Revista de Análisis Económico, Vol. 19, Nº 1, pp. 31-49 (Junio 2004)

EXPLAINING AND PREDICTING BANK FAILURE USING DURATION MODELS: THE CASE OF ARGENTINA AFTER THE MEXICAN CRISIS*

MARCELO DABOS
Universidad de San Andrés

WALTER SOSA ESCUDERO
Universidad de San Andrés

Abstract

This paper studies the role played by several bank specific financial indicators in determining the process of bank failure in Argentina after the Mexican crisis known as the "tequila effect". Due to the relative scarcity of previous studies, this paper priorizes the use of semiparametric and non-parametric methods which allow us to measure the effect of bank specific financial explanatory variables in the process of bank failure together with duration dependence effects without the need of arbitrary and possibly unrealistic assumptions. The dynamic of bank failures can be fairly characterized by observable factors, which discards the possibility that it had been governed by contagion processes solely. The non-monotonocity of the implicit hazard rate suggests that there were contagion effects, and that they had a strong influence in the first 200 days of the crisis.

I. Introduction

In December 20th 1994, the day in which the Mexican devaluation that originated the bank crisis known as "tequila effect" occurred, 206 financial institutions existed

* Corresponding author: Marcelo Dabós, Department of Economics, University of San Andres, Victoria, Argentina. E-mails: dabos@udesa.edu.ar, wsosa@udesa.edu.ar. We thank an anonymous referee for useful comments who helped to improve the final version considerably. The usual disclaimer applies.

in Argentina. In August 1995 this number had fallen to 156 as a consequence of the financial reorganization after the crisis. The fall in the number of financial institutions (defined as banks, financial companies, saving and loan mortgage companies, and loan companies), was around 25%. In September 1997, only 143 institutions existed.

Dabós (1998) analyzed this “mortality” pattern for a sample of cooperative banks, strongly rejecting the hypothesis that failures occur in a non-systematic way, since several observable bank specific factors are jointly statistically significant variables in a binary choice bank failure model. A peculiar result is that in spite of having a good predictive power, the model has troubles isolating the role played by competing factors. In this paper we tackle this issue by using a broader sample and by considering *timing* until failure instead of just whether a bank failed or not. This increase in the sample size and the fact that a binary indicator can be understood as a censored version of a duration add sufficient sampling variability to build and estimate a model that not only predicts but also has good explanatory power.

The use of duration models to explain and predict failures of financial institutions is relatively recent. Cole and Gunther (1995), González-Hermosillo, Pazarbasioglu and Billings (1996), Lane, Looney and Wansley (1986), Weelock and Wilson (1995) and Whalen (1991) have produced results in this direction using different duration techniques to analyze bank failures. Due to the absence of previous work using this methodology for the Argentine case, this paper priorizes the use of semiparametric and non-parametric methods, which allows to measure the effect of relevant variables that determine bank failures together with duration dependence effects, without the need of arbitrary and possibly not realistic assumptions. Hence a contribution of this paper is also methodological, establishing a framework that can be suitably extended to study other episodes.

Banking crises are persistent around the world. Glick and Hutchison (1999) found banking crises in 71 both developed and less developed countries during the 70’s, 80’s and 90’s. Kaminsky and Reinhart (1999) examined the empirical regularities and the sources and scope of problems in the onset of 76 currency crises and 26 banking crises. Argentina is only an example of a much broader problem of banking crises around the world. Argentina suffered several banking crises in 1980-1982, 1989-1990, 1995-1997 and 2001. McCandless *et al.* (2003) is useful explanation of the latest banking crisis in Argentina.

The paper is organized as follows. Section II describes the main facts to be studied, namely, the events that led to the bank failure process that took place in Argentina after the Mexican crisis. Section III is methodological and presents the techniques utilized to construct a model for bank failures. Section IV contains estimation results and some empirical exercises to check the validity of the model. Section V concludes.

II. The Bank Failure Process in Argentina

Two groups of banks were considered for the analysis: mutual, and private national banks. Public and foreign banks were not included mainly because it is not clear if they did not failed thanks to the backup of the state, independent of fundamentals, as in the case of the public banks, or to the international help in the case of foreign banks, altering the mechanism through which observable bank-specific factors impact on its failure probability. Non-bank institutions were not considered to keep the analysis constrained to banks, leaving an homogeneous sample that offers enough sampling variability for the analysis.

December 20^{th} 1996 is the last day considered in the analysis. Choosing the correct time window is not a trivial task since it involves a trade-off between homogeneity and sample accuracy. Once a period is chosen for the analysis, for those banks that did not fail in that period it is not possible to know if they will do so in the future, so a longer period would let us study banks' survival with more detail. On the other hand, a longer period would force us to given account of a more heterogeneous time span, possibly including structural changes, which, although interesting and relevant, are not the subject of this investigation.

In December 20^{th} 1994 there were 64 private national banks, as a result of taking out public and mutual banks, local offices of international institutions, and other four banks that, because of their special characteristics, were not included, out of the total financial institutions sample. These last four banks are: MBA Banco de Inversión S.A., Cofirene Banco de Inversión, Banco Federal Argentino and the Banco Caja de Ahorro S.A. From these 64 private national banks considered, only 3 (Crédito Comercial S.A., Extrader S.A. and Multicrédito S.A.) had problems before the crisis had ended (that is, before 146 days passed, or from 20 December 1994 to 15 May 1995). Between May 15^{th} 1995 (end of the crisis), and December 20^{th} 1995, another 11 private banks disappeared from the financial system. So 14 banks disappeared in one year from the beginning of the Tequila crisis. Between December 20^{th} 1995 and December 20^{th} 1996, 8 private national banks disappeared. Hence in December of 1996, from the 64 banks initially considered, only 42 were still operating. In total, the number of disappeared private banks from December 1994 to December 1996 was 22.

Looking at the mutual banks, from the 38 existing institutions in December 1994, 6 disappeared before the crisis ended. Among the remaining ones, another 23 were out of the market before 20 December 1995 (one year after the Mexican crisis began). Between 20 December 1995, and 20 December 1996, two more mutual banks disappeared. In total, out of the 38 mutual banks operating when the crisis started, 31 disappeared between December 20^{th} 1994 and December 20^{th} 1996. Table 1 summarizes this information.

III. Duration Methods for the Bank Failure Process

During the bank crisis generated by the Tequila effect many banks faced the following scenario. Suppose that a bank begins to lose deposits, so all its efforts will be aimed at continue operating. Suppose now that after a week, the bank is still losing deposits and operating. A relevant question that managers, directors, and control authorities would like to answer at that moment is what is the failure probability for the next week, period, or instant, considering that the bank is still operating. Another

TABLE 1

BANK FAILURES IN EACH GROUP

Bank Types	Operating in 12/20/94	Dissapeared			Operating in 12/20/96
		Before 5/15/95	Between 5/15/95 and 12/20/95	Between 12/20/95 and 12/20/96	
Mutual	38	6	23	2	7
Private National	64	3	11	8	42

interesting question is which is the estimated time until its potential failure, given the bank's characteristics.

Duration models allow us to handle these questions in a parsimonious and informative way. Even though a detailed presentation of the duration model exceeds the scope of this paper, in this section we briefly discuss the particular aspects of this model that are pertinent for our analysis. Kalbfleish and Prentice (1980), Klein and Moeschberger (1977), Miller (1981) and Parmar and Machin (1996), among others, are recognised sources with special reference to medical and biological problems. Lancaster (1990) presents a detailed analysis of duration models with emphasis in labor economics applications, and Van Den Berg (2001) and Neumann (1997) are useful surveys of recent results. The usual example where this technique is used is the unemployment duration analysis. Kiefer (1988) is a relevant reference on this case.

In general terms, the variable of interest in this type of models is the time it takes a system to change from one state to another one. Generally, such change is associated with an event (finding a job, a firm's bankrupt, the solution of a labor conflict, a product disappearance of the market, etc.), which indicates the ending of an event whose duration we try to model. This random variable is called a *duration*, it takes positive values, and we will call it T.

If the duration is understood as a continuous random variable that takes positive values, its probability distribution can be characterized by any of the following three functions, its distribution function, $F(t) = Pr\ (T < t)$, its survival function, $S(t) = Pr(T > t) = 1 - F(t)$, or its density function $f(t) = dF(t)/dt = - dS(t)/dt$. As mentioned before, a quantity of interest will be the probability that the event ends in an interval beginning at t given that it has not finished until that moment. Such conditional probability will be:

$$h(t, \Delta) = Pr\ [\ t \leq T \leq t+\Delta \mid T \geq t\ ]$$

The *hazard rate* is defined as:

$$h(t) = lim_{\Delta \to 0}\ h(t, \Delta) / \Delta$$

and it represents the instant probability that the event concludes at t, conditional on the fact that it lasted up to t. Intuitively, the hazard rate approximates the probability of bank failure for the next moment, given that the bank is still operating at t. The relation between the hazard rate and the previous functions is given by:

$$S(t) = exp(-\int_0^t h(s)ds)$$

so it is possible to get any of the three previous functions from this one, so the hazard rate can also be used to characterize the duration.

A duration model is a model for the random variable T, based on any of the functions that characterize it, as previously described. As it is usual in the literature, the analysis is mainly centered on the survival function $S(t)$ and on the hazard function $h(t)$. It is also important to consider the *cumulative hazard function*:

$$\Lambda(t) = \int_0^t h(t)dt = -\log S(t)$$

which will be useful to describe estimators used in the empirical section.

As a first step, it is interesting to analyze the unconditional distribution of T, that is, momentarily ignoring the possibility of using explanatory variables. At this stage, it is cautious to rely on non-parametric methods, which, as mentioned in the Introduction, avoid the risk of using unrealistic assumptions, while being informative about the basic features of the duration process.

The rest of this section shortly describes the estimation and inference techniques used in this paper. The survival function will be estimated using the standard Kaplan-Meier non-parametric estimator. Suppose that failures occur at the moments $t_1, t_2, \ldots, t_d$ and that at period t_i , d_i banks fail. Let Y_i be the number of banks that are operating at period t_i. The Kaplan-Meier estimator of $S(t)$ for $t_i < t$ is:

$$S_{KM}(t) = \prod_{t_i \le t} \left[1 - \frac{d_i}{Y_i} \right]$$

and $S_{KM}(t) = 1$ *if* $t < t_1$. The survival functions are estimated for both bank groups separately. To test the hypothesis of similarity between survival functions of both groups, we use a log-rank test and a Wilcoxon test, which is standard practice in the literature, see Klein and Moechberger (1997, pp. 191-200), for more details.

Estimation of the hazard function is a more delicate matter than that of the survival. We use an *asymmetric kernel* smoothing method similar to the one described in Klein and Moeschberger (1997, pp. 152-163.). Basically, raw estimations of the hazard function are computed based on the Nelson-Aaleen's cummulative hazard function estimator:

$$\Lambda_{NA}(t) = \sum_{t_i \le t} d_i / Y_i$$

for $t_1 \leq t$, and 0 for $t \leq t_1$. The raw hazard function estimates are non zero at the moments where bank failures are observed, and correspond to:

$$h_{NA}(t_i) = \Lambda_{NA}(t_i) - \Lambda_{NA}(t_{i-1})$$

The final estimation corresponds to a kernel smoothed version based on:

$$\hat{h}(t) = b^{-1} \sum_{i=1}^{D} K\left(\frac{t-t_i}{b}\right) h_{NA}(t_i)$$

where b is the bandwidth and $K(\ )$ the kernel function. We use the standard Epanichnikov kernel that becomes asymmetric at the extremes. This modification is necessary because negative durations are not observed, and therefore near to zero it is necessary to use an asymmetric kernel. All routines were implemented in Splus for Windows, and the relevant code is available from the authors.

The incorporation of explanatory variables is not trivial, and depends on the desired interpretation. A flexible modeling strategy was chosen, whose main properties do not depend on ex-ante theoretical or empirical arbitrary assumptions. We also preferred a simple model specification, that would lead to simple economic interpretations. Under these considerations, the semiparametric character of Cox's *proportional hazard model* seems to provide a good balance between analytical simplicity and functional flexibility.

Under this model, explanatory variables affect the hazard rate in a proportional way. More precisely, the hazard rate is modelled as:

$$h(t \mid x) = h_0(t) \exp(\beta' X)$$

where β is a p-vector of coefficients and X is vector of p explanatory variables. The $h_0(t)$ function is called the *baseline hazard function*.

It is important to note that the baseline hazard controls the hazard rate's time behavior since it depends on time only through the baseline hazard. The second factor introduces the effect of explanatory variables in a multiplicative way. Therefore, the coefficients in this representation can be interpreted as semielasticities of the hazard rates with respect to marginal changes in the explanatory variables.

$$\partial \ ln \ h(y|X) \,/\, \partial X_k = \beta_k$$

with $k = 1, \ldots, p$.

The parameters of this model (the β coefficients) will be estimated using Cox's partial likelihood technique (Cox, 1972). The proportional hazard assumption implies that the effect of explanatory variables on the hazard function is constant over time, and works by moving the baseline hazard rate up or down in a proportional way. This assumption can be formally evaluated following the procedure proposed by Harrell (1986), based on Schoenfeld's residuals (Schoenfeld, 1982). In very general terms, under the proportional hazards assumption Schoenfeld's residuals should not be related to the elapsed time, so Harrell (1986) proposes to use the correlation between these residuals and the time ranks to base a test for the null hypothesis of a proportional hazard.

As expected, this strategy has some limitations. The sampling variation in the model is esentially cross-sectional and, consequently, refer to a specific time-span corresponding to that where the Mexican crisis took place. The dynamic aspect of the model, and hence the effect of time varying covariates, is captured through the effect of time through the baseline hazard which, by assumption is common to all banks. Even though this is a testable assumption using the procedures described in the previous paragraph, more sophisticated dynamic mechanisms that allow for time varying covariates are left open as an interesting and relevant research agenda, even though we provide evidence in the next section in favor of the proportional hazards assumption.

IV. Econometric Results

4.1 Indicators definition and data base used

In this paper we use the structure of explanatory variables considered in Dabós (1998) in his bank failure probability analysis. The following coefficients are considered as indicators of the banks' situation (by November 1994):

1) Equity / Assets (IND1)
2) Liabilities / Equity (IND2)
3) Immediate liquidity = (Cash + Public Securities)/ Deposits (IND3)
4) Structural liquidity = (Equity – Fixed Assets) / Liabilities (IND4)
5) Operating expenses / Liabilities (IND5)
6) Arrears portfolio – Losses provisions / Equity (IND6)
7) Return on equity (ROE) (IND7)

These indicators follow the traditional school in the analysis of financial institutions default risk, looking at the principal variables to be considered in the CAMEL model (Capitalization, Assets, Management, Earnings, Liquidity). 1 and 2 are capitalization indicators (C), 6 is an asset quality indicator (A), 7 is a rentability indicator (E), and 3 and 4 are liquidity indicators (L). 5, an efficiency indicator, would be an approximation for (M) management. The data base used is the monthly information of the individual bank's balances released by Argentina's Central Bank

4.2 Basic characteristics of the bank duration process

As a first approximation, a descriptive analysis of the main characteristics of bank's duration process was performed, leaving momentarily aside the possibility of including explanatory variables to enrich the analysis. Table 2 shows some

TABLE 2

DESCRIPTIVE STATISTICS FOR GROUPS OF BANKS

Variable	Mean	Std. Dev.	Min.	Max.
a) Mutual banks (38 banks)				
duration	288.1	226.0	57.0	731.0
ind1	17.1	8.1	8.4	50.2
ind2	5.7	2.2	1.0	10.9
ind3	22.4	8.2	8.6	55.2
ind4	11.5	14.9	0.2	91.2
ind5	16.3	5.2	7.6	27.8
ind6	66.9	45.0	8.3	178.4
ind7	1.3	17.7	– 62.3	17.9
b) Private national banks (64 banks)				
duration	586.0	226.3	37.0	731.0
ind1	17.6	15.4	4.4	95.1
ind2	7.1	4.2	0.1	21.6
ind3	32.8	22.3	4.8	120.3
ind4	15.9	31.7	– 10.9	213.2
ind5	11.4	13.8	0.9	84.7
ind6	58.4	69.8	– 1.4	442.0
ind7	4.9	11.9	– 36.9	27.1

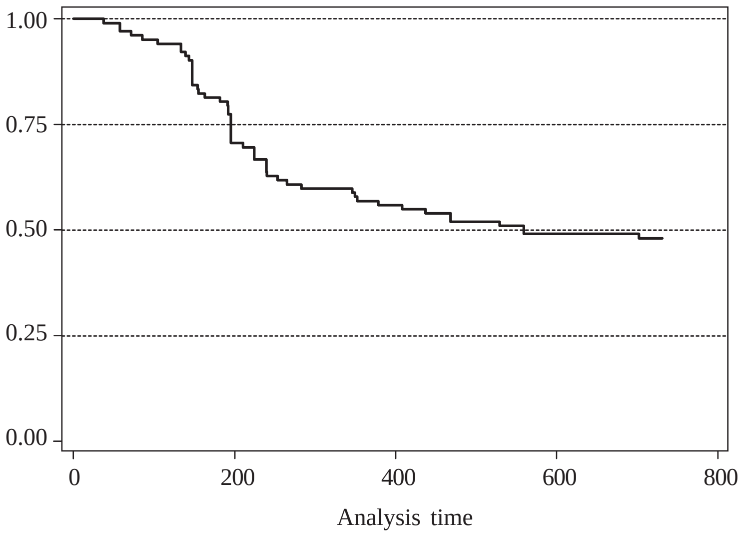
basic statistics of the bank's duration time and of the seven indicators that will later be used as explanatory variables. These statistics were calculated for both bank groups, separately.

Of the 38 mutual banks in the sample, 31 banks failed and 7 survived. During the period we are studying, the average duration of the whole sample in this category was 288 days, while the average duration for the banks that failed was 188 days. In the case of private bank, of the 64 banks in the sample, 22 failed, and 42 survived. The average duration of all the banks in this group was 586 days, and the average for those that failed was 309 days.

As mentioned in the previous section, the duration distribution can be characterized in several ways. Figure 1, shows graphically the estimation of the survival function for all the banks in the sample as a single group, based on the Kaplan-Meier estimator. Figure 2 shows these estimations for both bank groups separately. Table 3 presents a brief description of the bank failure dynamics. For each bank group, each row in the table shows the period in which the failure occurred, together with the information used to construct the previously mentioned Figures. These results suggest that bank mortality dynamics was notoriously different if it is a mutual or a private national bank. This observation is further validated by the survival function similarity tests shown in Table 4. In the same table results for

FIGURE 1

ESTIMATED SURVIVAL FUNCTION FOR ALL THE BANKS AS ONE GROUP



the log-rank and Wilcoxon tests are presented. Both statistics reject the null hypothesis of similarity between survival functions of both groups.

Two interesting effects can be observed. First, the survival function for private national banks is significantly higher than that of mutual banks: *at every moment of the crisis it is more probable to have a mutual bank failure than a private national bank failure*. Second, the survival function for private national banks decreases slower than that of mutual banks, which shows a strong acceleration approximately 200 days after the Tequila crisis began.

Figure 3 shows hazard function estimation results using the smoothing method described in the previous section, for the sample as one group, and for mutual and private national banks separately. As was mentioned in Section III, the hazard rate measures, at each moment, the instant probability that a bank fails, given that is was active until this moment. So this function can be interpreted as a failure risk indicator for the banks that are still operating. The results clearly show a phenomenon suggested by the survival functions estimated before. From the beginning of the crisis, bank failure risk increases monotonically and rapidly until approximately 200 days after the crisis began, and then it begins to fall.

This possible non-monotonicity of the hazard rates narrows the number of possible parametric models that could be chosen to represent bank failure dynamics. For example, the Weibull or exponential models have monotonic or constant hazard rates, which are not compatible with the results shown in this section.

FIGURE 2

ESTIMATED SURVIVAL FUNCTIONS FOR PRIVATE AND COOPERATIVE BANKS SEPARATELY

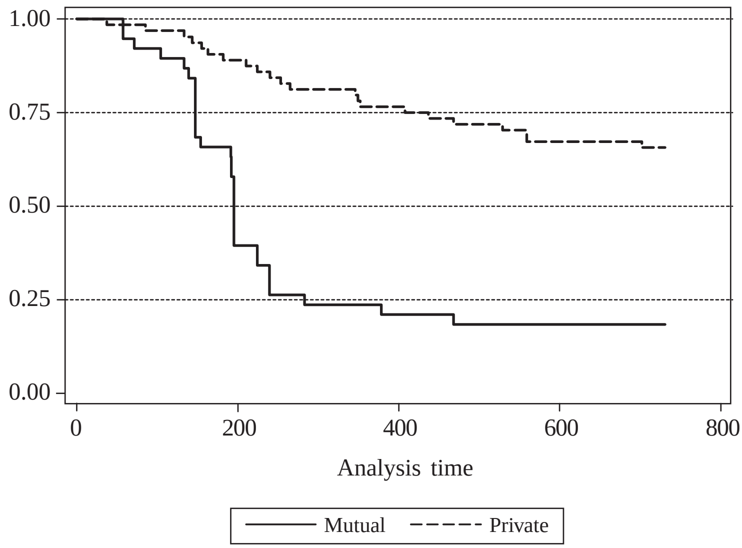


FIGURE 3

HAZARD FUNCTIONS

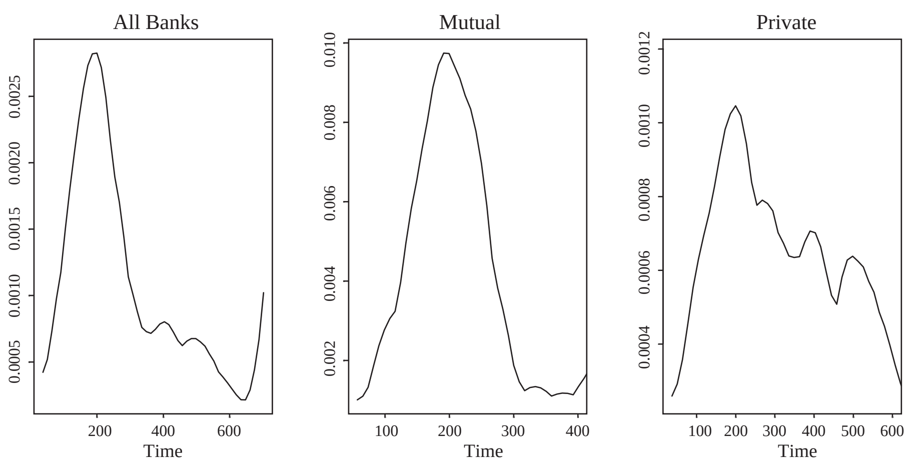


TABLE 3

BANK FAILURE DYNAMICS

Time (t) (in days)	Banks in t	Failures	Survival Function	Standard Error	95% Confidence Interval	
Mutual banks (type = 0)						
57	38	2	0.947	0.036	0.806	0.987
71	36	1	0.921	0.044	0.775	0.974
104	35	1	0.895	0.050	0.743	0.959
133	34	1	0.868	0.055	0.712	0.943
139	33	1	0.842	0.059	0.682	0.926
147	32	6	0.684	0.075	0.512	0.807
154	26	1	0.658	0.077	0.485	0.785
191	25	1	0.632	0.078	0.459	0.763
192	24	2	0.579	0.080	0.408	0.717
195	22	7	0.395	0.079	0.242	0.544
224	15	2	0.342	0.077	0.198	0.491
239	13	3	0.263	0.071	0.137	0.408
283	10	1	0.237	0.069	0.118	0.379
378	9	1	0.211	0.066	0.099	0.350
468	8	1	0.184	0.063	0.081	0.320
731	7	0	0.184	0.063	0.081	0.320
Private national banks (type = 1)						
37	64	1	0.984	0.016	0.894	0.998
85	63	1	0.969	0.022	0.881	0.992
133	62	1	0.953	0.026	0.862	0.985
143	61	1	0.938	0.030	0.842	0.976
155	60	1	0.922	0.034	0.822	0.967
163	59	1	0.906	0.036	0.803	0.957
182	58	1	0.891	0.039	0.784	0.946
210	57	1	0.875	0.041	0.766	0.935
224	56	1	0.859	0.044	0.747	0.924
240	55	1	0.844	0.045	0.729	0.913
253	54	1	0.828	0.047	0.711	0.901
265	53	1	0.813	0.049	0.694	0.889
346	52	1	0.797	0.050	0.676	0.877
349	51	1	0.781	0.052	0.659	0.864
352	50	1	0.766	0.053	0.642	0.852
408	49	1	0.750	0.054	0.625	0.839
437	48	1	0.734	0.055	0.608	0.826
468	47	1	0.719	0.056	0.591	0.813
529	46	1	0.703	0.057	0.575	0.799
559	45	2	0.672	0.059	0.542	0.772
702	43	1	0.656	0.059	0.526	0.758
731	42	0	0.656	0.059	0.526	0.758

TABLE 4

TESTS OF EQUALITY OF SURVIVAL FUNCTIONS

	Chi2(1)	Pr>Chi2
Long-Rank	32.43	0.0000
Wilcoxon	32.65	0.0000

Ho: Both survival functions are equal.
Ha: Different bank groups have different functions.

We present the estimation of the survival and hazard functions for all banks in Figures 1 and 3 to show that the estimation for all banks only can lead to serious errors, given the heterogeneity of the different types of banks.

4.3 A conditional model for banks duration

This section studies the possibility of using the banks' specific financial indicators as explanatory factors of the survival dynamics. Based on the results of the previous section, the effect of the bank specific indicators was allowed to change by bank group. Fourteen explanatory variables were included, which correspond to the seven previously mentioned indicators, and slope dummy variables per bank group. The *indk* (k=1,......,7) variables correspond to the X_k indicator while *ink* takes values equal to *indk* for private banks and zero for the rest of them, so that coefficients of this last group of variables are interpreted as differential effects by bank type. Under these considerations, Cox's model is specified as:

$$h(t \mid X) = h_0(t)\ exp\left\{ \sum_{k=1}^{7} (\beta_k X_k + \gamma_k DX_k) \right\}$$

where D is an indicator with value 1 if the bank is private national and 0 if it is mutual, and according to the definition, X_k = *indk*, and DX_k = *ink*. The effect of the *k-th* on the hazard rate can be evaluated as:

$$\frac{\partial \ln h(t \mid X)}{\partial X_k} = \beta_k + \gamma_k D$$

so that the semi-elasticity of the hazard rate with respect to the *k-th* indicator is β_k for mutual banks, and $\beta_k+\gamma_k$ for private national banks. It is important to remark that the inclusion of slope dummy variables for bank types gives estimates that are numerically equivalent to estimating different models for different bank types, with the computational advantage of being able to test for significant differences with simple "t" test statistics.

Estimation results for this model are shown in Table 5, where the corresponding columns show the estimated coefficients, their standard deviations, the *z*-statistic corresponding to the null hypothesis that the coefficient is equal to 0, the *p*-values, and a 95% confidence interval.

In general terms, the global fit is acceptable, and many variables are significantly different from zero. The χ^2 statistic corresponding to the null hypothesis that none of the explanatory variables is significantly different from zero is notoriously larger than the critical values commonly accepted. This result suggests that it is not correct to say that the bank failure process was caused by random, non-systematic factors, and that, on the contrary, it is significantly influenced and explained by financial bank's specific observable factors. This result agrees with those obtained by Schumacher (2000) in a bank failure multinomial choice model, and also confirms those suggested by Dabós (1998) who, maybe because of a strong collinearity between the explanatory variables, found that only one indicator is relevant to explain the bank failure process.

Dabós (1998) uses as explained variable a binary indicator equal to 1 if the bank failed during certain period, and 0 if it did not. Therefore, as Doksum and Gasko (1990) remark, the binary model is a censored version of the duration model, where the explained variable is the time elapsed until default. This suggests that the higher variability considered by the duration model allows a better identification of the factors that explain the development of the bank crisis.

The null hypothesis that all the *ink* variable (k=1,......,7) coefficients are all equal to zero is also rejected by a standard Wald test (not shown). This together

TABLE 5

COX'S MODEL ESTIMATION

Variable	Coefficient	Std. Err.	z	P > |z|	95% Confidence interval	
ind1	0.104	0.067	1.553	0.120	– 0.027	0.235
ind2	0.236	0.137	1.717	0.086	– 0.033	0.505
ind3	– 0.061	0.027	– 2.266	0.023	– 0.114	– 0.008
ind4	– 0.030	0.038	– 0.798	0.425	– 0.104	0.044
ind5	– 0.030	0.041	– 0.740	0.460	– 0.110	0.050
ind6	0.012	0.007	1.695	0.090	– 0.002	0.026
ind7	0.017	0.013	1.282	0.200	– 0.009	0.042
in1	– 0.092	0.071	– 1.298	0.194	– 0.230	0.047
in2	– 0.127	0.134	– 0.947	0.343	– 0.390	0.136
in3	0.004	0.038	0.099	0.921	– 0.071	0.079
in4	– 0.032	0.050	– 0.631	0.528	– 0.130	0.067
in5	0.197	0.073	2.704	0.007	0.054	0.339
in6	– 0.017	0.008	– 2.147	0.032	– 0.033	– 0.002
in7	– 0.042	0.022	– 1.905	0.057	– 0.086	0.001
Total banks = 102				Log likelihood		– 199.3553
Total defaults = 53				Chi2(14)		59.3100
				Prob > chi2		0.0000

with the rejection that the remaining variables are also jointly significant (also based on a Wald test-statistic) must be interpreted as indicating that it is important to distinguish between bank types in the proportional part of the model. It is important to stress the importance of considering joint tests as opposed to pooling individual significance tests since maybe due to high multicollinearity only few variables appear as significant.

Looking at individual effects, the capitalization measure given by indicator 1 (Equity/Assets) is not significantly different from zero for both groups, that is, its effect is null for cooperative and private banks. The capitalization effect, measured by the leverage level as Indicator 2 (Liabilities/Equity), is significantly different from zero, and with the expected sign. A marginal increase in its level increases the default risk in approximately 23% for both cooperative and private national banks.

Indicator 3, Immediate liquidity (Cash + Public securities/Deposits), also has the correct sign. An instant increase in marginal liquidity reduces the default risk in approximately 6% for both groups. Structural liquidity measured by Indicator 4 (Equity – Fixed assets/Liabilities) has no effect at all in both groups.

Indicator 5, (Operating expenses/Liabilities) presents a positive effect on the default risk in private national banks. A marginal increase in it increases the default risk in 20% in private national banks, and has no effect in mutual banks. The expected effect for the portfolio situation measured by Indicator 6 (Arrears portfolio – Losses provisions/Equity) has the expected sign in mutual banks but not for private national banks. If Indicator 6 increases marginally, the default risk in mutual banks increases 1%, and decreases 0,5% in private national banks. Maybe because of high colinearity problems Indicator 6 does not appear as an important variable to explain the banks default risk. Finally, the greater the return on equity (ROE), the lower the default risk (-4%), but only for private national banks. The return on equity effect is not important at all in mutual banks.

Testing the proportional hazards assumption

The proportional hazards assumption implies that the effect of explanatory variables on the hazard function is constant over time, that is, a marginal change in any of the explanatory variables induces a vertical shift along time. To test if the proportional hazards assumption is valid, Harrell's test was used, based on Schoenfeld's residuals, as discussed in the previous section. Intuitively, a model is estimated allowing Cox's model β coefficients to change over time. Under the proportional hazards assumption these coefficients should be constant, and therefore, under the null hypothesis, the graph showing those coefficients estimations along time should approximate a horizontal line. The Harrell test can be interpreted as a test in which the coefficients are constant. This is confirmed by the tests results, which, for a 95% confidence interval, suggest not to reject the constant coefficients null hypothesis in any of the cases. In general terms, the obtained results suggest that the proportional hazards assumption is a valid restriction for the studied case. Results are presented in the following Table.

TABLE 6

PROPORTIONAL HAZARDS HYPOTHESIS TEST

	Rho	Chisq	p
ind1	– 0.120	0.909	0.340
ind2	0.001	0.000	0.991
ind3	0.018	0.012	0.911
ind4	0.092	0.268	0.605
ind5	0.193	1.987	0.159
ind6	– 0.080	0.426	0.514
ind7	– 0.096	0.494	0.482
in1	0.066	0.239	0.625
in2	– 0.050	0.173	0.678
in3	0.059	0.170	0.680
in4	0.004	0.001	0.974
in5	– 0.105	0.755	0.385
in6	0.147	1.275	0.259
in7	0.092	0.402	0.526
Global		10.340	0.737

4.4 Predictive ability of the model

As an additional exercise, it is interesting to check the predictive ability of the estimated model. Predictions are computed in a similar way to those presented by Whalen (1991). For each period t, it is possible to obtain the estimated survival probability $S(t|X)$, using the relations between the hazard rate and the survivorship function described in Section III, applied to the Cox's proportional hazard's model:

$$\hat{S}(t \mid X) = \hat{S}_0(t)^{exp(\hat{\beta}'X)}$$

where:

$$\hat{S}_0(t) = exp\left\{-\int_0^t \hat{h}_0(u)du\right\}$$

where $h_0(\ )$ is the estimated baseline hazard function, and $\hat{\beta}$ is the coefficients vector estimated in Cox's model. Using these formulae, for any period t^*, it is possible to find the probability that a bank had survived until that period given its financial characteristics (the X vector). Then banks are classified as "survivors" if the estimated probability of doing so is higher than a pre-established cut-off value S^*. This exercise is done bank by bank, and then predicted survival status is compared to the observed one. According to this analysis there may be two possible classification errors: predicting that a bank would survive until t^* when it did not, or predicting that the bank would not survive until t^*, when it did.

Table 7 presents the resulting classifications from the previously described analysis. Predictions where made considering three periods of reference: 365, 547 and 731 days (one year, one year and a half, two years). The cut-off value to consider that the bank survived is taken as the proportion of banks that had survived until the reference period, as it is usually done.

The columns correspond to predicted values and the rows to observed values. For example, for the 365-day predictions, 34 banks failed before 365 days, and the model correctly predicted that. 44 banks that survived more than 365 days are correctly classified by the model. On the other side, 10 banks are classified in the survival group, while they failed before 365 days, and 14 banks were incorrectly classified in the default group while they survived during the reference period.

In general terms, the model's prediction ability is more than acceptable, with a correct classification rate of around 80%. The predictive performance does not seem to change with the prediction horizon, though a small improvement is observed. It is interesting to remark that Dabós (1988) produces correct classification rates of around 80% for cooperative banks, hence, the improvement offerd by the use of a suvival model in being able to identify explanatory factors does not come at a substantial price in terms of predictive power.

A second prediction exercise was to calculate survival functions for the different bank groups. That is, the predicted survival probabilities where calculated for three bank groups classified according to their observed performance. Follow-

TABLE 7

COX'S MODEL PREDICTIONS

	Predicted values		Total
	0	1	
365-days predictions			
0	34	10	44
1	14	44	58
Total	48	54	102
547-days predictions			
0	38	12	50
1	11	41	52
Total	49	53	102
731-days predictions			
0	41	12	53
1	9	40	49
Total	50	52	102

ing Wallen's (1991) classification, banks can be separated in: 1) seriously ill, 2) ill, and 3) sane. The first ones correspond to the group that failed before 365 days, the second ones correspond to those that failed between 365 and 547 days, and the last ones to those that didn't failed. For each bank group the survival function was calculated, using the financial indicators' averages. These average values are shown in Table 8.

Results are shown graphically in Figure 4.

These results show that bank dynamics are quantitatively different according to the banks' financial performance. Survival probabilities decrease slowly until approximately 150 days after the crisis began. After 200 days, they show a strong risk acceleration for banks with worse conditions, and then they continue to fall slowly. The probabilities, for all the banks, seem to be stabilized at the 400 days from the beginning of the crises.

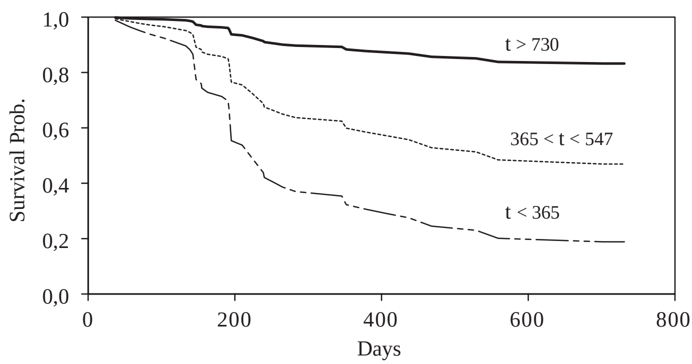
TABLE 8

AVERAGES OF EXPLANATORY VARIABLES FOR THE DIFFERENT BANK GROUPS

Failure	ind1	ind2	ind3	ind4	ind5	ind6	ind7
t < 365	15,033	6,773	21,073	7,944	14.748	85,231	0,017
365 < t < 547	22,616	6,271	37,301	20,237	21,243	46,090	3,361
t > 730	18,603	6,483	34,415	18,791	10,417	43,119	6,753

FIGURE 4

SURVIVAL FUNCTIONS PREDICTED FOR DIFFERENT BANK GROUPS



V. Conclusions

This paper presents improvements in many directions on the study of the dynamics of bank crisis in Argentina after the Mexican crisis. First, it provides a clear description of the way in which the bank failure process occur as a consequence of the events that happened after the Mexican crisis. Second, the use of the time to default as the explained variable allowed us to clearly identify important factors in the bank failure process. This is an advantage over standard binary choice models, which, in spite of their good predictive ability for the Argentine case (as in Dabós, 1998), have problems to isolate the individual effect of the explanatory variables.

The model used seems to provide a good balance between flexibility and parsimony. The semiparametric character of Cox's model allowed us to estimate the relevant parameters without relying on restrictive or unreal structures about the failure dynamics, as it happens with parametric models which are based on specific functional structures. This paper provides original evidence about the functional form of the hazard rate, which presents a clear non-monotonic behavior, which discards standard choices like the weibull or exponential specifications. Although based on a relatively small number of observations, the methodology used in this paper could provide useful information about the non-parametric estimation of that function, which is an interesting point for future research.

The dynamics of bank failures can be fairly characterized by observable factors, which strongly discards the possibility that it has been governed by imitation processes solely, where the bank financial situation does not play any role. It is interesting to note that the non-monotonicity of the hazard rate function suggests that there were imitation effects, and that they had a strong influence in the first 200 days of the crisis.

References

COLE, R. A. and J. W. GUNTHER (1995). "Separating the Likelihood and Timing of Bank Failure," *Journal of Banking and Finance,* 19 (6), September, pp. 1073-1089.

COX, D. R. (1972). "Regression Models and Life-Tables", *Journal of the Royal Statistical Society,* May/Aug. B 34, pp. 187-220.

DAHL, D. and M. F. SPIVEY (1995). "Prompt Corrective Action and Bank Efforts to Recover from Undercapitalization," *Journal of Banking and Finance,* 19 (2), May, pp. 225-243.

DABÓS, M. (1998). "Crisis Bancaria y Medición del Riesgo de Default: Métodos y el Caso de los Bancos Cooperativos en Argentina", *Desarrollo Económico,* Vol. 38, Special Number, Autumn, pp. 215-230.

DABÓS, M. and L. GOMEZ MERA (1999). "The Tequila Banking Crisis in Argentina," Chapter 9 in *International Banking Crises,* Benton E. Gup (ed.), Quorum Books, Westport, Connecticut. London, pp. 134-161.

DEMIRGUC-KUNT, A. (1989). "Deposit-Institution Failures: A Review of Empirical Literature," *Federal Reserve Bank of Cleveland Economic Review,* Fourth Quarter, pp. 2-18.

DOKSUM, K. and M. GASKO (1990). "On a Correspondence Between Models in Binary Regression Analysis and in Survival Analysis", *International Statistical Review,* 58 (3), pp. 243-252.

FLEMING, T. and D. HARRINGTON (1991). *Counting Processes and Survival Analysis,* Wiley, New York.

GLICK, R. and M. HUTCHISON (1999). "Banking and Currency Crises: How Common are Twins?, *Pacific Basin Working Paper Series 99-07,* Federal Reserve Bank of San Francisco.

GONZÁLEZ-HERMOSILLO, B.; C. PAZARBASIOGLU and R. BILLINGS (1996). "Banking System Fragility: Likelihood Versus Timing of Failure: An Application to the Mexican Financial Crisis," *IMF* Working Paper.

HARRELL, F. (1986). "The PHGLM Procedure". *SAS Supplemental Library User's Guide, Version 5.* SAS Institute, Inc., Cary, NC.

KAMINSKY, G. L. and C. M. REINHART (1999). "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems", *The American Economic Review,* 89 (3), June, pp. 473-500.

KALBFLEISCH, J. D. and R. PRENTICE (1980). *The Statistical Analysis of Failure Time Data,* Wiley.

KIEFER, N. M. (1988). "Economic Duration Data and Hazard Functions," *Journal of Economic Literature,* 26 (2), June, pp. 646-679.

KLEIN, J. P. and M. MOESCHBERGER (1997). *Survival Analysis: Techniques for Censored and Truncated Data,* Springer-Verlag New York Inc.

LANCASTER, T. (1990). *The Econometric Analysis of Transition Data,* Econometric Society Monographs No. 17, Cambridge University Press.

LANE, W. R.; S. LOONEY and J. WANSLEY (1986). "An Application of the Cox Proportional Hazards Model to Bank Failure," *Journal of Banking and Finance,* 10 (4), December, pp. 511-31.

MCCANDLESS, G.; M. GABRIELLI and M. ROUILLET (2003). "Determining the Causes of Bank Runs in Argentina during the Crisis of 2001", *Revista de Análisis Económico,* 18 (1), Junio, pp. 87-102

MILLER, R. (1981). *Survival Analysis,* Wiley, New York.

NEUMANN, G. (1997). "Search Models and Duration Data", in H. Pesaran y P. Schmidt, *Handbook of Applied Econometrics, Vol II: Microeconomics,* Blackwell Publishers, Oxford, pp. 300-351.

PARMAR, M. and D. MACHIN (1996). *Survival Analysis: A Practical Approach,* Wiley, New York.

RAMLAU-HANSEN, H. (1983). "Smoothing Counting Process Intensities by Means of Kernel Functions", *Annals of Statistics,* 11, pp. 453-466.

SCHOENFELD, D. (1982). "Partial Residuals for the Proportional Hazards Regression Model", *Biometrika,* 69 (1), pp. 239-241.

SCHUMACHER, L. (2000). "Bank Runs and Currency Run in a System without a Safety Net: Argentina and the 'Tequila' Shock", *Journal of Monetary Economics,* 46 (1), August, pp. 257-277.

VAN DEN BERG, G. (2001). "Duration Models: Specification, Identification and Multiple Durations", Chapter 55 in J. Heckman and E. Leamer (eds.), *Handbook of Econometrics,* Vol. 5, Elsevier, Amsterdam.

WEELOCK, D. and P. WILSON (1995). "Why do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions", *Federal Reserve Bank of Saint Louis,* Working Paper N° 950133, August.

WHALEN, G. (1991). "A Proportional Hazards Model of Bank Failure: An Examination of Its Usefulness as an Early Warning Tool," *Federal Reserve Bank of Cleveland Economic Review,* First Quarter, pp. 21-31.