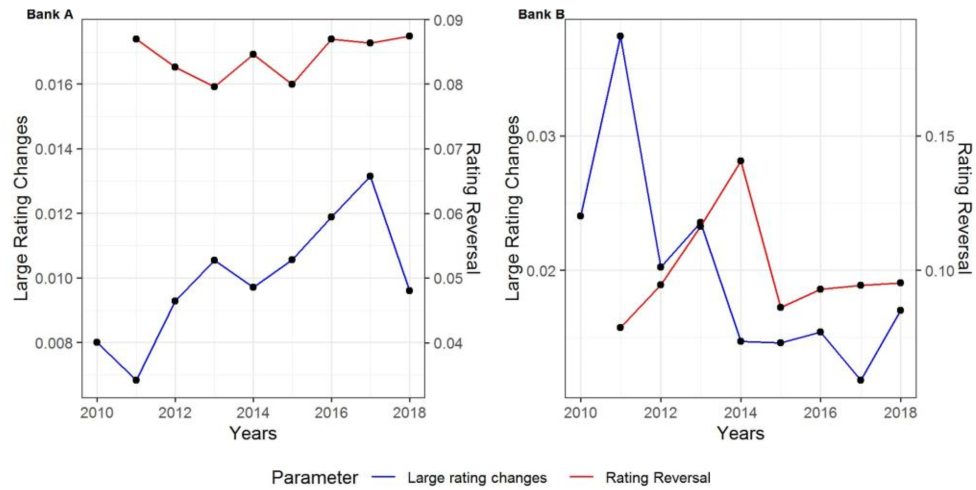

Figure 3: Large rating changes (LRC) and rating reversals (RR) for both Bank A and Bank B during the period 2009-2018. Note that the spacing for the values on the vertical axes on the two subplots are not of equal magnitude.

5.3 Relationship between Measures of Volatility

In order to compare our measures of volatility - *RatVol*, *RatVolU*, and *RatVolD* - with more conventional measures of volatility, we calculate the correlations between them on the data set for the two banks. The results in Table 5 and 6 show some intriguing relationships.

For Bank A, we observe high, significant correlations between LRC and our three measures of volatility: *RatVol*, *RatVolU*, and *RatVolD*. The correlation is slightly higher for *RatVolU* (0.81) than for *RatVolD* (0.68), suggesting that large rating changes are slightly more frequent or of a larger magnitude among upgrades than downgrades. We note small and insignificant correlations between our three volatility measures and RR. This indicates that rating reversals are not significantly more frequent in periods of higher total volatility, nor in periods of higher volatility due to upgrades or downgrades. From this, we conclude that credit rating reversals seem to occur independently of overall rating volatility.

For Bank B, we observe high, significant correlations between LRC and *RatVol* and *RatVolD*. As seen in Table 6, the correlation is very high for *RatVolD* (0.98). This suggests that large rating changes are quite frequent among downgrades. The lower, *insignificant* correlation between LRC and *RatVolU* (0.44) implies that large rating changes do not occur as frequently for upgrades. As was the case for Bank A, we note relatively small and insignificant correlations between our three volatility measures and RR.

Our analysis shows insignificant correlations between RR and the RatVol measures for both banks. If all measures contained mostly the same information regarding rating volatility, the correlations would ideally be high and significant. The lack of such a relationship suggests that our measures for volatility contain some information absent in traditional measures of stability. For instance, our measures include a full account of rating transitions because they include not only small or large rating changes, but also the direction of the adjustment. Therefore, it can to some extent be argued that our measures combine information present in the traditional stability measures and provide a more accurate account of rating stability.

	RatVol	*RatVolU*	*RatVolD*	LRC	RR
RatVol	-	0.93***	0.94***	0.79**	0.26
RatVolU	0.93***	-	0.75**	0.81***	0.06
RatVolD	0.94***	0.75**	-	0.68**	0.42
LRC	0.79**	0.81***	0.68**	-	-0.05
RR	0.26	0.06	0.42	-0.05	-

Table 5: Correlation between traditional measures of rating stability and our alternative measures of volatility for Bank A. Asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

	RatVol	*RatVolU*	*RatVolD*	LRC	RR
RatVol	-	0.91***	0.70**	0.77**	0.23
RatVolU	0.91***	-	0.34	0.44	0.07
RatVolD	0.70**	0.34	-	0.98***	0.45
LRC	0.77**	0.44	0.98***	-	0.32
RR	0.23	0.07	0.45	0.32	-

Table 6: Correlation between traditional measures of rating stability and our alternative measures of volatility for Bank B. Asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

5.4 Rating Quality

The Accuracy Ratio (AR) is a measure for the discriminatory power of rating systems. It measures how well a credit rating system captures and predicts the default risk of companies. An AR-value equal to 1 indicates that a CRA perfectly predicts which companies will default. Values closer to 0 indicate that a CRA's ratings do not possess any predictive power of which firms will default. In Figure 4, the resulting accuracy ratios for the two banks are presented, Bank A in blue and Bank B in red. When performing a KPSS test on AR for both banks separately, we find that the measures are trend stationary at a 10% significance level, fulfilling the condition for regression analysis.

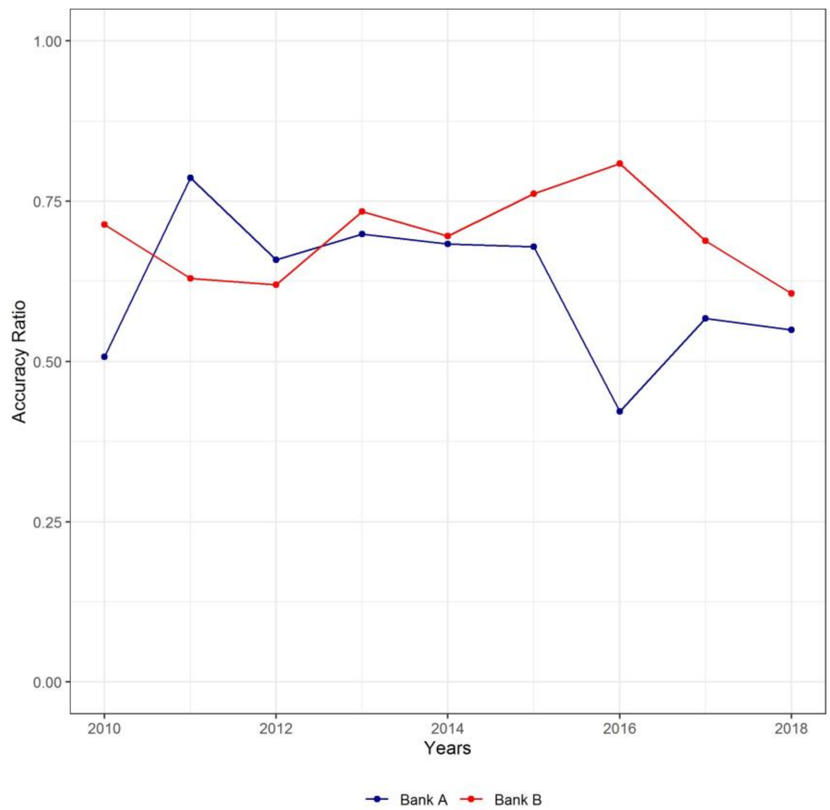

Figure 4: Accuracy ratio for both banks in the period 2010-2018. Values closer to 1 indicate more accurate predictions of default. Bank B outperforms Bank A in terms of more accurately predicting default rates in 7 out of the 9 years.

The Accuracy Ratio is higher than 0 for both banks throughout the nine-year period. For Bank A, the mean value is 0.62, while it is slightly higher for Bank B, at 0.70. In comparison, a previous study calculated the AR for Moody's and Standard & Poor's during the period 1994-2011 to be 0.71 and 0.73, respectively (Carvalho et al. (2014)). We can, therefore, conclude that both banks in question have relatively high values for AR, implying that their rating systems overall seem to predict defaults adequately.

In 7 out of the 9 years that our data spans, Bank B is more accurate at predicting defaults than Bank A. Nevertheless, the values for the two banks are reasonably similar up until 2015, at which point they begin to diverge. The Norwegian Central Bank identifies a clear correlation between Norway's GDP and the oil price (Solheim (2008)). In 2014-2016, the oil price declined sharply. Consequently, the state of the Norwegian business cycle in the same period was in a so-called trough, with negative deviation from the GDP trend.

The divergence between the Accuracy Ratio of Bank A and Bank B occurs at exactly the same time as the oil price collapses and the trough materializes. During this period, the default prediction accuracy of Bank A decreases. Meanwhile, despite its higher (indirect) exposure to oil price, the AR for Bank B improves. One possible explanation for this seemingly inverse relationship between the Accuracy Ratio of the two banks and the oil price, is the difference in the amount of experience with distressed oil-service companies that each of them have. Since Bank A has less customers in petroleum related industries, its default prediction models might not be as well-suited at assigning correct default probabilities to such companies as those of Bank B. Tables 1 and 2 in the Data section show that Bank B consistently has lower default rates as a percentage of total ratings than Bank A, ever since the reversal of the state of the business cycle in 2014. Consequently, another possible reason for Bank B's higher AR could be that it is less difficult to correctly predict which *oil-service* companies will default in periods of low oil prices, than it is to predict default probabilities in *non-oil* sectors in the same economic environment. If the companies associated with Bank B that did default were indeed mostly petroleum related companies, and considering its substantial experience with such companies, this could explain the higher prediction accuracy. A third possible reason mentioned above, is the fact that companies that defaulted prior to having received a default rating are not recorded as defaulted in the data set, resulting in a higher accuracy measure than is actually the case. Since we do not have access to sector-specific default rates of either bank, we do not examine these hypotheses further.

Comparing the accuracy ratio with ratings volatility for the two banks over time, we observe a potential trade-off of accuracy for rating stability. We explore this relationship more thoroughly in a section 5.6. While ratings are more stable for Bank A throughout our sample period, default prediction accuracy fell after the oil price collapse and the resulting economic slowdown. Bank B on the other hand has a more stable default prediction accuracy throughout this same period, perhaps due to the reasons mentioned in the paragraph above.

5.5 Business Cycle Effects

As noted by Altman & Rijken (2006), CRAs advertise a through-the-cycle methodology in their rating assignments. In this context, this means that credit ratings should not be significantly dependent on the state of the business cycle. In this paper, we use data from two banks. Consequently, our results could "legitimately" differ from that of previous studies performed on data from large CRAs with a long history (such as Standard & Poor's and Moody's). Nevertheless, we analyze the effect that the state of the business cycle has on the volatility and quality of credit ratings, in order to assess whether banks can claim to follow the same methodology as CRAs.

5.5.1 Business Cycle Effects on Volatility of Ratings

Previous studies find contradictory evidence for a through-the-cycle methodology (see, e.g., Carvalho et al. (2014)), even among the biggest, leading CRAs. Naturally, we expected the same to be the case for our data, with ratings volatility varying with the state of the business cycle. We initially expected positive *GDP.Dev* to be accompanied by a higher *RatVolU*, a lower *RatVolD*, and a lower or insignificant *RatVol*. The rationale behind these initial assumptions is that good economic times results in more upgrades, less downgrades and more stability in credit ratings, i.e. lower total volatility.

We run a comprehensive multivariate analysis of the effect that the business cycle might have on the volatility of ratings. In addition to *GDP.Dev*, we include five more business cycle proxies: *SwapRate*, *VIX*, *NewBonds*, *NewLoans*, and *RateDef*. When performing a multivariate regression with all six measures as independent variables, the results produce only one statistically significant coefficient for either bank (see

Table 12 in the Appendix). We suspect that this is due to the presence of multicollinearity among the independent variables. Therefore, we calculate the variance inflation factor (VIF) for both of the banks to assess which variables are likely to be exhibiting indications of multicollinearity, i.e. VIF values typically higher than 30, before removing these variables. We then proceed to run multivariate regressions on the remaining independent variables.

The VIF values for Bank A and B are presented in Table 7. Analyzing the VIF values of both banks jointly, we see that the values are high for *VIX* and *NewBonds*. These variables are highly correlated, with a correlation of -0.92 as seen in Figure 6 in the Appendix. The VIF value is also quite high for *SwapRate*, a variable that shows signs of correlation with *GDP.Dev*. Due to their high VIF values, these variables are removed as independent variables. We also believe that these variables (*VIX*, *NewBonds*, and *SwapRate*) are proxies for the business cycle that affect credit ratings of CRAs more than banks, so their removal from the model can be economically justified. The resulting multivariate regression is, therefore, composed of the explanatory variables *GDP.Dev*, *NewLoans*, and *RateDef* since they appear to be uncorrelated.

	Bank A	**Bank B**
	VIF	VIF
$GDP.Dev_t$	9.37	5.98
$SwapRate_t$	17.70	7.76
VIX_t	43.81	26.18
$NewBonds_t$	57.89	35.09
$NewLoans_t$	1.92	3.84
$RateDef_t$	3.66	8.79

Table 7: Variance inflation factor (VIF) values for the following measures: *GDP.Dev*, *SwapRate*, *VIX*, *NewBonds*, *NewLoans*, and *RateDef*. Higher VIF values indicate multicollinearity between independent variables.

After eliminating correlated independent variables, we run a multivariate regression with the remaining variables. The result is presented in Table 8. For Bank A, we see that all three volatility measures show significantly negative relationships with *GDP.Dev*. In other words, rating assignments appear to be affected by the business cycle with higher volatility during troughs. We can conclude that the stability of ratings is positively correlated with the state of the business cycle and that Bank A's credit ratings are point-in-time measures of credit risks.

Our results for Bank A are consistent with the result found in previous studies, such as Amato & Furfine (2004) and Carvalho et al. (2014). The evidence implies that rating adjustments are more intense during worse economic times and less so during better times. The stability of credit ratings in better times indicates a procyclical rating policy. In these periods, there will be less incentive to change credit assessments compared to bad times due to less overall default risk. A possible explanation for higher rating volatility in periods when the economy underperforms, is that defaults are more common in such periods. Thus, CRAs may attempt to improve default prediction by over-eagerly reassessing ratings leading to higher volatility and lower stability.

For Bank B on the other hand, only *RatVol* and *RatVolU* have significant coefficients. They both appear to be positively related to the rate of default, *RateDef*, suggesting higher total volatility and upward adjustments when the rate of default is higher. Higher default rates are presumably more common during troughs, implying more volatility in ratings during bad economic times. In other words, rating volatility appears to be countercyclical, as expected. However, as previously argued, the measure of the default rate for Bank B is likely to be too low. If we had access to the *real* default rate for this bank, this *could* alter the results of this regression analysis, but not necessarily. Since we do not have access to this data, we do not discuss it any further.

Bank A	$RatVol_t$	$RatVolU_t$	$RatVolD_t$
Intercept	1.46***	1.02***	1.04***
t-ratio	(14.22)	(13.21)	(10.65)
$GDP.Dev_t$	-1.71**	-1.24**	-1.18*
t-ratio	(-2.79)	(-2.68)	(-2.02)
$NewLoans_t$	0.00	0.00	0.00
t-ratio	(0.86)	(1.05)	(0.45)
$RateDef_t$	-0.07	-0.05	-0.05
t-ratio	(-0.82)	(-0.72)	(-0.65)
Adj.R^2	0.54	0.52	0.30

Bank B	$RatVol_t$	$RatVolU_t$	$RatVolD_t$
Intercept	1.08***	0.46	1.13**
t-ratio	(3.87)	(-1.03)	(3.54)
$GDP.Dev_t$	0.14	-1.43	1.75
t-ratio	(0.09)	(-1.03)	(0.95)
$NewLoans_t$	0.00	0.00	0.00
t-ratio	(0.30)	(1.19)	(-0.66)
$RateDef_t$	0.45**	0.53***	0.05
t-ratio	(2.87)	(3.98)	(0.28)
Adj.R^2	0.69	0.77	0.06

Table 8: Results from the multivariate regression examining the relationship between business cycle measures (*GDP.Dev*, *NewLoans*, and *RateDef*) and our volatility measures (*RatVol*, *RatVolU*, and *RatVolD*) for both banks. As before, asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

We performed several regularization regression techniques (Lasso, Ridge and Elastic Net), but neither of these yielded statistical significant results.

5.5.2 Business Cycle Effects on Quality of Ratings

Previous studies find evidence for increased rating quality during recessions (see Bar-Isaac & Shapiro (2013) and Bolton et al. (2012)). We investigate if this is the case for our data sets by examining the dependency that the state of the business cycle has on the quality of ratings with multivariate regressions.

Initially, we expected that firms with higher credit risks underperform when the economy underperforms, thus exposing their actual creditworthiness to CRAs. CRAs in turn end up with more accurate ratings in such periods. In other words, we expected rating quality to be countercyclical, implying a negative relation between accuracy and the state of the business cycle.

We run a comprehensive multivariate analysis of the effect that the business cycle might have on the quality of ratings. When running a multivariate regression with all seven independent variables, the results have some complications (see Table 13 in the Appendix). Several of the coefficients for Bank A seem to have a statistically significant effect on accuracy. However, an adjusted R^2 equal to 1 is unreasonable and suggests a spurious regression. The results for Bank B on the other hand yield *no* statistically significant coefficients. We suspect that this is due to the presence of multicollinearity among our independent variables, which could also be present in Bank A. Therefore, we first calculate the variance inflation factor (VIF) for both of the banks. Then, we remove the variables exhibiting indications of multicollinearity, i.e. variables with high VIF values, typically higher than 30, and run multivariate regressions on the remaining independent variables.

	Bank A	**Bank B**
	VIF	VIF
$GDP.Dev_t$	11.08	7.07
$SwapRate_t$	20.04	10.47
VIX_t	53.27	41.51
$NewBonds_t$	60.84	35.99
$NewLoans_t$	1.96	3.91
$RateDef_t$	3.68	10.67
$RatVol_t$	9.58	23.79

Table 9: Variance inflation factor (VIF) values for the following measures: *GDP.Dev*, *SwapRate*, *VIX*, *NewBonds*, *NewLoans*, *RateDef*, and *RatVol*. Higher VIF values indicate multicollinearity between independent variables.

The VIF values for Bank A and B are presented in Table 9. Higher VIF values indicate multicollinearity between independent variables. Analyzing the VIF values of both banks jointly, we see that *VIX* and *NewBonds* have the highest VIF values. These variables are too highly correlated, which is an issue in OLS regression. In addition, we believe that these measures are more relevant when analyzing credit ratings from CRAs and not banks. Consequently, we remove them as independent variables. This result implies that *GDP.Dev*, *SwapRate*, *NewLoans*, *RateDef*, and *RatVol* are the only sufficiently uncorrelated variables and thus included in the following multivariate regression.

After removing the correlated independent variables, we run a multivariate regression with the remaining variables for each bank. The results are presented in Table 10. For Bank A, there is a statistically significant *negative* relationship between *AR* and the independent variable *RateDef* at the 10% level. The negative relationship between the two variables implies that the accuracy, i.e. quality, of ratings is higher when the rate of default is lower. Generally, the rate of default is lower during better times, so the regression implies higher accuracy when the economy performs well. This could imply that the bank finds it more challenging to anticipate which companies will struggle in poor economic conditions than in more favorable conditions. This counter-intuitive result contradicts previous work such as Bar-Isaac & Shapiro (2013) and Bolton et al. (2012). The crucial difference between these studies and this paper is the source of the data sets. Bolton et al. (2012) note that, due to the conflict of interest inherent in CRAs business models, they have a tendency to understate risk to attract new business in periods when the economy performs well, leading to rating bias. This, in addition to due diligence potentially deteriorating in such periods, is a possible reason for decreased accuracy in such periods. Bank A on the other hand, does not face the same conflict of interest, as it only assigns ratings to its own customers. Therefore, the inverse conclusion is justifiable. We also find a surprising significant relationship between *AR* and *RatVol* at the 5% level. After controlling for business cycle effects, higher rating volatility seems to be associated with lower accuracy. We will briefly explore this in the next section. The remaining explanatory variables are insignificant.

For Bank B, the adjusted R^2 is 0.26 before removing variables suspected of multicollinearity (see Table 13 in the Appendix) and 0.57 afterwards, indicating an improved model. There is a statistically significant *negative* relationship between *AR* and the independent variable *GDP.Dev* at the 5% level. In other words, worse economic times seem to be associated with higher accuracy. The results for Bank B are, unlike Bank A, consistent with the results of Bar-Isaac & Shapiro (2013) and Bolton et al. (2012). As previously argued in this section, this *could* be due to Bank B: 1) being comparatively more experienced with distressed companies (particularly oil service companies) than Bank A, 2) being more conservative and cautious when extending debt, resulting in lower overall default rates than Bank A even in times of plunging oil prices, and therefore having better predictions of which companies *actually* do default during such periods, or 3) having a too high AR due to the bank's handling of some of its defaulting customers by removing them from the data set, so that these firms are not recorded as defaulted. The coefficient for *SwapRate* is *positive* at a significance level of 10%. This implies that the accuracy is higher when the difference between long (10 year) and short (2 year) swap rates are higher. A steep upward sloping swap curve is indicative of better times. The relationship between *AR* and *SwapRate*, therefore, could suggest that the accuracy is higher when the economy is doing well. However, considering the small value of this coefficient (0.32) combined with the fact that the value of *SwapRate* varies little during 2010-2018, the period that our data spans (between 0.75% and 1.5%), this is essentially treated as a constant when compared to the effect that the other significant variable, *GDP.Dev*, has on the regression. The latter explanatory variable has a coefficient of -2.82 and a much larger variation in its value during the period (varying between -5% and +3%). We can thus conclude that the result of the multivariate analysis is reasonably, although not perfectly, consistent with that of previous studies. The remaining explanatory variables are insignificant.

	Bank A		Bank B	
	AR_t	t-ratio	AR_t	t-ratio
Intercept	3.12**	(4.78)	0.56**	(2.81)
$GDP.Dev_t$	-0.90	(-0.91)	-2.82**	(-3.89)
$SwapRate_t$	0.25	(1.66)	0.32*	(2.53)
$NewLoans_t$	0.00	(1.89)	0.00	(-1.04)
$RateDef_t$	-0.29*	(-2.67)	-0.13	(-1.31)
$RatVol_t$	-1.74**	(-3.97)	0.02	(0.15)
$Adj.R^2$	0.82			0.57

Table 10: Results from the multivariate regression examining the relationship between business cycle measures (*GDP.Dev*, *SwapRate*, *NewLoans*, *RateDef*), a volatility variable (*RatVol*), and the rating quality measure (*AR*) for both banks. As before, asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

We performed several regularization regression techniques - Lasso, Ridge and Elastic Net - but they yield no statistical significant results and are thus not included.

5.6 Relationship between Accuracy and Stability

Cantor & Mann (2006) claim that CRAs follow a through-the-cycle methodology and trade off accuracy for increased stability. Therefore, if CRAs do indeed trade off accuracy in order to attain more stable ratings, a lower accuracy should be accompanied by a higher stability, i.e. lower volatility. If that is the case, we should observe an upward sloping relationship in figure 5 below. The reasoning behind this assumption is that rating changes of a higher frequency or magnitude should more accurately predict the relative risk of the firm. Table 11 displays the results of the univariate linear regressions for *AR* and the three volatility measures, *RatVol*, *LRC*, and *RR* for both banks.

For Bank A, we conclude that the results are significant for *RatVol* (Figure 5) and insignificant for *LRC* and *RR*. Moreover, for all three volatility measures, a higher volatility is associated with a lower AR - i.e. there is a negative correlation between the measures. In other words, accuracy is higher when the volatility is lower. This contradicts common sense, as you would expect that more frequent rating adjustments (higher *RatVol*) would result in more accurate ratings (higher *AR*). We speculate whether Bank A's increased stability during better times could arise from less pressure from management to accurately predicting defaults due to overall lower default rates. When the economy underperforms, higher default rates result in larger and more frequent rating adjustments as shown in the previous section. This might compel CRAs to attempt to more accurately predict defaults by reassessing ratings too frequently, resulting in higher volatility *and* lower accuracy in such periods and possibly explaining the inverse relationship. Our results do, however, not indicate that Bank A trades off accuracy for stability of its credit ratings.

For Bank B, the results indicate no statistically significant relationships. Nevertheless, we find the same negative relationship between *AR*, and *RatVol* and *LRC* as we did for Bank A, but a positive relation for *RR*.

We stress that the results for Bank B are insignificant, so no definitive unambiguous conclusions can be drawn. However, as a consequence of weak and insignificant relationships between accuracy and stability, one could claim that this indicates a lack of a tradeoff between the two measures. Thus, Bank B appears to *not* trade off stability for accuracy.

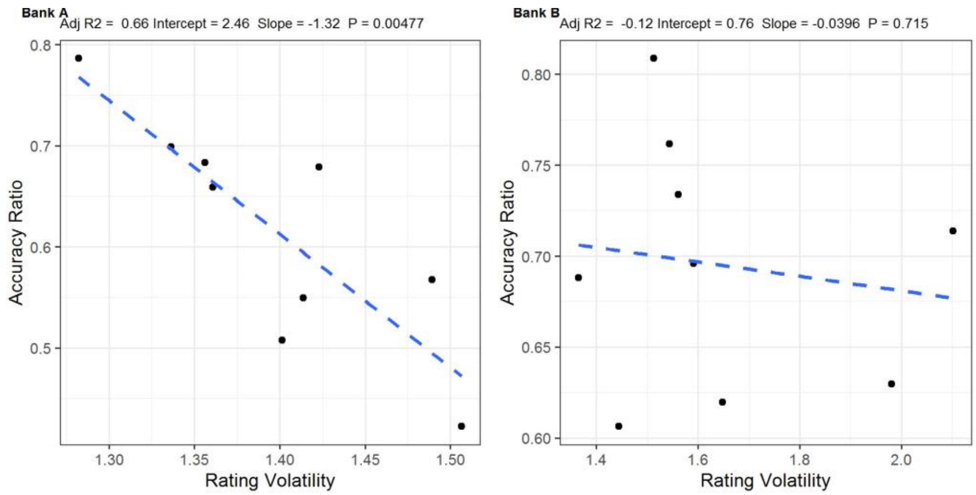

Figure 5: The relationship between the accuracy (AR) and rating volatility (*RatVol*) for both banks. Each point in the figure represents data for a given year.

	Bank A					Bank B				
	Intercept	$RatVol_t$	LRC_t	RR_t	R^2	*Intercept*	$RatVol_t$	LRC_t	RR_t	R^2
AR_t	2.46***	-1.32***			0.70	0.76***	-0.04			0.02
t-ratio	(5.42)	(-4.07)				(4.41)	(-0.38)			
AR_t	0.89***		-27.22		0.21	0.76***		-3.09		0.13
t-ratio	(4.39)		(-1.36)			(11.59)		(-1.00)		
AR_t	1.09			-15.00	0.18	0.65***			0.44	0.01
t-ratio	(1.74)			(-1.16)		(4.28)			(0.29)	

Table 11: OLS estimates of the regressions examining the relationship between *AR* and *RatVol*, *LRC*, and *RR* for both banks. Asterisks denote statistical significance at the 1% (***), 5% (**) and 10% (*) levels.

Even though we previously observed differences in the development of ratings volatility and accuracy ratio between the banks over time, we do not find a *statistically significant* trade-off between rating stability and accuracy. A possible explanation for the lack of quantitative evidence for such a trade-off, is that our data set consists of ratings from Norwegian savings and loans *banks* and not *CRAs*. Unlike CRAs which regularly update their credit ratings, the data sets from banks include ratings reassessed at a *fixed frequency of once a year*, regardless of firm-specific changes in creditworthiness or changes in the economic environment. As a result, we cannot make as definitive conclusions as previous studies have done, such as Cantor & Mann (2006). They conclude that there is a lack of quantitative evidence supporting their hypothesis of the CRAs' willingness to induce higher rating accuracy to achieve more stability. We only conclude that there is a negative relationship between AR and volatility. Whether this is intentional or not, we can only speculate and neither confirm nor deny due to the nature of our data sets. Since rating assessment frequency is fixed, however, it is very unlikely that this is in fact intentional.

6 Summary and Suggestions for Future Work

Credit ratings are a vital component of financial markets, providing information to investors and regulators about the riskiness of financial debt securities and their issuers. Credit rating agencies (CRAs) therefore have a very important responsibility of assigning credit ratings reflecting the "true" creditworthiness of obligors. At the same time, many CRAs claim that they attempt to apply a *through-the-cycle* methodology to assigning credit ratings. This means, among other things, that users of credit ratings expect that credit ratings are stable over time and that only permanent changes in credit risk should result in a credit rating adjustment. This methodology also implies that the state of the business cycle should not have a significant impact on credit ratings, because ratings are simply a measure of credit risk *relative* to other firms. However, stable ratings are achieved at the expense of accuracy. CRAs thus have to balance two competing goals: a high rating *stability* and a high rating *accuracy*.

6.1 Summary

In this paper, we apply different quantitative models and tests to quantify and analyze the stability and accuracy of credit assessments. We examine the influence of the business cycle on rating adjustments and analyze the trade-off between accuracy and stability. We perform these tests on two different data sets: 1) A large data set from a Norwegian savings and loans bank (Bank A), 2) another large data set from a somewhat smaller Norwegian savings and loans bank (Bank B), which is more exposed to the oil-service sector.

We aim at applying new rating volatility measures to Nordic credit ratings (scores), contributing to existing research in several ways. We examine how different business cycle variables affect rating stability and accuracy. Furthermore, we analyze credit scores from financial institutions (banks) rather than CRAs. CRAs and banks have different incentives for their credit ratings, and we thus contribute with new results not seen in previous credit rating studies. Lastly, we implement multivariate regression methods, some of which, to our knowledge, are not used in previous credit rating research.

At the core of our analysis is a measure for rating volatility and instability developed by Carvalho et al. (2014) known as *RatVol*. It summarizes the information contained in a two-dimensional transition matrix into a single scalar for each time period. This allows us to perform time-series tests rarely seen in the credit default modeling literature.

Our results indicate that the intensity of both upgrades and downgrades varies through time for both banks. Rating volatility for Bank A follows an upward trend during the period of our data sets, whereas Bank B appears to be following a downward trend. Interestingly, we observe that the intensity of both upgrades and downgrades for Bank A is higher during bad economic times, so-called troughs. This is not just a surprising result - it is also *inconsistent* with the claim that ratings are relative rankings of firms and largely independent of the business cycle. Despite investors' need for ratings to be stable and independent of the business cycle, we find contradicting evidence and the bank appears to unintentionally target absolute levels of risk at a specific point in time. Characterizing Bank A's rating methodology as through-the-cycle is thus problematic. Bank B displays more frequent rating changes than Bank A based on transition matrices, and a higher LRC, RR and AR over time. These observations indicate a point-in-time approach to credit ratings. We also observe that Bank B produces ratings volatility that appears to be independent of the business cycle and could, therefore, adhere to a through-the-cycle methodology.

Surprisingly, we find that for Bank A the accuracy of ratings is *procyclical*, i.e. higher during better economic times. This result contradicts previous findings of Bar-Isaac & Shapiro (2013) and Bolton et al. (2012) that rating quality is countercyclical. Rating accuracy for Bank B on the other hand appears to be consistent with these studies, with higher accuracy during worse economic times. During these periods, it *could* be easier for this bank to separate companies with high probability of default from companies with low probability of default, resulting in a higher accuracy. However, a more likely reason is that the accuracy measure (AR) for Bank B is higher than its real value. Evidence from the data set suggests that not all companies that default were assigned default ratings before they were removed as customers of the bank - some were removed as customers still with a non-default rating. Consequently, these companies will not show up with a default rating in the default statistics and will thus not influence the AR value.

Notably, we do not find evidence that higher ratings volatility leads to higher ratings accuracy. In fact, higher volatility is instead associated with *lower* accuracy for Bank A and uncorrelated with accuracy for Bank B. Unlike other studies that have performed similar analysis, we cannot make definitive conclusions as to why we see this unanticipated relationship in Bank A due to the fixed frequency of rating adjustments. However, we speculate that Bank A's increased stability during better times could arise from less pressure from management to accurately predicting defaults due to lower overall default rates. Conversely, when the economy underperforms, higher default rates possibly result in larger and more frequent rating adjustments. This compels banks to attempt to more accurately predict defaults by reassessing ratings too frequently, resulting in higher volatility *and* lower accuracy in such periods and possibly explaining the inverse relationship. The insignificant relationship between accuracy and stability for Bank B suggests that they do not actively trade off stability for accuracy.

6.2 Suggestions for Future Work

We have examined the relationship between different proxies for the business cycle and credit rating stability and accuracy. When indiscriminately including all our proposed business cycle proxies and running multivariate regressions, we get few statistically significant results. However, when adjusting our analyses to overcome suspected problems of multicollinearity, our results do yield some statistically significant relationships, although varying for the two banks. It is quite possible that this is due to their differing customer locations and thus different relations to the business cycle. However, a more likely reason is differing policies with regards to handling adjustments to their customers' credit ratings, thus affecting the data sets and our analyses. Although we do not have any clear evidence of such differing policies between the two banks, we suspect this to be the case. One way to handle this possible issue is to perform the same analyses on continuously updated credit rating data from the two banks and not on annual data. We were not able to obtain this data from the banks, so we leave this for future work.

The data sets we utilize are sufficiently large for the purposes of our analyses. However, performing the same analyses on more comprehensive data sets would yield more complete results and enable the possibility of making more definitive conclusions. Although we were not able to obtain data, which were updated more frequently and in a similar fashion as credit ratings from CRAs, many banks have such data in their possession. With such data, entities (either banks or CRAs) update their credit ratings not on a fixed basis, but whenever they consider the creditworthiness of a particular company as changed. With such data, the results from the methods provided in this paper would allow direct comparison to those of previous work on CRAs. We could also more confidently conclude on the Banks' and CRAs' underlying motivations for rating adjustments, rather than simply describing the relationship between rating adjustments and different measures. Using data with more frequent updates and stretching over a longer time period, would permit the use of additional quantitative methods such as Quantile regression. Also, a longer period will naturally include data from several business cycles, leading to a more complete analysis.

In this paper, we analyze the ratings from *regional* banks in Norway and compare how rating volatility and stability changes with the business cycle. An extension of our work could look at additional national and international banks or CRAs both in the Nordics and Europe, exposed to a wider geographical region and to different industry-specific risks, and examine whether this is reflected in their rating changes.

Appendix

A. Alternative Measures of Rating Volatility

A.1 Large Rating Changes

Credit ratings are based on fundamental data and hence are not expected to change frequently. Nevertheless, unexpected changes in the economy or within a company may lead to large changes in a company's credit rating - so-called multi-notch rating adjustments. A measure known as *Large Rating Changes* (*LRC*) is used to estimate the frequency of such events Cantor & Mann (2003). It estimates the stability of credit ratings and is defined as

$$LRC_t = \frac{\sum_{i=1}^{R_t} lrc_{it}}{R_t} \quad \forall\, t \in T \tag{14}$$

where

i is the index of the customers of the bank with credit rating

R_t is the total number of all credit ratings apart from defaults in period *t*

lrc_{it} is a binary variable equal to 1 if the rating change is three or more notches from period $t-1$ to period t, and 0 otherwise

Intuitively, this measure describes the ratio of ratings that received an update of 3 or more notches from one period to the next, to the total number of ratings.

A large and sudden increase in the frequency of large rating changes - i.e. a high LRC - could indicate that CRAs have been too slow to incorporate the change in the risk of a company in their credit ratings. However, there are instances where large rating changes are justified. Therefore, LRC is not an objective measure that should be used to infer definite statements about the quality of the rating process over time. Nevertheless, large rating changes should *only* occur if there is a substantial shift in a company's risk profile.

A.2 Reversal of Ratings

Another measure used to estimate the stability of credit ratings is known as *Rating Reversals* (*RR*). In the literature this measure is defined as cases of CRAs assigning both upward and downward rating changes within a 12 months period. We define a rating reversal as

$$RR_t = \frac{\sum_{i=1}^{R_t} rr_{it}}{R_t} \quad \forall\, t \in T \tag{15}$$

where

i is the index of the customers of the bank with credit rating

R_t is the total number of all credit ratings apart from defaults in period t

rr_{it} is a binary variable equal to 1 if a rating change is an upgrade preceded by a downgrade or vice versa, and 0 otherwise. In other words, a value of 1 is assigned if a rating in period $t-2$ is equal to a rating in period t and different than the rating in period $t-1$

Intuitively, our measure of rating reversals describes the ratio of ratings experiencing a change in one direction two periods ago and then a change in the opposite direction during the previous period, to the total number of ratings.

An increase in the frequency of rating reversals may indicate that CRAs are overly sensitive to temporary shocks, resulting in rating adjustments that are based on erroneous risk assessments. However, like LRC, there are instances where rating reversals are justified. Therefore, RR is not an objective measure that can be analyzed directly over time, and it should be carefully interpreted with reference to the underlying macroeconomic conditions. These measures of volatility are more applicable to CRAs where rating assessments happen when the rating entity deems it necessary.

B. Adjusting the Business Cycle Variable with Hodrick-Prescott filter

Let y_t denote the logarithms of Norway's real GDP for years $t = 1, 2, ..., T$. The variables y_t contain a short-term trend component τ_t and a cyclical component c_t. The value of y_t equals $y_t = \tau_t + c_t$. The HP-filter creates a measure for the smoothness of the path of τ_t and involves solving the following minimization problem:

$$min_\tau \left(\sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right) \tag{16}$$

The first term is the cyclical component, $c_t = y_t - \tau_t$. It penalizes deviations of τ_t from y_t. The second term is a multiple λ that penalizes variability in the second differences of the trend component. The parameter λ is a subjectively chosen positive scalar that places a relative weight on variability in the trend components as compared to the cyclical component. A lower λ means a lower cyclical component c_t, while a higher λ means a higher cyclical component. For quarterly data, Hodrick and Prescott (1997) Hodrick & Prescott (1997) suggest $\lambda = 1600$, and for annual data $\lambda = 100$ has been proposed. Since our data is annual, we use $\lambda = 100$.

For the other business cycle variables and financial indicators (swap rate, VIX index, balance sheet of banks' assets and new bond issuance) we annualized the data as explained in the previous section.

C. Testing multicollinearity

The variance inflation factor (VIF) is a measure used to identify multicollinearity (see, e.g., Rawlings et al. (1998)). It calculates the correlation between the independent variables and the strength of that correlation by quantifying how much the variance of the estimated coefficients is inflated. The VIF value for an estimated regression coefficient β_j is given by:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{17}$$

where R_j^2 is the R^2-value when regressing the variable j on the other variables. A VIF-value of 1 indicates no correlation between this particular variable and the remaining variables. A VIF-value greater than 1 suggests multicollinearity, i.e., the covariate vector X_j is not orthogonal to all other vectors in the matrix of explanatory variables (the design matrix; X). Suggestions for cut-off values of VIF are typical between 5 and 30. When multicollinearity is observed, mitigation should be attempted.

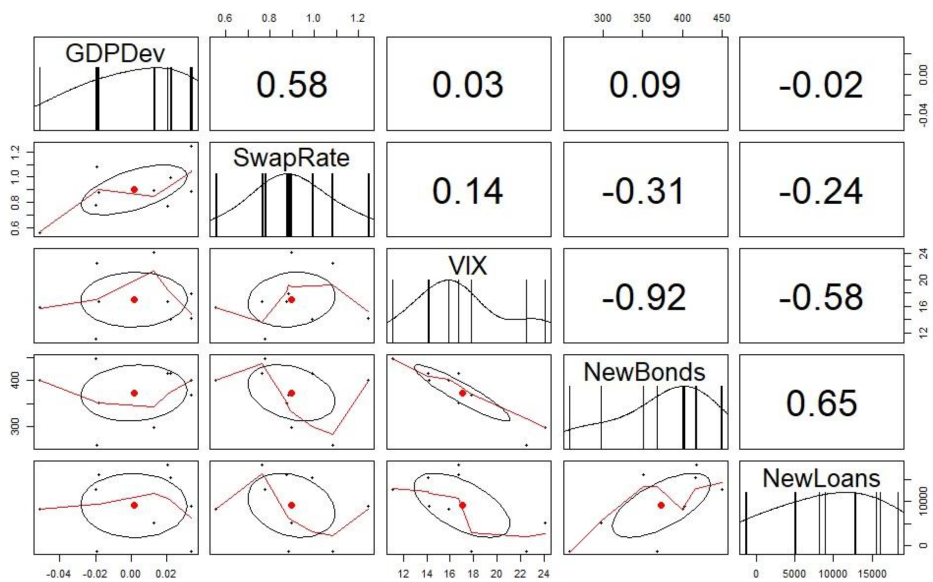

Figure 6: Combined correlation and scatter plots for the variables *GDP.Dev*, *SwapRate*, *VIX*, *NewBonds*, and *NewLoans*. Plots for the variables are along the diagonal. Correlations between different variables are shown in the top triangle of the matrix of plots. Scatter plots and correlation eclipses showing the relationship between different variables are shown in the bottom triangle of the matrix of plots.

In section 5.5.1 we noted that a multivariate regression with six measures of the business cycle as independent variables only produced one statistically significant coefficient for either bank (see Table 12 below). We suspected that this is due to the presence of multicollinearity among the independent business cycle variables. Figure 6 above further support our suspiction. A VIF analysis revealed which variables were likely to exhibit multicollinearity (table 7, section 5.5.1), and these variables were then removed from the regression analysis.

Bank A	$RatVol_t$	$RatVolU_t$	$RatVolD_t$	LRC_t	RR_t
Intercept	2.09	0.68	2.27	0.00	-0.01
t-ratio	(1.51)	(0.62)	(2.36)	(0.29)	(-0.11)
$GDP.Dev_t$	-1.02	-1.94	0.49	-0.04*	-0.06
t-ratio	(-0.61)	(-0.07)	(-0.71)	(-1.74)	(0.02)
VIF	9.37				
$SwapRate_t$	-0.18	0.18	-0.44	0.00	0.00
t-ratio	(-0.52)	(0.66)	(-1.79)	(1.45)	(0.08)
VIF	17.70				
VIX_t	-0.02	0.00	-0.03	0.00	0.00
t-ratio	(-0.66)	(0.11)	(-1.45)	(-0.64)	(1.04)
VIF	43.81				
NewBonds	0.00	0.00	0.00	0.00	0.00
t-ratio	(-0.32)	(0.38)	(-1.08)	(1.46)	(1.03)
VIF	57.89				
$NewLoans_t$	0.00	0.00	0.00	0.00	0.00
t-ratio	(-021)	(0.01)	(-0.44)	(-1.90)	(-0.05)
VIF	1.92				
$RateDef_t$	0.02	-0.11	0.13	0.00	0.00
t-ratio	(0.11)	(-0.96)	(1.29)	(-2.10)	(-0.17)
VIF	3.66				
Adj.R^2	0.58	0.53	0.66	0.96	-0.08

Bank B	*$RatVol_t$*	*$RatVolU_t$*	*$RatVolD_t$*	*LRC_t*	*RR_t*
Intercept	0.03	1.82	-1.74	-0.18	-0.07
t-ratio	(0.01)	(0.47)	(-0.75)	(-2.33)	(0.08)
$GDP.Dev_t$	-1.77	-0.34	-2.05	-0.17	0.03
t-ratio	(-0.61)	(-0.07)	(-0.71)	(-1.74)	(0.02)
VIF	5.98				
$SwapRate_t$	0.42	-0.14	0.73	0.05	0.05
t-ratio	(0.84)	(-0.17)	(1.49)	(2.85)	(0.20)
VIF	7.76				
VIX_t	0.05	-0.01	0.08	0.01*	0.00
t-ratio		(-0.20)	(1.94)	(3.51)	(0.00)
VIF	26.18				
NewBonds	0.00	0.00	0.00	0.00	
t-ratio	(0.23)	(-0.40)	(1.00)	(2.18)	(0.21)
VIF	35.09				
$NewLoans_t$	0.00	0.00	0.00	0.00	
t-ratio	(-0.19)	(0.71)	(-1.67)	(-2.10)	(0.04)
VIF	3.84				
$RateDef_t$	0.13	0.42	-0.29	-0.01	0.01
t-ratio	(0.65)	(1.27)	(-1.52)	(-1.82)	(0.2)
VIF	8.79				
Adj.R^2	0.83	0.53	0.62	0.82	-1.91

Table 12: Results from the multivariate regression examining the relationship between business cycle measures (*$GDP.Dev_t$*, *$SwapRate_t$*, *VIX_t*, *$NewBonds_t$*, *$NewLoans_t$*, and *$RateDef_t$*) and volatility measures (*$RatVol_t$*, *$RatVolU_t$*, *$RatVolD_t$*, *LRC_t*, and *RR_t*) for both banks. As before, asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

Table 12 shows the results from multivariate analyses of the effect that the business cycle might have on the *volatility* of ratings *before* removing variables suspected of multicollinearity.

	Bank A			Bank B		
	AR_t	t-ratio	VIF	AR_t	t-ratio	VIF
Intercept	5.40***	(61.11)		0.46	(0.25)	
$GDP.Dev_t$	-1.74**	(-21.65)	11.08	-3.06	(-1.26)	7.07
$SwapRate_t$	0.02	(1.48)	20.04	0.37	(0.84)	10.47
VIX_t	-0.02**	(-19.32)	53.27	0.01	(0.13)	41.51
$NewBonds_t$	0.00*	(-8.85)	60.84	0.00	(0.07)	35.99
$NewLoans_t$	0.00**	(18.07)	1.96	0.00	(-0.62)	3.91
$RateDef_t$	-0.26**	(-41.68)	3.68	-0.12	(-0.74)	10.67
$RatVol_t$	-2.73***	(-88.34)	9.58	-0.04	(-0.08)	23.79
$Adj.R^2$	1.00			-0.26		

Table 13: Results from the multivariate regression examining the relationship between business cycle measures ($GDP.Dev_t$, $SwapRate_t$, VIX_t, $NewBonds_t$, $NewLoans_t$ and $RateDef_t$), a volatility variable ($RatVol_t$), and the rating quality measure (AR_t) for both banks. As before, asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

Table 13 shows the results from multivariate analyses of the effect that the business cycle might have on the ratings quality measure AR *before* removing variables suspected of multicollinearity.

References

Altman, E. & Rijken, H. (2006), 'A point-in-time perspective on through-the-cycle ratings', *Financial Analysts Journal* **62**(1), 54–70.

Amato, J. D. & Furfine, C. H. (2004), 'Are credit ratings procyclical?', *Journal of Banking & Finance* **28**(11), 2641 – 2677. Recent Research on Credit Ratings.
URL: *http://www.sciencedirect.com/science/article/pii/S0378426604001049*

Bar-Isaac, H. & Shapiro, J. (2013), 'Ratings quality over the business cycle', *Journal of Financial Economics* **108**(1), 62–78.

Bernanke, B., Gertler, M. & Gilchrist, S. (1999), 'The financial accelerator in a quantitative business cycle framework', *Handbook of Macroeconomics* **1**, 1341–1393.

Bolton, P., Freixas, X. & Shapiro, J. (2012), 'The credit ratings game', *The Journal of Finance* **67**, 85 – 111.

Calverley, J. (2002), *The Investor's Guide to Economic Fundamentals*, Wiley.

Cantor, R. & Mann, C. (2003), 'Measuring the performance of corporate bond ratings', *Social Science Research Network (SSRN)* .
URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract$_i$d* = 996025

Cantor, R. & Mann, C. (2006), 'Analyzing the tradeoff between ratings accuracy and stability', *Social Science Research Network (SSRN)* .
URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract$_i$d* = 996019

Carvalho, P., Laux, P. & Pereira, J. (2014), 'The stability and accuracy of credit ratings', *Social Science Research Network (SSRN)* .
URL: *https://dx.doi.org/10.2139/ssrn.2504972*

Estrella, A. & Mishkin, F. S. (1998), 'Predicting u.s. recessions: Financial variables as leading indicators', *The Review of Economics and Statistics* **80**(1), 45–61. **URL:** *http://www.jstor.org/stable/2646728*

Hodrick, R. J. & Prescott, E. C. (1997), 'Postwar u.s. business cycles: An empirical investigation', *Journal of Money, Credit and Banking* **29**(1), 1–16. **URL:** *http://www.jstor.org/stable/2953682*

Isil, E., Brandon, J., Woojin, K. & Michael S., W. (2012), 'Macroeconomic conditions and capital raising', *The Review of Financial Studies* **25**(2), 341–376. **URL:** *http://www.jstor.org/stable/41407833*

J.Lobo, G., Paugam, L., Stolowy, H. & Astolfi, P. (2017), 'The effect of business and financial market cycles on credit ratings: Evidence from the last two decades', *Abacus* **53**(1), 59–93. **URL:** *https://ssrn.com/abstract=2929313*

Kruge, M. & Tysnes, H. (2011), 'Et makroøkonomisk perspektiv p˚a aksjemarkedet i norge og usa', https://openaccess.nhh.no/nhh-xmlui/handle/11250/168943. Master Thesis, NHH Norwegian School of Economics.
URL: *https://openaccess.nhh.no/nhh-xmlui/handle/11250/168943*

Nickell, P., Perraudin, W. & Varotto, S. (2000), 'Stability of rating transitions', *Journal of Banking & Finance* **24**, 203–227.

Rawlings, J., Pantula, S., Dickey, D. & MyiLibrary (1998), *Applied Regression Analysis: A Research Tool*, Vol. 41.

Schroeter, U. (2013), Chapter 35 - credit ratings and credit rating agencies, *in* G. Caprio, D. W. Arner, T. Beck, C. W. Calomiris, L. Neal & N. Veron, eds, 'Handbook of Key Global Financial Markets, Institutions, and Infrastructure', Academic Press, San Diego, pp. 379 – 390. **URL:** *http://www.sciencedirect.com/science/article/pii/B9780123978738000384*

Solheim, H. (2008), 'Effects of higher oil prices on the norwegian economy', *Economic Commentaries - Norges Bank Monetary Policy* .
URL: *https://norges-bank.brage.unit.no/norges-bank-xmlui/handle/11250/2558947*

SSB (2020*a*), 'Banks and mortgage companies', https://www.ssb.no/en/statbank/table/07880.

SSB (2020*b*), 'Real gdp growth', https://www.ssb.no/en/nasjonalregnskap-og-konjunkturer/statistikker/knr.

Trustee, N. (2018), 'Norwegian bond market report'.

Wong, S. S.-L., Puah, C.-H., Mansor, S. A. & Liew, V. K.-S. (2016), 'Measuring Business Cycle Fluctuations: An Alternative Precursor To Economic Crises', *Economic Computation And Economic Cybernetics Studies And Research* **50**(4), 235–248.
URL: *https://ideas.repec.org/a/cys/ecocyb/v50y2016i4p235-248.html*