The role of Credit Officers in the Performance of Micro Loans: evidence from Brazil

By : Isabelle Agier
Juliano Assunção

Cahier de recherche 02-2011
Cahier de recherche GIREF 01-2011

GIREF
Groupe international de recherche en éthique financière et fiduciaire

http://www.giref.uqam.ca/fr/index.php

Isabelle Agier: Paris school of Economics (France), CERMi (Belgium), PUC-Rio (Brazil), Email: isabelleagier@gmail.com, Isabelle Agier is grateful to CAPES and FAPERJ in Brazil for granting this research.

Juliano Assunção: Ponti□cal Catholic University of Rio de Janeiro, Department of Economics, Email: juliano@econ.puc-rio.br

1 Introduction

Asymmetric information is an essential issue of credit markets. There is a substantial literature on how different aspects of the credit contract can help in mitigating its consequences: screening (Stiglitz and Weiss, 1981), collateral (Bester and Hellwig, 1987), dynamic incentives (Morduch, 1999; Armendáriz and Morduch, 2000), group lending (Stiglitz, 1990; Banerjee et al., 1994; Besley and Coate, 1995; Ghatak, 1999).

Relationship between lenders and borrowers has been documented in the literature as a potential tool to solve asymmetric information problems. Relationship has two main advantages: on one hand, it allows to acquire information on the borrower and to uncover his type (Petersen and Rajan, 1995; Boot and Thakor, 2000), and, on the other hand, to discipline his behavior and align borrower's target to lender's one (Boot and Thakor, 1994). Value of relationship in lending comes from the dynamic aspect and may generate a gain for borrowers in terms of loan conditions: price and access (Petersen and Rajan, 1994, 2002). Technology to acquire information is documented by Berger et al. (2005b) (advantage of small structure) and Berger et al. (2005a) (credit scoring). Nonetheless, in all this literature, relationship is depicted as a black-box where individual actions and responsibilities from the loan officer and the rest of the lending institution are not distinguished.

Recent literature has considered explicitly the credit officer. Labie et al. (2010) depict their role in potential clients discrimination and Beck et al. (2009) questions whether officer's gender matters in credit performances. In this paper, we focus on the role of the credit officers as key players to deal with information asymmetry issues. They are responsible for collecting information about potential solvency of credit applicants and to verify the ex-post state of nature as in Townsend (1979) and Gale and Hellwig (1985). Our analysis is based on a model of credit performance in which the credit officer's ability is considered explicitly. The key parameters of the model are estimated with the Vivacred database.

The empirical setting provided by the Vivacred data is quite favorable for the sake of this study. First, there is information about who is the credit officer in charge of each contract. The sample contains both cross-section and time series variation to estimate the parameter of each officer. Second, all contracts in the sample are homogeneous: individual loans with a flat interest rate. Third, different from other studies reviewed by Hermes and Lensink (2007), we have a large sample of almost 32,000 contracts over an 11 years period.

Our analysis starts with a model of lending with costly state verification. The credit officer's ability affects both the probability of success of the project financed and the outcome from the auditing process for clients declaring failure. From this model, we derive the probability of the payment delay as a function of the loan size, the client's income and profile and the credit officer's ability. We then specify functional forms for the distribution of all random variables of the model in order to estimate the structural parameters of the model by maximum likelihood.

Our estimates depict substantial variation among credit officers' ability. The individual probability to detect misreporting varies from 66% to 82% between credit officers, and the probability of success of their portfolio ranges from 58.9% to 99.8%. Estimation suggests that they make more difference at the selection stage than at the audit.

Credit officers estimated ability is positively correlated with the experience gained at Vivacred, measured by the number of loans under supervision and number of month worked for the institution. It is not the case for the previous experience, measured in terms of professional length (in years) and the age at hiring. This evidence suggests either the existence of learning-by-doing or the importance of relationship as studied in Berger and Udell (1995), Petersen and Rajan (1995) and Carrasco and de Mello (2006).

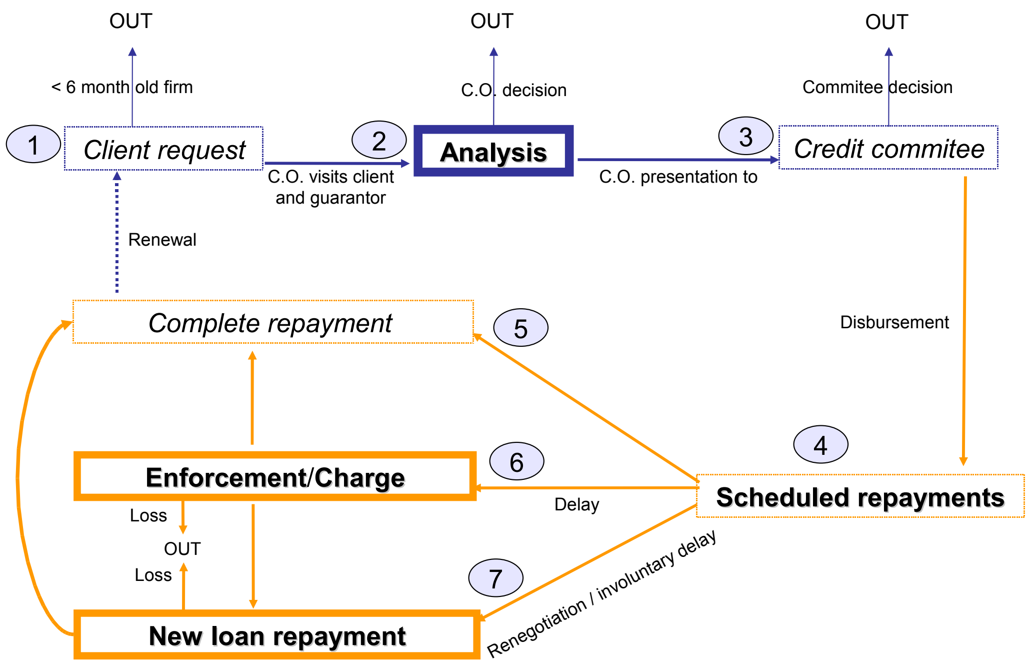
Based on the estimates, we have used the model to simulate different situations in order to illustrate the impact of credit officer's ability on the overall payment delay. The observed baseline probability of delay is 8.6%. Removing credit officers' heterogeneity (same ability level for all), this overall delay is 10.5% for median ability and 11.5% for average ability. Repeating this exercise for the whole range of ability levels estimated, the probability of delay vary from 41% to 2% applying from the worst to the best credit officer's ability level for everyone.

The remaining parts of this paper are organized as follows. Section 2 and 3 provide a description of Vivacred and the data. Section 4 presents the structural model and its implementation for estimation. Section 5 and 6 presents the estimation results and the relationship between the officers' profile and their estimated ability. Simulation exercises are presented in section 7 and section 8 concludes.

2 Credit officer's role in credit cycle

Credit officers have an important potential in order to reducing information asymmetry problems, handling two main tasks during the process: (i) collect information about applicant solvency and present it to the credit committee for approval, (ii) check the ex-post state of nature and enforce the contract in case of payment delay. To better understand which steps of the process are influenced by the officer, it is useful to explain, step by step, how the credit cycle works in Vivacred, as summarized in figure 1.

Figure 1: Credit cycle in Vivacred



The process starts with the client request (step 1). In order to be eligible, the borrower's firm cannot be created less than six month before the application and the client cannot be registered in the "Serviço de Proteção ao Crédito" - SPC[1]. To apply for a credit in Vivacred, the client has to present ID documents of both himself and the guarantor(s), along with the documents of the firm. In addition, the purpose of the loan has to be explained.

An important issue for the purpose of this study is how credit officers are

[1]SPC is a national database recording all the late payments declared by any institution delivering any kind of credit (including payments in credit).

assigned to each contract. There is no strategic consideration on the allocation of the credit officers according to the Vivacred staff. The match between clients and officers is made by geographic area, aiming at reducing operational costs.

The credit officer visit the client and the guarantor (step 2) for gathering the information on the collateral, solvency and business risk. The applicant is met at the business location. The client is asked to fill a questionnaire about his personal situation, the household budget and the financial situation of the firm (for more details see the data set description in section 3). The same questions are asked to the guarantor, although the financial information is not as detailed.

This process of data collection is particularly demanding for the credit officers. The typical client does not usually hold a formal accountability; neither for the business nor for the household budget. Thus, in order to evaluate the business balance sheet and the household budget, the credit officer makes indirect questions. Social skills, experience and prior knowledge of other business in favelas are very important to obtain reliable information about the business at this point. For example, the officer ask questions about the monthly and weekly amounts spent and received in different roundabout ways to check the coherence of declared values. Or the officer might ask for rough evaluations of stock and main items of the budget. After completing the basic information set, the officer helps the client to parameterize the proposal, establishing the size of the loan and the number of installments.

This application is then presented to the credit committee (step 3). The committee has the final word on the approval and the terms of the loan, which might differ from the initial proposal. In case of approval, the client receives the money and begins repayments in the following month (step 4). Installments can be paid on a monthly or semi-monthly basis, in cash or check. Sometimes the officer can even collect the money direct from the client.

In case of repayment (step 5), which means paying the whole value without significant delays, the client typically get access to another loan, possibly of bigger value. If there are no delay during the whole period of the contract, the registration fee (TAC) of the subsequent loan is reduced.

In case of delay (step 6), officers play another important role on the process. It starts with the respective officer visiting the client and trying to convince him/her to pay the installment. At the beginning of every workday, each credit officer receives a list of delayed contracts. In one or two days of delay, credit officers start to call clients. At this stage, the officers not only negotiate with the client but also help in finding ways of paying the debt. For example, a new loan repayment can be offered (step 7) if the client shows his good-faith. Credit officers put most of the effort to get the repayment within a delay period of 30 days which, along with credit origination, determines the variable portion of their monthly wage compensation. After 180 days of delay, the loan is considered lost. In case of default, the client is included on the SPC in case of negotiation failure.

3 Data description

We have data on all credit contracts of Vivacred in the period of 1997 to 2007, from all the six branches. This sample consists on about 32,000 actual contracts. Our sample comprises all the relevant dimensions of the credit contract - the client, the credit officer and the guarantor, the contract and the business characteristics. All the financial variables are deflated to January 1997, according to the consumer price index (IPC) for the city of Rio de Janeiro.

We restrict our sample to approved contracts - only the actual (disbursed) contracts are considered. We have removed from our data 79 contracts without credit officer identification, 146 contracts made by 6 trainee officers, 113 contracts in group and 25 contracts with null income. Contracts defined as "special loans", designed mainly to employees, were also removed because they fell outside of Vivacred's main activity.

The standard loan is individual (as opposed to group loans), restricted for firms with more than six months since creation. The duration varies according to the use of the money, being typically one year for treasury or two years for investments. In addition, there are short-term contracts, with less than 4 months, discount of receivables (mainly checks), and others. We have excluded joint liability group loans because they are about only 1% of our sample and are associated with different incentive mechanisms.

Each contract is classified as "paid", "delayed", or "defaulted". We create dummy variables indicating delay with more than 7, 15 and 30 days. A delay above 180 days is considered default - a situation in which the outstanding debt is considered lost, although being recoverable afterwards.

Table 1: Summary statistics

N = 31,692	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Delay 7 days	0.134	0.341	0	1
Delay 15 days	0.112	0.315	0	1
Delay 30 days	0.086	0.280	0	1
Delay 180 days	0.034	0.181	0	1

Descriptive statistics on delay are summarized in the table 1. The proportion of loans with at least one late repayment is respectively 13.4%, 11.2% and 8.6% considering the 7, 15 and 30 days of delay thresholds. Moreover, 3.4% of the loans are (at least partially) in default (delay above 180 days), representing 2.7% of the lent amount.

% of loans with at least N days of delay (by credit officer's portfolio).

Figure 2: 7 days vs 30 days

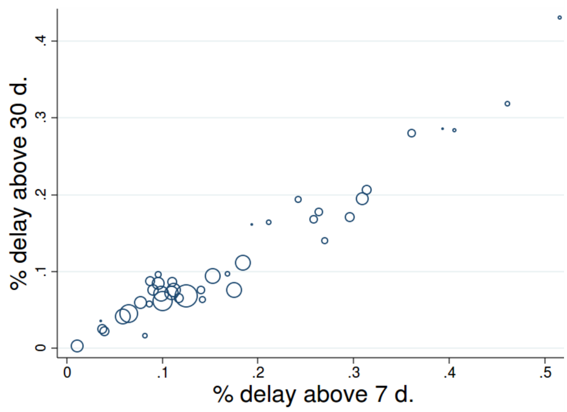


Figure 3: 30 days vs 180 days

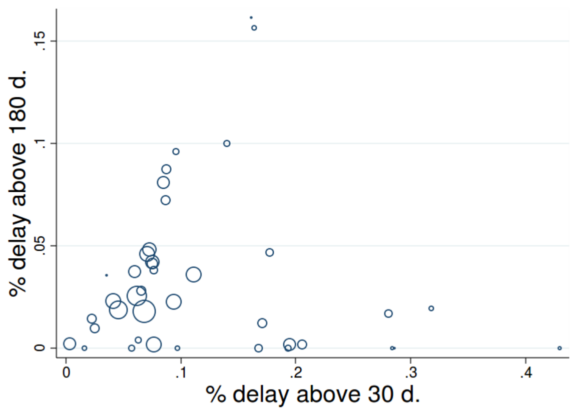


Figure 2 and 3 present, for each credit officer, the proportion of loans with at least N days of delay. Figure 2 depicts the 7 days threshold in x-axis and 30 days threshold in y-axis. Figure 3 represents the 30 and 180 days thresholds. Each circle represents a credit officer portfolio and its size is proportional to the number of loans attended.

The monotonic relationship between the 7 and 30 days threshold is obvious. Credit officers ranking does not change considering one threshold or another. The relationship between 30 and 180 days thresholds is not so clear but continue to be increasing when considering only the biggest portfolios. Some of the smallest portfolios with a high delay end up with a low default and some portfolios with medium default end up with a high delay relatively to the other credit officers. Thus, the capacity to avoid a default, conditional to observe a delay, is not heterogeneous among the credit officers. In order to compare the actual credit officers' ability, the portfolio composition will be considered later.

Nonetheless, we chose to use the 30 days delay threshold for two reasons. First, 30 days delay is an important threshold to determine credit officers' monthly bonus. Second, default is rare or does not occur at all for some credit officers. Thus, it would not be possible to apply the structural model and estimate the corresponding maximum likelihood for such data.

Table 2 presents descriptive statistics on the main variables for the whole sample and separating the loans by the 30 days of delay threshold.

Over the 11 years, the average approved loan size is R$ 1,014, to be repaid in 9 installments. 93% of the loans were signed jointly with a guarantor and 32% were financing investment (the remaining part been for treasury). Delayed contracts lies on a comparable loan size but are planned to be repaid in more installments. Half of the contracts (53%) were made for trade activities and 43% for services sector. Other activities like agriculture are marginal (4%).

Vivacred clients are equally distributed in terms of personal profile: gender, marital status, taking care of dependents or not. Different from other MFI, there is no pro-women policy. However, loans are more likely to suffer delay when taken by men (55.2% when delayed, 50.2% otherwise) and less likely when taken by married clients (32% against 43%) or clients taking care of dependents (50% against 53%). They late payers are 3 years younger.

The household extra income, household consumption and business profits are, respectively, R$292; R$525 and R$940, by month on average. The average guarantor's available income (income-consumption) is R$483. The client's extra income and the guarantor's available income are substantially lower

Table 2: Descriptive statistics

Delay	All credits	30 days <	30 days ≥	Dif.
Obs	31,692	28,964	2,728	
Credit characteristics				
Loan size	1014.2	1052.6	1030.0	22.54
Installments	9.04	9.17	9.47	−0.301***
Guarantor	0.931	0.929	0.953	−0.0247***
Capital investment	0.320	0.325	0.340	−0.0155
Loan repayment	0.095	0.090	0.126	−0.0359***
Client characteristics				
Men	0.504	0.502	0.552	−0.0498***
Married	0.476	0.503	0.373	0.129***
Have dependents	0.522	0.527	0.504	0.0232*
Age	42.4	42.4	39.2	3.2 * **
Extra Income	291.6	299.3	227.6	71.70 * **
Consumption	525.1	525.7	577.1	−51.45 * **
Current account	0.068	0.071	0.026	0.0458***
Guarantor characteristics (when required)				
Men	0.577	0.577	0.579	−0.00240
Married	0.417	0.426	0.323	0.103***
Age	46.4	46.39	46.06	0.339
Income-consumption	483.4	549.8	204.2	345.6 * **
Relationship (months and previous credits)				
Active (months)	17.3	17.58	8.435	9.149***
Total (months)	19.96	20.26	9.294	10.97 * **
# credits	2.31	2.30	1.12	1.182***
# demands	2.61	2.61	1.36	1.254***
Business characteristics				
Profit	940.4	953.2	988.2	−34.99
Agriculture	0.025	0.026	0.021	0.005
Services	0.432	0.433	0.429	0.004
Retail trade	0.528	0.530	0.512	0.018
Other sectors	0.014	0.011	0.039	−0.027***

for loans with delay, client consumption is lower and business profit is not affected by the threshold.

Other funding sources are sometimes available in the urban Brazilian context: consumption loans, credit card or credit in shops. Taking it into account may be important to understand client's repayment behavior (willingness to repay the loan). Reduced alternatives may turn more valuable the possibility of loan renewal. However, alternative access to credit is a difficult information to get for the lender. In Brazil, only information about delay and default in other institution are available through the SPC register. In Vivacred, an applicant can get a loan only if his SPC record is clean.

Bank account ownership is an available proxy for credit access in Vivacred database. Only 6.8% of the clients have a current account and those clients are far less likely to delay their repayment. Thus, current account ownership seem more likely to reflect an income effect than a relationship effect.

In addition, loans with delay arise more frequently (13% against 9%) after a loan terms renegotiation (loan repayment), and with a shorter relationship, measured both in terms of loans number (requested or approved) and months (total since the first loan and months of active repayment).

Data about the credit officers were collected directly from the Vivacred paper files. We have data on: age, gender, education, marital status, address, wage, hiring and layoff (if applicable) dates, experience (years since first job), previous positions (administrative, financial or sales experience). Table 3 depicts descriptive statistics on officers attributes.

The set of officers is equally distributed in terms of personal characteristics: 47.5% are men, 45% are married, one half has children and one half lives in a favela. Their educational level is quite high for Brazilian terms - 60% have

Table 3: Summary statistics : Credit officers profile

Variable	Mean	Std. Dev.	Min.	Max.	N
Men	0.475	0.505	0	1	40
Married	0.45	0.503	0	1	40
Have children	0.5	0.506	0	1	40
Living in favela	0.5	0.506	0	1	40
College	0.6	0.496	0	1	40
Exp. before Vivacred	7.64	5.75	0.5	27	36
Exp. in finance	0.47	0.51	0	1	36
Exp. in sales	0.58	0.5	0	1	36
Age at entrance	28.8	7.3	19.5	60.6	40
Age at exit	32.4	7.4	20.3	61.8	40
Adm. tasks before	0.125	0.334	0	1	40
Max. wage	936.8	368.1	450	2037	40
# months in Vivacred	44.6	33.3	1.6	130.9	40
# credits disbursed	792.5	805.7	29	3768	40

college. They had, on average, 7.6 years of experience before being hired by Vivacred. 47% had a previous occupation requiring financial skills and 58% in sales. Information about professional experience previous Vivacred is missing for 4 officers. Moreover, the typical officer was hired by Vivacred at almost 29 years old, stayed 45 months attending 793 actual credits (excluding denied applications). Vivacred administrative employees represent 12.5% of credit officers recruiting. The fixed part of the salary at the exit (or present one) is on average R$ 937.

Finally, there is one caveat regarding the data collecting process in Vivacred. First-time borrowers have more reliable data and a more complete information set. This is because there is only an update of the relevant issues in the subsequent contracts, when the borrower behaves well. As a consequence, the occurrence of missing information increases with the number of contracts each borrower sign. For this reason, we impute some missing variables considering the observed average of previous contracts.

4 Credit officers' ability and loan performance

4.1 Model

This model is widely inspired by Vivacred practices, where credit officers' ability affects both the selection and the state verification phase in the credit process. The model is used to provide the probability of payment delay as a function of the loan size, client characteristics and the credit officer's ability parameter.

We focus on the lender-borrower relationship. We assume the borrower has a project that requires a fix investment of size L, which can be financed at the (gross) interest rate $\rho > 1$. The return of the project is a binary random variable with values $r \in \{\underline{r}, \overline{r}\}$, where $\overline{r} > \rho > \underline{r}$. The borrower submit a credit request, providing a set of information x about the probability of success of the project. The officer, whose ability is parameterized by α, interacts with the credit committee and selects the approved projects. The financial and personal information on the client and his business, represented by x, is compiled by the credit officer and reported to the credit committee. There is subjective and potentially crucial information not reported but partially observed by the credit officer. For example, in order to compute the business balance-sheet, the credit officer asks different questions about revenues and expenditures, checking the coherence of the reported answers. An experienced credit officer better identify whether the client has cognitive problems or is trying to hide anything important. Thus, we assume that the probability of success of the approved projects is given by:

$$p_s(\mu_s, \alpha, x, b) = \frac{1}{1 + e^{-(\mu_s \cdot \alpha + xb)}},$$

where b, μ_s, α are parameters (b is a vector).

The true state of nature is not observed by the lender. The return of the project is observed only by the borrower, which reports a return $\hat{r}$ at the end of the contract. An announcement of $\hat{r} = \underline{r}$ is made through the delay of payment, which is interpreted as the first signal that the borrower is not able to repay the loan. In this case, the lender send the credit officer to visit the borrower and audit the true state of nature. The quality of the audit process also depends on the credit officer ability α - the probability of learning the true state of nature is represented by $p_a(\alpha)$. The borrower who is found misreporting the true state has to repay the total amount ρL plus additional costs $\tau L + \phi$, where τL represents a penalty that varies with the loan size and ϕ is a non-pecuniary cost such as reputation or the relationship with the guarantor.

If $r = \underline{r}$ or if the audit process fails, the borrower pays $\min(\underline{r}L + I, \rho L)$, where I is the additional resources put as collateral in the credit contract. Typically, I comes from other income sources of the borrower or even from former business profit.

The borrower payoffs in each situation are summarized as follows:

state of nature	announcement	audit	cost
$r = \overline{r}$	$\hat{r} = \overline{r}$	none	ρL
$r = \overline{r}$	$\hat{r} = \underline{r}$	success	$(\rho + \tau)L + \phi$
$r = \overline{r}$	$\hat{r} = \underline{r}$	fail	$\min(\underline{r}L + I, \rho L)$
$r = \underline{r}$	$\hat{r} = \overline{r}$	none	ρL
$r = \underline{r}$	$\hat{r} = \underline{r}$	success	$\min(\underline{r}L + I, \rho L)$
$r = \underline{r}$	$\hat{r} = \underline{r}$	fail	$\min(\underline{r}L + I, \rho L)$

The borrower always report the true state of nature when the project fails, paying $\min(\underline{r}L + I, \rho L)$. There is no incentive to misreport in this case. If the project is successful, on the other hand, there is a wrong announcement only if:

$$p_a(\alpha)\left[(\rho+\tau)L+\phi\right]+(1-p_a(\alpha))\min(\underline{r}L+I,\rho L)<\rho L,$$

which can be written as

$$\phi < \frac{[(\rho-\underline{r}-p_a(\alpha)(\rho-\underline{r}+\tau))L-(1-p_a(\alpha))\min(I,(\rho-\underline{r})L)]}{p_a(\alpha)} \equiv g(L,\rho,\underline{r},\alpha,\tau,I). \tag{1}$$

If the distribution of the reputation cost ϕ is represented by F, we have that

$$\Pr(\hat{r}=\underline{r}|r=\overline{r})=F(g(L,\rho,\underline{r},\alpha,\tau,I)). \tag{2}$$

There is delay in two situations - in case of project failure or in case of a misreport about the success of the project. If we define D as a binary variable indicating delay, we have

$$\Pr(D=1)=1-p_s(\mu_s,\alpha,x,b)+p_s(\mu_s,\alpha,x,b)\Pr(\hat{r}=\underline{r}|r=\overline{r})$$

and, substituting (2),

$$\Pr(D=1)=1-p_s(\mu_s,\alpha,x,b)+p_s(\mu_s,\alpha,x,b)F(g(L,\rho,\underline{r},\alpha,\tau,I)). \tag{3}$$

In summary, the ability of the credit officer α affects the performance of the credit contracts through two channels. First, better officers select better projects, with lower chance of failure. Second, borrowers are less willing to misreport when dealing with better officers because they anticipate a higher chance of punishment.

Notice that the model assumes there is no strategic allocation between credit officers and clients. We rule out, for instance, the possibility of having an officer with higher α (more efficient) assigned to a client with smaller ϕ (more likely to lie). The motivation behind this structure comes from the prevailing rule in Vivacred. The allocation of credit officers in Vivacred is primarily determined by the location of the business or the client home.

4.2 Estimation

Let's assume that we have a sample of N credit loans, indexed by $i = 1, ..., N$, with data on delay D_i, size of the loan L_i and the income used as collateral I_i . For each credit loan i, we also observe the identity of the credit officers, who are indexed by $j = 1, ..., M$. We denote by $j(i)$ the identity of the officer j who is assigned to client i. The interest rate and penalty are known and do not vary among borrowers ($\rho = 1.4$ and $\tau = 0.2$). The reputation cost ϕ is drawn from a log-normal distribution with parameters k_1 and k_2.

We can obtain the likelihood function from equation (3):

$$\begin{aligned}\mathcal{L}(\{\alpha_j\}_{j=1}^{M}, \mu_s, p_a(\cdot), \underline{r}, k_1, k_2, b|\mathbf{D}, \mathbf{L}, \mathbf{I}, \mathbf{x}, \rho, \tau) = \\ \prod_{i=1}^{N}[1 - p_s(\mu_s, \alpha_{j(i)}, x_i, b)(1 - F(g(L_i, \rho, \underline{r}, \alpha_{j(i)}, \tau, I_i)))]^{D_i=1} \times \\ \times [p_s(\mu_s, \alpha_{j(i)}, x_i, b)(1 - F(g(L_i, \rho, \underline{r}, \alpha_{j(i)}, \tau, I_i)))]^{D_i=0} \quad (4)\end{aligned}$$

where $p_a(\cdot)$ is a function $p_a : [0, 1] \rightarrow [0, 1]$ to be specified.

The probability of success depends on:

- credit officer ability,
- his portfolio composition: gender, age, family situation of the client and weather he has a bank account, business sector,
- economic and geographical context: Vivacred branch and year of attribution of the credit,
- variables of comparison between the client and credit officer's demographic characteristics to get a potential "discrimination effect".

Except the credit officer ability (α), all these variables are denoted as x.

Even if Vivacred has clear rules of selection, the credit officer ability can make a difference and has to be distinguished from his observable portfolio composition effect. A credit officer can select systematically clients with "better" observable characteristics and then get a better result. We do not get it with the estimation of α's but with b.

For the probability of detecting the true state of nature, we consider a functional form for $p_a(\cdot)$ which allows it to be positively or negatively related to the credit officer's ability at the selection stage. We take $p_a(\alpha)$ as a linear function of α with the restriction of having values in the interval $[0, 1]$. Thus,

$$p_a(\alpha) = \begin{cases} 0, & \text{if } \alpha < -\frac{\bar{p}}{\mu_a}; \\ \mu_a \alpha + \bar{p}, & \text{if } -\frac{\bar{p}}{\mu_a} \leq \alpha \leq \frac{1-\bar{p}}{\mu_a}; \\ 1, & \text{if } \alpha > \frac{1-\bar{p}}{\mu_a}. \end{cases} \tag{5}$$

There are three other issues to be considered. First, $\underline{r}$ is not observed for us (although it is for the client) and can vary from a client to another. Then, we consider $\underline{r} \in \{\underline{r}_1, ..., \underline{r}_K\}$ ($\underline{r} < \rho$) with a probability of $1/K$ in each possibility and integrate the parameter out of the likelihood function.

Second, although we have data on monthly extra income (EI) and business profit (BP), we do not know how much can be obtained by the lender in case of default. We define the income "available" as a proportion of the monthly extra income and of the business profit, both multiplied by the number of installments (m). The proportion of these two sources of income are not necessarily the same. $I = \beta_{EI} \cdot (EI \cdot m) + \beta_{BP} \cdot (BP \cdot m)$. We estimate these two coefficients in the likelihood maximization. Substituting (5) into (4) and considering the distributions of $\underline{r}$ and I, the estimation problem is given by:

$$\begin{array}{cc} \max\limits_{\substack{\{\alpha_j\}_{j=1}^{M},\mu_s,\mu_a,\bar{p},\\ k_1,k_2,\beta_{EI},\beta_{BP},b}} & \mathbb{E}_{\underline{r}}\mathcal{L}(\{\alpha_j\}_{j=1}^{M},\mu_s,\mu_a,\bar{p},k_1,k_2,\beta_{EI},\beta_{BP},b,\underline{r}|\mathbf{D},\mathbf{L},\mathbf{EI},\mathbf{BP},\mathbf{x},\rho,\tau) \\ \text{subject to} & \begin{array}{c} 0 \leq \mu_a \cdot \alpha_j + \bar{p} \leq 1 \\ \beta_{EI},\beta_{BP} \in [0,1] \\ k_1,k_2 \in \mathbb{R}^{+*} \end{array} \end{array}$$

When the credit officer worked in two different branches, we split the officer's dummy between the main and the other branch he worked and then estimate not one but two α's as if there were two distinct credit officers.

Vivacred offers a possibility to renegotiate the contract (more installments of smaller value each) when the client shows not to be able to repay according to the initial conditions. Then, a new credit contract is made. As 30 days of delay is our measure of loan performance, this tool can be an issue to compare officers. Some of them renegotiate more often than others: between 0% and 18% of initial portfolio. It may create a bias depending whether the loan repayment is made before or after the 30 days' delay.

To address this problem, we aggregate refinanced and refinancing contracts. We cumulate the days of delay between the two contracts, and consider the initial loan size and installments. The client characteristics are the same (except the income that is averaged between the two loans). When the credit officer changed, we assign the aggregated information for both (shared responsibility). This case is relatively rare.

Finally, we create a dummy indicating if the observation result from such an aggregation and include it in the probability of success (as an extra x in xb). The estimation results presented in the paper, are based on the data with aggregation (29,154 observations). The same estimation without aggregation treatment (31,692 observations) is in appendix. The results are quite similar.

5 Credit officers matters more for selection than for state verification

Table 4: Maximum likelihood estimates (with aggregation)

Estimated α 's

α_1	-0.0029***	(0.1077)	α_{14}	0.4477	(0.1652)	α_{28}	0.0032***	(0.1008)
α_2	-1.3482***	(0.3069)	α_{15}	0.1070***	(0.0668)	α_{29}	-0.5872***	(0.0433)
α_3	0.1412***	(0.1272)	α_{16}	-0.2911***	(0.0896)	α_{30}	1	
α_4	0.0037***	(0.0408)	α_{17}	0.0801***	(0.0791)	α_{31OB}	-1.2030***	(0.4712)
α_5	0.1443***	(0.0765)	α_{18}	0.1055***	(0.0802)	α_{31MB}	-0.6899***	(0.0849)
α_6	-0.2641***	(0.0320)	α_{19}	0.1738***	(0.0856)	α_{32}	0.5789**	(0.0443)
α_7	0		α_{20}	0.9063	(0.2458)	α_{33}	0.7357	(0.1248)
α_8	-0.0885***	(0.0453)	α_{21}	0.3470	(0.1592)	α_{34}	0.4241	(0.2940)
α_9	0.2327***	(0.0853)	α_{22OB}	-0.7004***	(0.1004)	α_{35}	-0.7520***	(0.1101)
α_{10OB}	0.7798*	(0.1718)	α_{22BM}	0.1781**	(0.1667)	α_{36}	0.5142	(0.1197)
α_{10MB}	0.8781	(0.2093)	α_{23}	-0.0905***	(0.0412)	α_{37}	0.0817**	(0.1856)
α_{11OB}	-0.2921***	(0.0427)	α_{24}	-0.6230***	(0.0584)	α_{38}	-1.0082***	(0.0578)
α_{11MB}	0.0136***	(0.0654)	α_{25}	0.4907	(0.1903)	α_{39}	-0.4239***	(0.0679)
α_{12}	-0.6231***	(0.1247)	α_{26}	0.6416	(0.5877)	α_{40OB}	-0.1649***	(0.0335)
α_{13}	-0.6491***	(0.0141)	α_{27}	0.2912**	(0.1107)	α_{40MB}	-0.1965***	(0.0373)

Probability of misreporting : Audit, Available Income and ϕ Distribution

μ_a	0.2543***	(0.0064)	β_{RE}	0.8385***	(0.0010)	k_1	67.5	(3466394)
$\bar{p}$	0.5973***	(0.0047)	β_{RB}	0.0457***	(0.0000)	k_2	42.2	(1354164)

Probability of success parameters (b,μ_s)

Context: Sector, Guarantor and Current account and Year

μ_s	1.6739**	(0.5695)	Repay	-1.6464***	(0.0049)	2002	-2.4877***	(0.0204)
Trade	-0.0004	(0.0025)	1997	-1.7589***	(0.0724)	2003	-1.5630***	(0.0210)
Agric.	-0.0906	(0.0267)	1998	-1.7491***	(0.0385)	2004	-0.8885***	(0.0229)
Other	-0.9718***	(0.0209)	1999	-2.1459***	(0.0295)	2005	-1.0360***	(0.0205)
No Guar.	0.2158*	(0.0138)	2000	-2.5227***	(0.0236)	2006	-1.1265***	(0.0203)
C. Ac.	0.6661***	(0.0220)	2001	-2.6172***	(0.0232)	2007	0	

Comparing client / credit officer: Gender, Marital Status, Have Dependents, Age

Is a Female			Is Married			Has Dependent(s)		
Yes/Yes	-0.2762	(0.0689)	No/No	-2.2440***	(0.1129)	No/No	0.7132**	(0.0990)
Yes/No	0.1211*	(0.0050)	Yes/No	-1.8885***	(0.1136)	Yes/No	0.8149***	(0.0992)
No/Yes	-0.4483*	(0.0683)	No/Yes	-0.5385***	(0.0056)	No/Yes	-0.0934	(0.0048)
Age	0.0227***	(0.00001)	< 10 y.	0.0440	(0.0037)	Cons.	4.2424***	(0.3173)

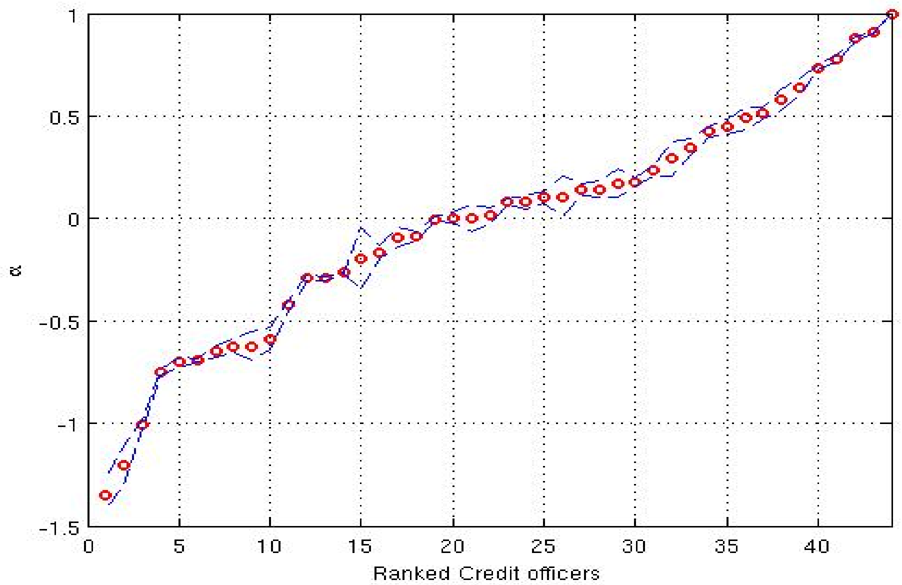
Wald test: α's are not compared to 0 but to the highest α (= 1).

*** p<0.01, ** p<0.05, * p<0.1.; Standard errors in parenthesis.

Tables 4 presents the officer's ability (α), probability of success parameters (b, μ_s) and probability of misreporting ones, namely, (μ_a, $\bar{p}$) related to the detection of misreporting, (k_1, k_2) for ϕ distribution parameters, and (β_{EI}, β_{BP}) defining the "available" income.

The estimated α's have an average of -0.016, and ranges from -1.35 to 1[2]. Figure 4 presents the estimated α's, ranked from the lowest to the highest.

Figure 4: Estimated α's.



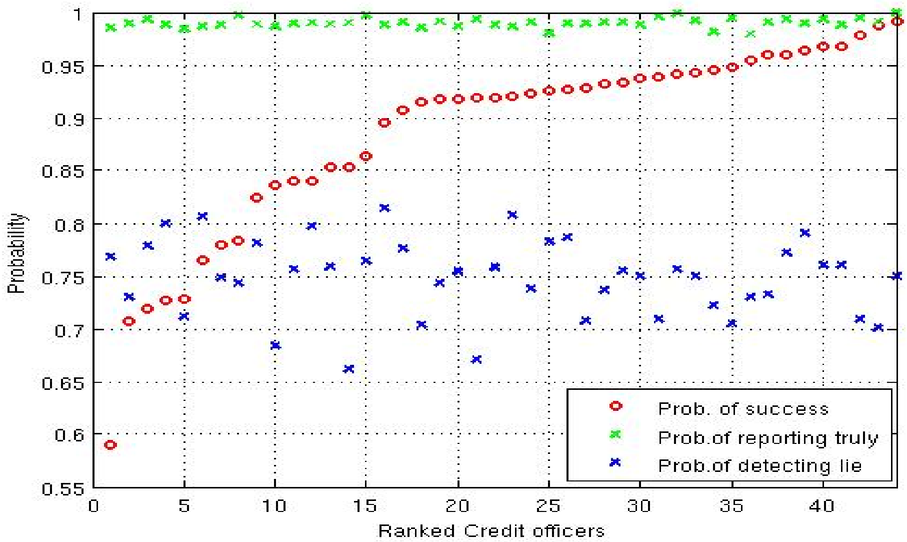
The probability of detecting misreporting depends on the credit officer's ability $\mu_a \cdot \alpha$ and on fixed part $\bar{p}$. The probability of success depends on the credit officers ability $\mu_s \cdot \alpha$ and their portfolios composition xb (including a constant). The estimated values related to the two stages of credit cycle are respectively 0.25 (μ_a), 0.59 ($\bar{p}$), 1.67 (μ_s) and 2.71 for the average xb. Thus, credit officers' ability (α) is making more difference in the probability

[2]As the α's are measured in relative terms (and are estimated in linear combination including a constant), we need to remove an officer's dummy (restrict an α to 0), to be able to identify the 44 other ones. Furthermore, in order to determine the relative weight of officer's ability in selection and audit probability, the α's are respectively multiplied by μ_s and μ_a. Thus, we need to fix another α to 1 for identification. α_{30} was chosen after numerous estimations pointing at it as the highest α.

of success than in the probability of detecting misreporting[3]. Moreover, the signal of μ_a and μ_s are not restricted allowing a credit officer to be efficient at one stage and not at the other (selection vs. audit). As both are positive, credit officers are ranked in the same order in both stages.

Nonetheless, α is not the only element explaining the heterogeneity of probability of delay among credit officers. Figure 5 presents the components of the delay probability ranked by the credit officer's average probability of success. These components, seen in the model, are the probability of success $(P_s(\mu_s, \alpha, x, b))$ and the probability for the client of reporting truly the state of nature $(F(g(\alpha, ...)))$ which includes the probability for the credit officer to detect misreporting at the audit $(\mu_a\alpha + \bar{p})$.

Figure 5: Probability of delay components



This is a way to illustrate why it is interesting to construct a model and not only estimate a probit or logit model of the probability of delay. It is very clear that credit officers are not ranked on the same (or even opposite) order

[3] $0.25/0.59 = 0.4237$ and $1.67/2.71 = 0.6162$

at the two stages. Moreover, even if the probability of detecting misreporting varies between 66% and 82%, this does not seem to make any difference for the probability of misreporting (97.8% to 99.6%). The probability of success is more variable, as it ranges from 58.9% to 99.8%. Credit officers matters much more at the selection stage than at the audit stage.

To compute the figure 5, we have decomposed the overall probability of delay in failure and misreporting (see equation 3). For each contract, we have predicted probabilities based on estimated parameters. On average, the probability of delay is 9.03% (close to the actual proportion 8.6%), the probability of success is 91.81% and the probability of misreporting is 0.93%. Thus, 10.63%[4] of the probability of delay is explained by misreporting. This illustrates the importance of officers expertise in auditing.

Back to table 4, the "available" income related to parameters β_{RE} and β_{RB} are respectively 0.83 and 0.04. In case of project failure, the income considered as "available" by the institution is composed almost exclusively of extra income. The potential income generated by the firm, before the investment, is practically neglected.

To conclude on the probability of misreporting, the estimation of ϕ distribution parameters k_1 and k_2 needs some clarifications. Their estimates, respectively 67.5 and 42.2, are not precise (see variance value). First, a multitude of possible couples (k_1, k_2) can be the solution to the log-likelihood maximization: Fixing one parameter, the optimal other parameter is always proportional to the first one. Optimization results exhibits an "optimal diagonal". We are able to improve the objective function increasing the two parameters together. But, after a certain threshold, increasing these two

[4] 0.9181*0.0093/0.0803=0.10633

parameters does not alter the objective function anymore (flat region). Second, a deeper inspection of the likelihood reveals that the distribution of ϕ matters only through the fraction of individuals with incentive to misreport, i.e., with $\phi < g(L, \rho, \underline{r}, \alpha, \tau, I)$. Parameters k_1 and k_2 are not completely identified.

Figure 6: ϕ and the misreporting threshold distributions.

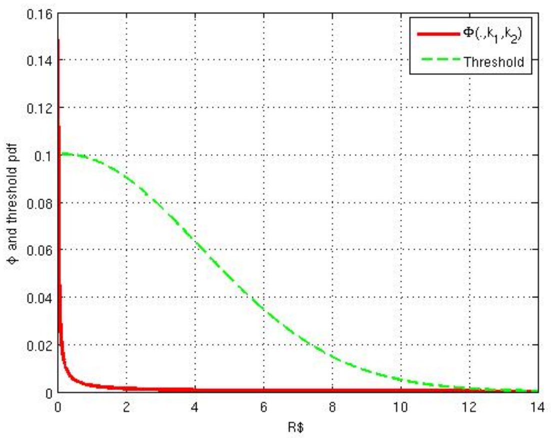


Figure 7: Income, debt and reputation cost distributions.

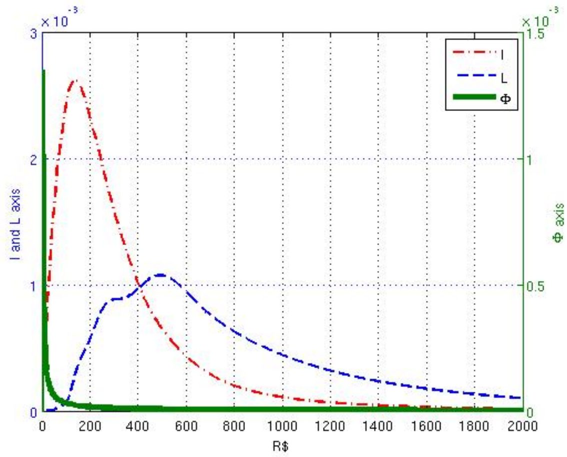


Figure 6 illustrates this point, presenting the distribution of the ϕ's and the thresholds to which it is compared to determine if misreporting is worthwhile or not. The distribution is mainly concentrated in 0 meanwhile the threshold take bigger values at the beginning of the distribution. What is important is the proportion of clients with almost no cost of lying. These clients will misreport. The other ones have an important cost of misreporting and will repay on time. Figure 7 depicts the estimated distribution of the reputation cost ϕ, to be compared to the empirical distribution of "available" income I and loan size L.

The probability of success ($p_s(\mu_s, \alpha, x, b)$) depends both on estimated α's (one or two for each credit officer) and a combination of observables (xb). Vector

x is composed of[5]: the sector ("services"), Year ("2007"), whether the client has a current account, lack of guarantor, and the set of comparison dummies between the client and the credit officer: gender (both male), marital status (both married), whether they have dependents (both with), client's age and a dummy indicating if the client-officer age difference is less than 10 years. Additionally, we include the aggregation dummy (renegotiate loan installments is the first signal of difficulties in the project).

The two main sectors attended by Vivacred are services and retail trade concerning respectively 42% and 53% of the loans. These two main sectors are not significantly different. Thus, the activity does not seem to matter in terms of probability of success. Moreover, as we could expect, renegotiated loans are less successful. The year dummies parameters are all significants and show an interesting evolution. From 1997 to 2001, the selection is getting worse and worse as the parameters are going down, but the trend reverts later (2001 to 2007). These dummies can capture the macro economic context or suggest a learning-by-doing effect.

The absence of a guarantor in the credit contract is not significant for the probability of success. Contracts without guarantor represent less than 10% of the sample and are allowed only for small values and well known clients (explaining the negative signal meanwhile not significant). In the model, the guarantor influence is implicitly included in the subjective cost of lying and thus appears in the misreporting condition. It could be tempting to introduce explicitly the guarantor dummy in the probability of misreporting but it would be redundant.

The probability of success is significantly higher when a client holds a current

[5]Omitted dummy is in bracket in case of qualitative variable.

account (6.7% of the loans). Current account ownership does not seem to sustain the credit access theory that would predict a more virtuous behavior for client with less alternative of funding. This variable may be a poor proxy for credit access or express an income effect. The average extra income and consumption are respectively R$454 and R$930 if the client has a current account and R$333 and R$808 otherwise.

The question of matching is not our central focus here but, we found important to control for it, not only comparing the clients and credit officers' gender like in Carter et al. (2007), but in a more general manner. Dummies comparing clients and credit officers' marital status, presence of dependents in the household and whether client-officer age difference is smaller than 10 years, are included in the probability of success. These dimensions are significant except the age difference. Nonetheless, older clients are more successful than younger. Even controlling for all of those dimensions, the probability of success vary between credit officers[6]. The α's are widely spread and the majority is significantly different to the maximum one.

6 Vivacred-specific experience vs previous skills

We have estimated the credit officers' ability (α) and found it more relevant in the selection stage than in the audit stage. In the present section, we are mainly interested in the relationship between credit officers' skills and their ability (α). Thus, to take in account their experience, we include alternatively, in the regressions of α, the age at entrance in Vivacred, the professional experience length (years) before Vivacred, the number of months

[6]Note that significance levels reported in table 4 are not in comparison to 0 but to the maximum α fixed to 1. Significance test from 0, would be a nonsense as 0 refers to an ordinary credit officer.

worked in Vivacred and finally the number of credits attended in the institution. Additionally, we control for three other potentially important skills dimensions: whether the credit officer is a Favela resident (field knowledge), whether he/she has a superior grade (education), whether he/she was hired to handle branch attendance tasks before being a credit officer (institution knowledge). Table 5 presents the OLS regressions of α on officers' observable characteristics.

We take into account the experience acquired before being hired by Vivacred, including alternatively the number of year of professional experience and the age at which the officer was hired. The two measures of previous experience are not significant, neither the education dummy (superior grade). To interpret this lack of significance, let's remember that α is the residual effect in the probability of success, as we have already taken into account the observable composition of the portfolio in xb. Thus, previous experience may be related to a better choice of observable characteristics rather than "subjective" (unobservable) attributes of the client.

To take into account the experience in Vivacred, we use alternatively the number of months worked for the institution and the number of contracts attended (in log). Additionally, living in a favela could be an advantage. Nonetheless, it is not significant. Maybe the relevant information would be to live in the *same* favela that the client, involving a complex social relationship theory that is not our point here. The Vivacred-specific experience appears to be a trait that matters for ability (significant in both cases).

The ability level estimates, on one hand, the subjective part of the selection efficiency and, on the other hand, the ability to figure out when a client is misreporting when an audit occurs. Thus, it makes completely sens that

Table 5: OLS Regressions: α on credit officers' profile

	(1)	(2)	(3)	(4)
Experience	Age at entrance	Year of exp.	Months in VC	# credits
measured as ...	-0.00367	0.00467	0.00524**	0.154***
	(0.0106)	(0.0144)	(0.00241)	(0.0556)
Favela resident	0.104	0.0856	-0.0886	-0.0384
	(0.153)	(0.179)	(0.170)	(0.148)
Superior Grade	0.0642	0.0744	-0.129	-0.0396
	(0.155)	(0.182)	(0.171)	(0.146)
Adm. tasks before	0.0313	0.120	0.0231	0.0981
	(0.220)	(0.243)	(0.192)	(0.186)
Constant	0.246	0.116	0.123	-2.217**
	(0.347)	(0.261)	(0.174)	(0.867)
Observations	45	40	45	45
R^2	0.556	0.424	0.605	0.631

Controls : Gender, Marital status, Has children.
Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

specific but not previous experience matters. This result may reflect a pure selection effect, in which only the most perceptive officers would remain working in Vivacred. But it may also suggest that the officers' task is subject to a learning-by-doing process. Unfortunately, we do not have the means of differentiating among the two potential effects.

6.1 Does the branch environment matter?

As the field experience seems to be important, we examine now if a branch switching matters for a credit officer's ability. To do so, we have estimated two α for each credit officer who worked in two branches (one α for the main branch and another one for the other branch). The main branch is the one where the credit officer handled the higher number of loans.

Table 6 recalls, for the five credit officers who switch, the ability (α) estimated for both branches, and presents their rank among the 45 estimated α. Table 6 depicts as well the number of loans handled in each branch and indicates which was the first workplace.

Table 6: Is the credit officer ability the same between branches?

C.O. id.	**# credits**		**Estimated α**		**α's Rank**		**Test**	
	Other	**Main**	**Other**	**Main**	**Other**	**Main**	χ^2	
10	233	322^+	0.78	0.88	4	3	0.53	
11	125^+	248	-0.29	0.01	34	23	1.91	
22	222	401^+	-0.7	0.18	41	15	3.85	**
31	253	283^+	-1.2	-0.69	44	40	0.33	
40	688^+	885	-0.16	-0.20	30	31	0.06	

$^+$ Branch where the credit officer began to work.

The branch switch has heterogeneous effect among credit officers. The first and last officer have the same ability among branches as the two estimated ability (α) are very close and successive in the ranking. In the three other cases, credit officers rank is better where they handled more loans whatever they began at that branch or not. This result seems to indicate than if learning process is at stake, it is at least partially, branch specific.

It is important to recall that this ranking is about the credit officers ability and not about their portfolio composition. Thus, if a credit officer is in a better area (safer clients' characteristics), his average delay can be lower than for another officer with a better estimated α working in a worse area. It is the case, as well, for a credit officer moving from a branch to another.

7 Simulations: Eliminating heterogeneity

Maximum delay above 30 days concerns 8.6% of our sample. As we have seen, loan performance is associated both to credit officers' ability and their portfolio composition (client demographic characteristics, available income and loan value). In this section, we use the model to simulate what would be the overall delay by fixing one of these elements and letting the others vary.

Figure 8: Simulated overall delay for homogeneous ability.

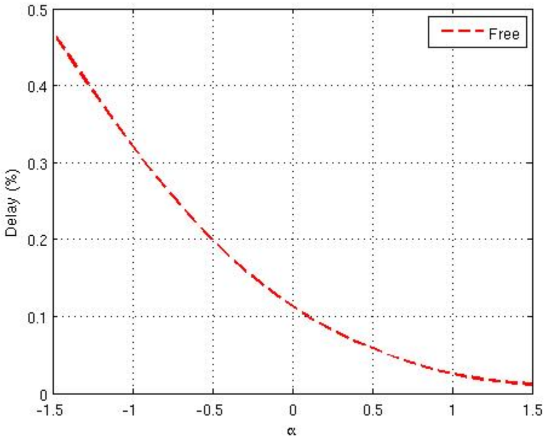
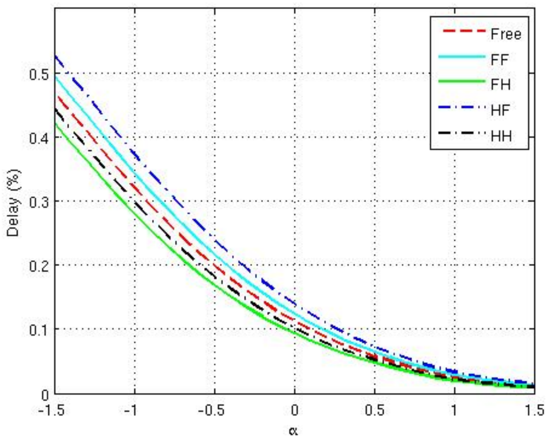


Figure 9: Simulated delay fixing client and c.o. gender.



First, we simulate what would be the overall delay if all the credit officers had the same ability (α), considering different values of this homogeneous in the $(\alpha_{min}, \alpha_{max})$ interval and maintaining the clients' heterogeneity observed in the data. Figure 8 summarizes this exercise. The horizontal axis represents the α level, while the vertical axis depicts the corresponding simulated overall delay. According to this simulation, the proportion of loans with a delay above 30 days is highly affected by the credit officers' ability. Considering the range of abilities estimated with our data, simulated delay varies from 2% (for $\alpha_{max} = 1$) to 41% (for $\alpha_{min} = -1.3$).

Second, we do the same exercise fixing some client characteristics. More specifically, we simulate the overall delay considering alternatively, for all the loans, the following clients' and credit officers' genders combinations: both male, both female, one male and one female. Figure 9 depicts this second exercise. The gender combination has a small impact compared to the α concerning the variation of delay generated. The credit officer's ability seems to be more relevant to explain the delay than client's and officer's gender.

Figure 10: Estimated b's related to contract year.

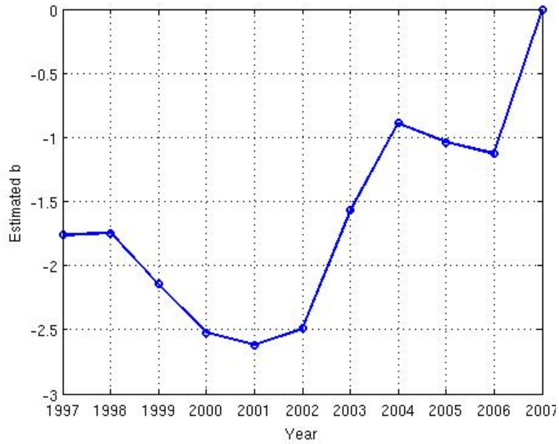


Figure 11: Simulated delay fixing the year of beginning.

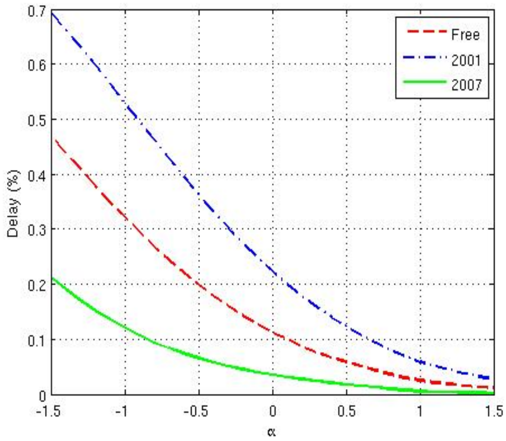


Figure 12: Average income and loan value by c.o.portfolio.

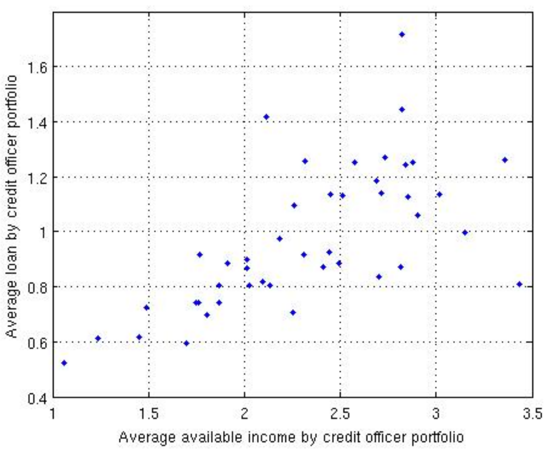
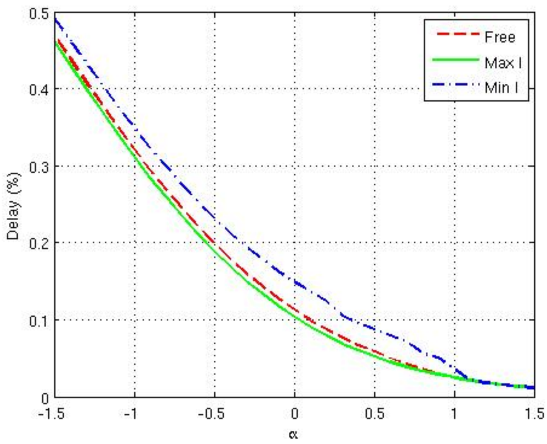


Figure 13: Simulated delay fixing the income.



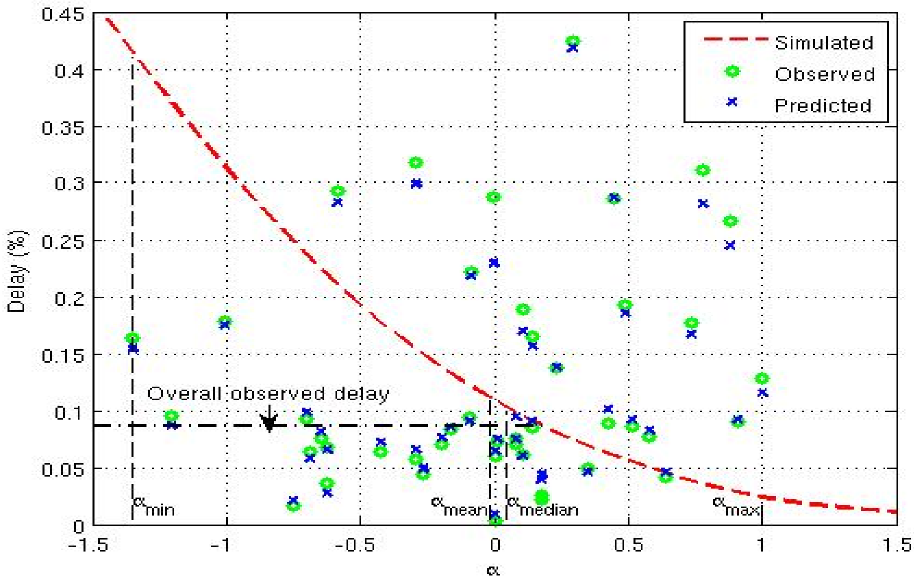
Third, we consider that all credits were made the same year. Back to table 4, the estimated parameters related to year's contract (included in vector b) are more widely ranged than the other b's components in the probability of success. Figure 10 presents the evolution such parameters. Year 2001 being the worst and 2007 the best, we choose this two extremes for the simulations. Figure 11 presents these simulations.

As we could expect, curve shifting is bigger than for the gender combinations exercise. Considering all the loans would be attended by the credit officer with the lower α (-1.3), the overall delay would be 70% if all credits would have been made in 2001 and 21% in 2007. Assuming all the loans would be attended by the credit officer with the higher α (1), the delay would be respectively 0.5% and 6%. Hence, the credit officer ability seem to have a smaller impact in more recent years. It is difficult to separate the learning-by-doing effect interpretation to the conjuncture one, but this figure tend to show some evidence for the first one. It seems that the institution as a whole, has improved its methodology turning the credit officer ability less relevant.

Fourth, we do the same exercise considering that all clients have the same available income. Figure 12 illustrates the diversity of credit officers' portfolio. Each point represents the average available income (horizontal axis) and loan value (vertical axis) of a credit officer's portfolio. Figure 13 depicts the simulation exercise, considering the worst situation (all the clients with the lower income) and the best one (all with the highest income). We would expect a bigger impact on the curve shifting. Nonetheless, as we have seen, selection matters more than audit and income is involved only as a kind of collateral (in the misreporting part of the story).

Finally, we predict what would be the overall delay for each actual credit officer using the estimated parameters. Then, we are able to compare this prediction to the first simulation exercise (varying an homogeneous officer's ability) and to the actual proportion of loans with a delay above 30 days for each credit officer portfolio. Figure 14 summarizes this last exercise. In

Figure 14: Simulated and observed delay by homogeneous α



practice, ability and portfolio composition varies together from a credit officer to another. Thus, the delay is not monotonic on α. Considering the range of estimated abilities, simulated delay vary from 2% to 41% while the observed and predicted delay vary from 0,5% to 42%. The spread around the red curve allow us to conclude that there is a distinct effect of the actual (subjective) ability and the portfolio composition effect on loan performance.

In practical terms, simulations allows to measure the effect of an intervention that would affect the average ability of a credit officers team. The α's vary in a range of 2.3 (to -1.3 to 1). For example, an increase in 10% of this range on the ability of Vivacred officers (+0.23 for each α) would reduce the expected delay of loans from 8.61% to 7% (a reduction of 10% would increase the delay to 11.5%). As the ability is related to the Vivacred-specific experience, there is potentially a substantial gain for the institution in reducing its officers turnover. The standard methodology (analysis questionnaire and enforcement rules) is not everything!

7.1 Discussion about identification

Only a partial information about repayment (delay above 30 days) has been used is this paper. However, the database contains a more detailed information which is the maximum number of days of delay for each loan. Thus, any delay threshold could be used to identify the credit officer's ability in the two stages: selection and audit.

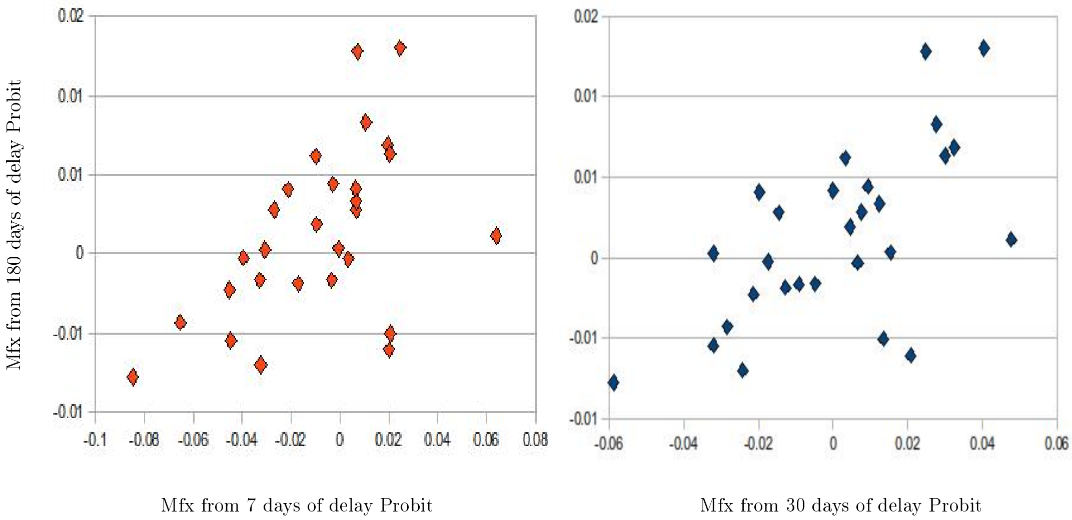
One could argue that an early delay (7 days) is a first signal for possible project failure before the audit take place. Thus, this event give an information about the selection made by the credit officer. Delay above 30 days is already a mix between selection and audit process as the credit officer already visited the client to understand what is happening. However the main loss of bonus for the credit officer occurs above 30 days of delay. Finally, the 180 days of delay represents the end of the process. Even if loan recovery may happen after this delay, this threshold determine the loss and imply the strongest penalty in the credit officer's bonus. Thus, default dummy is informative about the audit phase and the capacity for a credit officer to deal with bad payers.

In order to disentangle credit officer's ability in selection and audit phase, we can compare estimated coefficients associated to credit officers dummies in a probit regression of early delay (7 days) and default (180 days). If the coefficients are alike in the two regressions, thus, credit officers matters mainly in the selection. Their impact on repayment conditionally to observe a delay is homogeneous (audit methodology has the same effect for everyone). The second phase does not change ranking or distance among credit officers.

on the contrary, if credit officers are differently ranked in delay and default regression, they have an heterogeneous impact on the repayment phase.

We run three probit regressions: 7, 30 and 180 days of delay dummies on the credit officers dummies and all the controls[7] discussed in this paper. Figure 15 represents the marginal effects associated to credit officers' dummy coefficients from early delay (7 days) against those from default (180 days) on the left and delay (30 days) against default on the right.

Figure 15: Marginal effects (Mfx) associated to officers' dummy coefficients from probit regressions of 7, 30 and 180 days of delay.



Both graphics show points reasonably aligned crosswise[8]. Thus, credit officers ranking does not change so much between the two delay (7 and 30 days) regressions and the default regression. As the delay is much more likely to be affected by selection than default, results found estimating the structural model estimation are not invalidate by the probit regressions.

[7]Controls coefficients are not reported here as only officers' dummies coefficients are of interest in this section.

[8]Two points representing few observations were removed from figure 15 because of their very high value in the default probit regression.

Comparing figure 15 to figure 3 in descriptive statistics, allows us to measure the importance of each credit officer portfolio composition. Conditionally to their portfolio composition, credit officers has an heterogeneous impact on delay but not so much on default. Default does not depend so much on the credit officer identity while delay continue to be widespread among them after controlling for the portfolio composition. At a first glance, default could seem to be sufficient to evaluate the credit officer's ability. However, as the state verification is costly in time, delay matters for the institution efficiency.

8 Conclusion

This paper focuses on the role of the credit officer as an important way to deal with information asymmetry issue. It presents a structural model of credit provision with costly verification state. The credit officer ability is considered explicitly both in the probability of success of the project (adverse selection problem) and in the probability of misreporting (moral hazard problem).

Thanks to a wealthy individual dataset provided by the Brazilian NGO Vivacred, we are able to estimate such a model. The results emphasize a substantial heterogeneity in ability among credit officers mainly at the selection stage. Moreover, it is interesting to notice that officers' ranking is completely different at the two stages due to the portfolio composition. Such a result confirms the relevance of the use of a structural model.

Both credit officer's ability and portfolio composition have a significant effect on the probability of success. Concerning the former, gender and family compositions seem to matter, while client-officer age difference does not. Different from other studies, we find evidence neither that outcomes between female and male credit officers are different; nor that clients who are similar to their credit officer are privileged.

Furthermore, the experience prior to joining and at other function in Vivacred (before being an officer) are not related to the estimated ability, whereas Vivacred-specific experience as a credit officer is.

Beyond this conclusion, two more questions remain to be investigated. First, as we have showed before the probability of success depends strongly on the year dummies and the credit officer's ability is sometime varying through branches. It would be interesting then to know if these two facts are due to a learning-by-doing process and/or to the macroeconomic context. Second, as the credit officer is important for the selection process, the existence of group-specific attendance needs to be explored and, should the case arise, its legitimacy.

References

Armendáriz, B. and Morduch, J. (2000). Microfinance beyond group lending. *Economics of Transition*, 8(2):401–420.

Banerjee, A. V., Besley, T., and Guinnane, T. W. (1994). Thy neighbor's keeper: The design of a credit cooperative with theory and a test. *The Quarterly Journal of Economics*, 109(2):491–515.

Beck, T., Behr, P., and Guttler, A. (2009). Gender and banking: Are women better loan officers? *working paper*.

Berger, A. N., Frame, W. S., and Miller, N. H. (2005a). Credit scoring and the availability, price, and risk of small business credit. *Journal of Money, Credit and Banking*, 37(2):191–222.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., and Stein, J. C. (2005b). Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial Economics*, 76(2):237–269.

Berger, A. N. and Udell, G. F. (1995). Relationship lending and lines of credit in small firm finance. *Journal of Business*, 68(3):351–81.

Besley, T. and Coate, S. (1995). Group lending, repayment incentives and social collateral. *Journal of Development Economics*, 46(1):1–18.

Bester, H. and Hellwig, M. (1987). Moral hazard and equilibrium credit rationing: An overview of the issues. Discussion Paper Serie A 125, University of Bonn, Germany.

Boot, A. W. A. and Thakor, A. V. (1994). Moral hazard and secured lending in an infinitely repeated credit market game. *International Economic Review*, 35(4):899–920.

Boot, A. W. A. and Thakor, A. V. (2000). Can relationship banking survive competition? *Journal of Finance*, 55(2):679–713.

Carrasco, V. and de Mello, J. M. P. (2006). A relational theory of relationship lending under contractual incompleteness. Working Paper 520, Department of Economics PUC-Rio (Brazil).

Carter, S., Shaw, E., Lam, W., and Wilson, F. (2007). Gender, entrepreneurship, and bank lending: the criteria and processes used by bank loan officers in assessing applications. *Entrepreneurship Theory and Practice*, 31(3):427–444.

Gale, D. and Hellwig, M. (1985). Incentive-Compatible debt contracts: The One-Period problem. *Review of Economic Studies*, 52(4):647–63.

Ghatak, M. (1999). Group lending, local information and peer selection. *Journal of Development Economics*, 60(1):27–50.

Hermes, N. and Lensink, R. (2007). The empirics of microfinance: what do we know? *Economic Journal*, 117(517):F1–F10.

Labie, M., Méon, P., Mersland, R., and Szafarz, A. (2010). Discrimination by microcredit officers: Theory and evidence on disability in uganda. CEB Working Papers 10-007, Université Libre de Bruxelles.

Morduch, J. (1999). The microfinance promise. *Journal of Economic Literature*, 37(4):1569–1614.

Petersen, M. A. and Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance*, 49(1):3–37.

Petersen, M. A. and Rajan, R. G. (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443.

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570.

Stiglitz, J. E. (1990). Peer monitoring and credit markets. *World Bank Economic Review*, 4(3):351–66.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265 – 293.

A Computing the gradient

The problem to be solved is:

$$\max_{\substack{\{\alpha_j\}_{j=1}^M,\mu_s,\mu_a,\bar{p},\\ k_1,k_2,\beta_{EI},\beta_{BP},b}} \mathbb{E}_{\underline{r}}\mathcal{L}(\{\alpha_j\}_{j=1}^M,\mu_s,\mu_a,\bar{p},k_1,k_2,\beta_{EI},\beta_{BP},b,\underline{r}|\mathbf{D},\mathbf{L},\mathbf{EI},\mathbf{BP},\mathbf{x},\rho,\tau)$$

$$\text{subject to} \quad \begin{array}{c} 0 \le \mu_a \cdot \alpha_j + \bar{p} \le 1 \\ \beta_{EI},\beta_{BP} \in [0,1] \\ k_1,k_2 \in \mathbb{R}^{+*} \end{array}$$

Tacking the log-likelihood function as:

$$\ln \mathcal{L}(\{\alpha_j\}_{j=1}^M,\mu_s,\mu_a,\bar{p},k_1,k_2,\beta_{EI},\beta_{BP},b,\underline{r}|\mathbf{D},\mathbf{L},\mathbf{EI},\mathbf{BP},\mathbf{x},\rho,\tau) = \sum_{i=1}^N \ln(Pr(D=1)) \cdot \mathbb{I}_{D_i=1} + \ln(Pr(D=0)) \cdot \mathbb{I}_{D_i=0}$$

with $Pr(D=1) = 1-(\frac{1}{1+e^{-(\mu_s\cdot\alpha_{j(i)}+x_ib)}})[1-F(g(L_i,\rho,\underline{r},\alpha_{j(i)},\mu_a,\bar{p},\tau,I_i),k_1,k_2)]$,

$g(L,\rho,\underline{r},\alpha,\mu_a,\bar{p},\tau,I) \equiv \frac{[\rho-\underline{r}-(\mu_a\cdot\alpha+\bar{p})(\rho-\underline{r}+\tau)]L-(1-\mu_a\cdot\alpha-\bar{p})\min(I,(\rho-\underline{r})L)}{(\mu_a\cdot\alpha+\bar{p})}$,

$I = \beta_{EI} \cdot EI + \beta_{BP} \cdot BP$,

and ϕ has a log-normal distribution:

$$F(x,k_1,k_2) := \int^x \frac{\mathrm{e}^{-1/2\,\frac{(\ln(u)-k_1)^2}{k_2{}^2}}}{u\cdot k_2\sqrt{2\cdot\pi}}du; f(x,k_1,k_2) = \frac{\mathrm{e}^{-1/2\,\frac{(\ln(x)-k_1)^2}{k_2{}^2}}}{x\cdot k_2\sqrt{2\cdot\pi}}$$

Partial derivative of $\ln\mathcal{L}$ in $\theta \in \{\{\alpha_j\}_{j=1}^M,\mu_a,\bar{p},k_1,k_2,\beta_{EI},\beta_{BP},b\}$ is:

$$\frac{d\ln\mathcal{L}}{d\theta} = \mathbb{I}_{D_i=1}\cdot\frac{\frac{dPr(D=1)}{d\theta}}{Pr(D=1)} - \mathbb{I}_{D_i=0}\cdot\frac{\frac{dPr(D=1)}{d\theta}}{1-Pr(D=1)}$$

Partial derivatives of probability of success $p_s(\mu_s,\alpha,x,b) = \frac{1}{1+e^{-(\mu_s\cdot\alpha+xb)}}$ are:

$$\frac{dp_s(\mu_s,\alpha,x,b)}{d\mu_s} = \alpha\cdot\frac{e^{-(\mu_s\cdot\alpha+xb)}}{(1+e^{-(\mu_s\cdot\alpha+xb)})^2}$$

$$\frac{dp_s(\mu_s, \alpha, x, b)}{d\alpha} = \mu_s \cdot \frac{e^{-(\mu_s \cdot \alpha + xb)}}{(1 + e^{-(\mu_s \cdot \alpha + xb)})^2}$$

$$\frac{dp_s(\mu_s, \alpha, x, b)}{db} = x \cdot \frac{e^{-(\mu_s \cdot \alpha + xb)}}{(1 + e^{-(\mu_s \cdot \alpha + xb)})^2}$$

Partial derivatives of misreporting threshold $g(L, \rho, \underline{r}, \alpha, \mu_a, \bar{p}, \tau, I)$ are:

$$\frac{dg(L, \rho, \underline{r}, \alpha, \mu_a, \bar{p}, \tau, I)}{d\bar{p}} = \begin{cases} 0 & \text{if } I \geq L(\rho - \underline{r}) \\ \frac{I - L \cdot (\rho - \underline{r})}{(\mu_a \cdot \alpha + \bar{p})^2} & \text{if } I < L(\rho - \underline{r}) \end{cases}$$

$$\frac{dg(L, \rho, \underline{r}, \alpha, \mu_a, \bar{p}, \tau, I)}{d\alpha} = \mu_a \cdot \frac{dg(L, \rho, \underline{r}, \alpha, \mu_a, \bar{p}, \tau, I)}{d\bar{p}}$$

$$\frac{dg(L, \rho, \underline{r}, \alpha, \mu_a, \bar{p}, \tau, I)}{d\mu_a} = \alpha \cdot \frac{dg(L, \rho, \underline{r}, \alpha, \mu_a, \bar{p}, \tau, I)}{d\bar{p}}$$

Partial derivatives in μ_a and $\bar{p}$

The partial derivatives of the probability of delay in the probability of detecting misreporting parameters (μ_a and $\bar{p}$) are:

$$\frac{dPr(D = 1)}{d\mu_a} = p_s(\mu_s, \alpha_j, x, b) \cdot \mathbb{E}_{\underline{r}}[f(g(\mu_a, ...), k_1, k_2) \cdot \frac{dg(\mu_a, ...)}{d\mu_a}]$$

and

$$\frac{dPr(D = 1)}{d\bar{p}} = p_s(\mu_s, \alpha_j, x, b) \cdot \mathbb{E}_{\underline{r}}[f(g(\bar{p}, ...), k_1, k_2) \cdot \frac{dg(\bar{p}, ...)}{d\bar{p}}]$$

Partial derivative in α_1, ..., α_M

The partial derivatives of the probability of delay in the officers ability coefficient (α_1, ..., α_M) are:

$$\frac{dPr(D = 1)}{d\alpha_j} = -\frac{dp_s(\mu_s \cdot \alpha_j, x, b)}{d\alpha_j} \cdot (1 - \mathbb{E}_{\underline{r}} F(g(\alpha_j, ...), k_1, k_2)) +$$
$$p_s(\mu_s \cdot \alpha_j, x, b) \cdot \mathbb{E}_{\underline{r}}[f(g(\alpha_j, ...), k_1, k_2) \cdot \frac{dg(\alpha_j, ...)}{d\alpha_j}]$$

Partial derivative in k_1 and k_2

The partial derivatives of the probability of delay in the log-normal distribution parameters (k_1 and k_2) are:

$$\frac{dPr(D=1)}{dk_1} = p_s(\mu_s, \alpha_j, x, b) \cdot \mathbb{E}_{\underline{r}}[-\frac{\mathrm{e}^{-1/2\,\frac{(\ln(g(...))-k_1)^2}{{k_2}^2}}}{k_2\sqrt{2\cdot\pi}}]$$

and

$$\frac{dPr(D=1)}{dk_2} = p_s(\mu_s, \alpha_j, x, b) \cdot \mathbb{E}_{\underline{r}}[\frac{\mathrm{e}^{-1/2\,\frac{(\ln(g(...))-k_1)^2}{k_2^2}}}{k_2\sqrt{2\cdot\pi}} \cdot (\ln(g(...)) - k_1)]$$

Partial derivative in β_{EI} and β_{BP}

$$\frac{dPr(D=1)}{d\beta_n} = p_s(\mu_s, \alpha_j, x, b) \cdot \mathbb{E}_{\underline{r}}[f(g(I_i, ...), k_1, k_2) \cdot \mathbb{I}_{[I_i<(\rho-\underline{r})L_i]} \cdot \frac{(\mu_a \cdot \alpha_j + \bar{p} - 1)}{\mu_a \cdot \alpha_j + \bar{p}} \cdot n]$$

with $n = EI, BP$

Partial derivative in b and μ_s

$$\frac{dPr(D=1)}{dm} = (\mathbb{E}_{\underline{r}} F(g(L_i, \rho, \underline{r}, \alpha_j, \mu_a, \bar{p}, \tau, I_i), k_1, k_2) - 1) \cdot \frac{dp_s(\mu_s, \alpha_j, x, b)}{dm}$$

with $m = b, \mu_s$

B Maximum Likelihood without aggregation

Table 7: Maximum likelihood estimates (without aggregation)

Estimated α 's								
α_1	0.0002***	(0.0959)	α_{14}	0.3818	(0.2455)	α_{28}	-0.0507***	(0.0784)
α_2	-1.1437***	(0.2304)	α_{15}	0.1260**	(0.1332)	α_{29}	-0.4261***	(0.0391)
α_3	-0.0381***	(0.0953)	α_{16}	-0.1786***	(0.0906)	α_{30}	1	
α_4	0.0222***	(0.0290)	α_{17}	0.0896***	(0.0735)	α_{31OB}	-1.0277**	(0.7224)
α_5	0.1506***	(0.0902)	α_{18}	0.1324***	(0.0859)	α_{31MB}	-0.6214***	(0.0538)
α_6	-0.2067***	(0.0256)	α_{19}	0.2049**	(0.1091)	α_{32}	0.4076***	(0.0359)
α_7	0		α_{20}	0.7518	(0.2888)	α_{33}	0.5497	(0.0960)
α_8	-0.1009***	(0.0369)	α_{21}	0.4970	(0.3407)	α_{34}	0.3156	(0.2842)
α_9	0.1544***	(0.0953)	α_{22OB}	-0.6686***	(0.0488)	α_{35}	-0.6618***	(0.1042)
α_{10OB}	0.6840	(0.1878)	α_{22MB}	0.0642***	(0.1319)	α_{36}	0.4102*	(0.0995)
α_{10MB}	0.8109	(0.2536)	α_{23}	-0.1307***	(0.0300)	α_{37}	0.0015**	(0.1893)
α_{11OB}	-0.2542***	(0.0466)	α_{24}	-0.3412***	(0.1051)	α_{38}	-0.9503***	(0.0397)
α_{11MB}	-0.0075***	(0.0898)	α_{25}	0.4815	(0.2416)	α_{39}	-0.3939***	(0.0243)
α_{12}	-0.5469***	(0.0534)	α_{26}	0.3441	(0.4007)	α_{40OB}	-0.1586***	(0.0262)
α_{13}	-0.5881***	(0.0096)	α_{27}	0.2912*	(0.1315)	α_{40MB}	-0.1933***	(0.0273)
Probability of misreporting : Audit, Available Income and ϕ Distribution								
μ_a	0.2266**	(0.0101)	β_{RE}	0.8267***	(0.0614)	k_1	68.13	(3868197)
$\bar{p}$	0.6050***	(0.0084)	β_{RB}	0.0431***	(0.0000)	k_2	41.14	(1410992)
Probability of success parameters (b,μ_s)								
Context: Sector, Guarantor and Current account and Year								
μ_s	2.0890*	(1.5089)	Repay	-1.1298***	(0.0058)	2002	-2.1693***	(0.0175)
Trade	0.0301	(0.0024)	1997	-1.6295***	(0.0624)	2003	-1.2127***	(0.0179)
Agric.	-0.0081	(0.0274)	1998	-1.5671***	(0.0331)	2004	-0.5524***	(0.0201)
Other	-0.8295***	(0.0190)	1999	-1.9833***	(0.0270)	2005	-0.5895***	(0.0173)
No Guar.	0.1755	(0.0121)	2000	-2.2779***	(0.0202)	2006	-0.7499***	(0.0162)
C. Ac.	0.7057***	(0.0232)	2001	-2.3624***	(0.0196)	2007	0	
Comparing client / credit officer: Gender, Marital Status, Have Dependents, Age								
Is a Female			Is Married			Has Dependent(s)		
Yes/Yes	-0.2798	(0.0789)	No/No	-2.4118***	(0.2123)	No/No	0.7167*	(0.1680)
Yes/No	0.1561**	(0.0046)	Yes/No	-2.0459***	(0.2148)	Yes/No	0.8242**	(0.1676)
No/Yes	-0.4683*	(0.0789)	No/Yes	-0.5353***	(0.0054)	No/Yes	-0.0812	(0.0047)
Age	0.0229***	(0.00001)	$<$ 10 y.	0.0751	(0.0035)	Cons.	3.9424***	(0.4468)

Wald test: α's not compared to 0 but to the highest α (= 1).

*** p$<$0.01, ** p$<$0.05, * p$<$0.1.; Standard errors in parenthesis.