

University of Groningen

Credit ratings, private information, and bank monitoring ability

Nakamura, Leonard I.; Roszbach, Kasper

Published in:
Journal of Financial Intermediation

DOI:
[10.1016/j.jfi.2017.11.001](https://doi.org/10.1016/j.jfi.2017.11.001)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Nakamura, L. I., & Roszbach, K. (2018). Credit ratings, private information, and bank monitoring ability. *Journal of Financial Intermediation*, 36, 58-73. <https://doi.org/10.1016/j.jfi.2017.11.001>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Credit ratings, private information, and bank monitoring ability[☆]

Leonard I. Nakamura^a, Kasper Roszbach^{*,b,c}

^a Research Department, Federal Reserve Bank of Philadelphia, USA

^b Research Department, Norges Bank, Oslo, Norway

^c Faculty of Economics and Business, University of Groningen, The Netherlands



ARTICLE INFO

Keywords:

Monitoring
Banks
Credit bureau
Private information
Public information
Ratings
Regulation
Supervision
Overconfidence

JEL classification:

D82
G18
G21
G24
G32
G33

ABSTRACT

In this paper, we use credit rating data from two large Swedish banks to elicit evidence on banks' loan monitoring ability. For these banks, our tests reveal that banks' internal credit ratings indeed include valuable private information from monitoring, as theory suggests. Banks' private information increases with the size of loans.

However, our tests also reveal that publicly available information from a credit bureau is not efficiently impounded in the bank ratings: credit bureau ratings predict future movements in bank ratings and improve forecasts of both bankruptcy and loan default.

The inefficiency of bank ratings is greater for smaller loans. We investigate possible explanations for these findings. Our results are consistent with bank loan officers placing too much weight on their private information, a form of overconfidence. Risk analyses of the loan portfolios in our data could thus be improved by combining the bank credit ratings with public credit bureau ratings.

The methods we use represent a new basket of straightforward techniques that enable both financial institutions and regulators to assess the performance of credit rating systems.

1. Introduction

How can bank managers, investors, bank regulators, and other stakeholders know whether a bank is a good monitor? This question has become more important since the 2008–2009 financial crisis, during which a large number of banks around the world proved insufficiently attentive to risks within their portfolios. In this paper, we develop and test a method for quantifying the ability of a bank to monitor its loans. This method also provides the user with a test of whether banks collect private information.

If banks collect private information about the borrowers they monitor, as economic theory tells us, in addition to public information

from a credit bureau, and if credit ratings summarize the information included in them, then bank credit ratings should be able to forecast future changes in credit bureau ratings. To test this, we exploit a data set that contains both internal bank credit ratings and external credit bureau ratings of corporate borrowers. In this paper we present strong evidence that the banks in our data set do indeed have private information. This private information is greatest for large loans.

At the same time, if bank credit ratings summarize all public and private information included in them, credit bureau ratings should not be able to predict changes in bank ratings. We present evidence, however, that credit bureau ratings do predict bank ratings. This may be because of bank ratings' coarseness, for example, or because soft bank

[☆] We are indebted to Elif Sen, Fuyuo Nagayama, Gustav Alfelt, and Ethan Haswell for providing outstanding research assistance and grateful for comments from three anonymous referees, the editor, Bo Becker, Sreedhar Bharath, Martin Brown, Mikael Carlsson, Sonja Daltung, Evren Damar, Hans Degryse, Mark Flannery, Mark Flood, Tor Jacobson, Elizabeth Kiser, William Lang, Steven Ongena, Jose-Luis Peydro, Harvey Rosenblum, Frank Schorfheide, Norman Schürhoff, Daniele Terlizzate and seminar participants at the University of Zurich, Federal Reserve Bank of Philadelphia, Riksbank, Finansinspektionen, Svenska Handelsbanken, the National Bank of Serbia, the 2008 EEA annual meeting, the Probanker symposium in Maastricht, the Tor Vergata Conference on Banking and Finance, the Federal Reserve System Committee Meeting on Banking, the 2009 ASSA meetings, the 2009 FMA, the 3rd Swiss Winter Conference on Financial Intermediation, the 2010 Chicago Bank Structure conference, the CEPR-EIEF conference on Transparency, Disclosure and Market Disciplines in Banking Regulation, the 2011 Econometric Society European meeting, and the Banque de France Panel Data conference 2012. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of Norges Bank, the Executive Board of Sveriges Riksbank, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System. Earlier versions of this paper were titled "Credit Ratings and Bank Monitoring Ability." The working paper version is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers and www.ssrn.com

* Corresponding author at: Research Department, Norges Bank, Oslo, Norway.

E-mail addresses: leonard.nakamura@phil.frb.org (L.I. Nakamura), kasper.roszbach@norges-bank.no (K. Roszbach).

information is inefficiently impounded in the hard (credit bureau) information. We present tentative evidence that suggests that bank loan officers overestimate the precision of their private information - that is, they are overconfident in their information. This inefficiency is greater for small loans.

Diamond (1984) and Fama (1985) first put forth the hypothesis that banks were special relative to alternative lenders: Investors delegate the monitoring of borrowers to financial intermediaries because the latter are more efficient. Then, provided banks are sufficiently large and diversified, lending through such intermediaries dominates direct lending by investors. Empirical research in this area has been extensive. Lummer and McConnell (1989) and Mester et al. (2007) describe in detail how banks' monitoring activities, by using transaction account information that provides ongoing data on borrowers' activities, make these intermediaries better monitors of loans. Another strand of literature has studied which conditions may weaken banks' or other investors' monitoring efforts. Agarwal and Hauswald (2010) study the effects of distance on the acquisition and use of private information. Recent work has also shown that screening and monitoring quality by financial intermediaries dropped substantially in the wake of the recent financial crisis (Keys et al., 2009). However, the general notion that financial intermediaries are superior monitors relative to, for example, public alternatives and other investors, remains empirically unchallenged. In particular, the informational superiority of bank credit ratings over public alternatives has not been demonstrated empirically.

The ability of a bank to collect private information and thereby produce a superior judgment of borrowers' expected performance is of relevance not only for regulators and banks but potentially also for the industrial organization of borrowers and for business cycle theory. Dell'Ariccia and Marquez (2004), for example, have pointed out that informational asymmetries among lenders affect banks' ability to extract monopolistic rents by charging high interest rates. As a result, banks finance borrowers of relatively lower quality in markets characterized by greater information asymmetries. When forced to curtail lending, they reallocate their loan portfolios toward more creditworthy, more captured borrowers. Povel et al. (2007) investigate the relation between the cost of monitoring and reporting fraud incentives for companies over the business cycle. Their work has implications for how carefully financial institutions should scrutinize firms in which they invest and for the gains from more publicly available information. In addition, a now extensive literature has described the advantages that financial systems derive from the existence of credit bureaus that aggregate public credit information (e.g., Pagano and Japelli (1993)).

The focus of this paper is on proposing a new basket of straightforward techniques that enables both financial institutions and regulators to assess the performance of credit rating systems. We present a new test that emphasizes the forecasting power of informationally superior estimates of creditworthiness. We do so by carrying out quantitative tests of the relative information of banks and credit bureaus as revealed by their credit ratings.¹ In our theoretical model, we have two monitors: a private monitor, i.e., the bank, and a public monitor, i.e., the credit bureau. Both receive noisy signals of the borrower's creditworthiness. The public monitor receives a public signal, while the private monitor receives both the public signal and a private signal. We think of creditworthiness as being a monotonic transform of the probability of default and model it as a variate that follows a random walk with normal disturbances.² Each monitor processes its noisy signals to

make an optimal estimate of the borrower's creditworthiness using a Kalman filter. The output from this estimation, a continuous processed signal, is then reported in a coarsened form as a discrete categorical rating. A consequence of this coarsening is that some of the information in the continuous signal is lost.³

A closely related paper by Agarwal and Hauswald (2010) also makes use of public and private credit scores. It makes an important contribution to the literature on the role of distance and information in banking relationships by providing direct evidence on the location-specific nature of soft information for loan granting and pricing. Supportive of earlier work, they find that distance erodes a lender's ability to collect private information. While Agarwal and Hauswald's interest lies particularly in the impact of distance on the capacity of banks to collect private information, we complete their work and focus on the efficiency and quality of banks' monitoring. We study the relative value of public information, the extent to which it is optimally incorporated in banks' internal credit ratings, and possible explanations for the (in) efficiency of bank monitoring. We therefore concentrate on assessing whether the bank credit ratings are sufficient statistics for forecasting default or whether there is information in the public credit ratings that has not been impounded in the bank ratings.⁴ We also investigate whether credit ratings are able to forecast defaults, and we test the ability of public and private ratings to forecast default using Cox proportional hazard regressions; in particular, we ask if the public credit ratings add information to the bank credit ratings in forecasting default.

The technique we use here is related to the methodology in Berger et al. (2000), who use vector autoregressions and Granger-causality to compare market and supervisory assessments of bank performance. In particular, they examine bank supervisors' assessments of banks to test the relative information of supervisors and rating agencies. In this paper, we go further and imbed our tests in an explicit model of information updating. As a consequence, we obtain tighter tests that are more explicit about the sources of apparent violations of forecasting theory.⁵

When we apply this technique to a data set of matched bank and credit bureau data, we find that the ratings of both banks do forecast movements in the credit bureau rating. We take this to be evidence that each bank has some private information. However, we also provide evidence that credit bureau ratings can forecast the bank ratings and thus that bank ratings are inefficient measures of borrowers' creditworthiness. This finding can be interpreted in two ways: Either the banks fail to incorporate publicly available information optimally or information is lost by the banks in the process of setting their ratings. When we look into the causes of these results, we find that the occurrence of staggered updating of information by either the credit bureau or the banks does not account for them. We also present evidence that neither the discretization nor the coarsening of the credit bureau rating grades can explain our findings. We have strong prima facie evidence that at least one of the two banks has inefficient ratings, although we cannot rule out that the discretization of the bank ratings may be responsible for this apparent inefficiency of the information aggregation by the banks. The inefficiency we find is greater for small loans.

To assess the fundamental quality of our ratings, we also evaluate

³ To our knowledge, there is not yet any formalized rationale for why this coarsening takes place. One common rationale for coarsening is that ratings changes may require actions - for example, some investors may be required to divest bonds below investment grade. However, the need for action can also be satisfied by continuous ratings with cutoff points.

⁴ We do not investigate at length if credit ratings are indeed able to forecast defaults since there is already an extensive body of work on bond and other credit ratings that, for example, tests the value of bond ratings relative to other financial data in forecasting defaults, interest rate spreads, and portfolio governance. Cantor (2004) and Krahn and Weber (2001) contain a summary of and references to recent research in this area.

⁵ Claessens and Embrechts (2003) assess the consistency between bank internal and external sovereign ratings. They find both are driven by similar factors and underestimate "event risks."

¹ Grunert et al. (2005) present information on nonfinancial factors in internal credit ratings, which suggest that judgmental factors are valuable in bank credit ratings, but acknowledge that such information may be obtained by public monitors such as bond rating agencies.

² Löffler (2004) and Altman and Rijken (2004) argue that credit ratings may have a more complex objective than summarizing default risk. In our case, we know that the sole objective of the bank and credit bureau ratings is to predict counterparty default risk. We will later return to the exact definition of a default.

their accuracy in predicting fundamental events like bankruptcy and loan default. We find that adding the bank rating or the credit bureau rating increased predictive accuracy in a regression explaining loan default or bankruptcy.⁶ This finding reinforces our conclusion that the bank ratings contain some private information but that they are not sufficient statistics for their borrowers' creditworthiness. In other words, we find further evidence of an inefficiency in banks' aggregation of public and private information.

The information inefficiencies we identify can potentially have three different types of explanations: factors related to the rating process, characteristics of the bank, and characteristics of the borrower. Our results indicate that adding private information to public information in generating credit ratings may be more difficult than has been generally recognized. This view is consistent with comparable inefficiencies that have been identified in other areas of financial economics. [Chen and Jiang \(2006\)](#) have shown that equity analyst ratings are typically biased because analysts place too much weight on their private information. [Kahneman \(2011\)](#) has argued that overconfidence is a widespread phenomenon in human decision making. [Altman and Rijken \(2004\)](#) and [Cantor \(2004\)](#) demonstrate that bond ratings move too slowly relative to public information, which has been attributed to the raters' desire to smooth ratings on behalf of their clients. [Hertzberg et al. \(2010\)](#) point out that career concerns may cause loan officers' credit ratings to be biased optimistically.

When we extend our benchmark model to explicitly allow for overconfidence, we find some evidence that is consistent with the presence of overconfidence among loan officers. We build upon our benchmark Kalman filter model by allowing for the possibility that the loan officer is overconfident in the information from the private bank signal, modeled by assuming that the loan officer believes the precision of the private signal is greater than it is in fact. Under such conditions, the bank rating overshoots in the direction of the noise in the private signal and then mean-reverts. This results in an ability of the public signal to forecast the bank rating. We show that the bank rating empirically mean-reverts and that this mean reversion is greater when the public signal does a better job of forecasting the bank rating. This evidence and the lack of competing explanations do not clinch the case that these errors are in fact due to overconfidence, but it is consistent with this explanation. Our findings imply that it is not optimal for either the banks' risk managers or for their regulator to accept the bank's own private credit ratings as the single measure by which to evaluate portfolio credit risk. Instead, it would be beneficial to incorporate more information into a risk review. In particular, credit bureau ratings could be used to improve overall portfolio risk evaluation.

The remainder of this paper is organized as follows. In [Section 2](#), we set forth the theory, develop simulations to more closely mimic the underlying rating process, and enunciate our hypotheses. In [Section 3](#), we describe the data we use to test the theory. [Section 4](#) contains all the empirical analysis, including a series of robustness tests that seek to account for the possibility that credit ratings may not be linear in risk. [Section 4.4](#) discusses if overconfidence by loan officers can explain some of our results. [Section 5](#) concludes.

2. A model of information processing

A well-known theory of banking is that banks possess private information about the creditworthiness of borrowers ([Diamond, 1984](#)). This information may, for example, be derived from the transaction accounts of borrowers ([Mester et al., 2007](#)), which provide a bank lender with uniquely fresh information about the activities of its borrowers. If this theory is true, it follows that banks are uniquely suited to measuring the risks of their borrowers. Based on this line of reasoning,

⁶ Bank credit ratings' predictive power is slightly lower for small borrowers, while credit bureau ratings have less power for large borrowers.

bank examiners have been encouraged to use banks' internal credit ratings as the best available measure of the risk in the bank loan portfolio. In the language of statistical theory, these credit ratings are taken to be sufficient statistics for determining the creditworthiness of loans.

In this section, we discuss a simple theory of signal extraction that describes how producers of credit ratings optimally process different signals of a borrower's creditworthiness. The theory produces a number of testable implications for the relationship between ratings based on publicly available information and ratings based on both publicly and privately obtained information. In [Section 2.1](#), we discuss a simple theoretical model whose details are available in [Nakamura and Roszbach \(2013\)](#). [Section 2.2](#) contains a description of the testable hypotheses implied by the theoretical model.

2.1. Model

In our signal extraction model, we make three important assumptions. First, we postulate that bank credit ratings are measures of borrowers' creditworthiness, i.e., probability of default. Second, we assume that the creditworthiness of a borrower is unidimensional. Our third assumption is that the aim of bank and credit bureau ratings is to measure the objective underlying risk of default.

By means of our first assumption, we exclude cases where ratings are loan-specific. The banks and the credit bureau that provided data for our empirical analysis employ borrower ratings, not loan-specific ratings. The second assumption is a common one in credit risk analysis and implies that credit ratings, for example, do not aim at predicting the bank's potential loss experience once a borrower defaults (loss given default, or LGD). In nearly all models of default behavior, this has been a starting point because there are, to our knowledge, no formalized theories of loss experience. By the same assumption, we also exclude cases where ratings reflect not only risk but also potential profitability. The last assumption is important because different definitions of a default exist, both within the banking industry and between banks and credit bureaus. A reasonable justification for this assumption is that banks use the ratings of credit bureaus as acceptable measures of borrowers' probability of default (PD) and bank regulators accept them as such. Given these three assumptions and provided updating occurs at an appropriate frequency, we can then think of a bank's credit ratings as intended to capture the riskiness of its loan portfolio at any moment in time.

The theory we utilize models the borrower's creditworthiness as being a random walk.⁷ Private and public information are noisy signals of the underlying borrower creditworthiness. As in [Muth \(1960\)](#), all monitors update their optimal (Kalman filter) estimate of the unobserved state using the information available to them, and they incorporate their estimate into a current credit rating for the borrower.

In the theoretical model we utilize, the bank, unlike the credit bureau, not only observes the same public signal as the credit bureau but also gets a noisy, private signal of borrowers' actual creditworthiness. Other information about the borrower, including its relationships with other creditors, is obtained by a credit bureau. The bank obtains the credit bureau information costlessly and fully incorporates this information, in addition to the private information the bank receives, into its private rating of the borrower. It can be shown that, if a bank's credit ratings capture risk optimally given the private information available to the bank, those ratings should forecast movements in the public ratings of a credit bureau. On the other hand, the credit bureau ratings should not forecast movements in the bank's ratings. When the unobserved state, i.e., actual creditworthiness, follows a random walk with noise

⁷ A full derivation is in [Nakamura and Roszbach \(2013\)](#). There and in an on-line appendix we also provide additional tables on variable distributions and on the robustness checks noted in this article

and the signal of creditworthiness that a monitor receives itself is noisy too, we arrive at this result by applying the Kalman filter to obtain Muth's formula on exponentially weighted lags of past signals. This formula can be characterized by a period-by-period updating coefficient, d .

We shall call the filtered signals credit ratings. On the one hand, it is obvious that the public monitor's credit rating will not forecast the bank's credit rating. On the other hand, the bank's credit rating will forecast the public monitor's credit rating for two reasons. One is that the bank has a better fix on the true creditworthiness because it has private information that the credit bureau does not have. The other reason is more subtle: The bank incorporates the credit bureau signal more rapidly into its rating than does the credit bureau itself. This happens because the updating speed, d , depends on the relative precision in the total information of the bank, which, given that the bank has private information as well as the public information, is greater than for the credit bureau. That is, the bank is not simply updating with the credit bureau rating but is actually incorporating the information in the credit bureau signal faster than the credit bureau itself does.

If we were to translate this updating behavior into a regression model that aims to explain how credit ratings are revised using both bank ratings and credit bureau ratings, then the resulting fundamental regression equations would be:

$$y_{ct|t} = \alpha_{10} + \alpha_{11}y_{ct|t-1} + \alpha_{12}y_{bt|t-1} + e_{1t} \quad (1)$$

$$y_{bt|t} = \alpha_{20} + \alpha_{21}y_{ct|t-1} + \alpha_{22}y_{bt|t-1} + e_{2t} \quad (2)$$

Considering Eq. (1), we expect the constant term to be zero. Because of the private information encompassed by bank ratings, the sum of the coefficients of $\alpha_{11} + \alpha_{12}$ should be unity and $\alpha_{12} \geq 0$. For Eq. (2), we expect that the credit bureau's rating will not be able to forecast the bank rating, since the information underlying it is already embedded in the bank rating so that $\alpha_{21} = 0$. Because the underlying information follows a random walk, the coefficient on the lagged bank rating should be unity and the constant term should be zero: The forecasts are expected to be martingales.

In Section 4, we will test two weaker conditions for the optimality of credit ratings, given that credit ratings are not continuous, as in our theory, but coarsened: (i) that the bank's credit rating for borrowers forecasts the public monitor's credit rating, but (ii) that the public monitor's credit rating does not forecast the bank's credit rating. These are the standard Granger causality conditions, and we could test them using VARs with one lag on each equation, as in Eqs. (1) and (2). If we find that the bank's credit ratings are forecastable by the public monitor, then this will constitute prima facie evidence that the bank credit ratings are not sufficient statistics for determining the creditworthiness of the bank portfolio. It will also mean that an optimal measure of the risk in the bank portfolio should include measures of borrower quality from outside the bank's credit rating system.

When we test the above conditions in Section 4, we will also want some quantitative support for interpreting the goodness of fit of our estimated equations. We have, therefore, derived a general result on the maximum attainable improvement in R^2 in regression Eqs. (1) and (2) from the inclusion of the private information. To take Eq. (1), the change in the credit bureau's rating can be decomposed into contributions from the change to the underlying creditworthiness of the borrower, changes in the credit bureau's signal, and the error in the credit bureau's rating at time $t - 1$. The first two parts are clearly unforecastable noise terms. So the only part of the change in the credit bureau's rating that is potentially forecastable is the part due to the error in the past estimate, the previous rating. It can be shown that under the model assumptions, this reaches at most 0.25. In other words, the maximum reduction in the sum of squared errors one can obtain based on knowledge at $t - 1$ is 0.25, where the period is the time over which ratings normally change, that is, one year. This result will be

important in Section 4 when we need to evaluate the fit of our regressions.

2.2. Hypotheses

In this section, we summarize the implications that the simple model presented in Section 2.1 has for the relationship between public (credit bureau) and private (bank) borrower ratings. In Section 4, we will test these hypotheses.

The first question is whether banks possess private information. If banks do not possess private information not contained in the credit bureau rating, there is no rationale for separate bank ratings. The second question, then, is given that banks possess private information and low-cost access to the credit bureau's ratings, based on the model, we expect that the credit bureau's rating will not be able to forecast the bank rating because the information contained in credit bureau ratings is already embedded in the bank rating; in terms of Eq. (2), $\alpha_{21} = 0$. Because the underlying information follows a random walk, the coefficient on the lagged bank rating should be unity and the constant term should be zero. Hence, under rational expectations, forecasts of bank credit ratings should be martingales. Of course, conditioned on information outside the information set from which the forecast has been made, changes in the rating may no longer be unforecastable. As a consequence, one test of whether a particular forecast of default is based on a larger information set than another (or a refinement of the information set) is that it will be able to forecast the movements in that other forecast. From this argument follows:

Hypothesis 1. A bank's internal credit rating should contribute to forecasting changes in a public credit rating of the same borrower.

If a bank's internal credit ratings *do* forecast changes in public credit ratings *and* if the bank's future ratings are not forecastable by the public credit rating, it would appear likely that the bank has strictly superior information. We would then have no evidence against the hypothesis that the bank has private information it uses rationally. Moreover, we would have strong grounds for the belief that a bank supervisor should use the bank credit ratings in measuring the risk in the bank's loan portfolio.

Hypothesis 2. Changes in a bank's credit ratings should not be forecastable.

If the credit bureau ratings forecast the bank's future credit ratings, not only do we know that the bank's ratings are not sufficient statistics, but the proof is constructive: It tells us how to improve on the bank's ratings as a measure of risk.

Corollary 1. *If changes in a bank's internal credit ratings are forecastable, then the variables in the equation that predicts the change in the bank's credit ratings will improve estimates of the riskiness of bank borrowers.*

Corollary 1 also means that if bank credit ratings are forecastable, then an optimal measure of the risk in the bank portfolio should include measures of borrower quality from outside the bank's credit rating system.

Another way to think about this is the following. If agent I's forecast of some future event is superior to that of agent II, then, statistically speaking, this means that I will be accurate more often than II. Put another way, the future offers fewer surprises for I than for II. If the future event is more than one period away, and information is revealed in the meantime it is more likely that the new information will confirm I's view of the future than it will II's. The forecast of II is then more likely to approach that of I, assuming it is rational, than that I's forecast will move toward II's. As a consequence, I's current forecast will tend to forecast II's future forecast, taking into consideration II's current forecast. Even stronger, if I's forecast is optimal and I knows II's forecast, then II's forecast cannot be better than I's and will not forecast I's future forecast.

Table 1

Descriptive statistics on loans outstanding. The table contains descriptive statistics on actually utilized credit in Bank A and Bank B. All numbers are averages over four years, i.e., over the period 1997-Q3 to 2000-Q1. Large borrowers have maximum loan sizes of 5 million SEK or greater, small borrowers have maximum loan sizes of 500 thousand SEK or smaller.

	Bank A				Bank B			
	Total	Large	Medium	Small	Total	Large	Medium	Small
Total loan outstandings (billion SEK)	91.7	85.3	5.73	0.664	110	103	7.07	0.845
Mean loan size (million SEK)	4.397	20.8	0.639	0.085	10.4	25.9	1.141	0.204
Number of loans, quarterly average	20851	4103	8954	7794	10586	3979	6192	415

2.3. Testable implications

If we define r_{ct} as the rating of the credit bureau at t and r_{bt} as the rating of the bank at t then, under the assumptions in Section 2.1, Eqs. (1) and (2) translate into the regressions (3) and (5) below that we can estimate.

Hypothesis 1 in a strict sense implies that $\alpha_{1c} + \beta_{1c} = 1$ and $\alpha_{1c} > 0$. This is what we would expect of an optimal bank forecast if it were continuous. Unfortunately, the rating data we have available for testing Hypothesis 1 are discrete instead of continuous. Under these conditions, we can no longer be sure that both parameter restrictions will hold and we will test a weaker version of the hypothesis, namely, that $\alpha_{1c} > 0$. Under this hypothesis, the bank rating does forecast changes in the credit bureau rating and has a significant impact on the residual sum of squares (RSS) in regression equation (3).

$$r_{ct} = \beta_{1c} r_{ct-1} + \alpha_{1c} r_{bt-1} + \varepsilon_{1ct} \quad (3)$$

We will measure the decline in RSS when estimating regression (3) relative regression (4) and compare this with the theoretical maximum of 0.25; according to our baseline model, given that new information is available each period, last period's bank rating can at best forecast the credit bureau rating one-quarter of the time once the credit bureau's lagged rating is taken into account.

$$r_{ct} = \beta_{2c} r_{ct-1} + \varepsilon_{2ct} \quad (4)$$

Because we also wish to test for the marginal informational value of adding a lag of the credit bureau rating, we will as well estimate regression (5)

$$r_{bt} = \alpha_{1b} r_{bt-1} + \beta_{1b} r_{ct-1} + \varepsilon_{1bt} \quad (5)$$

and the simple autoregressive form

$$r_{bt} = \alpha_{2b} r_{bt-1} + \varepsilon_{2bt} \quad (6)$$

Hypothesis 2 in Section 2.2 implies that $\alpha_{1b} = 1$ and $\beta_{1b} = 0$. For the same reasons mentioned above for Hypothesis 1, we will allow for a constant in the empirical equivalent of Eq. (5) and test the weaker hypothesis that $\beta_{1b} = 0$ in Section 4.1.1. Under this hypothesis, the credit bureau rating does not forecast changes in the bank rating and has an insignificant impact on the RSS when moving from regression (5) to regression (6).

3. Data

The primary data sources we use are the credit registries of two of the four major Swedish commercial banks, which we shall call Bank A and Bank B, and the registry of the leading credit bureau in Sweden, Upplysningscentralen AB (UC), which we shall call the credit bureau. The two banks are both universal banks and sufficiently sophisticated that they now follow the Basel 2 Internal Ratings Based approach. However, during the sample period, capital requirements were not yet based on internal ratings. UC is an incorporated company that is jointly owned by the major Swedish banks. Ownership shares are related to bank size. Nonfinancial enterprises and all financial institutions reported data on loan applications, loans made, balances, and loan performance to UC every second month, with balances registered on the

last day of the month.⁸ Credit abuse and payment remarks were and are reported continuously. Credit abuse cannot be reported until a firm or household is at least 90 days behind on the agreed-upon repayment schedule. UC produces credit ratings for almost all Swedish businesses. The information on *corporate* credit balances was and is not incorporated in the credit ratings. The credit bureau ratings are not solicited, and the bureau's revenues from its rating activities come through the sale of various types of credit reports.

The data set covers the period starting in 1997-Q1, ending in 2000-Q1 for Bank A and in 2000-Q2 for Bank B. Because of a change in the credit bureau rating system, we will delete the first two quarters of the bank data sets. This gives us one to 11 quarterly observations for, on average, roughly 15,000 borrowers in Bank A and one to 12 quarterly observations on 8000 borrowers in Bank B. Borrowers, incorporated businesses or aktiebolag, have at least the legally required minimum of Swedish Kronor (SEK) 100,000 (approximately US \$12,500 at that time) in equity. Many of them, particularly for Bank A, are very small. Roughly 37 percent of Bank A's borrowers are small borrowers, defined as having maximum borrowing of less than SEK 500,000 (about US \$62,500 in the time period examined), adjusted for inflation from the first quarter of 1997. About 4 percent of Bank B's borrowers have borrowings this small. Although Bank B has roughly half as many borrowers, its number of large borrowers is nearly as large as that for Bank A, with large borrowers defined as having more than SEK 5 million in maximum borrowing (about US \$625,000). As Table 1 shows, small and medium-sized borrowers represent between 60 and 80 percent of all borrowers, but only a small proportion of the total loan portfolio of either lender.⁹

Both banks maintain an internal credit rating scheme: Bank A assigns each business customer to one of 15 credit rating grades, while Bank B uses seven classes. Higher numbers imply worse ratings, and rating grades 15 and 7 in the respective systems represent defaulted customers. Both banks employ the same definition of a default, namely that (i) the principal or interest payments are 60 days overdue and (ii) a bank official has to make a judgment and reach the conclusion that any such payment is unlikely to occur in the future. Both the credit bureau's and the banks' ratings are "borrower" ratings, not loan-specific ratings.

The credit bureau has five rating classes in addition to a default rating, and a numerically higher rating again implies worse creditworthiness. A firm is given a default status when any of the following events occurs: The firm is declared legally bankrupt, has suspended payments, has negotiated a debt composition settlement, is undergoing a reconstruction, or is distraint without assets. To keep track of these events, the credit bureau collects event data from Tingsrätten (District Court), Bolagsverket (the Swedish Companies Registration Office or SCRO), and Kronofogdemyndigheten (the Swedish Enforcement Authority). Once any of the above distress events occurs, the firm in question is at once registered as defaulted. This event is observed by us on the last day of that particular quarter. In the following quarter, we

⁸ To the best of our knowledge, the reporting frequency changed to monthly shortly after our sample period ended. We have, unfortunately, only been able to obtain indicative information from the credit bureau about this change of reporting frequency.

⁹ A more complete description of the bank and credit bureau data can be found in Jacobson et al. (2006, 2013).

Table 2

Empirical distribution of bank ratings for Bank A and Bank B borrowers. All numbers are over the period 1997-Q3 to 2000-Q1. Higher ratings imply worse creditworthiness. Observations are defined as quarter-borrower pairs. Defaults are defined by each bank and are excluded from totals to avoid attrition effects on the distribution.

Rating Bank A	Nobs	Percent	Rating Bank B	Nobs	Percent
1	3382	1.62	1	57	0.05
2	50,826	24.38	2	2835	2.43
3	109,655	52.59	3	29,764	25.56
4	30,003	14.39	4	70,987	60.96
5	9363	4.49	5	11,574	9.94
6	3589	1.72	6	1228	1.05
7	1696	0.81			
Default	873			708	
Total	208,514	100.00		116,445	99.99
Mean rating	3.04			3.82	
Std. Deviation	0.96			0.68	

then let the firm exit our data set. If more than one of these distress events is observed for a specific firm over our sample period, we assume the firm in question defaulted in the quarter during which the first of these events took place. For about 45 percent of the defaulting firms, one of the other default-triggering events occurs simultaneously (i.e., during the same quarter).¹⁰

Bank A's credit ratings are not single peaked in default frequency and worse yet, when regressed against the credit bureau ratings are not monotonic. We have therefore reconfigured Bank A's credit ratings into an eight-grade rating system that is monotonic in default frequency and with respect to the credit bureau ratings.¹¹ In most of our analysis, we will exclude observations where a counterparty has defaulted according to either the credit bureau or the bank, because the default rating reflects actual behavior rather than a bank's estimate of creditworthiness. The only exception will be regressions where a bank default dummy is our dependent variable. In those regressions we will omit observations where borrowers had a default rating grade at the credit bureau, i.e., they either filed for bankruptcy or were declared bankrupt. Credit ratings need to be updated by loan officers at least once every 12 months. Table 2 shows that the credit ratings for both lenders are highly concentrated, just as for U.S. large-bank credit ratings. Bank A has about 77 percent of its ratings in its two largest rating categories (this concentration is also very high in the original 15 rating system), while Bank B has roughly 86 percent of its ratings in its two largest rating categories.

The distribution of credit bureau ratings is shown in Table 3. While Bank A and Bank B's borrowers are concentrated in the center of their distributions, the credit bureau's ratings for these same borrowers are concentrated in the top rating. The two sets of ratings thus appear to be scaled quite differently.

The ratings of the credit bureau are available to the bank loan officers at near zero cost through an online computer system. That is, at the time that a loan officer establishes the credit rating, the latest available rating from the credit bureau and a set of background variables from the credit bureau are part of the loan officer's information set.

4. Empirical results

In this section, we present the results from our empirical analysis.

¹⁰ About 5 percent of the firms that experience a credit bureau default reemerge from their default status. We do not include these reemerged companies in our data. Nearly all reemerging companies default a second and final time, mostly in sample and some out of sample. The vast majority of all terminal credit bureau defaults concern legal bankruptcy declarations. For the firms that reemerge after a default, the first default involves a legal bankruptcy in less than half a percent of all cases and "distraint, no assets" in 98 percent. At their second default, these percentages are reversed.

¹¹ Details and the complete rationale may be found in Nakamura and Roszbach (2013).

Table 3

Empirical distribution of credit bureau ratings for Bank A and Bank B borrowers. All numbers are over the period 1997-Q3 to 2000-Q1. Higher ratings imply worse creditworthiness. Observations are defined as quarter-borrower pairs. Defaults are defined by the credit bureau as and are excluded from totals to avoid attrition effects on the distribution.

Rating	Bank A Borrowers		Bank B Borrowers	
	Nobs	Percent	Nobs	Percent
1	90,335	43.32	38,413	32.99
2	55,120	26.43	33,816	29.04
3	43,160	20.70	31,714	27.24
4	12,353	5.92	7770	6.67
5	7546	3.62	4732	4.06
Default	508		365	
Total	208,514	100.00	116,445	100.00
Mean rating	2.00		2.20	
Std. Deviation	1.10		1.09	

Some of the supporting robustness tests we refer to in the remainder of the paper are available in the working paper version, which contains an extensive appendix. Hereafter, we will explicitly refer to table numbers for results presented in the paper, while auxiliary results without such a reference can be assumed to be available only in the online working paper version.

We will make the hypotheses in Section 2.2 operational by testing the informational content of both a bank's internal credit rating and the external credit bureau rating. First, in Section 4.1, we present the results of a basic test of hypotheses 1 and 2 using OLS regression on Eqs. (3) to (6). We first run OLS regressions for the credit bureau ratings on their lagged values as in Eq. (4) and then add a bank's lagged credit rating as in Eq. (3). We check the linearity and monotonicity properties of the rating systems by using dummy variables for the ratings. Conversely, we also present the results of regressions for each bank's credit rating on its lagged values as in Eq. (6). In step two, we then add the credit bureau's lagged ratings as in Eq. (5). We verify if our results are sensitive to the selection of firms' size or to lag length because theory does not give us any guidance regarding the time advantage that the credit bureau or the banks have vis-a-vis each other. After we have established that bank ratings do in fact contain private information but also inefficiently impound public information so that public ratings can forecast them, we check in Section 4.1.3 that the credit bureau rating and the bank ratings not only have statistically significant coefficients in the regressions but also shift the mean and distribution of expected default or bankruptcy probabilities in a measurable way.

In Section 4.2, we verify if the credit bureau rating and bank ratings not only predict each other but also predict intrinsically important events. We do so by running Cox regression of borrower default and bankruptcy risk and find by means of exclusion tests that both ratings are needed to produce efficient forecasts of loan defaults and bankruptcies.

In Section 4.3, we investigate if our findings are an artifact of either staggered updating of information sets or of the loss of information due to the discrete nature of ratings. After finding that our results are robust when we take these things into account, we look at overconfidence as an alternative explanation. In Section 4.4, we show that the banks' ratings behave in a way that is consistent with overconfidence, i.e., loan officers placing too much weight on their own information.

4.1. Testing hypotheses 1 and 2

4.1.1. Hypothesis 1: Do bank ratings contain private information?

In this section, we examine whether bank ratings incorporate private information that allows them to predict credit bureau ratings. We find robust evidence that they do.

Our theoretical model argues that in regressions 3 and 5, the

Table 4

OLS regressions with all borrowers, credit bureau, Bank A, and Bank B. Sample period is 1997-Q3 to 2000-Q1; standard errors are robust. Lags are one quarter lags. Standard errors are displayed in parentheses. "Credit bureau rating dummies" and "Bank rating dummies" mean that the discrete rating variable was replaced by a set of dummies corresponding to the respective rating grades.

Explanatory variables	Dependent variable					
	Credit bureau rating			Bank A rating		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.480 (0.00494)	0.076 (0.00868)	0.632 (0.0102)	0.217 (0.00323)	0.550 (0.00816)	0.579 (0.0119)
Lag credit bureau rating	0.885 (0.00111)	0.861 (0.00131)	0.860 (0.00132)		0.0599 (0.00122)	
Lag Bank A rating		0.0612 (0.00141)		0.939 (0.00105)	0.907 (0.00130)	0.904 (0.00135)
Credit bureau rating dummies	No	No	No	No	No	Yes
Bank rating dummies	No	No	Yes	No	No	No
Residual sum of squares	55,575	55,021	55,001	26,540	25,831	25,735
Adj. R ²	0.7784	0.7806	0.7807	0.8610	0.8647	0.8652
Nobs	208,514	208,514	208,514	208,514	208,514	208,514

Explanatory variables	Dependent variable					
	Credit bureau rating			Bank B rating		
	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.449 (0.00593)	0.941 (0.0144)	0.700 (0.0476)	0.162 (0.00444)	0.286 (0.00703)	0.279 (0.00760)
Lag credit bureau rating	0.886 (0.00142)	0.858 (0.00169)	0.857 (0.00170)		0.0191 (0.00072)	
Lag Bank B rating		0.102 (0.00251)		0.960 (0.00116)	0.947 (0.00134)	0.947 (0.00134)
Credit bureau rating dummies	No	No	No	No	No	Yes
Bank rating dummies	No	No	Yes	No	No	No
Residual sum of squares	30,607	30,164	30,147	4981	4941	4939
Adj. R ²	0.7802	0.7834	0.7835	0.9079	0.9087	0.9087
Nobs	116,445	116,445	116,445	116,445	116,445	116,445

forecast residual sum of squares is minimized by the Kalman-filter optimal estimate of the underlying creditworthiness of the borrower. The economic significance of the Bank A and Bank B ratings can therefore be measured as the extent to which the residual sum of squares in the

credit bureau ratings is reduced when we move from Eq. (4) to (3). It can also be measured as the size of the coefficient of the lagged Bank A and B ratings in Eq. (3). Below, we present the results along these two dimensions.

Table 5

Sensitivity of the explanatory power of bank ratings and credit bureau ratings to lag length and borrower size in OLS regressions. Entries in the table reflect the percentage by which the residual sum of squares is reduced when a one-, two-, three-, or four-quarter lag of bank rating A (BRA), bank rating B (BRB) or the credit bureau rating (CBR) is added as an explanatory variable in addition to the lagged dependent variable. The different rows display the reduction in RSS given a fixed lag length for small, medium-sized and large borrowers, as defined in Table 1.

Explanatory variable added	Dependent variable							
	Credit bureau rating				Credit bureau rating			
	BRA(L1)	BRA(L2)	BRA(L3)	BRA(L4)	BRB(L1)	BRB(L2)	BRB(L3)	BRB(L4)
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All borrowers	1.00	1.17	1.33	1.49	1.44	2.00	2.31	2.40
Small borrowers	0.94	1.08	1.27	1.50	1.20	1.63	1.96	1.87
Medium-sized borrowers	1.04	1.20	1.35	1.44	1.39	1.95	2.21	2.31
Large borrowers	1.01	1.26	1.42	1.60	1.51	2.08	2.47	2.54

Explanatory variable added	Dependent variable							
	Bank A rating				Bank B rating			
	CBR(L1)	CBR(L2)	CBR(L3)	CBR(L4)	CBR(L1)	CBR(L2)	CBR(L3)	CBR(L4)
Sample	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
All borrowers	2.67	4.73	6.42	7.92	0.82	1.50	2.01	2.36
Small borrowers	3.00	5.35	7.33	9.20	0.43	0.73	0.96	1.47
Medium-sized borrowers	2.63	4.66	6.35	7.65	0.92	1.68	2.33	2.77
Large borrowers	2.08	3.64	4.82	6.09	0.70	1.30	1.62	1.85

Residual Sum of Squares Measures.

In Table 4, columns (1) and (2), we see that when we include the one-quarter lagged Bank A rating in an OLS regression of the credit bureau rating on its own lag, then the RSS drops from 55,575 (Table 4 column 1) to 55,021 (Table 4 column 2), a reduction of 1.0 percent. If we use bank rating dummies instead, to allow for the possibility that the ratings schemes may not be linearly related, the residual sum of squares shrinks by about the same percentage (Table 4, column 3), indicating that nonlinearities in the rating scheme are quantitatively not important. The percentage reduction in RSS from columns (1) to (2) is shown directly in Table 5, column (1), row 1. The remaining columns in the upper panel of Table 5 present these reductions with lag lengths from one to four quarters, while different rows display the reductions for varying loan sizes, as defined in the previous section. We include varying lag lengths because theory does not give us guidance regarding the time period during which private information of the banks has an edge relative to public information. A longer lag length provides a longer period of time for the public information to register the outcomes predicted by the bank’s private information. We expect the ratings to be increasingly predictive over longer periods, and that is what we observe. As the lag length increases, up to four quarters, the explanatory power rises. For all borrowers, the reduction in RSS over four quarters is about 1.5 percent, and this reduction does not vary a great deal across size of borrowers. The reduction in RSS, as we have said, should be compared with the maximum attainable of 0.25 described in Section 2.1. As such, it represents some 6 percent of the maximum. Bank A thus has private information that helps forecast public information four quarters in the future.

For Bank B’s ability to predict credit bureau ratings, we look at columns 7 and 8 in Table 4, where we see that the RSS falls from 30,607 to 30,164, or 1.4 percent. As for Bank A, replacing the bank rating by a vector of bank rating dummies has little impact on the RSS. Turning to Table 5, columns (5)–(8) we see that as the lag period for the Bank B

ratings rises, the RSS falls by more, and at four quarters, falls by 2.4 percent for all loans, with a similar drop for medium-sized and large borrowers. For small borrowers, the effect is lower; one should keep in mind, however, that Bank B has only few small clients. These reductions in the RSS for medium and large borrowers are nearly 10 percent of the maximum of 0.25. Thus Bank B, even more than Bank A, clearly holds substantial private information not included in the credit bureau rating and thus private information forecasts public information well into the future.

Ordered logit regressions, not shown here, show the same qualitative results. Although ordered logit regressions have the advantage that they allow one to take into account the discrete nature of credit ratings, we will focus on the OLS regressions because they are consistent and less sensitive to distributional assumptions than are the ordered logits. Since we have a large number of observations, consistency seems a more appropriate criterion than efficiency.

Coefficients.

The coefficient on Bank A rating in column (2) of Table 4 is 0.06 and is highly significant. This is further evidence that Bank A ratings possess some information about borrowers not included in the credit bureau ratings. If we turn to Table 6, columns (9) to (12), we see the coefficient on the Bank A rating for varying lag lengths and loan sizes rising to 0.113 over four quarters. A way to interpret this coefficient is to say that a one grade change in the Bank A rating will be associated with a one grade change in the credit bureau about one-ninth of the time, very roughly speaking.

The coefficient of the one-quarter lag Bank B rating in Table 4, column (8) is 0.10 and is highly significant as well. Turning to Table 6, we see in columns (13) to (16) that over four quarters, the Bank B coefficient rises to 0.2; a unit difference in the Bank B rating leads a unit change in the credit bureau rating roughly speaking one-fifth of the time. This is clearly an economically significant effect. Ordered logit regressions not displayed here show that the above findings are robust

Table 6

Sensitivity of regression coefficients to lag length and borrower size in OLS regressions with credit bureau, Bank A and Bank B rating. Entries in the table reflect the coefficient (and robust standard error) on the new explanatory variable when a one-, two-, three-, or four-period lag of bank rating A (BRA), bank rating B (BRB) or the credit bureau rating (CBR) is added as an explanatory variable in addition to the lagged dependent variable.

Explanatory variable added	Dependent variable							
	Bank A rating				Bank B rating			
	CBR(L1)	CBR(L2)	CBR(L3)	CBR(L4)	CBR(L1)	CBR(L2)	CBR(L3)	CBR(L4)
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All borrowers	0.060*** (0.001)	0.108*** (0.001)	0.151*** (0.001)	0.188*** (0.002)	0.019*** (0.001)	0.036*** (0.001)	0.049*** (0.001)	0.061*** (0.001)
Small borrowers	0.065*** (0.001)	0.118*** (0.002)	0.165*** (0.003)	0.206*** (0.003)	0.012*** (0.003)	0.023*** (0.005)	0.033*** (0.006)	0.048*** (0.008)
Medium-sized borrowers	0.059*** (0.001)	0.107*** (0.002)	0.148*** (0.002)	0.183*** (0.003)	0.020*** (0.001)	0.037*** (0.001)	0.052*** (0.002)	0.065*** (0.002)
Large borrowers	0.051*** (0.002)	0.093*** (0.003)	0.129*** (0.003)	0.164*** (0.004)	0.018*** (0.001)	0.034*** (0.002)	0.046*** (0.002)	0.056*** (0.002)

Explanatory variable added	Dependent variable							
	Credit bureau rating				Credit bureau rating			
	BRA(L1)	BRA(L2)	BRA(L3)	BRA(L4)	BRB(L1)	BRB(L2)	BRB(L3)	BRB(L4)
Sample	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
All borrowers	0.061*** (0.001)	0.082*** (0.002)	0.099*** (0.002)	0.113*** (0.003)	0.102*** (0.002)	0.150*** (0.003)	0.180*** (0.004)	0.197*** (0.005)
Small borrowers	0.060*** (0.002)	0.080*** (0.003)	0.098*** (0.004)	0.114*** (0.004)	0.100*** (0.013)	0.145*** (0.019)	0.174*** (0.023)	0.183*** (0.028)
Medium-sized borrowers	0.063*** (0.002)	0.084*** (0.003)	0.100*** (0.003)	0.112*** (0.004)	0.103*** (0.003)	0.152*** (0.004)	0.181*** (0.005)	0.199*** (0.006)
Large borrowers	0.059*** (0.003)	0.082*** (0.004)	0.097*** (0.005)	0.112*** (0.005)	0.099*** (0.004)	0.147*** (0.005)	0.179*** (0.006)	0.196*** (0.007)

to the estimation method used.

Overall, these findings offer strong and robust evidence in support of the hypothesis that the banks in our sample have private information and that their internal ratings predict credit bureau ratings. Generally, banks have more private information as loans become larger and have greater predictive power for longer forecasting horizons up to four quarters. We also note that Bank A ratings appear less informative than Bank B ratings.

4.1.2. Hypothesis 2: Are banks' credit ratings efficient?

In hypothesis 2, we consider whether there is information in the credit bureau ratings that helps predict bank ratings. Since the credit bureau ratings are part of the information that is freely available to loan officers, we would not expect these ratings to forecast changes in bank ratings. As we shall see, however, this is not the case: credit bureau ratings do forecast bank ratings. Given this finding, we then face the question of how we measure the economic significance of the credit bureau ratings – i.e., how large the implied inefficiency of the bank credit ratings is?

Residuals Sum of Squares Measures.

As with the bank ratings discussion under hypothesis 1 in Section 4.1.1, the economic significance of the credit bureau ratings can be measured by the extent to which the residual sum of squares is reduced when we move from Eqs. (6) to (5) and by the size and significance of the coefficient of the credit bureau rating in Eq. (5). Table 4, columns (4) and (5), summarize the regression output for Eqs. (5) and (6) for Bank A. Moving from column (4) to column (5), the RSS is reduced from 26,540 to 25,831, a decrease of 2.7 percent. Since the maximum possible reduction is 25 percent (on the null hypothesis that the bank rating is efficient), we see immediately that this is a large number, representing more than 10 percent of the maximum.¹² Recall that we found in Section 4.1.1 that Bank A's private information reduced the credit bureau RSS by 1.5 percent. Thus, over the entire portfolio, the credit bureau appears to have relatively better information than bank A since its ratings have a proportionally bigger downward impact on the forecast error: Credit bureau ratings forecast Bank A ratings better than vice versa. This is quite surprising. Another interesting observation to be made is that, particularly for the longer lag horizons, credit bureau ratings generally perform notably better than bank ratings when predicting ratings movements for small borrowers while bank ratings have greater predictive power for large borrowers. Notable as well are the results in Table 5, columns (9)–(12), where we see that over a four-quarter horizon, adding a lag of the credit bureau rating reduces the RSS of the Bank A ratings overall by 8 percent, varying from 6 percent for large loans to 9 percent for the smallest loans. Eight percent is nearly one-third of the maximum. Moving to Table 4, column (6), we see that there may be some nonlinearity in ratings, as the RSS falls by another 3 percent when we replace the rating variables by a set of dummies.

By contrast, appending the lag of the credit bureau rating to a regression on the Bank B rating in Table 4 reduces the RSS by only 0.8 percent, substantially less than the reduction in RSS that Bank B achieves in predicting the credit bureau rating at a one quarter lag. In Table 5, we see, however, that when we increase the lag length, the reduction in Bank B ratings achieved by credit bureau ratings rises to 2.4 percent, or roughly 10 percent of the theoretical maximum, about the same reduction achieved by Bank B on credit bureau ratings.

Coefficients.

The coefficient on the credit bureau rating is 0.06 in Table 4, column (5), comparable to the coefficient on the Bank A rating in column (2). However, the standard deviation of the credit bureau rating is 1.10 (Table 3) while that of Bank A's rating is 0.96 (Table 2). So if the data were normalized the credit bureau coefficient would be slightly

higher. Moreover, turning to Table 6 we see that the coefficient on the credit bureau rating at four lags rises to 0.19, substantially larger than the coefficient on the Bank A rating we found. At a four-quarter lag, a one notch rating difference in the credit bureau rating leads a change in the Bank A rating roughly one time in five.

In regressions on Bank B ratings, the coefficient of the credit bureau rating in Table 4, column (11), is approximately 0.02, statistically significant, but much smaller in magnitude than the corresponding coefficient in the regression on Bank A ratings (column (5)). With the standard deviation of the Bank B ratings, 0.68 (Table 2), being smaller than that of the credit bureau rating, 1.09 (Table 3), the coefficient on the lagged Bank B rating in column (11) is meaningfully larger on a standardized basis than that of the credit bureau in column (8).¹³ As we move from column 11 to column 12, we see that allowing for non-linearity of the credit bureau ratings does not improve the forecast ability appreciably.

4.1.3. Both ratings affect predicted default risk

We also provide complementary evidence that supports our assertion that credit bureau and bank ratings each make a substantial contribution to explaining the variable underlying the bank credit ratings and the credit bureau rating, respectively: namely, loan default and bankruptcy. We do so by testing the equality of the area under the ROC curve before and after expanding an autoregressive logit model of default with a second rating.¹⁴ We present these results in Table 7, panel I.¹⁵ A set of chi-square tests of equality of the areas under the receiver operating characteristics (ROC) curve show that the discriminatory accuracy of the model increases significantly, both in-sample and out-of-sample, in a logit default regression when a second rating is added.

Kolmogorov–Smirnov tests, displayed in Panel II of Table 7, show that not only the mean but also the distribution of predicted default probabilities shifts significantly when adding a second rating as an explanatory variable in an autoregressive logit model of default risk. This holds both in-sample and out-of-sample.

Overall, these findings constitute distinct evidence against the hypothesis that bank ratings are fully efficient and cannot be predicted by lagged credit bureau ratings. This finding is highly robust for variations in loan size and across different forecast horizons. Intuitively, the inefficiency is greater for smaller loans, where monitoring costs are arguably relatively high for banks. The results hold all the more for Bank A, which has ratings that are relatively less informative.

4.2. Do credit ratings predict fundamental events?

In the previous section, we found that bank ratings, which contain both public and private information, are only partially able to forecast credit bureau ratings that are produced using publicly available information. Vice versa, we showed that, somewhat surprisingly, credit bureau ratings are able to partially forecast internal bank credit ratings. From a research perspective, an intuitively attractive conclusion to be drawn from these results would be that credit bureau ratings are of

¹³ In additional regressions that are not presented in the paper, we find that the forecasting ability of credit bureau ratings is greatest for the riskiest loans for both Bank A and Bank B credit ratings. For Bank A it is monotonically increasing in risk; for Bank B forecasting ability is approximately equal for low- and medium-risk firms. As in Table 5 and 6, all increases in RSS are bigger when we add the credit bureau rating to an AR(1) for Bank A than for Bank B.

¹⁴ Because no comparable test is available for the Cox model, we instead use a logit model of default for this particular test.

¹⁵ Because the ROC curve and the ROC test can be computed only for discrete dependent variables, we re-estimate the specifications in the upper panel of Table 7 as a logit model. The ROC graph plots the trade-off between the benefits (true positives) and the costs (false positives) of a binary classification system. Note that an increase in the discriminatory accuracy of a model does not map 1:1 into a greater fit of a model. Vice versa, a better fit does not imply improved discriminatory accuracy.

¹² See Section 2.1.

Table 7

Panel I: Do both ratings contribute substantially to predicting in-sample and out-of-sample defaults? This table displays tests of the area under the ROC curve for in-sample and out-of-sample predicted logit probabilities. Logit models of default probabilities are estimated first (1) using only the lag of the dependent variable, then (2) by adding the lag of a second rating. For each model we estimate the area under the ROC curve. Then we test if the expanded model has the same explanatory power as estimated by the area under the ROC curve, against the alternative of the expanded model being better. We implement this test for both the credit bureau rating (CBR) and the bank ratings (BR) as the dependent variable. P-values are one-sided probabilities of the Null hypothesis that the estimated ROC area of the small and the expanded model are of equal size. For the out-of-sample test, the data sample is first split into two equally sized sub-samples but each covering the same time period. Next, using the first half of the data, the logit regression model is estimated. Then, using the second half of the data, the ROC curve is computed.

Bank rating used	Dependent variable: CB default		Dependent variable: Bank default	
	Regressors: (1) CBR, (2) CBR and BR		Regressors: (1) BR, (2) BR and CBR	
	Bank A	Bank B	Bank A	Bank B
	(1)	(2)	(3)	(4)
In-sample	0.0001	0.0000	0.0000	0.0000
Out-of-sample	0.0002	0.0003	0.0000	0.0000

Panel II: Does each credit rating significantly affect the distribution of implied (in- and out-of sample) default probabilities? This table contains "two-sample Kolmogorov-Smirnov tests" of equality of the implied distribution functions of in- and out-of sample default probabilities. The probabilities are generated from a logit regression model of loan default risk. Defaults are obtained from either the credit bureau data or the bank rating data. First (i) Defaults are regressed on the lag of the same rating, then in a second step (ii) the first lag of another rating is added. The distributions of the default probabilities implied by each of these two models are then compared using a K-S test. The two-sample K-S test first tests the hypothesis that the distribution for model 1 contains smaller values than for model 2, then tests the reverse. The corrected p-values displayed here are for a two-sided test of the Null hypothesis that the distributions are equal taking into account that each one sided test used a different sample size.

Bank rating used	Dependent variable: CB default		Dependent variable: Bank default	
	Regressors: (1) CBR, (2) CBR and BR		Regressors: (1) BR, (2) BR and CBR	
	Bank A	Bank B	Bank A	Bank B
	(5)	(6)	(7)	(8)
In-sample	0.0000	0.0000	0.0000	0.0000
Out-of-sample	0.0000	0.0000	0.0000	0.0000

higher quality than one would expect from theory, whereas bank ratings are less so. If this is indeed the case, then we should at least expect credit bureau ratings to also be better predictors of credit bureau defaults, i.e., bankruptcies, than bank ratings are. Since credit bureau ratings are constructed to predict bankruptcy, whereas bank ratings are designed to predict defaults in loan portfolios, any other finding would cast doubt on our conclusions in Section 4.1.

To verify if this proposition holds, we perform an additional test on the data: we compare the explanatory power of bank credit ratings and credit bureau ratings in a duration model setting and test if either the credit bureau rating or the bank credit rating can be excluded without a significant deterioration in the explanatory power. We implement the test by estimating the following Cox proportional hazards model:

$$\log h_i(t) = \alpha(t) + \beta x_{it} + \epsilon_{it} \tag{7}$$

or equivalently,

$$h_i(t) = h_0(t) \exp(\alpha(t) + \beta x_{it} + \epsilon_{it}) \tag{8}$$

for a number of competing specifications. Here, $h_i(t)$ is the hazard rate of firm i at time t , $\alpha(t) = \log h_0(t)$, and x contains all time-varying covariates. The Cox model leaves the baseline hazard function unspecified, thereby making relative hazard ratios both proportional to each other and independent of time other than through values of the covariates.

We run three sets of regressions to verify the above assertion. In the first group of regressions, the main variable of interest is a firm's bankruptcy hazard rate, or instantaneous risk of bankruptcy at time t conditional on survival to that time. First, we use a subsample of borrowers in Bank A and let $x_{it} = r_{c,t-1}$ to compute the explanatory power of lagged credit bureau ratings. Next, we take $x_{it} = r_{b,t-1}$ and finally $x_{it} = (r_{c,t-1}, r_{b,t-1})$. The log likelihood values for these regressions are collected in the first three rows of Table 8, column (1). In the last three

rows, we present results from regressions where we instead use dummy variables for the ratings and let the vector of explanatory variables in the respective regressions be

$$x_{it} = [DUMr_{c,t-1}^1, DUMr_{c,t-1}^2, \dots, DUMr_{c,t-1}^{G-1}], \tag{9}$$

$$x_{it} = [DUMr_{b,t-1}^1, DUMr_{b,t-1}^2, \dots, DUMr_{b,t-1}^{K-1}] \tag{10}$$

or

$$x_{it} = [DUMr_{c,t-1}^1, DUMr_{c,t-1}^2, \dots, DUMr_{c,t-1}^{G-1}, DUMr_{b,t-1}^1, DUMr_{b,t-1}^2, \dots, DUMr_{b,t-1}^{K-1}] \tag{11}$$

where $DUMr_{b,t-1}^g = 1$ if $r_{b,t-1}^g = g$ and zero otherwise, and G and K are the number of grades in the respective rating systems.

The log likelihood values in the first two rows of Table 8, column (1) show that the lagged credit bureau rating is better at explaining bankruptcy hazard rates than the lagged Bank A rating is. This finding is robust to exchanging the lagged rating for a set of lagged rating dummies (rows (4) and (5)). When we repeat the same Cox regressions for Bank B borrowers, we obtain similar results (Table 8, column (3)). We also performed an additional robustness test - not shown here - where we instead use a two-quarter lag instead of a one-quarter lag. This modification does not change the results qualitatively. As expected, the coefficients on the lagged rating dummies are monotonically increasing in risk for both the credit bureau and the bank ratings. This reflects the fact that higher ratings should be stronger indicators of future defaults and therefore should be associated with higher hazard rates.

In columns (2) and (4), we present the results from a similar set of Cox regressions where the dependent variable instead is the instantaneous risk of a loan default in Bank i , $i = \{A, B\}$, at time t , conditional on survival to that time. A similar comparison as above between the likelihoods in the first two rows makes it clear that, for

Table 8

Panel I: How good are bank credit ratings (BR) and credit bureau ratings (CBR) at predicting fundamental events (bankruptcies and loan defaults)?

The table displays the loglikelihood values for different Cox regression specifications on credit bureau defaults and bank defaults. The Breslow method has been used for tied observations. The upper three regressions use the continuous rating variable as an explanatory variable while in the lower three regressions the rating variable has been converted into a set of dummy variables.

Explanatory variables	Dependent variable		Dependent variable	
	CB default	Bank A default	CB default	Bank B default
	Bank rating used on RHS: Bank A		Bank rating used on RHS: Bank B	
	(1)	(2)	(3)	(4)
Lag(1) CBR	−1461.7	−2686.1	−1080.6	−1353.6
Lag(1) BR	−1507.9	−2695.9	−1096.5	−1372.1
Lag(1) CBR, Lag(1) BR	−1434.2	−2564.2	−1051.4	−1308.0
Lag(1) of Dummy CBR	−1459.0	−2714.7	−1079.1	−1375.5
Lag(1) of Dummy BR	−1504.8	−2687.5	−1121.7	−1369.4
Lag(1) of Dummy CBR, Lag(1) Dummy BR	−1434.4	−2580.5	−1057.6	−1321.3
Nobs	214,968	214,925	119,819	119,812
No. Failures	165	312	128	163
No. Subjects	31,621	31,635	17,475	17,490

Panel II: Exclusion tests to test if any of rating variables can be excluded on Cox regressions of bankruptcy or loan default.

Panel II contains exclusion test results of the hypothesis that either the bank rating or the credit bureau rating can be excluded as an explanatory variable from the Cox regressions with both the continuous bank rating and the continuous credit bureau rating on row 3 in panel I without significantly reducing the explanatory power of the model. The exclusion tests thus compare the log-likelihood in row 1 and row 2 respectively with row 3. Significance of the “additional” explanatory variable is shown at the 5(*), 1(**), and 0.1(***). In the likelihood ratio tests (lower panel), the value displayed is $2 \cdot \log(\text{likelihood ratio})$.

Explanatory variables	Dependent variable		Dependent variable	
	CB default	Bank A default	CB default	Bank B default
	Bank rating used on RHS: Bank A		Bank rating used on RHS: Bank B	
	(1)	(2)	(3)	(4)
Exclusion of Lag(1) BR	55.0***	243.8***	58.4***	91.2***
Exclusion of Lag(1) CBR	147.4***	263.4***	90.2***	128.2***

both banks, lagged credit bureau ratings are better at explaining loan default hazards than the bank ratings are themselves. These results are also robust to exchanging the first lag with the second lags of the explanatory ratings.

The results in Table 8 also illustrate how the nonlinearities in both bank and credit bureau ratings come into play in our analysis. Columns (2) and (4) show that if one implicitly imposes a restriction of equal marginal effects of rating grade changes on the loan default hazard, then both Bank A and Bank B rating adjustments have substantially less explanatory power for the hazard than credit bureau ratings do (row (2) vs row (1)). However, once the implicit equality constraint is relaxed by instead using dummy variables, this relationship reverses and adjustments of bank ratings are found to have the greater impact on the hazard rate (columns (2) and (4), row 5 vs. row 4). This reversal may be caused by the fact that default risk is very small for a nontrivial number of corporations. Deteriorations of these companies’ ratings thus lead to a very large increase in the hazard rate. Imposing that each one-notch change in the rating must have an equally sized effect on the hazard restricts the explanatory power of the ratings. When comparing columns (2), (4) with (1), (3), we can see that this loss of information is greater when using bank ratings than credit bureau ratings, most likely because the former are less persistent.

In panel II of Table 8, we present the likelihood ratio test statistics for exclusion of the bank rating in the credit bureau autoregression and for exclusion of the credit bureau rating in the bank rating autoregression. High values of the test statistic mean that adding an explanatory variable leads to a large increase in the loglikelihood and that the null hypothesis of equal explanatory power is rejected, i.e., the additional variable should not be excluded. For example, the log likelihood of the model with the credit bureau rating alone in the

regression explaining credit bureau default for all Bank A borrowers is -1461.7 . As the regression that uses both the credit bureau rating and the Bank A rating has a log likelihood of -1434.2 , twice the log likelihood ratio amounts to 55.0, making the Bank A rating very significant in a chi-square test with one degree of freedom. The fact that none of the variables can be excluded in panel II implies that neither the bank ratings nor the credit bureau ratings are on their own sufficient statistics of default. This holds for all definitions of default as well as for when we lag both ratings an additional period. This provides the most conclusive evidence that the credit bureau rating adds information to the bank rating even though the bank loan officers have ready access to the credit bureau ratings when they make their ratings.

4.3. Are staggering of information and rating coarseness driving the results?

In the theoretical model of Section 2.1, we made two assumptions about the format and updating frequency of the credit ratings. To start with, credit ratings were allowed to be continuous. Moreover, we treated the banks and the credit bureau as if they update their ratings simultaneously in each time period. The actual credit rating data we work with depart from these assumptions in two respects.

First, credit ratings are categorical, not continuous, variables. In moving from continuous variables to categorical variables, the bank rating may lose information, thereby making the credit bureau data more valuable. When bank credit ratings are categorical, some of the information in the public signal is not captured in the bank’s credit rating. If credit bureau ratings are continuous, the public monitor’s rating will contain information that has been lost in the aggregation. Then the public monitor’s rating may well predict the bank’s signal even though the bank is fully aware of the public signal and “processes” it optimally.

Table 9

Robustness tests: Controlling for staggered updating of the information and coarseness of the underlying credit ratings. This table shows the explanatory power of lagged bank ratings or credit bureau ratings in OLS regressions using only observations of lagged ratings variables when change of rating is observed. Entries in the table reflect the percentage by which the residual sum of squares is reduced when a four-period lag of bank ratings or credit bureau ratings is introduced as an explanatory variable in addition to the lagged dependent variable. Data are at yearly frequency. Therefore we provide the results for Table 5 re-run on four-quarter lags in columns (9) through (12) to facilitate a comparison with the other columns of this table.

Explanatory variable added	Dependent variable			
	Credit bureau rating		Bank A rating	Bank B rating
	Regressions Conditioned on Bank Rating Change			
	(1)	(2)	(3)	(4)
	Bank A rating	Bank B rating	Credit bureau rating	
All borrowers	1.69	6.92	8.84	2.80
Small borrowers	2.11	4.36	9.80	1.55
Medium-sized borrowers	1.78	6.86	8.19	2.41
Large borrowers	0.83	7.41	8.55	3.40
	Regressions conditioned on credit bureau rating change			
	(5)	(6)	(7)	(8)
All borrowers	1.59	2.98	9.56	1.91
Small borrowers	1.76	1.56	10.12	0.96
Medium-sized borrowers	1.42	2.93	10.04	1.96
Large borrowers	1.69	3.15	7.39	1.87
	Unconditioned regressions			
	(9)	(10)	(11)	(12)
All borrowers	1.49	2.40	7.92	2.36
Small borrowers	1.50	1.87	9.20	1.47
Medium-sized borrowers	1.44	2.31	7.65	2.77
Large borrowers	1.60	2.54	6.09	1.85
	Continuous credit bureau rating		Bank A rating	Bank B rating
	Bank A rating	Bank B rating	Continuous credit bureau rating	
Data selection	(13)	(14)	(15)	(16)
Conditioned on bank rating changes, all borrowers	1.51	3.35	5.53	1.98
Conditioned on credit bureau rating changes, all borrowers	3.99	4.14	3.22	0.63
Unconditioned, all borrowers	5.54	5.00	3.32	0.87

However, when both public and private monitors produce categorical ratings, we can no longer be sure what impact the loss of information due to converting continuous projections into categorical ratings will have on the mutual forecasting power of public and private ratings.

Second, our data set does not allow us to control for the exact time at which updating of information sets takes place. Hence, bank and credit bureau ratings changes may be staggered, without the data explicitly accounting for differences in information sets between monitors. The data-providing banks update their credit ratings at least once a year and, in practice, do so close to once per year on average. The credit bureau collects data from financial institutions, corporations, and official resources at a higher frequency. For payments remarks, this occurs more or less daily, while for other variables, this typically happens at a yearly and sometimes a quarterly or monthly frequency. In some instances, the credit bureau may thus have updated its credit rating more recently than the bank, which can create a potential for credit bureau ratings to forecast the bank ratings. At other times, banks may already have received parts of a company's financial statement before it was filed. In our regression results in Tables 4 to 6, the former would generate a downward bias while the latter would produce an upward bias in the estimated amount of private information.

To accommodate the above two deviations from our model assumptions, we relaxed the tests of Hypotheses 1 and 2 in Sections 4.1.1 and 4.1.2, i.e., by testing weaker parameter restrictions on the lagged

dependent variables.

To make sure that our finding that bank internal credit ratings contain private information but are inefficient is not a mere result of the staggering of information sets and the coarseness of rating grades, we perform two tests. First, we remove all observations for which we cannot be sure that the information sets impounded in the ratings were recently been updated. We can do this for both the credit bureau and the bank data. Second, we use continuous measures of creditworthiness rather than discrete ratings. We can do this for the credit bureau data only. In Sections 4.3.1 and 4.3.2, we describe these tests in greater detail.

4.3.1. Staggering of information

First, we repeat the regressions underlying columns (4), (8), (12), and (16) in Table 5 while restricting the data set to observations where bank ratings had just been modified. Our data set does not permit us to directly observe the quarter in which the bank loan officer collected information to review the credit ratings, but we can observe if bank ratings were just modified.¹⁶ Because credit ratings can be adjusted

¹⁶ We follow the approach of Bils et al. (2009), who study staggered prices on the assumption that menu costs prevent observed prices from equaling shadow prices. They use the observations when prices change to infer underlying shadow price movements. Because most bank clients are reviewed once a year, we use four-quarter lags for the right-hand-side variables in these regressions.

Table 10
 Panel I: Does mean reversion in Bank credit rating changes point at overconfidence as an explanation of rating inefficiency? The table displays regression coefficients and their standard errors (in parentheses) for OLS regressions of change in Bank A or Bank B ratings regressed on lagged change in the same Bank rating (all borrowers). Significance levels are shown at the 5 (*), 1 (**), and 0.1 (***) levels. Sample period is 1997:Q3 to 2000:Q1; standard errors are clustered by borrower. Changes computed over different time intervals.

Explanatory variable used	Dependent variable							
	Change in Bank A rating				Change in Bank B rating			
	Δ BRA(L1)	Δ BRA(L2)	Δ BRA(L3)	Δ BRA(L4)	Δ BRB(L1)	Δ BRB(L2)	Δ BRB(L3)	Δ BRB(L4)
Constant	0.03183*** (.00079)	0.006897*** (.00180)	0.11814*** (.00317)	0.18065*** (.00500)	0.01145*** (.00062)	0.01936*** (.00139)	0.02453*** (.00244)	0.03170*** (.00381)
Δ BR(k)(L [i]), k = A, B	-0.05598*** (.00335)	-0.06305*** (.00324)	-0.10286*** (.00516)	-0.15375*** (.00940)	-0.02370*** (.00470)	-0.04656*** (.00570)	-0.09519*** (.00780)	-0.13202*** (.01116)
Adj. R2	0.0034	0.0046	0.0106	0.0206	0.0005	0.0020	0.0080	0.0146
Nobs	174788	122346	79072	42696	96689	68692	45605	27234

Panel II: Does mean reversion in ratings occur only for certain borrower sizes or irrespective of borrower type (size)? This table displays OLS regressions of the change in Bank A and Bank B on the lagged change in the rating, for subsets of borrowers by size. The coefficients are those on the lagged change in the (compressed) Bank A rating or the Bank B rating. Significance levels are shown at the 10 (*), 5 (**), 1 (***), and 0.1 (****) levels. Sample period is 1997:Q3 to 2000:Q1, standard errors are clustered by borrower

Explanatory variable used	Dependent variable							
	Change in Bank A rating				Change in Bank B rating			
	Δ BRA(L1)	Δ BRA(L2)	Δ BRA(L3)	Δ BRA(L4)	Δ BRB(L1)	Δ BRB(L2)	Δ BRB(L3)	Δ BRB(L4)
Small borrowers	-0.0582*** (.0053)	-0.0642*** (.0056)	-0.1054*** (.0088)	-0.1625*** (.0166)	-0.0277*** (.0235)	-0.0226*** (.0304)	-0.1015*** (.0430)	-0.0948*** (.0669)
Medium-sized borrowers	-0.0527*** (.0046)	-0.0648*** (.0048)	-0.1022*** (.0078)	-0.1518*** (.0137)	-0.0204*** (.0052)	-0.0515*** (.0071)	-0.1027*** (.0104)	-0.1387*** (.0144)
Large borrowers	-0.0592*** (.0099)	-0.0567*** (.0066)	-0.1001*** (.0108)	-0.1428*** (.0196)	-0.0481*** (.0089)	-0.0411*** (.0098)	-0.0842*** (.0122)	-0.1250*** (.0182)

Table 11

Bank rating inefficiency related to measure of overconfidence. The dependent variable is Bank Inefficiency, the percent by which mean square error of predictions of bank ratings are reduced by lagged credit bureau ratings from Table 5. The explanatory variable is the coefficient of mean reversion (Reversion Coefficient) from Table 10, Panel II. The observations are borrower size-lag length-Bank combinations, with 3 sizes of borrower, 2 banks, and 4 lag-lengths, there are 24 observations. Errors are clustered by Size-Bank combination. Significance levels are shown at the 5 (*), 1 (**), and 0.1 (***) levels.

Explanatory variables	Dependent variable			
	Basic	Adding lag length dummies	Adding bank-size dummies	Adding both dummies
	(1)	(2)	(3)	(4)
Constant	-0.0003 (0.0032)	0.0261 * (0.0112)	-0.0348** (0.0091)	-0.0447** (0.0121)
Reversion coefficient	0.4017*** (0.0900)	1.0311*** (0.2266)	0.2812** (0.0930)	0.0117 (0.1746)
Size-bank dummies	No	No	Yes	Yes
Length of lag dummies	No	Yes	No	Yes
Degrees of freedom	22	19	17	14
R sq	0.45	0.66	0.88	0.89
Statistics		Bank rating inefficiency	Reversion coefficient	
Mean		0.031	0.070	
Standard deviation		0.033	0.049	

only after a loan officer has updated and filed client information, limiting regressions to these observations eliminates any chance that the credit bureau rating reflects more recent information than the bank rating. Here we use four-quarter lags because the changes in bank ratings occur most often at annual intervals. For Bank A, for example, rating changes that take place four quarters after an earlier rating change occur five times as often as one to three quarter intervals. For Bank B, this frequency is two to three times higher than in the other intervals. This reflects the timing of the bank credit review, which is typically annual. As a result of this timing, there is less likely to be a change in the bank rating in the first period or two after a rating review, thus giving less scope for the credit bureau rating to have explanatory power over movements in the bank credit rating when we condition on a rating change. Using four-quarter lags in this case ensures the greatest likelihood of detecting whether the credit bureau rating can truly predict a subsequent bank rating change, after the effect of “aging” of the bank rating has been filtered out. The results of these regressions are presented in Table 9, to which we have added a copy of the results for the regressions without conditioning on fresh rating changes from Table 5 to facilitate a comparison.

The results in columns (1) to (2) of Table 9 show that bank ratings have greater predictive power for credit bureau ratings when we restrict ourselves to observations where bank ratings were just renewed. This holds in particular for Bank B ratings, which on average contribute to a 6.92 percent decline in RSS (column 2) compared to 2.40 percent in unconditioned regressions (column 10). Columns (3) to (4) show that credit bureau ratings preserve their explanatory power for both Bank A and Bank B credit ratings. Coefficients tend to become substantially larger for Bank B ratings (not shown here), when we condition on Bank B’s rating being newly updated, and economically more meaningful, rising to roughly 0.3. The effect is smaller when doing the same experiment for Bank A ratings. In line with our earlier findings, the contribution of Bank B ratings to future movement in credit ratings is greater than the contribution of Bank A ratings, and the contribution of credit bureau ratings is greater with respect to Bank A ratings than to Bank B ratings. When we split up the data into small, medium-sized, and large businesses, the same pattern emerges as before: The predictive power of external ratings is manifest in the case of small businesses and least distinct with respect to larger businesses. For Bank B we should not draw any conclusions from the results for small businesses because of the very small subsample size.

Note that when we select only observations with fresh bank rating changes, the credit bureau ratings also become more predictive of bank credit ratings. One explanation for this may be that borrowers who experience rating changes at a bank are noisier than other borrowers. An alternative explanation could be that this is driven by overconfidence, which we will discuss in greater detail in Section 4.4.

In a second robustness test, we repeat the same set of regressions using only observations where the credit bureau rating had just been altered. Again, we find that restricting the data set does not bring about any qualitative changes in the results. Bank credit ratings still have predictive power with respect to credit bureau ratings. Columns (5) to (8) of Table 9 make it clear that, just as in Section 4.1.2, Bank B’s ratings are better predictors of future credit ratings than Bank A’s ratings are. Consistent with earlier results, Bank B appears to have a slight advantage in rating larger companies. More detailed results, provided in the working paper version, show that the coefficients on credit bureau ratings, conditional on the credit bureau’s rating being newly changed, do not increase for the regressions predicting Bank B ratings, although they do for Bank A ratings. The smaller impact from refreshing information in credit bureau variables is not so surprising, because credit bureau ratings are updated relatively continuously. Hence the lack of a credit bureau rating change does not imply that the rating maybe obsolete in the same way as it does for bank ratings.

Overall, these tests demonstrate that the staggering of information updating by the credit bureau and the banks in our dataset does not affect our conclusion that our banks’ internal credit ratings do contain private information, consistent with theory, but are nevertheless inefficient measures of creditworthiness.

4.3.2. Coarseness of the rating scale

For a last test, we investigate whether using discrete instead of continuous credit bureau ratings has altered the explanatory power that we attribute to lagged bank ratings. For this purpose, we exploit that the credit bureau has provided us not only with the actual credit rating but also with the near-continuous measure of creditworthiness that underlies its credit rating. This is a numerical rating that runs from 0 to 100 (from 0.5 to 1 and then by units up to 99). We take logarithms of these numerical ratings, and we rerun the regressions of columns (1) and (2) in Table 4 while using continuous measures of creditworthiness for the credit bureau rating. Columns (13) and (14) in Table 9 show that bank credit ratings have predictive power for credit bureau ratings even when

the creditworthiness information has not been made discrete. This is the strongest evidence so far that the banks produce private information.

Unfortunately, we cannot repeat this test for bank credit ratings because we do not have similar continuous creditworthiness measures for the banks. We can therefore not investigate further if the inefficiency in the bank ratings is caused by the discreteness of the ratings or by banks not efficiently impounding their private information into the public information. The above tests for the credit bureau ratings do help us understand better that the discrete nature of ratings does not per se result in substantial inefficiency. In Section 4.4, we will look further into the latter of these two alternative explanations for our findings and investigate if loan officers inefficiently impound private information into public information because of overconfidence.

4.4. Can overconfidence explain our findings?

In Section 4.1 we found that credit bureau ratings are able to predict bank ratings. One possible rationale for the existence of such an inefficiency in the literature, as we have mentioned, is that bank loan officers overvalue their private information relative to the information content of the credit bureau ratings. To account for such a possibility, we discuss in Section 4.4.1 what happens in the context of our model of information processing when we modify it to include the possibility that loan officers are overconfident. In Section 4.4.2 we will also test the implications from a model that incorporates overconfidence on our data.

4.4.1. Information processing with overconfidence

One can think of a bank loan officer as somebody who is overconfident in the private signal and believes that the precision of the private signal is greater than its true precision. The consequence of overconfidence will be that the loan officer will aggregate the two signals with a greater weight given to the private signal relative to the public signal. This adds more of the random noise in the bank's private signal to the bank's credit rating than is optimal. Moreover, believing this combined signal to be more accurate than is the case, the overconfident loan officer will construct an estimate of creditworthiness in which "updating" takes place too quickly. This implies that the bank rating will - because the weight on the private signal is too great - overshoot in the direction of the private signal. Over time, this will generate reversion toward the mean; such movements in the rating will be forecastable and produce negatively serial correlation in the first differences of credit ratings. Since these reversion movements arise from overaggressive ratings changes, overconfidence can explain the greater forecasting power on the part of credit bureau ratings following bank rating changes documented in Table 9.

One can use the same model to think about a bank loan officer who is underconfident in the private signal, making the weight on it too low. The noise in this underweighted signal will, however, never be worse than the noise of the signal of the credit bureau, and the Kalman filter coefficient in the loan officer's calculation will still be greater than that of the credit bureau. Thus, the credit bureau's estimate will theoretically not forecast the bank rating. Nor will there be mean reversion; in fact, because an underconfident loan officer places too little weight on the private signal, there will be a tendency for the bank rating to undershoot and to move further in the same direction when the next signal is received. Thus, movements in the bank rating with underconfidence will result in the rating changes being positively serially correlated.

Overconfidence of the bank loan officer in his or her private information thus results in future predictability of the bank credit rating as a consequence of the excessive weighting of the noisy private signal. Overconfidence also leads to excessive movement of the credit rating in the direction of the private signal, followed by a mean reversion. Such mean reversion is something we can directly detect in a difference regression: The previous period's change in rating should predict this period's change in rating with a sign reversal. We will test this in Section 4.4.2.

4.4.2. Testing for the presence of overconfidence

To test for the presence of overconfidence, we estimate the following difference regression:

$$\Delta r_{bt} = a + b\Delta r_{bt-1} + u_t \quad (12)$$

If we find that $\hat{b} < 0$, it will imply evidence of overconfidence, while $\hat{b} > 0$ would imply evidence of underconfidence. We run the difference regression (12) using different lags: the difference from the past quarter, the past two quarters, the past three quarters, and the past year. Columns (1) to (4) in Table 10 panel I, show these first for all Bank A borrowers, then columns (5) to (8) repeat this for Bank B borrowers. The coefficients are uniformly negative and highly significant, i.e., we find significant mean reversion in all cases, as the overconfidence theory in Section 4.4.1 predicts. We thus reject the possibility of underconfidence on two grounds. First, we observe mean reversion rather than positive serial correlation of bank rating changes. Second, as we found in Section 4.1.2, the credit bureau ratings are predictive of bank ratings.

To verify if overconfidence is specific to certain subsets of borrowers, we repeat the mean-reversion regressions for small, medium-sized, and large borrowers. The results displayed in Table 10, panel II, show that the mean-reversion coefficients are uniformly negative for the bank ratings. For Bank A's ratings (Table 10, columns (9) to (12)), mean reversion is somewhat stronger for the small borrowers and for the longer lags. This is consistent with our earlier observation that Bank A's ratings are more predictable (by means of public credit bureau ratings) for small borrowers and for longer lags.

We also find that the absolute size (signed size) of these mean reversion coefficients is positively (negatively) related to the inefficiency of the bank credit rating (i.e., the RSS improvement we obtained earlier on when we added the credit bureau rating to a regression that explained the bank rating). These are shown in the first column of Table 11. These results are robust to either including a set of five bank-size dummies or a set of three lag-length dummies but not to adding both. This provides some evidence that the degree of overconfidence - as measured by mean reversion - is associated with the predictability of the bank credit rating and therefore of the inefficiency of the bank rating process.

5. Conclusion

Using data from two large Swedish banks, we find strong evidence that these banks, relative to a credit bureau that produces ratings using public information only, obtain private information about their clients and incorporate it into their internal credit ratings. However, we also show that these banks' internal credit ratings do not contain all the information about borrowers that is incorporated in the credit bureau ratings, even though the credit bureau ratings are available to the bank loan officers. Banks generally collect more private information for larger clients in their portfolio, while their inefficiency with respect to incorporating public information is greater for smaller borrowers.

Our findings may be due to banks' failure to incorporate publicly available information optimally or to loss of information in the process of generating credit ratings. We investigate this departure of bank rating measures from optimality and show that it is not due to the staggered timing of rating information updating and is unlikely to be a consequence of the discrete nature of the ratings. We do find that bank ratings are mean reverting, which is consistent with overconfidence, and that mean reversion is related to the degree of predictability of the bank ratings. This finding is consistent with the presence of overconfidence, i.e., bank loan officers placing too much weight on their private information, although we cannot preclude the possibility that these banks face difficulties in aggregating different types of information.

These results imply that, for these banks, it would not be optimal for their risk managers or their regulators to accept the banks' own private credit ratings as the single measure by which to evaluate portfolio credit

risk. Instead, it would be beneficial to incorporate more information into a risk review. In particular, our findings imply that publicly provided credit bureau ratings contain information over and above what is contained in bank credit ratings. Such credit bureau ratings could thus be used to improve overall evaluation of portfolio risk.

The basket of straightforward techniques that we propose enables both financial institutions and regulators to assess the performance of banks' credit ratings systems. By using both internal bank credit ratings and external credit bureau ratings of corporate borrowers, one can investigate if bank credit ratings are able to forecast the ratings of a public monitor, such as a credit bureau. Through the use of these tests, banks may improve on the credit ratings that they employ to evaluate borrowers. The techniques can also be applied to bond ratings for larger commercial loans.

Our analysis raises new theoretical questions about how banks assess the creditworthiness of their customers. Why do banks use relatively crude rating gradations instead of continuous assessments of default risk? In what banking systems or structures is public information likely to be more valuable in the supervision of bank lending? What determines how much private information banks collect on their customers and how they aggregate private and public information? These questions are important issues for future research to address.

References

- Agarwal, S., Hauswald, R., 2010. Distance and private information in lending. *Rev. Financ. Stud.* 23 (7), 2757–2788.
- Altman, E.I., Rijken, H.A., 2004. How rating agencies achieve rating stability. *J. Bank. Finance* 28, 2679–2714.
- Berger, A., Davies, S., Flannery, M., 2000. Comparing market and supervisory assessments of bank performance - who knows what when. *J. Money Credit Bank.* 32 (3), 641–667.
- Bils, M., Klenow, P.J., Malin, B.A., 2009. Reset price inflation and the impact of monetary shocks. NBER Working Paper 14787, March, NBER.
- Cantor, R., 2004. An introduction to recent research on credit ratings. *J. Bank. Finance* 28, 2565–2573.
- Chen, Q., Jiang, W., 2006. Analysts' weighting of public and private information. *Rev. Financ. Stud.* 19, 319–355.
- Claessens, S., Embrechts, G., 2003. Basel 2, sovereign ratings and transfer risk external versus internal ratings. Manuscript, ssrn.com/abstract=386480.
- Dell'Ariccia, G., Marquez, R., 2004. Information and bank credit allocation. *J. Financ. Econ.* 72, 185–214.
- Diamond, D., 1984. Financial intermediation and delegated monitoring. *Rev. Econ. Stud.* 51, 393–414.
- Fama, E.F., 1985. What's different about banks? *J. Monetary Econ.* 15, 29–40.
- Grunert, J., Norden, L., Weber, M., 2005. The role of non-financial factors in internal credit ratings. *J. Bank. Finance* 29, 509–531.
- Hertzberg, A., Liberti, J.M., Paravisini, D., 2010. Information and incentives inside the firm: evidence from loan officer rotation. *J. Finance* 65, 795–828.
- Jacobson, T., Lindé, J., Roszbach, K., 2006. Internal ratings systems, implied credit risk and the consistency of banks' risk classification policies. *J. Bank. Finance* 30, 1899–1926.
- Jacobson, T., Lindé, J., Roszbach, K., 2013. Firm default and aggregate fluctuations. *J. Eur. Econ. Assoc.* 11 (4), 945–972.
- Kahneman, D., 2011. *Thinking Fast and Slow*. Farrar, Strauss and Giroux, New York.
- Keys, B.J., Mukherjee, T., Seru, A., Vig, V., 2009. Financial regulation and securitization: evidence from subprime loans. *J. Monet. Econ.* 56 (5), 700–720.
- Krahnhen, J.P., Weber, M., 2001. Generally accepted rating principles: a primer. *J. Bank. Finance* 25, 3–23.
- Löffler, G., 2004. Ratings versus market-based measures of default risk in portfolio governance. *J. Bank. Finance* 28, 2715–2746.
- Lummer, S.L., McConnell, J.J., 1989. Further evidence on the bank lending process and the capital market response to bank loan announcements. *J. Financial Econ.* 25, 99–122.
- Mester, L.J., Nakamura, L.I., Renault, M., 2007. Transactions accounts and loan monitoring. *Rev. Financ. Stud.* 20, 529–556.
- Muth, J.F., 1960. Optimal properties of exponentially weighted forecasts. *J. Am. Stat. Assoc.* 55 (290), 299–306.
- Nakamura, L. I., Roszbach, K., 2013. Credit Ratings and Bank Monitoring Ability. Federal Reserve Bank of Philadelphia Working Paper. 13–21.
- Pagano, M., Japelli, T., 1993. Information sharing in credit markets. *J. Finance* 48 (5), 1693–1718.
- Povel, P., Singh, R., Winton, A., 2007. Booms, busts, and fraud. *Rev. Financ. Stud.* 20 (4), 1219–1254.