Credit Line Usage, Checking Account Activity, and Default Risk of Bank Borrowers

Lars Norden

Rotterdam School of Management, Erasmus University

Martin Weber

University of Mannheim

We wish to thank Utpal Bhattacharya, Arnoud W.A. Boot, John H. Boyd, Hans Degryse, Michael Faulkender, Daniel Foos, Paolo Fulghieri (the editor), Jens Grunert, Florian Heider, Andrew Hertzberg, Kevin Keasey, Laetitia Lepetit, Ernst Maug, Loretta J. Mester, Sebastian Müller, Steven Ongena, Mitchell A. Petersen, Andreas Pfingsten, Kasper Roszbach, Sascha Steffen, Gregory F. Udell, Wolf Wagner, and two anonymous referees, as well as participants at the American Finance Association 2008 Meetings in New Orleans, the European Finance Association 2008 Meetings in Athens, the Midwest Finance Association 2008 Meetings in San Antonio, the Conference "Small Business Banking and Finance: A global perspective" in Cagliari, the European Summer Symposium in Financial Markets 2007 in Gerzensee, the Finance Brown Bag Seminar at the Kelley School of Business, the Finance Bag Lunch Seminar at the Kellogg School of Management, and the Research Seminar at the University of Mannheim for many useful comments and suggestions. In addition, we are grateful to the management and three anonymous employees of the bank that provided us with the data. Parts of this paper were completed when Norden was visiting Indiana University, Kelley School of Business, Finance Department, Bloomington, IN 47405, USA. Financial support from the German Research Foundation is gratefully acknowledged (Deutsche Forschungsgemeinschaft, grant: NO 804/1-1). Send correspondance to Lars Norden, Rotterdam School of Management, Erasmus University, Department of Finance, Burgemeester Oudlaan 50, 3000 DR Rotterdam, the Netherlands; telephone: +31 10 4082807; fax: +31 10 4089017. E-mail: lnorden@rsm.nl.

Information on borrower quality is a fundamental issue in debt contracting, corporate and consumer finance, and financial intermediation. We investigate the link between account activity and information production on the default risk of borrowers. Based on a unique data set, we find that credit line usage, limit violations, and cash inflows exhibit abnormal patterns approximately 12 months before default events. Measures of account activity substantially improve default predictions and are especially helpful for monitoring small businesses and individuals. We also find that the early warning indications from account activity result in higher loan spreads, and in a higher likelihood of limit reductions and complete write-offs. Our study shows that information on account activity provides a real-time window into the borrower's cash flows, thus explaining why banks have an advantage in providing certain types of debt financing.

Information on borrower quality is a fundamental issue in debt contracting, corporate and consumer finance, and financial intermediation. Lenders such as banks, finance companies, and suppliers consider different sources and types of information to assess a borrower's default risk, but little is known empirically about how information production actually works. In this paper, we investigate whether information on credit line usage and checking account activity is helpful for monitoring borrowers, and how banks use this information in managing their credit relationships.

Theoretical work on debt contracting considers signaling, screening, or monitoring mechanisms to overcome problems arising from asymmetric information, such as adverse selection and moral hazard. Corporate and consumer finance research both provide evidence on the determinants of credit availability, types of lenders, financial constraints, and bankruptcy. Modern research on financial intermediation has been analyzing the role of banks in information production (Diamond 1984; Ramakrishnan and Thakor 1984; Boyd and Prescott 1986). Later theoretical work laid the groundwork for a more focused examination of bank monitoring (see, e.g., Diamond 1991; Rajan 1992). Subsequent empirical research on relationship lending has produced evidence that focuses primarily on the benefits that stem from a banking relationship (see, e.g., Petersen and Rajan 1994; Berger and Udell 1995). However, almost entirely missing from these studies is a direct examination of the sources and types of information banks use for monitoring. This is an important issue, since access to private and timely information may explain why banks apply different types of lending technologies to certain borrowers, such as relationship lending.

Our study has also implications for the financial system architecture, since information synergies between the left-hand side of the balance sheet (loans) and right-hand side (deposits) explain, to some extent, the uniqueness of banks. These synergies are inconsistent with some policy prescriptions, such as "narrow banking" and the "originate-and-distribute model". The financial crisis of 2007-2009 illustrates that a separation of information production and risk taking, as implied by credit risk transfer, may lead to unfavorable consequences.

Some speculation and indirect evidence from corporate and consumer finance and from banking research indicate that offering a checking account along with a loan is important. By providing linked financial services, the bank can access information that is private, timely, quasi costless, and reliable. In particular, the combined activity in a borrower’s checking account and her credit line reveals significant information about her cash flows. That is, it provides the lender with information about the borrower’s “debits” (draws on the account that reflect cash outflows) and “credits” (receipts that reflect cash inflows). Thus, these debits and credits may be the key determinant of a borrower’s financial flexibility and debt capacity (e.g., Sufi 2009). Unlike accounting numbers, payment data are less likely to be influenced by rules and policies. However, account activity might be fragmented across different banks, which implies that main banks may receive the most benefit from this source of information. Nakamura's (1993) “checking account hypothesis” states that firms’ bank account activity is informative and that banks use this information in managing their relationships. To the best of our knowledge, there is no evidence on this hypothesis for individuals, and only a few papers such as those by Mester, Nakamura, and Renault (2007) and Jiménez, Lopez, and Saurina (2009) study the link between credit line usage and default risk of firms.

Thus, in this paper we investigate whether the combined information on credit line usage and checking account activity is helpful for monitoring, and how banks use this information. We exploit the fact that in Germany, a line of credit is usually offered jointly with a checking account. This practice is similar to a U.S. consumer having a checking account with an overdraft line attached to it. We apply a univariate event study and multivariate probit models to a unique data set from a German universal bank that includes more than three million account-month observations on firms and individuals for the period 2002-2006.

First, we find that credit line usage, limit violations, and cash inflows exhibit abnormal patterns approximately 12 months before default. Measures of account activity substantially improve the fit of default prediction models. Our analysis by borrower type indicates that account activity is particularly useful for monitoring small businesses and individuals. Moreover, the bank can observe abnormal patterns earlier in the case of close bank-borrower relationships, suggesting that this information is especially valuable for main banks.

Second, we show that banks use information on credit line usage and account activity in managing their lending relationships. We find that borrowers with a checking account pay higher spreads on subsequent loans if the bank has obtained early warnings indications from account activity during the year before the loan is granted. Our analysis also indicates that account activity is especially informative for severe and unexpected defaults, and that clear early warning indications increase the likelihood that the bank writes off the loan completely at default. These findings show that the combination of credit line usage and checking account activity effectively gives the lender a real-time window into the borrower's cash flows.

Our study contributes to the banking and finance literature in several ways. We provide universal evidence on a vital component of financing relationships, the link between account activity and information obtained about borrower quality. This evidence enables us to draw a clearer distinction between banks' information production for different borrower types and lending technologies. We provide insights on the value of account activity for monitoring firms and individuals, and for relationship lending, transaction-based private debt, and arm's-length capital markets debt. Our results also shed light on the events leading to financial distress, and have implications on bankruptcy resolution, especially debt restructuring. We believe that using Germany is particularly instructive, because in Germany, integrated checking accounts are a key financial service for firms and individuals, and the large majority of firms rely on main banks as relationship lenders (Elsas and Krahnen 1998; Elsas 2005).

The paper proceeds as follows. In Section 1 we review the related literature. In Section 2 we describe the institutional framework and the data. In Section 3 we present the method and our results on the relation between account activity and default risk of bank borrowers. In Section 4 we show how banks use this information in managing their lending relationships. In Section 5 we report findings from several tests of robustness. Section 6 concludes.

1. Review of Related Literature: Bank-Firm Relationships, Checking Accounts, and Lines of Credit

Our study relates to three strands of the banking and finance literature: the strength of bank-firm relationships, checking accounts as a source of information for banks, and lines of credit.

The first strand examines the idea that banks can gather information if the borrower uses financial services other than loans, such as checking accounts, payment services, savings and money market accounts, brokerage services, and underwriting activities. Thus, there may be an informational spillover from other financial services to the lending business.[1] Analyses such as those by Black (1975), Fama (1985), Petersen and Rajan (1994), Blackwell and Winters (1997), Chakravarty and Scott (1999), Petersen and Rajan (2002), Cole, Goldberg, and White (2004), Elsas (2005), Berger et al. (2005), and Bharath et al. (2007) hint at the source of banks' potential informational advantage and include dummy variables in empirical models to measure its impact. However, evidence on the usefulness of checking accounts for gathering information is mixed. These studies also show that borrowers with a checking account are closer to their bank, and that they communicate with their bankers on a more personal level (Petersen and Rajan 2002). Firms with checking accounts at small banks benefit from higher credit availability (Cole, Goldberg, and White 2004) and longer loan maturities (Kirschenmann and Norden 2008). Elsas (2005) shows that payments and information-sensitive financial services have a positive impact on the probability of a bank being a main bank. Berger et al. (2005) find that the existence of checking accounts has a negative impact on bank size, the physical distance between the bank and firm, and the probability of impersonal communication, and a significantly positive impact on the probability of having an exclusive lender and on the duration of the bank relationship. However, these studies provide only limited insights, because they include dummy variables for the existence of checking accounts but do not explain how banks obtain an informational advantage. Our paper explicitly looks inside the checking accounts to investigate the link between borrowers' account activity and default risk.

The second strand of literature goes one step further in identifying banks' sources of credit-relevant information. Black (1975) states that "if the individual routes most of his receipts and payments through his loan account, they can serve as a continuing source of credit information." Fama (1985) inspires Nakamura (1993) to propose the checking account hypothesis. Nakamura states that checking accounts are informative, and that banks use this information to monitor their borrowers. Vale (1993) proposes a theoretical model to study the complementary role of deposits as a source of funding and private information in lending. Solving Vale's model leads to two empirical implications. First, that borrowers that have been depositors receive better loan terms than others. Second, that exclusive borrowers are treated more favorably because they do not hide information from their banks.

There are also case studies and applied credit risk research based on small samples of transactions account data (e.g., Eisfeld 1935; Apilado, Warner, and Dauten 1974; von Stein 1983; Hackl and Schmoll 1990; Schlüter 2005). However, most of the evidence in these studies is limited to specific questions.

Mester, Nakamura, and Renault (2007) provide direct evidence on the checking account hypothesis. They analyze the usefulness of checking account information for a small sample of firms borrowing from a Canadian bank during the period 1988-1992. Apart from the sample size and composition, the most important difference between Mester, Nakamura and Renault (2007) and our study is that they consider a single lending technology, asset-based lending, i.e., secured lending that is based on receivables and inventories as collateral. Their analysis relies on annual and monthly data from 100 small firms, of which 50 are in financial distress. They find that if the borrower has an exclusive bank relationship, then monthly changes in receivables can be traced in checking accounts. Moreover, borrowings that exceed inventory and receivables help to predict rating downgrades and loan write-downs. In addition, the bank intensifies its monitoring as the credit quality decreases, i.e., loan reviews become more detailed and more frequent. However, there is no evidence on if and how this information effectively influences banks' lending decisions.

Our paper connects to the third strand of literature by examining lines of credit and loan commitments, in particular to the quantitative aspect of takedown risk (Ho and Saunders 1983; Melnik and Plaut 1986; Berger and Udell 1995; Shockley and Thakor 1997; Agarwal, Ambrose, and Liu 2006; Jiménez, Lopez, and Saurina 2009). Lines of credit are a "package of loan terms" that provide future financial flexibility during adverse credit market conditions and ensure credit availability in times of a credit quality deterioration. Moreover, some studies, such as those by Berger and Udell (1995) and Shockley and Thakor (1997) assert that lines of credit represent a formalization of the concept of bank-borrower relationship. In the context of repeated interaction such as relationship lending, these institutional arrangements can serve as a mechanism to reduce asymmetric information.

Using data on U.S. home equity lines of individuals, Agarwal, Ambrose, and Liu (2006) provide evidence that a decline in the credit quality of the borrowers is associated with an increase in credit line usage and a decrease in the probability of prepayment. Jiménez, Lopez, and Saurina (2009) analyze data from Spain during the period 1984-2005 to investigate the determinants of corporate credit line usage and its implications for exposure at default estimations. They show that the risk profile of the borrower, characteristics of the bank, and the business cycle all have a significant impact on annual credit line usage. Their study differs from ours because they do not analyze the impact of early warning indications from account activity or the behavior of banks.

Sufi (2009) takes a corporate finance perspective and examines credit lines by using 10-K SEC filings of U.S. firms for the period 1996-2003. Sufi's main finding is that to be able to comply with covenants, firms with credit lines must maintain a high cash flow level. In contrast, firms with low cash flows or covenant violations due to declines in cash flows have more difficulty accessing credit lines. These results suggest that the lack of access to a credit line is a robust measure for describing the financial constraints on firms.

2. Institutional Framework and Data

2.1 Institutional Framework

Our data relates to the credit line usage and checking account activity of borrowers from a large German universal bank. The line of credit and the checking account are typically offered as a bundle, i.e., one integrated account that can either display a credit or a debit balance. This account has an authorized credit limit, which we refer to throughout the paper as a "line of credit". Such an integrated account structure with alternating balances is the standard case for firms and individuals in Germany, continental Europe, and many other countries around the world. For example, 99% of firms' and individuals' non-cash transactions in Germany, measured by frequency and value of transactions, are processed through integrated checking accounts by means of wire transfers, direct debits, and debit cards (Deutsche Bundesbank 2008).[2] The account structure implies that the balance changes directly in response to cash inflows and outflows. The line of credit is drawn on immediately when debits exceed credits, and it is paid back gradually as the credits exceed debits. There are extreme cases in which credits are either always higher than debits (the line of credit is never used) or always lower than debits (the line of credit is permanently drawn). This practice contrasts with that of the U.S. in that checking accounts and lines of credit for businesses in the U.S. are provided through two separate accounts. In this institutional setting, the firm decides when and how much to draw from the line of credit and, in particular, when and how much to pay back.

Given this institutional framework, it is important to understand how different sources of information for banks, such as financial statements or checking accounts, reflect their clients' economic activities. Thus, we first consider the following situation of a small business. Typically, a bank obtains audited financial statements once a year and assigns an internal credit rating to the borrower that is based on both hard and soft facts (Grunert, Norden, and Weber 2005). Financial information is verified by a third party, but it is backward looking, for example, it might be four to 16 months out of date. Further, the information might be manipulated to optimize tax payments or dividends. The bank evaluates soft facts based on the firm's management quality and product market position, which are both relatively stable over time. In contrast, information on credit line usage and checking account activity are available to the bank in real time. The bank can observe this information long before receiving the financial statements, which may later confirm that the borrower is, in effect, in financial problems. Economic reasons for an increased use of the credit line are, for example, unusual decreases in sales, unsold inventories, slow turning receivables, accumulated losses, or unexpected withdrawals by the owners.

The link between economic activities and sources of information is analogous for individuals. Cash inflows typically come from the salary and other personal income. Cash outflows relate to consumption or savings. If the loan officer observes either a sharp decline in cash inflows or an increase in cash outflows in the integrated account, then she knows that the borrower is in trouble.

However, the value of information on either individual or institutional account activity may be limited because of unobservable cash transactions and the fragmentation of payments across different banks. In addition, the complexity of payments (e.g., the amount, frequency, and purpose) is higher for large firms than for individuals.

2.2 The Data

Our study is based on a unique data set that includes monthly observations on credit lines and checking accounts. The data are provided by a German universal bank, which among comparable banks is one of the largest 5% in total assets, as defined by the Deutsche Bundesbank. Our sample covers the period from January 2002 to December 2006.

We begin with a data set that comprises more than 3.7 millions of account-month observations. First, we exclude all borrowers that are financial institutions, state- or city-owned entities, and certain legal advisers. We do so because we intend to analyze banks' information production on the default risk of firms and individuals.

Second, to ensure that the sample is free of any survivorship bias that might arise from switching banks and default events, we consider all the bank's current and former borrowers for the full sample period.[3]

Third, we differentiate by borrower type (TYPE). TYPE equals one for large firms, which we define as firms with total annual sales above one million euros;[4] TYPE equals two for small firms, which are all firms with total annual sales equal or below one million euros; and TYPE equals three for individuals (e.g., employees, workers, retired persons). Since there is no information on total assets in our data set we follow the common practice and use total sales as a proxy for firm size. We take a firm's mean cumulative credits per month multiplied by twelve to calculate total annual sales. Classifying firms according to their mean credits may be a good proxy for size in general, but it may be biased when there are multiple bank relationships, because we cannot observe most of the credits. Nor do we have information about how many bank relationships a borrower has. This fact creates a conservative bias, making it more difficult to provide evidence on the usefulness of account activity measures for monitoring. However, since most small firms in Germany usually have only one or two checking accounts, information fragmentation is not a big problem. We report summary statistics by borrower type, i.e., for large firms, small firms, and individuals, and for the full sample in Table 1.

Insert Table 1 here

In Table 1, Panel A describes the structure of our final data set. The rightmost column of Panel A indicates that the full sample consists of 3,271,879 account-month observations from 86,945 accounts that belong to 67,215 borrowers. The columns in the middle of Panel A characterize the sample by borrower type and indicate that 29,650 of all account-month observations come from large firms, 430,611 from small firms, and 2,811,618 from individuals. With respect to size and ownership our sample is representative for all firms in Germany (see Federal Statistic Office 2006). For instance, 96.8% of the firms in our sample exhibit annual sales below two million euros, while this statistic is 94.9% for all firms in the German economy. Moreover, given the large share of individuals, it is not surprising that most of the borrowers (78%) have just one checking account. The data cover 48-60 months for 77% of all accounts, which allows for both cross-sectional and time-series analysis.

Panel B summarizes the main account activity measures and bank relationship variables for the full sample and by borrower type at the account level, and are reported in thousands of euros. For the full sample, the mean minimum balance per month (LOW) is -0.16 and the corresponding maximum (HIGH) is 4.15. Mean cumulative debits (DEBIT) and credits (CREDIT), i.e., the respective sums of monthly cash out- and inflows, are 8.04 and 8.22. We use CREDIT as our proxy for the monthly sales of firms and income of individuals, and use DEBIT to capture monthly expenses. The mean balance (MID) is 1.99 and the average line of credit (LIMIT, which we define as a negative number) is -8.12. Comparing variables by borrower type shows that large firms have the largest mean credits (CREDIT=560.43) and individuals the smallest (CREDIT=2.42). The same relation holds for LOW, DEBIT and

LIMIT, indicating that the classification based on TYPE is reasonable. In the full sample, the mean duration of the bank-borrower relationship, DUR, is 10.56 years. The mean distance, DIST, which we define as the aerial distance between the domicile of borrowers and the bank's head office, based on the first three digits of the ZIP code, is 7.21 kilometers. The average internal credit rating, RAT, is 2.87, measured on a rating scale of one to six, with one being the best rating, and ratings five and six comprising defaulted borrowers. The bank has different internal credit rating systems for firms and individuals and considers the rating for all loan approval decisions, loan pricing, and loan loss provisioning (Machauer and Weber 1998; Treacy and Carey 2000; Grunert, Norden, and Weber 2005). Banks base the internal credit ratings for firms on hard facts such as financial statements, and on soft facts such as the firm's product market position and management quality. Banks base the ratings for individuals on characteristics such as age, marital status, home status, and income. Information on account activity is not included in the bank's internal credit rating system but we note that loan officers may make discretionary use of this kind of information in their lending decisions. This institutional setup makes it possible for us to identify different sources of information.

Table 1, Panel C presents the frequency of rating changes and default events. During the sample period, we observe 12,803 changes of the internal credit ratings. Although we know the exact dates, we assign a rating change to the corresponding month because we measure all other variables at a monthly frequency. We note 5,515 rating upgrades and 7,288 rating downgrades. Moreover, there are 1,009 default events (DEF), i.e., these borrowers experience downgrades from ratings one to four to either rating five or rating six. The bank assigns a rating of five to borrowers that fulfill at least one of the following conditions: first specific loan loss provision, 90 days past due on any obligation, or foreclosure. The bank assigns a rating of six to borrowers that have filed for bankruptcy. A movement from rating five to rating six is a transition within the default state and does not represent a default event. As is common practice in the banking industry, the definition of default refers to the borrower level and meets the international regulatory requirements (Basel Committee on Banking Supervision 2006). The type of default is distributed almost uniformly with 541 downgrades to rating five and 468 downgrades to rating six. Although the bank's loan loss provisions follow strict regulations and are reviewed by the auditors, tax authorities, and federal banking supervisors, defaults to rating five are partly under the control of the bank management, but defaults to rating six (bankruptcy filings) are exogenous events. Separating by borrower type indicates that there are 32 large firms, 347 small firms, and 630 individuals that default during the sample period.

3. Empirical Analysis of the Relation between Account Activity and Default Risk

3.1 Methods and Univariate Event Study Results

We use two methods to investigate whether information on account activity is useful for monitoring bank borrowers. First, we perform a univariate event study. This approach has the advantage that we can easily visualize the results. Second, we estimate multivariate probit regression models in which the data are in calendar time for different time horizons. Thus, we can investigate which factors influence the probability of default in month t+1.

In the event study, we first identify default events in calendar time. Second, we transform calendar time into event time at a monthly frequency with an event time window of 48 months [event time = -36, -35, …, 0, 1, …, 12]. We observe 1,009 default events and assign these incidents to event time zero. Third, we calculate different measures of account activity for borrowers that default at event time zero in each month of the event time window. To make comparisons, we calculate the same variables in the same month for all borrowers in our sample. Thus, a bank can ex ante compare a single borrower with an appropriate benchmark (e.g., the median borrower, industry-specific, or size-based benchmarks). If the bank were to use nondefaulting borrowers as a benchmark, then the bank could only identify this group ex post, i.e., it does not help the bank to obtain ex ante warning indications.[5] In the remainder, we report results that are based on the median measures of account activity for defaulters and all borrowers in the full sample.

First, we examine the credit line usage and the cumulative number of credit limit violations. We measure credit line usage in percents by USAGE_LOW=LOW/LIMIT and USAGE=MID/LIMIT. Since we define the nominal amount of the creditline LIMIT as negative number, the ratio of a negative account balance and the nominal creditline results in a positive number, i.e., positive values for USAGE and USAGE_LOW correspond to a draw on the line of credit. A credit limit violation occurs if the minimum balance in any one month falls short of the credit limit. We define the cumulative number of credit limit violations, CVIOL, as the sum of these limit violations up to a particular month. Since we do not have daily data, we cannot observe when the bank "bounces" checks or wire transfers because of "no sufficient funds" to constrain violations. Since CVIOL refers to the observed number of overdrafts it understates the number of intended overdrafts, which makes it more difficult for us to use this measure as a predictor for borrower defaults. Figure 1 displays the findings for these variables.

Insert Figure 1 here

Panel A shows that the median credit line usage of defaulters is very different from that of nondefaulters. Although USAGE_LOW (USAGE) for defaulters starts above 60% (20%) 36 months prior to default, the corresponding values for nondefaulters are roughly zero (-30%). As explained above, a negative sign for credit line usage indicates that there is no usage at all, i.e., the checking account has a credit balance. Closer to default around event time 0, the value of USAGE_LOW (USAGE) for defaulters reaches almost 100% (80%), indicating a systematic increase of credit line usage. This observation implies that borrowers who subsequently default exhibit a gradually increasing need for liquidity, which is due to a decrease in credits such as sales or personal income. We will come back to this point when we look at the development of credits before default. Most of the run-up of USAGE_LOW occurs during the event time interval [-36, -16]. In addition, USAGE displays a sharp increase during the nine months prior to default. Conversely, the credit line usage (USAGE_LOW and USAGE) of all borrowers remains relatively stable around zero percent (-30%). We observe that there is no systematic increase or decrease in the benchmark. This finding is consistent with the results for Spanish firms by Jiménez, Lopez, and Saurina (2009), who detect a major increase in credit line usage during the 12 months prior to default.

Panel B indicates that limit violations, CVIOL, for defaulters differ considerably from those of all borrowers. Although the latter remains relative flat, the former increases continuously from a median of five at event time -36 to a median of 20 at event time 0. The slope of the curve for defaulters becomes considerably steeper around event time -18. This result accords with the findings on Canadian firms from Mester, Nakamura, and Renault (2007). We conclude that defaulters exhibit a significantly abnormal increase in credit line usage and in the number of limit violations before default.

Next, we analyze the absolute account amplitude, AMPLI, a measure of the account variation that to our knowledge has not been used in previous studies. We define this variable as AMPLI=HIGH–LOW and note that it always takes positive values regardless of the sign of LOW and HIGH. We hypothesize that borrowers who go bankrupt exhibit a systematic decrease in the absolute account amplitude. This may be the case for the following reasons. First, both firms and individuals who default later tend to face a decline in credits. In other words, firms enter financial distress mainly when cash inflows decrease because of a decline in sales. The same reasoning holds for individuals who become unemployed. Second, given a decrease in credits, borrowers become increasingly financially constrained. Therefore, at a certain level, to avoid overdrafts debits must decrease. Figure 2 depicts the event study results for the median of AMPLI.

Insert Figure 2 here

The median-amplitudes AMPLI of defaulters and all borrowers are relatively similar and in a range of 1,000 to 2,000 euros during the event time interval -36 to -6. Most important, from event time -5, we note that the AMPLI of defaulters drops sharply from 1,000 euros to less than 200 euros but the corresponding value for all borrowers does not change at all. The fact that AMPLI for defaulters is above the line for all borrowers during event time -36 to -30 is due to a relatively small number of observations in the first months. We conclude that the account amplitude provides a warning indication roughly five months prior to default.

In addition to account balances we also look at turnover information reflected by the cumulative monthly cash flows, i.e., credits (CREDIT) and debits (DEBIT). To examine the cash flow dynamics of defaulters compared to all borrowers, we define CLR (DLR) as the ratio of CREDIT (DEBIT) over the LIMIT (multiplied by -100). The inverse of CLR (DLR) represents the duration of the limit-credit (limit-debit), i.e., how many months it takes to pay back (reach) the limit. CLR and DLR are normalized by LIMIT to control for size effects and correspond to positive numbers. We present the results for the medians of both variables in Figure 3.

Insert Figure 3 here

Figure 3 indicates that both CLR and DLR of the defaulters vary in the range 60% to 80% during event time -36 to -18, and drop sharply afterwards. Most of this sudden decrease happens during event time -18 to -12. We see that CLR and DLR are highly correlated, as hypothesized above. The CLR of defaulters decreases as firms' face contracting sales and individuals' loose parts of the personal income. Interestingly, the DLR is also decreasing, indicating that firms are stretching out their suppliers and that the bank pays attention at potential overdrafts. For comparison, CLR and DLR of all borrowers change very little, remaining above 70% for the entire time. We conclude that the CLR indicates abnormal patterns roughly 12 to 18 months prior to default.

3.2 Multivariate Regression Analysis

We now estimate multivariate probit models to study which factors at calendar time t and t-12 influence the probability of default at calendar time t+1. The dependent variable, our indicator for jumps to default DEF, equals one if a borrower exhibits a downgrade from ratings one to four at time t to ratings five or six at time t+1, and zero otherwise. Explanatory variables are the internal credit rating (the rating level RAT or rating changes ΔRAT), changes in credit line usage (ΔUSAGE), changes in the absolute account amplitude (ΔAMPLI), changes in the credit-to-limit ratio (ΔCLR), and changes in the cumulative number of limit violations (ΔCVIOL) during the periods [t, t-12] and [t-12, t-24]. The number of observations decreases since we cannot include accounts with a time series of less than 13 or 25 months and without a credit line. We add the percentage change of the credit line (ΔLIMIT) as a control variable because, except for AMPLI, we base the account variables on the payment behavior and the credit limit. However, the bank can change the credit limit to manage the loan exposure. For example, USAGE might rise if the borrower increasingly draws a constant line of credit, or if the the bank reduces the LIMIT and the account balance remains constant. We expect to find a negative sign for the coefficient of limit changes because limit reductions are one response the bank can make concerning borrowers with financial problems.

We calculate all variable changes over 12 months, because within this period the bank has to perform at least one internal rating review. This approach is prudent because we do not consider monthly variable changes, which are more likely to occur in checking accounts than in internal ratings. We also use models with changes of all variables over consecutive three- and six-month intervals and obtain similar results.

We estimate the following set of models: Model I includes only the level of the internal rating in month t. This model uses the ordinal rating variable on a scale from one to six. However, we find that if we include dummy variables with either rating one or three as the reference category, all subsequent results remain unchanged. Model II uses only changes of the account activity variables over the last 12 months. We believe it is reasonable to use dynamic information (changes) rather than static information (levels) to fully exploit the account activity. When we consider models that include levels of checking account variables, we obtain basically similar results. Model III uses the level of the internal rating in t and changes of account activity variables over the last 12 months. Model IV shows the changes of the internal rating and the changes of the account activity variables over the last 12 months (Table 2, Panel A). Furthermore, we consider different prediction horizons, e.g., time t-12 in Panel B, and both time t and t-12 in Panel C. Both the internal credit rating levels and rating changes act as a benchmark against which we can test the usefulness of account activity information. Table 2 reports the results for the full sample, spanning firms and individuals.

Insert Table 2 here

Our two key findings are that measures of account activity have a significant impact on the probability of default that goes beyond the internal credit rating. Further, that incorporating these measures substantially increases the fit of default predictions. In Table 2, Panel A shows that the internal credit rating is significant and positively associated with future default events (Model I). Models II, III, and IV indicate that all checking account variables are correctly signed and highly significant. Consistent with previous findings, the change of the credit line usage and the cumulative number of limit violations display a positive coefficient and the change of the absolute amplitude and the credit-to-limit ratio shows a negative coefficient. The coefficient of the control variable ΔLIMIT is statistically significant and negative. Furthermore, a comparison of the goodness-of-fit of the models, i.e., the McFadden R^2 adjusted for the number of regressors, shows that using measures of account activity leads to a better fit than does the rating level alone. Strikingly, combining both sources of information more than doubles the McFadden R^2, which is 0.028 in Model I compared to 0.083 in Model III or 0.07 in Model IV. This finding suggests that banks can improve their monitoring by incorporating measures of account activity in their borrower monitoring systems, i.e., the information is complementary.

The corresponding regressions models in Table 2, Panel B, include the explanatory variables lagged by 12 months, i.e., the rating level at month t-12 and changes of variables during the period [t-12, t-24]. In Panel C, the regression models include information from one and two years before default. We note that we cannot estimate Model I by including rating levels in t and t-12, since these are highly correlated. Essentially, the results from Panels B and C are consistent with the hypothesis that measures of account activity are useful to predict borrower defaults.

One additional finding is that the account amplitude (AMPLI) is a good short-run default predictor, as can be seen in Figure 2. The change of this variable is statistically significant and correctly signed in the year before default, but it becomes nonsignificant over the period t-12 to t-24 (see Panels B and C). Finally, information on account activity is not only more timely, but also more volatile, than the constituents of the internal credit ratings. How this higher variation affects the value of this information is an empirical issue. Our results indicate that the benefits from an increased timeliness outweigh the disadvantage from a higher volatility.

3.3 Results by Borrower Type

We investigate whether the previous results hold if we differentiate by large firms, small firms and individuals based on TYPE, because the complexity of cash flows in the checking accounts, the likelihood of having multiple banking relationships, and the mechanism of default differ considerably across these borrower types. Moreover, banks apply different lending technologies (Berger and Udell, 2006). In addition, credit lines granted to firms are frequently secured by a collateral pool, the value of which is independent of the effective credit line usage, while credit lines attached to consumer checking accounts are typically unsecured. As noted above, in Germany, funding with credit cards does not play an important role for either businesses (similar to the U.S.) or for individuals (unlike in the U.S.). Consequently, information available to banks is not fragmented because of credit cards. The size and structure of our data set and, more important, these institutional differences across borrower groups make it possible for us to perform an extensive test of the usefulness of information on account activity.

We now conduct an event study in which we use borrower-specific benchmarks. We compare measures of account activity for defaulters in each of the categories of TYPE with the median of all borrowers in the same category. This analysis not only highlights differences across categories, but is also more accurate than the analysis in the previous section because the benchmarks are now borrower type-specific. Figure 4 illustrates the results for the account amplitude AMPLI, which is our measure for the account balance variation within a month.

Insert Figure 4 here

The event study confirms our findings on the full sample. In all subgroups we find that the amplitude of defaulters decreases prior to default, but the amplitude of all borrowers change very little. However, there are differences between borrower types. Differentiating by TYPE

(Figure 4, Panels A, B, and C) indicates that a systematic drop of AMPLI below the benchmark occurs six to twelve months before default at small firms (Panel B), five months before default for individuals (Panel C), and only slightly before default at large companies (Panel A). There are several plausible explanations for this result. The quality of hard information for large firms is better, due to financial reporting standards, disclosure rules, and auditors, which is consistent with the significant ability of the internal credit rating to predict defaults. Further, large firms also benefit from different sources of funding and a higher number of bank relationships. We repeat the same univariate event study for our other measures of account activity and obtain qualitatively similar results.

In a next step, we re-estimate the probit regression models from the previous section separately for the categories of TYPE. Table 3 summarizes the estimation results. To conserve space and because the other models lead to similar outcomes, we only report the results for the specification that corresponds to Model IV (Table 2, Panel C).

Insert Table 3 here

For large firms, only the change of the internal credit rating and the cumulative number of violations display a significant and positive coefficient. For small firms, ΔUSAGE, ΔAMPLI, and ΔCVIOL at t and t-12, and the internal credit rating change at t and t-12 have a significant influence on the probability of default. For individuals, we find that all contemporaneous account activity variables are correctly signed and highly significant. In addition, in contrast to firms, the internal rating change at t-12 has no impact.

We note that the change of the cumulative number of limit violations (ΔCVIOL) performs especially well. The change of CVIOL over the preceding 12 months exhibits a significant and positive coefficient for all categories of TYPE (except at t-12 for large firms). As hypothesized, the coefficients of ΔLIMIT are significantly negative at time t for all categories.

Our results on measures of account activity do not change if we exclude ΔLIMIT from the regression models.

Given these differences across borrower types, we conduct a further analysis to investigate what banks can gain in economic terms. To do so, we test whether the benefit is highest for monitoring small firms and individuals, as suggested by the previous analysis. We examine how the goodness-of-fit of default prediction models, measured by the adjusted McFadden R^2, is affected if we incorporate both information on account activity and internal credit ratings. In unreported regression analyses (model specifications as in Table 2, Panel C), we find that banks gain most if they use measures of account activity for monitoring small firms and individuals. The adjusted McFadden R^2 increases from 1.2% to 7.0% for small firms and from 1.9% to 6.6% for individuals, which is a substantial improvement in absolute and relative terms. For large firms, we find only a small increase from 9.2% to 9.7%. This result is consistent with the fact that default prediction models based on internal credit ratings or rating changes perform best for large firms. One explanation for this result is that input factors of ratings for large firms tend to be more sensitive to the deterioration of a borrower's credit quality than to information on account activity. As noted earlier, ratings for large companies are based on a broad set of public information, including various financial ratios and other accounting numbers, and private soft information. However, measures of account activity for large firms might be fragmented due to multiple bank relationships, but the opposite holds true for small firms and individuals. For these borrowers, the input factors of the credit rating are relatively sticky, but the account activity reflects cash flows in a timely manner, and therefore indicates incremental changes in a borrower's credit quality in real time.

We note that the question of how the value of account activity compares to the value of credit bureau information goes beyond the scope of this paper (for details see, e.g., Kallberg and Udell 2003). Account activity is one source of timely, first-hand private information for banks, while credit bureaus help sharing existing information.

3.4 Checking Account Information and Bank-Relationship Characteristics

Because the previous findings may be sensitive to characteristics of the bank-borrower relationship, we extend our analysis by considering two key variables that are frequently used in related studies such as those by Berger and Udell (1995), Petersen and Rajan (2002), and Degryse and Ongena (2005). These variables are the duration of the bank-borrower relationship and the physical distance between the domicile of the borrower and the bank's head office. Both variables are indicators of the intensity of the bank-borrower relationship and may also serve as proxies for the extent of asymmetric information in the bank relationship (Boot 2000; Elyasiani and Goldberg 2004). From a theoretical perspective, we expect to find that the usefulness of account activity for monitoring rises over time and falls with physical distance, because the bank gathers an increasing amount of private information. Single pieces of account activity are quantitative information, but the sum of these pieces may convey a qualitative message. The latter is analogous to the bank's general assessment that emerges from a clear, robust view on the borrower's cash flows. Table 4 presents the probit regressions results differentiated by duration and distance.

Insert Table 4 here

When we consider duration (Panel A), we find that for long-duration borrowers, account activity is more informative than are ratings. We note that changes of the internal credit ratings at time t and t-12 are only weakly related to future defaults. This finding suggests that the reliability of information on account activity increases over the course of a bank-borrower relationship. In contrast, account activity and rating information tend to be complementary for predicting defaults for short-duration borrowers. Panel B shows that credit ratings and account information are both significantly related to defaults of short-distance borrowers but less important for long-distance borrowers. Again, limit violations (ΔCVIOL) are especially informative in most cases.

These results show that this type of private information is particularly useful if there is a close bank-borrower relationship. Banks are better able to collect and validate private information on account activity the longer the bank borrower relationship and the lower the physical distance. This finding is consistent with the evidence that borrowers have their checking accounts at nearby banks and that the latter are usually the borrowers' main bank relationship (e.g., Berger et al. 2005). Stated differently, the longer the time series on account activity the more reliable the "learning effect" for the bank. Our findings are also consistent with findings from cross-sectional studies and imply that main banks can benefit most (see, e.g., Agarwal and Hauswald, 2007).

4. Empirical Analysis of Bank Behavior based on Account Activity

To examine how banks use account activity information in managing their relationships, we study the impact of warning indications on incremental lending decisions such as loan pricing, limit changes, and account closures. We also examine the account activity conditional on the type of default.

4.1 The Impact on Loan Pricing

We now investigate whether information on account activity influences the pricing of loans. We exploit the fact that loan officers may make discretionary use of account information in lending decisions.[6]

We merge our initial data with an additional data set on all new loans to firms granted by the same bank in 2005, comprising 643 straight loans with a total of 83 million euros. This data set displays the following median (mean) loan terms: Spreads of 227 bps (251 bps), loan amount of 30,000 euros (129,978 euros), collateral-to-loan ratio of 37% (44%), and a maturity of 50 (79) months. By using this combined data set, we can study the terms of the bank's incremental lending decisions conditional on the existence of checking accounts and lines of credit, and early warning signals from credit line usage and account activity.

From a theoretical perspective it is not clear if and how information on account activity might affect lending decisions. For instance, there might be no impact on loan terms if the internal credit rating is the only pricing factor. Moreover, as suggested by Vale (1993), a superior monitoring technology that results from the access to account information might lower the price of credit. However, the impact of account information on loan terms may be more differentiated. It may depend on whether early warning indications are positive or negative, and whether the bank exploits the information symmetrically or asymmetrically. If the bank uses the information asymmetrically, we expect to find higher loan rates when there are negative early warning indications, but no effect on loan rates when there is neutral or favorable account activity.

To address this issue, we estimate cross-sectional regressions using loan spreads as the dependent variable (SPREAD). We use credit ratings (RAT), a dummy variable indicating the existence of a checking account at the same bank (CHECK), and controls as explanatory variables. We then refine the model by replacing CHECK with indicator variables for the account activity (USAGE_HIGH, VIOL_HIGH) and by adding further borrower characteristics and loan terms. We note that including indicator variables for the individual ratings instead of the ordinal variable, RAT, does not change our findings for the impact of CHECK, USAGE_HIGH, and VIOL_HIGH. Table 5 reports our results.

Insert Table 5 here

In Table 5, Panel A shows that borrowers with checking accounts (CHECK) at the same bank pay higher loan spreads than do firms without checking accounts. The coefficient of

CHECK is highly significant and amounts to 40 bps in Model I and 89 bps in Model II, when we include a series of control variables and industry fixed effects. These findings are consistent with univariate results from a nonparametric Wilcoxon rank sum test (not reported here), indicating a significant loan spread difference of 60 bps between borrowers with and without checking accounts. In addition, the coefficient of CHECK becomes larger if we consider only unsecured loans (120 bps) or relatively small loans (135 bps). These are loans for which early warning signals are particularly valuable for the bank.

Panel B indicates that this result is driven mainly by higher loan spreads for borrowers for whom the bank observes a negative account activity, such as high credit line usage or high number of limit violations during the 12 months preceding the loan decision. Model III shows this result in the coefficient of USAGE_HIGH, which amounts to 44 bps (p-val. = 0.06). In addition, the result is consistent with the significantly positive coefficient of 72 bps for VIOL_HIGH in Model IV. There is no impact on loan spreads if there are positive or no signals. Thus, we find evidence that the bank exploits information on account activity asymmetrically, i.e., the pricing impact exists only for borrowers who exhibit increases in default risk. The latter result seems plausible, since the bank may temporarily earn rents instead of lowering the loan spreads when it receives positive signals. Furthermore, our finding is consistent with the view that loans to firms with no checking account at the bank are likely to be secondary relationships. These loans may be underpriced because of bank competition (see for pricing effects due to bank switching, e.g., Ioannidou and Ongena 2007). This interpretation is supported by the fact that firms without checking accounts exhibit a substantially lower duration of the bank-firm relationship (median is 1.2. years) than firms with checking accounts (median is 4.7 years).

In addition, if we restrict the analysis to firms that have checking accounts with lines of credit (n=247), we find that the impact is economically and statistically most significant for the best ratings 1 and 2, but the effects are marginally significant for rating 3 and not significant for rating 4. If the rating has already deteriorated to rating 4, then the marginal value of negative early warning signals from checking account activity is relatively low. In unreported results, we find that reductions of the credit limit in the year before the loan approval result in higher loan spreads for same-rated borrowers.

4.2. The Type of Default

To investigate whether the account activity differs by the type of default, we first distinguish between borrowers' account activity before "hard" and "soft" defaults. Hard defaults are downgrades from ratings 1-4 to rating 6 (bankrupcty filing), and soft defaults are downgrades from ratings 1-4 to rating 5 (first specific loan loss provision).

Hard defaults are publicly observable and exogenous, while the timing and occurrence of soft defaults are partly under control of the bank. This classification does not necessarily refer to the magnitude of economic losses, but rather to the timing and the legal status of financial distress. Typically, the first specific loan loss provision precedes a bankruptcy filing. Because of this sequencing, we expect to find clearer warning indications for hard defaults compared to soft defaults. Figure 5 shows the event study results for credit line usage (USAGE) for soft defaults (474 downgrades from rating 1-4 to rating 5 and no subsequent downgrades) and hard defaults (535 direct downgrades from rating 1-4 to rating 6 or temporary downgrades from rating 1-4 to rating 5 that end up in rating 6).

Insert Figure 5 here

During the event time -36 to -24 months, credit line usage (USAGE) is similar for both types of defaults but different from the nondefaulters. From event time -24, the credit line usage of borrowers who experience hard defaults is always above the line for soft defaults. Most important, the credit line usage prior to hard defaults increases sharply from 30% to almost 100% prior to default, while the line for soft defaults is relatively volatile and increases only slightly. The real-time character of account activity explains why this information is more informative for borrowers that file for bankruptcy (hard defaults) compared to those that remain pending in rating 5 (soft default).

Second, we investigate the relation between account activity and the bank's expectations about the loss given default. Our data set includes information on whether the bank keeps or closes an account at default. Since account closures at default coincide with complete write-offs, we can differentiate between "default and complete write-off" (n=268) and "default and partial write-off" (n=741).[7] We expect to find that the bank is more likely to write off the exposure completely at default if it has observed early warning indications for a long time before default. The univariate event study and the multivariate probit models show that all measures of account activity (USAGE, CVIOL, AM, and CLR) indicate an earlier and stronger deterioration of the credit quality if the bank writes off the loan completely. For example, USAGE exhibits an early and fast increase from zero at event time -36 to 100% at event time -18 for defaulters whose accounts are closed at default, but defaulters who experience no account closure at default exhibit a less steep increase from 30% to 50%. This finding supports the differentiation by hard and soft default, because we observe complete write-offs at default three times more often for hard defaults than for soft defaults.

Third, we compare the usefulness of account activity for monitoring before defaults that are relatively surprising compared to those that are less surprising. For this purpose, we condition the analysis on the rating at the beginning of the sample period and on the rating one month before default. The variable USAGE is very informative for defaulters from relatively good ratings (1, 2, or 3) in both tests, but it is not at all useful for rating 4. The same result holds true for the absolute amplitude, AMPLI. In an unreported multivariate probit regression, we find that account activity is especially useful if it provides supplementary information that goes beyond what is already incorporated in the internal credit ratings. This finding is in line with our result from Section 4.1, that spreads on new loans to high quality borrowers are adversely affected if the bank obtains early warning indications.

5. Tests of Robustness

5.1 Miscellaneous Empirical Checks

First, we estimate a two-stage regression model to analyze the overlap of information in credit ratings and account activity. In the first stage, we regress changes in the internal credit ratings on changes in the account activity variables. In the second stage, we explain future defaults by using the residuals from the first stage, i.e., the component in credit rating changes that is left unexplained by account activity. We find that the residuals from the first stage are significant and positively associated with subsequent borrower defaults, indicating that information on account activity has predictive power that goes beyond the information included in credit ratings. This result confirms the findings from Section 3.

Second, in our two-stage regression model, instead of using the rating changes over the period [t, t-12] to predict defaults at time t+1, we use the changes of the probability of default (PD) associated with each rating. This test explicitly takes into account that a move from rating 3 to rating 4 corresponds to a considerably larger change in the average PD than a move from rating 1 to rating 2. We find that the coefficient of ΔPD_t is positive and highly significant ($p<0.01$). The sign and the statistical significance of account activity measures do not change in comparison to those in Table 2. Thus, we conclude that our previous findings are robust if we control for nonlinearities in default risk changes.

Third, to determine if extreme observations influence the results, we winsorize the variables USAGE, AMPLI, CLR, and LIMIT at the 0.5% and 99.5% levels. Repeating all previous analyses with winsorized explanatory variables leads to slightly higher coefficients and similar findings. The impact of credit line usage (USAGE) is slightly reduced in terms of statistical significance in comparison to Table 2, but the results for the other account activity measures remain unchanged.

Fourth, we separately re-estimate Model IV from Table 2, Panel C, for the years 2002-2004 (378 defaults) and the years 2005-2006 (631 defaults). We find that rating changes have an impact on future defaults in both subsamples. Two out of four account activity variables (AM, CVIOL) are significant and correctly signed in the first half of the sample, and all four in the second half. Given that our data cover only five years, we cannot say whether this result is due to an increasing reliability of account information or because of the higher number of defaults in the second half. Most important, we observe that at least some measures of account activity are significantly related to subsequent defaults in both subsamples.

5.2 Checking Accounts with and without Lines of Credit

Our empirical results in Section 3 are based on data from checking accounts with a credit line, and cover 55% of all account-month observations and 49% of all borrowers in the original sample. Here, we investigate whether measures of account activity are also useful for monitoring borrowers who do not have a line of credit but do have other loans, such as investment loans, mortgages, or consumer loans. This test analyzes whether our results are driven by a selection bias.

To perform this investigation we modify the probit models from Table 2 as follows. We use DEF_{t+1} as the dependent variable and include a reduced number of explanatory variables, the rating change and the change of the account amplitude, and a new variable, the cumulative number of overdrafts (OVER). The latter variable counts the number of limit violations for borrowers who have a credit line and overdrafts for borrowers who do not have a credit line. We define an overdraft as one in which the minimum balance is below zero euros. We exclude all variables that can only be observed for borrowers with a credit line (USAGE, CLR, LIMIT). Thus, the number of observations available for model estimation increases by

72% (from 1.2 million to 2.2 million). First, we estimate the previously described model on all observations. Second, we consider only data from accounts without a credit line.

The regressions show that both the change of the internal rating (ΔRAT) and the cumulative number of overdrafts (ΔOVER) are highly significant and positively related to future defaults, but the account amplitude is not. In addition, differentiating by TYPE indicates that ΔOVER is significant for each category, and that the magnitude of the estimated coefficient increases from 0.063 (large firms) and 0.071 (small firms) to individuals (0.085). Furthermore, for accounts without credit lines, we observe that the cumulative number of overdrafts is the only variable that has a significant impact on the probability of subsequent defaults.

5.3 Single- and Multiple-Account Bank Relationships

The type of the bank relationship might also affect the extent to which the account activity is useful for monitoring: the number, the purpose, and the relative importance of the accounts in the case of multiple-account clients could be important. Firms may have several separate accounts for subsidiaries, products, or purposes, and couples can have joint or separate checking accounts. Hence, banks need to find a way to aggregate the account information.

We consider the structure of the bank relationship in two ways. We repeat the analysis for borrowers who have only one account with the bank, thus there is no fragmentation across accounts. This approach is restrictive, since we lose 40% of all observations. Therefore, we consider a second approach in which we rank all accounts per borrower based on the monthly mean credits and the length of time series. Spearman's rank correlation coefficient between both criteria is 0.77, indicating a positive, but not perfect, correlation. Based on these rankings, we compare default predictions for single-account borrowers with those of multiple-account borrowers, considering only their most important account. Following this approach, we drop only 10% of all observations. Table 6 presents the results.

Insert Table 6 here

In Table 6, Panel A shows that the results on the account activity variables for single-account relationships are similar to those in Table 2. For multiple-account bank relationships, we find that the rating is relatively more important; nevertheless, three out of four account activity measures display the expected sign and are statistically significant. In Panel B, we use two approaches to identify the most important accounts and obtain results similar to those in Table 2, Panel A.

6. Conclusion

We investigate the link between account activity and information production on borrower quality. Theoretical research assumes that banks have an informational advantage over nonbank lenders and capital markets, but almost entirely missing from the literature is direct evidence on the exact sources of this advantage. We attempt to fill this gap by examining borrowers' account activity as one important source of private information for banks.

For this purpose, we investigate whether credit line usage and cash flows in a borrower's checking account are helpful for monitoring, and how banks use this information. Our analysis is based on a unique data set that includes more than three million account-month observations for the period 2002-2006. This data set makes it possible for us to distinguish between firms and individuals, different lending technologies, and types of default.

The two principal results of our study are that account activity is very informative, and that banks use this information in managing their lending relationships. Early warning indications from account activity help significantly to predict future borrower defaults approximately one year in advance, and serve as a basis for banks' loan pricing, credit limit management, and loan loss provisioning. We document that incorporating information on account activity improves default prediction models substantially, and that it is especially useful for monitoring small businesses and individuals. The value of account activity for monitoring increases with the duration of the bank-borrower relationship and decreases with physical distance. This finding suggests that relationship lenders can benefit most from this source of information.

We also show that banks use the information on account activity in managing their lending relationships. Firms with a checking account pay higher credit spreads on loans if the bank has obtained early warning indications from credit line usage and checking account activity during the year before the new loan is granted. Moreover, integrated account information is more useful for monitoring if borrower defaults are severe events such as bankruptcies and/or total write offs, and relatively surprising events such as defaults of high grade borrowers. Both results show that banks benefit from the timely and proprietary nature of this information.

We identify a key source of private information that comprehensively explains why banks are "special" vis-à-vis nonbank lenders and capital markets, and why banks apply specific lending technologies to certain borrowers, such as relationship lending to small businesses. We provide universal evidence on the role of account activity in delivering relationship lending and thus augment the scarce literature in this field of research (Mester, Nakamura, and Renault 2007; Jiménez, Lopez, and Saurina 2009). Moreover, our findings have implications for financial system architecture, suggesting that information production and risk taking should not be separated. Future research may investigate in more detail both the bright and dark sides of banks' use of this kind of information in the context of keeping or stopping lending relationships, as well as the level of assistance banks provide to borrowers in financial distress.

References

Agarwal, S., B. Ambrose, and C. Liu. 2006. Credit Lines and Credit Utilization. *Journal of Money, Credit, and Banking* 38:1-22.

Agarwal, S., and R. Hauswald. 2007. The Choice between Arm's-Length and Relationship Debt: Evidence from eLoans. Working Paper.

Apilado, V., D. Warner, and J. Dauten. 1974. Evaluative Techniques in Consumer Finance -- Experimental Results and Policy Implications for Financial Institutions. *Journal of Financial and Quantitative Analysis* 9:275-283.

Basel Committee on Banking Supervision. 2006. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework.

Berger, A., N. Miller, M. Petersen, R. Rajan, and J. Stein. 2005. Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics* 76:237-269.

Berger, A., and G. Udell. 1995. Relationship lending and lines of credit in small firm finance. *Journal of Business* 68:351-381.

Berger, A., and G. Udell. 2006. A more complete conceptual framework for SME finance. *Journal of Banking and Finance* 30:2945-2966.

Berlin, M., and L. Mester. 1999. Deposits and relationship lending. *Review of Financial Studies* 12:579-607.

Bharath, S., S. Dahiya, A. Saunders, and A. Srinivasan. 2007. So what do I get? The bank's view of lending relationships. *Journal of Financial Economics* 85:368-419.

Black, F. 1975. Bank Funds Management in an efficient market. *Journal of Financial Economics* 2:323-339.

Blackwell, D., and D. Winters. 1997. Banking relationships and the effect of monitoring in loan pricing. *Journal of Financial Research* 20:275-289.

Boot, A. 2000. Relationship Banking: What do we know? *Journal of Financial Intermediation* 9:7-25.

Boyd, J., and E. Prescott. 1986. Financial intermediary-coalitions. *Journal of Economic Theory* 38:211-232.

Chakravarty, S., and J. Scott. 1999. Relationships and rationing in consumer loans. *Journal of Business* 72:523-544.

Cole, R., L. Goldberg, and L. White. 2004. Cookie-Cutter versus Character: The Micro Structure of Small-Business Lending by Large and Small Banks. *Journal of Financial and Quantitative Analysis* 39, 227-251.

Degryse, H., and S. Ongena. 2005. Distance, Lending Relationships, and Competition. *Journal of Finance* 60:231-266.

Deutsche Bundesbank. 2008. Statistics on Payment Systems in Germany. January 2008.

Diamond, D.W. 1984. Financial Intermediation and Delegated Monitoring. *Review of Economic Studies* 51:393-414.

Diamond, D.W. 1991. Monitoring and Reputation: The Choice Between Bank Loans and Directly Placed Debt. *Journal of Political Economy* 99:689-721.

Eisfeld, C. 1935. Kontenanalyse im Dienste der Kreditbeobachtung. Sparkasse 55:333-345.

Elsas, R., and J. Krahnen. 1998. Is relationship special? Evidence from credit-file data in Germany. *Journal of Banking and Finance* 22:1283-1316.

Elsas, R. 2005. Empirical Determinants of Relationship Lending. *Journal of Financial Intermediation* 14:32-57.

Elyasiani, E., and L. Goldberg. 2004. Relationship lending: a survey of the literature. *Journal of Economics and Business* 56:315-330.

Fama, E. 1985. What's different about banks? *Journal of Monetary Economics* 15:29-39.

Federal Statistic Office. 2006. Business register: Enterprises by employees liable to pay social insurance contributions and taxable turnover.

Grunert, J., L. Norden, and M. Weber. 2005. The role of non-financial factors in internal credit ratings. *Journal of Banking and Finance* 29, 509-531.

Hackl, P., and A. Schmoll. 1990. Kontodatenanalyse zur Früherkennung von Kreditrisiken. *Österreichisches Bank-Archiv* 38:870-881.

Ho, T., and A. Saunders. 1983. Fixed Rate Loan Commitments, Take-Down Risk, and the Dynamcis of Hedging with Futures. *Journal of Financial and Quantitative Analysis* 18:499-516.

Ioannidou, V., and S. Ongena. 2007. Time for a change: Loan Conditions and Bank Behavior when Firms Switch. Working paper, Tilburg University, forthcoming *Journal of Finance*.

Jiménez, G., J. Lopez, and J. Saurina. 2009. Empirical analysis of corporate credit lines. *Review of Financial Studies* 22:5069-5098.

Kallberg, J., and G. Udell. 2003. The value of private sector business credit information sharing: The US case. *Journal of Banking and Finance* 27:449-469.

Kashyap, A., R. Rajan, and J. Stein. 2002. Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking. *Journal of Finance* 57:33-73.

Kirschenmann, K., and L. Norden. 2008. The relation between borrower risk and loan maturity in small business lending. Working Paper, May 2008, University of Mannheim.

Machauer, A., and M. Weber. 1998. Bank Behavior Based on Internal Credit Ratings of Borrowers. *Journal of Banking and Finance* 22:1355-1383.

Melnik, A, and S. Plaut. 1986. Loan Commitment Contracts, Terms of Lending, and Credit Allocation. *Journal of Finance* 41:425-435.

Mester, L., L. Nakamura, and M. Renault. 2007. Transaction accounts and loan monitoring. *Review of Financial Studies* 20:529-556.

Nakamura, L. 1993. Commercial Bank Information: Implications for the Structure of Banking. In: Klausner, Michael, and Lawrence J. White (eds), Structural Change in Banking, Business One/Irwin, Homewood, IL.

Petersen, M. 2004. Information: Hard and Soft. Working Paper, July 2004.

Petersen, M., and R. Rajan. 1994. The Benefits of Lending Relationships: Evidence from Small Business Data. *Journal of Finance* 49:3-37.

Petersen, M., and R. Rajan. 2002. Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance* 57:2533-2570.

Rajan, R. 1992. Insiders and Outsiders: The Choice Between Informed and Arm's-Length Debt. *Journal of Finance* 47:1367-1406.

Ramakrishnan, R.T., and Thakor, A.V. 1984. Information Reliability and a Theory of Financial Information. *Review of Economic Studies* 51:415-432.

Schlüter, B. 2005. Empirische Analyse der Bedeutung interner Informationen von Kreditinstituten für die Bonitätsprüfung. Shaker: Aachen, 2005.

Shockley, R., and A.V. Thakor. 1997. Bank loan commitment contracts: Data, Theory, and Tests. *Journal of Money, Credit, and Banking* 29:517-534.

Sufi, A. 2009. Bank lines of credit in corporate finance: an empirical analysis. *Review of Financial Studies* 22:1057-1088.

Treacy, W., and M. Carey. 2000. Credit risk rating systems at large US banks. *Journal of Banking and Finance* 24:167-201.

Vale, B. 1993. The dual role of demand deposits under asymmetric information. *Scandinavian Journal of Economics* 95:77-95.

Von Stein, J. 1983. Früherkennung von Kreditrisiken mit quantitativen Methoden. *Betriebswirtschaftliche Blätter* 57:367-368.

Footnotes

[1] Our paper is about informational synergies. There is also evidence for liquidity and cost synergies, delivering rationales for the simultaneous supply of deposit taking and lending by banks (Berlin and Mester 1999; Kashyap, Rajan, and Stein 2002).

[2] For comparison, credit cards account for 14% of the relative frequency (25% of the total value) of all payment card transactions in the year 2006.

[3] We do not base our data selection on a particular point in time, but on the entire sample period. We include clients who started borrowing before 2002 but who are no longer in the sample because either they switched banks or went bankrupt. We also include borrowers who began borrowing from the bank during the 2002-2006 sample period. For example, there are 58,790 (51,069) accounts with a time series of at least 36 (48) months. We have data on exactly 60 months for 6,235 accounts.

[4] Our results are robust to alternative thresholds and classifications to distinguish between large and small firms. For example, if we distinguish between large and small firms based on total annual sales of 500,000 euros, which corresponds approximately to the mean annual sales of firms in our sample, we obtain similar results. A classification based on an exogenous variable, e.g., the firms' legal form (corporations vs. sole proprietorships), leads to the same conclusions since in Germany larger firms are more likely to be incorporated.

[5] We also calculate the corresponding checking account variables for all nondefaulters and all borrowers. Our results are almost identical to those that we report here, since the impact of the defaulters on the average variables of the entire sample is very small.

[6] We note that in mid-2007, the bank introduced new internal rating systems to comply with the new capital adequacy framework (Basel II). Measures of account activity are now important components of the ratings for small businesses and individuals, indicating that there has been a switch from a discretionary to a rules-based use of this information.

[7] Because of data limitations we are unable to conduct a more detailed analysis of the time-in-default and the recovery rate.

Table 1

Summary statistics

We obtain our data from a German universal bank. The sample period is January 2002 to December 2006. In this table, Panel A provides the number of borrowers, accounts, and account-month observations by borrower type (TYPE) and for the full sample. We differentiate between large firms (TYPE=1), small firms (TYPE=2), and individuals (TYPE=3). Large firms are defined as firms with total annual sales above one million euros (calculated as monthly mean cumulative credits x 12). Panel B shows account activity and bank-relationship variables. The variables LOW, …, LIMIT are reported in thousands of euros and refer to the account level. LOW and HIGH may be negative or positive. We define DEBIT and CREDIT as positive numbers. MID can be positive or negative. We define LIMIT as a negative number. The row titled "Duration" indicates the duration of the bank relationship in years and "physical distance" measures the distance between the bank and the borrower based on first three digits of the borrower's and bank's ZIP code. The internal rating RAT ranges from one (best) to six (worst). Panel C shows the frequency of rating changes and default events. We define default events as internal rating changes from ratings 1-4 to rating 5 or 6. The bank assigns a borrower to rating 5 when it establishes a specific loan loss provision for the first time. A rating 6 indicates that the borrower has filed for bankruptcy. All events refer to the account level

Panel A: Number of borrowers, accounts, and account-month observations

Statistic	Large firms	Small firms	Individuals	Full sample
Number of borrowers	529	7,374	59,673	67,215
Number of accounts	585	10,474	75,886	86,945
Number of account-months	29,650	430,611	2,811,618	3,271,879

Panel B: Account activity and bank-relationship variables

		Large firms		Small firms		Individuals		Full sample	
Variable	Variable description	Mean	Median	Mean	Median	Mean	Median	Mean	Median
LOW	Min account balance per month	-49.12	0.00	-5.06	0.00	1.10	0.00	-0.16	0.00
HIGH	Max account balance per month	142.36	36.00	0.99	1.00	3.18	1.00	4.15	1.00
MID	Monthly average account balance	46.61	10.00	-2.03	0.30	2.13	0.50	1.99	0.50
DEBIT	Monthly cumulative debits	552.57	152.28	8.21	1.00	2.31	1.00	8.04	1.00
CREDIT	Monthly cumulative credits	560.43	154.00	8.42	1.00	2.42	1.00	8.22	1.00
LIMIT	Credit limit	-244.29	-100.00	-21.72	-5.11	-3.34	-2.04	-8.12	-2.50
DUR	Duration of bank relationship (years)	11.72	10.02	9.29	6.57	10.75	8.23	10.56	7.98
DIST	Physical distance (kilometers)	8.86	0.00	6.46	0.00	7.30	0.00	7.21	0.00
RAT	Internal credit rating (1 to 6 scale)	2.47	2.00	2.94	3.00	2.87	3.00	2.88	3.00

Panel C: Frequency of rating changes and default events

Variable	Large firms	Small firms	Individuals	Full sample
Rating changes	418	2,786	9,599	12,803
hereof upgrades	179	1,231	4,105	5,515
hereof downgrades	239	1,555	5,494	7,288
Defaults	32	347	630	1,009
… hereof rating 5	14	160	367	541
… hereof rating 6	18	187	263	468

Table 2

Probit regression results for the entire sample

In this table, the dependent variable in the regressions is default at time t+1 (DEF_{t+1} equals one for jumps to default and zero otherwise). Explanatory variables are the internal credit rating (RAT), the rating change (ΔRAT), the change of the credit line usage (ΔUSAGE), the change of the absolute account amplitude (ΔAMPLI), the change of the credit-to-limit ratio (ΔCLR), the change of the cumulative number of limit violations (ΔCVIOL), and the relative change of the credit limit (ΔLIMIT) during the period [t, t-12] and the period [t-12, t-24]. We divide the variables ΔUSAGE and ΔCLR by 100 to scale the estimated coefficients. The regressions take into account the clustering of observations at the account level and are based on p-values from Huber-White robust standard errors.

Panel A: Estimation results, one year before default

	Model I			Model II			Model III			Model IV		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.	Coeff.		p-val.	Coeff.		p-val.
RAT_t	0.266281	***	0.000				0.208525	***	0.000			
ΔRAT_t										0.211167	***	0.000
$\Delta USAGE_t$				0.000144	*	0.063	0.000208	**	0.015	0.000153	*	0.059
$\Delta AMPLI_t$				-0.000146	***	0.001	-0.000172	***	0.000	-0.000131	***	0.000
ΔCLR_t				-0.000106	***	0.001	-0.000120	***	0.000	-0.000109	***	0.000
$\Delta CVIOL_t$				0.077724	***	0.000	0.071964	***	0.000	0.076983	***	0.000
$\Delta LIMIT_t$							-0.087448	***	0.000	-0.091155	***	0.000
Const.	-4.253245	***	0.000	-3.644547	***	0.000	-4.24081	***	0.000	-3.64982	***	0.000
McFadden Adj. R^2	0.0280			0.0660			0.0830			0.0700		
Obs.	1,276,045			1,276,045			1,276,045			1,276,045		

Table 2 (continued)

Panel B: Estimation results, two years before default

	Model I			Model II			Model III			Model IV		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.	Coeff.		p-val.	Coeff.		p-val.
RAT_{t-12}	0.236937	***	0.000				0.183783	***	0.000			
ΔRAT_{t-12}										0.223531	***	0.000
$\Delta USAGE_{t-12}$				0.000151	**	0.048	0.000214	**	0.015	0.000158	*	0.051
$\Delta AMPLI_{t-12}$				-0.000115		0.395	-0.000162		0.242	-0.000108		0.421
ΔCLR_{t-12}				-0.000821	***	0.005	-0.000092	***	0.002	-0.000079	***	0.006
$\Delta CVIOL_{t-12}$				0.070431	***	0.000	0.065913	***	0.000	0.069721	***	0.000
$\Delta LIMIT_{t-12}$							-0.115875	***	0.000	-0.123281	***	0.000
Const.	-4.121694	***	0.000	-3.59118	***	0.000	-4.111816	***	0.000	-3.596531	***	0.000
McFadden Adj. R^2	0.0200			0.0580			0.0700			0.0620		
Obs.	810,830			810,830			810,830			810,830		

Panel C: Estimation results, one and two years before default

	Model II			Model III			Model IV		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.	Coeff.		p-val.
RAT_t				0.215205	***	0.000			
ΔRAT_t							0.227892	***	0.000
$\Delta USAGE_t$	0.000413	***	0.000	0.000437	***	0.000	0.000425	***	0.000
$\Delta AMPLI_t$	-0.000188	***	0.009	-0.000256	***	0.000	-0.000155	**	0.028
ΔCLR_t	-0.000129	**	0.026	-0.000184	***	0.000	-0.000166	***	0.002
$\Delta CVIOL_t$	0.056748	***	0.000	0.051520	***	0.000	0.055201	***	0.000
$\Delta LIMIT_t$				-0.131189	***	0.000	-0.137544	***	0.000
RAT_{t-12}									
ΔRAT_{t-12}							0.245161	***	0.000
$\Delta USAGE_{t-12}$	0.000331	***	0.008	0.000443	***	0.003	0.000357	***	0.009
$\Delta AMPLI_{t-12}$	-0.000139		0.109	-0.000184	**	0.023	-0.000128		0.162
ΔCLR_{t-12}	-0.000035		0.452	-0.000072	**	0.011	-0.000065	**	0.015
$\Delta CVIOL_{t-12}$	0.034718	***	0.000	0.032850	***	0.000	0.034796	***	0.000
$\Delta LIMIT_{t-12}$				-0.013631		0.161	-0.013702		0.149
Const.	-3.63967	***	0.000	-4.254466	***	0.000	-3.652111	***	0.000
McFadden Adj. R^2	0.0750			0.0930			0.0840		
Obs.	810,830			810,830			810,830		

Table 3

Probit regression results by borrower type

In this table, we present probit regression results by borrower type (TYPE). We differentiate between large firms (TYPE=1), small firms (TYPE=2), and individuals (TYPE=3). Large firms are defined as firms with total annual sales above one million euros (calculated as monthly mean cumulative credits x 12). The dependent variable is default at time t+1 (DEF_{t+1} equals one for jumps to default and zero otherwise). The explanatory variables are the rating change (ΔRAT), the change of the credit line usage (ΔUSAGE), the change of the absolute account amplitude (ΔAMPLI), the change of the credit-to-limit ratio (ΔCLR), the change of the cumulative number of limit violations (ΔCVIOL), and the relative change of the credit limit (ΔLIMIT) during the period [t, t-12] and the period [t-12, t-24]. We divide the variables ΔUSAGE and ΔCLR by 100 to scale the estimated coefficients. Regressions take into account the clustering of observations at the account level and are based on p-values from Huber-White robust standard errors.

	Large firms (TYPE=1)			Small firms (TYPE=2)			Individuals (TYPE=3)		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.	Coeff.		p-val.
ΔRAT_t	0.365344	***	0.001	0.254244	***	0.003	0.154022	**	0.012
$\Delta USAGE_t$	0.011286		0.449	0.000312	*	0.092	0.000453	***	0.001
$\Delta AMPLI_t$	-0.000025		0.134	-0.003089	***	0.001	-0.000836	***	0.000
ΔCLR_t	-0.000236		0.897	0.000678		0.494	-0.000170	***	0.002
$\Delta CVIOL_t$	0.067634	**	0.039	0.059737	***	0.000	0.049968	***	0.000
$\Delta LIMIT_t$	-0.316655	***	0.000	-0.149881	***	0.000	-0.128109	***	0.000
ΔRAT_{t-12}	0.344095	**	0.050	0.382752	***	0.000	0.135809		0.209
$\Delta USAGE_{t-12}$	-0.006373		0.755	0.000340	**	0.042	0.000260		0.354
$\Delta AMPLI_{t-12}$	-0.000078		0.420	-0.001337	***	0.007	-0.000282	*	0.065
ΔCLR_{t-12}	0.001989		0.572	0.000407		0.238	-0.000072	***	0.008
$\Delta CVIOL_{t-12}$	0.005452		0.884	0.020503	*	0.078	0.037841	***	0.000
$\Delta LIMIT_{t-12}$	-0.144170		0.107	-0.004691		0.691	-0.017457		0.201
Const.	-3.415596	***	0.000	-3.509065	***	0.000	-3.681956	***	0.000
McFadden Adj. R^2	0.102			0.090			0.057		
Obs.	11,551			99,742			699,537		

Table 4

Probit regression results by duration and distance

In this table, the dependent variable is default at time t+1 (DEF_{t+1} equals one for jumps to default and zero otherwise). Explanatory variables are the rating change (ΔRAT) and the changes of the credit line usage (ΔUSAGE), the absolute account amplitude (ΔAMPLI), the credit-to-limit ratio (ΔCLR), the cumulative number of limit violations (ΔCVIOL), the relative change of the credit limit (ΔLIMIT) during the period [t, t-12] and the period [t-12, t-24], and TYPE2 (TYPE3) indicating small firms (individuals). Large firms (TYPE1) are used as reference category and defined as firms with total annual sales above one million euros (calculated as mean cumulative monthly credits x 12). We divide the variables ΔUSAGE and ΔCLR by 100 to scale the estimated coefficients. Duration DUR is differentiated by a median split (short if DUR < 7.9 years) and distance DIST is differentiated by the 90% quantile (small if DIST < 12.8 kilometers). Regressions take into account the clustering of observations at the account level and are based on p-values from Huber-White robust standard errors.

Panel A: Estimation results by duration of the bank relationship (DUR)

	Long duration			Short duration		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.
ΔRAT_t	0.121835	*	0.094	0.256945	***	0.000
$\Delta USAGE_t$	0.000294	***	0.003	0.001797	***	0.002
$\Delta AMPLI_t$	-0.000172	**	0.038	-0.000009		0.847
ΔCLR_t	-0.000162	***	0.003	-0.000423	*	0.075
$\Delta CVIOL_t$	0.053306	***	0.000	0.046406	***	0.000
$\Delta LIMIT_t$	-0.133436	***	0.000	-0.127969	***	0.000
ΔRAT_{t-12}	0.196287	*	0.081	0.259298	***	0.000
$\Delta USAGE_{t-12}$	0.000201		0.140	0.001083	***	0.000
$\Delta AMPLI_{t-12}$	-0.000057		0.503	-0.000111		0.321
ΔCLR_{t-12}	-0.000065	**	0.021	-0.000611	**	0.036
$\Delta CVIOL_{t-12}$	0.030831	***	0.002	0.033720	***	0.001
$\Delta LIMIT_{t-12}$	-0.014090		0.276	-0.015054		0.319
TYPE2	0.168918		0.353	-0.273631	**	0.018
TYPE3	-0.066713		0.711	-0.345299	***	0.001
Const.	-3.698015	***	0.000	-3.162164	***	0.000
McFadden Adj. R^2	0.072			0.079		
Obs.	613,563			197,267		

Panel B: Estimation results by bank-borrower distance (DIST)

	Small distance			Large distance		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.
ΔRAT_t	0.200086	***	0.000	0.304440	*	0.074
$\Delta USAGE_t$	0.000570	***	0.001	0.000871		0.100
$\Delta AMPLI_t$	-0.000037		0.553	-0.000254		0.826
ΔCLR_t	-0.001090	***	0.001	-0.000499		0.191
$\Delta CVIOL_t$	0.050279	***	0.000	0.064804	***	0.000
$\Delta LIMIT_t$	-0.129912	***	0.000	-0.152610	***	0.001
ΔRAT_{t-12}	0.216820	***	0.003	0.331247	***	0.009
$\Delta USAGE_{t-12}$	0.000337	**	0.053	0.000642	***	0.001
$\Delta AMPLI_{t-12}$	-0.000103		0.306	0.000874		0.408
ΔCLR_{t-12}	-0.000107		0.709	-0.000041		0.159
$\Delta CVIOL_{t-12}$	0.032939	***	0.000	0.028827		0.109
$\Delta LIMIT_{t-12}$	-0.015054		0.185	-0.014110		0.410
TYPE2	-0.154007		0.110	-0.709286	***	0.000
TYPE3	-0.301279	***	0.001	-0.680043	***	0.000
Const.	-3.366564	***	0.000	-2.357564	***	0.000
McFadden Adj. R^2	0.079			0.097		
Obs.	718,375			82,592		

Table 5

Account activity and loan pricing

In this table, the dependent variable (SPREAD) is the credit spread on all new commercial loans granted by the bank in 2005. Explanatory variables are the internal credit rating (RAT) and an indicator variable for firms with a checking account at the same bank (CHECK). DUR measures the duration of the bank-firm relationship in years, COLLAT is a dummy variable that indicates secured loans, MATURITY is the logarithm of the maturity of the loan in months, LOANSIZE is the logarithm of the nominal loan amount, and FLOAT indicates floating loan rates. In the full model we also add industry fixed effects (differentiation by ten industry codes, as used by the bank). USAGE_HIGH (VIOL_HIGH) are dummy variables that equal one if the average credit line usage (number of credit limit violations) is above the median of these variables during the year before the loan approval. The cross-sectional regressions take into account the clustering of observations at the borrower level and are based on p-values from Huber-White robust standard errors.

Panel A: Checking accounts and loan spreads

	Model I			Model II		
Dep. Var.: SPREAD	Coeff.		p-val.	Coeff.		p-val.
RAT	0.270182	***	0.008	0.312511	***	0.003
CHECK	0.409949	**	0.033	0.899674	***	0.000
DURATION				0.008518		0.524
COLLAT				-0.431499	**	0.034
MATURITY				-0.264941	**	0.012
LOANSIZE				-0.602559	***	0.000
FLOAT				-0.105774		0.669
Const.	1.641533	***	0.000	9.416698	***	0.000
Industry fixed effects	No			Yes		
Adj. R^2	0.0240			0.2710		
Obs.	643			643		

Panel B: Credit line usage, limit violations, and loan spreads

	Model III			Model IV		
Dep. Var.: SPREAD	Coeff.		p-val.	Coeff.		p-val.
RAT	0.450801	***	0.002	0.320331	**	0.028
USAGE_HIGH	0.447339	*	0.060			
VIOL_HIGH				0.723914	***	0.000
DURATION	-0.008084		0.629	-0.006406		0.674
COLLAT	-0.366744		0.202	-0.381741		0.186
MATURITY	-0.404458	**	0.021	-0.366098	**	0.024
LOANSIZE	-0.885591	***	0.000	-0.956302	***	0.000
FLOAT	-0.025617		0.937	0.097216		0.765
Const.	13.07123	***	0.000	14.00696	***	0.000
Industry fixed effects	Yes			Yes		
Adj. R^2	0.3630			0.3730		
Obs.	330			330		

Table 6

Probit regression results by structure of the bank-borrower relationship

In this table, the dependent variable is default at time t+1 (DEF_{t+1} equals one for jumps to default and zero otherwise). The explanatory variables are the rating change (ΔRAT), the change of the credit line usage (ΔUSAGE), the change of the absolute account amplitude (ΔAMPLI), the change of the credit-to-limit ratio (ΔCLR), the change of the cumulative number of limit violations (ΔCVIOL), and the relative change of the credit limit (ΔLIMIT) during the period [t, t-12]. We divide the variables ΔUSAGE and ΔCLR by 100 to scale the estimated coefficients. Regressions take into account the clustering of observations at the account level and are based on p-values from Huber-White robust standard errors.

Panel A: Results for single and multiple-account bank relationships

	Single account bank relationships			Multiple account bank relationships		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.
ΔRAT_t	0.050171		0.269	0.224470	***	0.000
$\Delta USAGE_t$	0.000284	***	0.008	0.000137	*	0.100
$\Delta AMPLI_t$	-0.000238	***	0.008	-0.000069		0.259
ΔCLR_t	-0.000091	***	0.003	-0.000341	*	0.060
$\Delta CVIOL_t$	0.077009	***	0.000	0.073649	***	0.000
$\Delta LIMIT_t$	-0.086478	***	0.000	-0.079814	***	0.000
Const.	-3.64889	***	0.000	-3.645819	***	0.000
McFadden Adj. R^2	0.067			0.063		
Obs.	772,873			508,354		

Panel B: Results by proxies for the most important account per borrower

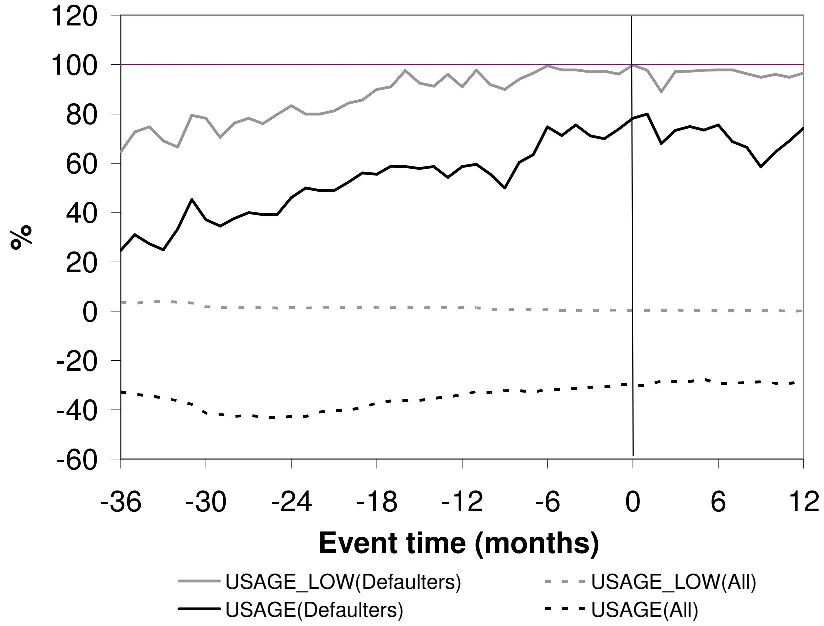
	Most important account (highest mean credit)			Most important account (longest time series)		
Dep. Var.: DEF_{t+1}	Coeff.		p-val.	Coeff.		p-val.
ΔRAT_t	0.095906	***	0.006	0.112403	***	0.002
$\Delta USAGE_t$	0.000283	***	0.010	0.000146	*	0.067
$\Delta AMPLI_t$	-0.000134	***	0.001	-0.000147	***	0.000
ΔCLR_t	-0.000109	***	0.001	-0.000108	***	0.001
$\Delta CVIOL_t$	0.078918	***	0.000	0.077557	***	0.000
$\Delta LIMIT_t$	-0.088597	***	0.000	-0.086922	***	0.000
Const.	-3.67367	***	0.000	-3.664562	***	0.000
McFadden Adj. R^2	0.072			0.070		
Obs.	1,142,540			1,142,717		

Figure 1

Credit line usage and limit violations

In this figure, panel A displays credit line usage (in %) as USAGE_LOW=LOW/LIMIT and USAGE=MID/LIMIT. USAGE_LOW for defaulters (all borrowers) is depicted by a solid (broken) gray line. USAGE for defaulters (all borrowers) is shown by a solid (broken) black line. In Panel B we calculate the cumulative number of credit limit violations CVIOL for each account in calendar time (a solid line for defaulters, a broken line for all borrowers). A violation occurs if the monthly minimum account balance falls short of the credit line, i.e., LOW < LIMIT. The event study is based on 1,009 defaults during the period January 2002 to December 2006. We measure event time in months and defaults occur at event time 0.

Panel A: Credit line usage



Panel B: Cumulative number of limit violations

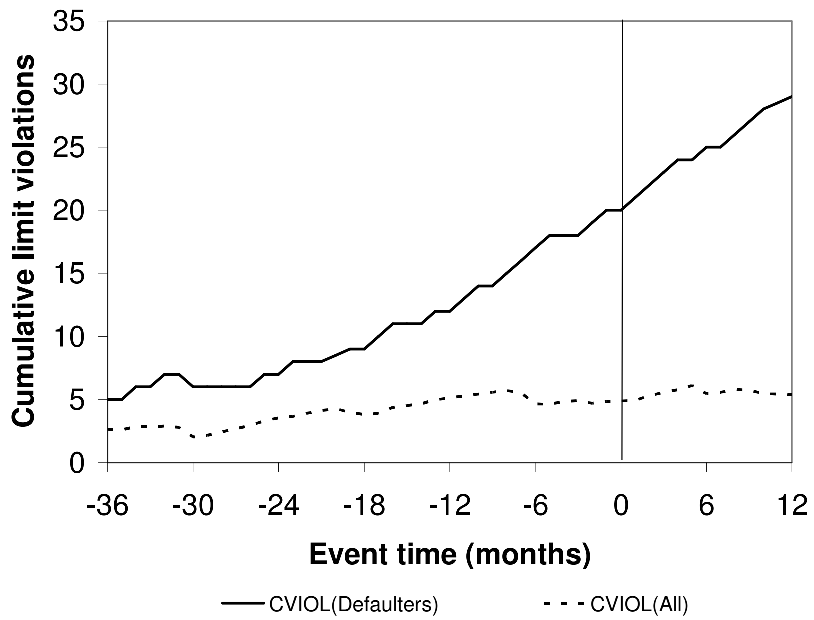


Figure 2

Absolute account amplitude

In this figure, we define the absolute account amplitude (in thousand euros) as AMPLI=HIGH-LOW. The solid (broken) line shows AMPLI for defaulters (all borrowers). The event study is based on 1,009 defaults during the period January 2002 to December 2006. We measure event time in months. Defaults occur at event time 0.

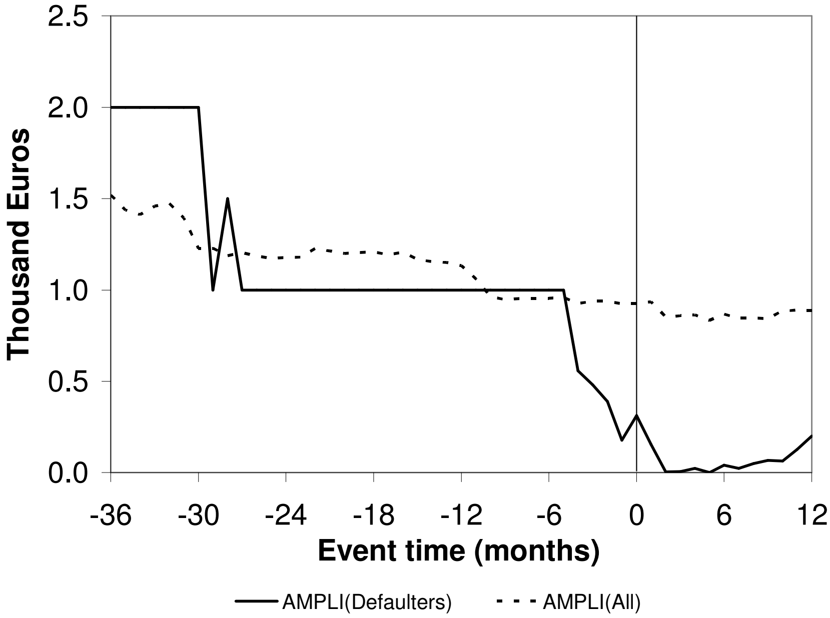


Figure 3

Monthly cumulative credits and debits relative to the credit line

In this figure, we define the credit-to-limit ratio (in %) as CLR=CREDIT/LIMIT x (-100). CLR for defaulters (all borrowers) is depicted by a solid (broken) black line. DLR for defaulters (all borrowers) is shown by a solid (broken) gray line. We define the debit-to-limit ratio (in %) as DLR= DEBIT/LIMIT x (-100). Both CLR and DLR can take only positive values, since we multiply the credit line LIMIT by (-100) and define DEBIT and CREDIT as positive numbers. The event study is based on 1,009 defaults during the period January 2002 to December 2006. Event time is measured in months. Defaults occur at event time 0.

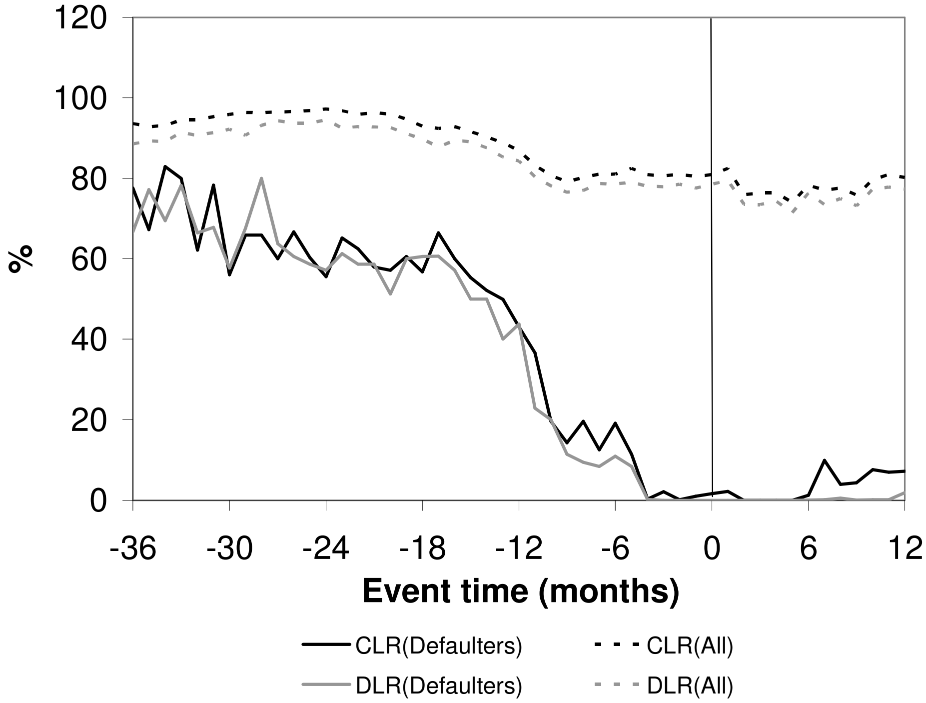
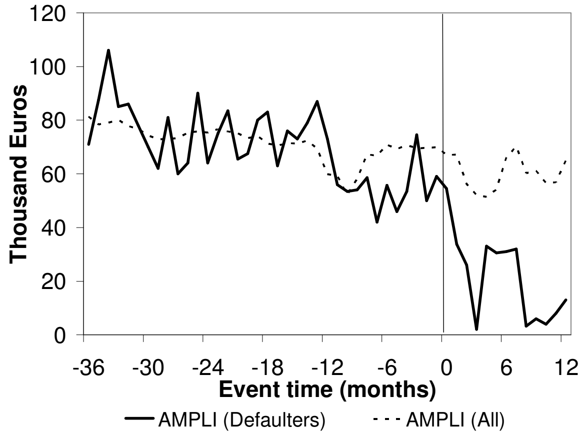


Figure 4

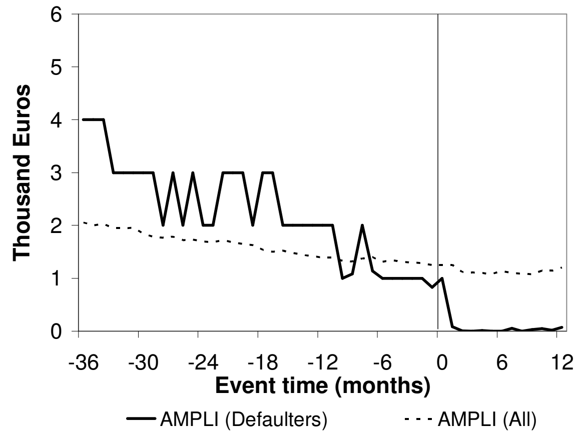
Absolute account amplitude by borrower type

In this figure, we define the absolute account amplitude (in thousand euros) as AMPLI=HIGH–LOW. AMPLI for defaulters (all borrowers) is displayed by a solid (broken) black line. Panel A, B, and C report AMPLI by borrower type (TYPE). We differentiate between large firms (TYPE=1), small firms (TYPE=2), and individuals (TYPE=3). Large firms are defined as firms with total annual sales above one million euros (calculated as mean cumulative monthly credits x 12). The event study is based on 1,009 defaults during the period January 2002 to December 2006. Event time is measured in months. Defaults occur at event time 0.

Panel A: Large firms (TYPE=1)



Panel B: Small firms (TYPE=2)



Panel C: Individuals (TYPE=3)

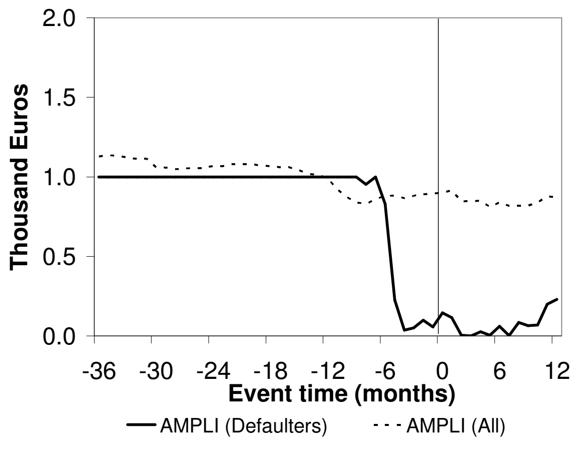


Figure 5

Credit line usage by default type

In this figure, we measure credit line usage (in %) as USAGE=MID/LIMIT. USAGE is depicted by a solid black (gray) line for borrowers who end up with a bankruptcy filing (loan loss provision). USAGE for all firms is shown by a broken black line. The event study is based on 474 defaults (rating 5 without subsequent downgrades) and 535 defaults (directly to rating 6 or indirectly to rating 6 via rating 5) during the period January 2002 to December 2006. Event time is measured in months. Defaults occur at event time 0.