4Asymmetric Information in an Expanding Credit Market: Evidence from Brazilian Car Loan Market

4.1.Introduction

The Brazilian credit market had never taken off. After coping with an adverse economic environment for several years, only after 2004 the vehicle credit market flourished. The total credit supplied for car loans grew on average 25% per year between 2004 and 2010. The yearly total credit concession almost tripled during that period. Moreover, this credit expanding was accompanied by better contract terms. For instance, the average down payment required by banks decreased along the years. Reflecting this credit boom, the total car production increased from 1.9 million unit in 2003 to 3.6 in 2010. Brazil became the fifth-largest car producer manufacturer in the world.

A crucial explanation for this fact was the implementation of the Lei de Alienacão Fiduciária (Fiduciary Law), a legal reform that improved credit environment by turning the judicial system faster and more efficient. As Assunção, Benmelech, and Silva (2012) pointed out, after the implementation of the law, average contract terms improved - spread deceased, maturity and loan size increased. Furthermore, the population of borrowers changed. They documented that the average borrower income decreased, and the proportion of high-risk borrowers in the population increased. Under the previous environment, this revolution would never have been possible.

A side effect of this sudden boost in the vehicle credit market is worse borrower's per- formance. As banks expanded credit toward riskier borrowers, these newly consumers had a higher probability of default. Also, Assunção, Benmelech, and Silva (2012) show that, after the legal reform, the delinquency rate increased. We hypothesize that this movement carried an increase in asymmetric information between borrowers and lenders.

Our goal in this paper is to empirically test the hypothesis that the above mention boom in the vehicle credit market lead to an increased asymmetry information between borrowers and lenders. This analyzes required micro-data with detailed information about contract terms, personal characteristics, and borrower performance. Fortunately, we access a rich data set from one of the three largest Brazilian private banks.[1] Moreover, we investigate how The Bank tried to tackle this issue using observable variables. Basically, a bank can screen borrowers based on contract terms or personal and car characteristics. We test both situations.

We use a standard model of lifetime data to predict borrower's performance. Assuming reasonable parametric forms, we estimate the hazard function, which is defined as the probability that a default occurs at time t, conditional on default not occurring up to t. If the probability of borrower's default decreases along the loan maturity, it means that default is more likely at the beginning. We interpret this as a sign of asymmetric information issue. The Bank granted credit to a borrower who would never be able to pay all installments or who is not willing to pay his debt. Thus, in both cases, borrowers will, on average, default at the beginning. Conversely, if the hazard function is not decreasing along time, we interpret as a sign that asymmetric information is not a significant issue. An estimative of the model provides us with at least three results.

First, we collect evidence that asymmetric information increased during the sample pe- riod. When we do not take into account observable variables, the estimated hazard function decreased year after year. After 2007, in particular, it is always above one. As The Bank expanded credit, the asymmetric information issue worsened.

Second, using all information available, The Bank (at least partially) succeeded in mit- igating this problem. When we estimate the model taking into account all information available to The Bank, which means contract terms as well as personal and car character- istics, the hazard function is not below one anymore, except for 2006. Thus, after properly controlling for all observable characteristics, we do not observe a higher proportion of bor- rower's default in the beginning of the loan. The Bank partially mitigated the asymmetric information issue that arose from its credit expansion policy.

Third, we promoted a horse race between contract terms, and personal and car charac- teristics in order to evaluate which one does a better job helping The Bank in the task of mitigating the asymmetric information. We re-estimate the model twice in a comparative way. First, adding only contract terms. Second, adding only personal and car characteristics. Results clearly shown that contract terms do a better job in screening borrowers.

Our paper related with current literature in at least three dimensions. First, a growing body of papers relates contract terms and asymmetric information. For instance, Einav, Jenkins, and Levin (2011) analyze the role that down payment can play in screening borrow- ers, and Adams, Einav, and Levin (2009) used a non-linearity in contract terms in order to disentangle moral hazard from adverse selection. Second, there is a literature that analyzes consequences of a credit expanding. For example, Mian and Sufi (2009) investigates the consequences of the U.S. mortgage credit expansion. Keys et at. (2010) and Mian and Sufi (2009) suggest that credit expansion leads to a subsequent waves of default and repossession. We go a step further and document the mechanisms. A credit expansion is usually associated with riskier borrower, causing an increase in the informational problems. Third, our paper use a lifetime model similar to Meyer (1990) and Meyer (1995).

The paper is organized as followed. In the next section, we describe the economic environment in the sample period. In Section 3, we develop the model. In Sections 4 and 5, we take a glance at data set and then describe the empirical results. Finally, we conclude.

4.2.Economic Environment

Brazil presented a modest level of financial development, the credit market had never taken off until 2004. Adverse macroeconomic environment can give a partial explanation. The country copied with high inflation for several years. For instance, the yearly average inflation between 1985 and 1994 was 750% when the Real Plan succeeded in bringing down inflation to a lower level. Indeed, high inflation inhibited the development of the credit market as pointed out Boyd, Levine, and Smith (2000). Furthermore, even when inflation was relatively low, the credit market remained struggling. Consensus among specialists showed the urgent need for micro-reform. On this issue, Beck (2000) note:

"*Brazil is trapped in a low-intermediation state, exhibiting both a low level and low ef- ficiency of financial intermediation (...). Compared to other countries, Brazil shows a low level of contract enforcement and limited information sharing. The weaknesses in these components of the legal and regulatory framework can partly explain the low-intermediation trap in Brazil.*"

The government undertook several credit reform, with a view to mitigating inefficiencies and boosting the credit market. Among them, we can underscore the 'crédito consignado' (2004) and 'lei de aliençao fiduciária' (fiduciary law) (2004). 'Credito consignado' is a law that allows borrowers to repay their loans with installments debited directly deducted from the consumer's paycheck, thus allowing for a lower interest rates and higher amount of credit supplied by banks for personal loans (Coelho, Funchal and Mello (2012)).

The fiduciary law impacted the vehicle credit market and mortgage by removing judicial inefficiency. Banks no longer need to wait until the end of the trial; they are allowed to sell the car (collateral) with a court injunction. This small modification had a huge impact. It takes three weeks between the recovery and sele of the car in case of default, instead of the previous minimum of two years. Faster justice dynamizes the financial sector.

The law represented a turning point for the vehicle credit market. Assunçao, Benmelech, and Silva (2012) present several dimensions and we will underscore two of them. First, after the law, contract terms improved. Not only did interest rates decrease, but also loan maturity (on average) became longer, and loan size increased. Second, the law led to a 'democratization of credit.' The proportion of high-risk borrowers increased, and borrowers(on average) had a lower income.

The vehicle credit market witnessed a true revolution. The concession of credit has grown monotonically for the whole sample period. Not only was the average growth 25%, but also the minimum growth was 4%. The monthly concession of credit almost tripled from R$2.5 billion (US$865 million) on Jan/2004 to R$6 billions (US$3.5 billion) by the end of the sample. As a result, the total credit supplied increased pari passu from R$31 billion (US$10.73 billion) to R$90 billion (US$52.94 billion) for the same period. As it became clear, credit has grown consistently, instead of a discontinuity right after the law. We conjecture that it takes time for all players to figure out the new environment setting, such as understanding the new legislation and justice behavior, delinquency rate of new borrowers with different contract terms, etc. Thus, financial institutions only fully internalized all benefits along the years.

The vehicle credit became less concentrated among states. The three most important states used to concentrate two-thirds of the total contracts half[2]. Likewise, the credit became more disseminated in less economically important states. In 2003, half of the states which received less contracts from The Bank jointly represented 2.71% of the total number of contracts. In 2010, this category represented 11.9%.

Finally, and not surprising, the automobile industry grew at the same pace. The total auto sales accumulated in 12 months remained quite stable in the interval of 1.3 and 2 million vehicles during the 9 years that preceded the sample period. Nevertheless, this number jumped from 1.5 million to 3.6 million from the beginning to the end of our sample period, respectively. The total production also increased from 1.9 million per year to 3.6 million.

This boom in vehicle modality was part of a boom in the whole credit market. Figure 1d shows the spectacular evolution of it. Remarkable, after it had been flat around 25% for a while, the total credit to GDP ratio grew from a modest 24% in January 2004 to 45.3% in January 2010. The credit reforms combined with a better macroeconomic environment provided conditions for the referred revolution. On this respect, Veja, the most popular weekly news magazine in Brazil, highlighted:

"*this revolution, propelling the country towards a modern society of mass consumption, was only possible thanks to three pillars: institutional reforms, competition and political and economic stability*"

The yearly average of economic growth during the 7 years encompassing the sample period was 4.5%, which is remarkably higher than the 2.3% of the immediate previous 7 years. Financial crisis hit the economy only after the bankruptcy of Lehman Brothers, when it decelerated from 5.2% in 2008 to -0.2% in the following year. Nevertheless, as we previously mentioned, the credit kept growing even during 2009.

Reflecting this growth, the unemployment rate sharply drop from 11.7% on average in 2004 to 6.7% seven years later. According to the Brazilian Central Bank, the delinquency rate of the whole system (including all banks and all modalities for individuals) remained quite stable, almost always between the interval 6% - 8%. In particular, the delinquency rate for vehicles also remained stable between 1.5% and 3.5%. Inflation behaved relatively well. It fluctuated inside the interval of 3.5% and 7.6%, usually around the Central Banks target of 4.5%.

4.3.Model

First, we briefly review the standard theory concerning analysis of lifetime data. Then, based on the model, we derive some implications relating borrower's predicted performance and the asymmetric information issue.

Let T be a random variable representing the time elapsed until a consumer defaults on his auto loan, and t be a given realization of T. The distribution of T is completely characterized by the hazard function, given by

$$\lambda(t) = \lim_{\Delta \downarrow 0} \frac{\Pr(t \leq T < t + \Delta | T \geq t)}{\Delta} \qquad (1)$$

In fact, the cumulative distribution function of T, given by $F(t) = \Pr[T \leq t]$, can be calculated from the hazard function using the following expression:

$$F(t) = 1 - \exp\left(-\int_0^t \lambda(s)\,ds\right) \qquad (2)$$

Intuitively, the hazard function $\lambda(t)$ is the probability that a default occurs at time t, conditional on default not occurring up to t. The hazard is duration dependent if it changes with t; there is positive duration dependence when $\lambda(t)$ increases with t, and negative duration dependence when $\lambda(t)$ decreases with t.

In our environment, it is natural to assume that the distribution of T varies across individuals. Let x and y be respectively vectors of observed and unobserved individual characteristics affecting the default hazard rate. We assume that the hazard for individual i

at time t is given by:

$$\lambda_i(t|x_i, v_i) = \lambda_0(t)e^{x_i\beta + y_i\Psi} = v_i\lambda_0(t)e^{x_i\beta}. \quad (3)$$

All unobserved heterogeneity is captured by an individual-specific random effect v_i representing the individual's frailty that enters multiplicative in the above expression. An individual who has a high value of v has a higher default hazard and is therefore a less creditworthy borrower. The term $\lambda_0(t)$, that remains constant across individuals, is called the baseline hazard.

Let G(v) be the distribution function of the unobserved term v. At any time t the com- position of the pool of borrowers with respect to the unobserved heterogeneity is determined by

$$dG(v|T > t, x) = \frac{(1-F(t|x,v))dG(v)}{1-F(t|x)} \quad (4)$$

Therefore, the composition of the sample of survivors changes with the passing of time. Because individuals with higher v tend to default earlier, the pool of borrowers still honoring their credit contract at high durations tend to have lower values of v. This phenomenon, which has been called "weeding out", has implications on the population hazard rate which is observed (unconditional on unobserved characteristics). It is given by

$$\lambda(t, x) = \lambda_0(t)e^{x'\beta}E(v|T > t, x). \quad (5)$$

The population hazard at t is the individual hazard times the mean frailty of survivors to t. Therefore, if we ignore the presence of unobserved heterogeneity, estimates of the default hazard will show duration dependence that is more negative than the true duration dependence of the population.

To solve the problem, we will assume a parametric form for the base line hazard. We assume in our baseline specification that the baseline hazard has a Weibull distribution so that

$$\lambda_0(t) = pt^{p-1}. \quad (6)$$

We are interested in the estimates of the parameter p , which measures duration dependence.[3] This parameter indicates how the hazard function varies along the loan maturity. Thus, interpreting the estimation of this parameter works as a bridge between the above model and the credit market.

If $p < 1$, it means that the probability that default at time t occurs, conditional on default not occurring up to t, decreases along the loan maturity. We interpret this as a sign of asymmetric information. We assume that any given negative shock that might hit the borrower leading him not to pay the loan, is uniformly distributed along the loan. A parameter p lower than one means that the probability of a default is higher at the beginning than at the end. Thus, some borrowers are not paying their loans for other reasons rather than being hit by a negative shock. We conjecture that either he would never be able to pay all monthly installments (adverse selection) or he is not willing to pay all his debt (moral hazard).

If $p = 1$, it means that the conditional probability that default is constant along the loan maturity. We interpret this as a sign that asymmetric information is not a considerable issue. It is evidence that the borrowers default in face of a negative shock, which has a uniform distribution by assumption. We also interpret $p > 1$ in the same way. Even though it might not be a perfect match between negative shocks and borrower's performance, the negative shock can affect borrower's repayment ability with a delay. For instance, a borrower loses his job, but still has income for a few months, due to unemployment insurance. If he expects to find a job in the short term, he would continue to pay his installments.

4.4.Data

We access data for the vehicle credit market from one of the three largest Brazilian private banks. According to the Central Bank, the combined assets value of Bradesco, Itaú Unibanco, and Santander was R$1.7 trillion (US$1 trillion), and the credit portfolio was R$573 billion (US$345 billion), representing more than 40% of the market share. The Bank also plays a significant role in vehicle credit market, having more than 15% of market share.

Our sample covers the years between 2004 and 2010.[4] We have about 86,000 observations with micro-level detailed information about contract terms as well as personal characteristics. The information about terms include interest rate (per months), maturity (in months), total due (in Reais), and installment value (in Reais). Furthermore, we have information about borrower's performance. For example, if the borrower delayed for more than 30 days any installment, and proportion of paid installments by him by the end of the sample. Finally, the data contains a rich set of borrower's characteristics, including type of risk, income, gender, presence of a third-part guarantor, type of job, type of residence, marital status, if the borrower is a client of The Bank and if the car is new.

In Table 1, we display the summary statistics. The average interest rate is 1.90% per month with median of 1.77%. Loan maturity is around four years (47.7 months), with 5th and 95th being two and five years, respectively. The total due and the installment value have means of R$22,472 (US$10,403) and R$482 (US$233) with standard deviation of 11,940 and 257, respectively. The high (proportion) standard deviation of both variables is partially explained by the fact that several variables directly affect them, such as interest rate, maturity, car value, down payment. Default - defined as delay in at least one installment for at least 30 days - has an average of 0.12. Finally, the proportion of paid installments - defined as the ratio of paid installments to total installments by December 2010 - is 0.72. As part of the loans made after 2006 is still open, even a not-defaulter might not have paid all installments. This explains why the proportion of paid installments is relatively low. If we restrict our attention to loans originated prior to 2006, the average of this variable is higher than 0.95.

Panel B provides information about borrower's characteristics. The average income is R$2,445 (US$1,199) associated with a high variance of 11,289. High-risk borrowers represent 3% of the sample, and contracts with a third part guarantor 10%. Two-third are male, and one-fourth is client of The Bank. About marital status, 37% are married, and 47%, single. Employees represent 60% and the self-employed or entrepreneurs, 16%. Most of the borrowers (81%) are home owners, indicating that they have another valuable equity, and only 16% live their with parents. Panel C shows that one-third of the consumers bought a new car and most of borrowers buy the car throught a dealer priority.

4.5.Empirical Analysis

4.5.1.What to expect from the model predictions

In Section 3, we develop a model about lifetime data. Then, we derive some prediction relating the model and the vehicle credit market. In this section, we estimate (3) under the assumption (6). We interpret the parameter p of the model as following: if $p < 1$, it is an evidence of asymmetric information in the credit market. On the other hand, if $p \geq 1$, it is evidence that asymmetric information is not so relevant.

We start by estimating equation (3) without controlling for observable variables, it is our benchmark. Then, we add all observable characteristics – contract terms as well as personal and car characteristics. We compare both results analyzing problems regarding asymmetric information. Finally, we do a horse race between contract terms, and personal and car characteristics to assess which set of variables is more helpfuk for more The Bank in tackling the asymmetric information.

4.5.2.Results

Let's turn to the estimations. Initially, we do it including no controls, which means that x is a null vector in terms of equation (3). Even though others variables affects the hazard rate, this estimation gives us a useful benchmark. Table 2 reports the result and, to simplify notation, we call the estimative of the parameter $\hat{p}^{nc}$.[5] For the first year of our sample, $\hat{p}^{nc}$ is 1.34. $\hat{p}^{nc}$ is 1.34. In the following years, $\hat{p}^{nc}$ presents a decline trend. The estimation point is $\hat{p}^{nc}$ presents a decline trend. The estimation point is 1.17 and 0.98 in 2007 and 2006, respectively. Nevertheless, from 2007 on, $\hat{p}^{nc}$ stabilizes around 0.90 and is always above one statistically significant at 5 percent. In terms of the model, asymmetric information worsened along the years. In particular, after 2007, $\hat{p}^{nc}$ is consistently above one, indicating that asymmetric information turned into a relevant issue.

This initial downward trend followed by a flat pattern is clearly shown by Figure 1. We plotted the estimation of parameter p as well as its interval of confidence at 5 percent. From 2004 to 2007, not only does the point estimation show a downward trend, but also its confidence interval almost always has no intersection. On the other hand, the period

2007 to 2010 is characterized by estimations around 0.90, and their confidence interval coincides in the range 0.87 - 0.91. Indeed, we cannot rule out the possibility that all these parameters (2007, 2008, 2009, and 2010) are jointly indistinguishable.

All these results have a close relationship with the evolution of the Brazilian credit market. According to Section 2, at the beginning of the sample, the vehicle credit was at a modest level. Credit was restricted. For instance, down payment can be viewed as a proxy for a barrier for a borrower to enter the market. Higher (proportional) down payment means a higher barrier to entry. In that year, the average of this variable was R$7,615 (US$ 2,599) – or 37% of the car value – and only 3 contracts had no down payment in 2004.

Then, the total credit supplied by banks for car loans started to increase consistently. The yearly average increase in total supplied was 25% during the sample period. Assunção, Benmelech, and Silva (2012) claimed to a "democratization of credit" that had begun in mid- August of 2004. In particular, the proportion of high-risk borrowers in the sample increased, and the average of borrower income decreased. Moreover, these "new riskier borrowers" have a higher delinquency rate.

We hypothesize that, as The Bank expanded credit toward riskier borrowers, the asymmetric information issue worsened. In terms of model, the parameter p decreased in the first half of the sample period reflecting this worsening. Moreover, after 2007, it is always the case that $\hat{p}^{nc} < 1$, confirming our hypothesis. The flat pattern in the following half might indicate that The Bank learned how to deal with this increasingly risky borrower. Thus a side effect of these movements seem to be an increase in the asymmetry of information between borrowers and lenders.

Back to the estimations, we include all controls (vector x contains both personal characteristics as well as contract terms). Table 3 presents results and Figure 2 shows the point of estimation as well as its confidence interval. We call the estimation $\hat{p}^{ac}$.[6] After an initially downward trend, the coefficient reverted and stabilized very close to one. In the beginning, the estimation of parameter p is very similar to the previous estimation. However, after 2007, $\hat{p}^{ac}$ and $\hat{p}^{nc}$ became be different. The parameter $\hat{p}^{ac}$ is almost always equal to or higher than one. Thus The Bank partially succeeded in mitigating the asymmetric information problem using observable characteristics.

Results from table 3 also have a close relation with economic background. As The Bank expanded credit, $\hat{p}^{ac}$ decreased, which is in line with our previous conjecture that asymmetric information issue worsened along the years. Nevertheless, in 2008 the downward trend was reverted, the parameter p became to close to one and remained stable until the end of the sample period. It might be the case that The Bank started to use observable characteristics to screen borrowers of different types of risk. Consequently, the asymmetric information issue diminished and estimated parameter p increased.

The results are in line with a growing body of empirical evidence that suggests credit expansion leads to subsequent waves of default The results are in line with a growing body of empirical evidence that suggests credit expansion leads to subsequent waves of default (Keys et at. (2010), Mian and Sufi (2009, 2010)), and with Assunção, Benmelech and Silva (2012) who point out that, after the implementation of the "Fiduciary Law ", borrowers increased the likelihood of a late payment. We go a step further to describe the mechanisms beyond this movement. A credit expansion is usually associated with a riskier borrower. This leads to a worsening in the asymmetric information issue.

Next, we do a horse race between contract terms, and personal and car characteristics. In Tables 2 and 3, we documented that there is asymmetric information in the credit market and The Bank deals with this using all observable variables. Now, we test which set of variables is more helpful for The Bank in mitigating the asymmetric information problem. Tables 4 and 5 show the results.

Comparing both tables shows an interesting fact. The $\hat{p}^{pc}$ has a very similar pattern of $\hat{p}^{nc}$, which is a naive parameter in the sense that The Bank does little to tackle the information problem.[7] This is evidence that The bank does not use personal characteristics to screen borrowers. Therefore, at least for the technology The Bank used at that time, personal characteristics were not helping them to mitigate problems arising from adverse selection. There are two possible complementary explanations. First, the quality of information is not good. People simply lie or give imprecise information. For instance, they overstate their real income. Alternatively, observable personal characteristics do not carry useful information about the borrower's type of risk.

On the other hand, the $\hat{p}^{ct}$ has a very similar pattern of $\hat{p}^{ac}$, which includes all efforts coming from The Bank to deal with the information problem. Thus, much of the effort that The Bank does to overcome asymmetric information comes through contract terms. The flexibility that borrowers have by the time they are signing the contract helps The Bank to screen the borrower's type of risk. High-risk borrowers probably prefer a different type of contract than low-risk borrowers. For example, according to an employee at The Bank, high-risk prefer longer maturity and lower down payment. Mapping their preference helps The Bank to screen borrowers.

Therefore, The Bank is able to access two sets of information about borrowers: personal characteristics and contract terms. Both may help it to mitigate the asymmetry in information. Results shown earlier points out that $\hat{p}^{pc}$ is closer to $\hat{p}^{nc}$ and $\hat{p}^{ct}$ is closer to $\hat{p}^{al}$. In a horse race between personal characteristics and contract terms, the latter is more helpful for

The Bank in tackling the asymmetric information problem.

Another additional result is the following. The Bank does not "map one for one" contract terms and personal characteristics. If this were the case, we would see $\hat{p}^{pc} = \hat{p}^{ct}$ for all years. That is not the case. This may indicate that either there is space for a negotiation about terms between borrowers and lenders or The Bank offers a menu of contracts.

We need to mention one caveat. In this section, we are not considering unobservable variables that might affect borrower's performance. As we see in Section 3, this might bias our estimative of parameter p. Nevertheless, this estimation is very useful for at least two reasons. First, we are observing the same set of variables that The Bank observes. Second, the results are interesting in their own. They are very useful even in a comparative way.

4.6.Conclusion

The aim of this paper is to empirically analyze problems arising from asymmetric infor- mation in the credit market. Using micro-data coming from one of the largest Brazilian banks, we investigate the car loan market. This particular market witnessed a sudden and sizable growth during the sample period. We conjecture that, as a side effect of this growth, asymmetric information issue worsened along the years.

To investigate our hypothesis, we estimate a model of lifetime data. Results confirm our conjecture. During the sample period, the asymmetric information problem became worsen. The Bank partially succeeded in tackling this problem using contract terms as well as personal and car characteristics. Finally, we run a horse race between contract terms, and personal and car characteristics in order to assess which set of variables is more helpful for The Bank dealing with the problem. We find that contract terms work better to tackle the problem.

4.7.Mathematical Appendix

In this section, we derive some of the relations that we used throughout the exposition

4.7.1.Relation between hazard and Distribution Function

$$\lambda(t) = \lim_{\Delta \downarrow 0} \frac{\Pr(t \leq T < t + \Delta | T \geq t)}{\Delta}$$

$$= \lim_{\Delta \downarrow 0} \frac{\Pr(t \leq T < t + \Delta)}{\Delta} \quad \frac{1}{1 - F(t)}$$

$$= \frac{f(t)}{1 - F(t)}$$

Now let $S(t) = 1 - F(t)$ be the survivor function. Then $\lambda(t) = -d \ln S(t)/dt$. The integrated hazard can be defined as integrated hazard can be defined as

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

So that so that $S(t) = \exp[\Lambda(t)]$

4.7.2.Evolution of Distribution of Unobserved Heterogeneity through time

Bayes rule implies that:

$$dG(v|T > t, x) = \frac{\Pr[T > t|x, v]dG(v)}{\int \Pr[T > t|x, v]dG(v)}$$

$$= \frac{(1 - F(t|x, v))dG(v)}{1 - F(t|x)}$$

4.7.3.Population Hazard

$\lambda(t|x) =$

$$\int \nu \lambda_0(t) e^{x\beta} dG(\nu \,|\, T > t, x)$$

$$= \lambda_0(t) e^{x\beta} E[\nu | T > t, x]$$

Appendix: Variable description and construction

For reference, the following is a list of the variables used, their sources, and a brief description of how each is constructed.

1. Spread : The difference between the monthly interest rate paid by the borrower and the federal fund rate (in percentage points).

2. Maturity : Loan maturity (in months).

3. Down payment : The amount paid by the buyer that was not financed (in R$).

4. Loan size: The total amount financed by The Bank (in R$).

5. Income : The borrower's (estimated) monthly income calculated by The Bank (in R$).

6. Client dummy: A dummy variable that takes the value of one if the borrower is a client of The Bank, and zero otherwise.

7. High risk dummy: A dummy variable that takes the value of one if the borrower is classified as a high risk, and zero otherwise.

8. Guarantor dummy: A dummy variable that takes the value of one if the loan has a guarantor, and zero otherwise.

9. Gender dummy: A dummy variable that takes the value of one if the borrower is a male, and zero otherwise.

10. Type of job: A five-category variable: employee, retired/pensioner, self-employed, entrepreneur, and other.

11. Type of residence : A four-category variable: homeowner, lives with parents, renter, and other.

12. Marital status : A five-category variable: single, married, divorced, widower, and other.

13. New car : A dummy variable that takes the value of one if the car is

new, and zero otherwise.

14. Car value: Car value (in R$).

15. Model: Car model.

16. Car age: The difference (in years) between the date that the loan was signed and the date that the car was manufactured.

17. Dealer priority dummy: A dummy variable that takes the value of one if the consumer bought the car from a priority dealer, and zero otherwise.

18. Federal fund rate: The federal fund interest rate.

19. Inflation: The inflation rate over the last 12 months.

20. GDP growth: Quarterly GDP growth.

21. Default: A dummy variable that takes the value of one if the borrower was at least 90 days late, and zero otherwise. (This the criteria used by the Central bank).

22. Proportion of installments paid: The ratio of paid installments to total installments by December 2010.