

Informational Synergies in Consumer Credit

Martin Hibbeln ^{a,*}, Lars Norden ^b, Piet Usselman ^c, Marc Gürtler ^d

^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*

^d *Braunschweig Institute of Technology, Abt-Jerusalem-Straße 7, 38106 Braunschweig, Germany*

Abstract

Lenders consider multiple sources of private information to assess consumer credit risk but little is known about informational synergies between these sources. Using 1.7 million monthly observations from checking accounts and credit card accounts of the same individuals during 2007-2014, we find that activity from both accounts contains information beyond credit scores and other controls. Checking accounts display warning indications earlier and more accurately than credit card accounts. Type I default prediction errors decrease by 33% when both information sources are considered. The evidence suggests sizeable informational synergies that are relevant for the supply and the allocation of consumer credit.

Key words: Credit risk, checking accounts, credit cards, consumer default, household finance

JEL classification: G20, G21, D12, D14

* Corresponding author: Martin Hibbeln; University of Duisburg-Essen, Lotharstr. 65, 47057 Duisburg, Germany; Phone: +49 203 379 2830; E-mail: martin.hibbeln@uni-due.de.

Informational Synergies in Consumer Credit

Abstract

Lenders consider multiple sources of private information to assess consumer credit risk but little is known about informational synergies between these sources. Using 1.7 million monthly observations from checking accounts and credit card accounts of the same individuals during 2007-2014, we find that activity from both accounts contains information beyond credit scores and other controls. Checking accounts display warning indications earlier and more accurately than credit card accounts. Type I default prediction errors decrease by 33% when both information sources are considered. The evidence suggests sizeable informational synergies that are relevant for the supply and the allocation of consumer credit.

Key words: Credit risk, checking accounts, credit cards, consumer default, household finance

JEL classification: G20, G21, D12, D14

1. Introduction

Financial institutions, telecommunication providers and retailers apply specific lending technologies in appraising consumer credit that require information production. Lenders gather hard and soft information from public and private sources to assess the credit quality of loan applicants and existing borrowers. In addition, there are credit information sharing mechanisms through which private information of individual lenders becomes collectively shared public information (e.g., Liberman, 2016; Doblas-Madrid and Minetti, 2013; Djankov, McLiesh and Shleifer, 2007; Pagano and Jappelli, 1993). There is consensus that a combination of different types of information is useful for credit risk assessment. However, private information production remains a black box, mainly due to the lack of data (Campbell, 2006). How useful is private information for assessing consumer credit risk? Are there synergies between different sources of private information? If yes, can lenders exploit these synergies to improve their credit decisions? Consumer credit is characterized by a large number of transactions, relatively small amounts 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). Understanding private information production is at the core of financial intermediation and potential informational synergies between sources of private information might affect the supply and the allocation of consumer credit in the economy.

In this paper we investigate whether there are informational synergies between the account activities from different credit products of the same individual. We focus on a subset of consumer credit products: credit card accounts and checking accounts. Most of the recent credit-related research in household finance focuses on mortgage lending, while there has been little attention to credit cards and/or checking accounts (for exceptions, see Stango and Zinman, 2014; Norden and Weber, 2010). Unlike mortgages, the latter credit products exhibit a time-varying risk exposure, no maturity, and no collateral. Monitoring their time-varying credit exposure is critical for the lender. Account activity measures such as overdrafts or the amplitude reflect private hard information that gives the lender a real time window into the borrower's cash in- and outflows. Our setting enables us to analyze why consumers default and how early lenders

learn about credit quality deterioration. We compare the account activity measures with traditional credit scores, borrower characteristics, and bank-borrower relationship characteristics.

We base the analysis on a large and unique panel dataset comprising more than 1.7 million monthly observations from checking accounts and credit card accounts of the same individuals. The data come from a large privately owned financial institution and span the period from 2007 to 2014. The widespread use and simultaneous presence of consumer checking accounts with lines of credit attached to these accounts¹ and credit card accounts enables us to examine informational synergies.

We obtain the following results. First, account activity measures from both credit products contain useful information for default prediction beyond traditional credit scores, borrower characteristics, and bank-borrower relationship characteristics. Such private information can be used for monitoring and screening. Second, account activity measures from checking accounts are more useful for default prediction than those from credit cards. Checking accounts display warning indications earlier and more accurately than credit card accounts. The main explanation is that the former reflect cash in- *and* outflows, while the latter reflect *only* cash outflows. We further observe that consumers default because of decreasing cash inflows, but not because of increasing cash outflows. Third, the private information derived from different sources displays an overlap, but none of the sources is redundant. Default predictions based on information from different sources exhibit a positive but imperfect correlation of 0.55. Fourth, the informational synergies we document here are sizeable. Lenders can improve their decision making, by lowering the type I error by around 33 percent when they consider checking account activity in addition to credit card account activity. We conduct several additional empirical checks and show that our previous results are robust and not the product of particular choices of samples or model specifications.

Our study provides novel evidence on the time-varying nature of consumer credit risk and demonstrates how lenders can exploit cross-product informational synergies. The informational synergies are likely to reduce informational asymmetries and improve the allocative efficiency in consumer credit

¹ It is one account that may have a positive or negative balance. The negative balance has to stay below the authorized external limit. Overdrafts are bounced by default but can be authorized manually by loan officers. Such consumer checking accounts are widespread in Europe, Canada, and many Asian countries.

markets because the supply and demand side are matched based on superior information. Studying private credit information production (our study), credit information sharing (e.g., Pagano and Jappelli, 1993, and several other studies) and their effects on allocative efficiency is complementary to studying whether there is credit over- or undersupply (e.g., Zinman, 2014). Furthermore, our findings suggest that, next to the “skin in the game effect”, lending and risk assessment should be carried out by the same institution because lending itself generates proprietary real time information 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 arising from volume-based compensation schemes, career concerns and credit reputation concerns (e.g., Liberman, 2016; Berg, Puri and Rocholl, 2014; Hertzberg, Liberti and Paravisini, 2010). A risk assessment based on different sources of private information is less likely to be biased by strategic behavior of the borrower and/or loan officer and thereby provides more comprehensive input to credit information sharing systems than a risk assessment based on a single source of private information.

We contribute to the literature on information production in banking and consumer credit in several ways. First, there are only a few studies about the link between account activity measures and information production, most of these focus on bank lending to small businesses (e.g., Norden and Weber, 2010; Jimenez, Lopez and Saurina, 2009; Mester, Nakamura, and Renault, 2007). One exception is the study of Norden and Weber (2010) that examines the credit line usage and checking account activity of large firms, small firms and consumers. They find that private hard information from checking accounts is particularly useful for monitoring consumers and small firms, but not for large firms. They also show how this information is used for loan pricing, credit limit management, and account closure decisions. For comparison, we investigate potential synergies between different sources of private information and focus on consumer credit. Agarwal et al. (2009) document that credit card customers with close relationships to their bank exhibit lower probabilities of default and have higher credit line usage than non-relationship customers. Whereas they focus on information about the bank-customer relationship, such as the existence and duration of other credit products, we analyze dynamic information about customers’ payment behavior on different accounts. Stango and Zinman (2014) analyze consumer panel data for checking

account activity and overdrafts from 2006 to 2008 to investigate how attention biases affect consumers' responses to financial shocks. In their study, they take the consumer's perspective, consider overdrafts but not authorized lines of credit and/or credit card accounts, and examine attention biases, while we take the lender's perspective and examine how private hard information from consumers' checking accounts and credit card accounts can be used to manage credit relationships.

Second, there is evidence suggesting a link between individual consumption patterns and consumer default risk (e.g., Vissing-Jorgensen, 2012; Stango and Zinman, 2016). Vissing-Jorgensen (2012) examines consumer purchase data from a Mexican retail chain to investigate whether the type of products a consumer buys provides information about his/her default risk. In her setting, consumer purchases are financed with consumer loans provided by the retailer. All purchases are financed with the same credit product. She shows that products that lead to high losses on consumer loans tend to be luxury goods purchased by consumers who buy abnormally large fractions of luxury goods given their income. She documents a link between consumer choice and the default risk of consumer loans, whereas we investigate potential synergies between private information from different credit products used by the same consumer. Stango and Zinman (2016) show that the dispersion of credit card rates, which different financial institutions offer to the same individuals, is significantly related to their shopping intensity, controlling for the individuals' credit risk.

Third, the informational synergies in consumer credit are not only relevant for credit risk management but also for the cross-selling of financial products, which can be beneficial for banks and bank customers. For instance, the marketing costs are lower, switching costs for customers are higher, and banks can learn more about their customers' risk preferences and consumption behavior (e.g., Kamakura et al., 2003). Customers with more valuable information for cross-selling purposes may pay lower prices (Akçura and Srinivasan, 2005). Studies on the cross-selling of financial products have mainly focused on which product a customer is most likely to buy next in order to offer the right product to the right customer at the right time (Kamakura et al., 1991; Li, Sun and Wilcox, 2005; Li, Sun and Montgomery, 2011). In our study, we document significant informational synergies that can be used for risk management and customer relationship management.

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. It is important that no brokerage services are offered by the bank so that we can exclude that checking accounts are used as clearing accounts for brokerage accounts.

Furthermore, we consider credit card accounts (Visa credit cards). Our dataset comprises two types of credit card accounts, which differ in their redemption schemes. There are accounts with full repayment of the credit card bill every month. These accounts are the majority of all credit card accounts. 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 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

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.

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-product information if irregularities occur, and overdrafts are manually approved or rejected by the relevant 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 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

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.

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. *Days Usage* is 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 one 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 $Def_{i,t+\tau}$ that equals one if a jump to default is observed at $t+\tau$ with $\tau \in \{1, 2, \dots, 12 \text{ months}\}$ for account i , and zero otherwise. Given the explanatory variables $Z_{i,t}$ that can be observed in month t , the estimated *PD* is:

$$PD_{i,t} := P(Def_{i,t+\tau} = 1 | Z_{i,t}) = f(Z_{i,t}). \quad (1)$$

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 ($=PD_{i,t}$) of the internal rating as $Rating_{i,t} = \log(PD_{i,t})$. 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 our results also hold when we use quadratic terms and other nonlinear functional forms, respectively. These tests confirm that the choice of the functional form of the explanatory variables is not critical for our main results.

3. Results

3.1. Monitoring with cross-product information

We first conduct a univariate event study of the balance and limit of checking accounts for defaulted and non-defaulted customers. We show the explanatory 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 ($\tau-24, \tau-23, \dots, \tau, \dots, \tau+12$). 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 ($\tau-24, \tau-23, \dots, \tau, \dots, \tau+12$) and compute the median of the monthly medians for each event time. Thus, we assume a hypothetical default event for the non-defaulted customers.

Figure 1 shows both variables for defaulted and non-defaulted checking accounts. On the x-axis, the time relative to the default event at τ is displayed in months, and on the y-axis the median *Limit* and the median of $(-1) \times$ *Balance* are shown in euros. For defaulted accounts, we see that the negative *Balance* is already increasing around 14 months before default. The negative *Balance* is slightly decreasing after default because the *Limit* at the time of default is often set to zero and no further increase in exposure is possible. *Limit* is typically reduced to zero in the month of default. Default is often due to overdrafts,

which is supported by the observation that *Limit* approximately equals the negative *Balance* one month before default. All findings hold for the 10% and 90% quantiles, as well as for credit card accounts (not reported). For non-defaulted customers, *Limit* (positive) and *Balance* are usually higher. This difference between non-defaulted and defaulted accounts can already be observed 24 months prior to default and increases as the time to default approaches. From Figure 1, we cannot see whether the usage of credit lines is increasing due to decreasing inflows (e.g., unemployment) or increasing outflows (e.g., higher consumption or unexpected expenses).

(Insert Figure 1 about here)

In Figure 2, we show the customers' payment behavior regarding cash inflows and outflows separately for defaulted and non-defaulted checking accounts. The gap between cash inflows and outflows, the *Net Inflow*, directly leads to the change of *Balance* shown in Figure 1. These inflows and outflows, as well as the *Net Inflow* are more informative for checking accounts than for credit card accounts. For credit cards only, the outflows are influenced by the customer, while the inflows correspond mechanically to the sum of the outflows because of monthly clearing.

(Insert Figure 2 about here)

Figure 2 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, 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, as shown in Figure 1. The findings indicate that for most customers, the reason for default on a checking account are decreasing cash inflows, but not increasing cash outflows. This result is consistent with those of Norden and Weber (2010), who use different data from a different time period. 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 the 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 3, 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

(Insert Figure 3 about here)

Panel A1 of Figure 3 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

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).

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 3 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 $\{t+1, t+2, \dots, t+12\}$, 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

⁶ *Days Usage* and *Days Overdrafts* are the percentage of days in the preceding 12 months, whereas we present in Figure 3 the median number of days per month with a negative balance or days with overdrafts. In Section 3.5 we show that our results are robust when the explanatory variables enter our model in different functional forms.

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. 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 columns (2) and (4) show that checking or credit card account defaults are due to information from both credit products.

We continue with 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 R² 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 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

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.

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.

(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)). The increase in the adjusted R^2 from column (7) to column (8) shows that the prediction of credit card defaults can be improved if account activity from checking accounts is added to the model.

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

3.2. 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. 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 the account activity variables from

credit card accounts and not from checking accounts (and vice versa). This analysis is based on clean cross-product information that can be used for screening.

(Insert Table 4 about here)

Checking account activity is more predictive of defaults on credit card accounts than credit cards variables (adj. R^2 of 0.25 vs. 0.20; see Table 3 (7)).

3.3. Consistency of information

How much does account activity information from different credit products overlap? How consistent is this information? Figure 4 informs us about the consistency of PD estimates derived from different credit products. In Panel A, we present the PD estimates for checking accounts, where the estimates are based on the account activity of checking accounts on the one hand and credit card accounts on the other hand. Similarly, in Panel B we show PD estimates for credit card accounts. In both cases the correlation between the estimates derived from checking account versus credit card account variables is around 50% (0.558 for checking accounts and 0.540 for credit card accounts). If we analyze the pairwise correlations between individual account activity measures, we find that the correlations are at maximum 0.35 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.

(Insert Figure 4 about here)

3.4. 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 improve their decision making significantly 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.⁸ The absolute number of the type I error decreases from 5,096 to 3,408 (minus 1,688), corresponding to a reduction by 33%.⁹ 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 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.

In addition to informational synergies regarding the probability of default, we also examine variables that indicate consumer credit risk prior to default. We expect that informational synergies are most pronounced when consumers are close to default because at that moment they might be forced to deviate from their typical payment activity. 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 long time before default tend to be temporary and unsystematic, but they become increasingly systematic when we get closer to default. Thus, we expect cross-product activity to be less informative long time before default rather than closely before default. 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 (adj. $R^2 = 0.239$ instead of 0.234) and credit card accounts (adj. $R^2 = 0.320$ instead of 0.319). In line with the reasoning above, informational synergies

⁸ 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.

⁹ 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.

concerning overdrafts are more pronounced for both, checking accounts (McFadden adj. $R^2 = 0.384$ instead of 0.373) and credit card accounts (McFadden adj. $R^2 = 0.301$ instead of 0.286).¹⁰

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

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 : $CLU_{i,t+12} = Exposure_{i,t+12}/Limit_{i,t}$. Thus, following Jimenez, Lopez, and Saurina (2009), we could estimate the *CLU* as follows:

$$E(CLU_{i,t+12} | X_{i,t}) = X_{i,t} \beta \quad (2)$$

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. To show that the rating is a key input for modeling *CLU*, we present the median *CLU* for checking accounts between $\tau-24$

¹⁰ 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.

and $\tau+12$ by rating in Figure 5. The rating at $\tau+12$ discriminates between high and low CLU prior to default. The better the rating, the lower the CLU in the period between $\tau+24$ and τ . This indicates that *Rating* is informative about the CLU at default, as found by Agarwal et al. (2006). Unreported results indicate that the results for credit card accounts are similar.

(Insert Figure 5 about here)

We can also see in Figure 5, though, 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: $(CLU_{i,t+12} | Def_{i,t+12} = 1, X_{it}) \neq \mathbb{E}(CLU_{i,t+12} | Rating_{i,t}, X_{it})$. 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:

$$\mathbb{E}(CLU_{i,t+12} | Def_{i,t+12} = 1, X_{i,t}, Z_{i,t}) = X_{i,t}\beta + \delta_j \lambda(Z_{i,t}, \gamma) \quad (3)$$

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 credit line usage at default

To measure credit line usage, it is possible to consider positive values only (conditional on usage), or to consider positive and negative values (unconditional on usage). From the lender's perspective, it is important to estimate potential losses, which only occur if the credit line usage is positive. We take restricted values of *CLU* and compare them with the unrestricted estimation.

(Insert Table 7 about here)

The results in Table 7 are similar to those from Table 6 in terms of sign, significance, magnitude, and model fit because the correlation between restricted and unrestricted *CLU* at default is rather high, especially for credit card accounts (0.62 for checking accounts and 0.97 for credit card accounts). However, the sign for *Amplitude* in checking accounts becomes positive, meaning that accounts with higher *Amplitude* have a higher restricted *CLU* at default. This is because accounts with higher amplitude have higher inflow, along with a higher variance in their in- and outflows. Customers with high amplitude take less credit, but in the event of default, the high amplitude often leads to a high credit line usage and, thus, to higher values of the (restricted) *CLU* at default. This interpretation is supported by the finding that the coefficient of *Amplitude* is significantly negative in regressions with a restriction of *CLU* at default to *negative* values. Overall, our main results hold both for restricted and unrestricted *CLU* at default.

4.3. 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.¹¹ Table 8 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 8 about here)

For the model of customer 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 aggregated customer variables, the adjusted R^2 of 0.506 is

¹¹ 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.

higher/lower than for the models with cross-product information for checking accounts (0.182) and credit card accounts (0.579).

4.4. 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 default and *CLU* at default prediction models and repeat the analyses corresponding to Table 3 and 6. 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.

4.5. Alternative default prediction horizons

We examine whether our results are sensitive to the default prediction horizon. For this purpose, we repeat the analyses with time horizons of one month and three months. At a one month time horizon, we achieve a higher adjusted R^2 when predicting credit card defaults compared to a time horizon of one year (adj. R^2 of 0.308 versus 0.278; Table 3, column (8)). However, the effects in terms of sign, significance, and magnitude are largely unchanged. Thus, our results also hold for shorter prediction horizons.

4.6. 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. 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 3.1. Thus, our results on informational synergies also hold for non-linear functional forms. The detailed results are available on request from the authors.

4.7. 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.

4.8. 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. 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, information from checking accounts is more useful than that from credit cards. Activity measures from checking accounts indicate credit quality deteriorations earlier and more accurately than those from credit cards. 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.
- 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.
- Campbell, J., 2006. Household Finance. *Journal of Finance* 61, 1553-1604.
- 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.
- 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.
- 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.
- Lieberman, A., 2016. The value of a good credit reputation: Evidence from credit card renegotiations. *Journal of Financial Economics*, forthcoming.
- 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.
- Stango, V., Zinman, J., 2014. Limited and Varying Consumer Attention Evidence from Shocks to the Saliency of Bank Overdraft Fees. *Review of Financial Studies* 27, 990-1030.

- 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.
- 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.
- Zinman, J., 2014. Consumer Credit: Too Much or Too Little (or Just Right)? *Journal of Legal Studies* 43, S2 (Special Issue on Benefit-Cost Analysis of Financial Regulation), S209-S237.

Figure 1. Limit and negative balance of checking accounts

The long-dashed line presents negative balances (median, in euros) for defaulted checking accounts from 24 months prior to default to 12 months after default. The dashed line presents negative balances (default-weighted median, in euros) for non-defaulted checking accounts. The short-dashed and solid lines present limits (median, in euro/default-weighted median, in euros) for defaulted/non-defaulted checking accounts from 24 months prior to default.

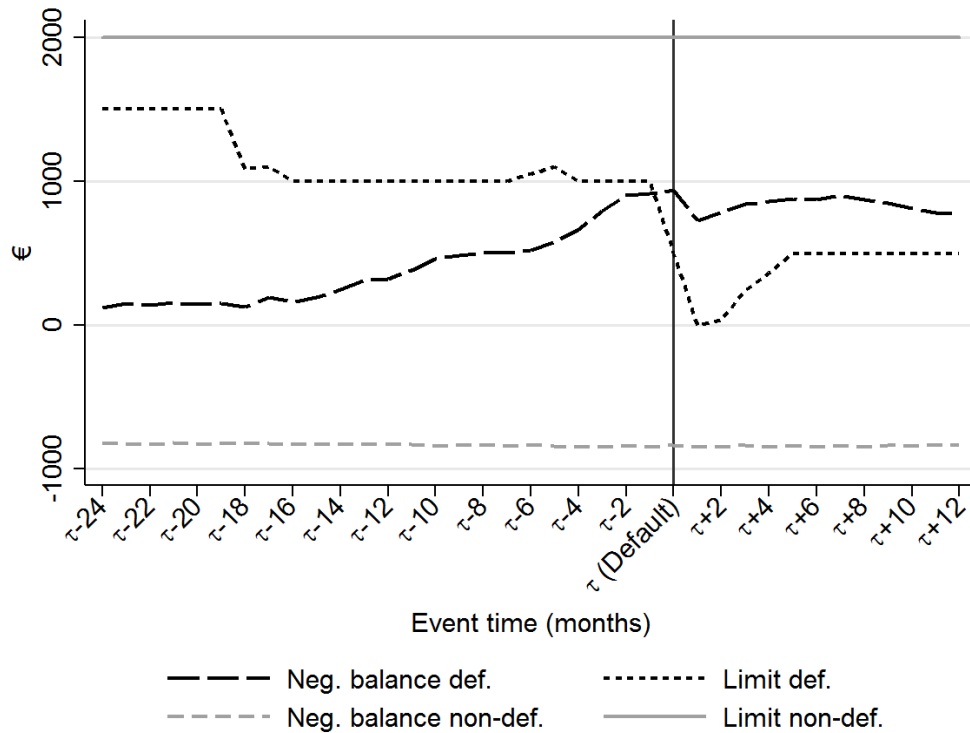


Figure 2. Cash inflows and outflows of checking accounts

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.

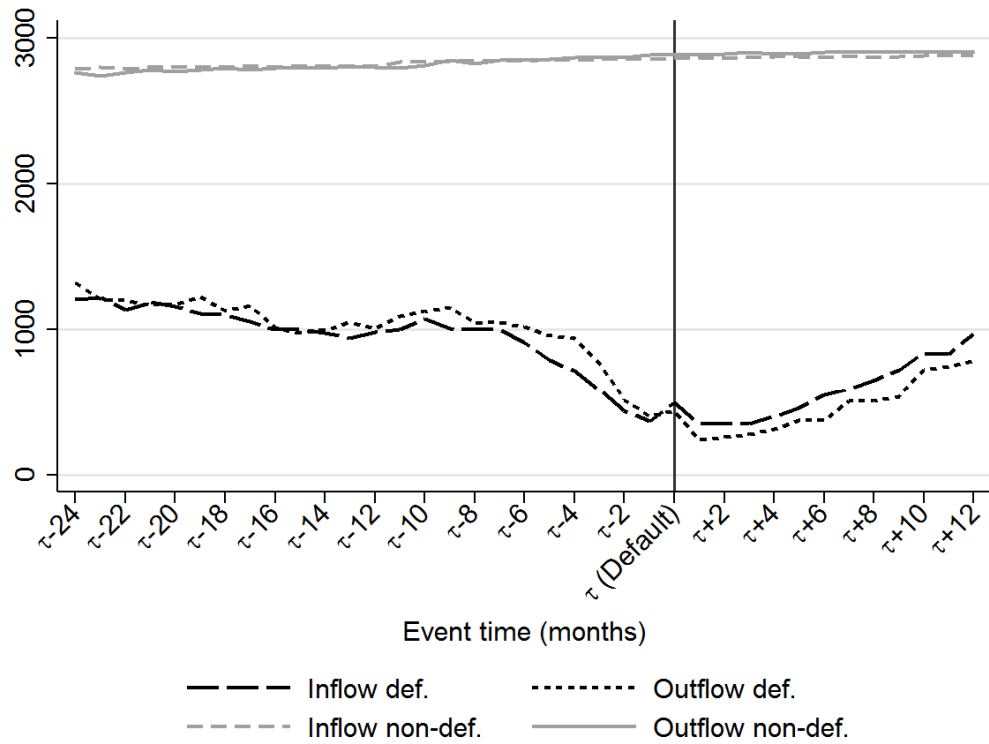
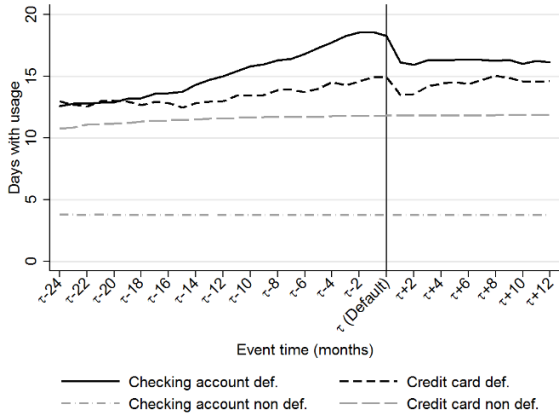


Figure 3. Days with credit line usage and overdrafts

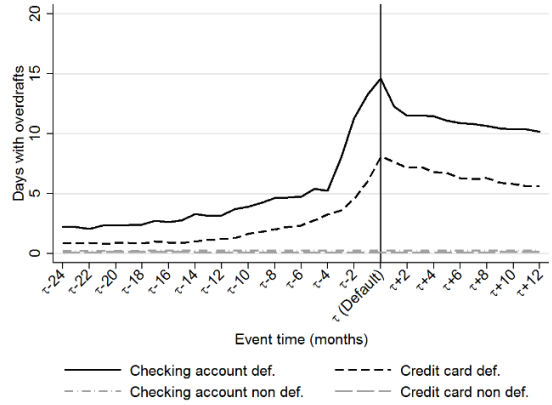
Panel A displays defaults of checking accounts, whereas Panel B displays defaults of credit cards. Days with usage are presented in Panel A1/B1 and days with overdrafts in Panel A2/B2. The solid and dash-dotted lines present days for defaulted (median) and non-defaulted (median, weighted by defaults)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

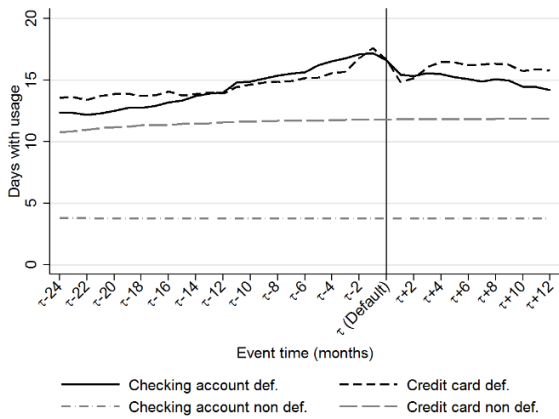


Panel A2: Days with overdrafts



Panel B: Default of credit cards at τ

Panel B1: Days with usage



Panel B2: Days with overdrafts

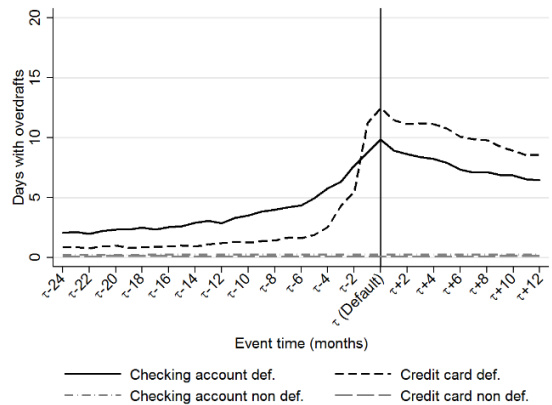
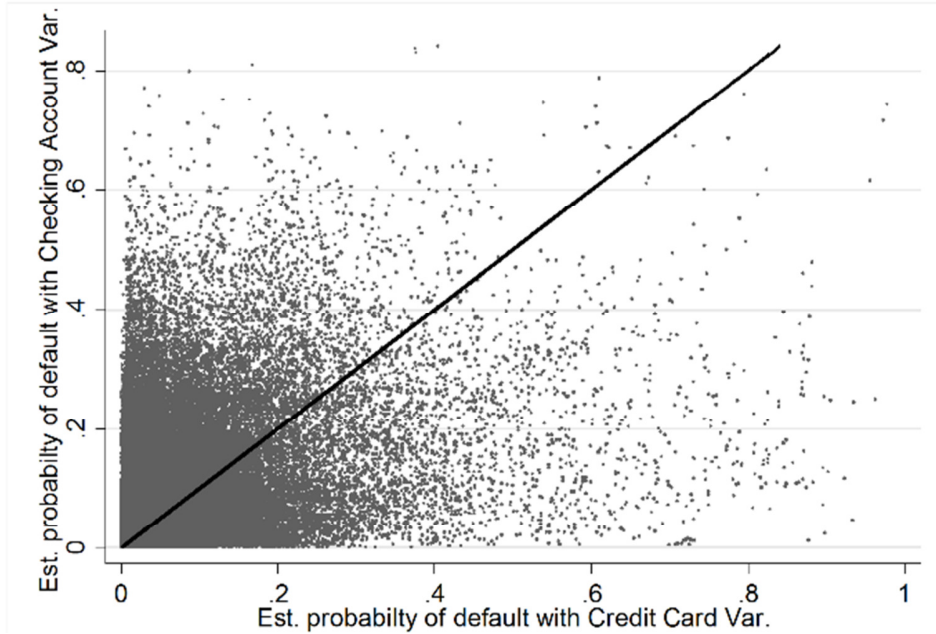


Figure 4. Consistency of information across credit products

Panel A shows estimates of the probability of default (*PD*) for checking accounts where the estimates are based on account activity of checking accounts versus credit cards. Panel B shows the corresponding *PD* estimates for defaults of credit card accounts.

Panel A: Default of checking accounts – PD estimates



Panel B: Default of credit cards – PD estimates

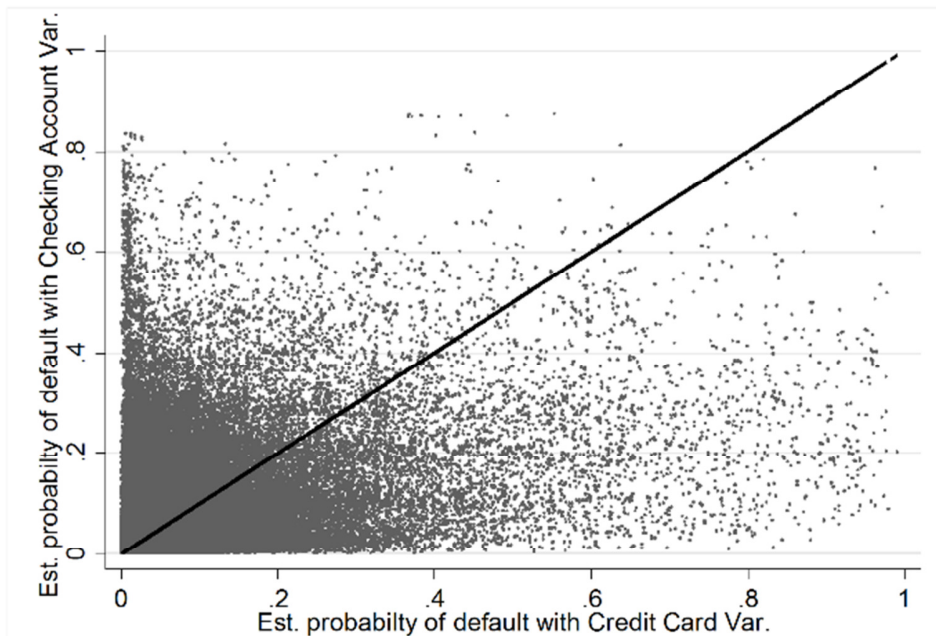


Figure 5. Credit line usage by credit rating

This figure displays the median of different rating categories for defaulted checking accounts from 24 months prior to default to 12 months after default. The five groups are formed at about 20%, 40%, 60%, and 80% quantiles with best ratings in category 1 and worst ratings in category 5.

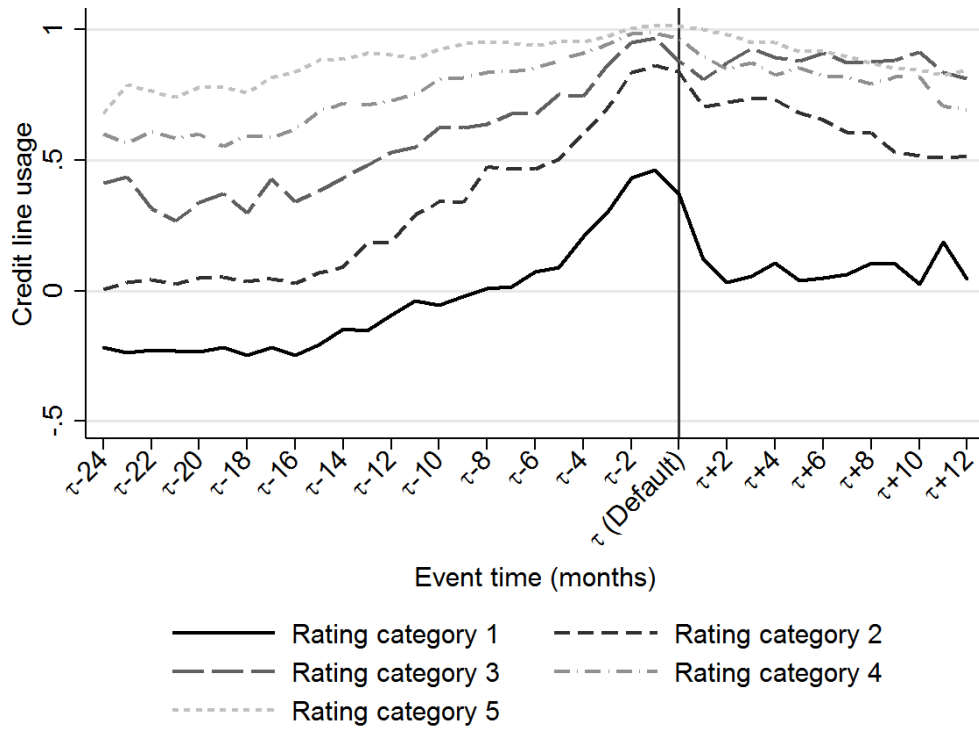


Table 1. Summary statistics

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

Variable	Checking Account		Credit Card		Checking Account Minus Credit Card	
	Mean	Median	Mean	Median	Pairwise Difference	<i>t</i> -stat.
Default Risk						
<i>Rating_t</i>	1.79	0.88	3.79	1.94	-1.939***	-428.25
Account Activity						
<i>Net Inflow/Lim_t</i>	0.02	0.00	0.00	0.00	0.0195***	19.18
<i>Amplitude_t</i>	1.80	1.03	0.39	0.24	1.407***	727.72
<i>Limit_t</i>	2437	2000	2581	2000	-143.7***	-163.29
<i>Bounced_t</i>	0.00	0.00	0.05	0.00	-0.0500***	-182.49
<i>Days Usage_t</i>	0.18	0.03	0.55	0.67	-0.367***	-1269.96
<i>Days Overdraft_t</i>	0.01	0.00	0.00	0.00	0.0070***	201.34
Relationship						
<i>Duration_t</i>	46.38	41.00	34.99	34.00	11.46***	616.02
Account Type						
<i>Full Payment_t</i>			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.Non-Default _{$t+12$} : Checking account		Default _{$t+12$} vs.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				
<i>Rating_t</i>	1.825*** (191.16)	1.079*** (129.85)	1.110*** (146.77)	1.729*** (197.96)
Account Activity				
<i>Net Inflow/Lim_t</i>	-0.0940*** (-9.97)	-0.0474*** (-16.07)	-0.0651*** (-24.21)	-0.0880*** (-10.23)
<i>Amplitude_t</i>	-0.814*** (-41.11)	0.128*** (29.70)	0.280*** (71.81)	-0.587*** (-32.51)
<i>Limit_t</i>	-0.716*** (-54.84)	-1.072*** (-83.33)	-1.063*** (-87.05)	-0.693*** (-58.19)
<i>Bounced_t</i>	0.00677*** (16.19)	0.336*** (120.29)	0.345*** (139.36)	0.00761*** (19.41)
<i>Days Usage_t</i>	0.456*** (218.64)	0.0355*** (14.03)	0.0644*** (27.93)	0.420*** (219.72)
<i>Days Overdraft_t</i>	0.109*** (309.61)	0.0464*** (216.29)	0.0482*** (304.21)	0.115*** (325.25)
Relationship				
<i>Duration_t</i>	-17.42*** (-71.83)	-12.22*** (-77.06)	-12.27*** (-84.87)	-16.51*** (-73.23)
<i>Prev. Def_t</i>		0.0210*** (51.18)		0.0419*** (93.13)
Account Type				
<i>Full Payment_t</i>		-0.263*** (-135.71)	-0.231*** (-130.24)	
Number of observations	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								
<i>Rating_t</i>	0.4626*** (52.184)		0.2215*** (19.633)	0.1703*** (14.915)				0.1314*** (13.178)
Account Activity								
<i>Net Inflow/Lim_t</i>		-0.0479*** (-14.394)	-0.0561*** (-15.847)	-0.0616*** (-17.407)				-0.0475*** (-18.548)
<i>Amplitude_t</i>		-0.0687*** (-7.902)	-0.0645*** (-7.470)	-0.0614*** (-7.348)				-0.0360*** (-6.242)
<i>Limit_t</i>		-0.1289*** (-12.371)	-0.1322*** (-12.231)	-0.0842*** (-6.678)				-0.0750*** (-6.681)
<i>Bounced_t</i>		0.2471* (2.489)	0.2288* (2.249)	0.1853† (1.779)				0.1649 (1.571)
<i>Days Usage_t</i>		1.2349*** (38.682)	0.8366*** (22.237)	0.9115*** (22.089)				0.7657*** (20.266)
<i>Days Overdraft_t</i>		2.1567*** (21.982)	1.8995*** (19.126)	1.4664*** (13.412)				1.1692*** (11.174)
Relationship								
<i>Duration_t</i>		-0.0084*** (-14.432)	-0.0055*** (-9.420)	-0.0027*** (-4.113)				-0.0018** (-3.006)
<i>Prev. Def_t</i>								-0.0191 (-0.236)
CREDIT CARD VARIABLES								
Default Risk								
<i>Rating_t</i>				0.1284*** (10.279)	0.3947*** (40.648)		0.2814*** (23.018)	0.1626*** (12.498)
Account Activity								
<i>Net Inflow/Lim_t</i>				-0.1078*** (-17.330)		-0.1226*** (-28.027)	-0.1493*** (-26.392)	-0.1418*** (-22.740)
<i>Amplitude_t</i>				0.0811*** (8.698)		0.1323*** (19.145)	0.1518*** (19.972)	0.1765*** (23.282)
<i>Limit_t</i>				-0.0241† (-1.802)		-0.1057*** (-11.646)	-0.0667*** (-7.303)	0.0053 (0.468)
<i>Bounced_t</i>				0.0343* (2.323)		0.1628*** (7.113)	0.1402*** (6.692)	0.0554*** (3.594)
<i>Days Usage_t</i>				-0.5843*** (-13.448)		0.1516*** (5.276)	-0.3906*** (-10.949)	-0.4539*** (-11.408)
<i>Days Overdraft_t</i>				0.8898*** (5.346)		4.7102*** (22.811)	4.1120*** (20.775)	2.2709*** (12.012)
Relationship								
<i>Duration_t</i>				-0.0068*** (-6.929)		-0.0120*** (-17.658)	-0.0104*** (-15.669)	-0.0080*** (-8.839)
<i>Prev. Def_t</i>				-0.1909† (-1.651)				
Account Type								
<i>Full Payment_t</i>				-0.1120*** (-3.627)		-0.5604*** (-22.549)	-0.2670*** (-9.105)	-0.0398 (-1.224)
<i>Constant</i>	-2.2031*** (-47.816)	-1.9650*** (-24.078)	-1.9854*** (-23.254)	-1.9658*** (-23.286)	-2.5644*** (-60.191)	-1.2397*** (-15.919)	-1.9346*** (-22.661)	-2.1022*** (-23.216)
<i>Customer controls</i>	No	Yes	Yes	Yes	No	Yes	Yes	Yes
<i>Year-month FE</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	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 Independent Variables	Default _{$t+12$} of Checking Account	
	Credit Card Account (1)	Checking Account (2)
Default Risk		
<i>Rating_t</i>	0.2715*** (24.155)	0.1718*** (16.769)
Account Activity		
<i>Net Inflow/Lim_t</i>	-0.1144*** (-21.253)	-0.0426*** (-15.034)
<i>Amplitude_t</i>	0.0387*** (3.950)	-0.0153** (-2.873)
<i>Limit_t</i>	-0.1012*** (-9.525)	-0.0981*** (-9.795)
<i>Bounced_t</i>	0.1225*** (7.837)	0.1986* (1.994)
<i>Days Usage_t</i>	-0.4614*** (-12.775)	0.7627*** (21.797)
<i>Days Overdraft_t</i>	2.1146*** (13.249)	1.7604*** (18.710)
Relationship		
<i>Duration_t</i>	-0.0094*** (-13.951)	-0.0052*** (-9.849)
<i>Prev. Def_t</i>	0.1383 (1.426)	0.1251† (1.799)
Account Type		
<i>Full Payment_t</i>	-0.3769*** (-14.280)	-0.1618*** (-5.984)
Constant	-1.8416*** (-21.021)	-1.6782*** (-20.378)
<i>Customer controls</i>	Yes	Yes
<i>Year-month FE</i>	Yes	Yes
Number of observations	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 *PDs* 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	Measure	Information Source		
		(1) Checking Account	(2) Credit Card Account	(3) Both Accounts
<i>Default_{t+12} of checking account</i>	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
<i>Default_{t+12} of credit card account</i>	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. CLU 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				
<i>Net Inflow/Lim_t</i>	-0.0797*** (-6.939)	-0.0772*** (-6.752)		0.0391*** (5.409)
<i>Amplitude_t</i>	-0.0269*** (-4.673)	-0.0431*** (-7.311)		-0.0670*** (-17.159)
<i>Limit_t</i>	-0.0448*** (-5.600)	-0.0565*** (-5.621)		0.0989*** (14.023)
<i>Bounced_t</i>	0.0726 (1.102)	0.0842 (1.282)		0.0093 (0.205)
<i>Days Usage_t</i>	0.7438*** (19.015)	0.7127*** (17.905)		0.0127 (0.480)
<i>Days Overdraft_t</i>	0.4762*** (6.842)	0.2343** (3.049)		0.3970*** (7.242)
Relationship				
<i>Duration_t</i>	-0.0037*** (-8.520)	-0.0022*** (-4.265)		0.0004 (1.149)
<i>Prev. Def_t</i>				0.2033*** (5.685)
CREDIT CARD VARIABLES				
Account Activity				
<i>Net Inflow/Lim_t</i>		-0.0384 (-1.582)	-0.2376*** (-15.200)	-0.2122*** (-13.784)
<i>Amplitude_t</i>		0.1228*** (9.919)	0.9247*** (122.519)	0.9308*** (123.307)
<i>Limit_t</i>		0.0074 (0.655)	-0.1519*** (-23.634)	-0.2062*** (-26.502)
<i>Bounced_t</i>		0.0037 (0.384)	-0.0077 (-1.077)	-0.0219** (-3.101)
<i>Days Usage_t</i>		0.2206*** (5.636)	0.0656* (2.087)	0.0349 (1.109)
<i>Days Overdraft_t</i>		0.0502 (0.406)	0.9162*** (10.823)	0.5432*** (6.080)
Relationship				
<i>Duration_t</i>		-0.0026** (-3.270)	-0.0027*** (-5.785)	-0.0039*** (-6.930)
<i>Prev. Def_t</i>		0.1705* (2.271)		
Account Type				
<i>Full Payment_t</i>		-0.0766** (-2.802)	-0.5989*** (-25.640)	-0.5521*** (-24.383)
<i>Constant</i>	2.3201*** (13.925)	2.1924*** (13.107)	2.2606*** (18.071)	2.2203*** (17.959)
<i>Lambda: Default</i>	-0.7217*** (-26.960)	-0.6953*** (-25.688)	-0.5551*** (-18.083)	-0.5242*** (-17.316)
<i>Customer Controls</i>	Yes	Yes	Yes	Yes
<i>Year-month FE</i>	Yes	Yes	Yes	Yes
Number of observations	19,149	19,149	22,964	22,964
Adj. R ²	0.173	0.182	0.562	0.579

Table 7. CLU at default restricted to positive values

This table reports estimates of the restricted credit line usage at default, which is defined as *restricted CLU at Default* $_{t+12} = \max(\text{CLU at Default}_{t+12}, 0)$. The estimates refer to the potential loss from a bank's perspective as the loss equals zero if the account balance is positive at time of default. Analogous to the results presented in Table 6, the estimates are based on a Heckman selection model. 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.)

	Restricted CLU at Default $_{t+12}$: Checking Account		Restricted CLU at Default $_{t+12}$: Credit Card	
	(1)	(2)	(3)	(4)
CHECKING ACCOUNT VARIABLES				
Account Activity				
<i>Net Inflow/Lim_t</i>	-0.0631*** (-11.364)	-0.0608*** (-11.036)		0.0362*** (5.321)
<i>Amplitude_t</i>	0.0783*** (28.156)	0.0697*** (24.533)		-0.0559*** (-15.239)
<i>Limit_t</i>	-0.0528*** (-13.622)	-0.0642*** (-13.238)		0.0819*** (12.353)
<i>Bounced_t</i>	-0.0169 (-0.531)	-0.0114 (-0.359)		0.0181 (0.427)
<i>Days Usage_t</i>	0.1716*** (9.062)	0.1461*** (7.614)		-0.0345 (-1.385)
<i>Days Overdraft_t</i>	0.7013*** (20.786)	0.5308*** (14.307)		0.4998*** (9.703)
Relationship				
<i>Duration_t</i>	-0.0004† (-1.684)	0.0008** (3.003)		-0.0007* (-2.227)
<i>Prev. Def_t</i>				0.1682*** (5.006)
Account Activity				
<i>Net Inflow/Lim_t</i>		-0.0527*** (-4.490)	-0.1865*** (-12.735)	-0.1632*** (-11.289)
<i>Amplitude_t</i>		0.0537*** (8.991)	0.9860*** (139.468)	0.9899*** (139.586)
<i>Limit_t</i>		0.0104† (1.916)	-0.1189*** (-19.749)	-0.1628*** (-22.268)
<i>Bounced_t</i>		0.0049 (1.047)	-0.0092 (-1.373)	-0.0229*** (-3.456)
<i>Days Usage_t</i>		0.1765*** (9.358)	0.0161 (0.547)	-0.0043 (-0.145)
<i>Days Overdraft_t</i>		0.0161 (0.270)	1.1643*** (14.681)	0.7640*** (9.101)
Relationship				
<i>Duration_t</i>		-0.0023*** (-5.870)	-0.0013** (-2.923)	-0.0013* (-2.474)
<i>Prev. Def_t</i>		0.1504*** (4.147)		
Account Type				
<i>Full Payment_t</i>		-0.0656*** (-4.975)	-0.8607*** (-39.343)	-0.8189*** (-37.633)
<i>Constant</i>	1.6152*** (20.010)	1.5424*** (19.110)	1.4379*** (12.268)	1.4024*** (12.072)
<i>Lambda: Default</i>	-0.2762*** (-21.167)	-0.2573*** (-19.594)	-0.1531*** (-5.323)	-0.1236*** (-4.346)
<i>Customer Controls</i>	Yes	Yes	Yes	Yes
<i>Year-month FE</i>	Yes	Yes	Yes	Yes
Number of observations	19,149	19,149	22,964	22,964
Adj. R ²	0.162	0.177	0.586	0.599

Table 8. 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 Independent Variables	Default _{t+12}			CLU at Default _{t+12}		
	Customer	Checking Acc. (part 1)	Credit Card Acc. (part 2)	Customer	Checking Acc. (part 1)	Credit Card Acc. (part 2)
	(1)	(2)	(3)	(4)	(5)	(6)
Default Risk						
<i>Rating_t</i>	0.2870*** (16.681)	0.1266*** (9.963)	0.1129*** (6.920)			
Account Activity						
<i>Net Inflow/Lim_t</i>	-0.0650*** (-12.227)	-0.0304*** (-11.037)	-0.0604*** (-6.333)	-0.0235 (-1.397)	-0.0351*** (-4.012)	-0.0713** (-3.010)
<i>Amplitude_t</i>	-0.0125 (-1.264)	-0.0068 (-1.496)	-0.0104 (-0.720)	-0.3231*** (-64.016)	-0.0346*** (-7.496)	0.0721*** (4.776)
<i>Limit_t</i>	-0.0332*** (-5.395)	-0.0581*** (-4.457)	-0.0024 (-0.184)	-0.0000*** (-6.540)	-0.0000 (-0.553)	-0.0279** (-2.914)
<i>Bounced_t</i>	0.0435 (1.119)	0.1571 (1.481)	0.0272 (1.355)	0.0230 (1.450)	0.0735 (1.526)	-0.0049 (-0.518)
<i>Days Usage_t</i>	-0.1101* (-2.311)	0.4570*** (9.817)	-0.2879** (-5.866)	0.3409*** (9.420)	0.5208*** (14.287)	0.1273*** (3.756)
<i>Days Overdraft_t</i>	2.5958*** (22.454)	1.8337*** (14.296)	1.0824*** (4.924)	0.5638*** (9.949)	-0.0974 (-1.376)	0.2214† (1.665)
Relationship						
<i>Duration_t</i>	-0.0020*** (-3.449)	-0.0004 (-0.529)	-0.0050** (-4.761)	-0.0029*** (-8.563)	-0.0007 (-1.607)	-0.0021** (-3.031)
Account Type						
<i>No Full_t</i>	-0.1517*** (-3.887)		-0.1750*** (-4.561)	0.0122 (0.541)		-0.0360 (-1.444)
Constant	-2.4611*** (-22.179)		-2.3155*** (-20.966)	2.5455*** (16.645)	3.1527*** (18.367)	
<i>Lambda: Default</i>				-0.6631*** (-22.396)	-1.0522*** (-31.304)	
<i>Customer controls</i>	Yes		Yes	Yes		Yes
<i>Year-month FE</i>	Yes		Yes	Yes		Yes
Number of observations	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