C.I.R.P.É.E

Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **11-19**

Is there Any Dependence Between Consumer Credit Line Utilization and Default Probability on a Term Loan ? Evidence from Bank-Level Data

Anne-Sophie Bergerès
Philippe d'Astous
Georges Dionne

Juillet/July 2011

Bergerès: Caisse de dépôt et placement du Québec
D'Astous: HEC Montréal, Department of Finance
Dionne: HEC Montréal, Department of Finance, CIRRELT and CIRPÉE

Financial support by SSHRCC and CGI is acknowledged. The authors thank Hind Diboune and Denise Desjardins for their contribution to this research, and Benoit Dostie for his very useful comments.

Abstract:
Recent studies of credit lines suggest a positive relationship between exposure at default and default probability on the line. In this paper, we consider another important dependence between two consumers' financial instruments by investigating the relationship between credit line utilization and default probability on a term loan. We model the variation of both financial instruments endogenously in a simultaneous equations system. We find strong evidence for a positive relationship between the two variables: individuals in the default state use their credit line about 59% more often, while credit line utilization has a positive marginal effect of 46% on loan default probability. Our results suggest that banks should monitor both financial instruments simultaneously.

Keywords: Consumer finance, consumer risk management, credit line, term loan, default probability, ability to pay, endogeneity, simultaneous equations

JEL Classification: D12, D14, G01, G21, G33

Is there any dependence between consumer credit line utilization and default probability on a term loan? Evidence from bank-level data

Consumer finance has been neglected for many years in the economics and finance literature (Campbell, 2006; Campbell et al. 2010; Tufano, 2009). The recent financial crisis clearly showed that consumer financial services are not well managed and deserve further study. The Canadian consumer credit line market expanded by 133% from 1999 to 2005. Its value currently far surpasses that of credit card debt and personal loans combined. In the United States, outstanding total revolving consumer credit, including credit card balances, represents 33% of total consumer debt in 2011. The presence of these relatively new and popular credit instruments should have a significant impact on consumer financial distress.

Recent studies of credit lines suggest a positive relationship between exposure at default, as measured by credit line utilization, and default probability on the credit line. For example, looking at Home Equity Lines of Credit (HELOC), Agarwal et al. (2006a) find that consumers with an *ex ante* higher expectation of credit deterioration use a smaller part of their line of credit at origination. Their findings also suggest that the drawing behavior after origination is inversely related to changes in the borrower's credit quality. This is reflected by higher credit line utilization for borrowers who end up defaulting on the line, and implies that the exposure at default should not be treated independently from the default probability on the credit line. Such findings show a quality reduction in the credit line market; consumers with degrading credit quality are the ones who use their credit lines the most. Studying corporate lines of credit, Jiménez et al. (2009) find similar results and evidence of a positive relationship between the probability of default and the amount drawn by a firm on its credit line. This relationship is strongest near the default event.

Agarwal et al. (2006b) study the differences between the actions taken on Home Equity Loans and HELOC. They find that, overall, the risk profile of these two financial instruments is heterogeneous. Their sample shows that borrowers using a term loan have a lower credit rating and contract a significantly smaller amount than individuals using a line of credit. They also assert that the probability of prepayment is higher on lines of credit than on term loans. This phenomenon could be explained partly by interest rate variations on credit lines, which push consumers to take advantage of lower interest rate periods. Their analysis also shows that the probability of default on credit lines is lower than that on term loans, the latter being more sensitive to changes in the value of the house and twice as sensitive to changes in interest rates.

Suffi (2009) addresses the determinants of credit line usage as an alternative to cash for liquidity management. He argues that the financial covenants imposed by banks on credit line accounts make this instrument more *contingent* in practice than the usual assumption of full commitment found in the theoretical literature. He finds that internal cash flows are a strong determinant of the probability that a firm uses a line of credit, the effect being most significant when the firm has a high probability of financial distress. Norden and Weber (2010) investigate checking accounts and the information they supply regarding credit line utilization for both firms and consumers. Using a German data set, they show that banks tend to monitor clients' credit line utilization through information on the use of their checking accounts.

Calem et al. (2011) perform panel data estimation and risk hazard analysis on HELOC borrower data to test for an adverse selection effect in the credit line market. They find that contrary to neoclassical theories of consumption, tightening economic conditions are associated with a negative shift in the credit quality of the borrower pool. Further analysis shows that drawdowns when the anticipation of future income is low are associated with a higher probability of borrower delinquency. They conclude that such evidence is consistent with an adverse selection problem in the market.

To our knowledge, no empirical research has focused on the relationship between credit line utilization and default probability of consumers on another credit instrument. Strahan (1999) argues that liquidity needs and default probability are interrelated, the former being higher when financing from other sources is harder, usually occurring when the borrower's default risk is above average. He also notes that a typical deal between a firm and its bank should include both a term loan and a credit line. Combining both short-term and long-term borrowing would help an institution divide monitoring costs between these two loans. It can hence gain a better understanding of the borrower's credit quality. Information on both credit instruments should be combined so that credit line drawdown behavior can be connected to the default risk assessment of a term loan and vice versa. We consequently investigate the statistical relationship between credit line utilization and default probability on a term loan.

The dataset for this study comes from a Canadian bank, and is composed of consumers who use both a line of credit and a term loan. The two variables of interest for each borrower are thus the percentage of a credit line drawn down relative to the commitment amount and the default state of the term loan. We build a simultaneous equation model in which credit line utilization is modeled by instrumented Tobit and default probability on the loan is modeled by instrumented Probit. Our assumptions are that both variables are endogenous and that their relationship is positive. First, borrowers who use the liquidity of their credit line to spend more than the amount of debt corresponding to their term loan credit risk, send a signal of bad liquidity management. Of course, this behavior is limited to the amount authorized by the bank on the line, but it may ultimately degenerate in a higher default probability on the term loan. Conversely, borrowers who are under financial distress and have a high default probability on their term loan can rely on the liquidity of their credit line to make payments on their loan instead of renegotiating their debt. Thus, a high default probability on the loan could affect positively credit line utilization. To our knowledge, this bivariate relationship has not been explored empirically; the major contribution of this study is to establish its statistical significance.

Our results show that the EAD, as modeled by credit line utilization, is endogenous in the default probability equation of the term loan and that the default probability variable is endogenous in the credit line utilization equation. We use valid instruments to overcome the endogeneity problem and to successfully estimate both equations. We find a positive relationship between the two variables and we also verify that the number of active credit lines plays a significant role in the borrower's probability of default on a term loan. Moreover, we propose a new measure of debt holder ability to pay back the loan that complements the credit score variable to explain the default probability. With the proposed definition, the borrower's ability to pay significantly affects credit line utilization and those with the worst ability to pay are more likely to default on the term loan.

The rest of the article is divided as follows. Section 1 presents consumer lines of credit and statistics from the Canadian and US markets. Section 2 summarizes our hypotheses and the econometric methodology used. We present the data, the variables, and their descriptive statistics in Section 3. Section 4 describes the major results of this study, and Section 5 concludes the paper.

1. Consumer lines of credit

1.1 General definition

In its most general definition, a line of credit is the maximum amount a bank has committed to grant a borrower for a predetermined period of time. Borrowers who are accepted can use funds up to the maximal amount authorized on the credit line. Two types of lines of credit are usually available to consumers: credit cards and revolving lines of credit (credit lines hereafter). One of the main differences between these two financial instruments is that credit cards require no collateral from the client, making them unsecured loans for the institution, whereas credit lines are sometimes secured by collateral. Some credit lines are secured by the equity owned on a house; they are called Home Equity Lines of Credit (HELOC). However, not every credit line is backed by collateral. For small consumer lines of credit, banks usually require a smaller asset as collateral or no collateral at all. Such lines make up our dataset. This feature is pertinent to the analysis.

Dey (2005) studies consumers' decisions regarding incurring debt through a credit card rather than through a personal line of credit. He finds that, theoretically, the collateral required on a line of credit can make consumers reluctant to use the full loan amount authorized by the bank. Depending on their utility function and corresponding risk aversion, borrowers may prefer to use an unsecured credit card, even if it bears a higher interest rate. This explains why some people carry a positive balance on their credit cards even though their credit line is not maxed out.

For consumers, a line of credit usually has no maturity; the bank renews the account periodically as long as it remains active. The interest charges are added to the amount drawn down by the borrower and no minimum monthly payment is required. If a client reaches the borrowing limit on the line, the institution can renegotiate the contract into a term loan on which interest and capital must be repaid monthly. Although there are many similarities between personal and corporate lines of credit, some differences are worth noting. Empirical evidence shows that firms often refinance their lines of credit as term loans before the maximum amount authorized is used, possibly to preserve their short-term ability to borrow (Agarwal et al., 2004). This happens because firms pay a percentage fee according to the unused portion of their line, whereas consumers do not. Only corporate credit lines are subject to material adverse change (MAC) clauses (Agarwal et al., 2006a).

1.2 Statistics on credit lines

Of all household financial products, lines of credit have experienced the greatest growth in recent years. The Statistics Canada 2005 Survey of Financial Security reports that credit lines accounted for 9.0% of total Canadian household debt in 2005, versus 5.7% in 1999. This increase represents a change of 133.2% in constant 2005 dollars. At that time, the personal credit line was the most important type of debt after the real estate mortgage, which accounted for 75.3% of total Canadian household debt. According to this study, the total amount of debt on credit cards and installment credit[1] has risen by 58.4% (to 3.4% of total family unit debt in 2005) during the same period, the second largest increase after that of credit lines.

[1] From Statistics Canada: “Instalment debt is the total amount owing on deferred payment or instalment plans where the purchased item is to be paid for over a period of time.”

These numbers are consistent with an upward trend in consumer lines of credit, which started in the 1990s. Concerning the assets of chartered banks used for consumer credit, the proportion allocated to personal loans did not increase as much as that allocated to lines of credit and credit cards. Figure 1 presents the changing balances of these three types of assets for Canadian chartered banks.[2] Since 1997, the total value of lines of credit has undergone a dramatic increase, far exceeding the total value of personal loans. In 2010, the portion of assets dedicated to personal lines of credit was the largest, almost double the value of credit card and personal loans combined. This suggests that the proportion of credit line debt used by households is now even higher than that reported by the Statistics Canada 2005 study.

We also observe an important rise in personal revolving credit utilization in the US since the mid-1990s. The latest numbers provided by the Federal Reserve[3] (FBR: G.19) show that in January 2011, outstanding revolving credit for consumers totaled $807.5 billion, representing 33% of the total outstanding consumer debt. This number represents all types of revolving consumer credit. The significant increase in revolving credit usage warrants the inclusion of these instruments in economic and financial models to fully understand the evolution in consumption decisions of households (Dey, 2005).

2. Hypotheses and methodology

Research on credit line utilization has usually focused on the relationship between borrowers' drawdown behavior and default probability on a credit line. The literature generally suggests a positive relationship between drawdown and the default state; lines that end up in default are

[2] Source: Series v36867, v36868 and v36869 of CANSIM.
[3] Federal Reserve Statistical Release G.19, available online: http://www.federalreserve.gov/releases/g19/current/g19.htm.

usually the ones that were used the most.[4] This has significant implications for the calculation of the exposure at default on the line since the relationship with the default probability must be taken into account.[5]

The main hypotheses of our study concern the relationship between credit line utilization and default probability on a term loan. Because consumers have a finite asset allocation, they must choose among various consumption and saving options, possibly resulting in a joint determination of their financial decisions. Borrowers, who use the liquidity provided by their credit line to spend more than the amount of debt corresponding to their credit risk, may increase their default probability. The effect of such behavior should ultimately send a signal to the bank of a higher default probability on the term loan because, once the credit line is fully used, the bank will require additional payments. In this sense, the relationship between credit utilization and default probability on the term loan would be positive. Conversely, borrowers who are under financial distress and who have a higher probability of default on the term loan can rely on the liquidity of the credit line to make payments on their term loan. Default probability on the loan would thus affect credit line utilization. We assume that both variables are endogenous and that the relationship between them is positive.

2.1 Specification of the econometric model

To consider the endogenous relations of the model, we build a simultaneous equation model in which one equation represents the default probability on the loan and the other represents the percentage use of the credit line. We establish the default state on the loan with a variable provided by the financial institution. A loan classified as "bad" in the database is presumed to be

[4] See Agarwal et al., 2006a, Jiménez et al., 2009b, and Norden and Weber, 2010.
[5] See, for example, Jiménez et al. (2009a).

in default. This rating is the worst the institution can attribute to a customer. Credit line utilization is calculated as the ratio of the amount drawn down on the line on the extraction date divided by the maximum amount allowed.

Neither variable can be assumed to follow the normal distribution, because the default probability is a dichotomous variable and the credit line utilization is bounded by zero and one (respectively a usage rate of 0% and of 100% of the line). With these specifications, the former variable is modeled by an instrumental Probit equation (1) and the latter variable by an instrumental Tobit equation (2). The model can be represented by the following system:

$$y_1 = \alpha_1 + \beta_1 y_2 + \boldsymbol{z}_{(\boldsymbol{1})}\, \delta_{(1)} + u_1 \tag{1}$$

$$y_2 = \alpha_2 + \beta_2 y_1 + \boldsymbol{z}_{(\boldsymbol{2})} \delta_{(2)} + u_2. \tag{2}$$

In such a system, the two endogenous variables are y_1, the default probability on the term loan, and y_2, the percentage of the line drawn down. The vectors $\boldsymbol{z}_{(\boldsymbol{i})}$ of explanatory variables and $\delta_{(i)}$ of parameters are used for the exogenous factors of the model. Variables u_1 and u_2 are respectively the random error terms for equations (1) and (2).

We estimate three sets of results for this system (see Table 7). The first set of results is estimated with Newey's Two-Step Efficient Estimator (Newey, 1987), a limited information procedure. Standard errors for this estimator are based on Amemiya's (1978, 1979) derivations of the efficient variance-covariance matrices. This procedure allows us to validate the chosen instruments for each endogenous variable by an Amemiya-Lee-Newey overidentification test (Lee, 1992). The second set is estimated from a full information maximum likelihood procedure and provides a simple Wald test of the exogeneity of an explanatory variable. It allows us to test the endogeneity of the default probability and the credit line utilization variables. For robustness, the last set of results is a joint two-step estimation of OLS and Probit models, known as Two-Step Probit Least Squares (2SPLS). All procedures are instrumented for the endogenous variables of the system. The second stage results of these three estimations are qualitatively identical and are presented in Table 7. We derive the marginal effects in Table 8 from the full information maximum likelihood procedure, which is the preferred set of results.

3. Data, variables and descriptive statistics

The data consist of observations with information ranging from January 1, 2005 to December 31, 2007. We first cleaned the database to retain the 16,370 clients who have been granted both a credit line and a term loan, and eliminate duplicates and observations for which information is missing. The final sample is a cross-sectional dataset of 14,827 observations on December 31, 2007 with monthly individual information back to January, 1, 2005, or to their first contract with the bank after that date. From this dataset, the bank considers 160 clients to be in default on the term loan, a proportion of 1.09% of the database. The extraction date paints a portrait of Canadian households when the financial crisis was already well under way, although the crisis did not affect the Canadian banking system significantly. The data cover information the bank acquired when it originally accepted the customers, along with updated information on the borrowers' credit quality and specific information. The mean (median) amount authorized on the credit line is \$5,051.37 (\$5,000) and the mean (median) value of the term loan is \$14,151.15 (\$12,878.79). Such financial instruments are not necessarily tied to the equity on the borrower's home (if any) and should be considered consumption credit. Although the institutions grant only one credit line per customer, we introduce a variable indicating how many more credit lines are active for the customer, possibly at other institutions. No other information on the additional lines is available.

Table 1 presents the variables used in the analysis; they can be divided into three categories:

(a) Variables that affect both the credit line utilization and the default probability on the loan

(b) Variables used as instruments for the credit line utilization

(c) Variables used as instruments for the default probability

3.1 Explanatory variables

To control for the characteristics of borrowers in the sample, we include information about their sex, credit score, seniority at the bank, working situation, number of dependents, age, and ability-to-pay ratio. The ability-to-pay variable is necessary to take into account the income and expenditures of borrowers, which can have economic significance beyond the information provided by the credit score. The ratio that we develop measures the financial constraint of the borrower, whereas the bank credit score we use is more of a credit delinquency measure than a capacity-to-pay measure. The definition proposed for the ability-to-pay variable is the ratio of the monthly payment on the loan to the borrower's capacity to pay. The income and expenditures are expressed in monthly amounts and the tangible assets are divided by the remaining months on the loan contract to fit the time frame. A negative value for this ratio implies that the expenditures are higher than the income and tangible assets added, while a value of more than one implies a monthly payment higher than the customer's ability to pay. More prudent borrowers have a small but positive ratio. The ability-to-pay ratio is expressed as:

$$\text{Ability-to-pay ratio} = \frac{Monthly\ payment\ on\ the\ loan}{Income + tangible\ assets - expenditures} \tag{3}$$

Because the dataset contains limited information on the customers' expenditures, we estimate these expenses based on a larger dataset compiled by Statistics Canada. By taking into account variables such as sex, age, income, type of housing, and dependents in an ordinary least squares estimation, we faithfully associate expenditures with each borrower in our sample and construct the ability-to-pay ratio. Results of this estimation are presented in Table 2. The ability-to-pay ratio has been segmented in different categories to better determine at which point it becomes economically significant. This segmentation is presented in Table 3. Category 5 is taken as the reference category in the next estimations.

3.2 Instrumental variables

To overcome the endogeneity bias of the model, we use significant variables to instrument the default probability and the credit line utilization ratio. As in all studies using instrumental variables, the dataset dictates the choice of instruments. The instrument sets have been validated by an Amemiya-Lee-Newey overidentification test.

For the default probability, the instruments used are the remainder term of the loan expressed in months, the number of active credit lines, and a dichotomous variable indicating whether the loan is secured by collateral. While these variables should be correlated to the default probability conditionally on all other exogenous variables of the model, neither instrumental variable should be correlated to credit line utilization. For the credit line utilization instruments, we use the amount drawn on the line and a dichotomous variable indicating the presence or absence of collateral on the credit line. In this case, the instruments should be correlated to the credit line utilization conditionally on all other exogenous variables of the model and should not be correlated to the default probability.

3.3 Descriptive statistics

Table 4 provides information on the composition of the sample. Slightly more than one percent of observations are considered in default on the term loan. Because individuals are either in default or not on the loan, we can use the Probit model to estimate the default probability. The table also presents the proportion of individuals with zero and total credit line usage, as well as the proportion of observations lying between these two extremes. This distribution justifies the Tobit modeling of the credit line utilization variable and our preference for the second set of results presented in Table 7. Some authors have also used OLS estimation (e.g. Agarwal et al., 2006a; Jiménez et al., 2009b), which is in line with the third set of results presented in Table 7.

Table 5 shows the proportion of the sample for which credit line utilization is either zero or total, depending on the default status of the loan. 38.8% of the observations for which the loan is considered in default have a total utilization of their credit line. Only 3.9% of the observations in the non-default group have such utilization.

Table 6 presents the mean and median of each variable, depending on whether the loan is in default or not. We test whether these statistics are significantly different across the default and non-default groups. The mean (median) credit line utilization is 78% (99%) for the default group compared with 45% (43%) for the non-default group. These numbers are in line with our hypothesis. In our sample, the higher the credit score, the riskier the borrower is considered by the institution. This is reflected across both groups, and the means and medians are statistically different. As for the age of the borrower and his seniority at the institution, the median test shows that defaulters are both significantly younger and have a shorter business relation with the bank. The sex and work variables are dummies and their median is 1; hence no value is greater than the median and the comparison test cannot be performed. The mean comparison test shows that most defaulters are male and employed. This suggests that borrowers holding a job take on more debt than they can manage, perhaps because they are overconfident about their financial situation. The median ratio of defaulters' ability to pay is significantly higher, meaning their monthly payment on the loan is higher in comparison to their ability to pay than for non-defaulters. Finally, individuals in default on the term loan have a shorter median of remaining term on their loan.

4. Estimation results

4.1 Endogeneity-Related Tests

Our first hypothesis about the relationship between credit line utilization and default probability on a term loan is that both variables are endogenous. To test this hypothesis, we perform a Wald test on the maximum likelihood estimation. Results, presented in Table 8, show that for both equations, the exogeneity hypothesis is rejected and each variable is considered endogenous in the equation for which it is used. To test the validity of the instruments, we use the Amemiya-Lee-Newey overidentification test available on the minimum chi-square estimation. This test concludes that both instrument sets are valid and confirms that we have eliminated the endogeneity bias in the system. The result of this test is also presented in Table 8.

4.2 Marginal Effects

We choose to present the marginal effects on average of the explanatory variables of the model obtained from the maximum likelihood estimation to better capture the conditional effects. For each of the equations studied, it is possible to derive different marginal effects. We retain the marginal effects of the linear prediction for the probability of default and the marginal effects of the latent variable for credit line utilization; they are presented in Table 8. Table 9 presents the Tobit decomposition according to McDonald and Moffitt (1980) for the credit line utilization equation.

4.3 Credit line utilization equation

As assumed, the default probability on the term loan is one of the most important determinants of credit line utilization. This variable is significant and is quantitatively the most important one in the model. This confirms the hypothesis that increasing the probability of default on a term loan leads borrowers to use their lines of credit more aggressively, even when the risk rating assigned by the bank is taken into account. The marginal effect of this variable indicates that individuals in default on a term loan use their line of credit about 59% more often than individuals who are not in default. The credit score variable suggests that borrowers who present a higher risk for the bank are the ones who use their lines the most.

As for the borrower's ability to pay, results show that in comparison to the fifth category (the best ability to pay), borrowers with the worst ability to pay are the ones who use their credit lines the most. This is very intuitive, because the variable measures the financial constraint of the borrower and thus serves as a measure of short-term liquidity needs that can be met by credit line utilization. Our results confirm that borrowers who have difficulty making the payments on their loan rely more on the use of their credit line. Further, the effect on credit line utilization for category 4 (Abil 4) is more than double that of category 1 (Abil 1), emphasizing the fact that borrowers with a worse ability-to-pay ratio need to rely more on the liquidity provided by the credit line.

McDonald and Moffitt (1980) decompose the marginal effects for a Tobit model with a lower limit into two categories: the intensive margin and the extensive margin. Their analysis can be applied to the Tobit model used in this paper. Such decomposition allows an analysis of the drawdown behavior of individuals depending on the usage category they are in. It provides insights about borrowers shifting from zero ($y_2 = 0$) to moderate $0 < y_2 < 1$ or total use $y_2 = 1$ of the credit line. Note that the unconditional expectation of the dependent variable can be written as:

$$E\ y_2\ = P\ y_2 = 0\ * E\ y_2 | y_2 = 0\ + P\ 0 < y_2 < 1\ * E\ y_2 | 0 < y_2 < 1\ + P\ y_2 = 1\ * E\ y_2 | y_2 = 1 \quad (4)$$

By taking the derivative with respect to the explanatory variables, we get:

$$\frac{\partial E\ y_2}{\partial X_k} = \frac{\partial P(0 < y_2 < 1)}{\partial X_k} * E\ y_2 | 0 < y_2 < 1\ + \frac{\partial E\ y_2 | 0 < y_2 < 1}{\partial X_k} * P\ 0 < y_2 < 1\ + \frac{\partial P(y_2 = 1)}{\partial X_k} \quad (5)$$

There are thus three effects to take into account: 1) the extensive margin; 2) the intensive margin; and 3) the change in the probability of using 100% of the credit line. We present the extensive and intensive margins in Table 9. The statistical significance for each variable is very similar; differences come from the estimation of the marginal effects used in the decomposition.

The first important result of such analysis is that the sign of the default probability variable is not the same for the extensive margin as for the intensive margin. The results show that the default probability has a positive impact on the marginal use of a credit line for individuals who already have a utilization rate between zero and one. Individuals facing default on the term loan thus decide to increase their credit line utilization to meet their liquidity needs. Even in situations of financial distress, individuals who already use their credit lines would therefore not hesitate to take on more debt to fulfill other financial obligations. However, for individuals who do not use their lines of credit or who already use the maximum amount authorized by the bank, the default situation has a negative impact on the latent utilization rate. Indeed, the effect of the probability of default at the extensive margin is negative. This shows that individuals who are not using their credit lines would be reluctant to start using them in the event of a default on the term loan. Perhaps such borrowers do not wish to aggravate their financial situation. This effect, however, also includes individuals who already use the line fully, although their proportion is very small in the sample. The marginal effects presented thus confirm the assumption of our model and suggest that the situation of default on a term loan affects individuals' decisions to draw on their lines of credit.

4.3 Default Probability Equation

The credit line utilization variable is highly significant in the default probability equation and is quantitatively the most important factor affecting default probability. The result suggests that a higher utilization of the credit line has a statistically significant impact on increasing the likelihood of default on a term loan. This, again, confirms our hypothesis and shows that it is imperative for financial institutions to assess consumer default probabilities jointly with the use of various financial instruments. This variable is even more important than the credit score assigned by the financial institution. The use of a line of credit is hence a potential signal of the default probability of a consumer term loan. The marginal effect of this variable on the default probability of the term loan is approximately 46%.

Further, the credit score is highly significant and positive. It shows that the bank has successfully managed to classify customers according to their perception of the risk they present. Obviously, a riskier consumer has a higher probability of default.

The first two categories of the ability-to-pay ratio are significant and show that individuals with a lower ability to pay have a higher probability of default. This significance disappears in categories 3 and 4, comprising individuals with better financial capacity. This leads us to believe that the ability to pay is especially crucial for the determination of the probability of default on a term loan when the borrower is very financially constrained. The economic effect is important and people with a very poor ability to pay would therefore send a signal of high probability of default to financial institutions, which should use this variable to assess the default risk of their customers in supplement of the credit score information. It indicates, indirectly, that the credit score variable is not sufficient to assess default probability. The credit score variable is, in fact, more of a proxy for credit delinquency than a measure of the customer’s liquidity constraint.

The number of active credit lines is particularly interesting in this equation; it is highly significant and its economic impact is very important. The results clearly show that borrowers with multiple lines of credit have a higher probability of default on their term loan. The sign of this variable allows us to conclude that individuals with a high number of active lines of credit do not benefit from better liquidity. Instead, each additional credit line increases the default probability by 26.23%. This result confirms that the various lines of credit individuals possess negatively affect their credit quality because they should be already heavily drawn down. It would thus be important for financial institutions to use this variable in their decision models to monitor borrowers’ credit activity and to be kept informed of other credit lines contracted.

5. Conclusion

This research studies two credit instruments simultaneously through joint modeling of credit line utilization and default probability on a term loan. Research on credit line utilization has generally focused on the relationship between the drawdown behavior of the borrower on a credit line and the default probability associated with the same line. We innovate by analyzing two different financial instruments simultaneously and by quantifying the effect of each instrument on the other. We estimate a simultaneous equation model in which the default status on the loan is modeled by an instrumental Probit equation, while the credit line utilization is modeled by an instrumental Tobit equation.

Our main hypotheses are that the two independent variables of the model are endogenous and that their relationship is positive. We use valid instruments to eliminate the econometric bias and affirm that an increased use of the credit line is associated with an increased default probability on a term loan, while an increased default probability is associated with an increased use of the line. The estimated marginal effects indicate that a default status on the term loan is associated with an increase of about 59% in credit line utilization. As for the credit line equation, the marginal effects show that for individuals moving from 0% to 100% utilization of their credit line, the probability of default on the loan increases by approximately 46%. We also find that the number of active lines of credit that an individual possesses is an important determinant of the likelihood of default on a term loan. The marginal effect of an additional line of credit for consumers leads to an increase of 26% in the probability of default. We also propose a new variable to measure the borrower's ability to pay. This variable complements the credit scoring variable developed by banks and credit agencies, which is more a measure of delinquency. Our results indicate that borrowers with the worst ability-to-pay ratios are the most likely to default on a term loan and to use their credit lines more extensively. These figures are reasonable and reflect the composition of the sample. The results thus highlight the need to evaluate simultaneously the credit risk for the consumer's portfolio of financial instruments instead considering each financial instrument individually.

5.1 Implications

Basel regulation requires banks to set aside a minimum capital reserve to avoid financial disasters. Such legislation has been adopted by several countries since 1988 and seeks to protect depositors of financial institutions. It is based primarily on banks' portfolio assessment of credit risk on bank loans or other assets such as private bonds. The Committee allows banks to develop an internal method for computing the capital to keep in reserve. Once their methodology is accepted, banks can use their own estimations of default probabilities, recovery rates, and loss given default. Given the results of our analysis, the amount borrowers draw on the credit line is correlated with the default probability on the term loan. The inclusion of such a correlation should allow the bank to manage risk diversification more effectively. A bank could therefore manage the risks of a borrower as a portfolio by taking into account the significant dependence across the borrower's various financial obligations. By creating portfolios of consumers' portfolios that allow greater diversification, financial institutions could reduce the minimum regulatory capital they need to keep in reserve to satisfy regulators.

5.2 Limits and possible extensions

Our analysis is limited to the data available. Studies on the use of a line of credit suggest that this instrument is positively correlated with changes in the creditworthiness of the borrower in a temporal relationship (Agarwal et al., 2006a; Jiménez et al., 2009b; Norden and Weber, 2010). It would thus be interesting to study our model in the context of panel data that could address the effects associated with different economic cycles. Such an analysis would provide more information on the dynamics of borrower behavior and could increase the potential correlation between the two financial instruments. Credit card balances could also be incorporated in our framework, if such data were available.

References

Agarwal, Sumit, Brent W. Ambrose, and Chunlin Liu (2006a). "Credit Lines and Credit Utilization." *Journal of Money, Credit, and Banking* 38(1), 1-22.

Agarwal, Sumit, Brent W. Ambrose, Souphala Chomsisengphet, and Chunlin Liu (2006b). "An Empirical Analysis of Home Equity Loan and Line Performance." *Journal of Financial Intermediation* 15, 444-469.

Agarwal, Sumit, Souphala Chomsisengphet, and John C. Driscoll (2004). "Loan Commitments and Private Firms." Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System (U.S.) 2004-27.

Amemiya, Takeshi (1978). "The Estimation of a Simultaneous Equation Generalized Probit Model." *Econometrica* 46(5), 1193-1205.

Amemiya, Takeshi (1979). "The Estimation of a Simultaneous Equation Tobit Model." *International Economic Review* 20(1), 169-181.

Calem, Paul, Matthew Cannon, and Leonard Nakamura (2011). "Credit Cycle and Adverse Selection Effects in Consumer Credit Markets – Evidence from the HELOC Market." Federal Reserve Bank of Philadelphia Working Paper 11-13.

Campbell, John Y. (2006). "Household Finance." *Journal of Finance* 61(4), 1553-1604.

Campbell, John Y., Howell E. Jackson, Brigitte C. Madrian, and Peter Tufano (2010). "The Regulation of Consumer Financial Products: An Introductory Essay with Four Case Studies." HKS Working Paper RWP10-40.

Dey, Shubhasis (2005). "Lines of Credit and Consumption Smoothing: The Choice between Credit Cards and Home Equity Lines of Credit." Bank of Canada Working Paper 2005-18.

Jiménez, Gabriel, Jose A. Lopez, and Jesús Saurina (2009a). "EAD Calibration for Corporate Credit Lines." Federal Reserve Band of San Francisco Working Paper 2009-02.

Jiménez, Gabriel, Jose A. Lopez, and Jesús Saurina (2009b). "Empirical Analysis of Corporate Credit Lines." *The Review of Financial Studies* 22(12), 5069-5098.

Lee, Lung-Fei (1992). "Amemiya's Generalized Least Squares and Tests of Overidentification in Simultaneous Equation Models with Qualitative or Limited Dependent Variables." *Econometric Reviews* 11(3), 319-328.

McDonald, John F., and Robert A. Moffitt (1980). "The Uses of Tobit Analysis." *The Review of Economics and Statistics* 62(2), 318-321.

Nelson, Forrest and Lawrence Olson (1978). "Specification and Estimation of a Simultaneous-Equation Model with Limited Dependent Variables." *International Economic Review* 19(3), 695-709.

Newey, Whitney, K. (1987). "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables." *Journal of Econometrics* 36, 231-250.

Norden, Lars, and Martin Weber (2010). “Credit Line Usage, Checking Account Activity, and Default Risk of Bank Borrowers.” *The Review of Financial Studies* 23(10), 3665-3699.

Statistics Canada (2006). “Assets and Debts Held by Family Units, Total Amounts.” Available online: http://www40.statcan.gc.ca/l01/cst01/famil109-eng.htm.

Strahan, Philip E. (1999). “Borrower risk and the price and nonprice terms of bank loans.” Federal Reserve Bank of New York Staff Report 90.

Sufi, Amir (2009). “Bank lines of credit in corporate finance: An empirical analysis.” The *Review of Financial Studies* 22(3), 1057-1088.

Tufano, Peter (2009). “Consumer Finance”. *Annual Review of Financial Economics* 1, 227-247

Wooldridge, Jeffrey M. (2002), *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

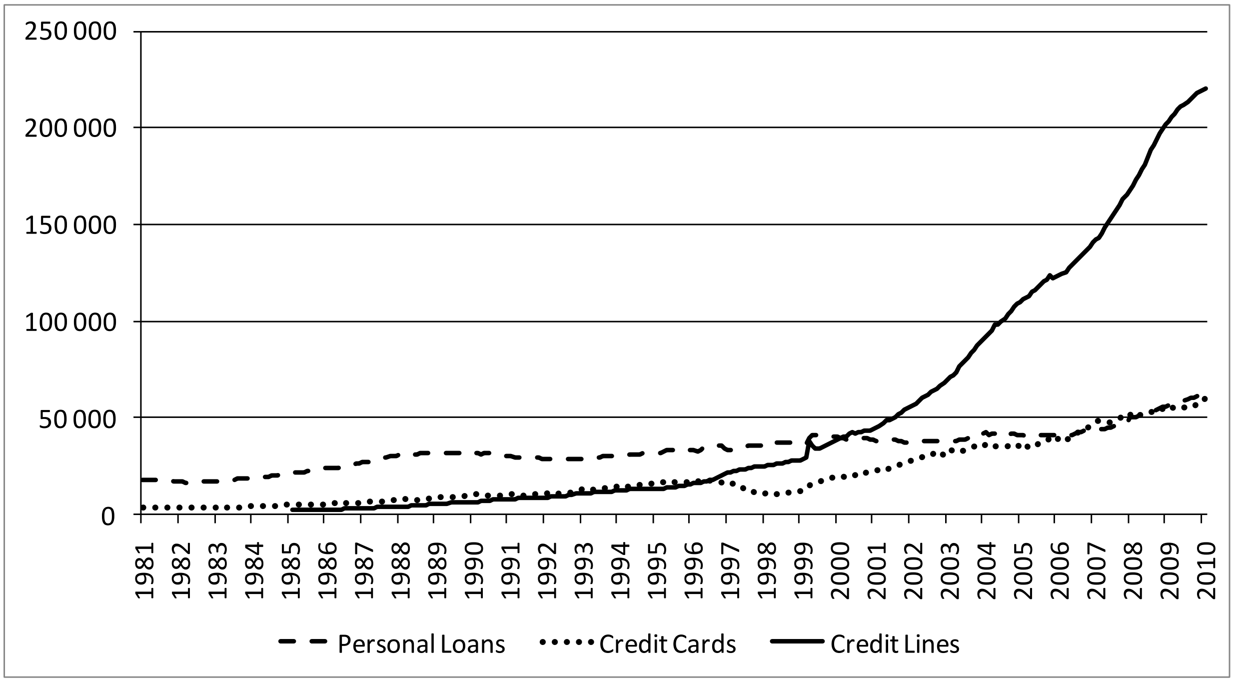

Figure 1
Assets of Canadian Chartered Banks ($ million), 1981-2010

The figure plots the monthly average value (in $ million) of personal loans, credit cards, and credit lines for Canadian Chartered Banks from 1981 to 2010. The data come from the series v36867, v36868 and v36869 of CANSIM.

Table 1
List of variables used in the econometric analysis

Variable	Code		Description
Default probability	Def	=	1 if the term loan is in default, 0 otherwise.
Credit line utilization	Util	=	$\frac{\text{Value of the credit line drawn on December 31, 2007}}{\text{Total amount authorized by the bank}}$.
Borrower sex	Sex	=	1 if the borrower is male, 0 otherwise.
Credit score	Score	=	Internal credit rating given to the client by the bank (1 to 8 with 1 being the lowest risk category).
Seniority	Sen	=	Seniority of the client at the institution (months).
Employment	Work	=	1 if the borrower is employed, 0 otherwise.
Dependents	Dep	=	Number of borrower's dependents
Borrower age	Age (categories 1 to 6 with 1 for the youngest group)	=	1 if consistent with the category, 0 otherwise.
Borrower ability to pay	Abil (categories 1 to 5 with 5 for the safer category)	=	1 if consistent with the category, 0 otherwise.
Remainder on the loan	Rem	=	Remaining term on the loan, in months.
Loan collateral	Loan_coll	=	1 if the loan is secured by collateral, 0 otherwise.
Additional credit lines	Lines	=	Number of additional active credit lines
Credit line used	Used	=	Amount used on the credit line
Credit line collateral	Line_coll	=	1 if the credit line is secured by collateral, 0 otherwise.

The table presents the variables, their code, and the definition used in the econometric analysis. The first two variables are used as dependent variables in the analysis; the other ones are used as independent variables.

Table 2
Estimates used for predicting expenditures

Variables	Male Sample	Female Sample
Intercept	601.44 (0.0001) ***	208.54 (0.0007) ***
Income	0.6126 (0.0001) **	0.7257 (0.0001) **
Dependents	469.97 (0.0001) ***	274.58 (0.0001) ***
Owner	147.77 (0.0028) **	201.17 (0.0002) ***
Age 1	33.08 (0.7117)	183.08 (0.0717) *
Age 2	64.57 (0.3269)	62.57 (0.3269)
Age 4	-129.91 (0.0323) **	143.76 (0.0213) **
Age 5	-151.74 (0.0421) **	286.32 (0.0007) ***
Age 6	-174.68 (0.2690)	-37.77 (0.7890)

The table presents the estimates used for predicting the expenditures in the ability-to-pay ratio. The first column reports the OLS results for the male sample while the second column reports the same estimation results for the female sample. Values in parentheses represent the p-value of the test statistic for the null hypothesis that the values are zero. *, **, and *** represent the significance at the 10%, 5%, and 1% levels respectively.

Table 3
Ability-to-pay ratio

Category	Ability to Pay
Abil 1	Ratio $\leq$ 0 or Ratio $\geq$ 1
Abil 2	$0.498 \leq \text{Ratio} < 1$
Abil 3	$0.249 \leq \text{Ratio} < 0.498$
Abil 4	$0.127 \leq \text{Ratio} < 0.249$
Abil 5	$0 < \text{Ratio} < 0.127$

The table presents the categories used for the ability-to-pay ratio. This ratio is expressed as the monthly payment on the loan divided by the sum of income and tangible assets minus the predicted expenditures. A negative value for this ratio implies that the expenditures are higher than the income and tangible assets added, while a value of more than one implies a monthly payment higher than the client's ability to pay. Safer borrowers (Abil 5) have a small but positive ratio.

Table 4
Dependent variables

Panel A: Loan Status	Observations	Proportion of data
Default	160	1.09%
Non-default	14,667	98.91%
Total	14,827	100%

Panel B: Credit Line Utilization	Observations	Proportion of data
0%	4,459	30.07%
0 % < Credit Line Utilization < 100%	9,726	65.60%
100%	642	04.33%
Total	14,827	100%

Panel A reports the number of observations in the default and non-default states, and their proportion in the data. Panel B reports the number of observations for a credit line utilization of 0%, between 0% and 100%, and 100%. It also shows the proportion of each category in the data.

Table 5
Credit line utilization by loan status

	Proportion of borrowers with	
Loan Status	0% utilization	100% utilization
Default	7.5 %	38.8 %
Non-default	30.3 %	3.9 %

The table presents the proportion of observations in the default and non-default states for borrowers using 0% and 100% of their credit line.

Table 6
Explanatory variables

	Default Group			Non-Default Group			Comparison Tests	
	Mean	Median	Standard deviation	Mean	Median	Standard deviation	Mean Comparison (T-test)	Median Comparison (Chi(2))
Util	0.78	0.99	0.39	0.45	0.43	0.33	0.0000 ***	0.000 ***
Score	4.55	4	1.82	2.88	3	1.67	0.0000 ***	0.000 ***
Age	38.94	37	10.84	41.74	42	11.77	0.9986	0.017 **
Sen	160.45	125	126.57	200.28	180	131.53	0.9999	0.000 ***
Dep	0.14	0	0.54	0.17	0	0.56	0.7115	0.622
Sex	0.72	1	0.45	0.66	1	0.47	0.0819 *	ψ
Work	0.93	1	0.26	0.89	1	0.31	0.0950 *	ψ
Abil	0 .96	0.38	7.149	0.75	0.25	43.568	0.4754	0.001 ***
Rem	17.96	15.63	19.132	23.68	19.23	22.62	0.9993	0.083 **
Loan_coll	0.22	0	0.41	0.24	0	0.43	0.7214	0.557
Lines	0.84	1	0.44	0.82	1	0.48	0.2359	0.726

The table presents the mean and median of each explanatory variable by the borrower’s loan status, along with a mean and median comparison test. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels respectively. Ψ means that the test statistic is not available because the variable is dichotomous and the median is 1 for each category.

Table 7
Estimation results

	(1) Newey's Two-Step Estimation		(2) Maximum Likelihood Estimation		(3) Two-Stage Probit Least Squares	
	Instrumented Tobit	Instrumented Probit	Instrumented Tobit	Instrumented Probit	Instrumented OLS	Instrumented Probit
Util		0.4616 (0.005) ***		0.4585 (0.006) ***		0.2765 (0.042) **
Def	8.5760 (0.000) ***		8.6685 (0.000) ***		0.1375 (0.000) ***	
Score	0.0267 (0.028) **	0.1829 (0.000) ***	0.0261 (0.034) **	0.1818 (0.000) ***	0.0318 (0.000) **	0.1804 (0.000) ***
Sen	0.0003 (0.000) ***	-0.0001 (0.793)	0.0003 (0.000) ***	-0.0001 (0.792)	0.0002 (0.000) ***	-0.0001 (0.849)
Work	0.0433 (0.130)	0.0278 (0.831)	0.0433 (0.133)	0.0282 (0.828)	0.0183 (0.340)	0.0200 (0.874)
Dep	0.0219 (0.133)	-0.0867 (0.185)	0.0220 (0.134)	-0.0862 (0.185)	0.0171 (0.084) *	-0.0870 (0.175)
Age 1	-0.0225 (0.654)	0.2341 (0.257)	-0.0225 (0.658)	0.2323 (0.257)	-0.0536 (0.086) *	0.2259 (0.256)
Age 2	-0.0405 (0.334)	0.3844 (0.018) **	-0.0411 (0.331)	0.3823 (0.018) **	-0.0433 (0.106)	0.3726 (0.017) **
Age 3	0.0140 (0.738)	0.4003 (0.013) **	0.0132 (0.754)	0.3983 (0.012) **	-0.0019 (0.943)	0.3855 (0.012) **
Age 4	-0.0391 (0.405)	0.6229 (0.001) ***	-0.0402 (0.396)	0.6196 (0.001) ***	-0.0420 (0.203)	0.5953 (0.001) ***
Age 5	-0.0877 (0.062) *	0.5731 (0.005) ***	-0.0886 (0.061) *	0.5700 (0.005) ***	-0.0699 (0.040) **	0.5233 (0.009) ***
Sex	-0.0825 (0.000) ***	0.2235 (0.003) ***	-0.0829 (0.000) ***	0.2224 (0.003) ***	-0.0544 (0.000) ***	0.2168 (0.004) ***
Abil1	0.1488 (0.000) ***	0.3904 (0.002) ***	0.1478 (0.000) ***	0.3882 (0.002) ***	0.1041 (0.000) ***	0.4125 (0.001) ***
Abil 2	0.1350 (0.000) ***	0.2998 (0.013) **	0.1343 (0.000) ***	0.2983 (0.013) **	0.0825 (0.000) ***	0.3196 (0.007) ***
Abil 3	0.1384 (0.000) ***	-0.0522 (0.696)	0.1385 (0.000) ***	-0.0516 (0.698)	0.0831 (0.000) ***	-0.0164 (0.900)
Abil 4	0.0602 (0.015) **	0.1686 (0.171)	0.0600 (0.017) **	0.1679 (0.170)	0.0268 (0.152)	0.1872 (0.118)
Rem		-0.0058 (0.004) ***		-0.0058 (0.004) ***		-0.0060 (0.002) ***
Loan_coll		-0.0369 (0.652)		-0.0374 (0.645)		-0.0239 (0.762)
Lines		0.2636 (0.000) ***		0.2622 (0.000) ***		0.2614 (0.000) ***
Used	0.0860 (0.000) ***		0.0860 (0.000) ***		0.0647 (0.000) ***	
Line_coll	-0.1027 (0.010) **		-0.1032 (0.010) **		-0.0580 (0.011) ***	
Intercept	-0.1071 (0.103)	-4.1746 (0.000) ***	-0.1049 (0.115)	-4.1506 (0.000) ***	0.5046 (0.000) ***	-3.9795 (0.000) ***
Observations	14, 827	14, 827	14, 827	14, 827	14, 827	14, 827
Adj. Pseudo R^2	0.0136	0.4378	-	-	0.4369	0.1138
P-value	0.000	0.000	0.000	0.000	0.000	0.000

The table presents the three sets of results estimated for the model. The first set of results is estimated with Newey's Two-Step Efficient Estimator (Newey, 1987), a limited information procedure. Standard errors for this estimator are based on Amemiya's (1978, 1979) derivations of the efficient variance-covariance matrices. The second set of results is estimated from a full information maximum likelihood procedure. For robustness, the last set of results is a joint two-step estimation of OLS and Probit models, known as Two-Step Probit Least Squares (2SPLS). All procedures are instrumented for the endogenous variables of the system and all results reported are for the second stage estimation. Values in parentheses represent the p-value of the test statistic for the null hypothesis that the values are zero. *, **, and *** represent the significance at the 10%, 5%, and 1% levels respectively.

Table 8
Marginal effect coefficients derived from the MLE

	Credit line utilization (Instrumented Tobit)	Default Probability (Instrumented Probit)
Util		0.4585 (0.006) ***
Def	0.5891 (0.000) ***	
Score	0.0104 (0.034) **	0.1818 (0.000) ***
Sen	0.0001 (0.000) ***	-0.0001 (0.792)
Work	0.0172 (0.131)	0.0282 (0.828)
Dep	0.0088 (0.134)	-0.0862 (0.185)
Age 1	-0.0090 (0.658)	0.2323 (0.257)
Age 2	-0.0164 (0.330)	0.3823 (0.018) **
Age 3	0.0053 (0.754)	0.3983 (0.012) **
Age 4	-0.0160 (0.396)	0.6196 (0.001) ***
Age 5	-0.0351 (0.060) *	0.5700 (0.005) ***
Sex	-0.0331 (0.000) ***	0.2224 (0.003) ***
Abil 1	0.0592 (0.000) ***	0.3882 (0.002) ***
Abil 2	0.0537 (0.010) ***	0.2983 (0.013) **
Abil 3	0.0554 (0.000) ***	-0.0516 (0.698)
Abil 4	0.0240 (0.017) **	0.1679 (0.170)
Rem		-0.0058 (0.004) ***
Loan_coll		-0.0374 (0.645)
Lines		0.2623 (0.000) ***
Used	0.0331 (0.000) ***	
Line_coll	-0.0343 (0.010) ***	
Amemiya-Lee-Newey overid. test	1.15 (0.283)	3.32 (0.854)
Wald test of exogeneity	22.17 (0.000) ***	4.64 (0.031) **
Observations	14827	14827
Prob > chi2		0.000
Prob > F	0.000	

The table reports the marginal effects of the independent variables derived from the Maximum Likelihood Estimation. This estimation allows a test of the exogeneity of the default probability and the credit line utilization variables by a Wald test of exogeneity. The null hypothesis of the test is the exogeneity of the variable. Newey's Two-Step Efficient Estimation allows a validation of the chosen instruments for each endogenous variable by an Amemiya-Lee-Newey overidentification test (Lee, 1992). The null hypothesis of the test is the validity of the instruments. Values in parentheses represent the p-value of the test statistic for the null hypothesis that the values are zero. *, **, and *** represent the significance at the 10%, 5%, and 1% levels respectively. Marginal effects reported for the Probit model are based on linear prediction.

Table 9
McDonald and Moffitt (1980) decomposition

	Credit line utilization (Tobit)	
	Intensive Margin	Extensive Margin
Def	0.1633 (0.000) ***	-0.1914 (0.000) ***
Score	0.0009 (0.034) **	0.0007 (0.036) **
Sen	0.0000 (0.000) ***	0.0000 (0.000) ***
Work	0.0015 (0.132)	0.0013 (0.178)
Dep	0.0008 (0.134)	0.0006 (0.135)
Age 1	-0.0008 (0.658)	-0.0006 (0.678)
Age 2	-0.0015 (0.331)	-0.0011 (0.365)
Age 3	0.0005 (0.754)	0.0003 (0.748)
Age 4	-0.0014 (0.396)	-0.0011 (0.426)
Age 5	-0.0031 (0.061) *	-0.0028 (0.120)
Sex	-0.0029 (0.000) ***	-0.0019 (0.000) ***
Abil 1	0.0053 (0.000) ***	0.0023 (0.000) ***
Abil 2	0.0048 (0.000) ***	0.0024 (0.001) ***
Abil 3	0.0049 (0.000) ***	0.0025 (0.000) ***
Abil 4	0.0021 (0.017) **	0.0013 (0.033) **
Used	0.0031 (0.000) ***	0.0022 (0.000) ***
Line_coll	-0.0037 (0.010) ***	-0.0036 (0.042) **

The table reports the McDonald and Moffitt (1980) decomposition for the Tobit equation of credit line utilization based on Equation (5). The intensive margin reports the marginal probability of utilization of the credit line for an individual who already has a utilization rate between 0% and 100%. The extensive margin reports the marginal utilization of the credit line for an individual who has a 0% or 100% utilization of the credit line. Values in parentheses represent the p-value of the test statistic for the null hypothesis that the values are zero. *, **, and *** represent the significance at the 10%, 5%, and 1% levels respectively.