DENNIS GLENNON
PETER NIGRO

Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach

In this paper we report that, although medium-maturity loans originated under the SBA 7(a) loan guarantee program are targeted to small firms that fail to obtain credit through conventional channels, the default experience is comparable to that of a large percentage of loans held by larger commercial banks. We establish that the default behavior of these loans is time sensitive—as a loan seasons, the likelihood of default increases initially, peaks in the second year, and declines thereafter. Using a discrete-time hazard framework, we show that the likelihood of default is conditional on borrower, lender, and loan characteristics, and changes in economic conditions.

JEL code: G21
Keywords: small business loans, default risk, survival analysis, hazard models.

CREDIT-RISK MODELS ARE used in a wide variety of lender applications, ranging from underwriting and pricing to account management and capital allocation. A number of techniques have been used to estimate one of the key components of these models—the likelihood of default *over time*. In this paper we demonstrate the advantages of explicitly capturing the effect of time on the probability of default for a sample of medium-maturity (i.e., seven years) Small Business Administration (SBA) guaranteed loans using a survival analysis/hazard model approach. We use a discrete-time hazard procedure—a computationally straightforward estimation procedure statistically similar to a Cox Proportional Hazard model—that is designed to make full use of the effect of changes in economic

The opinions expressed in this paper are those of the authors and do not necessarily reflect those of the Office of the Comptroller of the Currency or the SBA. We would like to thank Bill Lang, Bob DeYoung, and Fred Rubin for helpful comments on earlier drafts of this paper. Early work on this paper was completed while Glennon was a Visiting Fellow at the University of Technology, Sydney and Nigro was at the OCC.

DENNIS GLENNON *is a Financial Economist at the Risk Analysis Division, Office of the Comptroller of the Currency (E-mail:* dennis.glennon@occ.treas.gov). PETER NIGRO *is an Associate Professor, Sarkisian Chair at Department of Finance, Bryant University (E-mail:* pnigro@bryant.edu).

Received April 4, 2001; and accepted in revised form September 2, 2003.

Journal of Money, Credit, and Banking, Vol. 37, No. 5 (October 2005)

conditions over time. Our results show that the likelihood and timing of a medium-maturity, SBA loan defaulting is conditional on several key borrower, lender, and loan characteristics, as well as changes in the economic environment over time. Moreover, we find that as the medium-maturity loans season the likelihood of default increases initially, peaks in the second year after origination, and declines thereafter.

Although our analysis is limited to data on loans originated under the SBA 7(a) loan guarantee program, our results are likely to be of general interest for several reasons. First, although there have been several studies that focus on small business credit accessibility (Scott and Dunkelberg, 2003, Berger, Frame, and Miller, 2002, Cole, Wolken, and Woodburn, 1996, Peek and Rosengren, 1996, Berger and Udell, 1995), data availability problems have limited research on the performance of small-business firms receiving loans. As a result, little is known about the performance of loans that represent a significant portion of the total commercial and industrial loans held by U.S. commercial banks—accounting for roughly 11% of the total. SBA loans make up a large percentage of the total volume of small-business loans originated by commercial banks. The SBA data then allow us to examine the default behavior of a large segment of the small-business loan market. We show that medium-maturity SBA 7(a) loans are concentrated in the relatively more risky segment of the loan market but they are comparable in quality to a large percentage of loans held by larger commercial banks

Second, using survival analysis techniques, we show that not all business credits are of equal default risk and that a bank's exposure to loss due to default is not constant, but varies significantly over the life of the loan. This has important implications for the recent proposed changes to the Basel Capital Accord that would impose a uniform capital charge against small business credits.

Finally, we outline a method of predicting defaults that fits well within a net cash flow modeling framework. Within that framework, the time path of default is as important as the event itself for measuring a lender's exposure to losses. The discrete-time hazard modeling approach allows us to explicitly account for time and capture the impact of changing conditions on the likelihood of default using time-varying covariates. Although we develop our approach within the framework of the 1990 Federal Credit Reform Act (FCRA), the method we use to estimate the default model is germane to the development of any model used for risk-based loan pricing, allocation of reserves and capital, and the valuation of pools of securitized loans.[1]

The remainder of the paper is organized as follows. In Section 1, we present a brief overview of the SBA 7(a) loan guarantee program including a comparison of the cumulative default rate experience of SBA loans with rated corporate bonds

1. Under Section 504 of the Federal Credit Reform Act (FCRA) of 1990, SBA direct loans and loan guarantee commitments are allowed only to the extent that appropriations have been made in advance for their cost. The cost is defined as: "...the estimated long-term cost of the Government of a direct loan or loan guarantee, calculated on a net present value basis, excluding administrative costs and any incidental effect on government receipts or outlay" (see Section 502(5)). This requires the SBA to link the total dollars disbursed each year directly to the net present value of the expected long-term cost of the program.

over the same time horizon. We outline several unique features of the guarantee program that characterize the SBA's exposure to losses and motivate our specification of a conditional default model. We also summarize the default history by loan attributes and cohort segments, and use a simple non-parametric hazard technique to identify the general pattern in the timing of default unconditional on other factors. In Section 2, we outline the discrete-time hazard framework and estimate a hazard model controlling for borrower, lender, and loan characteristics, as well as the economic environment over the life cycle of the loan.[2] In the final section, we state our conclusions and discuss several areas of future research.

1. AN OVERVIEW OF THE SBA 7(A) LOAN GUARANTEE PROGRAM

The SBA lending programs constitute a substantial portion of the small-business loan market. For example, in 1999 U.S. commercial banks held roughly \$105 billion in small business loans (i.e., commercial and industrial (C&I) loans less than \$250,000) representing 11% of their total holdings of C&I loans.[3] In contrast, the SBA reported a combined managed guaranteed-loan portfolio of just over \$40 billion in 1999.[4]

Under the SBA 7(a) program, loan guarantees are made available to eligible small firms that meet specific SBA guidelines. Foremost among the requirements are: (1) the firm could not qualify for credit on reasonable terms without the SBA guarantee; (2) the prospects for repayment are sound; and (3) the firm must be considered "small." The first condition requires that the firm document that without the SBA guarantee it would not qualify for a bank loan of equivalent terms, pricing, and maturity.[5] In addition, the SBA requires that all 7(a) loans be secured by tangible assets of the businesses and, if necessary, the owners themselves to ensure that the prospects of recovery are financially sound. Finally, the firm must qualify as "small"—a relative measure based on employee count and annual revenue and defined specifically with respect to a firm's primary industry classification code (i.e., SIC code).

2. Although not addressed in this paper, the concept of competing risk is an issue that emerges directly from the development of the net present value model in which there are multiple exit categories (e.g., default, paid-as-agreed, and prepayment). This issue could be addressed in future research if additional data are available. See McDonald and Van de Gucht (1999) for more on the competing risk framework.

3. Banks report their volume of C&I lending *by loan size* once a year in the June Call Report.

4. This measure includes all SBA lending (e.g., 7(a), SBIC, and MICRO loan programs). The 7(a) program, however, is the largest representing over 80% of the SBA approved loans (measured in dollars) in 2001. We note, however, that our data do not represent the population of small business loans in general; only those that fail to obtain credit through conventional lending markets. As such, our results generalize only to the behavior of SBA small business loans, not necessarily the full market of small business loans.

5. In general, the longer maturity and lower price of SBA 7(a) loans distinguishes them from the unguaranteed small-business loans of commercial banks. The SBA's price cap of 350 basis points over prime makes these loans attractive to opaque borrowers who are unlikely to have access to this type of long-term financing at similar interest rate.

SBA-guaranteed loans may be used for a wide variety of purposes including working capital, expansion, equipment, and construction. In return for the guarantee, the SBA assesses a lender fee (which may be transferable to the borrower) that is a percentage of the guaranteed portion of the disbursed loan amount. The lender is responsible for loan servicing according to the terms and conditions of the loan agreement and for all administrative costs.[6] Regular payments are made until the loan is paid in full or the borrower defaults.

Following disbursement, lenders report quarterly loan-performance information to the SBA. Under the SBA's program guidelines, a lender holding an SBA loan that is in arrears for 60 days or more can, by contractual agreement, put the loan back to the SBA. The SBA is required to "purchase" the nonperforming loan at a value equal to the guaranteed portion of the remaining outstanding balance and delinquent interest—resulting in a pro rata loss sharing between the lender and the SBA determined by the contractual guarantee rate. Purchase, however, occurs only after the borrower defaults on its payments *and* the lender exercises the guarantee option. Loans that are in arrears, however, are not always put back immediately to the SBA.[7] In some cases, loans are placed into "liquidation" by the lender without first transferring ownership to the SBA. In most of these cases, the loans are eventually purchased by the SBA, usually after the servicer has exhausted their attempts to collect on the loan collateral over an extended loan recovery period. Because we are concerned with the timing of the default, we focus on the "in liquidation" measure of default and the corresponding date on which the loan is declared "in liquidation."[8]

1.1 A Preliminary Analysis of Credit Quality

SBA loan guarantees are only available to eligible small firms failing to qualify under the lender's conventional small-business underwriting guidelines. For that reason, they are likely to be, on average, of lower credit quality than the typical small-business loan. Although performance data for the population of small-business loans are not available, we use the cumulative default experience on corporate bonds evaluated by Moody's (2002) and Standard and Poor's (S&P) as reasonable—albeit crude—benchmark values to assess the relative credit quality of SBA 7(a) loans. Given the simplicity and familiarity of market participants with Moody's and S&P ratings (as well as their long history), several other researchers have used their ratings' transition information for similar purposes (e.g., Carey and Hrycay 2001).

6. The lender may securitize the guaranteed portion of the loan in the secondary market. In recent years, however, the SBA has assessed a 50 basis point fee against the value on these transactions. Interestingly, the spread on the securitized portion of the loans often generates a significant fee income to the originating bank that in some cases affects the pricing of the SBA loans to the borrower.

7. Technically, a single missed payment should trigger a default. In many of those cases, the payment is only delayed and the loan is soon restored to current status. Internally, banks often use past due measures (e.g., 30-, 60-, or 90-days past due) to track credit migration and default. These performance-related data, however, were not available in the SBA database.

8. Because not all lenders attempt to collect on loans in arrears greater than 60 days, not all loans pass through the "in liquidation" stage. Instead, they go directly to "purchased" status. In those cases, we identify the timing of the event as the date in which they were purchased. Therefore, the event date for default is either the purchased date or, if it exists, the liquidation date. Defining default this way affects the timing, not the recognition, of default.

In Panel A of Table 1, we report the cumulative default probabilities based on a 20% sample of seven-year maturity SBA 7(a) loans disbursed from 1983 to 1998 using a method consistent with that used by Moody's and S&P. We estimate the historical cumulative default rates using the approach developed by Moody's as outlined in Fons, Carty, and Girault (1993). Under that approach, the cumulative default rate is calculated by cohort group in which each group consists of all loans outstanding at the beginning of each cohort year. The default behavior of each cohort group is tracked over performance horizons spanning one to seven years and the weighted average default rate over all cohort groups is calculated for each performance horizon. We estimate, using this approach, an average life-of-loan cumulative default rate of almost 17% for our representative sample of medium-term SBA loans.

Table 1 shows that the pooled SBA cumulative default rates over the seven-year time horizon fall just below those of the average cumulative default rates on Moody's Ba-rated corporate bonds (Panel B), and between a BB and B rating based on S&P's cumulative default probabilities (Panel C).[9] As these ratings are near the upper end

TABLE 1

CUMULATIVE DEFAULT PROBABILITIES[a]

Panel A. Seven-Year Maturity, SBA-Guaranteed Loans[b]

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*	*Year 6*	*Year 7*
SBA loans (%)	4.18	8.27	10.50	12.29	13.84	15.28	16.68

Panel B. Moody's Cumulative Default Probabilities, 1920–2001[c]

Rating	*Year 1*	*Year 2*	*Year 3*	*Year 4*	*Year 5*	*Year 6*	*Year 7*
Baa	0.15	0.46	0.87	1.44	1.95	2.54	3.16
Ba	1.27	3.57	6.20	8.83	11.42	13.75	15.63
B	6.66	13.99	20.51	26.01	31.00	35.15	39.55
Inv. grade (%)	0.06	0.19	0.38	0.65	0.90	1.19	1.50
Spec. grade (%)	4.73	9.55	13.88	17.62	20.98	23.84	26.25

Panel C. S&P Cumulative Default Probabilities, 1981–2000[d]

Rating	*Year 1*	*Year 2*	*Year 3*	*Year 4*	*Year 5*	*Year 6*	*Year 7*
BBB	0.22	0.50	0.79	1.30	1.80	2.29	2.73
BB	0.98	2.97	5.35	7.44	9.22	11.11	12.27
B	5.30	11.28	15.88	19.10	21.44	23.20	24.77
CCC	21.94	29.25	34.37	38.24	42.13	43.62	44.40
Inv. grade (%)	0.08	0.19	0.31	0.51	0.72	0.95	1.17
Spec. grade (%)	4.14	8.34	11.93	14.67	16.84	18.64	19.98

[a]Ratings for SBA loans are not available. The SBA does not rate loans internally nor are there data on borrower/firm to develop external ratings. Although credit scores are sometimes used to rate loans, small-business scorecards were not generally available until 1997 precluding their use in this study.
[b]The cumulative default rates for the SBA loans were calculated using the approach developed by Moody's as outlined in Fons, Carty, and Girault (1993). The cumulative rates are weighted averages of the cohort-specific cumulative default rates.
[c]Source: Moody's (2002). The cumulative default rates for the top-tiered rating categories are not reported.
[d]Source: S & P (2001). The cumulative default rates for the top-tiered rating categories are not reported.

9. Data limitations prevent us from segmenting the SBA portfolio into risk categories (e.g., BBB, BB, etc.) similar to those developed by Moody's and S&P. More recently, credit scoring models have been used by lenders to rank applicants by expected performance for the purpose of defining risk-related categories. Unfortunately, wide spread use of small-business scoring models did not take place until after 1998—the year in which our sample ends. For that reason, we base our comparison of average performance on the pooled SBA loan portfolio.

of the speculative grade ratings (i.e., Ba or lower), we can conclude that SBA loans are concentrated in the relatively more risky segment of the loan market. Treacy and Carey (1998), however, find that banks associated with 26 of the 50 largest bank holding companies—representing more than 75% of aggregate banking industry assets—reported that roughly 45% of their rated assets fall into rating categories below Ba (i.e., speculative grades) at year-end of 1997. Their results suggest that although medium-maturity SBA-guaranteed loans originated under the 7(a) program are, on average, below investment grade, they are comparable in average quality to a large percentage of loans held by commercial banks.

The growth pattern in the SBA cumulative default rate in Table 1 suggests that the marginal default rates are time dependent. These results, however, are based on the average performance unconditional on program, loan, and macroeconomic-specific factors; factors that are likely to affect the expected likelihood and timing of default.

1.2 Program-Specific Factors

The SBA 7(a) loan guarantee program requires the SBA to share losses in event of default with the lender on a pro rata basis. For that reason, the SBA must estimate their loss exposure much the same way a lender measures its exposure for reserve allocation or capital purposes. There are, however, specific aspects of the SBA 7(a) program that differentiates it from more conventional lending programs.

Under the 7(a) program, the SBA offers loan guarantees to eligible businesses through qualified financial institutions. Both banks (i.e., commercial banks, thrifts, and savings banks) and nonbanks (i.e., finance companies) participate as lenders and play a critical role in the 7(a) program. The lenders are primarily responsible for choosing firms to receive loans, initiating SBA involvement, and monitoring the loans on an ongoing basis. The originating lender underwrites the loans within program guidelines, and the SBA undertakes various levels of review depending on the SBA-designated lender type: regular, certified, and preferred lenders.

Regular lenders submit complete SBA loan applications to the local SBA office after completing their own internal underwriting review. The SBA, however, re-underwrites an application submitted by a regular lender and makes a final decision to approve or deny the loan. The SBA's review duplicates or extends the underwriting analysis performed by the submitting lender. A large majority (i.e., roughly 70%) of the loans in our sample were underwritten in this manner.

The Certified Lender Program (CLP) accelerates the loan review process for lenders familiar with SBA policies and procedures. A CLP lender is expected to conduct a full review of the loan applicant that is later submitted to an SBA office for cursory review. The SBA makes the final approval decision, but its decision is based on a review of the lender's documentation, as opposed to a full, independent re-underwriting of the loan. CLP lenders originated almost 19% of the loans in our sample.

The Preferred Lender Program (PLP) further streamlines the underwriting process as the SBA delegates eligibility determination, loan approval, and most servicing and liquidation authority to these lenders. PLP lenders are chosen from among the best lenders in the Certified Lender Program and are re-certified every two years. They make up only 11% of the loans in our sample. In 1994, the SBA further streamlined its underwriting process by introducing the "low-doc" program for its regular lenders. This program reduces the time needed to underwrite smaller loans (i.e., less than $100,000) by requiring only minimal information from borrowers.[10]

1.3 Borrower, Lender, and Loan-Specific Factors

The SBA underwrites loan guarantees of varying terms to small businesses dispersed over a wide range of industry classifications, geographic locations, firm size, and corporate structures. Our data set includes information on: (1) loan-specific characteristics such as the guarantee percentage, loan amount, initial interest rate, interest-rate type, and low-documentation indicator; (2) lender characteristics such as SBA lender type, region (geographic location), and bank/nonbank status; and (3) borrower characteristics such as corporate structure, SIC division (industry classification), number of employees, and new/existing-firm status.[11] To the extent that a firm's solvency risk varies by these factors (e.g., firms in the retail sector tend to default more frequently, all else equal), we can use them to develop better estimates of the SBA's exposure to loss over time.

In Table 2 we report the descriptive statistics and default experience for a selected set of borrower or firm-specific characteristics (Panel A) and lender and loan characteristics (Panel B) for the pooled sample of SBA loans. In accordance with its mission, the SBA attempts to provide new firms with adequate startup capital to grow and expand. This is reflected in the relatively large percentage (i.e., nearly 35%) of the loans in our sample made to new firms.[12] Several studies (e.g., Good and Graves, 1993, Honjo, 2000) find that new firms fail at higher rates than established firms (all else equal). In Table 2, however, we show that for the pooled data, the cumulative default rate for SBA loans to new firms (i.e., 15.0%) is statistically the same as that for established firms (i.e., 14.8%). The 7(a) program also attempts to reach small firms irrespective of corporate structure. In Table 2, we show that almost 58% of the loan guarantees were made to small firms structured as corporations, 35% to individual proprietors, and the remaining 7% to firms structured as partnerships. The default rates, on a pooled basis, however, appear to be unaffected by the corporate structure of the firm.

Although firms representing almost all industry classifications obtain loans through the SBA 7(a) loan program, there is a high concentration of loans to firms in the manufacturing (i.e., 14.5%), service (i.e., 28.3%), and retail (i.e., 35.2%) industries.

10. The maximum low-doc loan amount was raised to $150,000 in 1998.

11. The primary strength of the data set is the coverage in both numbers of loans and the length of time that includes a full business cycle. The longitudinal nature of the data and the fact that it covers a substantial time period greatly increases the power of our statistical analysis. A primary weakness of the data set, however, is the lack of extensive borrower-level and firm-specific characteristics, such as credit history of the borrower and firm, as well as balance sheet and collateral information.

12. The SBA defines a "new" firm as one that has been in business for three years or less.

TABLE 2

DESCRIPTIVE STATISTICS FOR BORROWER LENDER AND LOAN CHARACTERISTICS

Panel A. Firm Characteristic

Attribute	*Number*	*% Total*	*% Default*	*F-test*[a]
1. Firm age				
New	4716	34.8	15.0	.7749
Existing	8834	65.2	14.8	
2. Corporate structure				
Corporation	7784	57.5	14.9	.5694
Partnership	966	7.1	16.1	
Individual	4800	35.4	14.7	
3. SIC division[b]				
Ag, fish, forest (A)	274	2.0	11.3	.0001
Mining (B)	48	0.4	18.8	
Construction (C)	763	5.6	16.4	
Manufacturing (D)	1970	14.5	16.0	
TCEG&S (E)	380	2.8	13.7	
Wholesale (F)	1315	9.7	16.2	
Retail (G)	4772	35.2	17.2	
FIRE (H)	147	1.1	10.2	
Service (I)	3829	28.3	11.3	
Non-classified (K)	52	0.4	15.4	
4. Firm sizes				
1	1710	12.6	12.8	.0014
2–5	5537	40.9	14.5	
6–25	4904	36.2	15.6	
26–100	1261	9.3	15.9	
100–500	128	0.9	17.2	
>500	10	0.1	50.0	

Panel B. Lender and Loan Characteristic

Attribute	*Number*	*% Total*	*% Default*	*F-test*
1. Originator type				
Bank	13,032	96.2	14.7	.0001
Nonbank	518	3.8	20.9	
2. SBA lender type				
Regular	9581	70.7	16.0	.0001
PLP	1447	10.7	8.0	
CLP	2522	18.6	14.8	
3. Low doc	4301	31.7	10.8	.0001
Full doc	9249	68.3	16.8	
4. Approval amount ($000)				
<50	3904	28.8	14.6	.2179
50–100	4750	35.1	14.7	
100–150	1635	12.1	15.6	
150–200	1024	7.6	12.8	
200–250	615	4.5	15.6	
250–500	1253	9.3	16.6	
>500	369	2.7	16.3	
5. Interest-rate type				
Fixed	2174	16	15.4	.5014
Variable	11,376	84.0	14.8	
6. Region				
Central	2296	16.9	15.0	.0026
West	3412	25.2	15.2	
Northeast	2266	16.7	14.2	
Midwest	1706	12.6	12.1	
Southeast	1925	14.2	15.5	
Southwest	1945	14.4	16.8	
7. Sold	2758	20.4	16.2	.0346
Held	10,792	79.6	14.6	

[a]F-test of the null hypothesis H_0: $\mu_1 = \mu_2 = ... = \mu_n$, where μ_i is the default rates by attribute.
[b]TCEG&S refer to transportation, communications, electric, gas, and sanitary services; FIRE refers to finance, insurance, and real estate.

There is a relatively large and statistically significant difference in the pooled default rates by industry classification. Although the statistical result might be partially affected by the small number of observations in several of the industry classification groupings—e.g., mining represents only 0.4% of the total, but has an exceptionally high, 18.8%, default rate—the differences in default rates for the industries with higher concentrations of loans are still relatively large (e.g., retail, 17.2% and service, 11.3%) and statistically significant. Finally, the size of firms qualifying for medium-maturity SBA loans is relatively small with almost 90% having fewer than 26 employees. Table 2 shows that the pooled default rate varies by firm size—the cumulative default rate tends to increase with firm size—and that the differences in default rates are statistically significant across size categories. Although this conclusion is affected by the small number of loans to firms with greater than 500 employees, the difference in default rates remains statistically significant after excluding this category.

Panel B of Table 2 provides the sample-specific descriptive statistics on both lender and loan characteristics. Elliehausen and Wolken (1990) and Cole and Wolken (1995) examine survey data from two separate National Survey of Small Business Finance during this time period and find that banks are the primary providers of credit to small businesses. Our sample is consistent with their results. The vast majority of SBA loans in our sample (more than 96%) were originated by banks. The pooled cumulative default rate of bank-underwritten loans is significantly lower than that for loans underwritten by nonbank financial institutions. This result supports previous research that banks produce valuable private information on borrowers (e.g., James, 1987, Petersen and Rajan, 1994) and that some of that information is unique to commercial banks through their provision of other services, such as checking accounts to small firms (Allen, Saunders, and Udell, 1991, Mester, Nakamura, and Rineault, 2001).

The pooled sample results in Table 2 also provide evidence of the success of the preferred lender program. PLP lenders, as a group, exhibit statistically lower default rates relative to both CLP and regular lenders. Loans originated under the PLP program experience a default rate that is half that of those originated under the other programs. Moreover, these results show that loans originated under the low-doc program exhibit lower overall default rates relative to "full-doc" loans. The low-doc results, however, should be considered preliminary given the program was introduced in 1994, and many of these loans, especially those originated in more recent years, have been "at risk" for a much shorter period. Approval amounts are concentrated in the under $150,000 loan categories and do not appear to affect default behavior.

More than 84% of the loans in our sample are variable rate loans. The interest rate re-pricing structure does not appear to affect the likelihood of default on a pooled basis. However, the large percentage of loans with variable interest rates are likely to have a significant effect on lenders' interest-rate risk, as opposed to credit risk, exposure as it pushes most of the interest-rate risk onto the borrowers.

In Table 2 we also report the distribution of borrowers by geographic location. We group borrowers by state of residence into six broadly defined regions with the West—which includes California—making up more than 25% of the sample.[13] The remainder of the originated loans are fairly well dispersed across the remaining regions ranging from 12.6% in the Midwest to 16.9% in the Central region. The difference in loan performance across regions can be attributed to numerous factors such as differences in economic conditions, application of underwriting guidelines, and the concentration of borrowers across industry classifications.

Finally, in Table 2 we show that the guaranteed portions of a relatively large percentage of SBA loans are sold in the secondary market. More than 20% of the loans in our sample had the guaranteed portion sold to investors. Interestingly, loans in which the guaranteed portion was sold have a statistically significant higher cumulative default rate (i.e., 16.2%) relative to those held by the originator (i.e., 14.6%).

The results of the pooled, univariate analyses summarized in Table 2 are useful only insofar as they identify factors that are highly correlated with default. However, without controlling for the impact of changes in program guidelines, the aging (or seasoning) of the loans, and the censoring of observations in 1998, these results may be misleading.

In Table 3 we report the number of loans, cumulative default rates, average dollar loan amounts, and mean SBA guarantee percentages by year of disbursement (i.e., loan cohorts). These results show that our sample is heavily weighted toward loans disbursed in the 1990s; a result that reflects the recent growth in SBA lending programs. The average nominal loan amounts increased until 1994 and declined thereafter. The decline coincides with the introduction of the SBA low-doc program in which the maximum loan amount was limited to $100,000. More recently, the SBA has taken steps to lower their exposure to loss through several minor structural changes to their program. Because lenders share the loss on the loans they place into the 7(a) loan guarantee program on a pro rata basis with the SBA, lowering the maximum allowable guarantee rate shifts more of the exposure onto the lenders. Although the average guarantee percentage of the SBA loans in our sample is 84.3%, the mean guarantee percentage has fallen below 78% for loans originated since 1996. These results suggest that the SBA's exposure to loss has been changing over time. Although the total dollar exposure has been increasing, their exposure to loss on individual loans has declined as they have lowered the average loan amount and transferred a larger percentage of the loss exposure onto the lenders—a shift in terms of the guarantee program that is likely to improve underwriting and strengthen loan monitoring by the lenders.

In addition, Table 3 shows that the cumulative default rate by cohort has been falling since 1987. This suggests that the recent changes in the 7(a) program have

13. The regions are defined as: Midwest: IL, IN, MI, OH, and WI; Central: IA, KS, MN, MO, NE, ND, and SD; the Southeast: DE, DC, FL, GA, MD, VA, NC, SC, WV, AL, KY, MS, and TN; the Southwest: AR, LA, TX, and OK; the West: AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, and WY; and the Northeast: CT, ME, MA, NH, NJ, NY, PA, RI, and VT.

TABLE 3

Descriptive Statistics for SBA 7(a) Loans by Loan Cohort[a]

	Totals		Cumulative Default Rate				Averages	
Year loan disbursed (cohort)[b]	Number of loans in cohort	% of Total in each cohort	Number of loans in default by cohort	Cumulative default rate by cohort	Cohort % of total defaults	Average annual default rate	Average loan amount by cohort	Average guarantee percentage by cohort
1983	160	1.2	47	29.4	2.3	4.2	$142,838	86.8
1984	568	4.2	170	29.9	8.4	4.3	$143,238	86.7
1985	344	2.5	95	27.6	4.7	3.9	$138,581	87.5
1986	489	3.6	159	32.5	7.9	4.6	$146,091	85.2
1987	481	3.6	131	27.2	6.5	3.9	$141,125	85.2
1988	410	3.0	100	24.4	5.0	3.5	$151,020	84.4
1989	493	3.6	112	22.7	5.6	3.2	$149,747	84.7
1990	538	4.0	113	21.0	5.6	3.0	$153,592	85.0
1991	511	3.8	96	18.8	4.8	2.7	$155,950	84.8
1992	801	5.9	125	15.6	6.2	2.6	$168,749	85.1
1993	821	6.1	107	13.0	5.3	2.6	$171,510	85.2
1994	1293	9.5	188	14.5	9.3	3.6	$139,328	85.2
1995	2459	18.1	339	13.8	16.8	4.6	$103,800	87.2
1996	1667	12.3	151	9.1	7.5	4.5	$104,535	79.4
1997	1859	13.7	82	4.4	4.1	4.4	$121,873	78.1
1998	656	4.8	3	0.5	0.1	0.9	$132,038	77.9
Total[c]	13,550	100	2018	14.9	100	3.5	$141,507	84.3

[a]Based on a 20% sample of seven-year maturity SBA 7(a) loans originated between 1983 and 1998.
[b]Loans originated since 1992 are right censored.
[c]The total cumulative default rate by cohort is the weighted average over the full sample.

affected the overall credit quality of the SBA loan program. The more recent decline in the cumulative default rate, however, is at least partially the result of censoring the seven-year maturity loans originated after 1991. We calculate the average annual default rate as a method of adjusting for the shorter exposure time of the censored loans. We find that the average annual default rate also declines after 1987; however, it begins to rise again after 1993.

The pattern in the average annual default rates for the more recent cohort groups is consistent with that expected if, in fact, default is time dependent. Loans disbursed after 1991 are censored in 1998. As a result, the average annual default rates for each year from 1992 to 1998 provide a time line of the default intensity over various levels of seasoning. In the absence of widespread fraud, few borrowers are likely to default soon after loan disbursement. However, immediately after disbursement, the firm is likely to be highly leveraged and, as a result, more dependent on the cash flow of the firm to make monthly payments. During this phase, even minor fluctuations in cash flow can endanger the solvency of weaker borrowers. Although a borrower might have retained a portion of its loan as working capital available to avoid early default, eventually these funds will be exhausted and the rate of default will begin to accelerate. As a result, we expect the rate of default to increase over the first several years after disbursement. However, as survival time increases and a more stable financial base is established, a borrower is better positioned to withstand minor fluctuations in the market. As a result, we expect the rate of default to fall off as the loans approach maturity.

The results in Table 3 are consistent with this characterization of the time path of default. The average annual default rates are significantly higher for loans more highly censored. The 1995–98 cohort groups represent loans censored during the first three years after disbursement—the period in which loan default is expected to be greatest. The relatively low default rate for the 1998 cohort group reflects the fact that few borrowers default soon after disbursement. However, the jump in the average annual default rate for the 1997 cohort is consistent with the expected acceleration of the rate of default over the first several years after disbursement. The relatively high average annual default rate for the 1995 and 1996 cohort groups suggests that the rate of default remains high into the third year. We also find that as the survival time approaches maturity, the average annual default rates of censored cohort groups (i.e., loans disbursed in 1992–94) decreases as expected.

The pattern in the average annual default rate in Table 3 suggests that the marginal default rates are time dependent. In Figure 1, we plot the unconditional distribution of the "time to default" for our sample of loans. Using a non-parametric hazard modeling approach, we generate the empirical hazard rates—defined as the ratio of the number of events (e.g., defaults) *in a given time period* to the number of accounts

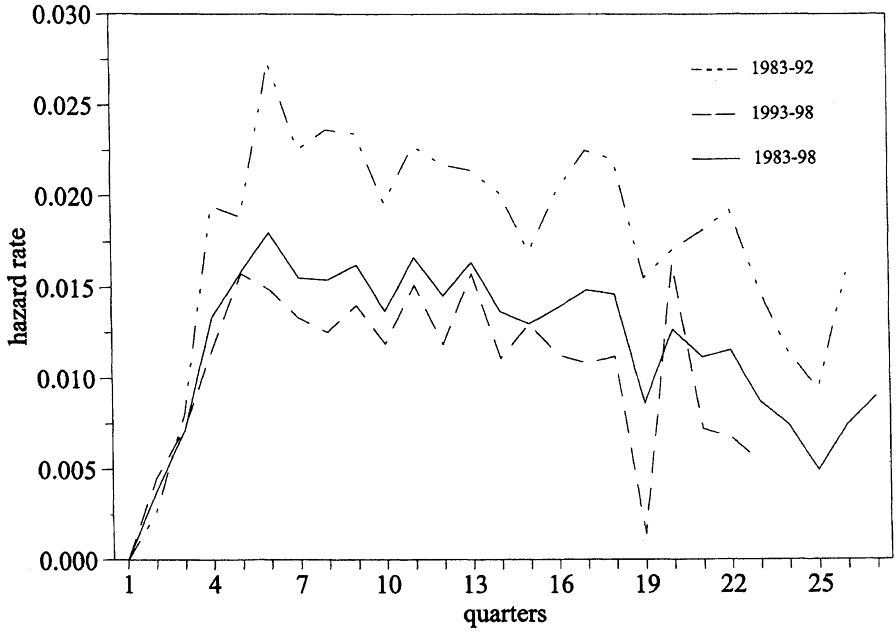

FIG. 1. Non-parametric hazard curves. - - - -, 1983–92; – –, 1993–98; —, 1983–98. (Note: The series were truncated to eliminate spikes in the graphs that result from (1) a rapid decline in the denominator of the hazard rate as the loan approaches maturity (i.e., a result that is attributed to the method used to define the effective sample size) and (2) an increase in the relatively small number of defaults in the last few quarters due to lender forbearance—a nuisance factor that greatly distorts the shape of the curves at, or near, maturity.)

"at risk" *at the beginning of the time period* (Allison 1995)—directly from the performance data over the full sample (i.e., 1983–98) and two sub-periods (i.e., 1983–92 and 1993–98). Under this approach, the duration of the data defines the shape of the hazard function, not the parameters of a specific distribution. For that reason, borrower, lender, and loan characteristics were omitted.

The graphs in Figure 1 show that, in general, the hazard curve is a concave function of time: increasing initially, then falling off toward zero as the loans season. The general shape of the hazard, and the time dependency of default implied from this result, is consistent with the expected effect of loan seasoning on default behavior.[14]

Although the general shape of the hazard curve is consistent across sub-periods, the hazard rates vary by sub-period. For example, the hazard function for the later period (i.e., 1993–98) appears to peak earlier, and at a much lower hazard rate, relative to the pooled hazard function based on the behavior of the 1983–92 cohort groups. The more recent cohort groups are performing better, on average, relative to the pre-1993 cohorts. The default rates, at similar points in time after disbursement, are much lower for loan-cohort groups originated after 1992.

The non-parametric results suggest that, although the general shape of the hazard function is concave, the specific form of the curve is sensitive to the choice of the observation period; a result that suggests that the hazard function may be conditional on the program, loan, and economic-specific conditions over time. Because the non-parametric hazard approach generates only unconditional estimates, we use a discrete-time hazard modeling approach that allows us to estimate the hazard rate conditional on both time-*invariant* factors (i.e., covariates that reflect credit-specific characteristics at time of origination), as well as covariates that vary over time (e.g., time-*varying* economic conditions).

2. ESTIMATING DEFAULT USING A HAZARD MODEL APPROACH

Loans originated under the SBA 7(a) program are primarily fixed maturity, amortizing loans. As a result, the cost of the default guarantee on a net present value basis depends not only on the percentage of the loans that default, but also the timing of the defaults. For that reason, we model the default process as a time-dependent event using a hazard model that explicitly takes into account the timing of default over the life cycle of the loan.

Under this approach, it is assumed that the time to default is a realization of a random process in which the event (i.e., default) time, T, is a random variable having a probability distribution (Allison 1995). The probability distribution of the random variable T can be characterized by the cumulative distribution function $F(t) = \Pr(T \leq t)$; or, more commonly by the survivor function:

14. The shape of the empirical non-parametric hazard is also consistent with the timing of defaults observed by Rhyne (1988) in her analysis of seven-year maturity SBA loans originated in 1973–83.

$$S(t) = \Pr(T > t) = 1 - F(t)\,,$$

where t is time, $S(t)$ the survivor function, $\Pr(T > t)$ the probability at the timing of the event T, is greater than some value t, and $F(t)$ is the cumulative density function (with a corresponding pdf denoted as $f(t)$) that represents the probability that the event time will be less than or equal to any value t.[15] More simply, the survivor function identifies the probability as loan survives past time t.

Alternatively, we can describe the same distribution of time to default using a hazard function (Kiefer 1988). The hazard rate is a measure of the probability that a loan will default in time t, given that it has survived up until that time. The hazard function is defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{F(\mathrm{t} + \Delta t) - F(t)}{\Delta t S(t)} = \frac{f(t)}{S(t)} = \frac{- d \ln S(\mathrm{t})}{dt}\,, \tag{1}$$

where $\Pr(.|.)$ is the conditional probability that the event takes place between t and $\mathrm{t} + \Delta t$ as Δt become small (i.e., approaches 0), and $F(.)$, $f(t)$, and $S(t)$ are defined above (Greene 1997). The purpose of this definition is to "... quantify the instantaneous risk that an event will occur at time t" (Allison 1995). There are several methods for estimating an empirical hazard model. A Cox Proportional Hazard (CPH) and a discrete-time hazard framework are two of the more common methods of incorporating time-varying covariates into a hazard model (e.g., Helwege, 1992, Shumway, 2001). Allison (1995), Shumway (2001), Brown and Goetzmann (1995), and Deng (1997) show that these two methods are statistically similar; however, a discrete-time hazard approach is computationally more efficient.

A discrete-time hazard model, however, requires that the data be structured in an "event-history" format. That is, for every loan there is an observation for each period (i.e., each quarter) the loan is active. For example, suppose the data set contains five loans all of which have contractual two-year maturity dates. Of these five loans, two pay as agreed (eight quarters of payment history for each loan), one is censored after one year (four quarters of payment history), one defaults after six months (two quarters of payment history), and the fifth loan defaults after 18 months (six quarters of payment history). Instead of five observations with a dependent variable assigned a value equal to the number of quarters to termination (either by default or censoring), we would have a total of 28 observations under an "event-history" sample design.[16] The dependent variable is a binary indicator variable that equals one if the loan defaulted in that period and zero otherwise. More specifically, the dependent variable is derived as a series of (stacked) binary variables, $D_i(1),\ldots D_i(t)$ for each loan i over t periods. Each observation $D_i(t_j)$ is assigned the

15. The pdf is defined as $f(t) = dF(t)/dt$, where $F(t) = \int_0^1 f(s)\, ds$. Moreover, $dS(t)/dt = -f(t)$.

16. In this example, the first two loans generate 16 observations, the third loan four observations, the fourth loan two observations, and the fifth loan an additional six observations for a total of 28 observations.

value 0 if the loan survives over the period (t_{j-1},t_j); and a value 1 if the loan defaults within the specific interval. More specifically, for a five-year (i.e., 20 quarters) maturity loan that defaulted in the fourth quarter after origination will contribute a total of only four (out of a possible 20) observations to the model-development sample in which the dependent variable takes the values $D_{it} = 0$, for $t = 1,2,3$; and $D_i,4 = 1$ in the default period. This data design captures the conditional probability of default within the interval (t_{j-1},t_j), given that the loan survived to t_{j-1}.

Our sample of 13,550 SBA 7(a) loans generates an "event-history" sample of 175,261 loan-quarters observations, including 2,018 loan-quarters in which a loan defaulted. The event-history sample design allows us to estimate the hazard model using qualitative dependent variable estimation techniques. More specifically, if we define D^*_{it} as a latent index value that represents the unobserved propensity to default conditional on the covariates, we can model the default propensity as:

$$D^*_{it} = \mathbf{x}_i\beta + \mathbf{z}_{t,t}\phi + \varepsilon = \mathbf{w}\theta + \varepsilon\,, \tag{2}$$

where $\mathbf{x}_i$ is a vector of time-invariant covariates (e.g., borrower, loan, and lender characteristic), $\mathbf{z}_{i,t}$ a vector of time-varying covariates (e.g., loan-specific and economic indicator variables), and β and ϕ are the corresponding vectors of time-invariant and time-varying parameters (Gross and Souleles 1998). We rewrite Equation (2) more compactly with $\mathbf{w} = [\mathbf{x}_i,\mathbf{z}_{i,t}]$ and $\theta = [\beta,\phi]'$. If we assume ε is distributed as a standard logistic and define:

$$D_{it} = 1 \quad \text{if} \quad D^*_{it} > 0\,,$$
$$D_{it} = 0 \quad \text{if} \quad D^*_{it} \leq 0\,,$$

then the probability that $D_{it} = 1$ is

$$\Pr(D^*_{it} > 0) = \Pr(\mathbf{w}\theta + \varepsilon > 0)\,,$$
$$\Pr(D^*_{it} > 0) = \Pr(\varepsilon > -\mathbf{w}\theta)\,,$$
$$\Pr(D_{it} = 1) = \Lambda(\mathbf{w}\theta)\,, \tag{3}$$

where $\Lambda(.)$ indicates the logistic cumulative distribution function. We estimate Equation (3) as a standard logit model (Allison, 1995, Gross and Souleles, 1998, Blank, 1989, Shumway, 2001).

An important feature of the discrete-time hazard approach is its ability to incorporate general economic indicators and loan-specific time-varying covariates (i.e., $\mathbf{z}_{i,t}$) directly into the model. The time-varying covariates capture the impact of changing conditions on the likelihood of default over the life of the loan. For example, we utilize aggregated economic time-series data specific to the region (i.e., state) and industry (i.e., SIC division) of the borrower to capture the impact of changes in the borrower's local economic and market conditions on the hazard rate. More specifically, we include the deviation in the state and industry-specific income (i.e., State/Industry Income: Change) from the value at time of origination, and the quarterly growth in employment (i.e., State/Industry Employment: Growth) specific to the state and industry of the borrower. In this way, the economic indicators are used to capture the condition of the local market in which the small business operates.[17]

In Table 4 we report the results of the discrete-time hazard model for several specifications. In Panel A, we use a sixth-order polynomial in age to capture the impact of loan seasoning on the hazard rate.[18] The results of incorporating various lender, borrower, and loan attributes are reported in Panels B–E. The additional loan-specific information improves the model fit as reflected in the increase in the log-likelihood ratios (LR). The parameter estimates are, in general, significant and robust across model specifications and are discussed in detail in the following sections.

2.1. Lender Attributes

We find that the impacts of both SBA lender type (i.e., PLP, CLP, and regular) and loan originator (i.e., bank/nonbank) on the likelihood of default are consistent with the univariate results in Table 2. Using the results in Panel E of Table 4, we find that loans originated by PLP lenders, on average, have a 21.7% (i.e., $e^{-0.2453} = 0.782$) lower odds of default, and those originated by CLP lenders have a 24.4% lower odds, over the life of the loan relative to loans submitted by regular lenders.[19] More interesting, however, is the result that after controlling for other factors, CLP lenders underwrite loans that are of equal or better credit quality as those submitted by PLP lenders. As expected, bank lenders appear to have a comparative advantage in lending to small firms. They have, on average, a 32.2% lower odds of default than nonbank lenders, all else equal. These results are robust across the alternative specifications (i.e., Panels B–E).

The results on the low-doc program are mixed across model specification. After controlling for a limited number of other lender, borrower, and loan attributes, low-doc status is statistically unimportant (Panels B–D). However, after controlling for structural changes (i.e., time dummies) and differences in regional economic conditions, loans originated under the low-doc program are more likely to default. These results reflect the fact that the low-doc program, introduced in 1994, coincided with a general improvement in overall credit quality of SBA loans in the 1990s. As a result, without controlling for those factors that reflect the shift in performance behavior, the models in Panels B–D could not isolate the specific impact of the low-doc program. The results in Panel E show that, after controlling for other factors, low-doc loans experience a 26.4% (i.e., $e^{0.2345} = 1.264$) higher odds of default.

17. For example, a five-year loan originating in the third quarter of 1990 to a wholesaler operating in New York would be matched with the income and employment trends in the wholesale trade industry, in the state of New York, for the full 20-quarters beginning in the third quarter of 1990.

18. We tested several other transformations of time-since-origination (i.e., age) to capture the concave shape of the hazard function identified in Figure 1. These included a piecewise transformation using annual and quarterly dummies, a quadratic function of age, and several higher-order polynomials of the age variable. We found that a sixth-order polynomial fits the data best based on a comparison of LR.

19. The odds ratios for the discrete covariates reported in Table 4 are calculated as e^{β}. It can be shown that the odds ratio corresponding to an increase in a continuous covariate X from, say, a value x_a to x_b $(b > a)$ is $\psi = [\exp(\beta)]^c$ where $c = x_b - x_a$. ψ is interpreted as the change in the odds for any increase of c-units in the corresponding risk factor X.

TABLE 4

DISCRETE-TIME HAZARD MODEL

	Panel A		Panel B		Panel C		Panel D		Panel E	
$-2 \ln L$ ($\beta = B_i$)	21485.2		21426.8		21328.9		21321.6		21229.8	
$-2 \ln L$ ($\beta = 0$)	22030.1		22030.1		22030.1		22030.1		22030.1	
LR-test: $-2(\ln L_{\beta=0} - \ln L_{\beta=B})$ (χ^2-test, p-value)	544.8 (*.0001*)		603.2 (*.0001*)		701.2 (*.0001*)		708.4 (*.0001*)		800.2 (*.0001*)	
$-2 \ln L$ ($\beta = B_j$)[a]			21485.2		21426.8		21328.9		21321.6	
LR-test: $-2(\ln L_{\beta=Bi} - \ln L_{\beta=Bj})$ (χ^2-test, p-value)			58.4 (*.0001*)		97.9 (*.0001*)		7.3 (*.1209*)		91.8 (*.0001*)	
Variables	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value
Intercept	−9.800	*0.0001*	−9.221	*0.0001*	−9.226	*0.0001*	−10.010	*0.0001*	−9.774	0.0001
1. Loan Seasoning										
Age	2.840	*0.0001*	2.848	*0.0001*	2.856	*0.0001*	2.859	*0.0001*	2.865	*0.0001*
Age^2	−53.010	*0.0001*	−53.179	*0.0001*	−53.270	*0.0001*	−53.360	*0.0001*	−53.427	*0.0001*
Age^3	471.700	*0.0001*	473.600	*0.0001*	474.000	*0.0001*	474.600	*0.0001*	476.900	*0.0001*
Age^4	−2127.9	*0.0003*	−2137.0	*0.0003*	−2135.6	*0.0003*	−2136.4	*0.0003*	−2157.9	*0.0002*
Age^5	4598.3	*0.0077*	4614.3	*0.0075*	4602.7	*0.0077*	4598.9	*0.0077*	4676.0	*0.0068*
Age^6	0.00000	*0.0603*	0.00000	*0.0599*	−0.000000	*0.0614*	0.00000	*0.0621*	0.0000	*0.0556*
2. Lender Attributes										
PLP			−0.382	*0.0001*	−0.3851	*0.0001*	−0.3141	*0.0034*	−0.2453	*0.0228*
CLP			−0.287	*0.0001*	−0.2957	*0.0001*	−0.2996	*0.0001*	−0.2797	*0.0001*
Bank lender			−0.522	*0.0001*	−0.4816	*0.0001*	−0.4382	*0.0001*	−0.3879	*0.0002*
Low doc			0.002	*0.9749*	−0.0121	*0.8374*	−0.0104	*0.8648*	0.2345	*0.0010*
3. Borrower Attributes										
Employees					0.00052	*0.0001*	0.00053	*0.0001*	0.00051	*0.0001*
New business					0.1300	*0.0077*	0.1169	*0.0178*	0.0929	*0.0614*
Corporation					0.0217	*0.6681*	0.0307	*0.5546*	0.0456	*0.3814*
Partner					−0.0512	*0.5692*	−0.0538	*0.5502*	−0.0753	*0.4029*
Industry: retail (G)					0.1623	*0.0018*	0.1557	*0.0031*	0.1722	*0.0012*
Industry: service (I)					−0.2810	*0.0001*	−0.2842	*0.0001*	−0.2047	*0.0020*
Northeast					−0.1358	*0.0954*	−0.1292	*0.1133*	−0.1459	*0.0759*
Midwest					−0.3136	*0.0006*	−0.3051	*0.0009*	−0.3492	*0.0001*
Central					−0.2133	*0.0084*	−0.2039	*0.0124*	−0.2474	*0.0025*
Southwest					0.0418	*0.6068*	0.0474	*0.5625*	0.0502	*0.5401*
West					−0.0624	*0.4035*	−0.0692	*0.3572*	−0.0890	*0.2375*

Variables	Panel A		Panel B		Panel C		Panel D		Panel E	
	Estimate	*p*-Value	Estimate	*p*-Value	Estimate	*p*-Value	Estimate	*p*-Value	Estimate	p-Value
4. Loan Attributes										
Approval amt.							0.000000	*0.8599*	0.0000	*0.5665*
Var. interest rate							−0.0127	*0.8372*	0.0299	*0.6303*
SBA guar. %							0.8574	*0.0371*	0.7442	*0.0872*
Sold							0.0878	*0.1234*	0.0905	*0.1107*
5. Time Dummies/Macroeconomic Conditions										
D91									−0.4859	*0.0001*
D96									0.0804	*0.3543*
Region/industry income: deviation (t_0)									−0.5270	*0.0491*
Region/industry employ: growth (t)									−2.1019	*0.0429*

[a]The ln L values are those from the previous panel. Because each successive model specification contains the previous model, they are nested which allows us to use a LR test to identify the better model.

2.2. Borrower Attributes

Although our data set does not contain information on the firm's financials, we do have several important borrower attributes that identify aspects of the firm's structure or that of the market/industry in which the firm operates. We evaluate the effect of firm size (proxied by the number of employees), new-firm status, corporate structure (i.e., corporation, partnership, or individual proprietor), and industrial classification (i.e., SIC Division), as well as the geographic location of the borrowers.

The results of including borrower-specific attributes in the model are reported in Panel C of Table 4. Contrary to the literature on small firms, the odds of default increases as the size of the firm increases, all else equal.[20] The magnitude of this effect, however, is very small. For each five person increase in the number of employees, on average, the odds of default increases by only 2.6%.[21] Consistent with previous research, new firms have a significantly higher odds of default relative to established firms, all else equal. This result is at odds with the univariate analysis in Table 2. However, the univariate results were derived without controlling for differences in lender, borrower, and loan attributes. Under the hazard model design, these other factors are held constant. As a result, the hazard model provides a better measure of the covariate-specific effect. On average, we find that new firms have a 9.7% higher odds of default (Panel E) relative to established firms with similar loan-specific attributes.

Similar to the results in Table 2, we find that the corporate structure does not affect the odds of default. The odds of default for loans to corporations or to partnerships are statistically the same (all else equal) as those to individuals irrespective of model specification. A borrower's industrial classification, however, does influence the odds of default. Loans to firms in the retail industry sector are more likely, and those to firms in the service industry sector are less likely, to default (all else equal) relative to firms in other industries—a result that reflects the higher default rate for retail firms, and the lower default rate for service firms, reported in Table 2. More specifically, from the results in Panel E, firms in the retail sector have odds of default that are 18.8% higher than firms in other industries, and firms in the service industry have odds of default that are 18.5% lower than firms in non-service industries. There are also differences in performance by geographic location. The Northeast, Central, and Midwest regions appear to have lower odds of default relative to the Southeast (and other) regions. This result may reflect factors that are difficult to measure such as differences in non-specific economic factors, industry concentrations, and the application of SBA-imposed underwriting guidelines across regions.

20. This result is due in part to "selection bias." Larger small firms relying on SBA guarantees are very risky since after attaining this size they should have migrated to the traditional bank loan market. Thus, larger firms in the SBA program are likely to be among the riskiest in this size category.

21. Using the derivation of the change in odds for continuous covariates outlined the previous footnote, we show, for an increase in a covariate X from, say, a value x_a to x_b ($b > a$), that the change in odds is $\psi = [\exp(\beta)]^c$ where $c = x_b - x_a$. Therefore, the change in the odds for an increase in the number of employees by five is $\psi = \exp(0.00051)^5 = 1.00255$, or approximately a 0.26% increase.

2.3 Loan Attributes

We find that neither approval amount nor interest-rate type has an impact on the odds of default; results are consistent with univariate analysis. However, in Panels D and E, we find that the higher the guarantee percentage, the greater the odds of default. Using the estimate of the impact of an increase in the guarantee rate from Panel E, we calculate that a 5% increase in the guarantee rate increases the odds of default by 3.8%. That is, for every 1% increase in the guarantee rate, the odds of default increases by roughly the same percentage.

Loans in which the SBA-guaranteed portion is sold in the secondary market have, on average, a higher odds of default (i.e., 9.5%), all else equal. This result is consistent with the argument that loans in which the guaranteed portion is sold are less likely to be closely monitored after disbursement. Moreover, given that the SBA guarantee increases the spread to the seller on the sale of the guaranteed portion, lenders have an incentive to underwrite a larger volume of SBA loans; a process that may involve trading quality for quantity.

Finally, although we generally expect loans with high, fixed interest rates to be relatively more risky, it is doubtful that loans booked under the SBA 7(a) program were risk-based priced over our sample period given the institutional restrictions (i.e., pricing caps) that limited a lender's ability to price for risk.[22] Moreover, as noted above, lenders can generate additional income from the spread between the loan rate paid by the borrower and the rate paid to the investors who purchase the securities enhanced by the SBA guarantee of the underlying assets. This provides an incentive for lenders to lower their offer rate to the borrower as a method of increasing loan volume for the purpose of generating larger loan pools. As a result, some lenders may lower quality at the same time they are lowering the offer rates to increase loan volume. These factors suggest that the expected inverse relationship between the interest rate (i.e., offer rate) and the likelihood of default for loans priced for risk will not hold in general for loans originated under the SBA program.

2.4. Time/Macroeconomic Conditions

In addition to the borrower, lender, and loan-specific factors discussed above, there were several qualitative changes in institutional factors introduced in the 1990s that were likely to have affected the overall credit quality (and default experience) of the SBA portfolio. For example, following the implementation of the 1990 FCRA, the SBA modified its lending programs to be consistent with the goal of streamlining risk assessment of government credit programs. As these changes were taking place, commercial banks were implementing the new risk-based capital (RBC) rules outlined in the 1988 Basel Accord.[23]

22. Lenders were more likely to have used a quantity adjustment instead of price adjustment mechanism to account for increased risk. That is, they were more likely to have denied lower quality applicants (i.e., higher risk applicants) for a given price, than adjust the price to reflect the higher risk. Given the requirement that the terms of an SBA loan correspond to those the applicant could not obtain under conventional underwriting guidelines, lenders were more likely to have made loans without pricing for risk and relied on the SBA guarantee to reduce their exposure to losses.

23. The risk-based capital rules outlined in the 1988 Basel Accord were phased in over several years. The transition period ended, and the rules were fully implemented, by December 31, 1992.

Under RBC, small-business loans in general are treated for capital adequacy purposes the same as all other C&I loans. As a result, the relatively lower regulatory capital charge on potentially higher risk small-business loans relative to loans to, say, AAA-rated corporations provided banks with the incentive to shift the composition of their loan portfolio toward small-business lending. Moreover, under RBC, the guaranteed portions of SBA small-business loans are subject to lower regulatory capital requirements. This condition provided banks with the additional incentive to underwrite a larger percentage of their small-business loans through the SBA program. To the extent that banks had previously (i.e., pre-RBC) placed their more risky small-business loans through the SBA program, the incentive to increase the percentage of their small-business loans underwritten through the SBA program should have resulted in an increase in the average credit quality of SBA loans. Finally, in 1996, the SBA modified their programs again to extend the coverage of their appropriations over a wider base of applications that included a reduction of the maximum guarantee percentage on loans less than 100,000 to 80% and all other loans to 75%.

Loans originated after the implementation of FCRA and the introduction of the RBC rules have a significantly lower odds of default. The estimated 38.5% lower odds of default in Table 4 is consistent with the lower hazard curve (i.e. 1993–98) in Figure 1. The institutional changes—primarily a reduction in the maximum guarantee percentages—implemented since 1996, however, have had no statistically significant effect on the odds of default.

The introduction of time-varying economic conditions using a combination of local economic/industry-specific variables also has had an impact on the odds of default. We include information on the deviation of region/industry-specific income away from the level at time of loan origination and the growth in region/industry-specific employment growth. Employment growth is used to capture the impact of industry and region-specific conditions on loan performance. To the extent that small businesses are more sensitive to local and industry economic trends, a slow down in the local market will increase the likelihood, and the timing, of default. However, if a period of an economic slowdown follows an extended period of growth, the impact of the slowdown on the survival (or default) probability is likely to be small. A decline in local economic activity following a period of growth would have to be relatively large to reverse the positive effect of the previous growth on survival. For that reason, we include the deviation in local and industry income from its level at time of origination (t_0). If the slowdown is small relative to the expansion since origination, the deviation will remain positive. Only for large declines in income will the deviation fall below its original level following an extended period of growth and contribute to the decline in the odds ratio. Our results are consistent with the above arguments. In Panel E we show that increases in both the deviation in the region/industry-specific income at time of origination and growth in region/industry-specific employment decrease the likelihood of default and increase survival time. If income deviates from its initial level by its average change (i.e., 0.1298), the odds of default are reduced by 6.6%

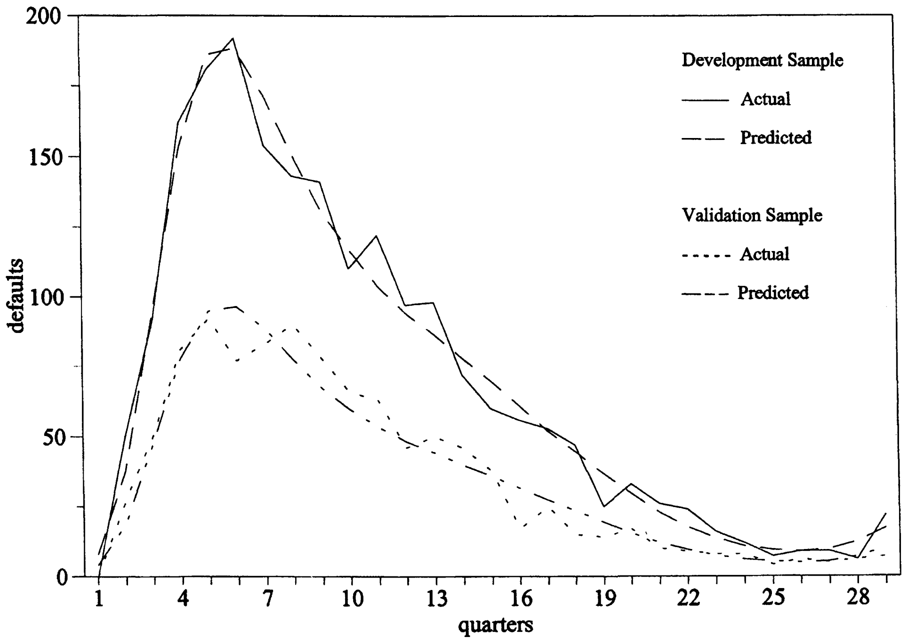

FIG. 2. Discrete-time hazard function. Development sample: —, actual; – –, predicted. Validation sample: - - -, Actual; – –, Predicted.

(i.e., $\exp(-0.5270)^{0.1298} = 0.9338$). Growth in region/industry employment by its average growth rate (i.e., 0.0063), lowers the odds of default by 1.3%.

2.5 Model Fit

A comparison of the log-likelihood values across model specifications is used to assess the improvement in within-sample fit as more information is introduced. The relatively large increase in the LR-test statistics between specifications implies that each model improves upon the other, although the improvement associated with the introduction of the loan attributes (Panel D) is statistically weak.[24]

In Figure 2, we show that the full model (Panel E) performs well at predicting the number of defaults for both in- and out-of-sample borrowers. The in-sample (i.e., development sample) results show that the predicted number of defaults closely

24. The LR test in Panel D is calculated as minus two times the difference in the LR between the models in Panel C (restricted model) and Panel D (unrestricted model). That is, the LR statistic is calculated as:

$$LR = -2(\ln L_r - \ln L_u) = 21328.9 - 21321.6 = 7.3 .$$

LR is distributed as a chi-square with four degrees of freedom (Greene 1997). The probability that an observation from a chi-square distribution with four degrees of freedom is greater than 7.3 is 0.1209 (a value slightly greater than $\alpha = .10$).

follows actual defaults over the life of the loan. The out-of-sample (i.e., validation sample) results show a similar pattern.

3. CONCLUSIONS

In this paper, we examine the loan performance of small firms receiving SBA-guaranteed loans. We place the performance history of SBA loans into perspective by comparing their default experience to that of rated corporate bonds. We find that the historical cumulative default rates of SBA-guaranteed loans fall between the Ba/BB and B rated corporate-bond grades as reported by Moody's and S&P. Although as a group SBA loans are below investment grade, the historical behavior places them in the upper end of the speculative grade category, and of similar credit quality as a large percentage of loans held by large commercial banks.

The growth path of the cumulative default rates, and the results of a non-parametric analysis of the timing of default, strongly suggest that the likelihood of default for medium-maturity SBA 7(a) guaranteed loans is time dependent. This is important for the SBA because, under FCRA, they are required to estimate the expected payment flows net of losses, which requires modeling the time path of defaults over the life of the loans. As a result, the timing of default plays an important role in assessing the SBA's exposure (in dollars) to losses on a net present value basis. More generally, however, the timing of default is an important feature of credit-risk modeling of interest to all lenders, especially under the proposed changes to bank capital rules, since they must estimate their exposure to credit losses over the life of the loans to allocate reserves against expected loan losses.

We use a discrete-time hazard model conditional on borrower, lender, and loan characteristics, as well as changes in the economic environment over time, to capture the expected default rates at each point in time over the life of the loan. We find that the SBA's exposure to default increases initially, peaks in the second year after origination, and declines thereafter. The inverted U-shaped hazard function for the 1983–98 observation period is consistent with previous research on SBA loan performance (Rhyne 1988).

Unlike previous research, however, we find that the hazard of default for SBA loans is conditional on several key borrower, lender, and loan characteristics. More specifically, we find that loans to new businesses have a greater hazard of default than established firms. We find, somewhat surprisingly, that larger SBA-qualified firms, as measured by the number of employees at time of loan origination, experience a greater hazard of default than smaller ones. There are several possible reasons for this effect including borrower reputation effects, different loan purpose, or a greater degree of commingling firm and personal assets as collateral for smaller firms that reduces the propensity to default.

In terms of lender characteristics, SBA specialized lender programs (i.e., CLP and PLP) have lower hazard rates relative to its Regular Lender Program. The performance of loans underwritten by the specialized lenders—lenders usually with the most experienced with SBA-qualified borrowers—shows that small-business lending is unique and may require a high degree of expertise gained over time. We also find that loans with higher guarantee percentages are associated with a greater hazard of default. This result is consistent with economic intuition, since lenders with less at stake are more likely to make riskier loans and have less incentive to monitor them on an ongoing basis. The SBA has been addressing this issue by lowering the maximum guarantee percentage over the past decade. We also find that loans underwritten under the low-documentation program initiated in 1994 experience higher hazard rates. Interestingly, however, we found that SBA lenders did not, in general, risk-base price during the sample period—although there are indications in the market that small-business lenders have recently begun to use credit scores as part of their underwriting analysis (a first step toward implementing a risk-based pricing approach).

Finally, we identify an important link between both regional and industry economic conditions on the hazard of default. Similar to previous research on other loan types, we find that the success and failure of small loans are closely tied to both the regional and industry-specific economic conditions in which the borrower operates. Moreover, we find support for a regime switch following the implementation of FCRA and the Basel risk-based capital rules in 1991.

Future research should extend our analysis to include prepayment estimates within the framework of a competing risk model. In addition, the model should be extended to include a measure of loss given default—a component of the net present value analysis that links the unit default rate to dollar losses through the estimation of the expected recovery rate.

LITERATURE CITED

Allen, Linda, Anthony Saunders, and Gregory F. Udell (1991). "The Pricing of Retail Deposits: Concentration and Information." *Journal of Financial Intermediation* 1, 335–362.

Allison, Paul D. (1995). *Survival Analysis Using the SAS System: a Practical Guide*. Cary, North Carolina: SAS Institute.

Berger, Allen N., W. Scott Frame, and Nathan H. Miller (2002). "Credit Scoring and the Availability, Price and Risk of Small Business Credit." Working Paper 2002–26, Board of Governors of the Federal Reserve System.

Berger, Allen N., and Gregory F. Udell (1995). "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business* 68, 351–382.

Blank, Rebecca (1989). "Analyzing the Length of Welfare Spells." *Journal of Public Economics* 29, 245–278.

Brown, Stephen, and William Goetzmann (1995). "Performance Persistence." *Journal of Finance* 50, 679–698.

Carey, Mark S., and Mark Hrycay (2001). "Parameterizing Credit Risk Models with Ratings Data." *Journal of Banking and Finance* 25, 197–270.

Cole, Rebel A., and John D. Wolken (1995). "Financial Services Used by Small Businesses: Evidence from the 1993 National Survey of Small Business Finances." *Federal Reserve Bulletin* 81, 629–667.

Cole, Rebel A., John D. Wolken, and R. Louise Woodburn (1996). "Bank and Nonbank Competition for Small Business Credit: Evidence from the 1987 and 1993 National Surveys of Small Business Finances." *Federal Reserve Bulletin* 82, 983–995.

Deng, Yongheng (1997). "Mortgage Termination: an Empirical Hazard Model with Stochastic Term Structure." *Journal of Real Estate Finance and Economics* 14, 309–331.

Elliehausen, Gregory, and John D. Wolken (1990). "Banking Markets and the Use of Financial Services by Small and Medium Size Businesses." *Federal Reserve Bulletin* 76, 801–817.

Fons, Jerome S., Lea V. Carty, and Denis Girault (1993). *Corporate Bond Defaults and Default Rates 1970–1992*." Moody's Special Report, New York, NY.

Good, Walter, and Robert Graves (1993). "Small Business Support Programs: the Views of Failed Versus Surviving Firms." *Journal of Small Business and Entrepreneurship* 10, 66–76.

Greene, William H. (1997). *Econometric Analysis*. Upper Saddle River, New Jersey: Prentice Hall.

Gross, David, and Nicholas Souleles (1998). "Explaining the Increase in Bankruptcy and Delinquency: Stigma vs. Risk Composition." Working Paper 98-28-B, Wharton Business School.

Helwege, Jean (1992). "Determinants of Savings and Loan Failure Rates: Estimates of a Time-varying Proportional Hazard Function." Federal Reserve Bank Finance and Economics Discussion Papers Series 207.

Honjo, Yuji (2000). "Business Failure of New Firms: an Empirical Analysis Using a Multiplicative Hazard Model." *International Journal of Industrial Economics* 18, 557–574.

James, Christopher (1987). "Some Evidence on the Uniqueness of Bank Loans." *Journal of Financial Economics* 28, 217–235.

Kiefer, Nicholas M. (1988). "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26, 646–679.

McDonald, Cynthia. G., and Linda M. Van de Gucht (1999). "High-yield Bond Default and Call Risks." *The Review of Economics and Statistics* 81, 409–419.

Mester, Loretta, Leonard Nakamura, and Micheline Rineault (2001). "Checking Accounts and Bank Monitoring." Working Paper 01-3/R. Federal Reserve Bank of Philadelphia.

Moody's (2002). "Default and Recovery Rates of Corporate Bond Issuers, 1920–2001." Moody's Investor Service, Global Credit Research, New York.

Peek, Joe, and Eric S. Rosengren (1996). "Small Business Credit Availability: How Important is Size of Lender?" In *Financial System Design: the Case for Universal Banking*, edited by Anthony Saunders and Ingo Walter, pp. 628–655. Burr Ridge, IL: Irwin.

Petersen, Mitchell A., and Raghuram Rajan (1994). "The Benefits of Lender Relationships: Evidence from Small Business Data." *Journal of Finance* 49, 3–37.

Rhyne, Elisabeth (1988). *Small Business, Banks and SBA Loan Guarantees*. Westport, CT: Greenwood Press.

Scott, Jonathan, and William Dunkelberg (2003). "Bank Mergers and Small Firm Financing." *Journal of Money, Credit, and Banking* 35, 999–1018.

Shumway, Tyler (2001). "Forecasting Bankruptcy More Accurately: a Simple Hazard Approach." *Journal of Business* 74, 101–124.

Standard and Poor (2001). "Corporate Defaults: Will Things Get Better Before They Get Worse." Standard and Poor's, New York.

Treacy, William F., and Mark S. Carey (1998). "Credit Risk Rating at Large U.S. Banks." *Federal Reserve Bulletin* 84, 897–921.