Informational Synergies in Consumer Credit

Martin Thomas Hibbeln [a], Lars Norden [b, *], Piet Usselmann [c], Marc Gürtler [c]

[a] *University of Duisburg-Essen, Lotharstraße 65, 47057 Duisburg, Germany*

[b] *Brazilian School of Public and Business Administration, Getulio Vargas Foundation, Praia de Botafogo 190, 22250-900 Rio de Janeiro, Brazil*

[c] *Braunschweig Institute of Technology, Abt-Jerusalem-Straße 7, 38106 Braunschweig, Germany*

Abstract

The production of private information about borrower risk is a core function of financial intermediation. However, little is known about synergies between different sources of private information. We focus on consumer credit and investigate the existence and magnitude of cross-product informational synergies, using 1.7 million monthly observations from checking accounts and credit card accounts of the same individuals during 2007-2014. We find that activity measures from both accounts contain significant information about default risk beyond credit scores, borrower characteristics, and bank-borrower relationship characteristics. We also find that checking accounts display warning indications earlier and more accurately than credit card accounts. Type I default prediction errors decrease by 33% when checking account information is added to credit card information, but only by 2% when credit card information is added to checking account information. The evidence suggests important informational synergies that are relevant for the supply and the allocation of consumer credit.

Key words: Asymmetric information, credit risk, checking accounts, credit cards, bankruptcy

JEL classification: G20, G21, D12, D14

---

[*] Corresponding author: Lars Norden; Brazilian School of Public and Business Administration, Getulio Vargas Foundation, Praia de Botafogo 190, 22250-900 Rio de Janeiro, Brazil. Phone: +55 21 3799 5544.
E-mail: lars.norden@fgv.br.

The authors are grateful to Tobias Berg, Alejandro Bernales, Régis Blazy, Geraldo Cerqueiro, Oliver Entrop, Moritz Lukas, Steffen Meyer, Tomasz Piskorski, Lei Zhao and participants at the European Economic Association Conference 2015, the Paris December 2015 Finance Meeting, the Southern Finance Association Conference 2015, the Banking Workshop 2015 in Münster, the Midwest Finance Association Conference 2015 (Outstanding Paper Award in Financial Institutions), the German Finance Association Conference 2014, and seminars at the Catholic University of Portugal, the University of Chile, Aarhus University, Goethe University Frankfurt, the University of Münster, the University of Hamburg, the University of Bremen, the University of Duisburg-Essen, the University of Giessen, the University of Innsbruck, the University of Leipzig, and the Braunschweig Institute of Technology.

Informational Synergies in Consumer Credit

Abstract

The production of private information about borrower risk is a core function of financial intermediation. However, little is known about synergies between different sources of private information. We focus on consumer credit and investigate the existence and magnitude of cross-product informational synergies, using 1.7 million monthly observations from checking accounts and credit card accounts of the same individuals during 2007-2014. We find that activity measures from both accounts contain significant information about default risk beyond credit scores, borrower characteristics, and bank-borrower relationship characteristics. We also find that checking accounts display warning indications earlier and more accurately than credit card accounts. Type I default prediction errors decrease by 33% when checking account information is added to credit card information, but only by 2% when credit card information is added to checking account information. The evidence suggests important informational synergies that are relevant for the supply and the allocation of consumer credit.

Key words: Asymmetric information, credit risk, checking accounts, credit cards, bankruptcy

JEL classification: G20, G21, D12, D14

1. Introduction

Financial intermediaries are special because they produce private information to screen and monitor borrower risk. The theory of financial intermediation considers private information production, which results in a reduction of asymmetric information, as the raison d'être for banks (Stiglitz and Weiss, 1981; Diamond, 1984; Ramakrishnan and Thakor, 1984; Boot, 2000). In empirical research, however, private information production of financial intermediaries has largely remained a black box, mainly due to the lack of data (Campbell, 2006). Some studies show the importance of banks' private soft information for lending (Berger and Udell, 1995), whereas little is known about banks' private hard information. Almost entirely missing is systematic evidence on banks' different sources of private information, whether there are synergies between these sources, and if yes, how large they are. Synergies are defined as "*a mutually advantageous conjunction or compatibility of distinct business participants or elements (as resources or efforts)*.[1]" In the context of financial intermediation, informational synergies can arise if stand-alone information is useful to assess borrower risk and if it is pairwise complementary with one another, that is, one source fills gaps that may exist in other sources and vice versa. Stated differently, private information from different sources is similar to pieces of a puzzle that, only when it is fully assembled, shows the whole picture, as in the saying about synergies "*the whole is greater than the sum of its parts.*"

In this paper, we investigate the existence and magnitude of potential informational synergies in consumer credit. Understanding the dynamic interplay between private information from different sources is important because it might affect the supply and the allocation of consumer credit in the economy. Specifically, we examine the link between consumer defaults and private information from checking accounts with lines of credit and credit card accounts. For these credit

[1] Merriam-Webster's Dictionary of English Usage.

products, it is critical for lenders to produce dynamic private information about borrowers because, unlike mortgages, they are unsecured and their risk exposure is time-varying.

The market for consumer credit is characterized by a large number of transactions, relatively small volumes and standardized products. The main credit products are checking accounts with lines of credit, credit cards, consumer loans and mortgages. Lenders rely on standardized public and private hard information and this reliance is even stronger when consumer debt is securitized (Rajan, Seru and Vig, 2015). There is evidence that the rise in credit supply, consumer bankruptcies and increased dispersion of interest rates in the U.S. during the 1980s and 1990s are related to innovations in information production that have facilitated the lending to riskier borrowers (Livshits, Mac Gee and Tertilt, 2016). Consumers default because they may face either negative income shocks or positive expenditure shocks that are usually due to unemployment, credit misuse, martial disruptions, health-care issues or lawsuits (Gross and Souleles, 2002; Chatterjee et al., 2007).

A key feature of our study is that these shocks are reflected by consumer credit products. Checking accounts provide the lender with dynamic information about consumer income and expenditures, whereas credit card accounts provide dynamic information about consumer expenditures. To monitor the default risk of their borrowers, lenders want to obtain early warning indications about shocks and want to know whether shocks are temporary or permanent. Activity measures such as overdrafts, cash inflow-to-limit ratios, or the account balance amplitude reflect private hard information that gives the lender a real-time window into the consumer's cash inflows and outflows. Our setting enables us to analyze the account activity from different credit products of the same consumer, controlling for credit scores, borrower characteristics, and bank-borrower relationship characteristics.

We contribute to the literature on financial intermediation and consumer credit in following three ways. First, there are only a few studies that examine the link between account activity and information production. These studies consider one source of private information and consequently neither investigate the existence of informational synergies nor their magnitude. These studies focus on lending to firms and analyze information from checking accounts (Mester, Nakamura, and Renault, 2007; Jimenez, Lopez and Saurina, 2009; Norden and Weber, 2010). Norden and Weber (2010) examine whether activity measures from checking accounts provide useful information about borrower defaults. Using data on large firms, small firms and individuals from the period 2002-2006, they find that account activity is useful for monitoring small firms and individuals, but not for large firms. Our study makes the next step. Norden and Weber (2010) consider checking accounts as single source of information, while we analyze defaults and account activity, respectively, from two different sources (checking accounts and credit card accounts of the same customer), focusing on consumer credit. This setup makes it possible for us to investigate the existence and magnitude of cross-product informational synergies.

Furthermore, there are studies on information production in consumer credit, but none of these studies examines cross-product informational synergies (Karlan and Zinman 2009; Agarwal et al., 2009; Allen, Damar and Martinez-Miera 2016). Agarwal et al. (2009) document that credit card customers with close relationships to their bank exhibit lower probabilities of default and display higher credit line usage than non-relationship customers. The analysis is based on indicators about prior characteristics of the bank-consumer relationship, such as the existence and duration of other credit products. For comparison, we consider dynamic private information that the lender collects simultaneously from different sources about the same customer. Allen, Damar and Martinez-Miera (2016) examine bank mergers and find that consumer default rates increase in markets in which the merging banks' branch networks overlapped pre-merger. The result indicates post-merger disruptions in bank-consumer relationships and temporarily degradation of banks' information production. The study does not examine cross-product informational synergies, but the results are consistent with the view that such synergies exist and that they affect outcomes in credit markets.

Second, evidence suggests a link between individual consumption patterns and consumer default risk (e.g., Vissing-Jorgensen, 2012; Stango and Zinman, 2016). Vissing-Jorgensen (2012) documents a link between consumer choice and consumer default risk. She examines credit-financed consumer purchases at a Mexican retail chain to investigate whether the type of purchased products provides information about default risk. The main finding is that credit risk is particularly high for consumers who buy abnormally large fractions of luxury goods relative to their income. Stango and Zinman (2016) show for the U.S. that the dispersion of credit card rates, which different financial institutions offer to the same individuals, is related to their shopping intensity, controlling for the individuals' credit risk. Our study is complementary since we investigate potential synergies between different sources of financial information resulting from consumer choices.

Third, informational synergies in consumer credit may be relevant for the cross-selling of financial products. Cross-selling can lower marketing costs and product prices, increase customers' switching costs, and lenders can learn more about their customers' risk preferences and consumption behavior (Kamakura et al., 1991; Akçura and Srinivasan, 2005; Li, Sun and Wilcox, 2005; Li, Sun and Montgomery, 2011). Hence, our study is not only relevant for credit risk management but also for customer relationship management.

Our analysis is based on a large and unique panel dataset comprising more than 1.7 million monthly observations from checking accounts and credit card accounts, respectively, of the same individuals. The data come from a large privately owned bank and span the period from 2007 to 2014.

We obtain two main results. First, activity from checking accounts and credit card accounts contain significant information about default risk beyond credit scores, borrower characteristics, and bank-borrower relationship characteristics. Lenders can improve default predictions when they combine the dynamic private information from both sources. Second, the activity measures from checking accounts are more useful for default prediction than activity measures from credit card accounts. Checking accounts display warning indications earlier and more accurately than credit card accounts. Interestingly, type I default prediction errors decrease by 33% when checking account information is added to credit card information, and only by 2% when credit card information is added to checking account information. The main explanation is that most consumers default because of negative income shocks that are reflected by checking accounts, but not by credit card accounts. For short prediction horizons, though, credit card information is informative, too. We conduct several additional empirical checks and show that our previous results are robust and not the product of particular choices of samples, methods or model specifications.

To the best of our knowledge, this is the first study that provides comprehensive on cross-product informational synergies in consumer credit. These synergies help lenders to reduce informational asymmetries and improve the allocative efficiency in consumer credit markets. Our findings suggest that, next to the well-known "skin in the game effect", consumer lending and risk assessment should be performed by the same institution because lending generates private information in real time that a non-lender cannot produce. We note that lenders can take advantage of private information only if the information production is not distorted by incentive problems due to compensation schemes, career concerns and credit reputation concerns (e.g., Hertzberg, Liberti and Paravisini, 2010; Berg, Puri and Rocholl, 2014; Liberman, 2016). A risk assessment based on different sources of private information is less likely to be biased by strategic behavior of the borrower or loan officer and is thereby more reliable than a risk assessment based on a single source of private information.

The remainder of this paper is organized as follows. In Section 2, we describe the institutional background, the data and the research method. In Section 3, we report the results of the analysis of informational synergies. In Section 4, we summarize the findings from further empirical checks and robustness tests. Section 5 concludes.

2. Institutional background, data and methodology

2.1. Institutional background

We consider data on consumer checking accounts and credit card accounts from a large privately owned European financial institution.[2] Checking accounts are used for receipts (e.g., salary, rental income, etc.), and expenses (e.g., rent, subscriptions, insurance, consumption expenses). The balance of a checking account can be positive or negative because by default a credit line is attached to it. If the balance is positive, consumers (may) receive interest from the bank, and if the balance is negative, consumers have to pay interest to the bank for using the credit line. The credit line is pre-authorized up to a specific limit and can be redeemed at any point in time. Consumers use a checking account to withdraw cash at ATMs, to make payments with debit cards, to use electronic direct debits or to pay bills with electronic wire transfers. We can exclude that checking accounts are used as clearing accounts for brokerage accounts because no brokerage services are offered by the bank.

Furthermore, we consider credit card accounts (Visa credit cards). The balance of a credit card account can be positive or negative, too. In the case of a positive balance, consumers receive an interest rate. If the balance is negative, it depends on the type of credit card account whether consumers have to pay interest to the bank. Our dataset comprises two types of credit card accounts, which differ in their redemption schemes. The majority of credit card accounts is with full repayment of the credit card bill every month. This monthly bill is charged on the individual's checking account on a fixed day per month and there is no interest to be paid. The payment day is one of the first days in a month for almost all customers. In addition, there are accounts that allow the consumer to stretch out the redemption over time and make only minimum repayments per

[2] To ensure confidentiality, the customer name, account number and customer number have been anonymized by the financial institution that provided us with the data.

month. These monthly minimum repayments correspond to 5% of the outstanding credit card debt and at least 50 euro. Customers pay interest rates of more than 10% p.a. on the outstanding credit card debt.

Checking accounts and credit card accounts are unsecured credit. All accounts have an initial line of credit of 1,000 € and consumers may ask for an increase of the limit later on. The limits of a customer's checking account and credit card account are set independently of each other. A negative account balance has to stay within the account limit, but we also observe overdrafts. Accounts can be overdrawn by using the checking account or the credit card offline, manual approvals by loan officers, or interest debit charged by the bank. Cross-product information is not used by the bank in an automated manner and not part of the bank internal rating system. Instead, the financial institution makes discretionary use of cross-production information if irregularities occur, and overdrafts are manually approved or rejected by a loan officer on a daily basis. This is consistent with Agarwal et al. (2009), who document that the information used to determine the internal credit score is traditionally limited to activity measures from one specific account.

The presence of the credit products we study here is widespread. For example, in 2014 more than 99% of the payment instruments in countries that are member of the Committee on Payment and Settlement Systems (CPSS) were credit transfers, direct debits, checks, debit cards, and credit cards (Bank for International Settlements, 2015). Hence, virtually all payments by individuals are made through either checking accounts or credit card accounts.

We further use internal bank information on borrower defaults. According to the default definition, complying with Basel III and EU Regulation No. 575, a borrower default can occur if the borrower is "90 days past due" or if the bank expects that the individual will not pay back all of his obligations. The latter occurs if the bank observes a limit violation, a bankruptcy, or receives negative information from a credit bureau. The default status refers to the account and not to the account owner. Thus, a default of one account does not necessarily lead to a default of other accounts of the same individual. This is standard practice in retail banking and in compliance with the Basel III regulatory rules for retail exposures.

2.2. Data

We base our analysis on dynamic information from checking accounts and credit card accounts. The raw dataset comprises 5,958,534 account-months observations of individuals who have credit cards and checking accounts at the same bank. The sample period is from December 2007 to January 2014.

We apply the following filter rules. First, in our baseline analysis we consider one credit card and one checking account per customer. Most of the customers with both products have exactly one checking account and one credit card (88%). For the 12% of customers where we observe more than one checking account or credit card, we focus on the most important account, so we drop 237,880 observations.[3] Second, we further drop observations with missing data so that we cannot compute the account activity measures (1,158,501 observations). Third, we only consider account data when we have observations for at least one year,[4] which reduces the dataset by 967,795 observations. Fourth, we consider only observations up to the first default, so that the observed account activity is not influenced by a previous default of the same account, reducing the dataset by 33,594 observations. We winsorize all the variables at the 1% and 99% quantiles. The final sample consists of 3,560,764 account-month observations from 46,925 customers. The characteristics of this dataset make it possible for us to investigate whether there are informational

[3] We assume that the account with the highest mean cash inflow per month is the most important one because it is likely that the customer's salary is paid to this account. For credit card accounts, inflows normally have to equal the outflows; hence, the most important credit card can also be identified using the highest mean inflow. The number of customers where we have to identify the most important account is rather low; hence, the choice of the procedure does not affect our analysis. In a robustness check, we repeat all subsequent analyses with customers who have exactly one checking account and one credit card. Our results show that the aggregation rule does not affect our results.

[4] If the observation period is shorter than the forecasting horizon, the credit risk can be underestimated (Gürtler and Hibbeln 2013). Thus, we do not consider observations of the last 12 months of our sample period, because we only observe defaults until the end of the sample period, but not within the full forecasting horizon of one year.

synergies between the account activities from different credit products of the same individual.

(Insert Table 1 about here)

Panel A of Table 1 reports the number of observations and defaults. There are 1,639 and 2,101 defaults at checking accounts and credit cards, respectively. Customers are more frequently in default on credit cards than on checking accounts. This result is not due to differences between customers (e.g., a lower average age or income for credit card accounts) because the checking accounts and credit cards come from the *same* customers. Moreover, individuals can default on multiple accounts or on one account. Possible reasons for a joint default are bankruptcy or negative credit information from a credit bureau. An individual default on a checking account can occur if the monthly expenses sufficiently exceed monthly receipts. Similarly, a customer can default on a credit card account. A high credit card bill could lead to an overdraft of the checking account, so that the bank decides to bounce (not authorize) the debit of the credit card bill. Such a bounced debit does not cause a default on the checking account, but it might do on the credit card account.

Panel B of Table 1 shows that about 50% of the monthly default observations (13,887 observations) are joint defaults of both accounts of the same person, whereas the remaining 50% occur either on the checking account (5,857 observations) or on the credit card account (7,565). These numbers confirm that both cases occur frequently and are therefore relevant for our study.

Panel C of Table 1 reports summary statistics of the main variables. We define the variable *Rating* as the logarithm of the probability of default according to the internal rating system of the bank. *Rating* does not contain any cross-product information and is account specific (the correlation between the corresponding credit scores per customer is 0.337). The mean *Rating* for checking accounts is significantly better than for credit card accounts, which is consistent with a lower number of defaults for checking accounts.

Net Inflow/Lim is the difference between monthly cash inflows and outflows as a percentage of the external limit. To measure the account variation, we use *Amplitude* as the difference between the maximum and minimum exposure in each month as a percentage of *Limit*. We observe higher amplitude for checking accounts than for credit card accounts as checking accounts are used for income and expenses. The mean and median of the external limit in euros indicates that most customers choose a higher limit than the initial limit; moreover, most customers choose a similar limit for both products (the correlation between the limit of both accounts is 0.767). *Bounced* is the average number of bounced (not authorized) debits in the preceding 12 months (per month). In most cases, this number is zero for checking accounts, whereas for credit cards bounced debits can be observed more frequently. We define *Days Usage* as the percentage of days with a negative balance in the preceding 12 months. This number is significantly higher for credit cards because for this product we usually observe no credit balance. We see a relatively active use of credit cards since 55% of the days in the preceding 12 months have a negative balance. Similar to *Days Usage*, we define *Days Overdrafts* as the percentage of days with overdrafts in the preceding 12 months. Only a small percentage of checking accounts exhibits overdrafts.

Moreover, we have information on bank-customer relationship characteristics. *Duration* indicates the length of the relationship in months measured separately for each account. For checking accounts, *Duration* is on average 46 months (about four years), which is slightly longer than for credit card accounts. Seven percent of the credit card observations belong to accounts without full monthly repayment (*Full Payment* = 0). Around 38% of observations with defaults belong to this product type, which indicates that riskier customers self-select into this redemption mode. We therefore use the redemption mode (*Full Payment*) as control variable. We also use information about the customer's age, gender, job, marital status, number of children, nationality, online versus offline banking, and academic degrees (not reported in Table 1).

2.3. Methodology

We analyze informational synergies on the probability of default (*PD*) of checking accounts and credit card accounts. To estimate the *PD* of one of these accounts, we first consider the account activity from the same account (where default occurs), then the "cross-product activity" from the other account, and finally information from both accounts at the same time. We define *PD* as the probability of default of an account within a one year time horizon. Thus, for a given month t, we define a default variable $Def_{i,t+12}$ that equals one if a jump to default is observed at $t+\tau$ with $\tau \in \{1, 2, \ldots, 12 \text{ months}\}$ for account i, and zero otherwise. Given the explanatory variables $Z_{i,t}$ that can be observed in month t, the estimated *PD* is: $PD_{i,t} := \mathrm{P}(Def_{i,t+12} = 1 \mid Z_{i,t}) = f\left(Z_{i,t}\right)$.

The explanatory variables contain account activity variables, which can include information from the particular account only, or we consider cross-product information as well. Furthermore, $Z_{i,t}$ includes bank-borrower relationship and borrower-specific variables. In addition, we use the internal credit scores to account for variables that can be observed by the bank but are not available to us (e.g., information from a credit bureau). We integrate the credit score of the internal rating ($=PD_{i,t}^{Bank}$) as $Rating_{i,t}$ = log($PD_{i,t}^{Bank}$). We perform pooled probit regressions and cluster the standard errors at the customer level. In our baseline analysis, the explanatory variables enter the model linearly. In robustness tests, we show that that the choice of the functional form of the explanatory variables is not critical for our main results.

3. Results

3.1. Account activity and cross-product information

We first conduct a univariate event study of cash inflows and cash outflows of checking accounts for defaulted and non-defaulted customers. We show the development of these variables from 24 months before default to 12 months after default; the default time is $\tau = 0$ (event time). We calculate the median of the explanatory variables for defaulted customers at a monthly frequency for a time window of 37 months (τ-24, τ-23, ..., τ, ..., τ+12). For non-defaulted customers, we calculate medians, weighted by the number of defaults, to allow a direct comparison between the explanatory variables for defaulted and non-defaulted customers: We first calculate the median account variable for each month (t = 12/2007 to 1/2014) for all non-defaulted customers that will not default in the next 12 months ($Def_{i,t+12} = 0$). Then, we determine the number of defaults in each month and calculate on this basis how often each month must be considered at each event time (τ-24, τ-23, ..., τ, ..., τ+12), and compute the median of the monthly medians for each event time. Thus, we assume a hypothetical default event for the non-defaulted customers.

In Figure 1, we show the customers' payment behavior regarding cash inflows and outflows separately for defaulted and non-defaulted checking accounts. On the horizontal axis, the time relative to the default event at τ is displayed in months, and on the vertical axis the median *Cash Inflow* and *Cash Outflow* are shown in euros. If the gap between cash inflows and outflows - the *Net Inflow* - is negative, the balance of the account decreases. Such a decrease will eventually lead to overdrafts and ultimately to a default of the account. These inflows and outflows, as well as the *Net Inflow* are expected to be more informative for checking accounts than for credit card accounts because for checking accounts the customer can choose the magnitude of in- and outflows largely independently. For credit cards, typically only the outflows are influenced by the customer, while the inflows often correspond mechanically to the sum of the outflows because of monthly clearing.

(Insert Figure 1 about here)

Figure 1 shows that the inflows and outflows of non-defaulted checking accounts are similar and rather stable, implying that *Net Inflow* is close to zero. For customers who default, we observe that inflows and outflows start to decrease about 8 months before default, which is similar to the finding of Norden and Weber (2010) who use different data from a different time period. We find that inflows decrease earlier and faster than outflows. This inflow decrease starts about six months before default and continues until default. Outflows, on the other hand, only slightly decrease, but around three months before default, they decrease rapidly. The latter is mainly a consequence of the customers approaching the limit. The findings indicate that for most customers, the reason for default on a checking account are decreasing cash inflows (e.g., unemployment), but not increasing cash outflows (e.g., higher consumption or unexpected expenses). The evidence suggests that a default is rather a consequence of a reduced income and not to increased expenses. Hence, the *Net Inflow* is negative for defaulted accounts and increases in the months before default; however, at the time of default, the inflows and outflows are almost identical; after default, inflows are even greater than outflows, which is a consequence of customers being constrained by a limit reduction at the time of default.

We further consider limit violations as a predictor for customer defaults. However, unlike earlier studies, we use dynamic private information from different sources at a daily frequency. We observe the number of days with a positive usage of the credit line and with overdrafts in each month. These two variables are informative for predicting defaults and hint at informational synergies across different credit products. In Figure 2, we present the median number of days with positive credit line usage (Panels A1 and B1), as well as the median number of days with overdrafts (Panels A2 and B2) for defaulted and non-defaulted accounts. In Panel A, we present these variables for defaults of checking accounts, whereas Panel B refers to defaults of credit cards.[5]

---

[5] In each panel, we plot four lines. In Panel A1, for example, we present the days with usage of checking accounts that defaulted (checking account default) and days with usage of checking accounts that did not default (checking account non default). To visualize the informational synergies we also plot the information of the corresponding credit card accounts: We present the days with usage of the credit card account when the borrowers default on their checking accounts (credit card default) and days with usage of the credit card account when the borrowers do not default on their checking account (credit card non default).

(Insert Figure 2 about here)

Panel A1 of Figure 2 shows for checking accounts that the number of days with positive credit line usage is higher for accounts with subsequent default. This difference is visible 24 months prior to default, but the difference increases as default approaches. We also show the number of days with positive credit line usage of *credit cards* prior to a default of the *checking account*. This number slightly increases before default too, meaning that cross-product information is informative. Similarly, Panel A2 indicates that checking account overdrafts strongly increase in the three months prior to default, peaking at the default event. The number of days with overdrafts is much higher for defaulted checking accounts than for non-defaulted accounts, which confirms that overdrafts provide useful early warning indications. Interestingly, the number of days with overdrafts on credit card accounts is also substantially increasing in the period prior to default of checking accounts. Hence, account activity from other credit products is informative for default prediction.

Panel B1 and B2 of Figure 2 show the number of days with positive credit line usage and overdrafts regarding the default events of credit card accounts. The findings for Panel B1 are similar to the defaults of checking accounts (Panel A1). The median number of days with positive usage is lower for defaulted credit card accounts than for non-defaulted accounts. Panel B2 indicates that overdrafts on credit card accounts strongly increase four months prior to default of the credit card account with a peak at default, while increasing overdrafts on checking accounts of the same customers can be observed substantially earlier (Panel A2), which confirms that cross-product information is informative.

We continue with the analysis of customers' payment activity before a default event. Specifically, we examine at time *t* whether a default event will occur in the next year, that is in

$\{t+1, t+2, \ldots, t+12\}$, to obtain information regarding the probability of default within a time horizon of one year. We analyze additional account activity variables that may influence the probability of default (next to *Net Inflow*, *Days Usage*, and *Days Overdrafts*)[6] and report the average differences between defaults and non-defaults in Table 2. Columns (1) and (2) refer to checking account defaults, whereas columns (3) and (4) refer to credit card account defaults. The results without cross-product information are shown in columns (1) and (3); the results for cross-product information are shown in columns (2) and (4).

(Insert Table 2 about here)

Table 2 shows that checking accounts and credit card accounts that defaulted have significantly worse ratings. We find different values for *Rating* depending on the information source due to different weights of account and customer characteristics in the scoring function. Moreover, defaulted checking accounts have lower *Net Inflows*, *Amplitude*, and *Limit*, along with a higher number of bounced credits (*Bounced*), a more frequent usage of credit line (*Days Usage*), more frequent *Days Overdrafts*, and a shorter relationship (*Duration*). As an additional variable to evaluate the bank-customer relationship, we use the dummy variable *Prev. Def.* to measure whether a previous default occurred on a different account of the same customer. For defaulted checking accounts, the percentage of customers with a previous default on a credit card account is significantly higher, which shows that previous defaults on other accounts can be used as early warning signals. The overall findings for credit card accounts are similar except for *Amplitude*, where *Amplitude* is higher for defaulted accounts. This is plausible because high activity on a credit

[6] *Days Usage* and *Days Overdrafts* are the percentage of days in the preceding 12 months, whereas we present in Figure 2 the median number of days per month with a negative balance or days with overdrafts. In Section 4.4 we show that our results are robust when the explanatory variables enter our model in different functional forms.

card is mainly due to consumption, whereas for checking account this could also be due to a high income. There is a significantly higher percentage of customers without full monthly payments for defaulted accounts than for non-defaulted accounts. The cross-product analysis results in column (2) shows that for defaulting versus non-defaulting checking accounts not only the checking account activity but also on credit card account activity is significantly different. The same holds for defaults of credit card accounts, see column (4). This suggests that cross-product account activity can be a useful source of information.

3.2. Monitoring with cross-product information

We now investigate whether informational synergies stemming from customers' payment activity can be used to monitor customers. We estimate multivariate probit regression models for the probability of default, in which we consider customers' payment activity from different accounts.[7] Table 3 reports the results. Columns (1)-(4) show the results for checking account defaults and columns (5)-(8) for credit card defaults. The upper half of the table refers to the checking account variables and the lower half to the credit card variables. We first discuss the defaults on checking accounts. In the probit regression (1), we use *Rating* as independent variable to predict defaults in the next 12 months, showing that the rating system of the bank is effective. Column (2) shows that all the coefficients are statistically significant and the signs of all the variables are identical to the univariate findings, which confirms that account activity variables are useful early warning indicators for default risk. The main findings remain stable if we additionally control for *Rating* in column (3). The adjusted R^2 slightly increases when the internal rating of the bank is added, which indicates that there is additional information contained in the rating system. Column (4) reports the

[7] We also use logistic and rare events logistic regression models (King and Zeng 2001). The choice of the model does not affect our results. Moreover, we perform all subsequent analyses with absolute account activity measures instead of using account activity ratios. We find that all main results are robust.

results for cross-product information. The influence of rating, account activity, and bank-customer relationship variables of credit cards on checking account defaults mostly points in the same direction as the impact of the corresponding checking account variables, which means that the information is widely consistent. We find that most of these variables are highly significant. Moreover, the explanatory power is higher if cross-product information is considered, and a likelihood ratio (LR) test confirms that cross-product information is jointly statistically significant ($p = 0.000$).[8]

(Insert Table 3 about here)

Columns (5)-(8) of Table 3 report the results of credit card defaults. The results are similar to checking accounts except for *Amplitude*, as a larger amplitude on checking accounts is a positive signal but on credit cards it is rather a negative signal, confirming the above univariate results. As discussed above, a likely reason is that high activity on a credit card is mainly due to consumption, whereas for checking account this could also be due to a high income. Furthermore, regarding cross-product information on credit card defaults, we confirm that the information is consistent (column (8)) and that cross-product information is jointly significant ($p = 0.000$). The increase in the adjusted R^2 from column (7) to column (8) shows that the prediction of credit card defaults can be substantially improved if account activity from checking accounts is added to the model.[9]

We find evidence for informational synergies that lenders can use for monitoring consumer credit risk. The additional cross-product information helps to improve the prediction of defaults

[8] The Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) speak in favor of the model with cross-product information, too.

[9] Other criteria of model selection such as AIC and BIC confirm this finding.

beyond traditional credit scores, customer characteristics, and bank-customer relationship characteristics.

3.3. Screening with cross-product information

We further investigate whether informational synergies stemming from customers' payment activity can be used to screen potential customers. When an individual opens a new bank account, the bank can gather information about their characteristics but it does not have information about his or her past account activity. However, if the individual has already a different account at the bank, the bank can observe the past activity on that account and use this source of information for screening. In our data set, we observe that only a small share of customers opens both accounts at the same time. Hence, there are many potential customers with one existing account, for whom this additional source of information could be used for screening. In Table 4, we present the results when we use only the account activity variables from credit card accounts to predict checking account defaults (and vice versa). This analysis is based on clean cross-product information that can be used for screening.

(Insert Table 4 about here)

We find that cross-product information is very useful for screening; checking account activity is even more predictive of credit card defaults than credit cards variables (adj. R^2 of 0.25 vs. 0.20; see Table 3 (7)).

3.4. Consistency of information

How much does account activity information from different credit products overlap? How consistent is this information? For answering these questions, we compare the *PD* estimates when these are based on account activity of only checking accounts versus only credit card accounts. First, we examine the *PD* estimates for checking accounts. We find that the correlation between the estimates derived from the account activity of checking accounts versus credit card accounts is 56%. Similarly, the correlation for the corresponding *PD* estimates for credit card accounts is 54%. If we analyze the pairwise correlations between individual account activity measures, we find that the correlations are at maximum 35% for all variables except for *Days Overdraft* (0.483) and *Limit* (0.777). Thus, the account activity of both credit products exhibits a positive but imperfect correlation, suggesting that there is useful non-redundant information.

3.5. Magnitude of the informational synergies

In the analysis above, we have documented synergies between information about customers' payment activity from checking and credit card accounts. We now examine the magnitude of these synergies. How much can financial institutions gain if they exploit cross-product account activity information? What are the marginal benefits of using additional sources of information?

Quantifying the synergies in form of a monetary equivalent is challenging because we would have to make various (crude) assumptions about the size and composition of the portfolio, the types of loans and loan terms, the account activity and control variables used in the model, the empirical default rate, the empirical loss given default, and others. We therefore refrain from following this route. Instead, we assess the economic and statistical significance by computing standard measures of prediction accuracy and goodness of fit to assess the economic and statistical significance of the cross-product informational synergies (e.g., Grunert, Norden and Weber, 2005). These measures indicate the economic significance in the sense that we can measure how many false credit decisions a lender can avoid when exploiting cross-product informational synergies.

Table 5 reports the accuracy of *PD* prediction results, which are based on the account activity of checking accounts (column 1), credit card accounts (column 2), and of both accounts (column

3). For default prediction accuracy, we report the adjusted McFadden R^2, the value of the area under the receiver operating characteristic (ROC), type I errors (predicting no default for consumers that do default) and type II errors (predicting default for consumers that do not default). We calculate the type I and II errors based on binary predictions using the empirical default rate in the respective sample as cut-off point. A value of one for the R^2 or ROC indicates a perfect prediction. For a random prediction these statistics are R^2=0 and ROC=0.5.

(Insert Table 5 about here)

We find evidence for sizeable informational synergies in both accounts. First, lenders can significantly improve their decision making when they consider information from both accounts rather than information from one account. Second, the improvement in the adjusted McFadden R^2, ROC and type I and type II errors is particularly strong when we compare information from credit card accounts (column 2) with information from both accounts (column 3). For defaults of credit card accounts, the adjusted McFadden R^2 increases from 0.204 to 0.278, the ROC from 0.845 to 0.892, the type I error decreases from 23.76% to 15.89%, and the type II error decreases from 22.24% to 18.48%. Strikingly, when lenders additionally consider checking account information they can reduce the type I error by around eight percentage points.[10] The absolute number of the type I error decreases from 5,096 to 3,408 (minus 1,688), corresponding to a reduction by 33%.[11] We explain this strong finding with the fact that changes in consumer income (e.g., due to unemployment) are identifiable in checking accounts, but not in credit card accounts. As shown

[10] We note that the type I error is significantly more costly than the type II error. The loss of the credit granted to a consumer who will subsequently default is higher than the loss of the interest income of a rejected loan to a consumer who will subsequently not default.

[11] The effect is even stronger for predicting the defaults of checking accounts: lenders can lower the type I error by around ten percentage points - corresponding to a 42% reduction of the type I error - when they consider checking account activity in addition to credit card account activity.

earlier, warning indications about decreases of cash inflows are critical for the prediction of defaults. Third, there is an improvement of all measures when we compare information from checking accounts (column 1) with information from both accounts (column 3) but the benefits are moderate. This finding shows again that checking accounts are more informative about changes in borrower quality than credit card accounts.

We focused on a time horizon of one year so far. We now analyze whether the informational content of cross-product account activity varies for different time horizons. Figure 3 shows the accuracy of PD estimates for different information sources and time horizons from 1 up to 12 quarters. As expected, the accuracy decreases for longer time horizons for each information source. Panel A shows that default prediction for checking accounts can slightly be improved when account activity information of credit cards is added, and the extent of the informational advantage is almost irrespective of the considered time horizon. Panel B shows the corresponding figure for predicting credit card defaults. Interestingly, the information content from the different accounts depends heavily on the considered time horizon. For short time horizons, account activity from checking accounts and credit cards is similarly informative, and a combination of both information sources is highly beneficial. For long time horizons of up to 3 years, however, checking account activity – which means pure cross-product information here – is the superior source of information. Our interpretation is that checking accounts usually provide more accurate early-warning indications about the consumer credit risk and possible abnormal patterns regarding, for example, the personal income; but if the account activity of *both* accounts shows distinct abnormal patterns at the same time, this is a strong indicator for a systematic behavior leading to a default of the credit card account.

(Insert Figure 3 about here)

In addition to informational synergies regarding the probability of default, we examine variables that indicate consumer credit risk prior to default. If we consider the account activity prior to default, we first observe negative net inflows (caused by a shortfall in cash inflows), then increased credit line usage and occasional overdrafts, and finally the default. Potential changes in individuals' account activity can be temporary and unsystematic, but they become increasingly systematic when we get closer to default because at that moment consumers might be forced to deviate from their typical payment activity. Thus, we expect cross-product activity to be increasingly informative for predicting net inflows versus credit line usage, overdrafts, or defaults. The findings confirm this expectation. The predictions of net inflows on checking accounts and credit card accounts prior to default cannot be improved significantly when cross-product information is considered. The prediction of future credit line usage can be improved slightly by considering cross-product information for checking accounts (adjusted $R^2 = 0.239$ instead of 0.234) and credit card accounts (adjusted $R^2 = 0.320$ instead of 0.319). In line with the reasoning above, informational synergies concerning overdrafts are more pronounced for both, checking accounts (adjusted McFadden $R^2 = 0.384$ instead of 0.373) and credit card accounts (adjusted McFadden $R^2 = 0.301$ instead of 0.286).[12]

Overall, there are sizeable benefits due to informational synergies concerning the prediction of consumer defaults. The marginal benefits are highest when lenders consider checking account activity in addition to credit card activity.

4. Additional empirical checks and robustness tests

4.1. Credit line usage at default

[12] For this analysis, we define a dummy variable which equals one if an overdraft can be observed in the subsequent 12 months, and zero otherwise.

In addition to the probability of default, banks estimate the credit line usage at default (or the absolute value of exposure at default) for credit risk management and regulatory capital requirements. We therefore also analyze potential informational synergies concerning the variable credit line usage at default.

The current credit line usage (*CLU*) of account i in month t is defined as $CLU_{i,t} = Exposure_{i,t}/Limit_{i,t}$. For estimation of the *CLU*, we choose a time horizon of one year. The *CLU* of account i in month t+12 can be written as $CLU_{i,t+12} = Exposure_{i,t+12}/Limit_{i,t+12}$. However, we would have to know the limit in t+12 to calculate the expected exposure in t+12 based on the estimated *CLU*. Thus, for predicting the *CLU*, we calculate the target variable $CLU_{i,t+12}$ as the ratio of the (at time t unknown) exposure in t+12 months and the (known) limit in t: $CLU_{i,t+12} = Exposure_{i,t+12}/Limit_{i,t}$. Following Jimenez, Lopez, and Saurina (2009), we could estimate the *CLU* as $\mathrm{E}(CLU_{i,t+12} | X_{i,t}) = X_{i,t}\beta$, where the explanatory variables $X_{i,t}$ includes account activity with or without cross-product information, bank-borrower relationship and borrower characteristics. In addition, we consider the credit rating (*Rating*) because it is likely that the *CLU* depends on the customer's default risk. We find that the rating discriminates between high and low *CLU* prior to default. The better the rating, the lower the *CLU* in the period between τ-24 and τ. This indicates that *Rating* is a key input for modeling the *CLU* at default, as found by Agarwal et al. (2006). Unreported results indicate that the results for credit card accounts are similar.

Though, we find that the *CLU* does not only depend on the rating, but the *CLU* also highly depends on the default status: The closer the default event, the higher is the *CLU*. If we ignore the dependency on the default status in the estimation, the estimates will be biased because the rating does not fully capture differences in the credit line usage between defaulters and non-defaulters: $(CLU_{i,t+12} \mid Def_{i,t+12} = 1, X_{it}) \neq \mathbb{E}(CLU_{i,t+12} \mid Rating_{i,t}, X_{it})$. For this reason, we implement a Heckman selection model (see Heckman, 1976, 1979) to account for a possible selection bias when estimating the *CLU* at default: $\mathrm{E}(CLU_{i,t+12} \mid Def_{i,t+12} = 1, X_{i,t}, Z_{i,t}) = X_{i,t}\beta + \delta_{\lambda}\lambda(Z_{i,t}\gamma)$, where the inverse Mills ratio λ is determined in a probit regression.

Next, we examine the consistency of information from different sources for estimating the *CLU* at default and whether cross-product information improves the estimate. In Table 6, we report the results for the Heckman selection models with different sources of information. Columns (1) and (2) provide the results for checking accounts, while columns (3) and (4) provide the results for credit cards.

(Insert Table 6 about here)

The results show that most coefficients for account activity and bank-customer relationship variables have the same sign for checking accounts and credit card accounts. However, a high *Amplitude* value is related to lower *CLU* at default for checking accounts, but to higher *CLU* at default for credit card accounts. A substantially larger fraction of variance can be explained for credit card accounts than for checking accounts (the adjusted R^2 is 0.562 vs. 0.173). This is contrary to our results for default prediction, where we find early warning indicators for checking accounts to be more informative. We further find that using information from different accounts can slightly improve the estimate of *CLU* at default. However, the informational benefit for credit line usage prediction is lower compared to default prediction.

4.2. Alternative definition of default

The previous analyses refer to the account level, as stipulated in the Basel III accord and EU Regulation No 575. We now repeat the analysis with default at the customer level. Under this definition, a default occurs if at least one account of the customer is in default. To estimate the *PD*, we use the variables of both accounts individually or aggregate the variables.[13] Table 7 reports the results. Analyses with aggregated variables are shown in columns (1) and (3) and analyses with variables of both individual accounts in columns (2) and (4).

(Insert Table 7 about here)

For the model of defaults with aggregated customer variables (column (1)), the adjusted R^2 is lower than for account-specific default models that were based on cross-product information (adj. R^2 of 0.191 versus 0.296 for checking accounts and 0.278 for credit card accounts). The sign of coefficients and statistical significances are as expected. The estimates using variables of both accounts individually (column (2)) are similar compared to the estimates for aggregated customer variables except for *Rating* and *Days Usage:* The coefficients for *Rating* in column (2) must be summed to be comparable with the *Rating* in column (1) because this is defined as the mean of both ratings. *Days Usage* in column (1) is slightly significant because of opposite effects for individual variables in column (2).

Estimating the credit line usage at customer default using aggregate customer variables is superior to estimation with individual account variables (adj. R^2 of 0.506 vs. 0.383). There are some differences regarding *Amplitude* and *Days Overdraft*. While high *Amplitude* on checking accounts is a positive signal, it is a negative signal on credit card accounts, which is consistent with our previous findings. We also find that the aggregated overdraft provide better early warning indications than the separate information. For estimating credit line usage at default with

[13] For the aggregated variables, we define the *Rating* as the mean of both ratings, *Net Inflow/Lim* as the difference between monthly total inflows and outflows as a percentage of the total external limit, *Amplitude* as the difference between the maximum and minimum exposure in each month as a percentage of *Limit* at both accounts, *Limit* and *Bounced* as the sum of limits and average number of bounced debits, and *Days Usage*, *Days Overdraft*, and *Duration* as the maximum of the corresponding variables. The *CLU* at default is limit weighted.

aggregated customer variables, the adjusted R^2 of 0.506 is higher/lower than for the models with cross-product information for checking accounts (0.182) and credit card accounts (0.579).

4.3. Interaction effects of cross-product account activity measures

We further investigate whether similar information from different sources leads to an amplification of early warning indications, which could be seen as further evidence for cross-product informational synergies. To address this issue, we add interaction terms of the activity measures from both accounts (*Net Inflow/Lim*, *Amplitude*, *Limit*, *Bounced*, *Days Usage*, and *Days Overdraft*) to the prediction models of default as well as *CLU* at default and repeat the analyses corresponding to Table 3 and 6 (unreported). Our key results are unchanged and the adjusted McFadden R^2 increases only marginally. The findings suggest that adding the interaction terms to the individual terms from both accounts results only in a marginal improvement.

4.4. Alternative functional forms for account activity variables

We examine whether our results are robust to the functional form in which the variables enter the models. For this purpose, we repeat the analyses of default prediction by including additional quadratic terms. Our unreported analyses show that the effect regarding the informational synergies remains the same. As an alternative model specification, we repeat analyses by including 49 dummy variables for the 50 quantiles for each interval-scaled variable to allow highly non-linear functional forms. We achieve higher adj. McFadden R^2 than presented in Table 3, but the effect of including informational synergies is similar or even slightly stronger. For example, the consideration of cross-product information increases the adj. McFadden R^2 by 8.0% for credit cards accounts compared to an increase by 7.4% for the linear specification from Section 3.2. Thus, our results on informational synergies also hold for non-linear functional forms.

4.5. Fixed effects models

Our estimation of the probability of default is based on pooled probit regression models. Although we included various control variables, our estimates could be biased due to unobserved customer characteristics. We therefore add customer fixed effects to our regression models. In unreported analyses we find that the coefficients of *Rating*, *Net Inflow/Lim, Amplitude*, *Bounced*, *Days Usage*, and *Days Overdrafts* are similar to the pooled probit estimates for both types of account. The coefficients for bounced debits and for an increased percentage of days with negative balance or overdrafts are even higher in the fixed effects model. Moreover, we find that cross-product information results in substantially increased estimates of the probability of default. Overall, the fixed effects regression results confirm our previous findings.

4.6. Impact of the recent financial crisis

The recent financial crisis could affect our results because customers are more likely to default and they have a higher demand for credit. Indeed, when we compare 2009-2010 (crisis) with 2011-2014 (post-crisis), we find higher default rates during the crisis for checking accounts (1.05% versus 0.83%, $p<0.001$) and credit card accounts (1.36% versus 1.00%, $p<0.001$). In addition, we find that the credit line usage at default is significantly higher in 2009-2010 ($p<0.001$); this shows that in a financial crisis, customers have a higher demand for credit. We repeat all analyses using observations from 2011-2014 (unreported). The findings for the *PD* and *CLU at default* are similar, indicating that they are robust in good and bad times.

5. Conclusion

We investigate potential synergies between different sources of private information in consumer credit. Informational synergies are important because they affect the supply and the allocation of credit in the economy. Our setting enables us to analyze why consumers default, how much credit they take, and how lenders can obtain early warning indications that capture the time-varying nature of consumer credit risk.

We provide evidence for significant informational synergies between different credit products of the same individuals. We find that the activity from checking accounts and credit card accounts contains information beyond credit scores, borrower characteristics, bank-borrower relationship characteristics and many other controls. Interestingly, in most situations information from checking accounts is more useful than that from credit cards, even when predicting defaults of credit cards. Activity measures from checking accounts indicate credit quality deteriorations earlier and more accurately. Particularly for short time horizons, however, a combination of both sources of information is highly beneficial. We also show that consumers default because of decreasing cash inflows, but not because of increasing cash outflows. Finally, lenders can lower the type I error by 33%, when they consider checking account activity in addition to credit card activity, suggesting sizeable benefits. Our results on cross-product informational synergies suggest that financial institutions can realize significant economies of scope in credit risk management and customer relationship management when they simultaneously offer different services to the same customer.

References

Agarwal, S., Ambrose, B. W., Liu, C., 2006. Credit lines and credit utilization. Journal of Money, Credit and Banking 38, 1-22.

Agarwal, S., Chomsisengphet, S., Liu, C., Souleles, N., 2009. Benefits of relationship banking: evidence from consumer credit markets. Working Paper, Federal Reserve Bank of Chicago, No. 2010-05.

Akçura, M.T., Srinivasan, K., 2005. Research note: customer intimacy and cross-selling strategy. Management Science 51, 1007-1012.

Allen, J., Damar, H., Martinez-Miera, D., 2016. Consumer Bankruptcy, Bank Mergers, and Information. Review of Finance 20, 1289-1320.

Bank for International Settlements, 2015. Statistics on payment, clearing and settlement systems in the CPMI countries.

Berg, T., Puri, M., Rocholl, J., 2014. Loan officer incentives, internal ratings and default rates. Working Paper, September 2014.

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

Boot, A. 2000. Relationship Banking: What Do We Know? Journal of Financial Intermediation 9, 7-25.

Campbell, J., 2006. Household Finance. Journal of Finance 61, 1553-1604.

Chatterjee, S., Corbae, D., Nakajima, M., Rios-Rull, J., 2007. A Quantitative Theory of Unsecured Consumer Credit with Risk of Default. Econometrica 75, 1525-1589.

Diamond, D., 1984. Financial Intermediation and Delegated Monitoring. Review of Economic Studies 51, 393–414.

Djankov, S., McLiesh, C., Shleifer, A., 2007. Private credit in 129 countries. Journal of Financial Economics 84, 299-329.

Doblas-Madrid, A., Minetti, R., 2013. Sharing information in the credit market: Contract-level evidence from U.S. firms. Journal of Financial Economics 109, 298-223.

Gross, D., Souleles, N., 2002. Do Liquidity Constraints and Interest Rates matter for Consumer Behavior? Evidence from Credit Card Data. Quarterly Journal of Economics 117, 149-185.

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

Gürtler, M., Hibbeln, M., 2013. Improvements in loss given default forecasts for bank loans. Journal of Banking and Finance 37, 2354-2366.

Heckman, J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 5, 475-492.

Heckman, J., 1978. Dummy endogenous variables in a simultaneous equation system. Econometrica 46, 931-959.

Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47, 153-161.

Hertzberg, A., Liberti, J., Paravasini, D., 2010. Information and incentives inside the firm: Evidence from loan officer rotation. Journal of Finance 65, 795-828.

Jiménez, G., Lopez, J., Saurina, J., 2009. Empirical Analysis of Corporate Credit Lines. Review of Financial Studies 22, 5069-5098.

Kamakura, W., Ramaswami, S., Srivastava, R., 1991. Applying latent trait analysis in the evaluation of prospects for cross-selling of financial services. International Journal of Research in Marketing 8, 329-349.

Kamakura, W., Wedel, M., De Rosa, F., Mazzon, J., 2003. Cross-selling through database marketing: a mixed data factor analyzer for data augmentation and prediction. International Journal of Research in Marketing 20, 45-65.

Karlan, D., Zinman, J., 2009. Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment. Econometrica 77, 1993-2008.

King, G., Zeng, L., 2001. Logistic Regression in Rare Events Data. Political Analysis 9, 137-163.

Li, S., Sun, B., Montgomery, A., 2011. Cross-selling the right product to the right customer at the right time. Journal of Marketing Research 48, 683-700.

Li, S., Sun, B., Wilcox, R., 2005. Cross-selling sequentially ordered products: An application to consumer banking services. Journal of Marketing Research 42, 233-239.

Liberman, A., 2016. The value of a good credit reputation: Evidence from credit card renegotiations. Journal of Financial Economics, forthcoming.

Livshits, I., Mac Gee, J., Tertilt, M., 2016. The Democratization of Credit and the Rise in Consumer Bankruptcies. Review of Economic Studies 83, 1673-1710.

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

Norden, L., Weber, M., 2010. Credit line usage, checking account activity, and default risk of bank borrowers. Review of Financial Studies 23, 3665-3699.

Pagano, M., Jappelli, T., 1993. Information Sharing in Credit Markets. Journal of Finance 43 1693-1718.

Rajan, U., Seru, A., Vig, V., 2015. The failure of models that predict failure: Distance, incentives, and defaults. Journal of Financial Economics 115, 237-260.

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

Stango, V., Zinman, J., 2016. Borrowing High versus Borrowing Higher: Price Dispersion and Shopping Behavior in the U.S. Credit Card Market. Review of Financial Studies 29, 979-1006.

Stiglitz, J., Weiss, E., 1981. Credit Rationing in Markets With Imperfect Information. American Economic Review 71, 393-410.

Vella, F., Verbeek, M., 1999. Estimating and Interpreting Models with Endogenous Treatment Effects. Journal of Business and Economic Statistics 17, 473-478.

Vissing-Jorgensen, A., 2012. Consumer Credit: Learning Your Customer's Default Risk from What (S)he Buys. Working Paper, April 2012.

Figure 1. Cash inflows and outflows of checking accounts

The long-dashed line presents inflows (median, in euros) for defaulted checking accounts from 24 months prior to default to 12 months after default. The dashed line presents inflows (median, weighted by defaults, in euros) for non-defaulted checking accounts. The short-dashed and solid lines present outflows (median/default weighted median, in euros) for defaulted/non-defaulted checking accounts up to 24 months prior to default.

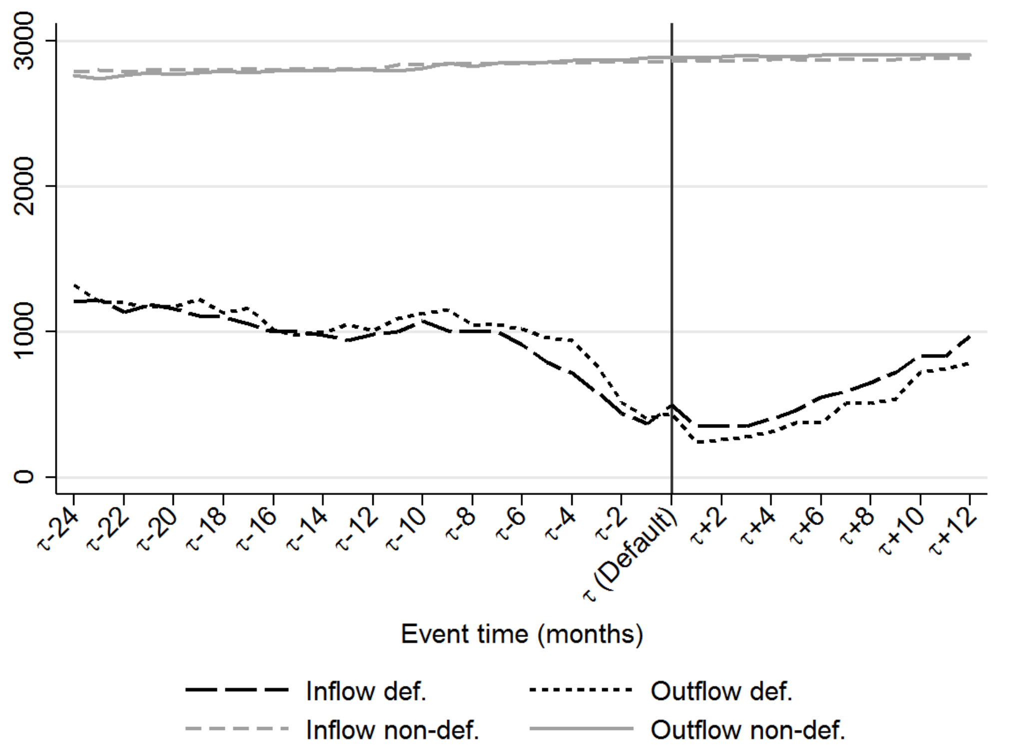
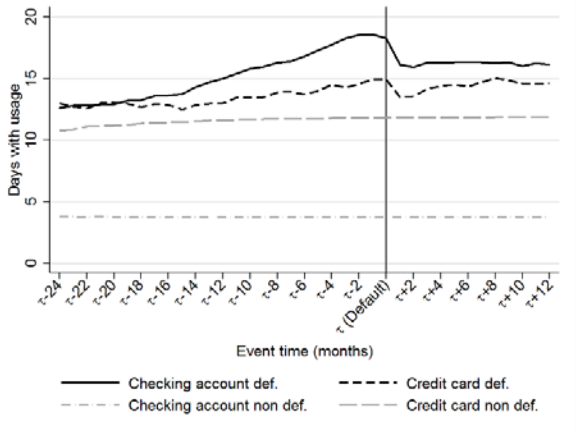


Figure 2. Days with credit line usage and overdrafts

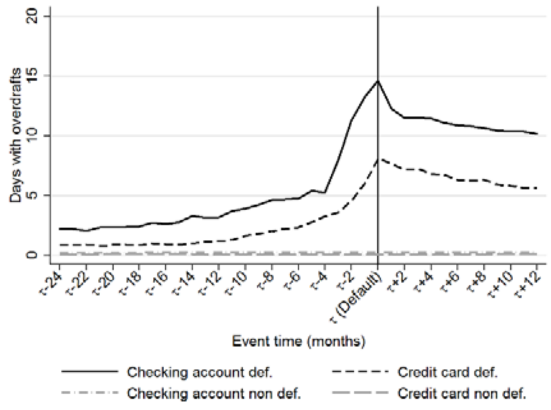
Panel A displays defaults of checking accounts, whereas Panel B displays defaults of credit cards. Days with usage are presented in Panel A1/B1 and days with overdrafts in Panel A2/B2. The solid and dash-dotted lines present days for defaulted (median) and non-defaulted (median, weighted by defaults) checking accounts, whereas the dashed and long-dashed lines present days for defaulted (median) and non-defaulted (median, weighted by defaults) credit card accounts.

Panel A: Default of checking accounts at τ

Panel A1: Days with usage

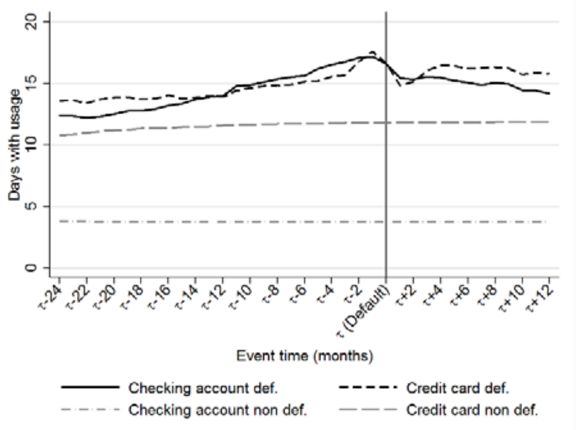


Panel A2: Days with overdrafts



Panel B: Default of credit cards at τ

Panel B1: Days with usage



Panel B2: Days with overdrafts

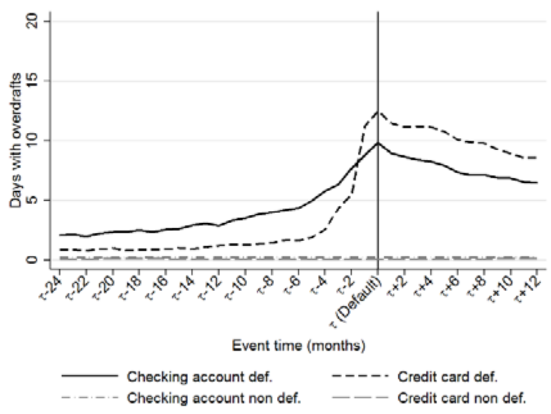
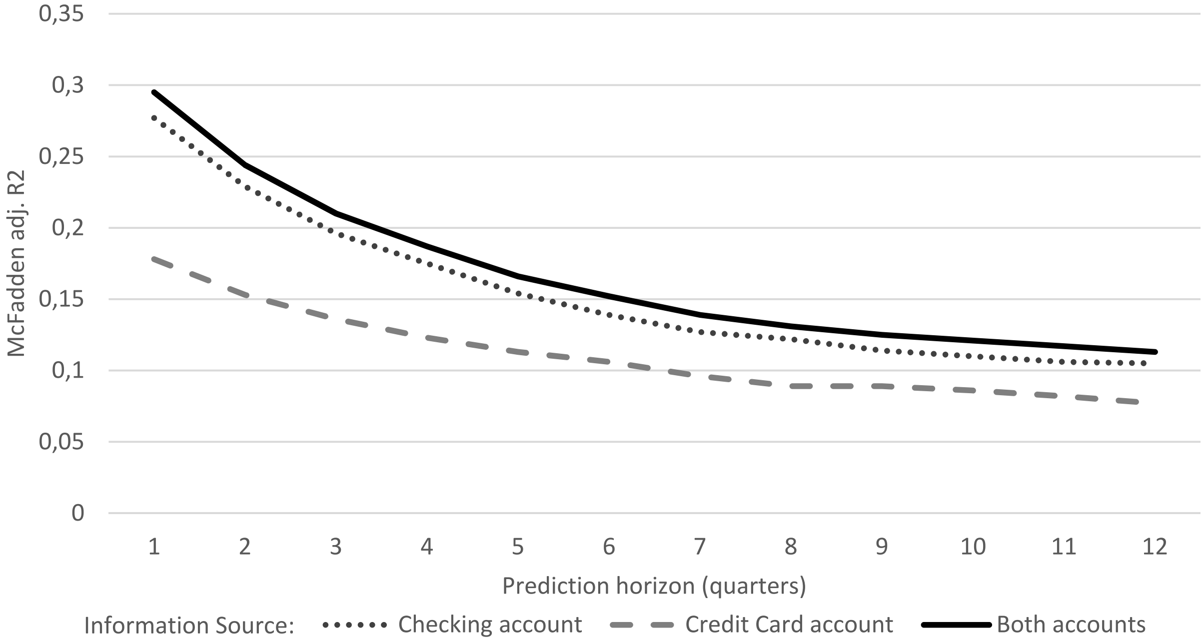


Figure 3. Accuracy of PD estimates for different time horizons by information source

This figure displays the accuracy of predicting defaults in 1, 2, …, 12 quarters, based on account activity information from checking accounts, credit card accounts, or both accounts. Panel A displays the accuracy for predicting defaults of checking accounts, whereas Panel B displays the corresponding accuracy for default prediction of credit cards.

Panel A: Default of checking accounts



Panel B: Default of credit cards

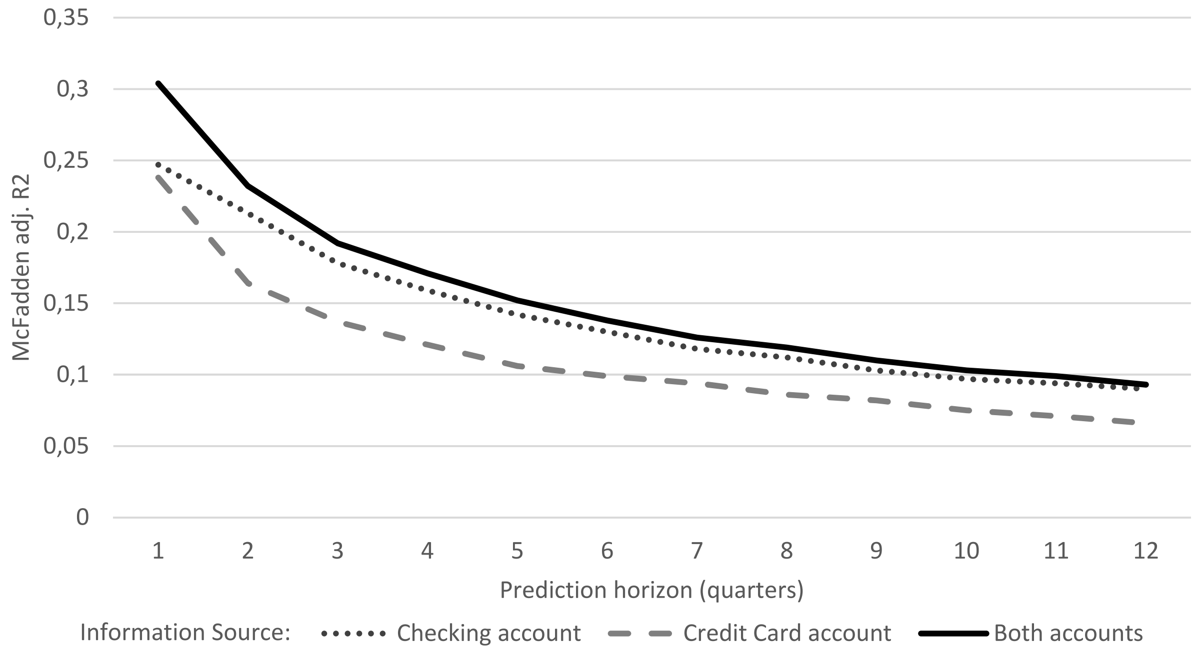


Table 1. Summary statistics

The sample spans the period from December 2007 to January 2014. Panel A reports the number of account-month observations and the frequency of default events. The default events are at account level. Panel B reports the number of monthly observations without default, with default of only checking accounts or credit cards, or with joint defaults of both accounts. Panel C provides summary statistics of default risk, account activity, relationship, and account type variables. *Rating* is the probability of default estimate of the bank's internal rating system. *Net Inflow/Limit* is the ratio of monthly inflows minus outflows to the external limit and can be positive or negative. *Amplitude* is the difference of the monthly maximum and minimum balance per Limit and hence positive. *Limit* is reported in Euro values. *Bounced* is the average number of bounced credits in the previous year. *Days Usage* and *Days Overdrafts* are the average percentage of days in the previous year with positive credit line usage and overdrafts, respectively. *Duration* is defined as the time period since account opening in months. *Full Payment* refers to credit cards with monthly full payments instead of delayed payments. The last column reports the pairwise comparison of checking accounts and credit cards on the customer level. ***: Statistically significant at the 0.1% level.

Panel A: Number of account-month observations and default events

Statistic	Checking Account	Credit Card
Number of account-month observations	1,779,356	1,781,408
Number of account-months with default in the subsequent 12 months	19,149	22,964
Number of defaults	1,639	2,101

Panel B: Number of account-month observations with individual vs. joint defaults

	Checking Account: No Default	Checking Account: Default
Credit Card: No default	1,754,099	5,857
Credit Card: Default	7,565	13,887

Panel C: Summary statistics and comparisons of checking accounts and credit cards

	Checking Account		Credit Card		Checking Account Minus Credit Card	
Variable	Mean	Median	Mean	Median	Pairwise Difference	*t*-stat.
Default Risk						
$Rating_t$	1.79	0.88	3.79	1.94	-1.939***	-428.25
Account Activity						
$Net\ Inflow/Lim_t$	0.02	0.00	0.00	0.00	0.0195***	19.18
$Amplitude_t$	1.80	1.03	0.39	0.24	1.407***	727.72
$Limit_t$	2437	2000	2581	2000	-143.7***	-163.29
$Bounced_t$	0.00	0.00	0.05	0.00	-0.0500***	-182.49
$Days\ Usage_t$	0.18	0.03	0.55	0.67	-0.367***	-1269.96
$Days\ Overdraft_t$	0.01	0.00	0.00	0.00	0.0070***	201.34
Relationship						
$Duration_t$	46.38	41.00	34.99	34.00	11.46***	616.02
Account Type						
$Full\ Payment_t$			0.93	1.00		

Table 2. Characteristics of accounts in default versus non-default

This table reports the average differences between the accounts in default and non-default for several explanatory variables. In addition to the variables presented in Table 1, models (2) and (4) contain the variable *Prev. Def*, which is one if a previous default is observed on a different account of the same customer, and zero otherwise. The explanatory variables are observed in month t, and the default variable is equal to 1 if a default occurs in the period [t, t+12 months]. Columns (1) and (2) refer to defaults versus non-defaults of checking accounts, and columns (3) and (4) refer to defaults versus non-defaults of credit card accounts. The explanatory variables in models (1) and (4) are checking account variables, whereas the explanatory variables in models (2) and (3) are from credit card accounts. Thus, models (2) and (4) refer to cross-product information. We report t-statistics in parentheses. *, **, and ***: Statistically significant at the 5%, 1%, and 0.1% levels, respectively. (Note: Limit in 1,000 euros.)

	Default_{t+12} vs.$\text{Non-Default}_{t+12}$: Checking account		Default_{t+12} vs.$\text{Non-Default}_{t+12}$: Credit card	
	Inform. Source: Checking Account	Inform. Source: Credit Card	Inform. Source: Credit Card	Inform. Source: Checking Account
	(1)	(2)	(3)	(4)
Default Risk				
$Rating_t$	1.825^{***}	1.079^{***}	1.110^{***}	1.729^{***}
	(191.16)	(129.85)	(146.77)	(197.96)
Account Activity				
$Net\ Inflow/Lim_t$	-0.0940^{***}	-0.0474^{***}	-0.0651^{***}	-0.0880^{***}
	(-9.97)	(-16.07)	(-24.21)	(-10.23)
$Amplitude_t$	-0.814^{***}	0.128^{***}	0.280^{***}	-0.587^{***}
	(-41.11)	(29.70)	(71.81)	(-32.51)
$Limit_t$	-0.716^{***}	-1.072^{***}	-1.063^{***}	-0.693^{***}
	(-54.84)	(-83.33)	(-87.05)	(-58.19)
$Bounced_t$	0.00677^{***}	0.336^{***}	0.345^{***}	0.00761^{***}
	(16.19)	(120.29)	(139.36)	(19.41)
$Days\ Usage_t$	0.456^{***}	0.0355^{***}	0.0644^{***}	0.420^{***}
	(218.64)	(14.03)	(27.93)	(219.72)
$Days\ Overdraft_t$	0.109^{***}	0.0464^{***}	0.0482^{***}	0.115^{***}
	(309.61)	(216.29)	(304.21)	(325.25)
Relationship				
$Duration_t$	-17.42^{***}	-12.22^{***}	-12.27^{***}	-16.51^{***}
	(-71.83)	(-77.06)	(-84.87)	(-73.23)
$Prev.\ Def_t$		0.0210^{***}		0.0419^{***}
		(51.18)		(93.13)
Account Type				
$Full\ Payment_t$		-0.263^{***}	-0.231^{***}	
		(-135.71)	(-130.24)	
Number of obs.	1,779,356	1,779,356	1,781,408	1,781,408

Table 3. PD of checking accounts, credit cards and cross-product information

This table reports probit estimates of the probability of default (*PD*) for several explanatory variables, using different sources of information. The explanatory variables are observed in month *t*, and the default variable is equal to 1 if a default occurs in the period [t, *t*+12 months]. Columns (1)-(4) refer to defaults of checking accounts, and columns (5)-(8) refer to defaults of credit card accounts. The explanatory variables in the upper half of the table are checking account variables, whereas the explanatory variables in the bottom half are from credit cards. Thus, columns (4) and (8) include cross-product information. We report *t*-statistics clustered at the customer level in parentheses. †, *, **, and ***: Statistically significant at the 10%, 5%, 1%, and 0.1% levels, respectively. (Note: Limit in 1,000 euros.)

	Default$_{t+12}$: Checking Account				Default$_{t+12}$: Credit Card			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CHECKING ACCOUNT VARIABLES								
Default Risk								
Rating$_t$	0.4626***		0.2215***	0.1703***				0.1314***
	(52.184)		(19.633)	(14.915)				(13.178)
Account Activity								
Net Inflow/Lim$_t$		-0.0479***	-0.0561***	-0.0616***				-0.0475***
		(-14.394)	(-15.847)	(-17.407)				(-18.548)
Amplitude$_t$		-0.0687***	-0.0645***	-0.0614***				-0.0360***
		(-7.902)	(-7.470)	(-7.348)				(-6.242)
Limit$_t$		-0.1289***	-0.1322***	-0.0842***				-0.0750***
		(-12.371)	(-12.231)	(-6.678)				(-6.681)
Bounced$_t$		0.2471*	0.2288*	0.1853†				0.1649
		(2.489)	(2.249)	(1.779)				(1.571)
Days Usage$_t$		1.2349***	0.8366***	0.9115***				0.7657***
		(38.682)	(22.237)	(22.089)				(20.266)
Days Overdraft$_t$		2.1567***	1.8995***	1.4664***				1.1692***
		(21.982)	(19.126)	(13.412)				(11.174)
Relationship								
Duration$_t$		-0.0084***	-0.0055***	-0.0027***				-0.0018**
		(-14.432)	(-9.420)	(-4.113)				(-3.006)
Prev. Def$_t$								-0.0191
								(-0.236)
CREDIT CARD VARIABLES								
Default Risk								
Rating$_t$				0.1284***	0.3947***		0.2814***	0.1626***
				(10.279)	(40.648)		(23.018)	(12.498)
Account Activity								
Net Inflow/Lim$_t$				-0.1078***		-0.1226***	-0.1493***	-0.1418***
				(-17.330)		(-28.027)	(-26.392)	(-22.740)
Amplitude$_t$				0.0811***		0.1323***	0.1518***	0.1765***
				(8.698)		(19.145)	(19.972)	(23.282)
Limit$_t$				-0.0241†		-0.1057***	-0.0667***	0.0053
				(-1.802)		(-11.646)	(-7.303)	(0.468)
Bounced$_t$				0.0343*		0.1628***	0.1402***	0.0554***
				(2.323)		(7.113)	(6.692)	(3.594)
Days Usage$_t$				-0.5843***		0.1516***	-0.3906***	-0.4539***
				(-13.448)		(5.276)	(-10.949)	(-11.408)
Days Overdraft$_t$				0.8898***		4.7102***	4.1120***	2.2709***
				(5.346)		(22.811)	(20.775)	(12.012)
Relationship								
Duration$_t$				-0.0068***		-0.0120***	-0.0104***	-0.0080***
				(-6.929)		(-17.658)	(-15.669)	(-8.839)
Prev. Def$_t$				-0.1909†				
				(-1.651)				
Account Type								
Full Payment$_t$				-0.1120***		-0.5604***	-0.2670***	-0.0398
				(-3.627)		(-22.549)	(-9.105)	(-1.224)
Constant	-2.2031***	-1.9650***	-1.9854***	-1.9658***	-2.5644***	-1.2397***	-1.9346***	-2.1022***
	(-47.816)	(-24.078)	(-23.254)	(-23.286)	(-60.191)	(-15.919)	(-22.661)	(-23.216)
Customer controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs.	1,779,356	1,779,356	1,779,356	1,779,356	1,781,408	1,781,408	1,781,408	1,781,408
McFadden adj. R^2	0.185	0.259	0.277	0.296	0.101	0.185	0.204	0.278

Table 4. Screening with cross-product information

This table reports estimates of the probability of default for several explanatory variables. The explanatory variables are observed in month t, and the default variable is equal to 1 if a default occurs in the period [t, t+12 months]. Column (1) refers to the estimation of probability of default of checking accounts, and column (2) refers to the corresponding estimation of credit card accounts. The explanatory variables for checking accounts defaults are credit card variables, whereas the explanatory variables for credit card defaults are from checking accounts. Thus, this table includes pure cross-product information. We report t-statistics clustered at the customer level in parentheses. †, *, **, and ***: Statistically significant at the 10%, 5%, 1%, and 0.1% levels, respectively. (Note: Limit in 1,000 euros.)

Dependent Variable	Default_{t+12} of Checking Account	
Independent Variables	Credit Card Account	Checking Account
	(1)	(2)
Default Risk		
$Rating_t$	0.2715^{***}	0.1718^{***}
	(24.155)	(16.769)
Account Activity		
$Net\ Inflow/Lim_t$	-0.1144^{***}	-0.0426^{***}
	(-21.253)	(-15.034)
$Amplitude_t$	0.0387^{***}	-0.0153^{**}
	(3.950)	(-2.873)
$Limit_t$	-0.1012^{***}	-0.0981^{***}
	(-9.525)	(-9.795)
$Bounced_t$	0.1225^{***}	0.1986^{*}
	(7.837)	(1.994)
$Days\ Usage_t$	-0.4614^{***}	0.7627^{***}
	(-12.775)	(21.797)
$Days\ Overdraft_t$	2.1146^{***}	1.7604^{***}
	(13.249)	(18.710)
Relationship		
$Duration_t$	-0.0094^{***}	-0.0052^{***}
	(-13.951)	(-9.849)
$Prev.\ Def_t$	0.1383	$0.1251^{\dagger}$
	(1.426)	(1.799)
Account Type		
$Full\ Payment_t$	-0.3769^{***}	-0.1618^{***}
	(-14.280)	(-5.984)
Constant	-1.8416^{***}	-1.6782^{***}
	(-21.021)	(-20.378)
Customer controls	Yes	Yes
Year-month FE	Yes	Yes
Number of obs.	1,779,356	1,781,408
McFadden adj. R^2	0.185	0.250

Table 5. Accuracy of PD estimates by information source

This table reports the accuracy for predicting defaults of checking accounts and credit card accounts. The estimates are based on (1) checking account information, (2) credit card information, and (3) information from both accounts. The accuracy of *PD*s is measured by the adjusted McFadden R^2, the area under the receiver operating characteristic (ROC) and the type I and type II errors. We use the empirical default rate in the checking account (credit card account) sample of 1.00% (1.20%) as cut-off point to calculate the type I and type II errors. The checking account sample comprises 1,779,356 observations and the credit card account sample 1,781,408 observations.

Dependent Variable		Information Source		
	Measure	(1) Checking Account	(2) Credit Card Account	(3) Both Accounts
Default$_{t+12}$ of checking account	Adj. McFadden R^2	0.277	0.185	0.296
	ROC	0.906	0.843	0.913
	Type I error	13.99	23.81	13.69
	Type II error	18.03	23.63	17.38
Default$_{t+12}$ of credit card account	Adj. McFadden R^2	0.250	0.204	0.278
	ROC	0.882	0.845	0.892
	Type I error	16.36	23.76	15.89
	Type II error	19.55	22.24	18.48

Table 6. Credit lines usage at default and cross-product information

This table reports estimates of the credit line usage at default for checking accounts, credit card accounts, and with cross-product information. Estimates are based on a Heckman selection model. The explanatory variables are observed in month *t*. Columns (1) and (2) refer to defaulted checking accounts and columns (3) and (4) refer to defaults of credit card accounts. The explanatory variables in the upper half of the table are checking account variables, whereas the explanatory variables in the bottom half are observed on credit card accounts. Thus, columns (2) and (4) include cross-product information. We report *t*-statistics clustered at the customer level in parentheses. †, *, **, and ***: Statistically significant at the 10%, 5%, 1%, and 0.1% levels, respectively. (Note: Limit in 1,000 euros)

	CLU at Default$_{t+12}$: Checking Account		CLU at Default$_{t+12}$: Credit Card	
	(1)	(2)	(3)	(4)
CHECKING ACCOUNT VARIABLES				
Account Activity				
Net Inflow/Lim$_t$	-0.0797***	-0.0772***		0.0391***
	(-6.939)	(-6.752)		(5.409)
Amplitude$_t$	-0.0269***	-0.0431***		-0.0670***
	(-4.673)	(-7.311)		(-17.159)
Limit$_t$	-0.0448***	-0.0565***		0.0989***
	(-5.600)	(-5.621)		(14.023)
Bounced$_t$	0.0726	0.0842		0.0093
	(1.102)	(1.282)		(0.205)
Days Usage$_t$	0.7438***	0.7127***		0.0127
	(19.015)	(17.905)		(0.480)
Days Overdraft$_t$	0.4762***	0.2343**		0.3970***
	(6.842)	(3.049)		(7.242)
Relationship				
Duration$_t$	-0.0037***	-0.0022***		0.0004
	(-8.520)	(-4.265)		(1.149)
Prev. Def.$_t$				0.2033***
				(5.685)
CREDIT CARD VARIABLES				
Account Activity				
Net Inflow/Lim$_t$		-0.0384	-0.2376***	-0.2122***
		(-1.582)	(-15.200)	(-13.784)
Amplitude$_t$		0.1228***	0.9247***	0.9308***
		(9.919)	(122.519)	(123.307)
Limit$_t$		0.0074	-0.1519***	-0.2062***
		(0.655)	(-23.634)	(-26.502)
Bounced$_t$		0.0037	-0.0077	-0.0219**
		(0.384)	(-1.077)	(-3.101)
Days Usage$_t$		0.2206***	0.0656*	0.0349
		(5.636)	(2.087)	(1.109)
Days Overdraft$_t$		0.0502	0.9162***	0.5432***
		(0.406)	(10.823)	(6.080)
Relationship				
Duration$_t$		-0.0026**	-0.0027***	-0.0039***
		(-3.270)	(-5.785)	(-6.930)
Prev. Def$_t$		0.1705*		
		(2.271)		
Account Type				
Full Payment$_t$		-0.0766**	-0.5989***	-0.5521***
		(-2.802)	(-25.640)	(-24.383)
Constant	2.3201***	2.1924***	2.2606***	2.2203***
	(13.925)	(13.107)	(18.071)	(17.959)
Lambda: Default	-0.7217***	-0.6953***	-0.5551***	-0.5242***
	(-26.960)	(-25.688)	(-18.083)	(-17.316)
Customer Controls	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes
Number of obs.	19,149	19,149	22,964	22,964
Adj. R^2	0.173	0.182	0.562	0.579

Table 7. Default at the customer level

This table reports estimates of the probability of default and credit line usage at default at the customer level (instead of account level) for several explanatory variables. The explanatory variables are observed in month *t*, and the default variable is equal to 1 if a customer default occurs in the period [*t*, *t*+12 months]. Estimates for credit line usage at default are based on the Heckman selection model. Columns (1)-(3) refer to the estimation of probability of default, and columns (4)-(6) refer to credit line usage at default. Models (1) and (3) refer to aggregated customer variables as independent variables, whereas models (2) and (4) contain estimates with non-aggregated variables. Part 1 refers to checking account variables and part 2 refers to credit card variables. (Note: For brevity, we do not present these models in one column.) We report *t*-statistics clustered at the customer level in parentheses. †, *, **, and ***: Statistically significant at the 10%, 5%, 1%, and 0.1% levels, respectively. (Note: Limit in 1,000 euros.)

Dependent Variable	Default$_{t+12}$			CLU at Default$_{t+12}$		
Independent Variables	Customer	Checking Acc. (part 1)	Credit Card Acc. (part 2)	Customer	Checking Acc. (part 1)	Credit Card Acc. (part 2)
	(1)	(2)		(3)	(4)	
Default Risk						
Rating$_t$	0.2870***	0.1266***	0.1129***			
	(16.681)	(9.963)	(6.920)			
Account Activity						
Net Inflow/Lim$_t$	-0.0650***	-0.0304***	-0.0604***	-0.0235	-0.0351***	-0.0713**
	(-12.227)	(-11.037)	(-6.333)	(-1.397)	(-4.012)	(-3.010)
Amplitude$_t$	-0.0125	-0.0068	-0.0104	-0.3231***	-0.0346***	0.0721***
	(-1.264)	(-1.496)	(-0.720)	(-64.016)	(-7.496)	(4.776)
Limit$_t$	-0.0332***	-0.0581***	-0.0024	-0.0000***	-0.0000	-0.0279**
	(-5.395)	(-4.457)	(-0.184)	(-6.540)	(-0.553)	(-2.914)
Bounced$_t$	0.0435	0.1571	0.0272	0.0230	0.0735	-0.0049
	(1.119)	(1.481)	(1.355)	(1.450)	(1.526)	(-0.518)
Days Usage$_t$	-0.1101*	0.4570***	-0.2879***	0.3409***	0.5208***	0.1273***
	(-2.311)	(9.817)	(-5.866)	(9.420)	(14.287)	(3.756)
Days Overdraft$_t$	2.5958***	1.8337***	1.0824***	0.5638***	-0.0974	0.2214†
	(22.454)	(14.296)	(4.924)	(9.949)	(-1.376)	(1.665)
Relationship						
Duration$_t$	-0.0020***	-0.0004	-0.0050***	-0.0029***	-0.0007	-0.0021**
	(-3.449)	(-0.529)	(-4.761)	(-8.563)	(-1.607)	(-3.031)
Account Type						
No Full$_t$	-0.1517***		-0.1750***	0.0122		-0.0360
	(-3.887)		(-4.561)	(0.541)		(-1.444)
Constant	-2.4611***	-2.3155***		2.5455***	3.1527***	
	(-22.179)	(-20.966)		(16.645)	(18.367)	
Lambda: Default				-0.6631***	-1.0522***	
				(-22.396)	(-31.304)	
Customer controls	Yes	Yes		Yes	Yes	
Year-month FE	Yes	Yes		Yes	Yes	
Number of obs.	1,767,269	1,767,269		11,621	11,621	
McFadden adj. R^2/adj. R^2	0.191	0.207		0.506	0.383	