WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 13-21
CREDIT RATINGS AND BANK MONITORING ABILITY

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

Kasper Roszbach
Sveriges Riksbank and University of Gronigen

May 14, 2013

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Credit Ratings and Bank Monitoring Ability*†

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

Kasper Roszbach
Sveriges Riksbank and University of Groningen

May 14, 2013

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' credit ratings indeed include valuable private information from monitoring, as theory suggests. However, our tests also reveal that publicly available information from a credit bureau is not efficiently impounded in the bank ratings: The credit bureau ratings not only predict future movements in the bank ratings but also improve forecasts of bankruptcy and loan default. 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. To the extent that overconfidence results in placing too much weight on private information, risk analyses of the bank loan portfolios in our data could be improved by combining the bank credit ratings and 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.

Keywords: Monitoring, banks, credit bureau, private information, ratings, regulation, supervision, overconfidence.

JEL codes: D82, G18, G21, G24, G32, G33

*We are indebted to Elif Sen, Fuyuo Nagayama, and Gustav Alfelt for providing outstanding research assistance and grateful for comments from Sreedhar Bharath, Martin Brown, Mikael Carlsson, Sonja Daltung, Hans Degryse, Mark Flannery, Mark Flood, Tor Jacobson, Elizabeth Kiser, William Lang, Steven Ongena, Harvey Rosenblum, Frank Schorfheide, Norman Schürhoff, and seminar participants at the 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 the Executive Board of Sveriges Riksbank, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/.

†Nakamura: Research Department, Federal Reserve Bank of Philadelphia; E-mail: LEONARD.NAKAMURA@PHIL.FRB.ORG. Roszbach: Research Division, Sveriges Riksbank and University of Groningen; E-mail: KASPER.ROSZBACH@RIKSBANK.SE

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 onset of the recent 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.

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 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.

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, Nakamura, and Renault (2007) describe in detail how banks' monitoring activities, by using transaction account information that provides ongoing data on borrowers' activities, make these intermediaries superior 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, Singh, and Winton (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.

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 informativeness 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

[1] Grunert, Norden, and Weber (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.

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, our focus is 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, Davies, and Flannery (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 demonstrate 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

[3] 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 Krahnen 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."

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. Although we cannot rule out that the discretization of the bank ratings may be responsible for the apparent inefficiency of the information aggregation by the banks, we have strong prima facie evidence that at least one of the two banks has inefficient ratings.

To assess the fundamental quality of our ratings, we also evaluate their predictive accuracy with respect to fundamental events like bankruptcy and loan default. We find that including both the bank rating and the credit bureau rating in a regression increases models' predictive accuracy.[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 "soft" and "hard" 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. We look into the first and second explanations. Our results indicate that adding soft information to hard 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, Liberti, and Paravisini (2010) point out that career concerns may cause loan officers' credit ratings to be biased optimistically.

We extend our benchmark model to explicitly allow for overconfidence and find some evidence that is consistent with the presence of overconfidence among loan officers. We build upon our Kalman filter model by allowing for the possibility that the loan officer is overconfident in

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

the information from the private bank signal, modeled by assuming that the loan officer believes that the precision of the private signal is greater than it is in fact. As a consequence, the bank rating overshoots in the direction of the noise in the private signal and then mean-reverts. This results in a mean reversion in the rating and an ability of the public signal to forecast the bank rating. We show that the bank rating in fact mean-reverts and that this mean reversion is greater for cases in which the public signal does a better job of forecasting the bank rating. This evidence, and the lack of other possible causes, does 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 most of the empirical analysis, including a series of tests that seek to account for the possibility that credit ratings may not be linear in risk. Section 5 presents the overconfidence extension of our theoretical model and corresponding empirical analysis. Section 6 concludes.

2 A model of information processing

A well-known theory of banking is that banks possess private information about the credit-worthiness of borrowers. This information may, for example, be derived from the transaction accounts of borrowers (Mester, Nakamura, and Renault, 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, 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 set forth a simple theory of signal extraction that describes how producers of credit ratings optimally process different signals of a borrower's creditworthiness. The theory will produce 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 formulate a simple theoretical model. 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 bank and credit bureau ratings measure the objective underlying risk of default.

By means of our first assumption we exclude cases where ratings are loan-specific. 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 that 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.

In the theoretical model we set up below, banks will have private information about the creditworthiness of their borrowers. This information is modeled as a noisy signal that the bank receives. We then show 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 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. Stated differently, a monitor's expectation of creditworthiness turns out to be an exponentially weighted lag of its past signals, with a base coefficient, d_i, on the current period's signal. The size of this base coefficient is determined by the relative precision of the monitor's signal, q_i.

We assume that each borrower j has some actual measure of creditworthiness, y_{jt}, that follows a random walk. For notational simplicity we will, however, suppress the subscript j. Each period, the noise term $u_t \sim N\left(0, \sigma^2\right)$ permanently shifts the underlying creditworthiness y_t:

$$y_t = y_{t-1} + u_t \tag{1}$$

There are two monitors indexed by i, $i \in \{b, c\}$, where b is a bank and c is a credit bureau. The signal of the underlying creditworthiness that each monitor i receives contains a temporary, normally distributed noise term $\eta_{it} \sim N\left(0, \sigma_{i\eta}^2\right)$. If we define the precision of monitor i's observation q_i relative to the disturbances of the actual creditworthiness, i.e., $q_i \equiv \sigma^2/\sigma_{i\eta}^2$, then it follows that $\sigma_{i\eta}^2 = \sigma^2/q_i$.

The credit bureau c observes a noisy public signal, s_{ct} of a borrower's creditworthiness y_t:

$$s_{ct} = y_t + \eta_{ct} \tag{2}$$

We introduce the following notation: $y_{ct|t-1} \equiv E\left(y_t|s_{ct-1}\right)$ is the credit rating before receiving the current period's signal, while $y_{ct|t} \equiv E\left(y_t|s_{ct}\right)$ is the filtered signal, which we will interpret as the updated credit rating.

If the noise terms are normally distributed, then the process by which the bank updates its credit ratings is linear in the past period's rating and the current signal. Then the updating process by the credit bureau is captured by the following regression equation:[7]

$$y_{ct|t} = \left(1 - d_c\right)\, y_{ct-1|t-1} + d_c\ s_{ct} \tag{3}$$

where d_c is a regression coefficient. Since $s_{ct} = y_{t-1} + u_t + \eta_{ct}$, this estimate incorporates in each period a proportion d_c of the current shock u_t and a proportion $1 - d_c$ of the past shocks incorporated in $y_{ct-1|t-1}$. In (3) we can use repeated substitution to obtain Muth's formula:

$$y_{ct|t} = d_c \sum_{i=o}^{\infty} \left(1 - d_c\right)^i s_{ct-i} \tag{4}$$

[7]Some intermediate steps inderiving the implications of the signal extraction model have been omitted from the main text and are available in Appendix 2.

It can be shown that the stationary solution is (Chow, 1975):

$$d_c = \frac{q_c}{2}\left(\sqrt{1+4/q_c}-1\right) \tag{5}$$

and that $\frac{\partial d_c}{\partial q_c} > 0$. A monitor thus updates his expectation of creditworthiness more slowly as the noise of the signal increases.

Moreover, the expected mean square error of the credit rating, $V_{ct|t} \equiv E\left(y_t - y_{ct|t}\right)^2$, equals

$$V_{ct|t} = \frac{\sigma^2}{2}\left(\sqrt{1+4/q_c}-1\right) \tag{6}$$

which we will use below. Table 1 displays how the updating coefficient d_c varies with the precision of the monitor's signal, q_c. The table shows that d_c falls faster in ranges where q_c is very small. For example, quadrupling the precision of the noise doubles the updating speed when precision is below. say, 0.27. In what may be considered the relevant ranges of precision for a monitor (between 3 and .05), a quadrupling of the relative noise in a signal reduces d_i by approximately 20 percentage points. In the relevant range, a quadrupling thus goes between doubling to adding one-third to d_c so, in this range, the rise in d_c rises is slower than the rise in precision.

Table 1: Values of d_c as a function of q_c

All entries have been constructed using equation (5)

q_c	3.2	1	0.27	.05	.011	.0026	.00064
d_c	.800	.618	.402	.200	.100	.050	.025

The bank, unlike the credit bureau, not only observes the same public signal as the credit bureau, but gets a noisy, private signal, s_{pbt}, of borrowers' actual creditworthiness:

$$s_{pbt} = y_t + \eta_{pbt} \tag{7}$$

where

$$\eta_{pbt} \sim N\left(0, \sigma^2/q_{pb}\right) \tag{8}$$

After receiving the signals, the bank aggregates them in proportion to their respective precision, q_{pb} and q_c, to form a composite signal s_{bt}:

$$\begin{aligned} s_{bt} &= (q_{pb}\ s_{pbt} + q_c\ s_{ct}) / (q_{pb} + q_c) \\ &= y_t + \eta_{bt} \end{aligned} \tag{9}$$

where

$$\begin{aligned} \eta_{bt} &= \left(q_{pb}\ \eta_{pbt} + q_c\ \eta_{ct}\right) / (q_{pb} + q_c) \\ &\sim N\left(0, \sigma^2/q_b\right) \end{aligned} \tag{10}$$

is the noise term of the bank's composite signal, and

$$q_b = q_{pb} + q_c \tag{11}$$

is the precision of the composite signal.

The composite signal can then be treated just like the public signal in Muth's formula, i.e.:

$$y_{bt|t} = d_b \sum_{i=o}^{\infty} (1 - d_b)^i\, s_{bt-i} \tag{12}$$

and

$$d_b = \frac{q_b}{2}\left(\sqrt{1 + 4/q_b} - 1\right) \tag{13}$$

We shall call the filtered signals credit ratings. 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 ($d_b > d_c$). 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. It can do so because overall its information is more precise. Note that in equation (12) the bank rating has a lower overall weight on the credit bureau signal than does the credit bureau's rating owing to the aggregation in equation (9), but that the rate at which the credit bureau signal fades from the bank rating is faster to the extent that d_b exceeds d_c, creating a correlation between the bank rating at earlier dates and the credit bureau rating at later dates solely through the public signal.

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:

$$y_{bt|t} = a_{10} + \alpha_{11} y_{ct|t-1} + \alpha_{12} y_{bt|t-1} + e_{1t} \tag{14}$$

$$y_{ct|t} = a_{20} + \alpha_{21} y_{ct|t-1} + \alpha_{22} y_{bt|t-1} + e_{2t} \tag{15}$$

Considering equation (14), 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 $\alpha_{11} = 0$. 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. For equation (15), we again expect the constant term to be zero. However, because of the private information encompassed by bank ratings, the sum of the coefficients of $\alpha_{21} + \alpha_{22}$ should be unity and $\alpha_{22} \geqslant 0$.

In Section 4 we will test two necessary, but not sufficient, conditions for the optimality of credit ratings: that the bank's credit rating for borrowers forecasts the public monitor's credit rating, but 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 equations (14) and (15). 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 therefore derive a general result on the maximum attainable improvement in R^2 in regression equations (14) and (15) from the inclusion of the private information.

The change in the credit bureau's rating can be decomposed into contributions from the new shock to the underlying creditworthiness, u_t; the new shock to the signal, η_{ct}; and the error in the credit bureau's rating at time t-1, $V_{t-1|t-1}$. 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 $V_{t-1|t-1}$. Using (4) and (5), we obtain:

$$d_c^2 V_{t|t} \quad = \quad \tfrac{1}{8} q^2 \sigma^2 \left(\sqrt{1 + 4/q} - 1 \right)^3 \tag{16}$$

Expression (16) implies that the proportion of the movement in the credit bureau's credit rating that can be forecasted based on knowledge of y_{t-1} is $d_c^2 V_{t|t}/\sigma^2$. It can be shown that for $q = .5$, $d_c^2 V_{t|t}$ reaches its maximum at $.25\sigma^2$. This means that if the variance of a public signal is twice that of the underlying creditworthiness, then the maximum reduction in the sum of squared errors one can expect based on knowledge at $t-1$ is .25. For all other values of q, the maximum reduction will be smaller. 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.

In the model, we treat borrower credit ratings as a forecast of the likelihood of default or of the loan's expected value. 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 equations (14) and (15), $\alpha_{11} = 0$. 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 is based on a larger information set than another (on a refinement of the information set) is that it will be able to forecast the movements in the other.

Hypothesis 1. *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.

If a bank has private information, then its ratings should be capable of forecasting the credit bureau's future rating. If it did not do so, then we would have evidence against the joint hypothesis that the bank (i) has private information and (ii) rationally uses this information.

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

Hypothesis 2 *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.

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 (IRB) 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 most of the 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 at the last day of the month.[8] UC produces credit ratings for almost all Swedish businesses. The information on corporate credit balances was and is not incorporated in the credit ratings because the legal framework that governs information reported by banks differs from that for other sources of information. 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.

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.

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 (CB) rating system, we will delete the first two quarters of the bank data sets. This gives us between one and 11 quarterly observations for, on average, roughly 15,000 borrowers in Bank A and one to 12 quarterly observations on 8,000 borrowers in Bank B. Borrowers, incorporated businesses or aktiebolag, have at least the legally required minimum of 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 2 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]

[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, Lindé and

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 adopted the following definition of a default. A firm is given a default status once 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 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 has 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]

In most of our analysis, we will exclude observations where a counterpart has defaulted 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 3 shows that the credit ratings for both lenders are highly concentrated, just as for U.S. large-bank credit ratings. Bank A has some 60 percent of its ratings in its two largest rating categories, while Bank B has roughly the same amount in its largest rating category. The first three columns

Roszbach (2006, 2013).

[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.

of Table 3 demonstrate that Bank A's ratings are not single-peaked. Later on, in Table 6, we will also show that the order of Bank A's ratings does not reflect their risk ranking. Because of these properties, and to bring the system of Bank A more in line with that of Bank B, we have converted the 14 non-bankruptcy grades into a system of seven ratings that is single peaked by grouping ratings 1 to 4, 5 to 7, and 8 to 10, while leaving the remaining, high-risk grades unaffected. This regrouping is shown in the second set of three columns in Table 3.

The credit bureau has five rating classes in addition to a default rating, and a numerically higher rating again implies worse creditworthiness. The default rating is assigned if bankruptcy occurs as defined by the credit bureau above. The distribution of credit bureau ratings is shown in Table 4. 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. Table 4 contains a mapping of bank ratings into credit bureau ratings and vice versa. Between 65 and 80 percent of firms that default on a bank loan have been assigned one of the three worse credit bureau ratings. Reversely, between 34 and 48 percent of firms that go bankrupt have also defaulted on their bank loans.

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. We will make the hypotheses in Section 2.2 operational by testing the informational content of both the bank's internal credit rating and the external credit bureau rating. In doing so, we rely on the fact that the informational content can be normalized because both ratings are efforts to estimate the same underlying variable: the borrower's creditworthiness. In terms of the theoretical model in Section 2.1, this means that the underlying filtered signals will have the same variance if the signals are being optimally forecasted.

In the body of the paper we display only the results from OLS regressions; in the Appendix we also present results from ordered logits. Although the latter are attractive because they allow one to take into account the discrete nature of credit ratings, we 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. Tables 5 through 7 summarize the results from two sets of regressions.

In Section 4.1, we first run OLS regressions for the credit bureau ratings on their lagged values and then add a bank's lagged credit rating. We also check the linearity 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. We then also add the credit bureau's lagged credit rating. Table 9 provides an example of running the same set of variables as in Tables 5 to 7, using an ordered logit model instead of OLS. In Section 4.2, we display the results from several Cox regressions on the default hazard. The results from a series of robustness tests are discussed, but the tabulated results are presented only in Appendix 1.

4.1 Testing hypotheses 1 and 2

If we define r_{bt} as the rating of the bank at t and r_{ct} as the rating of the credit bureau at t then, under the assumptions in Section 2.1, equations (14) and (15) translate into the following regressions we can estimate:

$$r_{bt} = \alpha_{1b} r_{bt-1} + \beta_{1b} r_{ct-1} + \varepsilon_{1bt} \tag{17}$$

Because we explicitly wish to test for the marginal informational value of adding a lag of the credit bureau rating, we will also estimate the simple autoregressive form

$$r_{bt} = \alpha_{2b} r_{bt-1} + \varepsilon_{2bt} \tag{18}$$

In a similar fashion, we will estimate two regressions explaining the credit bureau rating updating process:

$$r_{ct} = \beta_{1c} r_{ct-1} + \alpha_{1c} r_{bt-1} + \varepsilon_{1ct} \tag{19}$$

$$r_{ct} = \beta_{2c} r_{ct-1} + \varepsilon_{2ct} \tag{20}$$

In a strict sense, Hypothesis 1 in Section 2.2 implies that $\alpha_{1b} = 1$ and $\beta_{1b} = 0$. This is what we would expect of an optimal bank forecast if it were continuous. Unfortunately, the rating data we have available for testing Hypotheses 1 are discrete instead of continuous. Under these conditions, we can no longer be sure that both parameter restrictions will hold. We will therefore allow for a constant in the empirical equivalent of equation (17) 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 residual sum of squares (RSS) in the regression (17).

Hypothesis 2 in a strict sense implies that $\alpha_{1c} + \beta_{1c} = 1$ and $\alpha_{1c} > 0$. For the same reasons mentioned above for Hypothesis 1, we will test a weaker rather than the stricter version of the hypothesis, namely that $0 < \alpha_{1c} \leq 1$. Under this hypothesis, the bank rating does forecast changes in the credit bureau rating and has a significant impact on the RSS in regression equation (19).

Each of the three panels in Table 5 shows the main results we will discuss in sections 4.1.1 and 4.1.2, using data on borrowers from both Bank A and Bank B. Of the six regressions in each panel, four are the empirical equivalents of equations (17) to (20), i.e., they take into account the above mentioned limitations of discrete rating data. The remaining columns (3, 6; 9, 12; and 15, 18) are variations where we have included dummy explanatory variables for the lagged credit ratings instead of a simple one-period lag, in order to allow for nonlinearities in the impact on the dependent variable. To verify that our results are robust to variations in firm size, we also repeat the regressions while grouping data by small, medium-sized, or large firms. These results are presented in Appendix Tables $1A - C$, $2A - C$, and $3A - C$. In Table 9, we verify the robustness of our findings in Tables 5 to 7 with respect to estimation method by applying ordered logit instead of OLS.[11] Thereby we allow the ordering of the relevant dependent rating variable to occur in a nonlinear fashion with respect to the information in the explanatory variables. By also including dummy variables in the ordered logit models, we attempt to control for the widest range of nonlinearities in the data. Later, in Section 4.1.3, we present some additional robustness tests that verify to what extent the discrete nature and staggered updating of ratings influence our results.

Hereafter we will focus on results from the "full" regressions and refer to the subsets only when differences occur. When contrasting the results in each part of Table 5, we will compare differences in the RSS across regressions.

[11] We ran additional ordered logits on Bank B and by size of borrower for both banks. These results are available upon request.

4.1.1 Hypothesis 1

When we consider the results for equations (17) and (18), it is clear that, with between 12,000 and 200,000 observations, even small coefficients are significant. For both banks, we obtain highly statistically significant coefficients for the first lag of the credit bureau rating in regressions with a bank credit rating as the dependent variable (Table 5, columns 5, 11, and 17). This result is robust to transformations of the rating scale (i.e., moving from panel 1 to panel 2 of Table 5) and to variation in firm size (Table 7) and is independent of the estimation method (Table 9).[12] We also ran regressions where we replace the lagged dependent variable with lagged dummy variables. However, doing so invariably worsened the fit of the regression (results are not displayed here, but are available upon request).

The coefficients on the lag of the credit bureau ratings are in the order of .01-0.2 in the OLS regressions for Bank B (Table 5, column 17) and in the range of 0.05 to 0.10 for Bank A (Table 5, columns 5 and 11). Even taking into account the different scales that the two banks employ, this suggests that credit bureau ratings are more informative for predicting ratings in Bank A than in Bank B. In columns (4) to (6) of Table 5, we see that Bank A credit ratings remain relatively forecastable even when they are compressed, although not as much as the uncompressed ratings. Typically, adding lagged credit bureau ratings to the regression (moving from column 4 to column 5 in Table 5) reduces the RSS by more than when a lag of Bank A's rating is added to a regression on the credit bureau rating (moving from column 1 to column 2 in Table 5).

The observation that Bank A ratings are less informative is confirmed by the results in Table 7. There, we summarize the additional explanatory power of lagged credit bureau ratings when these are added to a regression of bank credit ratings on their own one-quarter lag. For example, the number 2.67 in column 5 of Table 7 equals the percentage decrease in RSS when moving from column 4 to column 5 in Table 5. Depending on the size of the borrowers, adding credit bureau ratings explains an additional 2.08 to 3.00 percent of the RSS for Bank A (Table 7, column 3), compared with .44 to .92 percent for Bank B (Table 7, column 4). When used to explain Bank A ratings (Table 7, column 3), credit bureau ratings are most informative in predicting small business ratings. An inspection of the corresponding results for Bank B in Table 7 reinforces this picture. Adding one lag of the Bank B rating lowers the RSS of the

[12] The full firm size OLS and ordered logit regressions are presented in Appendix Tables 1 to 3 and 4 to 6. The Appendix is available at www.riksbank.se/research/roszbach and www.phil.frb.org/research-and-data/economists/nakamura/.

credit bureau regression substantially more than adding the same lag of the credit bureau rating lowers Bank B's rating RSS (compare column 2 with column 4). This holds both for the full sample of borrowers and for all three subsamples. Columns (1) and (2) of Table 7 also make it clear that Bank B ratings are more informative than Bank A ratings with respect to the credit bureau ratings, as adding the former reduces the RSS by more than adding the latter does. The ordered logit regressions in columns (4) to (6) of Table 9 broadly confirm the findings in the OLS regressions, i.e., the results are robust to varying the estimation method.[13]

When reading Table 7, marginal contributions in a range between .44 and 3.00 percent may at first sight suggest that neither bank nor credit bureau ratings are particularly informative and that any conclusions from these ratings should be downplayed. However, bank and credit bureau ratings, both being predictors of future default risk, are constructed using a set of risk factors that is – or at least should be – substantially overlapping.[14] Public credit ratings are or should be based on all publicly available information, while internal bank credit ratings are based on public information and private information. As a consequence, a regression of any of these ratings on a lag of itself or the other rating will by construction produce only a relatively small marginal increase in the R^2 or Pseudo-R^2 when the lag of the other rating is added. The size of the marginal increase in the R^2 or Pseudo-R^2 when adding the bank rating to an autoregression of the credit bureau rating can be thought of as the contribution of private information. When adding the credit bureau rating to an autoregression of the bank rating, it can be thought of as the efficiency loss of the bank rating. The relative size of these two marginal effects provides a means to benchmark efficiency gains and losses in the collection and processing of information in the production of credit ratings.

Unfortunately, there is no statistical test to evaluate whether the increases in RSS presented in Table 7 are statistically significant. We do provide some complementary evidence that supports our assertion that credit bureau and bank ratings 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

[13] In additional regressions that 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 7, 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] See Jacobson et al. (2006) for evidence on bank ratings and Jacobson et al.(2013) for evidence on bankruptcy data.

a second rating. We present these results in panels I to IV of Table 8.[15] A set of chi-square tests of equality of the areas under the ROC curve show that the discriminatory accuracy of the model increases significantly in all regressions for the full samples and for seven out of nine sub-samples used in Table 7.

Overall, the above findings constitute distinct evidence against the hypothesis that bank ratings are fully efficient and are not predicted by lagged credit bureau ratings. Moreover, the results indicate that this holds all the more for Bank A, which has ratings that are relatively less informative.

4.1.2 Hypothesis 2

When examining the robust t-statistic on the lag of the bank rating in a regression of the contemporaneous credit bureau rating, we again find highly significant positive coefficients in all cases. As before, this finding is robust to variations in firm size, to transformations of the rating scale (first part to the second part of Table 5) and to varying the estimation method (Table 5 versue Table 9) and is stable across banks (panels 1 and 2 of Table 5 versus panel 3 of Table 5).[16]

As in Section 4.1.1 we verify that the results are robust to an exchange of the lagged bank rating by a set of lagged rating dummies. The results of these regressions are shown in columns (3), (9), and (15) of Table 5, and the individual coefficients on the Bank A rating dummies are displayed in Table 6. Evidently, there is nonlinear information in the Bank A ratings. Unfortunately, the coefficients in Table 6 turn out to be non-monotonically increasing in the rating. In other words, the improvement in the regression RSS is caused in part by the fact that the order of the ratings does not properly reflect the risk ranking as measured by the credit bureau ratings. The coefficients for Bank A rating grades 5 and 8 are, for example, significantly greater than for the two following ratings, i.e., grades 6 to 7 and 9 to 10 respectively. The additional explanatory power of the Bank A rating dummies is thus due to rating differences that do not correspond to their ordinal rank! This is strong prima facie evidence that Bank A's ratings are not adequately capturing relative risk and that worse bank credit ratings sometimes correspond to improved credit bureau ratings. It can then hardly be expected that these bank

[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.

[16] Firm-size regressions are available in Appendix Tables 1 to 3.

credit ratings are strictly ordinally related to an underlying optimal measure of creditworthiness in any appropriate way. Thus our decision to compress the ratings seems justified.

Some interesting differences can be observed between the banks. For example, if we add the lagged Bank A rating in an OLS regression of the credit bureau rating on its own lag, then the RSS drops from 55575 (column 1, Table 5) to 55237 (column 2, Table 5), a reduction of less than 0.6 percent. Interestingly, when adding the credit bureau rating to a regression of the Bank A credit rating on its own lag, the RSS falls from 174853 to 172284, a decrease of 1.5 percent. Thus, over the entire portfolio, the credit bureau appears to have better information than the bank since its ratings have a proportionally bigger impact on the error.

Above, we already argued that the uncompressed Bank A ratings suffer from some sub-optimality. The large degree of forecastability of the Bank A credit ratings offered additional evidence in this direction. As mentioned earlier, columns (5) to (6) in Table 5 show that Bank A credit ratings are relatively well forecastable by public credit bureau ratings. By contrast, appending the lag of the credit bureau rating to a regression on the Bank B rating in Table 5 reduces the RSS by only 0.8 percent. However, adding the lagged Bank B rating reduces the RSS of the credit bureau rating regression by 1.3 percent. Bank B thus has relatively better information than the credit bureau. Ordered logit regressions presented in the Appendix Tables $A4$, $A5$, and $A6$ show that these findings are not sensitive to the estimation method used. Even here, Bank B appears as a relatively better rater.[17]

On the whole, the above findings offer strong evidence in support of the hypothesis that the banks in our sample have private information and that their internal ratings predict credit bureau ratings. We also corroborate our earlier conclusion that Bank A ratings appear less informative than Bank B ratings. The fact that Bank A ratings are not monotonically increasing in risk provides a possible rationale for the differences between Bank A and Bank B.

Thus far, we have used only a one-quarter-lag prediction period. We next verify if our findings are robust to increasing the prediction period to two quarters, three quarters, or one

[17] The results in the ordered logit regressions resemble those in the OLS regressions. Consistent with our earlier findings, we see in Appendix Tables $4A-D$, $5A-D$, and $6A-D$ that Bank A is not as apt a rater as Bank B is. For small borrowers, a regression of the credit bureau rating on its own lag gives a pseudo-R^2 of .5032, and adding the lag of the Bank A compressed rating raises the pseudo-R^2 by .0025 to .5057 (Table A5, columns 1 and 2). By comparison, the regression of Bank A's compressed rating on its own lag gives a pseudo R^2 of approximately .6815. Adding the lagged information present in the bureau rating improves the fit, by .0059 to .6874 (columns 4 and 5). Although the contrast is not as clear as in the OLS regressions, the ordered logit regressions offer little evidence that Bank A's information collection and processing are superior to that by the credit bureau. As in the OLS regressions, the same image that Bank B is a relatively better rater emerges from Appendix Tables $6A-D$. Adding its lag increases the pseudo-R^2 of the autoregression forecasting the credit bureau rating by .0047, from .4985 to .5032 (Table $A6-A$, columns 1 and 2). By contrast, adding the credit bureau lag to the regression forecasting the Bank B credit rating raises it only by .0022.

year. In Table 10, we compare the information in Bank A's credit rating for predicting the credit bureau rating two or more quarters ahead, and vice versa, as measured by the reduction in mean square error. We have included the one-quarter-lag results from Table 7 as a baseline comparison. As we extend the prediction period from one quarter to four quarters, increasing by steps of a quarter, the same qualitative pattern continues to hold and, indeed, becomes exaggerated: Bank A looks ever worse vis-a-vis the credit bureau in rating both small, medium-sized, and large borrowers as the prediction horizon becomes longer. On the one hand, we see in the upper panel of Table 10 that Bank A maintains its ability to predict the credit bureau rating as we lengthen the prediction horizon. On average, the RSS improvement from using Bank A ratings rises, but only modestly and never doubles (Table 10, column 5 versus column 1). On the other hand, the value of the credit bureau ratings in predicting Bank A ratings rises markedly and more than doubles when we lengthen the prediction horizon to four quarters (Table 10, column 8 versus column 5). In particular, note that for small borrowers, the credit bureau rating is able to reduce the RSS of prediction of the Bank A rating by nearly 9 percent! The predictability of Bank A credit ratings for small- and medium-sized borrowers rises markedly; Bank A continues to do a good job of credit-rating large borrowers.

If we look at the superiority of Bank B's ratings over the credit bureau ratings, as shown in Table 11, columns 1 and 5, we see that the superiority remains more or less intact but diminishes somewhat. Indeed, the credit bureau's ratings appear modestly superior to Bank B's for medium-sized borrowers at lags of two quarters and longer. Overall, the results of extending the prediction period confirm our results for one-quarter lags: The bank ratings continue to have predictive power for the credit bureau ratings, but the credit bureau ratings' predictive power for the bank ratings becomes even stronger.

4.1.3 Staggering of information and rating coarseness

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. 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 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 7, 10, and 11, this would generate 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 observations for which it is possible that information sets have not 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. This we can do for the credit bureau data only.

Staggering of information First, we repeat the regressions underlying columns (4) and (5) 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 has collected information to review the credit ratings, but we can observe if bank ratings have just been modified. [18] Because credit ratings can be adjusted only after a loan officer has updated and filed client information, limiting regressions to these observations eliminates any risk 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, rating changes follow four quarters after a rating change, five times as often as other intervals. For Bank B, rating changes follow four quarters after a rating change, two to three times as often. 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 predict a subsequent rating change.

The results in columns (1) to (4) of Table 12 show that lagged credit bureau ratings still have explanatory power for both Bank A and Bank B credit ratings. In line with our earlier findings, 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 sub-sample size. Columns (9) to (12) show the four-quarter-lag rating changes unconditioned, for comparison.

For a second test, we repeat the regressions underlying columns (1) and (2) in Table 5 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 12 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.

Overall, these tests demonstrate that the staggering of information updating by the credit

[18] We follow the approach of Bils, Klenow, and Malin (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.

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 inefficient measures of creditworthiness.

Coarseness of the rating scale For a last test, we investigate whether using discrete instead of continuous credit bureau ratings alters 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 re-run the regressions of columns 1 and 2 in Table 5 using the continuous measure of creditworthiness as a dependent variable and its lag and the lagged discrete bank rating data as explanatory variables. In Table 12, columns (13) and (14) we see that bank credit ratings continue to have predictive power for credit bureau ratings, even when the latter are continuous. We take this as strong evidence that our earlier finding that banks have private information not embedded in the credit bureau ratings is not an artifact of credit bureau ratings being discrete transforms of continuous risk measures.

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 may not result in substantial inefficiency. In Section 5 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.

4.2 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 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 the above proposition holds, we perform an additional test on the data and compare the explanatory power of bank credit ratings and credit bureau ratings in a duration model setting. We implement the test by estimating the following Cox proportional hazards model:

$$log\ h_i(t) = \alpha(t) + \beta x_{it} + \varepsilon_{it} \tag{21}$$

or equivalently

$$h_i(t) = h_0(t)\ exp(\alpha(t) + \beta x_{it} + \varepsilon_{it}) \tag{22}$$

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 $\mathbf{x}$ 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, displayed in Table 13, 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 let $x_{it} = r_{c,t-1}$ to compute the explanatory power of lagged credit bureau ratings for borrowers in both Bank A and Bank B (Table 13, columns 1 and 7). Next, we take $x_{it} = r_{b,t-1}$ (columns 2 and 8) and $x_{it} = (r_{c,t-1}, r_{b,t-1})$ (columns 3 and 9). In columns (4) to (6) and (10) to (12) we present results from regressions where we instead use dummy variables for the ratings and let

$$\mathbf{x}_{it} = \left[DUM\_r^1_{c,t-1}, DUM\_r^2_{c,t-1}...., DUM\_r^{G-1}_{c,t-1}\right] \tag{23}$$

,

$$\mathbf{x}_{it} = \left[DUM\_r^1_{b,t-1}, DUM\_r^2_{b,t-1}...., DUM\_r^{K-1}_{b,t-1}\right] \tag{24}$$

or

$$\mathbf{x}_{it} = \Big[DUM\_r^1_{c,t-1}, DUM\_r^2_{c,t-1}...., DUM\_r^{G-1}_{c,t-1},$$
$$DUM\_r^1_{b,t-1}, DUM\_r^2_{b,t-1}...., DUM\_r^{K-1}_{b,t-1}\Big] \tag{25}$$

respectively, where $DUM\_r^g_{b,t-1} = 1$ if $r^g_{b,t-1} = g$ and zero otherwise, and G and K are the number of grades in the respective rating systems.

The log likelihood values in columns (1) and (2) of Table 13 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 (columns 4 and 5). The table also shows that qualitatively equal results are obtained when using Bank B ratings instead (columns 7, 8, 10, and 11). The Appendix Table $A7$ contains output from an additional robustness test, where we instead use a second lag instead of the first lag. This does not change the results qualitatively. As one would expect, 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.

Next, in Table 14, we present the results from a similar set of Cox regressions where the dependent variable is the instantaneous risk of a *loan* default in Bank i, $i = \{\mathrm{A}, \mathrm{B}\}$, at time t, conditional on survival to that time. A similar comparison between columns (1) and (2) and (7) and (8) makes it clear that, for both banks, lagged credit bureau ratings are better at explaining bank default hazards than are bank ratings themselves. In the Appendix Table $A8$ we again find these results are robust to exchanging the first lag with the second lags of the explanatory ratings. When we replace the lagged explanatory variables with a set of dummy variables the credit bureau ratings maintain their edge (Table 14, columns 4, 5, 10, and 11). The predictive power of the Bank B ratings may be somewhat blurred, however, by the fact that several dummies drop out of the regression because of the small number of bankruptcies that occur in transitions from rating grades 2, 3, and 6.

The results in Table 14 also illustrate how the nonlinearities in both bank and credit bureau ratings come into play in our analysis. Columns 1 and 2 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. Together with the results in Table 13, this might have suggested that credit bureau ratings are *in general* more informative than bank credit ratings.

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 4 and 5). 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. Columns (4) and (5) show that this loss of information is greater when using bank ratings than credit bureau ratings, most likely because the former are less persistent.

In Table 15, we present the log likelihoods of the regressions that include the credit bureau ratings alone, the bank ratings alone, and both the credit bureau ratings and the bank ratings together. We have marked the significance of the likelihood ratio tests for exclusion of the bank rating in the credit bureau autoregression, as well as for exclusion of the credit bureau rating in the bank rating autoregression. 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 is 55.0, making the Bank A rating very significant in a chi-square test with one degree of freedom. As can be seen, 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. In particular, this provides striking 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. In Appendix 1, Tables $A9$ to $A11$, we provide additional results on the log likelihoods and exclusion tests for subsets of small, medium, and large borrowers. An interesting observation to be made from those tables is that, particularly for the longer-lag horizon, credit bureau ratings perform notably better than bank ratings when predicting defaults for small borrowers (Table $A9$, columns (5) and(6)), while bank ratings have greater predictive power for large borrowers (Table $A11$, columns (5) and (6)).

Finally, we provide supporting evidence that all the above mentioned increases in explanatory power, when a second rating is added to the regression, are not only statistically but also economically significant. The results in Table 8, panels I and III, make clear that adding the first lag of a second rating to an autoregressive logit model increases the area under the ROC curve significantly, both in sample and out of sample.[19] An additional way to assess the quanti-

[19]Because no comparable test is available for the Cox model, we instead use a logit model of default for this

tative importance of using a second rating when predicting default risk is provided by any shifts in the implied default probability distribution when adding a second rating. Table 8, panels II and IV, shows that the default probability distribution also changes significantly when we add a second rating as an explanatory variable in an autoregressive logit model.

5 Overconfidence

In Section 4 we found that credit bureau ratings are able to predict bank ratings. One 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. In order to account for such a possibility, we will in Section 5.1 modify our benchmark model to include the possibility that loan officers believe that their information is more valuable than it in fact is– which amounts to the loan officer ascribing an overly high precision to the private signal. In Section 5.2 we will also test the implications from a model that incorporates overconfidence on our data.

5.1 Information processing with overconfidence

A bank loan officer who is overconfident in the private signal will believe its precision is αq_{pb} with $\alpha > 1$, although the true precision is q_{pb}.

The consequence of overconfidence will be that the two signals will be aggregated with

$$\begin{aligned} \widetilde{s}_{bt} &= \left(\alpha q_{pb} s_{pbt} + q_\eta s_{ct}\right) / \left(\alpha q_{pb} + q_\eta\right) \\ &= y_t + \widetilde{\eta}_{bt} \end{aligned}$$

where $\widetilde{\eta}_{bt} = \frac{\alpha q_{pb}\eta_{pbt} + q_\eta \eta_{ct}}{\alpha q_{pb} + q_\eta}$. This will deviate from the optimal signal

$$\begin{aligned} s_{bt} &= \frac{q_{pb} s_{pbt} + q_\eta s_{ct}}{q_{pb} + q_\eta} \\ &= y_t + \eta_{bt} \end{aligned}$$

particular test.

where $\eta_{bt} = \frac{q_{pb}\eta_{pbt}+q_\eta\eta_{ct}}{q_{pb}+q_\eta}$ by the following amount

$$
\begin{aligned}
\widetilde{\eta}_{bt}-\eta_{bt} &= \frac{\alpha q_{pb}\eta_{pbt}+q_\eta\eta_{ct}}{\alpha q_{pb}+q_\eta}-\frac{q_{pb}\eta_{pbt}+q_\eta\eta_{ct}}{q_{pb}+q_\eta} \\
&= \frac{\left(\alpha q_{pb}\eta_{pbt}+q_\eta\eta_{ct}\right)\left(q_{pb}+q_\eta\right)-\left(q_{pb}\eta_{pbt}+q_\eta\eta_{ct}\right)\left(\alpha q_{pb}+q_\eta\right)}{\left(\alpha q_{pb}+q_\eta\right)\left(q_{pb}+q_\eta\right)} \\
&= \frac{\alpha q_{pb}^2\eta_{pbt}+\alpha q_{pb}q_c\eta_{pbt}+q_\eta^2\eta_{ct}+q_bq_\eta\eta_{ct}-\left(\alpha q_{pb}^2\eta_{pbt}+q_{pb}q_\eta\eta_{pbt}+\alpha q_{pb}q_\eta\eta_{ct}+q_\eta^2\eta_{ct}\right)}{\left(\alpha q_{pb}+q_\eta\right)\left(q_{pb}+q_\eta\right)} \qquad (26)\\
&= \frac{\alpha q_{pb}q_c\eta_{pbt}+q_bq_\eta\eta_{ct}-\left(q_{pb}q_\eta\eta_{pbt}+\alpha q_{pb}q_\eta\right)}{\left(\alpha q_{pb}+q_\eta\right)\left(q_{pb}+q_\eta\right)} \\
&= \frac{(\alpha-1)\left(\eta_{pbt}-\eta_{ct}\right)}{\left(\alpha+\frac{q_\eta}{q_{pb}}\right)\left(\frac{q_{pb}}{q_\eta}+1\right)} \qquad (*)
\end{aligned}
$$

We will rewrite $(*)$ in equation (26) as $\gamma\left(\eta_{pbt}-\eta_{ct}\right)$, where $\gamma = \frac{\alpha-1}{\left(\alpha+\frac{q_\eta}{q_{pb}}\right)\left(\frac{q_{pb}}{q_\eta}+1\right)}$. It can be shown that γ is increasing in α, for all values of α, so that the random noise added to the loan officer's estimate of creditworthiness is increasing in overconfidence, as one would expect.

The overconfident loan officer will construct an estimate of creditworthiness in which this error is further weighted too highly by setting the Kalman filter coefficient ($\widetilde{d}_b$) too high. The loan officer's estimate of creditworthiness will initially have impounded into it the product of the loan officer's misaggregated signal $\widetilde{s}_{bt}$, multiplied by the Kalman filter coefficient. The misestimated creditworthiness in the period the signal is received will contain $\widetilde{d}_b\widetilde{\eta}_{bt}$ and in the next period $\widetilde{d}_b\left(1-\widetilde{d}_b\right)\widetilde{\eta}_{bt}$ so that this measure of creditworthiness will fall by $\widetilde{d}_b^2\widetilde{\eta}_{bt}$, Of the latter expression, a part $\widetilde{d}_b^2\gamma\left(\eta_{pbt}-\eta_{ct}\right)$ is pure noise relative to the optimal measure of creditworthiness. This implies there will be reversion toward the mean in case of overconfidence; it will also be forecastable and the first differences of credit ratings will be negatively serially correlated.

The aggregated signal will have variance equal to

$$
\begin{aligned}
E(y_t-\ \widetilde{s}_{bt})^2 &= E(y_t-\ (y_t+\widetilde{\eta}_{bt}))^2 \\
&= E\left(\frac{\alpha q_{pb}\eta_{pbt}+q_\eta\eta_{ct}}{\alpha q_{pb}+q_\eta}\right)^2 \\
&= E\left(\frac{\alpha q_{pb}\eta_{pbt}}{\alpha q_{pb}+q_\eta}\right)^2+E\left(\frac{q_\eta\eta_{ct}}{\alpha q_{pb}+q_\eta}\right)^2 \\
&= \left(\frac{\alpha q_{pb}}{\alpha q_{pb}+q_\eta}\right)^2\frac{1}{q_{pb}}+\left(\frac{q_\eta}{\alpha q_{pb}+q_\eta}\right)^2\frac{1}{q_\eta} \\
&= \frac{\alpha^2q_{pb}+q_\eta}{\left(\alpha q_{pb}+q_\eta\right)^2}
\end{aligned}
$$

This noise will be greater than the variance of the optimal weighting:

$$\begin{aligned}
E\left(y_t - \ s_{bt}\right)^2 &= E\left(y_t - \ \left(y_t+\eta_{bt}\right)\right)^2 \\
&= E\left(\frac{q_{pb}\eta_{pbt}+q_\eta\eta_{ct}}{q_{pb}+q_\eta}\right)^2 \\
&= E\left(\frac{q_{pb}\eta_{pbt}}{q_{pb}+q_\eta}\right)^2 + E\left(\frac{q_\eta\eta_{ct}}{q_{pb}+q_\eta}\right)^2 \\
&= \left(\frac{q_{pb}}{q_{pb}+q_\eta}\right)^2\frac{1}{q_{pb}} + \left(\frac{q_\eta}{q_{pb}+q_\eta}\right)^2\frac{1}{q_\eta} \\
&= \frac{1}{q_{pb}+q_\eta}
\end{aligned}$$

However, the loan officer will believe that the signal variance is actually:

$$\frac{1}{\alpha q_{pb}+q_\eta} < \frac{1}{q_{pb}+q_\eta} < \frac{\alpha^2 q_{pb}+q_\eta}{\left(\alpha q_{pb}+q_\eta\right)^2}$$

The extra noise, relative to the optimal signal, is:

$$\begin{aligned}
Extra\ noise &= \frac{\alpha^2 q_{pb}+q_\eta}{\left(\alpha q_{pb}+q_\eta\right)^2} - \frac{1}{q_{pb}+q_\eta} \\
&= \frac{\left(\alpha^2 q_{pb}+q_\eta\right)\left(q_{pb}+q_\eta\right)-\left(\alpha q_{pb}+q_\eta\right)^2}{\left(\alpha q_{pb}+q_\eta\right)^2\left(q_{pb}+q_\eta\right)} \\
&= \frac{\alpha^2 q_{pb}^2+\alpha^2 q_\eta q_{pb}+q_\eta q_{pb}+q_\eta^2-\left(\alpha^2 q_{pb}^2+2\alpha q_{pb}q_\eta+q_\eta^2\right)}{\left(\alpha q_{pb}+q_\eta\right)^2\left(q_{pb}+q_\eta\right)} \\
&= \frac{\alpha^2 q_\eta q_{pb}+q_\eta q_{pb}-2\alpha q_{pb}q_\eta}{\left(\alpha q_{pb}+q_\eta\right)^2\left(q_{pb}+q_\eta\right)} \\
&= \frac{(\alpha-1)^2 q_\eta q_{pb}}{\left(\alpha q_{pb}+q_\eta\right)^2}\left(\frac{1}{q_{pb}+q_\eta}\right) \\
&= \frac{(\alpha-1)^2\frac{q_\eta}{q_{pb}}}{\left(\alpha+\frac{q_\eta}{q_{pb}}\right)^2}\left(\frac{1}{q_{pb}+q_\eta}\right) \qquad (**)
\end{aligned} \tag{27}$$

The fractional part of (**) in equation (27) has a maximum less than $\frac{q_\eta}{q_{pb}}$ since $\frac{(\alpha-1)^2}{\left(\alpha+\frac{q_\eta}{q_{pb}}\right)^2}$ obviously can never attain 1. The first derivative is positive, but above a certain number the second derivative is negative.

The true noise of the overconfident officer's signal will be worse than the noise of the signal of the credit bureau if

$$\begin{aligned}
\frac{\alpha^2 q_{pb}+q_\eta}{\left(\alpha q_{pb}+q_\eta\right)^2} &> \frac{1}{q_\eta} &&\Rightarrow \\
\left(\alpha q_{pb}+q_\eta\right)^2 &< \alpha^2 q_{pb}q_\eta+q_\eta^2 &&\Rightarrow \\
\alpha^2 q_{pb}^2+2\alpha q_{pb}q_\eta &< \alpha^2 q_{pb}q_\eta &&\Rightarrow \\
\alpha q_{pb}+2q_\eta &< \alpha q_\eta &&\Rightarrow \\
\frac{q_{pb}}{q_\eta} &< \frac{\alpha-2}{\alpha}
\end{aligned}$$

This will happen only if the loan officer thinks the private signal's precision is more than twice as large than it actually is, and if the precision of the private signal is less than that of the public signal.

The loan officer will set the Kalman filter coefficient $\widetilde{d_b}$ to equal

$$\begin{aligned} \widetilde{d_b} &= \tfrac{\widetilde{q_b}}{2}\left(\sqrt{1+4/\widetilde{q_b}}-1\right) \\ &= \tfrac{\alpha q_{pb}+q_\eta}{2}\left(\sqrt{1+4/\left(\alpha q_{pb}+q_\eta\right)}-1\right) \end{aligned}$$

which is too large, with the size relative to the optimal d_b depending on θ, α and q_η The ratio $\widetilde{d_b}/d_b$, where d_b is the optimal Kalman filter coefficient, is not monotonic in θ, and for $q_\eta > 0.01$, does not rise above 2.

An implication from the model is thus that credit ratings of overconfident loan officers, by including too much noise, will tend to mean-revert toward the true signal because the rating is overshooting in the direction of the private signal. Hence there will tend to be negative serial correlation in the changes of the bank rating. Another way to think of this is that the over-confident officer places too much weight on current information – and subsequent information will cause a swing back toward the true underlying signal.

One can use the same model, but with $\alpha < 1$, if the bank loan officer is underconfident in the private signal and the weight on it will be too low. The true noise 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 be greater than that of the credit bureau. Thus the credit bureau's estimate will not forecast the estimate of the bank. 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.

In short, the theory of overconfidence that we laid out above implies that the overconfidence of the bank loan officer in his or her private information results in future predictability of the bank credit rating, as a consequence of the weighting of the noisy private signal. Such overconfidence leads to excessive movement of the credit rating in the direction of the private signal, followed by a mean reversion. This mean reversion is something we can directly detect in a difference regression: The last period's change in rating should predict this period's change in rating with a sign reversal. We will test this in Section 5.2.

5.2 Testing for the presence of overconfidence

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

$$\Delta r_{bt} = a + b\Delta r_{bt-1} + u_t \tag{28}$$

If we find that $\widehat{b} < 0$, this will imply evidence of overconfidence, while $\widehat{b} > 0$ would imply evidence of underconfidence.

We run the difference regression (28) 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 16 show these for the four lag lengths, for all Bank A and 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 5.1 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, 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 17 show that the mean-reversion coefficients are uniformly negative for the bank ratings. For Bank A's compressed ratings (Table 17, upper panel), mean reversion is strongest for the small borrowers and for the longer lags. It is also strong for medium-sized borrowers at longer lags, and weakest for the large borrowers. This is consistent with our earlier observation that Bank A's ratings are more predictable for small borrowers and for longer lags, as well as relatively more predictable for medium-sized borrowers than for large borrowers.

We also find that the absolute size (signed size) of these coefficients is positively (negatively) related to the inefficiency of the bank credit rating (i.e., the MSE improvement when we add the credit bureau rating). 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.

6 Conclusion

Using data from two large sophisticated 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.

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 for both groups 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? What determines how much soft information banks collect on their customers and how they aggregate soft and hard information? These questions are important issues for future research to address.

References

[1] Agarwal, Sumit, and Robert Hauswald, 2010, "Distance and Private Information in Lending," *Review of Financial Studies* 23(7), pp. 2757-2788.

[2] Altman, Edward I. and Herbert A. Rijken, 2004, "How Rating Agencies Achieve Rating Stability," *Journal of Banking and Finance* 28 (November), pp. 2679-2714.

[3] Berger, Allen, Sally Davies and Mark Flannery, 2000, "Comparing Market and Supervisory Assessments of Bank Performance - Who Knows What When," *Journal of Money, Credit and Banking* 32 (3) (August), pp.641-667.

[4] Bils, Mark, Peter J. Klenow and Benjamin A. Malin, 2009, "Reset Price Inflation and the Impact of Monetary Shocks," NBER Working Paper 14787, March.

[5] Cantor, Richard, 2004, "An Introduction to Recent Research on Credit Ratings," *Journal of Banking and Finance* 28 (November), pp. 2565-2573.

[6] Carey, Mark, and Mark Hrycay, 2001, "Parameterizing Credit Risk Models with Rating Data," *Journal of Banking and Finance* 25, pp.197-201.

[7] Chen, Qi, and Wei Jiang, 2006, "Analysts' Weighting of Public and Private Information," *Review of Financial Studies* 19, pp. 319-355.

[8] Chow, Gregory C., 1975, *Analysis and Control of Dynamic Economic Systems*, New York: John Wiley and Sons.

[9] Claessens, Stijn, and Geert Embrechts, 2003, "Basel 2, Sovereign Ratings and Transfer Risk External versus Internal Ratings," manuscript, ssrn.com/abstract=386480.

[10] Das, Sanjiv, Darrell Duffie, Nikunj Kapadia, and Leandro Saita, 2007, "Common Failings: How Corporate Defaults are Correlated," *Journal of Finance* 62(1), pp. 93-118.

[11] Dell'Ariccia, Giovanni, and Robert Marquez, 2004, "Information and Bank Credit Allocation," *Journal of Financial Economics* 72, pp. 185–214.

[12] Diamond, Douglas, 1984, "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies,* 51, pp. 393-414.

[13] Fama, Eugene F., 1985, "What's Different About Banks?" *Journal of Monetary Economics*, 15, pp. 29-40.

[14] Grunert, Jens, Lars Norden, and Martin Weber, 2005, "The Role of Non-financial Factors in Internal Credit Ratings," *Journal of Banking and Finance* 29, pp. 509-531.

[15] Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini, 2010, "Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation," *Journal of Finance 65*, (June), pp. 795-828.

[16] Jacobson, Tor, Jesper Lindé, and Kasper Roszbach, 2006, "Internal Ratings Systems, Implied Credit Risk and the Consistency of Banks' Risk Classification Policies," *Journal of Banking and Finance 30*, pp.1899-1926

[17] Jacobson, Tor, Jesper Lindé, and Kasper Roszbach, 2013, "Firm Default and Aggregate Fluctuations," *Forthcoming in: Journal of the European Economic Association.*

[18] Kahneman, Daniel, *Thinking Fast and Slow,* 2011, New York: Farrar, Strauss and Giroux.

[19] Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2009, "Financial Regulation and Securitization: Evidence From Subprime Loans," *Journal of Monetary Economics*, 56 (5) (July), pp. 700-720.

[20] Krahnen, Jan Pieter, and Martin Weber, 2001, "Generally Accepted Rating Principles: A Primer," *Journal of Banking and Finance* 25, pp. 3-23.

[21] Löffler, Gunter, 2004, "Ratings versus Market-based Measures of Default Risk in Portfolio Governance," *Journal of Banking and Finance* 28 (November), pp. 2715-2746.

[22] Lummer, Scott L., and John J. McConnell, 1989, "Further Evidence on the Bank Lending Process and the Capital Market Response to Bank Loan Announcements," *Journal of Financial Economics* 25, pp.99-122.

[23] Mester, Loretta J., Leonard I. Nakamura, and Micheline Renault, 2007, "Transactions Accounts and Loan Monitoring," *Review of Financial Studies* 20 (May), pp. 529-556.

[24] Povel, Paul, Rajdeep Singh, and Andrew Winton, 2007, "Booms, Busts, and Fraud," *Review of Financial Studies* 20 (4), pp. 1219-1254.

Table 2: Descriptive statistics on loans outstanding

The table contains descriptive statistics on actually utilized credit in Banks A and Bank B. All numbers are averages over four years, i.e., over the period 1997-Q3 to 2000-Q1

	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

Table 3A: 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. Default classes are excluded and shown in Table 4.

Rating Bank A	Obs	Percent	Compressed Bank A Rating	Obs	Percent	Rating Bank B	Obs	Percent
1	157	0.08						
2	505	0.24						
3	887	0.43	1	3 382	1.62	1	57	0.05
4	1 833	0.88	2	50 826	24.38	2	2 835	2.43
5	17 817	8.54	3	109 655	52.59	3	29 764	25.56
6	26 532	12.73	4	30 003	14.39	4	70 987	60.96
7	6 477	3.11	5	9 363	4.49	5	11 574	9.94
8	26 843	12.87	6	3 589	1.72	6	1 228	1.05
9	61 346	29.42	7	1 696	0.81			
10	21 466	10.29						
11	30 003	14.39						
12	9 363	4.49						
13	3 589	1.72						
14	1696	0.81						
	208 514	100.00		208 514	100.00		116 445	99.99
Mean rating		8.63		3.04			3.82	
Std. Deviation		2.17		0.96			0.68	

Table 3B: 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. Default classes are excluded and shown in Table 4.

	Bank A Borrowers		Bank B Borrowers	
Rating	Observations	Percent	Observations	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	7 770	6.67
5	7 546	3.62	4 732	4.06
	208 514	100.00	116 445	100.00
Mean rating	2.00		2.20	
Std. Deviation	1.10		1.09	

Table 4, panel A: Mapping compressed Bank A ratings into credit bureau ratings

Numbers are percentage shares of full sample over the period 1997-Q3 to 2000-Q1. Higher ratings imply worse creditworthiness. Observations are defined as quarter-borrower pairs. Numbers in the last line and last column display transitions into the default state as a percentage share of total defaults only.

	Credit bureau ratings							
	1	2	3	4	5	6	Total	Rating 6 only
Bank A rating								
1	1,0%	0,4%	0,2%	0,0%	0,0%	0,0%	1,6%	0,0%
2	16,4%	5,6%	2,3%	0,2%	0,0%	0,0%	24,5%	3,3%
3	20,8%	15,4%	12,7%	2,8%	1,1%	0,0%	52,8%	18,7%
4	3,0%	3,7%	4,7%	1,5%	0,8%	0,0%	13,7%	8,7%
5	0,4%	0,8%	1,2%	1,1%	0,8%	0,0%	4,4%	8,5%
6	0,1%	0,2%	0,4%	0,4%	0,6%	0,0%	1,7%	5,7%
7	0,0%	0,1%	0,2%	0,2%	0,4%	0,0%	0,8%	6,5%
8	0,0%	0,0%	0,0%	0,1%	0,1%	0,1%	0,3%	48,6%
Total	41,7%	26,3%	21,7%	6,3%	3,8%	0,2%	100,0%	100,0%
Rating 8 only	3,9%	5,8%	11,8%	16,8%	33,3%	28,3%	100,0%	

Table 4, panel B: Mapping Bank B ratings into credit bureau ratings

Numbers are percentage shares of full sample over the period 1997-Q3 to 2000-Q1. Higher ratings imply worse creditworthiness. Observations are defined as quarter-borrower pairs. Numbers in the last line and last column display transitions into the default state as a percentage share of total defaults only.

	Credit bureau ratings							
Bank B rating	1	2	3	4	5	6	Total	Rating 6 only
1	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
2	1,8%	0,5%	0,2%	0,0%	0,0%	0,0%	2,5%	0,0%
3	15,3%	6,5%	3,3%	0,4%	0,1%	0,0%	25,6%	0,8%
4	14,3%	19,3%	20,2%	4,6%	1,9%	0,1%	60,4%	23,8%
5	0,7%	2,1%	3,7%	1,6%	1,6%	0,1%	9,9%	30,4%
6	0,0%	0,1%	0,2%	0,2%	0,4%	0,0%	1,1%	10,7%
7	0,0%	0,1%	0,1%	0,1%	0,1%	0,1%	0,5%	34,2%
Total	32,2%	28,7%	27,7%	7,0%	4,1%	0,3%	100,0%	100,0%
Rating 7 only	3,0%	13,1%	18,2%	22,2%	25,8%	17,7%	100,0%	

Table 5: OLS regressions with all borrowers, credit bureau, Bank A, and Bank B
Sample period is 1997-Q3 to 2000-Q1; standard errors are robust.

	Dependent variable					
	Credit bureau rating			Bank A rating uncompressed		
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
Constant	.212 (.00214)	.071 (.00437)	.155 (.0352)	0.859 (.0110)	.815 (.0107)	1.470 (.0206)
Lag credit bureau rating	.885 (.00111)	.870 (.00123)	.856 (.00135)		.110 (.00225)	
Lag Bank A rating uncompressed		.020 (.00057)		.908 (.00115)	.887 (.00133)	.887 (.00133)
Credit bureau rating dummies	No	No	No	No	No	Yes
Bank rating dummies	No	No	Yes	No	No	No
Residual sum of squares	55575	55237	54889	174853	172284	172059
Adj. R^2	.7784	.7798	.7811	.8226	.8252	.8255
Nobs	208514	208514	208514	208514	208514	208514

	Dependent variable					
	Credit bureau rating			Bank A rating		
Explanatory variables	(7)	(8)	(9)	(10)	(11)	(12)
Constant	.212 (0.00214)	.076 (.00356)	.209 (.0082)	0.217 (.00323)	.191 (.00314)	.579 (.0119)
Lag credit bureau rating	.885 (.00111)	.861 (.00131)	.860 (.00132)		.0599 (.00122)	
Lag Bank A rating		.0612 (.00141)		.938 (.00105)	.907 (.00130)	.904 (.00135)
Credit bureau rating dummies	No	No	No	No	No	Yes
Bank rating dummies	No	No	Yes	No	No	No
Residual sum of squares	55575	55021	55001	26540	25831	25735
Adj. R^2	.7784	.7806	.7807	.8610	.8647	.8652
Nobs	208514	208514	208514	208514	208514	208514

	Dependent variable					
	Credit bureau rating			Bank B rating		
Explanatory variables	(13)	(14)	(15)	(16)	(17)	(18)
Constant	.237 (.00314)	-.088 (.00794)	.159 (.04700)	0,162 (.00444)	.172 (.00451)	.279 (.00760)
Lag credit bureau rating	0.886 (.00142)	0.858 (.00169)	0.857 (.00170)		.0191 (.00072)	
Lag Bank B rating		0.102 (.00251)		.960 (.00116)	.947 (.00133)	.947 (.00134)
Credit bureau rating dummies	No	No	No	No	No	Yes
Bank rating dummies	No	No	Yes	No	No	No
Residual sum of squares	30607	30163	30147	4981	4940	4939
Adj. R^2	.7802	.7833	.7835	.9079	.9087	.9087
Nobs	116445	116445	116445	116445	116445	116445

Table 6: Regressions of credit bureau with all borrowers; and dummies for Bank A uncompressed ratings

Details of regression in column 3 of Table 5. Refers to regression of credit bureau rating on its lag and dummies of uncompressed Bank A ratings, 1997-Q3 to 2000-Q1, robust standard errors. (*) indicates a coefficient is significantly different from following two ratings at .01 confidence level.

Variable	Coefficient		S.e.
constant	.1552		.0352
lagged credit bureau rating	.8557		.0014
dummy Bank A rating 2	.0552		.0405
dummy Bank A rating 3	.0705		.0383
dummy Bank A rating 4	.0615		.0369
dummy Bank A rating 5	.0823	*	.0354
dummy Bank A rating 6	.0306		.0352
dummy Bank A rating 7	.0278		.0354
dummy Bank A rating 8	.1438	*	.0353
dummy Bank A rating 9	.1180		.0352
dummy Bank A rating 10	.0597		.0353
dummy Bank A rating 11	.1788		.0357
dummy Bank A rating 12	.2537		.0357
dummy Bank A rating 13	.3012		.0365
dummy Bank A rating 14	.3907		.0373

Table 7: Explanatory power of lagged bank ratings or credit bureau ratings in OLS regressions

Entries in the table reflect the percentage by which the residual sum of squares is reduced when a one-period lag of bank ratings or credit bureau ratings is introduced as an explanatory variable in addition to the lagged dependent variable in Table 5.

	Dependent variable			
	Credit bureau rating		Bank A rating compressed	Bank B rating
Explanatory variable added	Bank A rating compressed	Bank B rating	Credit bureau rating	
Data sub-sample	(1)	(2)	(3)	(4)
All borrowers	1.00	1.45	2.67	0.81
Small borrowers	0.94	1.21	3.00	0.44
Medium-sized borrowers	1.04	1.40	2.63	0.92
Large borrowers	1.01	1.52	2.08	0.70

Table 8, panel I: Tests of the area under the ROC curve for predicted logit probabilities (in-sample)
Logit models are estimated first using only the lag of the dependent variable. Then we add 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 a greater explanatory power as estimated by the area under the ROC curve, We implement this test using both the credit bureau rating and the bank ratings as the dependent variable.

	ROC area	Std. Error	P-value	ROC area	Std. Error	P-value
	Bank rating used			*Bank rating used*		
Dependent variable: credit bureau default.						
Regressors:	Bank A			Bank B		
	(1)	(2)	(3)	(4)	(5)	(6)
Credit bureau rating (1)	0.8545	0.0158		0.8259	0.0179	
Credit bureau rating & bank rating (2)	0.8890	0.0131		0.8602	0.0154	
H_0: Area1=Area2			0.0001			0.000
	Bank rating used			*Bank rating used*		
Dependent variable: bank default.						
Regressors:	Bank A			Bank B		
	(7)	(8)	(9)	(10)	(11)	(12)
Bank rating (1)	0.9552	0.0043		0.9545	0.0047	
Bank rating & credit bureau rating (2)	0.9682	0.0034		0.9697	0.0037	
H_0: Area1=Area2			0.000			0.000

Table 8, panel II: Kolmogorov-Smirnov tests of equal distributions for predicted logit probabilities (in-sample)
Two-sample Kolmogorov-Smirnov tests for equality of distribution functions. Predicted probabilities are generated from a logit regression model where default events, as defined by a specific rating sytem, are first regressed on the lag of that specific rating. In a second step, the first lag of another rating was added. Then the predicted probabilities implied by these two models are compared using a K-S test. Corrected p-values take into account sample size. F(.) and H(.) are the distribution functions derived from logit regressions.

Hypothesis tested	D-statistic	P-value	Corrected P-value	D-statistic	P-value	Corrected P-value
	Variable added on right-hand side					
Dependent variable: credit bureau default. Regressors: credit bureau rating and bank rating	Bank A			Bank B		
	(1)	(2)	(3)	(4)	(5)	(6)
F(credit bureau rating & bank rating) < H(credit bureau rating)	0.2093	0.000		0.2461	0.000	
F(credit bureau rating & bank rating) > H(credit bureau rating)	-0.2396	0.000		-0.2199	0.000	
Combined Kolmogorov-Smirnov	0.2396	0.000	0.000	0.2461	0.000	0.000
	Variable added on right-hand side					
Dependent variable: bank default. Regressors: bank rating and credit bureau rating	Bank A			Bank B		
	(7)	(8)	(9)	(10)	(11)	(12)
F(bank rating & credit bureau rating) < H(bank rating)	0.3513	0.000		0.3307	0.000	
F(bank rating & credit bureau rating) > H(bank rating)	-0.1663	0.000		-0.2736	0.000	
Combined Kolmogorov-Smirnov	0.3513	0.000	0.000	0.3307	0.000	0.000

Table 8, panel III: Tests of the area under the ROC curve for predicted logit probabilities (out-of-sample)
Two-model tests for equality of the area under the ROC curve. First, the data sample is split into two equally sized sub-samples but each covering the same time period. Next, using the first half of the data, a logit regression model is estimated where default events, as defined by a specific rating system, are first regressed on the lag of that specific rating. Then, using the second half of the data, the ROC curve is computed. Then we add the lag of a second rating and re-compute the ROC curve. Then we test if the expanded model has a greater explanatory power as estimated by the area under the ROC curve. We implement this test using both the credit bureau rating and the bank ratings as the dependent variable.

	ROC area	Std. Error	P-value	ROC area	Std. Error	P-value
	Bank rating used			*Bank rating used*		
Dependent variable: credit bureau default.						
Regressors:	Bank A			Bank B		
	(1)	(2)	(3)	(4)	(5)	(6)
Credit bureau rating (1)	0.8583	0.0211		0.8440	0.0237	
Credit bureau rating & bank rating (2)	0.8921	0.0176		0.8890	0.0181	
H_0: Area1=Area2			0.0002			0.0003
	Bank rating used			*Bank rating used*		
Dependent variable: bank default.						
Regressors:	Bank A			Bank B		
	(7)	(8)	(9)	(10)	(11)	(12)
Bank rating (1)	0.9591	0.0060		0.9585	0.0060	
Bank rating & credit bureau rating (2)	0.9703	0.0049		0.9774	0.0037	
H_0: Area1=Area2			0.000			0.000

Table 8, panel IV: Kolmogorov-Smirnov tests of equal distributions for predicted logit probabilities (out-of-sample)
Two-sample Kolmogorov-Smirnov tests for equality of distribution functions. First, the data sample is split into two equally sized sub-samples but each covering the same time period. Next, using the first half of the data, a logit regression model is estimated where default events, as defined by a specific rating sytem, are first regressed on the first lag of that specific rating. Then, using the second half of the data, predicted probabilities are generated model. from the estimated. In the second stage, the lag of another rating was added and the same procedure as above was applied. Finally, the predicted probabilities implied by these two models are compared using a K-S test. Corrected p-values take into account sample size. F(.) and H(.) are the distribution functions derived from logit regressions.

Hypothesis tested	D-statistic	P-value	Corrected P-value	D-statistic	P-value	Corrected P-value
	Variable added on right-hand side					
Dependent variable: credit bureau default. Regressors: credit bureau rating and bank rating	Bank A			Bank B		
	(1)	(2)	(3)	(4)	(5)	(6)
F(credit bureau rating & bank rating) < H(credit bureau rating)	0.1799	0.000		0.2527	0.000	
F(credit bureau rating & bank rating) > H(credit bureau rating)	-0.2371	0.000		-0.2193	0.000	
Combined Kolmogorov-Smirnov	0.2371	0.000	0.000	0.2527	0.000	0.000
	Variable added on right-hand side					
Dependent variable: bank default. Regressors: bank rating and credit bureau rating	Bank A			Bank B		
	(7)	(8)	(9)	(10)	(11)	(12)
F(bank rating & credit bureau rating) < H(bank rating)	0.3500	0.000		0.3303	0.000	
F(bank rating & credit bureau rating) > H(bank rating)	-0.1645	0.000		-0.2784	0.000	
Combined Kolmogorov-Smirnov	0.3500	0.000	0.000	0.3303	0.000	0.000

Table 9: Ordered logit regressions with all borrowers; credit bureau and Bank A (compressed)
Bank A ratings have been compressed from 15 to 8 grades. Sample period is 1997-Q3 to 2000-Q1; standard errors are robust.

	Dependent variable					
	Credit bureau rating			Bank A rating		
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
Constant	5.393 (0.026)	5.932 (0.023)	5.435 (0.048)	6.653 (0.038)	7.094 (0.040)	5.205 (0.05)
Lag credit bureau rating	3.307 (0.011)	3.240 (0.011)	3.236 (0.011)	5.428 (0.022)	0.398 (0.006)	5.347 (0.022)
Lag Bank A rating		0.235 (0.006)			5.347 (0.022)	
Credit bureau rating dummies	No	No	No	No	No	Yes
Bank rating dummies	No	No	Yes	No	No	No
Pseudo-R2	0.5053	0.5080	0.5085	0.6981	0.7034	0.7035
McKelvey & Zavoina's R^2	0.799	0.802	0.802	0.889	0.894	0.894
BIC	273945	272477	272292	160754	157963	157949
Nobs	208514	208514	208514	208514	208514	208514

Table 10: Sensitivity of explanatory power to lag length in OLS regressions with credit bureau and Bank A rating

Entries in the table reflect the percentage by which the residual sum of squares is reduced when a one-, two-, three-, or four-period lag of bank ratings or credit bureau ratings is introduced as an explanatory variable in addition to the lagged dependent variable.

	Dependent variable			
	Credit bureau rating			
Explanatory variable added	Bank A rating compressed			
Lag period	1 Quarter	2 Quarters	3 Quarters	4 Quarters
	(1)	(2)	(3)	(4)
All borrowers	1.00	1.17	1.33	1.49
Small borrowers	0.94	1.08	1.27	1.50
Medium-sized borrowers	1.04	1.20	1.35	1.44
Large borrowers	1.00	1.25	1.41	1.59
	Bank A rating compressed			
Explanatory variable added	Credit bureau rating			
Lag period	1 Quarter	2 Quarters	3 Quarters	4 Quarters
	(5)	(6)	(7)	(8)
All borrowers	2.67	4.72	6.42	7.92
Small borrowers	3.01	5.34	7.32	9.20
Medium-sized borrowers	2.63	4.65	6.35	7.65
Large borrowers	2.08	3.64	4.82	6.09

Table 11: Sensitivity of explanatory power to lag length in OLS regressions with credit bureau and Bank B rating

Entries in the table reflect the percentage by which the residual sum of squares is reduced when a one-, two-, three-, or four-period lag of bank ratings or credit bureau ratings is introduced as an explanatory variable in addition to the lagged dependent variable.

	Dependent variable			
	Credit bureau rating			
Explanatory variable added	Bank B rating			
Lag period	1 Quarter	2 Quarters	3 Quarters	4 Quarters
	(1)	(2)	(3)	(4)
All borrowers	1.44	2.00	2.31	2.39
Small borrowers	1.18	1.60	1.93	1.83
Medium-sized borrowers	1.39	1.95	2.21	2.31
Large borrowers	1.51	2.08	2.47	2.54

	Bank B rating			
Explanatory variable added	Credit bureau rating			
Lag period	1 Quarter	2 Quarters	3 Quarters	4 Quarters
	(5)	(6)	(7)	(8)
All borrowers	0.82	1.50	2.00	2.36
Small borrowers	0.41	0.69	0.93	1.42
Medium-sized borrowers	0.92	1.68	2.32	2.77
Large borrowers	0.69	1.29	1.61	1.85

Table 12: Explanatory power of lagged bank ratings or credit bureau ratings in OLS regressions using only observations of lagged ratings variable when change of rating is observed

Entries in the table reflect the percentage by which the residual sum of squares is reduced when a one-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 comparision with the other columns of this table.

	Dependent variable			
	Credit bureau rating		Bank A rating	Bank B rating
	Regressions Conditioned on Bank Rating Change			
	(1)	(2)	(3)	(4)
Explanatory variable added	Bank A rating	Bank B rating	Credit bureau rating	
All borrowers	1,69	7,16	8,84	2,76
Small borrowers	1,53	0,46	14,72	0,28
Medium-sized borrowers	2,00	6,10	4,68	3,39
Large borrowers	1,75	9,68	2,67	2,08
	Regressions Conditioned on Credit Bureau Rating Change			
	(5)	(6)	(7)	(8)
All borrowers	1,61	2,95	9,46	1,70
Small borrowers	1,34	3,42	15,45	1,04
Medium-sized borrowers	1,73	2,66	4,93	1,84
Large borrowers	1,94	3,18	1,61	1,67
	Unconditioned Regressions on Credit Bureau Rating Change			
	(9)	(10)	(11)	(12)
All borrowers	1,59	2,61	8,05	2,45
Small borrowers	1,43	3,13	12,20	2,15
Medium-sized borrowers	1,70	2,13	5,56	2,98
Large borrowers	1,48	3,26	1,45	1,50
	Continuous credit bureau rating		Bank A rating	Bank B rating
Explanatory variable added	Bank A rating	Bank B rating	Continuous credit bureau rating	
Data selection	(13)	(14)	(15)	(16)
Conditioned on bank rating changes, all borrowers	1,62	6,54	8,69	2,66
Conditioned on credit bureau rating changes, all borrowers	1,26	2,24	9,71	1,95
Unconditioned, all borrowers	1,23	2,10	8,39	2,63

Table 13: Cox regressions on credit bureau defaults

The Breslow method has been used for tied observations. The (*) sign indicates that the variable had to be dropped because no defaults occur for the dependent the relevant lag variable at or because of high correlation with another explanatory variable. The "-" sign indicates that the particular RHS variable is not available for this regression.

	Dependent variable: Credit bureau default											
	RHS: Lag 1, Bank A or CB						RHS: Lag 1, Bank B or CB					
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lag credit bureau rating	3.26		2.43				2.97		2.20			
	(0.22)		(0.19)				(0.23)		(0.195)			
Lag bank rating		2.51	1.61					4.92	2.72			
		(0.12)	(0.10)					(0.54)	(0.35)			
Lag, Dummy CB rating = 1				0.015		0.031				0.014		0.030
				(0.0046)		(0.011)				(0.0068)		(0.0148)
Lag, Dummy CB rating = 2				0.024		0.048				0.046		0.079
				(0.0073)		(0.016)				(0.014)		(0.025)
Lag, Dummy CB rating = 3				0.071		0.133				0.098		0.148
				(0.015)		(0.031)				(0.023)		(0.036)
Lag, Dummy CB rating = 4				0.31		0.445				0.440		0.558
				(0.063)		(0.093)				(0.101)		(0.130)
Lag, Dummy bank rating = 2					*	0.214					*	*
						(0.087)						
Lag, Dummy bank rating = 3					1.88	0.186					*	*
					(0.64)	(0.054)						
Lag, Dummy bank rating = 4					5.83	0.334					2.39	1.29
					(2.04)	(0.096)					(0.79)	(0.44)
Lag, Dummy bank rating = 5					18.70	0.542					17.07	4.40
					(6.52)	(0.153)					(5.58)	(1.53)
Lag, Dummy bank rating = 6					44.03	0.910					*	*
					(15.63)	(0.261)						
Lag, Dummy bank rating = 7					66.45	*					-	-
					(24.77)							
Residual Sum of Squares												
Number of subjects	31621	31621	31621	31621	31621	31621	17475	17475	17475	17475	17475	17475
Number of failures	165	165	165	165 #	165	165	128	128	128	128	128	128
Nobs	214968	214968	214968	214968	214968	214968	119819	119819	119819	119819	119819	119819
Log likelihood	-1461.7	-1507.9	-1434.2	-1459.0	-1504.8	-1434.4	-1080.6	-1096.5	-1051.4	-1079.1	-1121.7	-1057.6

Table 14: Cox regressions on bank defaults

The Breslow method has been used for tied observations. The (*) sign indicates that the variable had to be dropped because no transitions occur from that particular grade into the default state or because of high correlation with another explanatory variable. The "-" sign indicates that the particular RHS variable is not available for this regression.

	Dependent variable: Bank A default						Dependent variable: Bank B default					
	RHS: Lag 1, Bank A or CB						RHS: Lag 1, Bank B or CB					
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lag credit bureau rating	3.69		2.37				3.14		2.26			
	(0.18)		(0.14)				(0.21)		(0.17)			
Lag bank rating		3.69	2.03					5.64	2.99			
		(0.18)	(0.09)					(0.54)	(0.34)			
Lag, dummy CB rating = 1				0.0076		0.0421				0.015		0.070
				(0.0022)		(0.0136)				(0.0057)		(0.028)
Lag, dummy CB rating = 2				0.0242		0.0993				0.045		0.128
				(0.0051)		(0.0232)				(0.011)		(0.034)
Lag, dummy CB rating = 3				0.0627		0.1962				0.090		0.197
				(0.0097)		(0.0334)				(0.018)		(0.042)
Lag, dummy CB rating = 4				0.2567		0.4778				0.221		0.349
				(0.0384)		(0.0738)				(0.052)		(0.084)
Lag, dummy bank rating = 2					*	*					*	*
Lag, dummy bank rating = 3					2.14	1.19					*	*
					(0.70)	(0.40)						
Lag, dummy bank rating = 4					10.14	3.59					9.52	5.59
					(3.22)	(1.21)					(5.63)	(3.38)
Lag, dummy bank rating = 5					40.36	8.15					69.81	22.61
					(12.57)	(2.78)					(41.14)	(13.88)
Lag, dummy bank rating = 6					80.48	12.46					249.67	49.89
					(25.76)	(4.41)					(151.96)	(31.92)
Lag, dummy bank rating = 7					216.85	29.40					-	-
					(67.73)	(10.26)						
Residual sum of squares												
Number of subjects	31635	31635	31635	31635	31635	31635	17490	17490	17490	17490	17490	17490
Number of failures	312	312	312	312	312	312	163	163	163	163	163	163
Nobs	214925	214925	214925	214925	214925	214925	119812	119812	119812	119812	119812	119812
Log likelihood	-2686.1	-2695.9	-2564.2	-2714.7	-2687.5	-2580.5	-1353.6	-1372.1	-1308.0	-1375.5	-1369.4	-1321.3

Table 15: Log likelihoods in Cox proportional hazards model; All borrowers

Log likelihood values for models with the first lag of the RHS variable are taken from columns (1)-(3) and (7)-(9) in Tables 13-14 and for lag 2 from Appendix Tables A7 - A8. Significance of an additional RHS variable is shown at the 10 (*), 5 (**), 1 (***), and 0.1 (****) levels. In the likelihood ratio tests (lower panel), the value displayed is 2*log(likelihood ratio).

	Dependent variable			
	Credit bureau default		Bank default	
	(1)	(2)	(3)	(4)
Data sample	Bank A	Bank B	Bank A	Bank B
Explanatory variables				
Lag of CB rating	-1461,7	-1080,6	-2686,1	-1353,6
Lag of bank Rating	-1507,9	-1096,5	-2695,9	-1372,1
Lag of CB and bank Rating	-1434,2	-1051,4	-2564,2	-1308,0
Lag 2 of CB rating	-1402,6	-910,7	-3169,3	-1544,0
Lag 2 of bank rating	-1439,1	-930,8	-3258,5	-1585,7
Lag 2 of CB and bank rating	-1385,6	-894,3	-3104,9	-1508,9
Likelihood ratio tests for exclusion of particular lags				
First Lag Only				
Exclusion of lag of bank rating	55,0 ****	58,4 ****	243,8 ****	91,2 ****
Exclusion of lag of CB rating	147,4 ****	90,2 ****	263,4 ****	128,2 ****
Second Lag Only				
Exclusion of lag 2 of bank rating	34,0 ****	32,8 ****	128,8 ****	70,2 ****
Exclusion of lag 2 of CB Rating	107,0 ****	73,0 ****	307,2 ****	153,6 ****

Table 16: OLS regressions with all borrowers, Bank A and Bank B; change in ratings regressed on lagged change in ratings
Sample period is 1997-Q3 to 2000-Q1; standard errors are clustered by borrower. Changes computed over different time intervals

	Dependent variable							
	Change in Compressed Bank A Rating							
	1 Quarter		2 Quarters		3 Quarters		4 Quarters	
	(1)		(2)		(3)		(4)	
Explanatory variables								
Constant	0.03183	***	0.006897	***	0.11814	***	0.18065	***
	(.00079)		(.00180)		(.00317)		(.00500)	
Lag change in compressed Bank A rating	-0.05598	***	-0.06305	***	-0.10286	***	-0.15375	***
	(.00335)		(.00324)		(.00516)		(.00940)	
Adj. R^2	0.0034		0.0046		0.0106		0.0206	
Nobs	174788		122346		79072		42696	

	Dependent variable							
	Change in Bank B Rating							
	1 Quarter		2 Quarters		3 Quarters		4 Quarters	
Explanatory variables								
Constant	0.01145	***	0.01936	***	0.02453	***	0.03170	***
	(.00062)		(.00139)		(.00244)		(.00381)	
Lag change in Bank B rating	-0.02370	***	-0.04656	***	-0.09519	***	-0.13202	***
	(.00470)		(.00570)		(.00780)		(.01116)	
Adj. R^2	0.0005		0.0020		0.0080		0.0146	
Nobs	96689		68692		45605		27234	

Table 17: OLS regressions with small, medium-sized, and large borrowers, Bank A and Bank B, change in ratings regressed on lagged change in ratings

Sample period is 1997-Q3 to 2000-Q1, standard errors are clustered by borrower

This table shows the coefficients on the lagged change in "Compressed Bank A" and Bank B rating (robust standard error).

	Dependent variable							
	Compressed Bank A rating							
	1 Quarter (1)		2 Quarters (2)		3 Quarters (3)		4 Quarters (4)	
Data sub-sample								
Small borrowers	-0.0582	***	-0.0642	***	-0.1054	***	-0.1625	***
	(.0053)		(.0056)		(.0088)		(.0166)	
Medium-sized borrowers	-0.0527	***	-0.0648	***	-0.1022	***	-0.1518	***
	(.0046)		(.0048)		(.0078)		(.0137)	
Large borrowers	-0.0592	***	-0.0567	***	-0.1001	***	-0.1428	***
	(.0099)		(.0066)		(.0108)		(.0196)	

	Dependent variable							
	Bank B Rating							
	1 Quarter		2 Quarters		3 Quarters		4 Quarters	
Data sub-sample								
Small borrowers	-0.0277	***	-0.0226	***	-0.1015	***	-0.0948	***
	(.0235)		(.0304)		(.0430)		(.0669)	
Medium-sized borrowers	-0.0204	***	-0.0515	***	-0.1027	***	-0.1387	***
	(.0052)		(.0071)		(.0104)		(.0144)	
Large Borrowers	-0.0281	***	-0.0411	***	-0.0842	***	-0.1250	***
	(.0089)		(.0098)		(.0122)		(.0182)	

Table 18. 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 Tables 8 and 9. The explanatory variable is the coefficient of mean reversion (Reversion Coefficient) from Table 11. The observations are borrower size-lag length-Bank combinations, with 3 sizes of borrower,   types of banks, there are 24 observations. Errors are clustered by Size-Bank combination.

	Dependent variable			
	Overall	Lag Length Dummies Bank Rating Inefficiency	Bank-Size Dummies	Full
	(1)	(2)	(3)	(4)
Explanatory variables				
Constant	-0.0003	0.0261 *	-0.0348 **	-.0447 **
	(0.0032)	(0.0112)	(0.0091)	(0.0121)
Reversion Coefficient	0.4017 ***	1.0311 ***	0.2812 **	0.0117
	(.0900)	(.2266)	(.0930)	(.1746)
Size-Bank Dummies	No	No	Yes	Yes
Length of Lag Dummies	No	Yes	No	Yes
Degrees of Freedom	22	19	17	15
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	