BANK LOAN LOSSES-GIVEN-DEFAULT,

A Case Study

J. Dermine and C. Neto de Carvalho*

First draft: October 20, 2003
Current draft: March 10, 2005

* INSEAD (Fontainebleau) and Universidade Catolica Portuguesa (Lisbon), respectively.
The authors are grateful to the comments of two referees, to Professor J. Santos Silva for his insights into the econometrics of fractional responses; to P. Jackson, M. Massa, A. Sironi and M. Suominen for comments on a first draft; to J. Cropper for editorial assistance; and to Banco Comercial PortuguLs for access to internal credit data.

BANK LOAN LOSSES-GIVEN-DEFAULT,

A Case Study

Abstract

The empirical literature on credit risk has relied mostly on the corporate bond market to estimate losses in the event of default. The reason for this is that, as bank loans are private instruments, few data on loan losses are publicly available. The contribution of this paper is to apply mortality analysis to a unique set of micro-data on defaulted bank loans of a European bank. The empirical results relate to the timing of recoveries on bad and doubtful bank loans, the distribution of cumulative recovery rates, their economic determinants and the direct costs incurred by that bank on recoveries on bad and doubtful loans.

JEL Classification: G21, G28
Keywords : credit risk, bank loan, losses-given-default (LGD)

Corresponding author: Jean Dermine, INSEAD, Boulevard de Constance, F-77300 Fontainebleau, France. E-mail : jean.dermine@insead.edu

Introduction

During the past 10 years, marked progress has been made to measure credit risk. Most approaches involve the estimation of three parameters: the probability of default on individual loans or pools of transactions, the estimate of the losses-given-default (LGD) and the correlation across defaults (Crouhy *et al*., 2000 ; Duffie and Singleton, 2003).
Although several empirical academic studies have analyzed credit risk on corporate bonds, very few studies have been applied to bank loans. The reason for this is that, as bank loans are private instruments, few data are publicly available. This paper contributes to our understanding of bank loan credit risk by providing a methodology to analyze the loss severity rate after a credit event. This is then applied to a unique set of data on losses on loans to small and medium-size firms over the period 1995-2000. The data were provided by Banco Comercial PortuguLs (BCP), the largest private bank in Portugal. The use of data from a specific bank is a limitation. But, given the absence of publicly available data on cash flow recovery on distressed bank loans, it is a step in our understanding of the determinants of bank loan losses-given-default.
In terms of contribution to the literature, and to the best of the authors' knowledge, this is the first paper to apply mortality analysis to defaulted bank loan recovery rates, to carefully document the timing of recoveries (a useful piece of information to calculate interest rate risk and credit risk provisions over time), and to test empirically the determinants of recovery rates, such as the impact of guarantees and collateral, loan size, the industry factor, and the age of the borrower. Moreover, it provides information on the direct costs incurred by a bank in recoveries on bad and doubtful loans. Finally, it is the first empirical paper on bank loan losses-given-default in Europe.
Empirical evidence on loan losses-given-default needs to be collected by banks and their supervisors. In the new capital accord agreed upon in June 2004 (Basel Committee, 2004), financial institutions are invited, in the internal rating-based (IRB) approach, to estimate the one-year probability of default and the expected LGD.

The paper is structured as follows. In Section 1, it is argued that bank loans are likely to have some characteristics significantly different from those of corporate bonds. This justifies the need for specific studies on bank loans, to complement existing studies on corporate bonds. A summary of the literature follows. In Section 2, the database on individual loans losses is presented. The mortality-based approach to analyze recovery rates on bad and doubtful loans is discussed in Section 3. Empirical evidence on cumulative recovery rates is presented in Section 4, and, in Section 5, a statistical analysis of the determinants of loan losses-given-default is developed. In Section 6, estimates of direct costs incurred in recoveries are presented. Section 7 concludes the paper.

Section 1. Literature Review

Altman (1989) applied actuarial analysis to study mortality rates of US corporate bonds. This was followed by extensive empirical literature on credit risk in the bond market (see, for instance, Nickell *et al.*, 2000, or Acharya *et al.,* 2003_a). In view of this vast literature, one needs to justify a study on recovery on bank loans. Several arguments are proposed. The first two are that small firms are informationally more opaque, and that the relationship between the owner/manager of the firm and the bank is often very close (Allen *et al.*, 2004). This has two implications. In the case of distress, the owner/manager has relatively more to lose because his or her skills will be firm-specific. Efforts to repay a loan could be greater than in a large public firm. Second, the close relationship with the bank might imply that a bank will hesitate to foreclose a loan, hoping to capture the option value of the future relationship (Dewenter and Hess, 2004). A third argument is that a bank might hesitate to foreclose large loans if it believes that this can have local spillover effects on other firms (i.e., clients of the bank). In the arm's-length transactional corporate bonds markets, these 'option' or 'macro' considerations will be ignored. A fourth reason is that a large number of distressed loans may create bottlenecks in the workout unit of the bank, with effect on recovery rates. Finally, in the case of distressed bank loans, knowledge of the timing of cash flow recoveries will be useful to measure interest rate risk, with computation of repricing gaps or duration measure, as well as to provide information to calculate dynamic loan loss provisions over time.

Studies on corporate bonds reported information on the probability of default over time for different bond ratings, on recovery rates based on market prices at the time of default, on estimates of rating transition matrices, and on the degree of correlation between default frequencies and recovery rates (see, for instance, Frye $2000_{a,b}$ and 2003; Allen and Saunders, 2003; Altman *et al.*, 2003; or Acharya *et al.*, 2003_a). Altman *et al.* (2003) report an average recovery rate (price after default) of 37 % for the US corporate bond market over the period

1982-2001. They report that the aggregate recovery rate on defaulted bonds is affected negatively by the supply of defaulted bonds. Acharya *et al.* (2003_a) report an average recovery rate of 48% for senior secured bonds, and 51% for senior unsecured bonds for the period 1982-1999. They report that recovery on individual bonds is affected not only by seniority and security, but also by the industry conditions at the time of default. These last two empirical studies validate the theoretical study by Shleifer and Vishny (1992), who examine the impact of industry conditions on liquidation values. The above studies relied on publicly traded bond data. Carey (1998) analyzed credit risk in privately placed bonds over the period 1986-1992, an asset category that resembles loans in that they are monitored private debt. As mentioned above, many fewer studies have focused on the bank loan markets because of the private nature of these transactions.

Asarnow and Edwards (1995) examined 831 defaulted loans at Citibank over the period 1970-1993. They reported an average cumulative recovery rate of 65%, based on the present value of future cash flows received after the default date. A significant result of that study was that the distribution of recovery rates was bi-modal, with a concentration of recovery rates on either the low or the high end of the distribution.[1] Carty and Lieberman (1996) measured the recovery rate on a sample of 58 bank loans. Based on secondary market prices for defaulted bank loans for the period 1989-1996, they reported an average defaulted bank loan price of 71%. They did not observe a bi-modal distribution, but reported skewness toward the high end of the price scale. In the same study, the authors measured the recovery rate on a sample of 229 small and medium-size loans in the US. They reported an average recovery rate of 79%, based on the present value of cash flows. Again, the distribution was highly skewed toward the high end of the scale. Grossman *et al.* (1998) analyzed recovery rate on 60 syndicated bank loans over the period 1991-1997. Based on secondary market prices after the credit event, they reported an average recovery figure of 82% with a standard deviation of 24%. No information was provided on the shape of the distribution of recovery rates. Altman and Suggitt (2000) analyzed the probability of default for US publicly rated

[1] The issue of the relevant distribution for recovery rates is important in portfolio credit risk modeling. For instance, Altman et al. (2003) assume that it follows a beta distribution.

bank loans with a size of at least US$ 100 million over the years 1991-1996, but did not report data for recovery rates. All the above studies focused on the US market. Two papers concern Latin America and Mexico. Hurt and Felsovalyi (1998) analyzed 1,149 bank loan losses in Latin America over the period 1970-1996. They reported an average recovery rate of 68.2%, calculating the present value of recovered cash flows. They showed that loan size as a contributing factor to loss rates, with large loan default exhibiting lower recovery rates. They attributed this to the fact that large loans, often not secured, were made to economic groups that were family owned. As in Asarnow and Edwards' study, they reported a bi-modal distribution. La Porta, Lopez-de-Silanes and Zamarripa (2003) analyzed loan default and losses-given-default in Mexico in the context of 'related lending', that is, lending to shareholders or directors of the bank. They reported an average recovery rate of 46% for 'unrelated' loans, and 27% for 'related' loans over the period 1995-1999. Evidence of skewness toward the high end of the distribution was also reported. None of the above studies provided information on the timing of recoveries.

In Europe, although several authors have analyzed the determinants of the aggregate level of banks' loan losses (for instance, Acharya *et al.* ,2003_b; Dahl and Logan, 2002; and Salas and Saurina, 2002), none, to the best of the authors' knowledge, have analyzed losses-given-default on individual loan transactions.

Thanks to access to a unique data-base on loans to SMEs, this paper provides some empirical evidence on the timing of recovery on individual bad and doubtful loans, on cumulative recovery rates, on their determinants and on the direct costs incurred in recovery.

Section 2. Bank Loan Losses, Database and Measurement Issue

The database used in this study was provided by the largest private bank in Portugal, Banco Comercial Português (BCP). It consists of 10,000 short-term loans granted to small and medium-size companies[2] over the period June 1995 to December 2000. All these companies, based in the south of Portugal, including Lisbon, have a turnover of more than €2.5 million. As discussed above, the use of data from a specific bank raises the issue of the generality of the empirical results. As shown in Table 1, the loan portfolio distribution used in this study

[2] The data gathered for this study do not include any reference to the identity of the clients or any other information that, according to Portuguese banking law, cannot be disclosed.

is fairly similar to the loan portfolio distribution of the entire Portuguese banking system. Differences are minor for most business sectors, textile and services excluded. One observes that the textile sector is under-represented in the sample (1.3% portfolio share vs. 7.6% for the country), and that the service sector is over-represented (33.3% in the sample vs. 19.5% for Portugal). The distribution of the sample reflects the importance of the capital Lisbon, and that the textile sector is strongly represented in the north of Portugal. The loan data used in this study appear to be quite representative of the entire country. However, it should be recognized that the data on recovery on distressed bank loans could capture some of the BCP workout unit's idiosyncrasies.

Panels A and B of Table 2 provide information on the number of defaults per year and the amount of debt outstanding at the time of default. Panels C and D of Table 2 show the number of loans with personal guarantee or collateral, the age of the firm and the number of years of relationship with the bank. One observes that the series of 374 default cases is distributed evenly over the six years, and that the distribution of debt outstanding at the time of default is highly skewed towards the low end. Half of the debt is less than €50,000.[3] The available information in the database includes the industry classification, the interest rate charged on the loan, the history of the loan after a default has been identified, the type of collateral or guarantees, the internal rating attributed by the bank, the age of the firm and the length of relationship with the bank.[4]

In Panel C of Table 2, the various forms of guarantees or collateral are reported. These include:

- Personal guarantees
- Real estate collateral
- Physical collateral (inventories)
- Financial collateral (bank deposits, bonds or shares).

In 35.6**%** of the cases, there is no guarantee or collateral. Personal guarantees, which are

[3] This sample includes loans of a much smaller size than those used in the reported US bank studies. The face value of loans was higher than US$ 100 million in Altman and Suggitt (2000), and the average commercial and industrial loan was US$6.3 million in Asarnow and Edwards (1995). Basel II (Basel Committee, 2004) gives banks the option to include loans of less than € 1 million in their retail portfolio.

[4] The database used in this study did not exist in the appropriate format. The bank provided help to identify all these variables, and to recover the data located in various databases. The history of each loan, after a default had occurred, was carefully analyzed.

used in 58.3% of the cases, refer to written promises made by the guarantor (often the owner or the firm's director) that allows the bank to collect the debt against the personal assets pledged by the guarantor. Collateral is used in 15% of the cases.[5]

Panel D of Table 2 shows the age of the firm and the number of years of its relationship with the bank. Companies have, on average, a life of 17 years, with extremes going from 6 months to 121 years. The average relationship with the bank is six years.[6] Table 3 reports the concentration of default cases in different business sectors and the use of guarantee/collateral across these sectors. Fifteen business sectors have been created, with reference to the European Union's NACE economic activity codes. Further aggregation, used in the econometric tests, leads to four activity sectors: real sector (activities with well-identified real assets, such as land, mines or real estate property, which could be used for security), manufacturing, trade and services. Default cases are observed in all business sectors, with a concentration in construction (13% of default cases), wholesale and retail trade (44%) and services (10%). The relative use of personal guarantee or collateral seems to be uniformly spread across the four aggregated activity sectors.

Any empirical study of credit risk raises two measurement issues. Which criterion should be used to define the time of a default event? Which method should be used to measure the recovery rate on a defaulted transaction?

The criterion used for the classification of a loan in the 'default' category is critical for a study on recovery rates, as a different classification would lead to different results. Three 'default' definitions are used in the literature:

i) A loan is classified as 'doubtful' as soon as "full payment appears to be questionable on the basis of the available information".[7]

ii) A loan is classified as 'in distress' as soon as a payment (interest and/or principal) has been missed.

iii) A loan is classified as 'in default' when a formal restructuring process or bankruptcy procedure is started.

---

[5] Jimenez and Saurina (2002) also observe, in the case of Spain, that a very large proportion of bank loans are not collaterized.

[6] The relatively short average relationship is due to the fact that the bank was created in 1985, after the deregulation of the Portuguese banking system.

In this study, because of data availability, we adopt the second definition; that is, a loan is classified as 'in default' as soon as a payment is missed.[8] For information, the reporting to the Central Bank of Portugal takes place after 30 days, if the loan remains unpaid or unrestructured.

The second methodological issue relates to the measurement of recovery on defaulted loans. There are two methodologies:

i) The price of the loan at the default date, defined most frequently as the trading price one month after the default. This approach has been used in studies on recoveries on corporate bond defaults.

ii) The discounted value of future cash flows recovered after the default date.

As no market price data are readily available for defaulted bank loans in Portugal, the second methodology -the present value of actual recovered cash flows- is the only feasible alternative. This approach was adopted by Asarnow and Edwards (1995), Carty and Lieberman (1996) and Hurt and Felsovalyi (1998). While these authors did not have access to the interest rate charged on individual loans and had to rely on an approximation of credit-risk adjusted yield curve, data on interest rates charged on the loans are available in this study. The present value of cash flows recovered on impaired loans allows the measurement of the proportion of principal and interest that is recovered after the default date. This approach has the advantage that, if a loan is fully repaid, the present value of actual cash flows recovered will be equal to the outstanding balance at the default date. It should be noted that this amount could differ from the price of the loan at a time of default, which would incorporate the expected cash flows and adequate risk or liquidity premia.

---

[7] Hurt and Felsovalyi (1998).

[8] For the sake of comparison, the definition of default adopted by the Basel Committee is as follows (Basel Committee, 2004, p.92): "A default is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place:
a)The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held).
b) The obligor is past due more than 90 days on any material credit obligation to the banking group. Overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstanding.

In order to measure the cash flows recovered after a default event, we tracked, each month, the post-default credit balances. Capital recovery is a reduction in the total balance. The total cash flow recovered is this capital recovery plus the interest on the outstanding balance. If, at first glance, the tracking of cash flows after a default event would appear a relatively simple (but time-consuming) exercise, special cases did require some adjustment. Two such cases are discussed hereafter.

First, there are 52 cases of multiple defaults. This refers to situations in which a company enters into a default category, returns to 'performing' status at a later date after paying the loan fully, and then falls back into the default category. A first option could have been to consider each default as a separate event, but this might have biased the econometric analysis with several defaults linked to the same borrower. In the econometric results that will be reported, we have kept the 'first default' cases. To check the robustness of the results based on this choice, we conducted similar tests, while keeping all of the 'multiple default' cases.

Second, in the case of formal loan restructuring, with the loan returning to a 'performing' status, we did not consider these cases as a 100% recovery. Indeed, in many cases, these loans subsequently fell back into the default category. The cash flows received on restructured loans were identified carefully.[9]

In Table 4, we report the 12-, 24-, 36- and 48-month cumulative recovery rates for the total sample, as well as the 48-month cumulative recovery rates for three categories: loans with no guarantee/collateral, loans with personal guarantee only, and loans with collateral. The mean cumulative recovery rate of 71% is of the same order of magnitude as those reported by Asarnow and Edwards (1995) and Hurt and Felsovalyi (1998) for Latin America. The 48-month recovery on loans with collateral is 92%, but the average recovery on loans with personal guarantee, 64%, appears lower than on loans with no guarantee/collateral, 76%.

In Table 5, we report the 48-month cumulative recovery rates for the 15 activity sectors. One observes higher recovery rates in the first five sectors, which belong to the aggregated 'real' sector. For instance, a 98% recovery rate is observed for the hotel/restaurant sector. Lower

[9] Note that this approach is more conservative than the one adopted by Hurt and Felsovalyi (1998).

recovery rates are observed in the wholesale and retail trade sectors (65% and 55% respectively). Data on the one-year frequency of default are also given for the different sectors. One observes that the value-weighted frequency of default is substantially lower than for the unweighted frequency, indicating a size effect, with large loans much less likely to default. A similar result was observed in Spain (Jimenez and Saurina, 2003).

Finally, Table 6 reports, for each year of the time period 1995-2000, the real rate of growth of GDP, the frequency of default for the total sample and for the four aggregated sectors, and the 12-month recovery rate. The first three years (1995-1997) exhibit a real rate of GDP growth lower than the average, while the last three years (1998-2000) show higher economic growth. Data on the frequency of default show that these are negatively correlated with the real rate of growth of GDP. Frequencies of default are higher in the first three years, 1995-1997, a period of lower economic growth. This is valid for the total sample and for each of the four aggregated sectors. However, one fails to observe a correlation between GDP growth and loan recoveries. Recoveries were high in 1997 (66%), while they were much lower in 1999 (32%).

Two approaches will be used to further analyze the recovery rates, a univariate mortality-based approach and a multivariate statistical analysis of the determinants of recovery.

Section 3. A Mortality-based Approach to Analyze Recovery Rates

Having access to the history of cash flows on these loans after default, we can study the time distribution of recovery. With reference to the studies by Altman (1989) and Altman and Suggitt (2000), we apply the mortality approach. It must be noted that the mortality approach was applied to measure the percentage of bonds or loans that defaulted *n* years after origination. The application of mortality to loan recovery rates is, to the best of our knowledge, novel. It examines the percentage of a bad and doubtful loan that is recovered *n* months after the default date. This methodology is appropriate because the population sample is changing over time. For some default loans, those of June 1995, we have a long recovery history (66 months), while for the '2000' loans in default, we have an incomplete history of recovery. The actuarial-based mortality approach, based on the Kaplan-Meier estimator (Greene, 1993) adjusts for changes over time in the size of the original sample.[10]

For an individual loan *i* in default, we define three concepts,[11] t denoting the number of periods after the initial default date:

$MRR_{i,t}$ = Marginal Recovery Rate in period t
= Cash $flow_i$ paid at the end of period t / $Loan_i$ outstanding at time t

$PULB_{i,t}$ = Percentage Unpaid Loan Balance at the end of period t = 1 – $MRR_{i,t}$

$CRR_{i,T}$ = Cumulative Recovery Rate T periods after default = $1 - \prod_{t=1}^{T} PULB_{i,t}$

Similar to Carey (1998), Asarnow and Edwards (1995), Carty and Lieberman (1996), Hurt and Felsovalyi (1998) and La Porta *et al.* (2003) the *Cumulative Recovery Rate* at time T, CRR_T, represents the proportion of the initial default loan that has been repaid (in present value terms) T periods after default. Note that these authors report only the total cumulative recovery rate over a long (unidentified) period, whereas we, adopting the Altman mortality-based approach, report the extent of cumulative recovery over time.[12]

Having computed the cumulative recovery rate on individual loans, one can compute an unweighted average cumulative recovery rate for the sample of loans. This is the approach adopted in the losses-given-default literature. Alternatively, one can compute a sample weighted average recovery rate that will take into account the size of each loan. This is defined as follows:

$SMRR_t$ = Sample (weighted) Marginal Recovery Rate at time t ,

---

[10] The different approaches (cohort analysis used by Moody's, static pool employed by Standard & Poor's, and mortality rate) to measure credit risk, are discussed and contrasted in Caouette, Altman, and Narayan (1998).
[11] A numerical example is provided in Appendix 0ne.
[12] Information on the timing of recovery can be used to calculate how the level of provisions should be adjusted over time. Indeed, information that little recovery takes place after four years is indicative of a need to sharply increase loan provisions in that period.

$$= \frac{\sum_{i=1}^{m} Cashflow\ received_{i,t}}{\sum_{i=1}^{m} Loan\ outs\tan ding_{i,t}}$$, where i stands for each of the m loan balances outstanding in the sample, t periods after default.

$SPULB_t$ = Sample (weighted) Percentage Unpaid Loan Balance at period t

$= 1 - SMRR_t$

$SCRR_T$ = Sample (weighted) Cumulative Recovery Rate T periods after the default

$$= 1 - \prod_{t=1}^{T} SPULB_t$$

A comparison of the Sample (weighted) Cumulative Recovery Rate with the average of recovery rates on individual loans will be indicative of a size effect.

Section 4. Cumulative Recovery Rates, Empirical Results

The sample marginal and cumulative recovery rates for the sample T periods after a default date, $SMRR_T$ and $SCRR_T$, respectively, are reproduced in Figures 1 and 2. One observes in Figure 1 that most of the marginal recovery rates in excess of 5% occur in the first five months after the default, and that the cumulative average recovery is almost completed after 48 months. The timing of cash flow recovery is useful information to compute interest rate risk on bad and doubtful debt with repricing gaps or duration.

As indicated in Figure 2, the unweighted cumulative recovery rates[13] after 36 and 48 months are, respectively, 67.3% and 70%. One observes that the sample weighted average cumulative recovery rates ($SCRR_T$) are 53.5% and 56.3%, 36 months and 48 months, respectively, after the default event. The differences between the unweighted and weighted

[13] Note that, as is the case with other studies, these are gross recovery rates. Estimates of the bank's internal recovery costs, which include the cost of the workout units and legal and accounting costs, are discussed in section 6.

average cumulative recovery rates are indicative that recovery on large loans is significantly lower. The average cumulative recovery rates reported in Figure 2 were calculated with the mortality-based approach. For the sake of comparison, we reported in Table 4 the average cumulative recovery rates using a different methodology. For each horizon (12, 24, 36 and 48 months), we include in the pool of loans those data with recovery available for that horizon. For instance, in the case of the 48-month cumulative recovery, we consider only the loans from 1995 and 1996. With this pool of loans, one calculates an unweighted average cumulative recovery rate over 48 months of 71%, to be compared to the 70% obtained with the mortality-based approach.

It is also of interest to analyze the distribution of cumulative recovery rates across the sample of loans. The distribution of cumulative recovery rates after 48 months is reproduced in Figure 3. This figure shows a bi-modal distribution with many observations with low recovery and many with complete recovery. These results are quite similar to those reported by Asarnow and Edwards (1995) and Schuermann (2004) for the US, and Hurt and Felsovalyi (1998) for Latin America. Loan portfolio models that incorporate a probability distribution for recovery rates should take into account this bi-modal distribution.[14]

We next attempt to analyze the determinants of the recovery rate.

Section 5. The Determinants of Recovery Rates, a Statistical Analysis

In this section, we attempt to estimate empirically the determinants of recovery rates. A discussion of the choice of explanatory variables and the econometric specification is followed by the empirical results.

Explanatory variables and econometric specification

Explanatory variables include the size of the loan, the type of guarantee/collateral support,

[14] Current commercial credit portfolio models do not incorporate the bi-modal distribution. For instance, CreditRisk+, developed by Credit Suisse Financial Products, assumes a fixed expected recovery rate within each band, while CreditMetrics uses a beta distribution (Crouhy *et al*., 2000).

the industrial sectors, the default year and the age of the firm.[15] The size of the loan is included because some empirical studies and the sample univariate weighted- and unweighted average cumulative recovery data have pointed out the effect of the loan size. The year dummy is included to provide a better understanding of the volatility of the recovery rate over time. In the case of Banco Comercial PortuguLs, it was indicated by the bank that a reorganization of the workout unit in 1999 could have had an impact on recoveries. Finally, it is of interest to know the magnitude of the effect of guarantee/collateral, as, if statistically significant, this variable can be taken into account in calculating loan loss provisions on bad and doubtful loans.

Additional explanatory variables have also been tested: the number of years of the client's relationship with the bank, the annual GDP rate of growth, the frequency of default in the industry sector, the rating of the borrower and the interest rate on the loan. The number of years of relationship could have an effect on the effort of a distressed borrower to repay his or her debt, to protect the information-based value created by the relationship. GDP growth or the frequency of default in the industry could affect the level of recovery, as some studies (for instance, Frye, $2000_{a,b}$ and 2003; Altman *et al.,* 2003; and Acharya *et al.* 2003_{a}) have found a negative correlation between economic activity and recovery level.

The size of loan, GDP growth, age of the firm, and length of relationship excepted, the explanatory variables will be represented by 'dummy' variables. The dependent variable, the cumulative loan recovery rate, is a continuous variable over the interval [0-1]. Due to the boundaries of the dependent variable, one cannot use the ordinary least square (OLS) regression,

$$E(y^* x) = \beta_1 + \beta_2 x_2 + ... + \beta_k x_k = x \beta \quad \textbf{(1)}$$

as it cannot guarantee that the predicted values from the model will lie in the bounded

[15] As we had access to information on the bank's internal rating on a subset of loans, we also attempted to test the impact of ratings with this subset of data. As the rating explanatory variable was found not to be statistically significant, the results are not reported.

interval (Greene, 1993). A common econometric technique is to use a transformation G(y) that maps the [0-1] interval onto the whole real line [- ω , + ω] (McCullagh and Nelder, 1989). There are several possible functional forms, but the most common ones are the cumulative normal distribution, the logistic function, or the log-log function. The logistic function is defined as:

$$G(x\beta) = \frac{\exp(x\beta)}{1+\exp(x\beta)} \quad (2),$$

and the log-log function is defined as:

$$G(x\beta) = e^{-e^{-x\beta}} \quad (3).$$

The cumulative normal distribution and the logistic function are symmetrically distributed, while the log-log function is asymmetric.[16] This might be more appropriate with our data, since there is a significant concentration of observations near the extreme value '1'. Following up on Papke and Wooldridge (1996), the non-linear estimation procedure maximizes a Bernoulli log-likelihood function:

$$l_i\,(b) = y_i\,[\log G\,(x_i\, b)] + (1-y_I)\log\,[1-G(x_i\, b)] \quad (4)$$

The quasi-maximum likelihood estimators (QMLE) are consistent and asymptotically normal (Gourieroux, Montfort and Trognon, 1984).

Empirical results

In Table 7, we report the empirical results for the base case of the log-log model. The explanatory variables include the size of the loan, the types of collateral/guarantees, the year and the industry sector. In the category of firms with multiple defaults, we have retained the first default case. In the robustness checks to be discussed later, we evaluate the effect of

[16] La Porta *et al.* (2003) use a Tobit approach. Developed for censored data, this methodology can ensure that the fitted values will be bounded downwards at 0, but this approach cannot ensure that they would be bounded upward at unity (Greene, 1993). Acharya *et al.* (2003_a) report OLS estimates.

including the multiple-default cases. The model is applied to cumulative recovery rates after 12, 24, 36 and 48 months.[17] From a loan pricing or capital regulation perspective, the estimate of the long-run cumulative recovery is the relevant variable. However, it is also of interest to understand the determinants of interim recoveries. Given the timing of the data, 1995-2000, the number of observations varies as we have 316 cases with a recovery rate twelve months after the default event, and 154 cases with a recovery rate after 48 months.

A first observation in Table 7 is that the size of the loan has a statistically significant negative impact on the recovery rate at all horizons (12-, 24-, 36-, and 48-month). This confirms the sample univariate unweighted and weighted average cumulative recovery rates reported earlier and the results of Hurt and Felsovalyi (1998), according to whom the recovery on large loans is lower than that on smaller loans. A tentative explanation is that, for this bank, the 'option' value of the relationship or the 'macro' considerations discussed earlier delay the foreclosure of larger loans, with a negative effect on future recoveries. A second observation is that, as expected, the collateral variables (real estate, physical or financial) have a statistically significant positive effect on recovery for the 48-month horizon. For shorter horizons, the effect is positive but not statistically significant, possibly because collateral will not be seized in the short term. As was the case with the univariate figures, personal guarantees have a negative (although not statistically significant) effect on recovery at all horizons.[18] Some of the year dummies are significant, highlighting a volatility of recovery over time. The year 1997 exhibits good recoveries, while the year 1999 exhibits lower recoveries. This confirms the univariate finding discussed earlier and the expectation that, due to a reorganization of the workout unit, recoveries would be lower in 1999. If there was a positive impact of the economy in that year, it was dominated by the reorganization of the workout unit mentioned earlier. The industry dummies are significant and negative in most cases, confirming the observation that recoveries in the hotel/restaurant sector -the base case- are higher. With regard to the overall estimation, the Wald test for a null hypothesis, that the vector of estimated coefficients is not different from zero, is strongly rejected. The low R^2 is expected for this type of analysis. As many loans are not collaterized, there is a large

[17]The 48-month period was chosen in order to allow us to keep two years of data (1995 and 1996) to run the statistical tests.

[18] A director of the bank explained that this result is likely to be due to two factors. First, guarantee or collateral support is not usually requested from good clients, so that the existence of a guarantee is an indicator of greater risk. Second, some borrowers are able to shift ownership of personal assets to other persons, so that, when the bank tries to execute the debt, there is not much left.

variation in recovery rates for the non- collaterized loans that is not explained by our limited set of explanatory variables.

Given the large number of explanatory variables, an alternative specification is reported in Table 8. Collateral, whatever its source, is represented by one dummy variable, and the 15 industrial sectors are grouped into the four aggregate sectors (real, manufacturing, trade, and services). Finally, the age of the firm is added. The results are confirmed. In other words, the loan size has a negative effect, the collateral variable is positive, the manufacturing and trade sectors have lower recoveries than the real sector, and the age of the firm has a positive impact on recoveries. This last effect can be interpreted as follows. Older firms may display less opaqueness on the quality of management and value of assets, helping, *ceteris paribus*, to obtain better recoveries at the time of distress.

As indicated above, additional explanatory variables have been tested: annual GDP growth, frequency of default in the industry sector, the number of years of relationship with the bank and the level of interest rates. The first variables are included to test for the effect of the economy on recoveries. None of these variables were significant. This could be due to the fact that, during the period 1995-2000 , there was no significant recession and that the idiosyncratic event of 1999 -reorganization- was the significant variable. The other two variables, the number of years of relationship with the bank and the level of interest rate were non-significant.

Robustness tests

Four types of robustness test have been done. A first test was to estimate the symmetrical logistic function. The next two tests involved the random selection of 90% of the observations, and the inclusion of the multiple-default cases in the data set. Finally, to ensure that the size effect was not driven by low recoveries on a few large loans, the 10% largest loans were eliminated from the sample. The results are fully consistent with those of the base specifications, confirming the statistical significance of the loan size, collateral, yearly dummy, industry sector, and age of the firm. For the sake of space, the estimated parameters are not reported.[19]

Marginal effects

The partial effect of x_i on *E(y|x)* is for the log-log function equal to:

$$\frac{\partial E(y)}{\partial x_i} = \frac{\partial}{\partial x_i}(e^{-e^{x\beta}}) = \frac{\partial}{\partial x_i}(-e^{-x\beta}).e^{-e^{x\beta}} = \beta.e^{-x\beta-e^{-x\beta}}.$$

Given the non-linearity of the functional specification, the marginal effects of explanatory variables on recovery rates are not constant. They can be calculated for specific values of the explanatory variables. For discrete value, such as collateral/no collateral, the loan recovery log-log function, G (XB), is calculated with and without collateral. The marginal effect is then calculated as the relative increase in loan recovery rate when a loan is collaterized. The estimation of marginal effects is reported in Table 9. For instance, in the case of an average loan size of €142,180 with no guarantee/collateral, the loan recovery rate increases by a relative 31% when a collateral is added. With regard to the size effect, the (negative) relative difference in recovery on the largest loan in the sample *vis-B-vis* recovery on the average loan is - 70%. Similarly, recovery on the youngest firm *vis-B-vis* average is - 12%, while recovery in the manufacturing sector is 26% lower than in the hotel/restaurant sector.

Section 6. Workout Costs Incurred with Recoveries

So far, the analysis has been concerned with *gross* recoveries. For loan pricing, the calculation of LGD and capital requirement, or the calculation of loan loss provision, one would need to know the recoveries *net* of the cost incurred by the bank to recover these cash flows. Although the argument that the data are specific to one bank applies in this case too, the figures on direct recovery costs are reported, because, to the best of our knowledge, no such information is available in the literature.

At Banco Comercial PortuguLs, two departments handle bad and doubtful loans: the *internal*

[19] The estimates of the robustness tests are available from the authors upon request.

restructuring department and the *contentious* department. Internal restructuring refers to effort by the bank to recover cash on its own. Contentious actions refer to the use of external lawyers or law courts to recover cash. The internal restructuring department is itself divided into two units : the *standardized* department, which deals with loans with a value below €75,000, and the *specialized* department, which deals with larger loans. Data on direct costs incurred in recovery have been collected for the year 2002. They are reported in Table 10. For reasons of confidentiality, all figures have been scaled by a common multiplicative factor. Only percentage figures are therefore relevant. One observes that the average cost of internal recovery amounts to 1.2% of the amount recovered. Not surprisingly, recovery costs on smaller loans are substantially higher than on large loans, 4.1% vs. 0.9%. Once the contentious department has to rely on external lawyers, the recovery costs rise to 10.4%. This higher figure reflects the complexity of cases sent to external lawyers. On average, the total recovery costs incurred by the bank, in 2002, amount to 2.6% of the amount recovered or restructured. A base of comparison is the direct recovery costs incurred on US bankruptcies; White (1996) reports that cost estimates range from 3% to 7% of total assets.[20]

Section 7. Conclusion

Loan losses-given-default were estimated for a sample of 374 corporate loans of a European bank over the period 1995 to 2000. The estimates were based on the discounted value of cash flows recovered after the default event. A univariate mortality-based approach was applied to measure cumulative recovery on bad and doubtful loans. The average recovery estimate of 71% was in the same order as that obtained in US studies. A multivariate approach was then applied to analyse the determinants of recovery rates. Three main conclusions can be drawn from this empirical case study. The first is that the frequency distribution of loan losses-given-default appears bi-modal, with many cases presenting 0% recovery and other cases presenting 100% recovery. Loan portfolio models, based on a fixed recovery estimate or a beta distribution, will not capture this characteristic. The second conclusion is that a

[20] Note that the base of comparison is not totally identical. We report direct cost as a percentage of amount recovered, while many bankruptcy studies report costs as a percentage of the face value of the debt,or as a percentage of the face value of debt plus market value of equity.

multivariate analysis of the determinants of loan losses allows us to identify several statistically significant explanatory variables. These include the size of the loan, collateral, industry sector, year dummies and age of the firm. Third, estimates of direct costs incurred by the bank in recovery are of the same order as those obtained in studies on US bankruptcies. A word of caution is that this study, being based on a dataset of one single bank, can capture some of the bank's idiosyncracies. Given the absence of publicly available data on cash flow recovery, this paper is a step towards a better understanding of the distribution and determinants of bank loan losses-given-default. Finally, the availability of data on recovery over time should open the way to the development of a dynamic measure of loan loss provisioning. This is the object of further work.

Appendix 1: Mortality-based Approach to Evaluate Recovery Rate, an Example

To define the concepts used to measure loan recovery rate, it is, for expository reasons, useful to refer to a simple example. Consider a loan of €100 that enters the 'default' category in December 2000. We track the subsequent payments on this loan, assuming, for expository convenience, that all payments take place at the end of the year. The interest rate is 10%.

	Dec. 2000	Dec. 2001	Dec. 2002	Dec. 2003
Loan outstanding (before cash payment)	100	110	66	44
Cash payment	0	50	26	14
Loan balance (after cash payment)	100	60	40	30

Let us define the *Marginal Recovery Rate* at December 2001, MRR_1 , as the proportion of the outstanding loan in December 2001 that is being paid, one period (in the example, one year) after default:

$$MRR_1 = \text{Cash flow paid}_1 / \text{Loan balance}_1$$
$$= 50/110 = (50 / 1.10) / 100 = 5/11$$

The marginal recovery rate can also be interpreted as the percentage repayment on the loan outstanding, in present value terms.

Let us define the *Percentage Unpaid Loan Balance* after payment in December 2001, $PULB_1$, as the proportion of the December 2001 loan balance that remains to be paid one period after default:

$$PULB_1 = 1 - MRR_1 = 1 - 5/11 = 6/11$$

Similarly, one can define the *Marginal Recovery Rate* at December 2002 as:

$$MRR_2 = \text{Cash flow paid}_2 / \text{Loan}_2$$
$$= 26/66$$

The *Percentage Unpaid Loan Balance* after payment in December 2002, two periods after default is equal to:

$PULB_{\,2} = 1 – MRR_2 \; = 1 – 26/66 = 40/66$

and the *Cumulative Recovery Rate* in December 2002, CRR_2, is defined as:

$CRR_{\,2} = (1 – (PULB_1 \text{ x } PULB_2)$

$\quad = (1 – 6/11 \text{ x } 40/66) = (1 – 240/726) = 1 - 40/121 = 81/121$

$\quad = (81/1.1^2) / 100.$

CRR_T, represents the proportion of the initial default loan that has been repaid (in present value terms), T periods after default.

**Table 1. Loan Portfolio Distribution by Business Sectors,
Sample vs. Banco Comercial PortuguLs, vs. Portuguese Banking System (%)**

	Sample	**Banco Comercial PortuguLs**	**Portuguese Banking System**
1.Agriculture/Fishing	2.7%	3.4%	5.0%
2.Mining	3.0%	1.1%	0.7%
3.Construction	12.1%	8.6%	14.5%
4.Hotel/Restaurant	2.2%	1.6%	na
6.Food/Beverages	6.1%	3.3%	5.9%
7.Textiles	1.3%	8.9%	7.6%
8.Chemicals	2.5%	1.5%	3.3%
9.Machinery	3.2%	4.8%	6.1%
10.Paper/Printing	2.3%	1.4%	1.9%
11.Other Non-mineral	1.3%	na	2.3%
12.13Retail and Wholesale Trade	27.9%	28.1%	28.4%
14. Transport	2.0%	1.9%	4.8%
15. Other Services	33.3%	35.3%	19.5%

Note: Portfolio distribution as of December 31,1995. The sample used in this study includes the loans to Southern Portugal, including Lisbon.

Sources: Banco Comercial PortuguLs Annual Reports; Banco of Portugal Annual Reports.

Table 2. Descriptive Statistics for the Sample of Bad and Doubtful Loans

Panel A: Number of Defaults per year	
1995	65
1996	89
1997	59
1998	57
1999	47
2000	57
Total	374

Panel B: Debt Outstanding at the Time of Default (€)		
	Number of Observations	Percentage
0<Debt<50,000	186	49.7%
50,000<Debt<100,000	79	21.1%
100,000<Debt<150,000	35	9.4%
150,000<Debt<200,000	12	3.2%
200,000<Debt<250,000	15	4.0%
250,000<Debt<300,000	7	1.9%
300,000<Debt<350,000	4	1.1%
Debt> 350,000	36	9.7%
Total	374	100%

Panel C: Number of Loans with Personal Guarantee or Collateral		
	Number of Observations	Percentage
No Guarantee/Collateral	133	35.6%
Personal Guarantee	218	58.3%
Real Estate Collateral	26	7.0%
Physical Collateral	7	1,9%
Financial Collateral	23	6.1%

Panel D: Age of Borrowing Firm and Age of Relationship with the Bank (Years)				
	Mean	Median	Min	Max
Age of Borrowing Firm	17	12.3	0.5	121
Age of Relationship	6	6	0.5	14

Note: In the cases of firms with a history of multiple defaults, only the first default case is included.

Table 3. Number of Default Cases, and Use of Collateral/Guarantee by Industrial Sectors (1995-1999)

Sectors of Activities	Number of defaults		Value-weighted Default Distribution	Number of Defaults with Guarantees (% of number of defaults in that industry)	Number of Defaults with Collateral (% of number of defaults in that industry)
1.Agriculture/Fishing	6	2%	8%	4 (67%)	2 (33%)
2.Mining	8	2%	1%	3 (38%)	1 (13%)
3. Construction	50	13%	15%	28(56%)	4 (8%)
4.Hotel/Restaurant	9	2%	3%	6(67%)	2 (22%)
5.Real Estate	10	3%	4%	6 (60%)	3 (30%)
6.Food/Beverages	12	3%	13%	6 (50%)	2 (17%)
7.Textiles	15	4%	2%	8 (53%)	0 (0%)
8.Chemicals	3	1%	3%	2 (66%)	1 (33%)
9.Machinery	20	5%	2%	9 (45%)	5 (25%)
10.Paper/Printing	12	3%	3%	8 (75%)	4 (33%)
11.Other Non-mineral	15	4%	3%	8 (53%)	2 (13%)
12.Wholesale Trade	122	33%	23%	85 (70%)	13 (11%)
13.Retail Trade	41	11%	9%	23 (56%)	7 (17%)
14.Transport	15	4%	3%	6 (40%)	2 (13%)
15.Other Services	36	10%	10%	16 (44%)	8 (22%)
Aggregated Sectors					
I. Real	83	22%	31%	47 (57%)	12 (14%)
II. Manufacturing	77	21%	25%	41 (53%)	14 (18%)
III. Trade	163	44%	32%	108 (66%)	20 (12%)
IV. Services	51	14%	12%	22 (43%)	10 (20%
Total	374	100%	100 %	218 (58%)	56 (15%)

Notes : With reference to European Union economic activities codes (NACE), the 15 sectors are defined as follows : Sector 1 (Agriculture/Fishing): 1, 2, 5, 20 ; Sector 2 (Mining): 11,13,14 ; Sector 3 (Construction): 45 ; Sector 4 (Hotel/Restaurant): 55 ; Sector 5 (Real Estate): 70 ; Sector 6 (Food/Beverages): 15, 16 ; Sector 7 (Textiles): 17, 18, 19 ; Sector 8 (Chemicals): 23, 24, 25 ; Sector 9 (Machinery): 26 to 37 ; Sector 10 (Paper/Printing): 21, 22 ;Sector 11 (Other mineral; cement) : 26 ; Sector 12 (Wholesale trade) : 50,51 ; Sector 13 (Retail Trade): 52 ; Sector 14 (Transport): 60 to 64 ; Sector 15. (Other Services) : 71 to 93.

Aggregated sectors : Real (sectors 1 to 5) ; Manufacturing (sectors 5 to 11) , Trade, (sectors 12+13), Services (sectors 14+15).

Table 4. Univariate Statistics on Recovery Rates

Sample Unweighted Cumulative Recovery Rates, a Pool-based Approach					
	Mean	Median	Minimum	Maximum	Standard Deviation
12-month cumulative recovery	52%	49%	0%	100%	44%
24-month cumulative recovery	65%	91%	0%	100%	41%
36-month cumulative recovery	69%	.95%	0%	100%	38%
48-month cumulative recovery (total sample)	71%	95%	0%	100%	37%[21]
48-month cumulative recovery (loans with no guarantee/collateral)	76%	92%	0%	100%	33%
48-month cumulative recovery (loans with personal guarantee only)	64%	93%	0%	100%	42%
48-month cumulative recovery (loans with collateral)	92%	98%	55%	100%	12%

Note: The pool-based approach includes for each horizon (12 months, 24 months, 36 months and 48 months) the subset of loans with recovery data available for that horizon.

[21]This standard deviation of cumulative loan recovery can be compared to lower estimates of 32.7% calculated by RiskMetrics (1997) with Asarnow and Edwards'data, the 29% of Carty and Lieberman (1996), and the 28.8% of Hurt and Felsovalyi (1998).

Table 5. Recovery and Frequency of Default by Industrial Sectors

	48-month Unweighted (Weighted) Average Cumulative Recovery (1995-1996)	12-month Unweighted (Weighted) Average Cumulative Recovery (1995-1999)	One-year Frequency of Default for the Number (volume) of Loans (1995-1999)
1.Agriculture/Fishing	98.1 % (96.5%)	64.2% (0.04%)	3.2% (3.6%)
2.Mining	91.3% (87%)	16.5% (0.007%)	6.8% (1%)
3.Construction	85 % (71%)	54% (33%)	4.5% (1%)
4.Hotel/Rrestaurant	98% (98%)	81% (36%)	3.8% (0.6%)
5.Real Estate	(87% (84%)	66% (77%)	2.5% (0.2%)
6.Food/Beverages	79.4% (68%)	41.1% (52.6%)	3.7% (2.4%)
7.Textiles	66.1 % (63%)	48.3% (36.4%)	10.2% (1.8%)
8.Chemicals	74.2 % (84.1%)	22% (30%)	1.1% (1.2%)
9.Machinery	63% (65%)	59% (39%)	3.4% (0.6%)
10.Paper/Printing	80% (92%)	44% (37%)	4.5% (1.7%)
11.Other Non-mineral	72% (78%)	62% (50%)	5.9% (2.2%)
12.Wholesale Trade	65% (36%)	53% (35%)	3.7% (1.1%)
13.Retail Trade	55% (52%)	48% (51%)	5.0% (2.4%)
14.Transport	70% (74%)	40% (45%)	3.6% (0.7%)
15. Other Services	88% (86%)	52% (25%)	2.8% (0.2%)

Note: The pool-based approach includes for each horizon (48 months and 12 months) the subset of loans available for that horizon.

Table 6. Annual Rate of Growth of GDP, Frequency of Default, and 12-Month Recovery Rates (1995-2000)

	1995	1996	1997	1998	1999	2000
Real (nominal) GDP Growth	2.4% (8.1%)	3.6%(6.3%)	3.7%(7.3%)	4.2% (7.9%)	3.8%(6.5%)	3.9% (6.5%)
Frequency of default for the number (volume) of loans (total sample)	4.3% (1.2%)	5.9% (1.3%)	4.0%(2.2%)	3.4%(0.4%)	2.7% (0.3%)	3% (0.2%)
12-month unweighted (weighted) average recovery (total sample)	50% (26.5%)	55% (45%)	66% (33%)	52% (49%)	32% (24%)	n/a
Frequency of default for the number (volume) of loans for Real sector	5.9% (1.5%)	4.1 % (2.0%)	3.4% (4.7%)	3.6% (0.5%)	3.2% (0.3%)	4.1% (0.2%)
12-month unweighted (weighted) average recovery for Real sector	63% (25%)	69% (66%)	44% (5%)	51% (62%)	44% (48%)	n/a
Frequency of default for the number (volume) of loans for Manufacturing sector	5.4% (1.1%)	6.5% (11.9%	3.5% (0.6%)	4.6% (1.1%)	3.0% (0.4%)	2.5% (0.5%)
12-month unweighted (weighted) average recovery for Manufacturing sector	57% (61%)	54% (50%)	74% (65%)	41% (18%)	30 % (18%)	n/a
Frequency of default for the number (volume) of loans for Trade sector	3.9% (2.2%)	7.1% (3.2%)	4.8% (3.4%)	3.0% (0.9%)	2.4% (0.9%)	3.0% (0.4%)
12-month unweighted (weighted) average recovery for Trade sector	40% (18%)	55% (26%)	64% (68%)	62% (62%)	27% (24%)	n/a
Frequency of default for the number (volume) of loans for Services sector	1.9% (0.2%)	4.7% (1.2%)	3.6% (0.7%)	3.1% (0.2%)	2.4% (0.2%)	2.7% (0.1%)
12-month unweighted (weighted) average recovery for Services sector	40% (19%)	44% (31%)	88% (55%)	45% (43%)	25% (5%)	

n/a : not available

Table 7. Log-log estimates of Cumulative Recoveries

Explanatory variable	12-month cumulative recovery (p-value)	24-month cumulative recovery (p-value)	36-month cumulative recovery (p-value)	48-month cumulative recovery (p-value)
Constant	1.65 (0.03*)	1.62 (0.02*)	3.24 (0.00*)	4.57 (0.00*)
Loan Size	- 0.58 (0.01*)	-0.86 (0.00*)	-1.18 (0.00*)	-1.26 (0.00)
Personal Guarantee	- 0.16 (0.31)	-0.21 (0.29)	-0.05 (0.83)	-0.35 (0.23)
Real Estate Collateral	0.28 (0.4)	0.29 (0.59)	0.67 (0.28)	1.98 (0.00)
Physical Collateral	- 0.36 (0.5)	0.58 (0.60)	-0.02 (0.99)	2.93 (0.00)
Financial Collateral	0.43 (0.30)	0.38 (0.42)	0.39 (0.5)	2.09 (0.02)
Year 1996	0.34 (0.10)	0.44 (0.06)	0.41 (0.1)	0.34 (0.20)
Year 1997	0.69 (0.01*)	0.54 (0.06)	0.52 (0.08)	
Year 1998	0.09 (0.69)	0.11 (0.68)		
Year 1999	- 0.47 (0.04*)			
1.Agriculture/Fishing	- 0.68 (0.43)	0.16 (0.84)	- 0.78 (0.49)	-0.56 (0.64)
2.Mining	- 2.19 (0.01*)	- 1.49 (0.06)	-1.49 (0.19)	-2.09 (0.02*)
3.Construction	-1.11 (0.16)	-0.82 (0.24)	-2.15 (0.04*)	-2.58 (0.00*)
5.Real Estate	-0.59 (0.53)	0.32 (0.76)	-1.70 (0.17)	-3.21 (0.00*)
6.Food/Bbeverages	-1.13 (0.21)	0.14 (0.89)	-0.35 (0.78)	-1.73 (0.08)
7.Textiles	-1.31 (0.11)	-1.14 (0.12)	- 2.53 (0.02*)	-3.57 (0.00*)
8.Chemicals	- 1.72 (0.04*)	-0.41 (0.71)	-1.65 (0.21)	-2.38 (0.04*)
9.Machinery	-1.09 (0.18)	-0.89 (0.23)	-2.62 (0.02)	-4.05 (0.00*)
10.Paper/Printing	-1.10 (0.16)	-1.51 (0.08)	-1.74 (0.28)	-4.32 (0.00*)
11.Other Non-mineral	-1.10 (0.18)	-0.84 (0.26)	-2.23 (0.05*)	-3.46 (0.00*)
12.Wholesale Trade	-1.21 (0.11)	-0.91 (0.16)	-2.60 (0.01*)	-3.65 (0.00*)
13.Retail Trade	-1.48 (0.05*)	-1.08 (0.12)	-2.83 (0.006*)	-4.00 (0.00*)
14.Transport	-1.81 (0.02*)	-1.51 (0.03*)	-2.69 (0.01*)	-3.62 (0.00*)
15.Other Services	-1.27 (0.10)	- 0.41 (0.57)	-1.42 (0.21)	-2.47 (0.01*)
Wald (Qui-squared) Test (p-value)	55.98 (0.00*)	34.7 (0.042*)	46.44 (0.00*)	390.8 (0.00*)
Reset Test (p-value)	0.58 (0.56)	-0.21 (0.83)	-1.03 (0.30	-2.58 (0.01)
Pseudo R^2	0.13	0.10	0.13	0.20
Number of Observations	317	270	213	154

* Represents significance at the 5% level.
The table presents the estimation of the log-log regression for the cumulative recovery rates at four horizons, respectively 12-, 24-, 36-, and 48-months. Cumulative recoveries are measured in cents per euro. The loan size is €1 million. Collateral, year, and industry sectors are represented by dummies.

Table 8. Log-log estimates of Cumulative Recoveries

Explanatory variable	12-month cumulative recovery (p-value)	24-month cumulative recovery (p-value)	36-month cumulative recovery (p-value)	48-month cumulative recovery (p-value)
Constant	0.49 (0.04*)	0.72 (0.01*)	1.02 (0.00*)	1.79 (0.00*)
Loan Size	-0.66 (0.00*)	-0.76 (0.00*)	-0.84 (0.00*)	-0.77 (0.00*)
Personal Guarantee	-0.17 (0.26)	-0.27 (0.16)	-0.14 (0.53)	-0.38 (0.15)
Collateral	0.31 (0.18)	0.61 (0.07)	0.69 (0.08)	1.72 (0.00*)
Year 1996	0.23 (0.26)	0.34 (0.13)	0.35 (0.16)	0.31 (0.24)
Year 1997	0.63 (0.01*)	0.49 (0.07)	0.41 (0.15)	
Year 1998	0.10 (0.67)	0.12 (0.63)		
Year 1999	-0.48 (0.04*)			
II.Manufacturing Sector	-0.30 (0.21)	-0.36 (0.19)	-0.33 (0.35)	-1.20 (0.02*)
III.Trade Sector	-0.25 (0.25)	-0.25 (0.33)	-0.66 (0.04*)	-1.25 (0.01*)
IV.Services Sector	-0.29 (0.26)	-0.00 (0.99)	0.09 (0.83)	-0.42 (0.48)
Age of firm	0.01 (0.03*)	0.01 (0.02*)	0.01 (0.04*)	0.01 (0.04*)
Wald (Qui-squared) test (p-value)	37.6 (0.00*)	31.31 (0.00*)	38.48 (0.00*)	43.43 (0.00*)
Reset Test (p-value)	1.59 (0.11)	0.25 (0.80)	0.87 (0.39)	0.23 (0.82)
Pseudo R^2	0.11	0.08	0.11	0.18
Number of Observations	316	269	212	153

* Represents significance at the 5% level.

The table presents the estimation of the log-log regression for the cumulative recovery rates at four horizons, respectively 12-, 24-, 36-, and 48-months. Cumulative recoveries are measured in cents per euro. The loan size is €1 million. Collateral, year, and industrial sectors are represented by dummies. The age of the firm is in number of months.

Table 9. Marginal Impact on 48-Month Cumulative Recovery

Variable	Marginal Impact
Collateral	31%
Manufacturing	- 26%
Largest vs. Mean Size	-70%
Smallest vs. Mean Size	2.4%
Oldest vs. Average Age of Firm	33%
Youngest vs. Average Age of Firm	-12%

Note: The marginal impact is calculated with the parameters of the regression reported in Table 8. The base case is a company in the Real sector, with average age and size, in the year 1996, with no personal guarantee. The marginal impact is the relative change in recovery rate when one variable is changed *vis-B-vis* the base case.

Table 10. Workout Costs incurred in Recovery (2002)

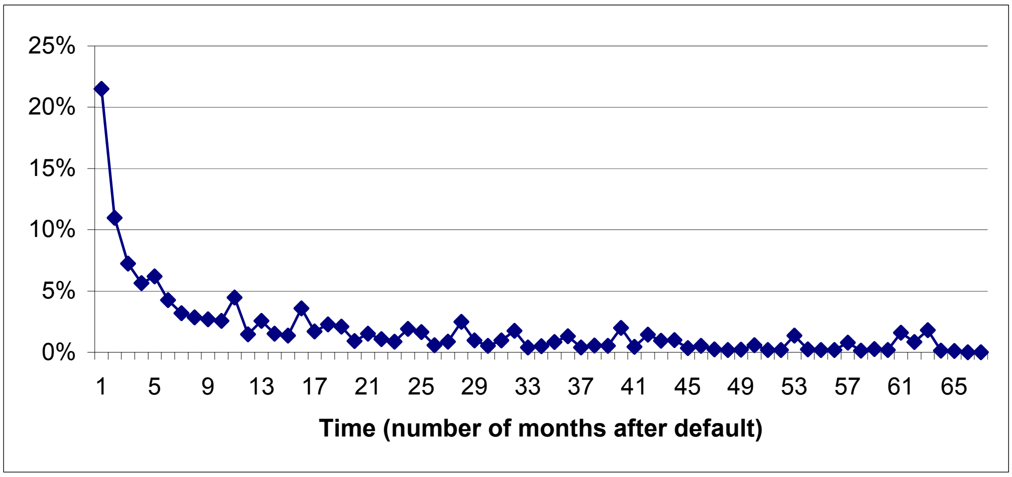
Internal Recovery Department			
	Standardized Unit	Specialized Unit	Total
Total internal costs	296	727	1023
Amount Recovered/Restructured During the Year	7252	78 ,000	85,252
Internal Recovery Cost per Euro Recovered (%)	4.1%	0.9%	1.2%

Contentious Department	
Internal Contentious Cost	278
External Lawyers and Court Expenses	1257
Total Internal and External Cost	1535
Cash Flows Recovered	14748
Contentious Recovery Cost per euro (%)	10,4%

Total Direct Cost (Internal and Contentious)	
Total Internal and External Cost	2,558
Total Amount Recovered	100,000
Average Recovery Cost per Euro (%)	2,6%

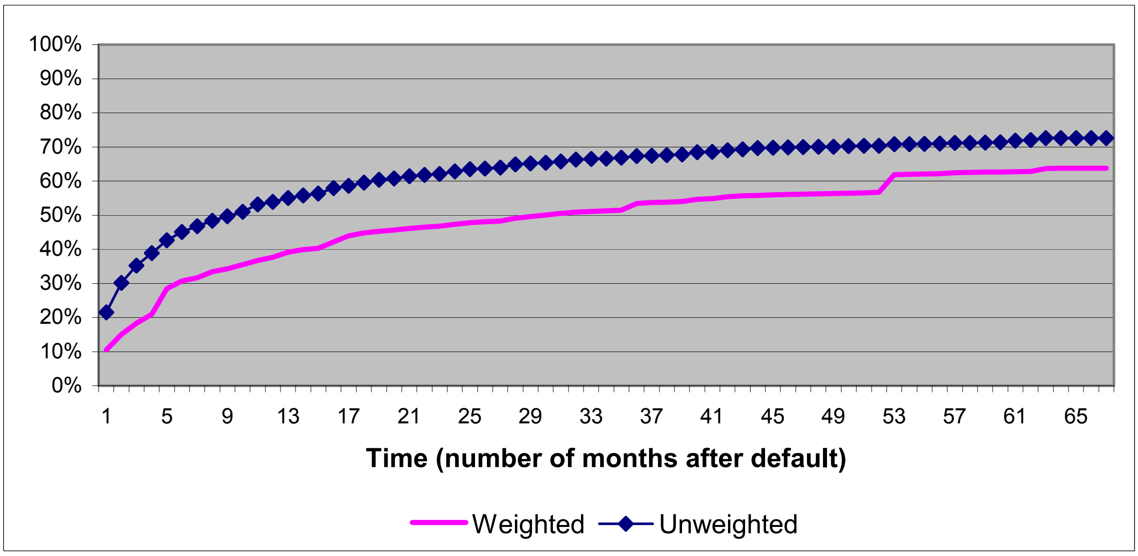
Note: This table reports the workout direct cost incurred in recovery by Banco Comercial Portugues in 2002. For reasons of confidentiality, the absolute figures have been scaled by a common factor. Only percentage figures are relevant. The standardized unit deals with loans with a value below €75,000, and the specialized unit deals with larger loans. The contentious department refers the cases to external lawyers or law courts.

Figure 1. Sample Unweighted Marginal Recovery Rate at time t+n ($SMRR_{t+n}$)



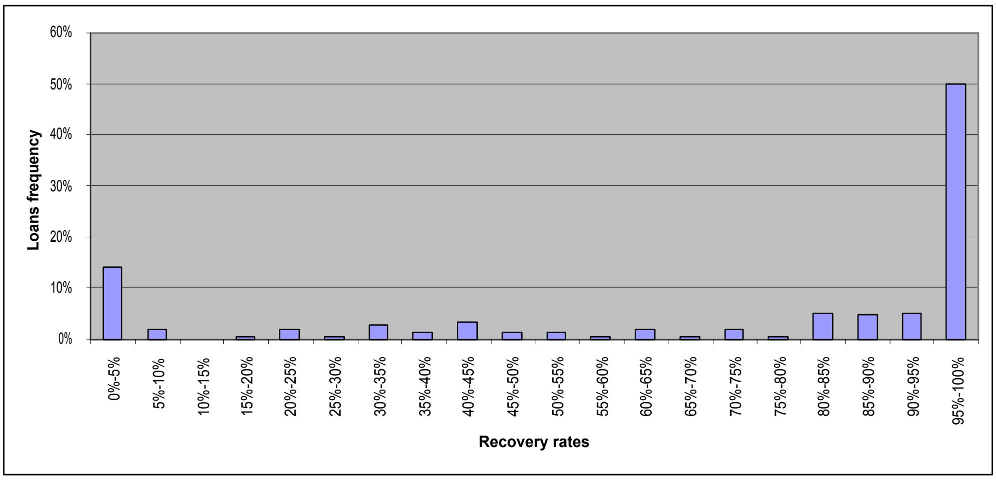
Note: This figure presents the marginal recovery *n*-months after default. The mortality-based approach is used to calculate the marginal recoveries.

Figure 2: Sample Unweighted and Weighted Cumulative Recovery Rate at time t+n (SCRRt+n)



Notes: The figure presents the cumulative weighted and unweighted recovery rates n-months after default. They have been calculated with the mortality-based approach.

Figure 3. Distribution of Cumulative Recovery Rates 48 Months after Default



Note: The figure presents the frequency of cumulative recovery rates on individual loans. Due to data limitation (five years), the cumulative recovery is calculated up to 48 months after default. This does not seem restrictive as Figure 2 indicates that most of the recovery is achieved 48 months after default.

References

Acharya, V.V., S.T. Bharath, A. Srinivasan, 2003_a . Understanding the Recovery Rates on Defaulted Securities. Mimeo, 1-35.

Acharya, V., I. Hasan, A. Saunders, 2003_b. The Effects of Focus and Diversification on Bank Risk and Return: Evidence from Individual Loan Portfolios. Journal of Business, forthcoming.

Altman, E., 1989. Measuring Corporate Bond Mortality and Performance. Journal of Finance, 44, 909-922

Altman, E.J. B. Brady, A. Resti, A. Sironi, 2003. The Link between default and Recovery rates: Theory, Empirical Evidence and Implications. Journal of Business, forthcoming.

Allen, L., A. Saunders, 2003. A Survey of Cyclical Effects in Credit Risk Measurement Models. BIS Working Papers 126, 1-32

Allen, L., G. DeLong, A. Saunders, 2004. Issues in the Credit Risk Modeling of Retail Markets. Journal of Banking and Finance, 28 (4), 727-752.

Altman, E., H. Suggitt, 2000. Default Rates in the Syndicated Bank Loan Market: A Mortality Analysis. Journal of Banking and Finance, 24, pp 229-253.

Altman, E., B. Brady, A. Resti, A. Sironi, 2003. The Link between Default and Recovery Rates: Implications for Credit Risk Models and Procyclicality. Mimeo, 1-46.

Asarnow, E., D. Edwards, 1995. Measuring Loss on Defaulted Bank Loans: a 24-Year Study. The Journal of Commercial Lending, 77, 11-23.

Basel Committee on Banking Supervision, 2004. International Convergence of Capital Measurement and Capital Standards. Basel.

Caouette, J., Altman E., J., P. Narayan, 1998. Managing Credit Risk : The Next Great Financial Challenge. New York: John Wiley & Sons.

Carey, M., 1998. Credit Risk in Private Debt Portfolios. The Journal of Finance, LIII, August, 1363-1387.

Carty, L., D. Lieberman, 1996. Defaulted Bank Loan Recoveries. *Moody's Investors Service*, November.

Credit Metrics, 1997. Technical Document. JP Morgan.

Credit Suisse Financial Products, 1997. CreditRisk+, A Credit Risk Management Framework, Technical Document.

Crouhy, M., D. Galai, R. Mark, 2000. A Comparative Analysis of Current Credit Risk Models. Journal of Banking & Finance, 24, 59-117.

Dahl, D., A. Logan, 2002. Granularity and International Diversification: An Empirical Analysis of Overdue Claims at Banks. Mimeo, 1-28.

Dewenter, K.L., A.C. Hess, 2004. Are Relationship and Transactional Bank Different ? Evidence from Loan loss Provisions and Write-Offs. Mimeo, 1- 46.

Duffie D., K.J. Singleton, 2003. Credit Risk. Princeton Series in Finance.

Frye J., 2000_a .Collateral Damage. Risk, April, 91-94.

Frye J., 2000_b. Depressing Recoveries. Risk, November, 108-111.

Frye J., 2003. A False sense of Security. Risk, August, 63-67.

Gourieroux, C., A. Monfort, A. Trognon, 1984. Pseudo-Maximum Likelihood Methods : Theory. Econometrica, 52, 681-700.

Greene W.H., 1993. Econometric Analysis. MacMillan.

Grossman, R., W.T. Brennan, J. Vento, 1998. Syndicated Bank Loan Recovery. CreditMetrics Monitor, First Quarter, 29-36.

Hurt, L., Felsovalyi A., 1998. Measuring Loss on Latin American Defaulted Bank Loans, a 27-Year Study of 27 Countries". The Journal of Lending and Credit Risk Management, 80, 41-46.

Jimenez, G., J. Saurina, 2002. Loan Characteristics and Credit Risk. Bank of Spain, 1-34.

Jimenz, G., J. Saurina, 2003. Collateral, Type of Lender and Relationship Bank as Determinants of Credit Risk. Journal of Banking & Finance, 27,1-21

La Porta, R., Lopez-deSilanes F., G. Zamarripa, 2003. Related Lending. The Quarterly Journal of Economics, February, 231-268.

McCullagh, P., J.A. Nelder, 1989. Generalized Linear Models. Sccond edition, London: Chapman and Hall.

Nickell, P., W. Perraudin, S. Varotto, 2000. Stability of Rating Transitions", in Special Issue on Credit Risk Modelling and Regulatory Issues. Journal of Banking & Finance, 24 (1/2), 203-228.

Papke, L.E., J. M. Wooldrigdge, 1996. Econometric Methods for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates. Journal of Applied Econometrics, 11, 619-632.

Ramsey J.B., 1969. Tests for Specification Errors in Classical Linear Least Squares Regression

Analysis. Journal of the Royal Statistical Society, Series B 31, 350-371.

Salas, V., J. Saurina, 2002. Credit Risk in Two Institutional Regimes : Spanish Commercial and Savings Banks. Journal of Financial Services Research, 22, 203-224.

Schuermann, T., 2004. What do We Know about Loss Given Default?. In D. Shimko (ed.), Credit Risk Models and Management, 2nd Edition. London : Risk Books.

Shleifer, A., R.W. Vishny, 1992. Liquidation Values and Debt Capacity: A Market Equilibrium Approach. Journal of Finance, XLVII (4), 1343-1366.

White, A. L., 1996. Bankruptcy Resolution: Direct Costs and Violation of Priority Claims. In Corporate Bankruptcy, J.S. Bandhari and L.A. Weiss editors. Cambridge: Cambridge University Press.